stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE BC 15/75 NOVEMBER

NETWORK FLOW THEORY

E.L. LAWLER

Not for distribution; for internal use only.

2e boerhaavestraat 49 amsterdam

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

AMS(MOS) subject classification scheme (1970): 90B10, 90C35

Network flow theory

by

E.L. Lawler

These notes are extracted, with minor adaptations, from chapter 4 of the book, *Combinatorial Optimization: Networks and Matroids*, to be published by Holt, Rinehart, and Winston, New York.

1. INTRODUCTION

Intuitively, a "flow network" is a directed graph in which an incompressible fluid flows, subject to certain physically plausible constraints. In a "network flow problem", one wishes to find an "optimal" flow in a given flow network. Mathematically, such a problem is simply a linear programming problem.

There are three very important and closely interrelated characteristics of the special class of linear programming problems which we call network flow problems: They possess optimal solutions in integers. They admit particularly efficient methods of solution. They provide an opportunity for formulating and solving a large number of interesting and important combinatorial problems, some of which have little, if any, obvious connection to the physical reality of flows.

These notes are intended to provide an introduction to the "classical" results of network flow theory: problem formulations, algorithms, theory, and practical applications. In so doing, we shall expect that the reader has at least an elementary knowledge of graph theory (acquaintance with nodes, arcs, paths, cutsets) and of linear programming (problem formulations, duality theory).

2. MAXIMAL FLOWS

Suppose that each arc (i,j) of a directed graph G has assigned to it a nonnegative number c_{ij}, the *capacity* of (i,j). This capacity can be thought of as representing the maximum amount of some commodity that can "flow" through the arc per unit time in a steady-state situation. Such a flow is permitted only in the indicated direction of the arc, i.e. from i to j.

Consider the problem of finding a maximal flow from a *source node* s to a *sink* node t, which can be formulated as follows. Let

$$x_{ij} = \text{the amount of flow through arc } (i,j).$$

Then, clearly,

(2.1) $$0 \le x_{ij} \le c_{ij}.$$

A *conservation law* is observed at each of the nodes other than s or t. That is, what goes out of node i,

$$\sum_j x_{ij},$$

must be equal to what comes in,

$$\sum_j x_{ji}.$$

So we have

(2.2) $$\sum_j x_{ji} - \sum_j x_{ij} = \begin{cases} -v, & i = s \\ 0, & i \neq s,t \\ v, & i = t \end{cases}$$

We call any set of numbers $x = (x_{ij})$ which satisfy (2.1) and (2.2) a *feasible flow*, or simply a *flow* and v is its *value*. The problem of finding a maximum value flow from s to t is a linear program in which the objective is to maximize v subject to constraints (2.1) and (2.2).

Let P be an undirected path from s to t. An arc (i,j) in P is said to be a *forward* arc if it is directed from s toward t and *backward* otherwise. P is said to be a *flow augmenting path* with respect to a given flow $x = (x_{ij})$ if $x_{ij} < c_{ij}$ for each forward arc (i,j) and $x_{ij} > 0$ for each backward arc in P.

Consider the network shown in Figure 2.1. The first number beside each arc (i,j) indicates its capacity c_{ij} and the second number indicates the arc flow x_{ij}. It is easily verified that the flow satisfies conditions (2.1) and (2.2), with s = 1 and t = 6, and that the flow value is 3.

An augmenting path with respect to the existing flow is indicated in Figure 2.2. We can increase the flow by one unit in each forward arc in

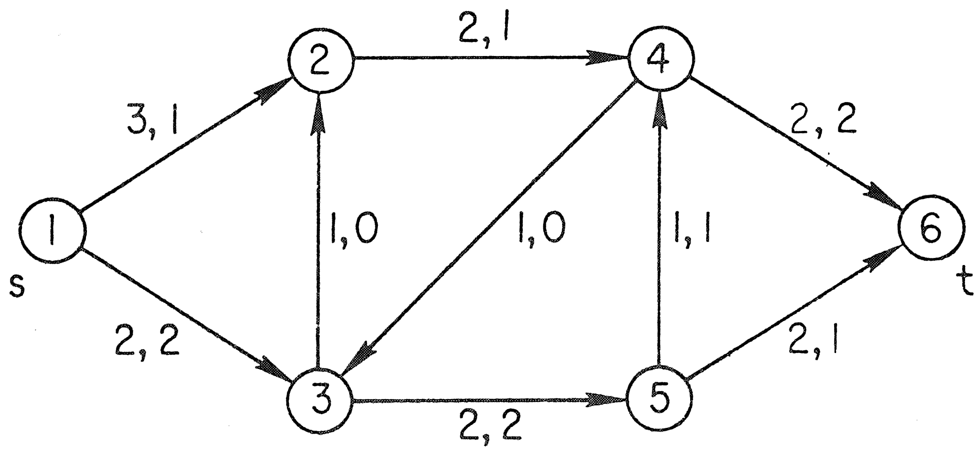


Figure 2.1. Feasible Flow

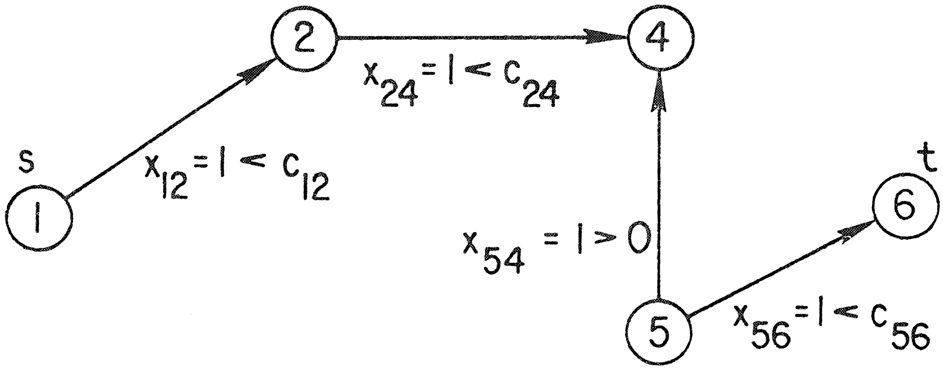


Figure 2.2. Augmenting Path

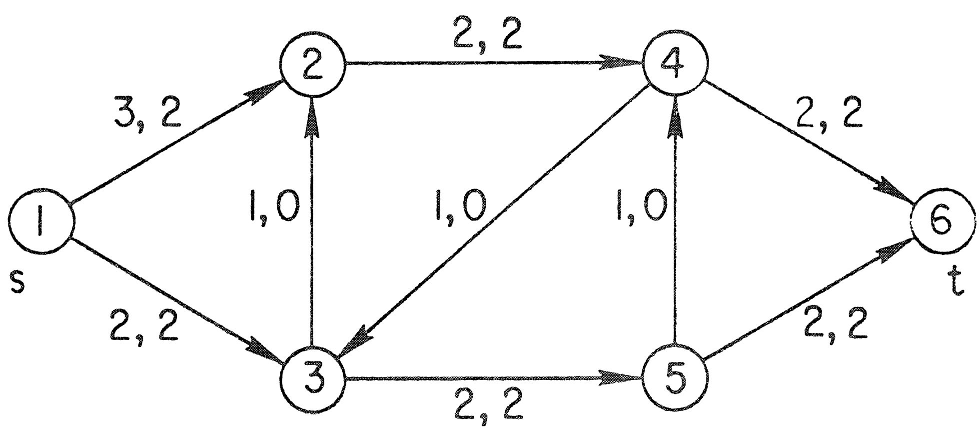


Figure 2.3. Augmented Flow

this path and decrease the flow by one unit in each backward arc. The result is the augmented flow, with a value of 4, shown in Figure 2.3. Note that the conservation law (2.2) is again satisfied at each internal node.

An (s,t)-*cutset* is identified with a pair (S,T) of complementary subsets of nodes, with $s \in S$ and $t \in T$ and consists of all arcs extending between S and T (in either direction). The *capacity* of the cutset (S,T) is defined as

$$c(S,T) = \sum_{i \in S} \sum_{j \in T} c_{ij},$$

i.e. the sum of the capacities if all arcs which are directed from S to T.

The value of any (s,t)-flow cannot exceed the capacity of any (s,t)-cutset. Suppose $x = (x_{ij})$ is a flow and (S,T) is a cutset. Sum the equations (2.2) identified with nodes $i \in S$ to obtain

$$(2.3) \qquad v = \sum_{i \in S} \left(\sum_j x_{ij} - \sum_j x_{ji}\right) = \sum_{i \in S} \sum_{j \in S} (x_{ij} - x_{ji}) + \sum_{i \in S} \sum_{j \in T} (x_{ij} - x_{ji}) = \sum_{i \in S} \sum_{j \in T} (x_{ij} - x_{ji}).$$

That is, the value v of any flow is equal to the net flow through any cutset. But $x_{ij} \leq c_{ij}$ and $x_{ji} \geq 0$, so

$$(2.4) \qquad v \leq \sum_{i \in S} \sum_{j \in T} c_{ij} = c(S,T).$$

In the case of the augmented flow shown in Figure 2.3, there is an (s,t)-cutset with capacity equal to the flow value. For example, $S = \{1,2\}$, $T = \{3,4,5,6\}$. It follows from the preceding analysis that the flow is maximal and that the cutset has minimal capacity. Notice that each arc (i,j), is *saturated*, i.e. $x_{ij} = c_{ij}$, if $i \in S$, $j \in T$ and *void*, i.e. $x_{ij} = 0$, if $i \in T$, $j \in S$.

We now state and prove three of the principal theorems of network flow theory. They will later be applied to prove other combinatorial results and to yield good algorithms for maximal flow problems.

THEOREM 2.1 (Augmenting Path Theorem). *A flow is maximal if and only if it admits no augmenting path.*

PROOF. Clearly, if an augmenting path exists the flow is not maximal. Suppose x is a flow that does not admit an augmenting path. Let S be the set of all nodes j (including s) for which there is an augmenting path from s to j, and let T be the complementary set. From the definition of augmenting path and from the definition of S and T, it follows that for all $i \in S$ and $j \in T$, $x_{ij} = c_{ij}$ and $x_{ji} = 0$. It follows from (2.3) that $v = \sum_{i \in S} \sum_{j \in T} c_{ij}$, the capacity of the cutset (S,T). From (2.4) it follows that the flow is maximal. □

THEOREM 2.2 (Integral Flow Theorem). *If all arc capacities are integers there is a maximal flow which is integral.*

PROOF. Suppose all capacities are integers and let $x^0_{ij} = 0$, for all i and j. If the flow $x^0 = (x^0_{ij})$ is not maximal it admits an augmenting path and hence there is an integral flow x^1 whose value exceeds that of x^0. If x^1 is not maximal it admits an augmenting path, and so on. As each flow obtained in this way exceeds the value of its predecessor by at least one, we arrive eventually at an integral flow that admits no augmenting path and hence is maximal. □

THEOREM 2.3 (Max-Flow Min-Cut Theorem). *The maximum value of an* (s,t)*-flow is equal to the minimum capacity of an* (s,t)*-cutset.*

PROOF. The proofs of the previous two theorems, together with (2.4), are sufficient to establish the max-flow min-cut result for networks in which all capacities are integers and hence for those in which all capacities are commensurate (that is, there exists some $c > 0$ such that every c_{ij} is an integral multiple of c).

To complete the proof of the max-flow min-cut result, we must show that every network actually admits a maximal flow. (Note that the existence of a minimum capacity cutset is not open to question. There are only a finite number of (s,t) cutsets, and at least one of them must be minimal.) We shall present an algorithm for computing maximal flows in the next section, and (although we do not provide the proof here) it can be proved that the algorithm always obtains a maximal flow in a finite number of steps, for any real number capacities. This demonstration is sufficient to complete the proof. □

3. MAXIMAL FLOW ALGORITHM

The problem of finding a maximum capacity flow augmenting path is evidently quite similar to the problem of finding a shortest path, or more precisely, a path in which the minimum arc length is maximum.

We propose a procedure in which labels are given to nodes. These labels are of the form (i^+,δ_j) or (i^-,δ_j). A label (i^+,δ_j) indicates that there exists an augmenting path with capacity δ_j from the source to the node j in question, and that (i,j) is the last arc in this path. A label (i^-,δ_j) indicates that (j,i) is the last arc in the path, i.e. (j,i) will be a backward arc if the path is extended to the sink t. Initially only the source node s is labeled with the special label $(-,\infty)$. Thereafter, additional nodes are labeled in one or the other of two ways:

If node i is labeled and there is an arc (i,j) for which $x_{ij} < c_{ij}$ then the unlabeled node j can be given the label (i^+,δ_j), where $\delta_j =$ $= \min\{\delta_i, c_{ij}-x_{ij}\}$.

If node i is labeled and there is an arc (j,i) for which $x_{ji} > 0$, then the unlabeled node j can be given the label (i^-,δ_j), where $\delta_j = \min\{\delta_i, x_{ji}\}$.

When the procedure succeeds in labeling node t, an augmenting path has been found and the value of the flow can be augmented by δ_t. If the procedure concludes without labeling node t, then no augmenting path exists. A minimum capacity cutset (S,T) is constructed by letting S contain all labeled nodes and T contain all unlabeled nodes.

A labeled node is either "scanned" or "unscanned". A node is scanned by examining all incident arcs and applying labels to previously unlabeled adjacent nodes, according to the rules given above.

MAXIMAL FLOW ALGORITHM

Step 0 (Start). Let $x = (x_{ij})$ be any integral feasible flow, possibly the zero flow. Give node s the permanent label $(-,\infty)$.

Step 1 (Labeling and Scanning).

