# CHAPTER 12

The contents of this chapter are derived from the preliminary study of Gomory and Hu. Because the ideas are new and seem to be of interest, they are reported here in their most preliminary form. Credits, if any, should be given to Dr. Gomory and the author, but the author is wholly responsible for any errors in the chapter. For readers interested in functional analysis, it may be better to read Section 3 briefly before reading Sections 1 and 2.

## §1. Relative Minimum Cuts

In this chapter we shall introduce some new concepts which have not appeared in the literature. These concepts are important especially for the approximation of a continuous media by a finite network which will be discussed in §3.

Throughout this chapter, we shall consider only undirected networks. Two arcs are said to be neighboring arcs if they have an end node in common. Two cuts $(X, \overline{X})$ and $(Y, \overline{Y})$ are said to be neighboring cuts if every arc of $(X, \overline{X})$ either also belongs to $(Y, \overline{Y})$ or is a neighboring arc of an arc that belongs to $(Y, \overline{Y})$ and the same relation is true for every arc of $(Y, \overline{Y})$.

Consider the network in Figure 12.1 with the cuts represented by dotted lines