1.1. If all labeled nodes have been scanned, go to Step 3.

1.2. Find a labeled but unscanned node i and scan it as follows: For each arc (i,j), if $x_{ij} < c_{ij}$ and j is unlabeled, give j the label (i^+,δ_j), where $\delta_j = \min\{c_{ij}-x_{ij},\delta_i\}$. For each arc (j,i), if $x_{ji} > 0$ and j is unlabeled, give j the label (i^-,δ_j), where $\delta_j = \min\{x_{ji},\delta_i\}$.

1.3. If node t has been labeled, go to Step 2; otherwise go to Step 1.1.

Step 2 (Augmentation). Starting at node t, use the index labels to construct an augmenting path. (The label on node t indicates the second-to-last node in the path, the label on that node indicates the third-to-last node, etc.) Augment the flow by increasing and decreasing the arc flows by δ_t, as indicated by the superscripts on the index labels. Erase all labels, except the label on node s. Go to Step 1.

Step 3 (Construction of Minimal Cut). The existing flow is maximal. A cutset of minimum capacity is obtained by placing all labeled nodes in S and all unlabeled nodes in T. The computation is completed.

4. COMBINATORIAL IMPLICATIONS OF MAX-FLOW MIN-CUT THEOREM

A number of combinatorial results can be viewed as consequences of the max-flow min-cut theorem. In order to show this, it is helpful to provide a generalization of the original theorem.

Let us consider a flow network in which there are arc capacities $c_{ij} \geq 0$ and, in addition, node capacities $c_i \geq 0$. Flows are required to satisfy not only the conservation conditions and arc constraints ($0 \leq x_{ij} \leq$ $\leq c_{ij}$) but also the node constraints,

$$\sum_j x_{ij} \leq c_i, \; i \neq s, \; t.$$

That is, the outflow (and hence the inflow) at any interior node does not exceed the capacity of the node. (If all node capacities are infinite, the situation is as before.)

It is natural to impose node capacities in certain applications. For example, nodes might be points of trans-shipment (transportation of goods), supply points (movement of troops), cleansing stations (overland pipelines) or relay stations (communication networks).

For a node having node capacities as well as arc capacities, we define an (s,t)-*cut* as a set of arcs and nodes such that any path from s to t uses at least one member of the set. The *capacity* of a cut is the sum of the capacities of its members.

As this notion of an (s,t)-cut appears to be different from the previous one of an (s,t)-cutset, it is necessary to show that in a network whose node capacities are all infinite, the minimum cut capacity in the new sense is equal to the minimum cutset capacity in the old sense. Let (S,T) be a cutset and let C be the set of all arcs directed from a node in S to a node in T. Then C is a cut in the new sense and its capacity is equal to that of (S,T). Let C be a cut, consisting entirely of arcs, let S be the set of all nodes that can be reached by directed paths from s not using any member of C and let T be the remaining nodes. Then (S,T) is a cutset and C contains every arc from S to T, so the capacity of (S,T) is at most that of C.

THEOREM 4.1 (Generalized Max-Flow Min-Cut Theorem). *In a network having node capacities as well as arc capacities, the maximum value of an* (s,t)-*flow is equal to the minimum capacity of an* (s,t)-*cut. Moreover, if all capacities are integers, there is a maximal flow that is integral.*

PROOF. Expand the network by replacing each interior node i by an *in-node* i', an *out-node* i'', and an arc (i',i'') of capacity c_i. For each arc (i,j) of the original network, there is an arc (i'',j') of capacity c_{ij} in the expanded network. (Let $s' = s'' = s$, $t' = t'' = t$.) An example of such an expansion is shown in Figure 4.1.

In the expanded network, nodes are uncapacitated and hence the original version of the max-flow min-cut theorem applies. As all flow entering i' must go to i'', and all flow leaving i' must come from i', there is a natural one-one correspondence between flows in one network and flows in the other. The theorem follows readily by applying the original max-flow min-cut theorem to the expanded network. □

A celebrated result of graph theory, and a precursor of many other duality theorems, is a theorem of K. Menger. This theorem was originally stated in terms of undirected graphs, but for convenience we give a formu-

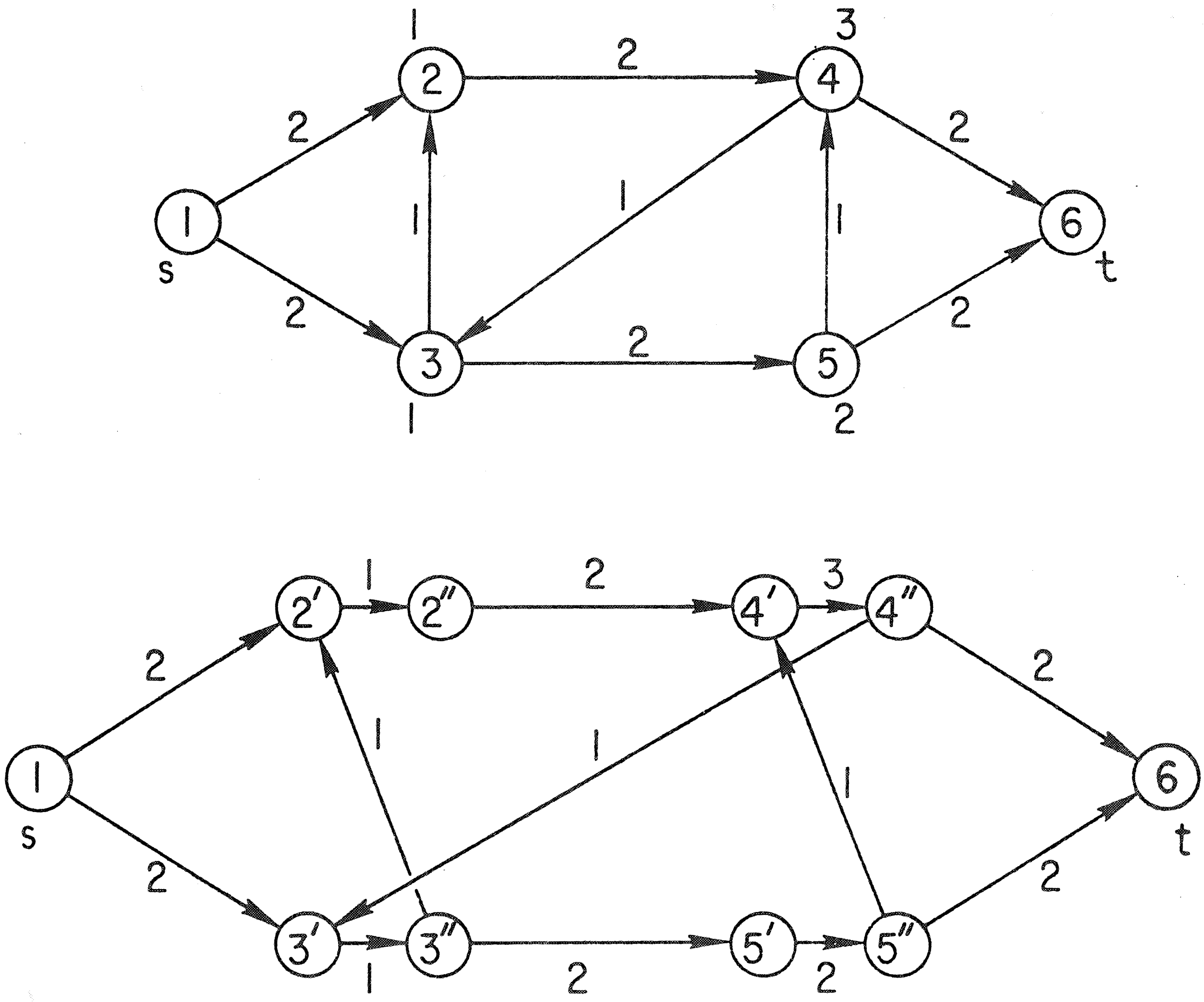


Figure 4.1. Example Network.

lation in terms of digraphs.

A digraph G is said to be k-*connected from* s *to* t if for any set C of k - 1 nodes missing s and t there is a directed path from s to t missing C. In other words, it is not possible to disconnect s from t by removing any fewer than k nodes.

Two (s,t) paths are said to be *independent* if they have no nodes in common except s and t.

<u>THEOREM 4.2</u> (Menger). *If digraph* G *is* k-*connected from* s *to* t *and does not contain arc* (s,t), *then* G *admits* k *independent directed paths from* s *to* t.

PROOF. Give each node a capacity of one and each arc an infinite capacity. Because of the nonexistence of arc (s,t), the minimum cut capacity is finite. From the k-connectivity of the digraph, it follows that the minimum cut capacity is at least k.

From Theorem 4.1, it follows that there is an integral maximal flow of value at least k. The structure of the flow network is such that this flow yields k pairwise independent directed paths from s to t and the theorem is proved. □

Although network flow theory appears to be concerned solely with digraphs, it also yields a good deal of information about the structure of undirected graphs.

THEOREM 4.3. *The maximum number of arc-disjoint* (s,t) *paths in an undirected graph* G *is equal to the minimum number of arcs in an* (s,t)*-cutset.*

PROOF. Construct from G a flow network in which for each arc of G there is a symmetric pair of arcs (i,j) and (j,i), each with unit capacity. There exists an integral maximal (s,t) flow in which at least one arc of each symmetric pair is void. Accordingly, such a flow yields a maximum number of disjoint (s,t) paths, in G. Application of the max-flow min-cut theorem completes the proof. □

By applying Theorem 4.3 to the dual of G and reinterpreting the results in the original graph, we obtain the following.

THEOREM 4.4. *If* G *is* (s,t) *planar, then the minimum number of arcs in an* (s,t) *path is equal to the maximum number of disjoint* (s,t)*-cutsets.*

5. LINEAR PROGRAMMING INTERPRETATION OF MAX-FLOW MIN-CUT THEOREM

The max-flow min-cut theorem can be viewed as a consequence of linear programming duality. The primal linear programming problem is:

$$\text{maximize } v$$

subject to

$$\sum_j x_{ji} - \sum_j x_{ij} = \begin{cases} -v, & i = s \\ 0, & i \neq s, t \\ +v, & i = t. \end{cases}$$

$$x_{ij} \leq c_{ij}$$

$$x_{ij} \geq 0.$$

Let u_i be a dual variable identified with the i^{th} node equation and w_{ij} be a dual variable identified with the capacity constraint on arc (i,j). Then the dual problem is

$$(5.1)\quad \begin{cases} \text{minimize} \\ \quad \sum_{i,j} c_{ij} w_{ij} \\ \text{subject to} \\ \quad u_j - u_i + w_{ij} \geq 0 \\ \quad u_s - u_t \geq 1 \\ \quad w_{ij} \geq 0 \\ \quad u_i \quad \text{unrestricted} \end{cases}$$

For any (s,t)-cutset there is a feasible solution to the dual problem whose value is equal to the capacity of the cutset. Let (S,T) be such a cut-set, and let

$$\begin{aligned} u_i &= 1, && \text{if } i \in S \\ &= 0, && \text{if } i \in T \\ w_{ij} &= 1, && \text{if } i \in S,\ j \in T \\ &= 0, && \text{otherwise.} \end{aligned}$$

Moreover, there is an optimal solution to the dual problem which corresponds to an (s,t)-cutset. For such an optimal solution, we may assume that $u_t = 0$. This is equivalent to dropping the redundant equation for node t from the primal problem. Also assume $u_s = 1$. (The reader can verify that there is no reason for u_s to be greater.) Then the remaining variables are forced to take on 0,1 values. For each arc (i,j), it is the case that $w_{ij} = 1$ if and only if $u_i = 1$ and $u_j = 0$. (Note that $c_{ij} > 0$.) Then let

$$S = \{i \mid u_i = 1\},$$

$$T = \{j \mid u_j = 0\}.$$

The capacity of the cutset (S,T) is exactly equal to the value of the optimal dual solution.

Thus, the dual problem, in effect, finds a minimum capacity (s,t)-cutset. The max-flow min-cut theorem follows immediately from the well-known fact that the optimal values of the objective functions for dual linear programming problems are equal.

It is also a well-known result of duality theory that primal and dual solution are optimal if and only if

$$x_{ij} > 0 \Rightarrow u_j - u_i + w_{ij} = 0$$

$$w_{ij} > 0 \Rightarrow x_{ij} = c_{ij}.$$

Suppose we view u_i as a "node potential", e.g. altitude or fluid pressure. Then for an optimal pair of primal and dual solutions exactly one of three cases exists for each arc (i,j):

Case 1. The potential at i is less than at j. There is zero flow in (i,j).

Case 2. The potential at i is equal to that at j. There may or may not be positive flow in (i,j).

Case 3. The potential at i is greater than at j. The flow in (i,j) is equal to its capacity c_{ij}.

These conditions correspond very well indeed with our intuitive notion of the relationships that should exist between node potentials and arc flows.

These ideas, in generalized form, are the basis for the out-of-kilter method presented later.

It is just as important to be able to recognize combinatorial problems which can be formulated as min-cut problems as it is to be able to recognize those which can be formulated in max-flow form. Generally speaking, one should be on the look out for problems with constraints involving sums or differences of pairs of variables. The problem below is an excellent example.

A PROVISIONING PROBLEM

In the well-known "knapsack" problem, it is assumed that the benefit to be gained from the selection of any given item is independent of the selection of the other items. This is clearly a simplistic view of utility. For example, the benefit to be gained from a gas stove without fuel is rather small.

A more sophisticated view can be taken, as follows. Suppose there are n items to choose from among, where item j costs $c_j > 0$ dollars. Also suppose there are m sets of items, $S_1, S_2, \ldots, S_m$, which are known to confer special benefits. If all of the items in set S_i are chosen, then a benefit of $b_i > 0$ dollars is gained. The sets are arbitrary and need not be related in any particular way, e.g. a given item may be contained in several different sets.

There is no restriction on the number of items that can be purchased, i.e. there is no limiting knapsack. Our objective is simply to maximize net benefit, i.e. total benefit gained minus total cost of items purchased.

Even without any constraints on the selection of items the problem appears to be unreasonably difficult. Yet it can be cast into the mold of a min-cut problem and can therefore be solved quite easily.

Let

$$\begin{aligned} v_j &= 1 \qquad \text{if item j is purchased} \\ &= 0 \qquad \text{otherwise,} \end{aligned}$$

and let

$u_i = 1$ if all of the items in set S_i are purchased
$\quad = 0$ otherwise.

Then the problem is to

maximize

(5.2) $$Z = \sum_i b_i u_i - \sum_j c_j v_j$$

subject to

(5.3) $$v_j - u_i \geq 0 \qquad \text{for each pair } i,j \text{ such that } j \in S_i,$$

and

$$u_i, v_j \in \{0,1\}.$$

Because of the 0,1 restrictions on the variables and constraints (5.3), it is not possible for a benefit b_i to be earned unless all items j in the set S_i are purchased.

Let us complexify matters by introducing $m + n$ new variables, $w_1, w_2, \ldots, w_m$, and $z_1, z_2, \ldots, z_n$.

Consider the problem:

minimize

(5.4) $$z = \sum_i b_i w_i + \sum_j c_j z_i$$

subject to

$$v_j - u_i \geq 0, \qquad j \in S_i$$

(5.5) $$u_i + w_i \geq 1, \qquad i = 1,2,\ldots,m$$

(5.6) $$-v_j + z_j \geq 0, \qquad j = 1,2,\ldots,n$$

$$u_i, v_j, w_i, z_j \in \{0,1\}.$$

Suppose $\bar{u} = (\bar{u}_i)$, $\bar{v} = (v_j)$ is a feasible solution to the original problem. Let $\bar{w} = (1-\bar{u}_i)$, $\bar{z} = \bar{v}$. Then $\bar{u}$, $\bar{v}$, $\bar{w}$, $\bar{z}$ is a feasible solution to the new problem. Moreover,

$$z = \sum_i b_i\bar{w}_i + \sum_j c_j\bar{z}_j = \sum_i b_i(1-\bar{u}_i) + \sum_j c_j\bar{v}_j = \sum_i b_i - Z.$$

Now suppose $\bar{u}$, $\bar{v}$, $\bar{w}$, $\bar{z}$ is a minimal solution to the new problem. From (5.5) and $b_i > 0$ is follows that $\bar{w}_i = 1 - \bar{u}_i$. From (5.6) and $c_j > 0$ it follows that $\bar{z}_j = \bar{v}_j$. Clearly $\bar{u},\bar{v}$ is a feasible solution to the original problem and again $z = \sum_i b_i - Z$.

It follows that a minimal solution to the new problem yields a maximal solution to the original problem.

We need to make a few more changes to put the problem into the form of a min-cut problem. We introduce two new variables u_0 and v_{n+1} and mn new variables w_{ij}. Let K be a large number. Consider the problem

$$(5.7)\quad \begin{cases} \text{minimize} & \\ \quad z = \sum_i b_iw_i + \sum_j c_jz_j + \sum_{i,j} Kw_{ij} & \\ \text{subject to} & \\ \quad v_j - u_i + w_{ij} \geq 0, & j \in S_i \\ \quad u_i - u_0 + w_i \geq 0, & i = 1,2,\ldots,m \\ \quad v_{n+1} - v_j + z_j \geq 0, & j = 1,2,\ldots,n \\ \quad u_0 - v_{n+1} \geq 1 & \\ \quad u_i,v_j,w_i,z_j,w_{ij} \in \{0,1\}. & \end{cases}$$

These changes make no essential difference in the problem. Because u_0 and v_{n+1} are restricted to 0,1 values, the constraint $u_0 - v_{n+1} \geq 1$ can be satisfied if and only if $u_0 = 1$, $v_{n+1} = 0$. If K is a sufficiently large number, all the variables w_{ij} are zero in a minimal solution.

Except for the 0,1 restrictions on the variables, (5.7) is in the same form as the min-cut problem (5.1). There is only a superficial difference in the designations of variables and their indices. But we know that problem (5.1) admits an optimal solution with 0,1 values for its variables. It follows that we can drop the 0,1 restrictions from (5.7), retaining only nonnegativity constraints on w_i, z_j, w_{ij}.

The network for the min-cut formulation of the provisioning problem is shown in Figure 5.1.

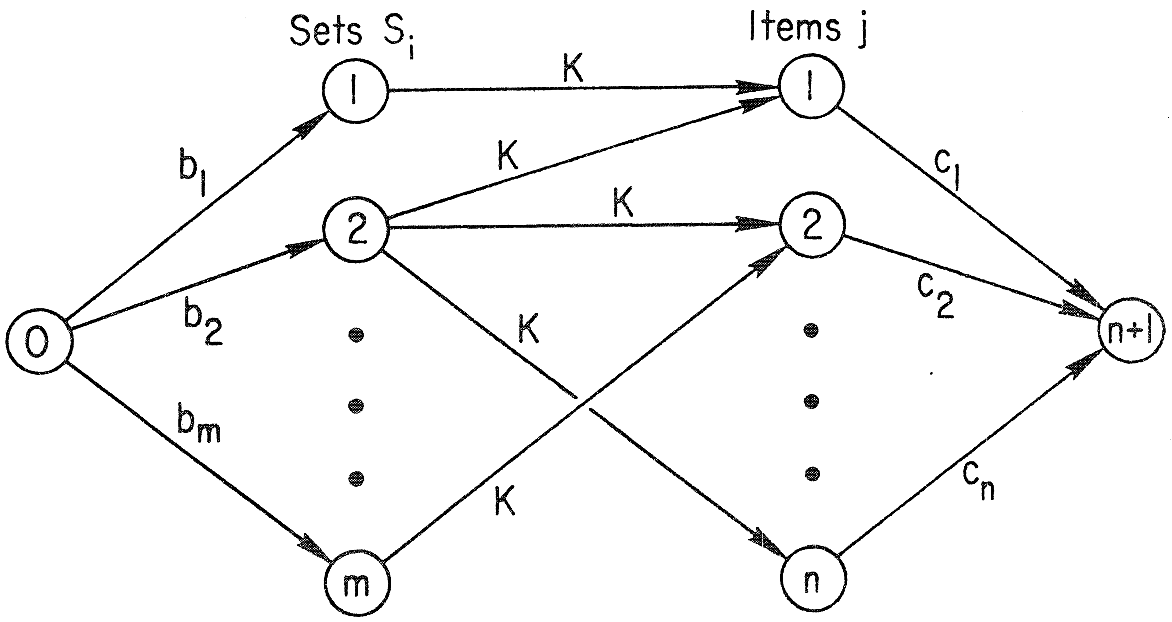


Figure 5.1. Network for Provisioning Problem.

6. MINIMUM COST FLOWS

Suppose in addition to a capacity c_{ij}, each arc of a flow network is assigned a *cost* a_{ij}. The *cost of a flow* $x = (x_{ij})$ is

$$\sum_{i,j} a_{ij} x_{ij}.$$

We now pose the problem of finding a minimum cost flow for a given flow value v.

ASSIGNMENT PROBLEM

There are n men and n jobs. The cost of assigning man i to job j is a_{ij}. For what man-job assignment is the total cost minimized?

Construct a directed bipartite graph with n nodes in each of its parts, and give arc (i,j) cost a_{ij} and infinite capacity. Add a source node s with an arc (s,i) to each node in the first part, and a sink node t with an arc (j,t) from each node in the second part. Set $c_{si} = 1$, $a_{si} = 0$, for all i, and $c_{jt} = 1$, $a_{jt} = 0$, for all j. A minimum cost integral flow of value n yields a solution to the problem.

The flow network for the assignment problem is shown in Figure 6.1. The first number of each arc represents its capacity and the second number is its cost.

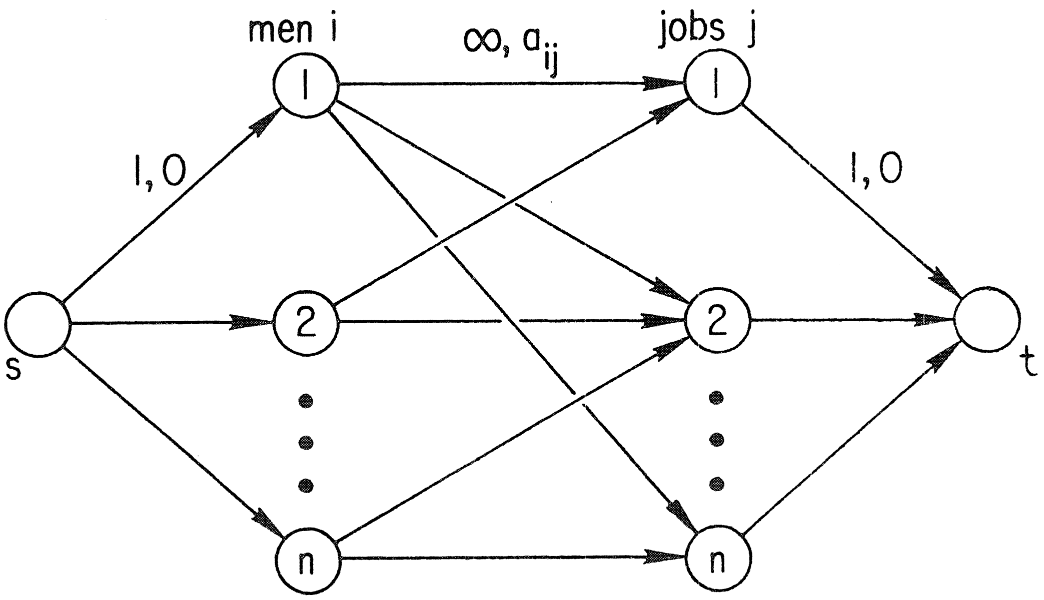


Figure 6.1. Network for Assignment Problem.

Let us define the *cost of an augmenting path* to be the sum of the costs of forward arcs minus the sum of costs of backward arcs. Thus the cost of a path is equal to the net change in the cost of flow for one unit of augmentation along the path. An *augmenting cycle* is a closed augmenting path. The cost of an augmenting cycle is computed in the obvious way, with respect to a given orientation of the cycle, i.e. clockwise or counterclockwise.

THEOREM 6.1. *A flow of value* v *is of minimum cost if and only if it admits no flow augmenting cycle with negative cost.*

PROOF. The only if part of the theorem is obvious. For the converse, suppose that $x^0 = (x^0_{ij})$ and $x^1 = (x^1_{ij})$ are two flows, both of value v, where x^0 is less costly than x^1. The difference between these two flows, $y = x^0 - x^1$, can be expressed as a sum of flow augmenting cycles with respect to x^1. Because of the cost of x^0 is less than that of x^1, at least one of these cycles must have negative cost. □

THEOREM 6.2 (Busacker and Gowan). *The augmentation by* δ *of a minimum cost flow of value* v *along a minimum cost flow augmenting path yields a minimum cost flow of value* $v + \delta$.

PROOF. By Theorem 6.1, it suffices to show that the flow resulting from augmentation along a minimum cost augmenting path does not admit a negative augmenting cycle. Suppose such a cycle C were introduced. Then C must contain at least one arc (i,j) of the minimum cost augmenting path P. But then $P \cup C - (i,j)$, or some subset of it, would be an augmenting path with respect to the original flow, and would be less costly than P, contrary to the assumption that P is minimal. □

A minimum-cost augmenting path can be found by means of a shortest path computation. Specifically, for a given flow $x = (x_{ij})$ and arc costs a_{ij}, let

$$\bar{a}_{ij} = \begin{cases} a_{ij}, & \text{if } x_{ij} < c_{ij},\ x_{ji} = 0 \\ \min\{a_{ij}, -a_{ji}\}, & \text{if } x_{ij} < c_{ij},\ x_{ji} > 0 \\ -a_{ji}, & \text{if } x_{ij} = c_{ij},\ x_{ji} > 0 \\ +\infty, & \text{if } x_{ij} = c_{ij},\ x_{ji} = 0, \end{cases}$$

where we understand that $a_{ij} = +\infty$ if (i,j) is not an arc of the flow network. A shortest (s,t) directed path with respect to arc lengths $\bar{a}_{ij}$ corresponds to a minimum-cost (s,t) augmenting path. A negative directed cycle corresponds to an augmenting cycle with negative cost.

We can now outline an algorithm for solving the minimum cost flow problem. This algorithm combines ideas of Klein and of Busacker and Gowen.

MINIMUM-COST FLOW ALGORITHM

Step 0 (Start). Let $x = (x_{ij})$ be any (s,t) flow with value $v' \leq v$, where v is the desired flow value. This initial flow can be the zero flow, or a flow of value v, perhaps determined by the max-flow algorithm. Or if a flow $x' = (x'_{ij})$ of value $v' > v$ is known, one can let $x = (v/v')\, x'$.

Step 1 (Elimination of Negative Cycles).

1.1. Apply a shortest path algorithm with respect to arc lengths $\bar{a}_{ij}$ with the objective of detecting negative cycles. If no negative cycle exists, go to Step 2.

1.2. Augment the flow around the corresponding augmenting cycle to obtain a less costly flow of the same value v', then return to Step 1.1.

Step 2 (Minimum Cost Augmentation).

2.0. If the existing flow value $v' = v$, the existing flow is optimal and the computation is completed. Otherwise, proceed to step 2.1.

2.1. Apply a shortest path algorithm with respect to arc lengths $\bar{a}_{ij}$ with the objective of finding shortest path from s to t. If no shortest path exists, there is no flow of value v and the computation is halted.

2.2. Augment the flow by δ, where $v' + \delta \leq v$, along a minimum-cost (s,t) augmenting path as determined by the shortest path computation. Return to Step 2.0.

Note that the procedure has two phases. In the first phase negative cycles are eliminated and in the second phase a succession of minimum-cost augmentations are made, until the desired flow value v is achieved. If one begins with the zero flow and no negative cycles exist with respect to the arc costs a_{ij}, then at most v augmentations are required, provided all capacities are integers. Each augmentation requires a shortest path computation.

7. LOWER BOUNDS AND CIRCULATIONS

Some combinatorial problems can be successfully formulated as network flow problems only if lower bounds on arc flow are imposed. That is, in addition to a capacity c_{ij} for each arc (i,j) we may designate a *lower bound* ℓ_{ij} and require that $\ell_{ij} \le x_{ij} \le c_{ij}$.

As an example, consider the following problem.

AIRCRAFT SCHEDULING

An airline wishes to use the smallest possible number of planes to meet a fixed flight schedule. A digraph is constructed with two nodes i,i' and an arc (i,i') for each flight. An arc (i',j) is provided if it is feasible for a plane to return from the destination of flight i to the starting point for flight j and be ready in time for its scheduled departure. (Planes are assumed to be identical and capable of making any of the flights.) In addition, there are dummy nodes s' and t, with arcs (s',i) and (i',t), for all i and i'.

Set $\ell_{ii'} = c_{ii'} = 1$, for all arcs (i,i') and $\ell_{i',j} = 0$, $c_{i',j} = 1$ for all other arcs (i',j). The minimum number of airplanes required to meet the flight schedule is determined by an integral (s',t) flow of *minimum* value.

Up to this point in our study of network flows we have not had to be concerned with the existence of feasible flows. The zero flow, if no other, always satisfied arc capacity constraints. Now, however, the nonexistence of a feasible flow is a distinct possibility. For example, a network with only two arcs, (s,1), (1,t), with $c_{s1} < \ell_{1t}$, has no feasible (s,t) flow.

It is useful to approach the feasibility problem through the study of "circulations". A *circulation* is simply a flow in a network in which conservation conditions are observed at all nodes. That is, there is no source or sink.

To convert a conventional flow problem to circulation form, add an arc (t,s) to the network, with $\ell_{ts} = 0$, $c_{ts} = +\infty$. Then a maximal (s,t) flow is simply a circulation for which x_{ts} is maximum.

Here is how to find a feasible circulation in a network with both lower bound and capacities, if such a circulation exists. Begin with the zero circulation. If all lower bounds are zero, this circulation is feasible. Otherwise, find an arc (p,q) for which $x_{pq} < \ell_{pq}$. Construct a flow augmenting path from q to p where this path is of the conventional type, except that we require $x_{ij} > \ell_{ij}$ for each backward arc and δ is chosen such that $\delta \leq \ell_{ij} - x_{ij}$. Augment the flow from q to p by δ, and repeat until $x_{pq} \geq \ell_{pq}$. Then repeat for another arc for which the arc flow is infeasible. Eventually a feasible circulation is obtained, if the network admits such a circulation.

But suppose at some point an augmenting path cannot be found. Let (t,s), with $x_{ts} < \ell_{ts}$, be the arc for which the augmenting path cannot be found. Let S be the set of nodes which can be reached from s by an augmenting path, and T those which cannot. For each arc (i,j) directed from S to T, $x_{ij} = c_{ij}$, and for each arc directed from T to S, $x_{ij} \leq \ell_{ij}$. See Figure 7.1.

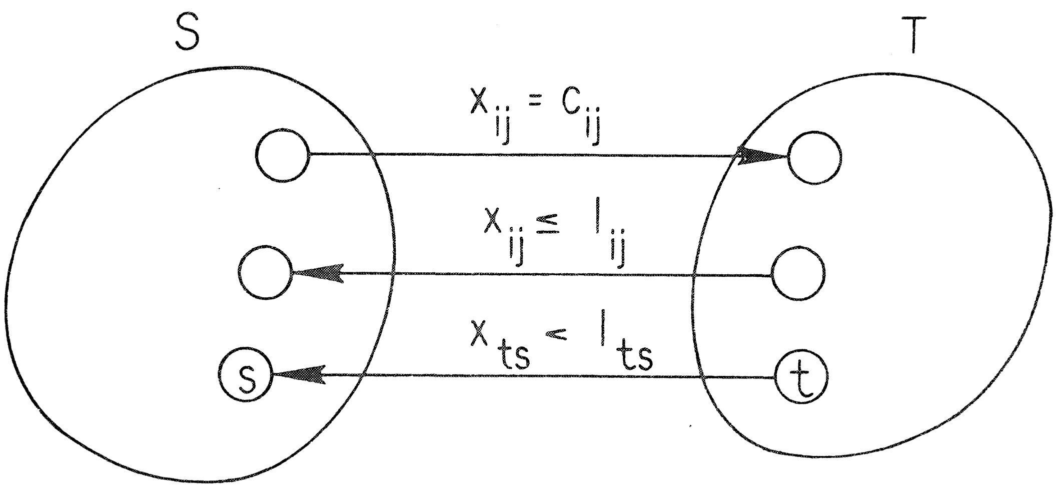


Figure 7.1. Infeasibility of Circulation.

The net flow across the cutset (S,T) is zero, i.e.:

$$\sum_{i\in S, j\in T} x_{ij} = \sum_{i\in T, j\in S} x_{ij}.$$

But

$$\sum_{i\in S, j\in T} x_{ij} = \sum_{i\in S, j\in T} c_{ij}$$

and

$$\sum_{i\in S, j\in T} x_{ij} < \sum_{i\in T, j\in S} \ell_{ij},$$

with strict inequality because of arc (t,s). We have constructed a cutset (S,T) for which

$$\sum_{i\in T, j\in S} c_{ij} < \sum_{i\in S, j\in T} \ell_{ij}.$$

We have thus proved the following theorem.

THEOREM 7.1. *In a network with lower bounds and capacities a feasible circulation exists if and only if*

$$\sum_{i\in S, j\in T} \ell_{ij} \le \sum_{i\in T, j\in S} c_{ij} \tag{7.1}$$

for all cutsets (S,T).

COROLLARY 7.2 (Generalized Max-Flow Min-Cut Theorem). *Let* G *be a flow network with lower bounds and capacities and which admits a feasible* (s,t) *flow. The maximum value of an* (s,t) *flow in* G *is equal to the minimum capacity of an* (s,t) *cutset, where the capacity of cutset* (S,T) *is defined as*

$$c(S,T) = \sum_{i\in S, j\in T} c_{ij} - \sum_{i\in T, j\in S} \ell_{ij}.$$

PROOF. Convert the flow problem to circulation form by adding an arc (t,s) to the network, with $\ell_{ts} = v$, $c_{ts} = +\infty$. Because a feasible (s,t) flow exists in the original network, a feasible circulation exists in the new network for sufficiently small (s,t) flow values v. By Theorem 7.1, the largest value of v for which there exists a feasible circulation is that which satisfies the inequalities (7.1) for all (s,t) cutsets, with strict equality in the case of at least one (s,t) cutset. But this value of v is precisely the minimum capacity of an (s,t) cutset, as defined in the statement of the theorem. □

As we noted in the statement of the aircraft scheduling problem, it is sometimes desired to find a *minimum* value flow, rather than a maximum value flow.

COROLLARY 7.3 (Min-Flow Max-Cut Theorem). *Let* G *be a flow network with lower bounds and capacities and which admits a feasible* (s,t) *flow. The minimum value of an* (s,t) *flow in* G *is equal to the maximum of*

$$\sum_{i\in S,\, t\in T} \ell_{ij} - \sum_{i\in T,\, j\in S} c_{ij}$$

over all (s,t) *cutsets* (S,T), *or equivalently, the negative of the minimum capacity of a* (t,s) *cutset.*

PROOF. Repeat the construction for the preceding corollary, this time letting $\ell_{ts} = 0$, $c_{ts} = v$. □

We can use Corollary 7.3 to prove a well-known theorem of Dilworth. This theorem concerns the minimum number of paths in an acyclic directed graph which are sufficient to cover a specified subset of arcs. (A set of paths "covers" a set of arcs A if each arc in A in contained in at least one path.)

THEOREM 7.4 (Dilworth). *Let* G *be an acyclic directed graph and let* A *be a subset of its arcs. The minimum number of directed paths required to cover the arcs in* A *is equal to the maximum number of arcs in* A, *no two of which are contained in a directed path in* G.

PROOF. Add nodes s and t to G, and arcs (s,i), (i,t), for all i ≠ s,t. For each arc $(i,j) \in A$, set $\ell_{ij} = 1$, $c_{ij} = +\infty$, and for all other arcs set $\ell_{ij} = 0$, $c_{ij} = +\infty$. A minimum value (s,t) flow yields the minimum number of directed paths required to cover all the arcs in A. (Note that if the graph contained directed cycles, some of the arcs in A could be covered by flow circulating around those cycles.) Apply Corollary 7.3 and the result follows immediately. □

When the Dilworth Theorem is applied to the aircraft scheduling problem, it yields the result that the minimum number of planes required by the flight schedule is equal to the maximum number of flights, no two of which can be made by the same plane.

Let A be the entire set of arcs of G, apply the Dilworth Theorem to the dual of G and reinterpret the results in the original graph. Then the following theorem is obtained, parallel to Theorem 4.4.

THEOREM 7.5. *If* G *is an acyclic,* (s,t) *planar digraph, then the maximum number of arcs in an* (s,t) *directed path is equal to the minimum number of* (s,t) *directed cutsets covering all the arcs of* G.

8. THE OUT-OF-KILTER METHOD

We shall now describe a general computational procedure, developed independently by Fulkerson and by Minty, for finding minimum cost circulations.

The minimum-cost circulation problem is to

$$(8.1)\quad \begin{cases} \text{minimize} & \\ \quad \sum_{i,j} a_{ij}x_{ij} & \\ \text{subject to} & \\ \quad \sum_j x_{ji} - \sum_j x_{ij} = 0, & \text{all } i \\ \quad 0 \le \ell_{ij} \le x_{ij} \le c_{ij}, & \text{all } i,j. \end{cases}$$

All of the flow problems we have studied so far, and many others, can be cast into the form of (8.1). For example:

MAXIMAL FLOW PROBLEM

To the given flow network with source s and sink t add a return arc (t,s) with $\ell_{ts} = 0$. $c_{ts} = +\infty$ and $a_{ts} = -1$. For all other arcs (i,j), the lower bounds (if any) and capacities are as given and $a_{ij} = 0$. (For a minimum flow problem, set $a_{ts} = 1$.)

MINIMUM COST FLOW PROBLEM

Add a return arc (t,s) with $\ell_{ts} = 0$, $c_{ts} = +\infty$ and $a_{ts} = 0$. The lower bounds, capacities and costs of all other arcs are as given.

FEASIBLE CIRCULATION PROBLEM

Set $a_{ij} = 0$ for all arcs (i,j).

SHORTEST PATH PROBLEM

To find a shortest path from s to t in a network with arc lengths a_{ij}, add a return arc (t,s) with $\ell_{ts} = c_{ts} = 1$. For all other arcs (i,j), let $\ell_{ij} = 0$, $c_{ij} = +\infty$, and a_{ij} is as given.

To find shortest paths from s to all other nodes, add return arcs (j,s) from all nodes $j \neq s$, with $\ell_{js} = c_{js} = 1$.

The out-of-filter algorithm is a primal-dual linear programming method. The problem dual to (8.1) is:

$$
(8.2)\quad \begin{cases}
\text{maximize} \\
\quad \sum_{i,j} \ell_{ij}\lambda_{ij} - \sum_{i,j} c_{ij}\gamma_{ij} \\
\text{subject to} \\
\quad u_j - u_i + \lambda_{ij} - \gamma_{ij} \leq a_{ij} \\
\qquad\qquad \lambda_{ij}, \gamma_{ij} \geq 0 \\
\qquad u_i \text{ unrestricted.}
\end{cases}
$$

The dual variables λ_{ij} and γ_{ij} are identified with the primal constraints $x_{ij} \geq \ell_{ij}$ and $-x_{ij} \geq -c_{ij}$. (The variable γ_{ij} is analogous to w_{ij} in (5.1), but there the primal constraints were of the form $x_{ij} \leq c_{ij}$, hence the change in sign in the inequalities of (8.2).) The dual variables u_i are identified with primal node equations, as in (5.1).

Applying duality theory of linear programming, we obtain the following orthogonality conditions which are necessary and sufficient for optimality of primal and dual solutions:

$$x_{ij} > 0 \Rightarrow u_j - u_i + \lambda_{ij} - \gamma_{ij} = a_{ij}$$

$$\lambda_{ij} > 0 \Rightarrow x_{ij} = \ell_{ij}$$

$$\gamma_{ij} > 0 \Rightarrow x_{ij} = c_{ij}.$$

The nonnegative variables λ_{ij} and γ_{ij} can effectively de dispensed with by noting that the above conditions are equivalent to the following:

$$(8.3) \qquad \begin{cases} x_{ij} = \ell_{ij} \Rightarrow u_j - u_i \le a_{ij} \\ \ell_{ij} < x_{ij} < c_{ij} \Rightarrow u_j - u_i = a_{ij} \\ x_{ij} = c_{ij} \Rightarrow u_j - u_i \le a_{ij} \end{cases}$$

For example, suppose $x = (x_{ij})$ is a primal solution and for some arc (i,j), $0 < \ell_{ij} = x_{ij} < c_{ij}$. Then

$$x_{ij} > 0 \Rightarrow u_j - u_i + \lambda_{ij} - \gamma_{ij} = a_{ij}.$$

But

$$x_{ij} < c_{ij} \Rightarrow \gamma_{ij} = 0$$

and from the nonnegativity of λ_{ij} it follows that $u_j - u_i \le a_{ij}$. A similar analysis of other cases establishes that conditions (8.2) are satisfied if and only if the primal and dual solutions are optimal.

We refer to conditions (8.2) as *kilter conditions* and represent them by a *kilter diagram* for each arc as shown in Figure 8.1. Points (x_{ij}, u_j-u_i) on the crooked line are *in kilter* and those which are not are *out-of-kilter*. To each point (x_{ij}, u_j-u_i) we assign a kilter number $K(x_{ij})$ equal to the absolute value of the change in x_{ij} necessary to bring the arc into kilter. Thus,

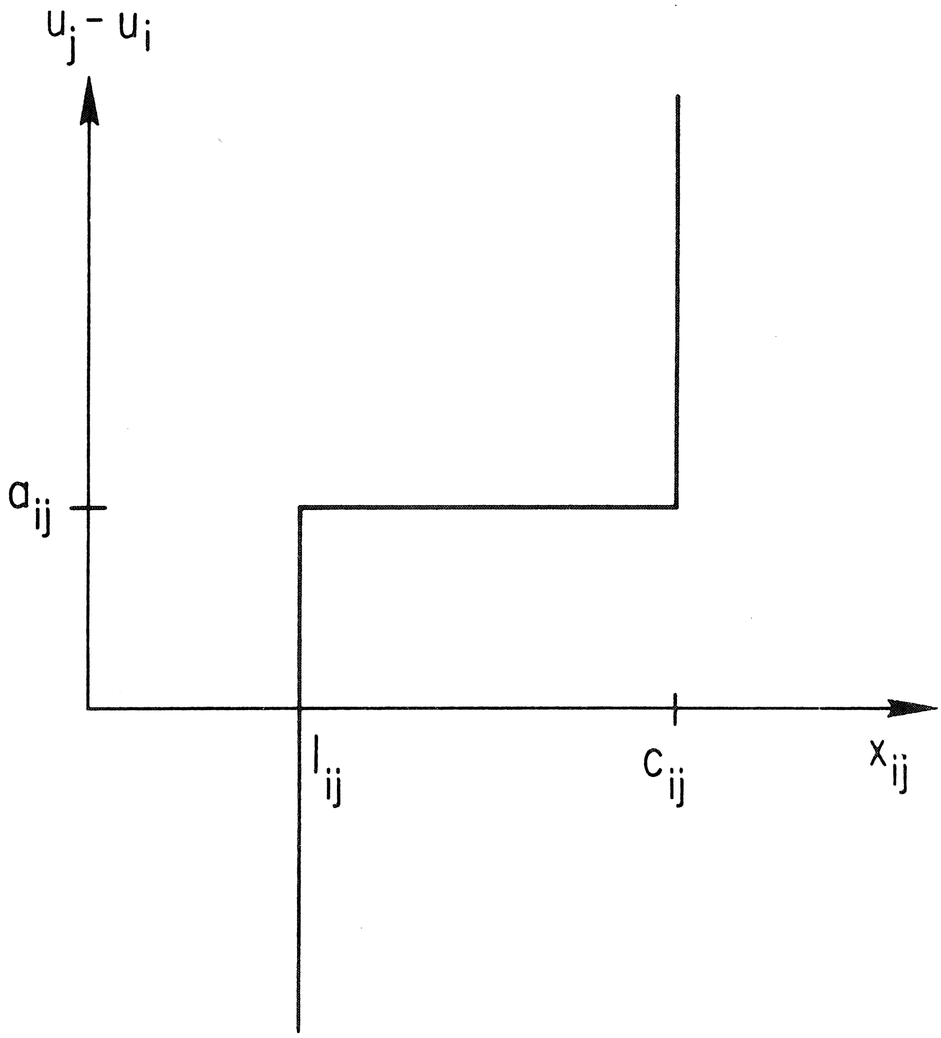


Figure 8.1. Kilter Diagram

$$K(x_{ij}) = \begin{cases} |x_{ij} - \ell_{ij}|, & \text{if } u_j - u_i < a_{ij} \\ \ell_{ij} - x_{ij}, & \text{if } x_{ij} < \ell_{ij},\ u_j - u_i = a_{ij} \\ x_{ij} - c_{ij}, & \text{if } x_{ij} > c_{ij},\ u_j - u_i = a_{ij} \\ 0, & \text{if } \ell_{ij} \le x_{ij} \le c_{ij},\ u_j - u_i = a_{ij} \\ |x_{ij} - c_{ij}|, & \text{if } u_j - u_i > a_{ij}. \end{cases}$$

The objective of the out-of-kilter method is to obtain a circulation

$x = (x_{ij})$ and a set of node numbers $u = (u_i)$ for which the kilter conditions (8.2) are satisfied. As conditions (8.2) are satisfied if and only if all kilter numbers are zero, the sum of the kilter numbers can be used as a measure of progress toward an optimal pair of solutions.

The out-of-kilter computation is begun with any circulation, feasible or not, provided node conservation conditions are satisfied, and with any set of node numbers whatsoever. At each iteration a change is made either in the circulation or in the node numbers. The type of change that is made is determined by the application of Minty's "Painting" Theorem, as described below.

THEOREM 8.1 (Minty). *Let* G *be a directed graph with a distinguished arc* (t,s). *Then for any painting of the arcs green, yellow and red, with* (t,s) *painted yellow, exactly one of the following alternatives holds:*

(1) (t,s) *is contained in a cycle of yellow and green arcs, in which all yellow arcs have the same direction.*

(2) (t,s) *is contained in a cutset of yellow and red arcs, in which all yellow arcs have the same directions.*

PROOF. Think of the graph as a network of streets, in which green arcs are two-way streets, yellow arcs are one-way streets (according to the directions of the arcs), and red arcs are streets blocked to traffic. Now starting at the street intersection represented by node s, either it is possible for traffic to move from s to t, or it is not. If there is some way, then there exists a minimal (s,t) path of yellow and green arcs, with all yellow arcs directed from s to t. This path, together with the arc (t,s), forms a cycle satisfying condition (1).

If there is no way for traffic to get from s to t then a cutset satisfying the description (2) can be constructed as follows. Let S be the set of all nodes accessible to traffic from s and let T be the complementary set. There can be neither yellow arcs directed from S to T nor green arcs between T and S in either direction. Otherwise, one or more of the nodes in S would be accessible to traffic from S, contrary to assumption. It follows that all arcs between T and S must be red arcs, in either direction, or yellow arcs including (t,s) directed from T to S. □

In applying the theorem, we shall color the arcs according to a scheme described below and then focus our attention on an out-of-kilter yellow arc (t,s). Then if we find a yellow-green cycle, we shall modify the circulation around that cycle. If we find a yellow-red cutset, we shall use that cutset as a basis for revising the node numbers.

Here is how we propose to color the arcs, and also change the directions of some of them:

(8.4) Paint an arc *green* if it is in kilter and it is possible to either increase or decrease the arc flow without throwing the arc out of kilter. For such an arc,

$$\ell_{ij} < x_{ij} < c_{ij} \qquad \text{and} \qquad u_j - u_i = a_{ij}.$$

(8.5) Paint an arc *yellow* if it is possible to increase the arc flow, but not to decrease it, without increasing the arc kilter number. For such an arc, either

$$x_{ij} < c_{ij} \qquad \text{and} \qquad u_j - u_i > a_{ij}$$

or

$$x_{ij} \leq \ell_{ij} \qquad \text{and} \qquad u_j - u_i = a_{ij}$$

or

$$x_{ij} < \ell_{ij} \qquad \text{and} \qquad u_j - u_i < a_{ij}.$$

(8.6) Paint an arc *yellow* and also reverse its direction if it is possible to decrease the arc flow, but not to increase it, without increasing the arc kilter number. For such an arc, either

$$x_{ij} > c_{ij} \qquad \text{and} \qquad u_j - u_i > a_{ij}$$

or

$$x_{ij} \geq c_{ij} \qquad \text{and} \qquad u_j - u_i = a_{ij}$$

or

$$x_{ij} > \ell_{ij} \qquad \text{and} \qquad u_j - u_i < a_{ij}.$$

(8.7) Paint an arc *red* if the arc flow can be neither increased nor decreased without increasing the kilter number. For such an arc either

$$x_{ij} = c_{ij} \qquad \text{and} \qquad u_j - u_i > a_{ij}$$

or

$$x_{ij} = \ell_{ij} \qquad \text{and} \qquad u_j - u_i < a_{ij}.$$

These cases account for all possibilities and are summarized in Figure 8.2. Note that all green and red arcs are in kilter. A yellow arc (i,j) is in kilter only if (x_{ij}, u_j-u_i) is a "corner" point in the kilter diagram for the arc.

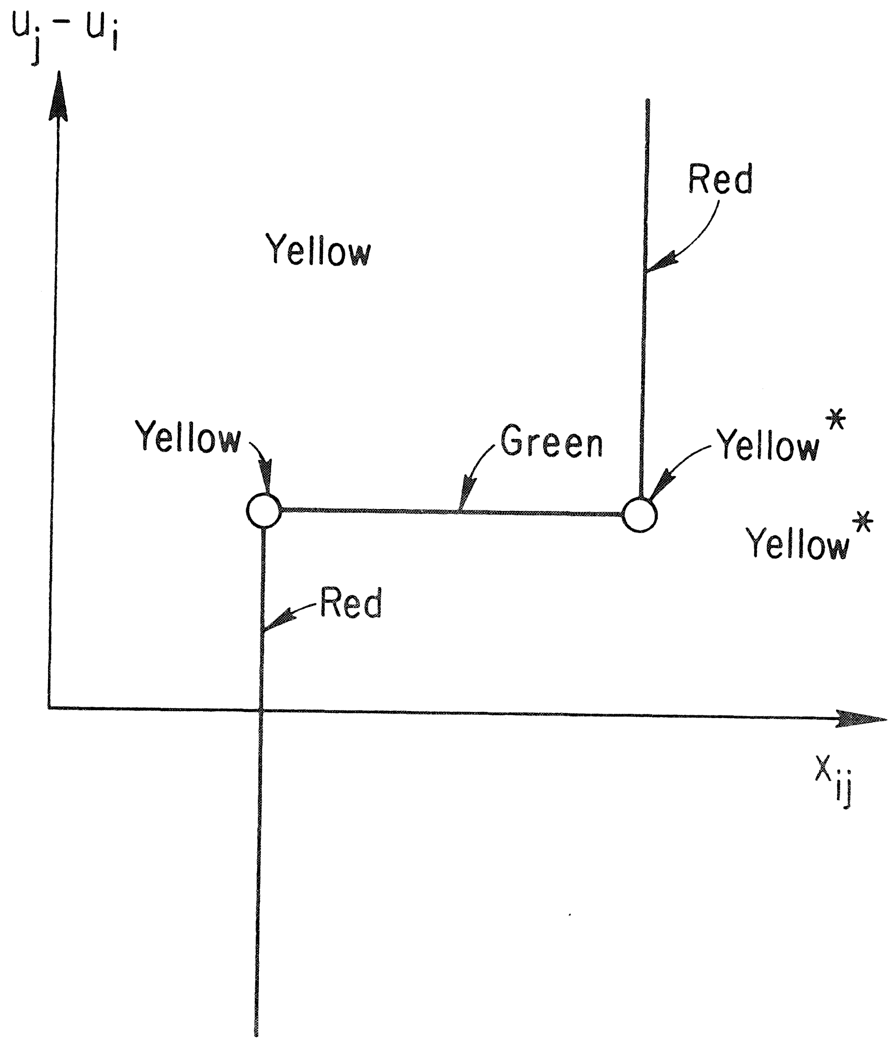
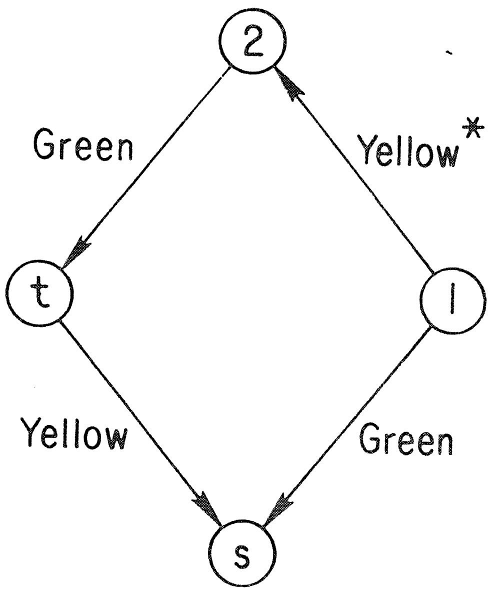


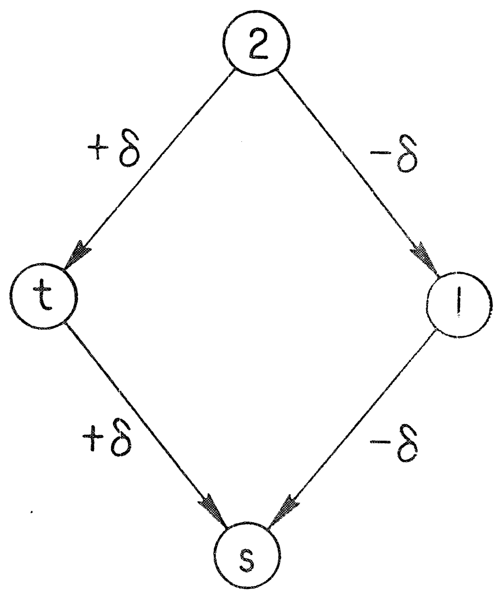
Figure 8.2. Painting of Arcs.

Let us focus attention on an out-of-kilter yellow arc (t,s) and apply the painting theorem. Suppose there is a yellow-green cycle C, in which all yellow arcs are oriented in the same direction as (t,s). Reorient all arcs whose directions were reversed at the time they were painted yellow. An increase by a small amount $\delta > 0$ in the flow through (t,s) will decrease its kilter number by a like amount, assuming the kilter number is finite. (If (t,s) is one of the yellow arcs whose directions was reversed, we mean to *decrease* the flow through (s,t), and the discussion below must be appropriately modified.) An increase by δ in the flow through the arcs of C oriented in the same direction as (t,s) and a decrease by δ in the other arcs will not increase the kilter number of any arc, and may decrease the kilter numbers of some. In other words C - (t,s) describes an augmenting path from s to t.

As an example, consider the cycle shown in Figure 8.3 (a). After reorientation of the yellow arc (1,2), the cycle is as shown in Figure 8.3 (b). Changes in the kilter diagrams for arcs in this cycle are indicated in Figure 8.4. Note that the largest permissible value for δ is determined by the green arc (2,1), which will be colored yellow the next time it is painted.



(a) Typical Yellow-Green Cycle



(b) Flow Increments after Reorientation

Figure 8.3.

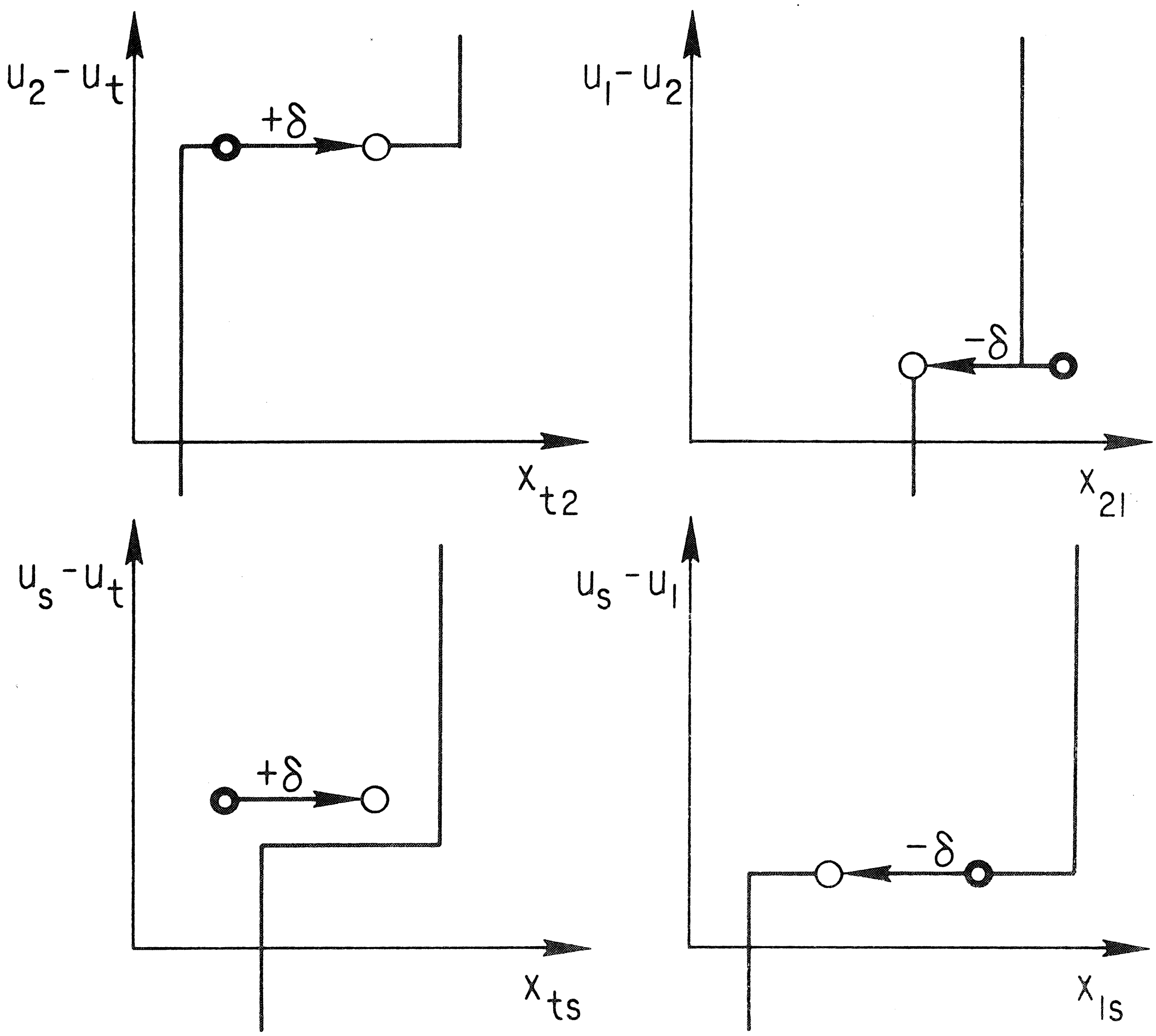


Figure 8.4. Kilter Diagrams for Yellow-Green Cycle.

An analysis of cases shows that the kilter diagrams of the yellow and green arcs in the cycle can be affected only in the manner suggested by the arrows in Figure 8.5. There is no increase in the kilter number of the arc, provided δ is sufficiently small. Let us now consider such a choice of δ.

For a given yellow-green cycle C, let Y,G denote the subsets of yellow and green arcs in C. Let superscripts +,- indicate subsets of Y,G for which arc flow is to be respectively incremented and decremented by δ. No in-kilter arc will be thrown out of kilter if δ is no greater than δ_1, δ_2, where

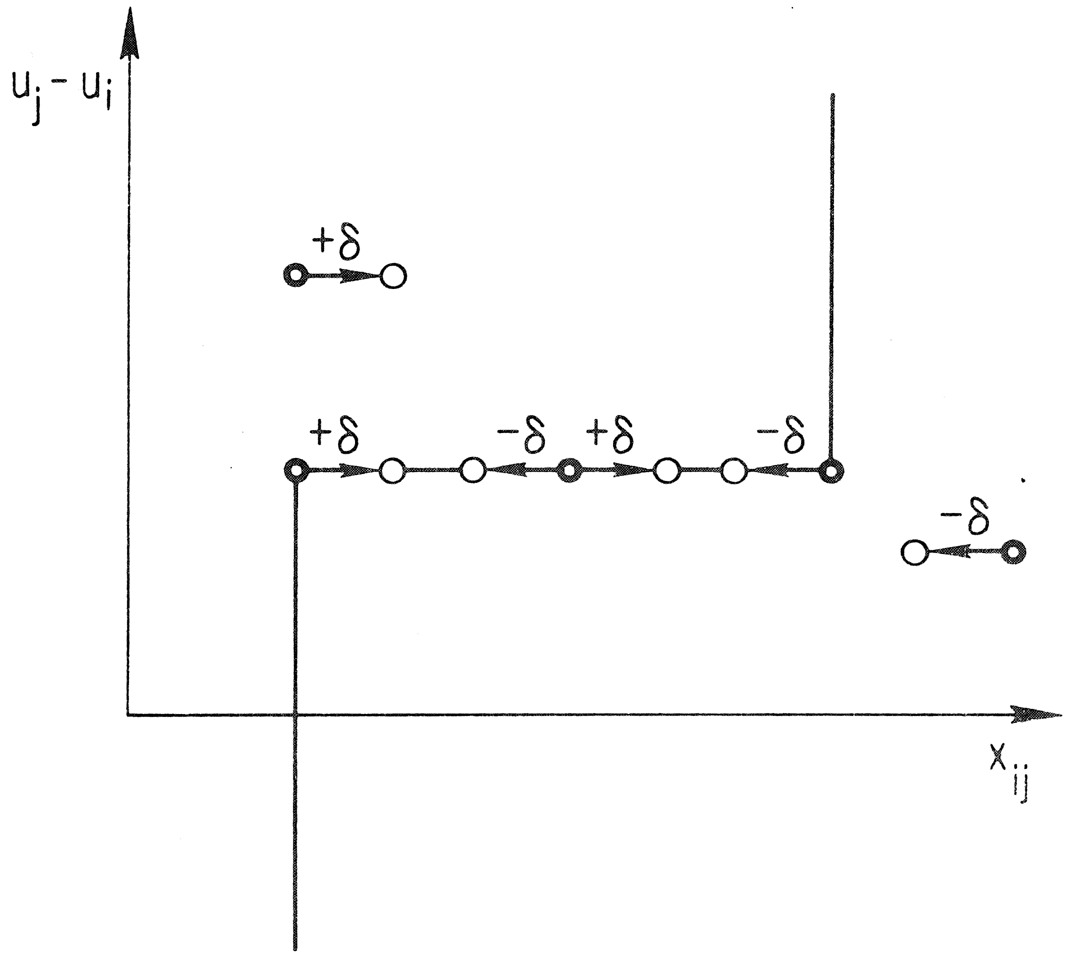


Figure 8.5. Possible Changes in Kilter Diagram for Yellow-Green Cycle.

$$\delta_1 = \min\{c_{ij} - x_{ij} \mid (i,j) \in Y^+ \cup G^+,\ u_j - u_i = a_{ij}\},$$

$$\delta_2 = \min\{x_{ij} - \ell_{ij} \mid (i,j) \in Y^- \cup G^-,\ u_j - u_i = a_{ij}\}.$$

The increment δ will not be any greater than necessary to bring an out-of-kilter arc into kilter if δ is chosen to be no greater than

$$\delta_3 = \min\{|c_{ij} - x_{ij}| \mid (i,j) \in Y^+ \cup Y^-,\ u_j - u_i > a_{ij}\},$$

$$\delta_4 = \min\{|x_{ij} - \ell_{ij}| \mid (i,j) \in Y^+ \cup Y^-,\ u_j - u_i < a_{ij}\}.$$

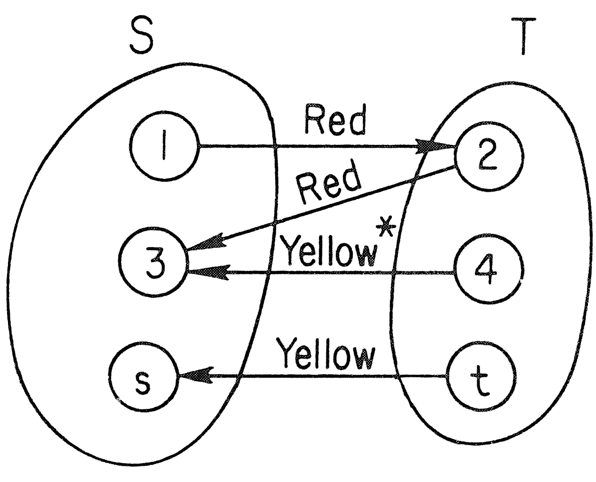
Accordingly, we choose

(8.8) $\delta = \min\{\delta_1, \delta_2, \delta_3, \delta_4\}$.

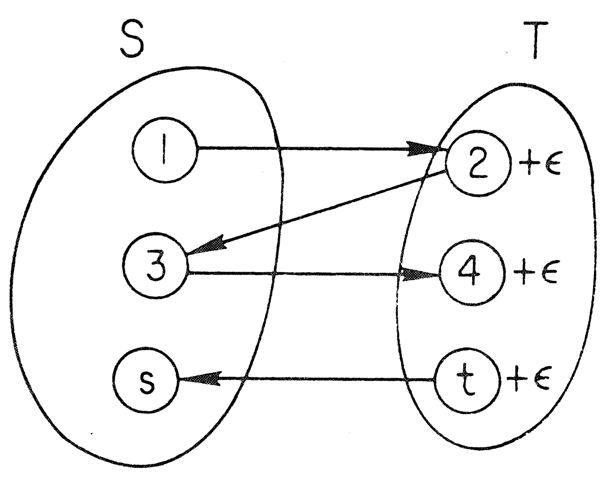
If in (8.8) δ is unbounded, i.e. each of $\delta_1, \ldots, \delta_4$ is determined by minimization over an empty set, then there is no finite optimal circulation. This can occur when capacities of arcs in the cycle are infinite and the net cost of circulation around the cycle is negative.

Now suppose there is a yellow-red cutset (S,T) with $s \in S$, $t \in T$, in which all yellow arcs are oriented in the same direction as (t,s). Reorient all arcs whose directions were reversed at the time they were painted yellow. An increase by a small amount $\varepsilon > 0$ in the node numbers of all nodes i in T affects the value of $u_j - u_i$ only for arcs in the cutset. Moreover, such a change will not increase the kilter number of any arc, and may decrease the kilter numbers of some.

As an example, consider the cutset shown in Figure 8.6 (a). After reorientation of the yellow arc (4.3), the cutset is as shown in Figure 8.6 (b). Changes in the kilter diagrams for arcs in this cutset are indicated in Figure 8.7. Note that the largest permissible value for ε is determined by the red arc (2.3), which will be colored yellow, and its direction reversed, the next time it is painted.



(a) Typical Yellow-Red Cutset



(b) Reorientation of Yellow Arc and Increments to Node Numbers.

Figure 8.6.

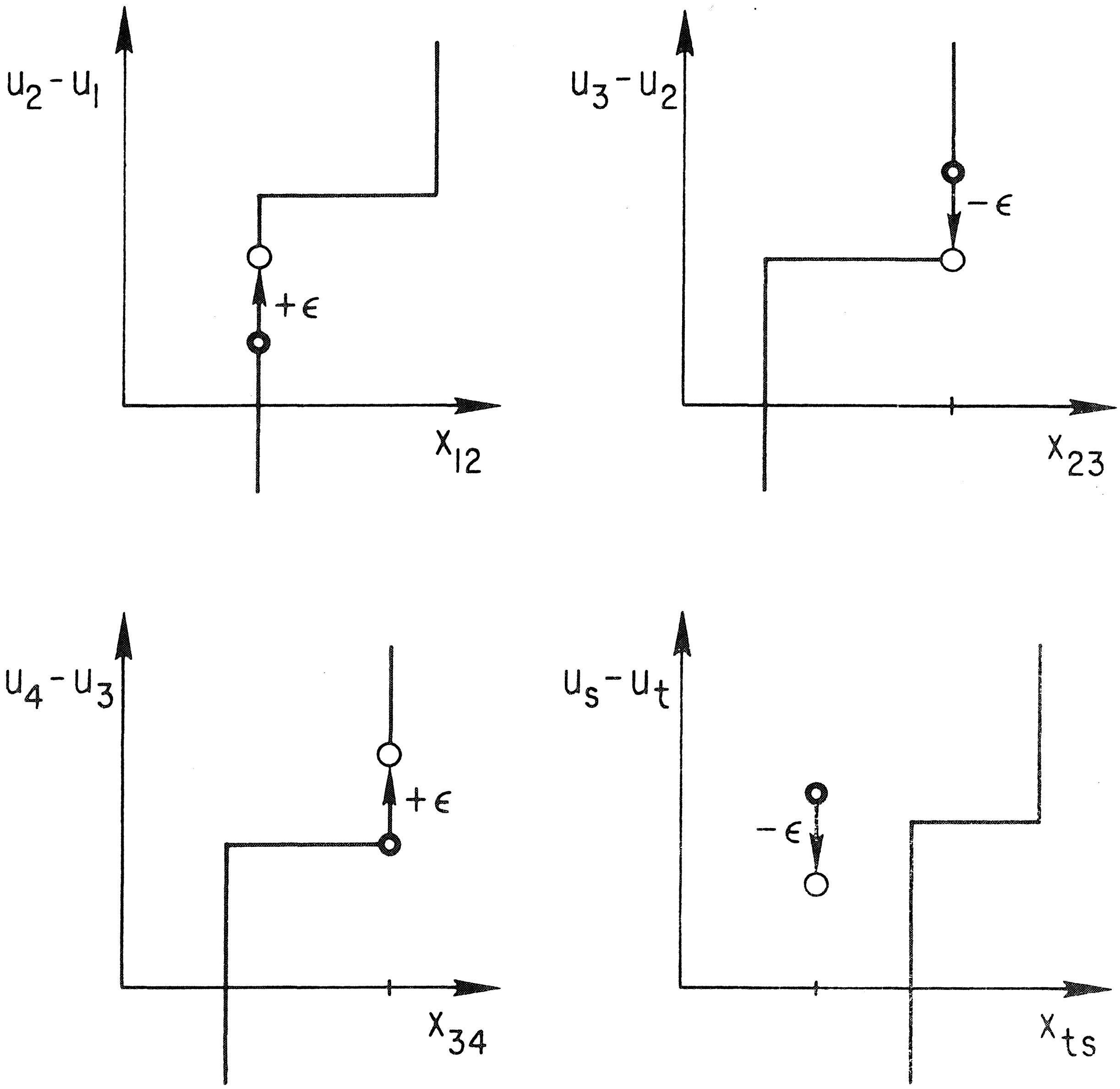


Figure 8.7. Kilter Diagrams for Yellow-Red Cutset.

An analysis of cases shows that the kilter diagrams of the yellow and red arcs in such a cutset can be affected only in the manner suggested by the arrows in Figure 8.8. In each case, there is no increase in the kilter number of an arc, provided ε is chosen sufficiently small. Let us now consider such a choise of ε.

For a given yellow-red cutset C let Y,R denote the subsets of yellow and green arcs in the cutset. Let superscripts +,- indicate subsets of arcs

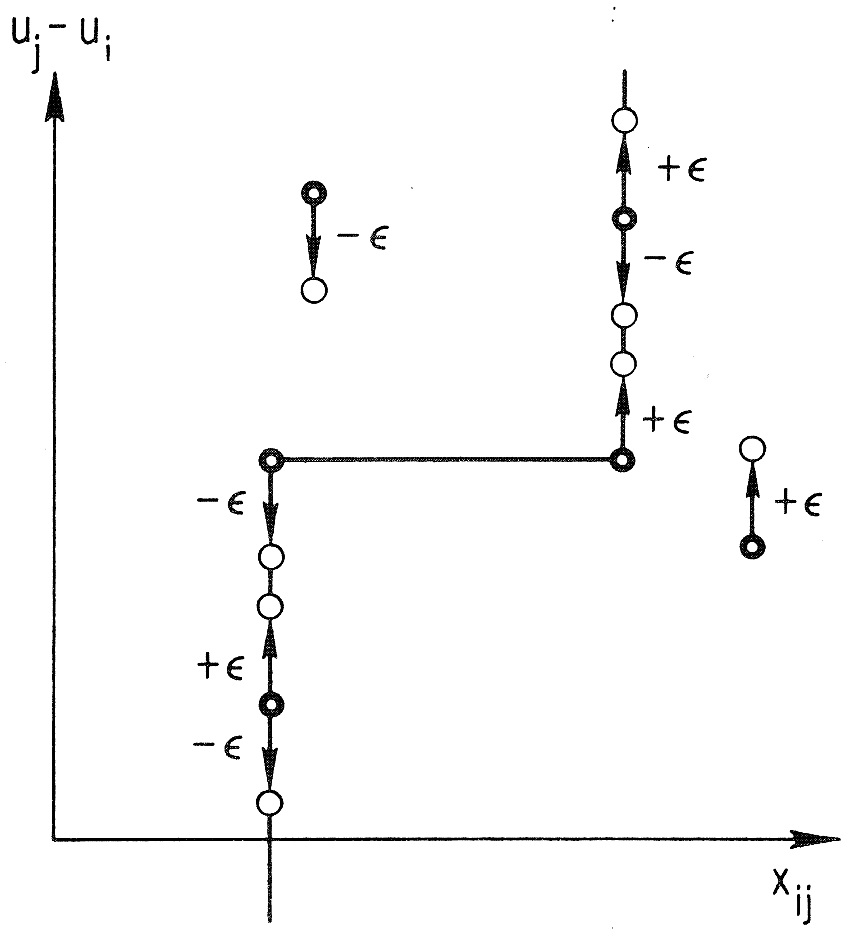


Figure 8.8. Possible Changes in Kilter Diagram of Arcs in Yellow-Red Cutset.

for which $u_j - u_i$ will be respectively increased and decreased by the ε-increment to the node numbers. No in-kilter arc will be thrown out-of-kilter if ε is no greater than $\varepsilon_1, \varepsilon_2$, where

$$\varepsilon_1 = \{u_j - u_i - a_{ij} \mid (i,j) \in R^-, \ x_{ij} = c_{ij}\}$$

$$\varepsilon_2 = \{a_{ij} - u_j + u_i \mid (i,j) \in R^+, \ x_{ij} = \ell_{ij}\}.$$

The increment ε will not be any greater than necessary to bring an out-of-kilter arc into kilter if ε is chosen to be no greater than $\varepsilon_3, \varepsilon_4$, where

$$\varepsilon_3 = \{u_j - u_i - a_{ij} \mid (i,j) \in Y^-, \ell_{ij} \leq x_{ij} < c_{ij}\}$$

$$\varepsilon_4 = \{a_{ij} - u_j - u_i \mid (i,j) \in Y^+, \ell_{ij} < x_{ij} \leq c_{ij}\}.$$

Accordingly, we choose

$$\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}. \tag{8.8}$$

There are three possible cases:

Case 1. ε is unbounded, i.e. each of $\varepsilon_1, \ldots, \varepsilon_4$ is determined by minimization over an empty set. This can occur only if $x_{ij} \geq c_{ij}$ for all arcs from S to T and $x_{ij} \leq \ell_{ij}$ for all arcs from T to S and $x_{ts} < \ell_{ts}$. Net flow from S to T is zero, so

$$\sum_{i \in S, j \in T} \ell_{ij} > \sum_{i \in T, j \in S} c_{ij}.$$

It follows from Theorem 7.1 that no feasible circulation exists.

Case 2. ε is finite and equal to either ε_3 or ε_4. At least one out-of-kilter arc is brought into kilter. No kilter numbers are increased and some may be decreased.

Case 3. ε is finite and greater than both ε_3 and ε_4. No out-of-kilter arc is brought into kilter. No kilter numbers are increased and some may be decreased. At least one red arc will be colored yellow the next time it is painted. For such an arc (i,j), if $i \in S$, $j \in T$, then $\ell_{ij} = x_{ij} < c_{ij}$ and if $i \in T$, $j \in S$, then $\ell_{ij} < x_{ij} = c_{ij}$. In addition, some arcs may change color from yellow to red. For each of these arcs, $i \in S$, $j \in T$ implies $\ell_{ij} < x_{ij} = c_{ij}$ and $i \in t$, $j \in S$ implies $\ell_{ij} = x_{ij} < \ell_{ij}$. No green arcs, of course, are affected.

A labeling procedure can be used, as in the proof of the painting theorem, to construct a yellow-green cycle or a yellow-red cutset. The node s is initially labeled, and all nodes reachable from s are successively labeled. To use the analogy of the proof of the painting theorem, green arcs are viewed as two-way streets, yellow arcs as one-way streets, and red arcs as streets blocked in both directions. If t is reachable from s, backtracing from the label on t yields a yellow-green cycle. If t is not reachable, let S contain all labeled nodes and T the remaining nodes. The desired yellow-red cutset is (S,T).

We are now ready to establish the convergence of the algorithm, provided all lower bounds and capacities are integers and the initial circulation is integral.

Each discovery of a yellow-green cycle results in the reduction of at least one kilter number by some $\delta \geq 1$. Thus, no more than K revisions of the circulation are necessary, where K is the kilter number for the initial circulation.

Assuming a feasible circulation exists, each time a yellow-red cutset is discovered, either an out-of-kilter arc is brought into kilter (Case 2) or at least one red arc changes color to yellow (Case 3). The former case reduces at least one positive kilter number to zero, so this cannot occur more than K times in all. The latter case cannot occur more than n - 1 times in succession, by the following reasoning.

Suppose the same arc (t,s) is used for the application of the painting theorem until a yellow-green cycle is discovered. Then each time a cocycle is discovered and Case 3 occurs, at least one red arc changes color to yellow in such a way that an additional node i in T will become reachable from s the next time the labeling procedure is applied. All nodes reachable from s remain reachable. (Changes from yellow to red are of no consequence.) Thus Case 3 can occur at most n - 1 times in succession before either a cycle is discovered or else an out-of-kilter arc is brought into kilter (Case 2).

Each application of the labeling procedure requires O(m) steps. Labels can be preserved between the discovery of cutsets, according to the previous reasoning. Thus, essentially only K complete applications of the labeling procedure are required. It is asserted that the computation is finite, even if the sum of the kilter numbers is not finite or if the lower bounds and capacities are not integers.

The procedure is summarized as follows.

OUT-OF-KILTER ALGORITHM

Step 0 (Start). Let $x = (x_{ij})$ be any circulation, possibly infeasible, but satisfying conservation conditions, and let $u = (u_i)$ be any set of node numbers. It is desirable to start with x,u such that the sum of the kilter numbers is small, but $x = 0$, $u = 0$ will do.

Step 1 (Painting and Labeling).

1.0. Choose any arc (t,s) which is out-of-kilter and apply an initial label to s. If there is no out-of-kilter arc, the existing circulation is optimal and the computation is completed.

1.1. Paint the arcs green, yellow and red and reorient them as necessary. Apply the labeling procedure to find nodes reachable from s, making use of existing labels. If t is reachable, go to step 2; otherwise go to step 3.

Step 2 (Change in Circulation). Identify a yellow-green cycle C by using the label on t to backtrace to s. Determine δ by (8.7). If δ is unbounded, there is no finite optimal solution and the computation is terminated. Otherwise, increment or decrement the flow in each arc in C by δ (reorienting yellow arcs as described in the text). Erase all labels on nodes and go to Step 1.0.

Step 3 (Change in Node Numbers). Let S contain all nodes reachable from s (labeled nodes) and T contain the remaining nodes. (S,T) is a yellow-red cutset. Determine ε by (8.8). If ε is unbounded, no feasible circulation exists and the computation is terminated. Otherwise increment u_i by ε for each node $i \in T$ and go to Step 1.1.

The out-of-kilter method is easily adapted to handle piecewise linear convex arc costs. A typical arc cost curve of this type and its corresponding kilter diagram are shown in Figure 8.9. It is left to the reader to determine how the algorithm should be generalized and to show that the order of complexity of the computation is unaffected, provided lower bounds, capacities and breakpoints b_{ij} are integers.

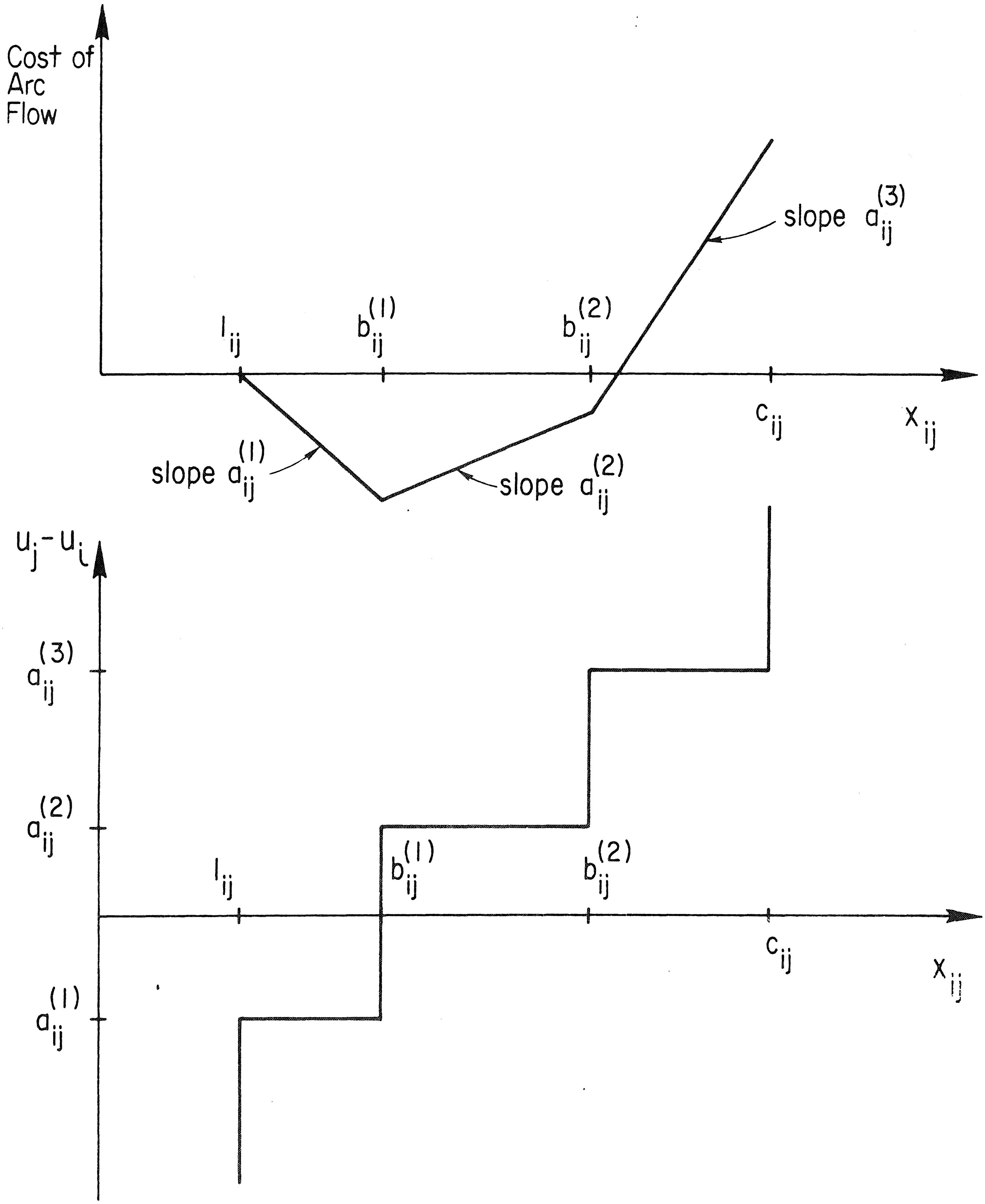


Figure 8.9. Typical Arc Cost Function and Its Kilter Diagram.

9. THEORETICAL IMPROVEMENT IN EFFICIENCY OF OUT-OF-KILTER METHOD

It is possible to establish a bound of $O(Km)$ on the number of steps for the out-of-kilter method, where K is the sum of the arc kilter numbers for the initial primal and dual solutions. If $x = 0$, $u = 0$ are taken as initial solutions, then K may be as large as the sum of all arc capacities, which are assumed to be integers.

In order to qualify as a bona fide polynomial bounded computation, the number of steps required by the out-of-kilter method should be polynomial not in the magnitudes of the arc capacities but in their logarithms, i.e. the number of bits required to specify them as input data. A similar observation holds for the minimum cost flow computation of Section 6, for which a bound of $O(mv)$ can be obtained. It is quite possible that the desired flow value v could approximate the sum of the arc capacities.

We shall not show that either algorithm is polynomial bounded (in fact, they are not). Instead we shall describe a "scaling" technique due to Edmonds and Karp whereby the out-of-kilter algorithm is applied to a series of problems which provide successively closer approximations to the given problem. A polynomial bound of the desired type is then obtained.

Suppose we wish to apply the out-of-kilter method to a problem with integer lower bounds and capacities and for which the maximum arc capacity is no great than 2^p. We first replace the original problem by a problem 0 in which

$$c_{ij}^{(0)} = \left\lceil \frac{c_{ij}}{2^p} \right\rceil,$$

$$\ell_{ij}^{(0)} = \left\lfloor \frac{\ell_{ij}}{2^p} \right\rfloor,$$

and arc costs are as given. (Here "$\lceil\ \rceil$" means "least integer no less than" and "$\lfloor\ \rfloor$" means "greatest integer no greater than".) All lower bounds and capacities are 0 or 1.

This 0-order approximation of the original network admits a feasible circulation, if a feasible circulation was possible in the original, for note that

$$2^p c_{ij}^{(0)} \geq c_{ij},$$

$$2^p \ell_{ij}^{(0)} \leq \ell_{ij}.$$

If we take u = 0, x = 0 as an initial circulation, in this crude approximation of the original network, all kilter numbers are 0 or 1. Hence $K \leq m$, where m is the number of arcs. Accordingly, the out-of-kilter method requires no more that $O(m^2)$ steps to obtain optimal primal and dual solutions x^0, u^0.

We now construct a problem (1) in which

$$c_{ij}^{(1)} = \left\lceil \frac{c_{ij}}{2^{p-1}} \right\rceil,$$

$$\ell_{ij}^{(1)} = \left\lfloor \frac{\ell_{ij}}{2^{p-1}} \right\rfloor,$$

and arc costs remain as given. All lower bounds and capacities are either 0, 1, or 2. If we take $2x^{(0)}, u^{(0)}$ as an initial primal and dual solutions, all arc kilter numbers are again 0 or 1 and again $K \leq m$. The out-of-kilter method requires no more that $O(m^2)$ steps to obtain primal and dual solutions $x^{(1)}, u^{(1)}$.

We continue in this way, passing from problem (k) to problem (k+1), taking $2x^{(k)}, u^{(k)}$ as initial solutions for problem k + 1. Finally, problem p is for a network identical to the original and we will have obtained a circulation for it in $O(m^2 p)$ steps overall. Since $p = \lceil \log_2 c_{ij} \rceil$ for the largest c_{ij}, we have obtained the desired result.

Kilter diagrams for a typical arc with $\ell_{ij} = 7$, $c_{ij} = 20$ are shown in Figure 9.1. The diagrams for successive problems are rescaled so as to best display their relationship with the original. The reader can verify that the $\ell_{ij}^{(k)}$ and $c_{ij}^{(k)}$ values are easily determined from the binary representation of ℓ_{ij} and c_{ij}.

It does not seem possible to apply this scaling technique to the minimum-cost flow algorithm, unless the algorithm is generalized in some way. That is, if $x^{(k)}$ is an optimal solution to problem k, then $2x^{(k)}$ may exceed capacity constraints for problem (k+1). Some technique must be used

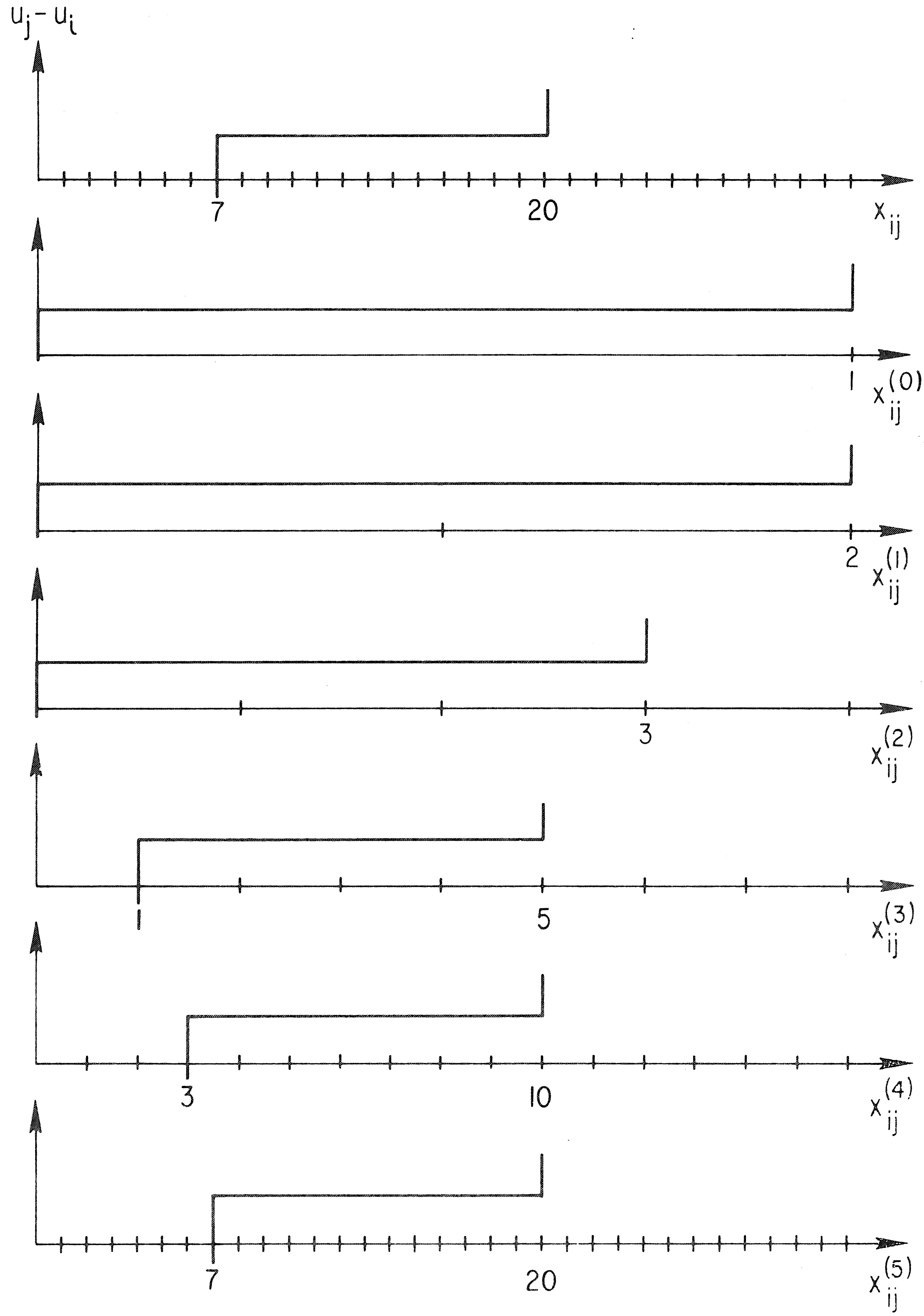


Figure 9.1. Scaled Kilter Diagram.

to restore feasibility before problem (k+1) can be solved. Edmonds and Karp proposed using a limited number of iterations of the out-of-kilter method for this purpose, but this seems a bit like cheating.

We should conclude by saying that this scaling technique, although easy enough to implement, is probably of very limited practical importance. Its significance appears to be largely theoretical, but in this realm it provides a very satisfying result.

COMMENTS AND REFERENCES

The standard reference on network flow theory is

FORD Jr., L.R. & D.R. FULKERSON, *Flows in Networks*, Princeton University Press, Princeton, N.J., 1962.

in which a number of results of combinatorial significance are emphasized. See also:

BERGE, C. & A. GHOUILA-HOURI, *Programming, Games and Transportation Networks*, John Wiley, New York, 1965.

BUSACKER, R.G. & T.L. SAATY, *Finite Graphs and Networks*, McGraw Hill, New York, 1964.

FRANK, H. & I. FRISCH, *Communication, Transportation and Flow Networks*, Addison Wesley, Reading, Mass., 1972.

HU, T.C., *Integer Programming and Network Flows*, Addison-Wesley, Reading, Mass., 1969.

and

IRI, MASAO, *Network Flows*, Transportation and Scheduling, Academic Press, New York, 1969.

Section 2

The max-flow min-cut theorem was established independently by:

ELIAS, P., A. FEINSTEIN & C.E. SHANNON, *Note on Maximum Flow Through a Network*, IRE Trans. on Inform. Theory, IT-2 (1956), 117-119.

and by

FORD Jr., L.R. & D.R. FULKERSON, *Maximal Flow Through a Network*, Canad. J. Math. 8 (1956), 399-404.

Section 8

The out-of-kilter method was developed independently by:

MINTY, G.J., *Monotone Networks*, Proc. Roy. Soc. London, Ser. A, 257 (1960), 194-212.

and by

FULKERSON, D.R., "An Out-of-Kilter Method for Minimal Cost Flow Problems", J. Soc. Indust. Appl. Math. 9 (1961), 18-27.

It appears that many of the ideas of the out-of-kilter method were anticipated by:

YAKOVLEVA, M.A., "Problem of Minimum Transportation Expense", in: *Applications of Mathematics to Economics Research*, ed. by V.S. Nemchinov, Moscow (1959), 390-399.

Minty formulated his version of the algorithm for convex cost functions, and illustrated the application of his method to steady-state electrical network calculations in

MINTY, G.J., "Solving Steady-State Nonlinear Networks of 'Monotone' Elements", *I.R.E. Trans. Circuit Theory*, CT-8 (1961), 99-104.

For other discussions of electrical analogs of network flows see:

DENNIS, J.B., *Mathematical Programming and Electrical Networks*, John Wiley, New York, 1959.

PRAGER, W., "Problems of Network Flow", *J. of Appl. Math. and Physics (ZAMP)* 16 (1965), 185-190.

Section 9

EDMONDS, J. & R.M. KARP, "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems", *J.A.C.M.* 19 (1972), 248-264.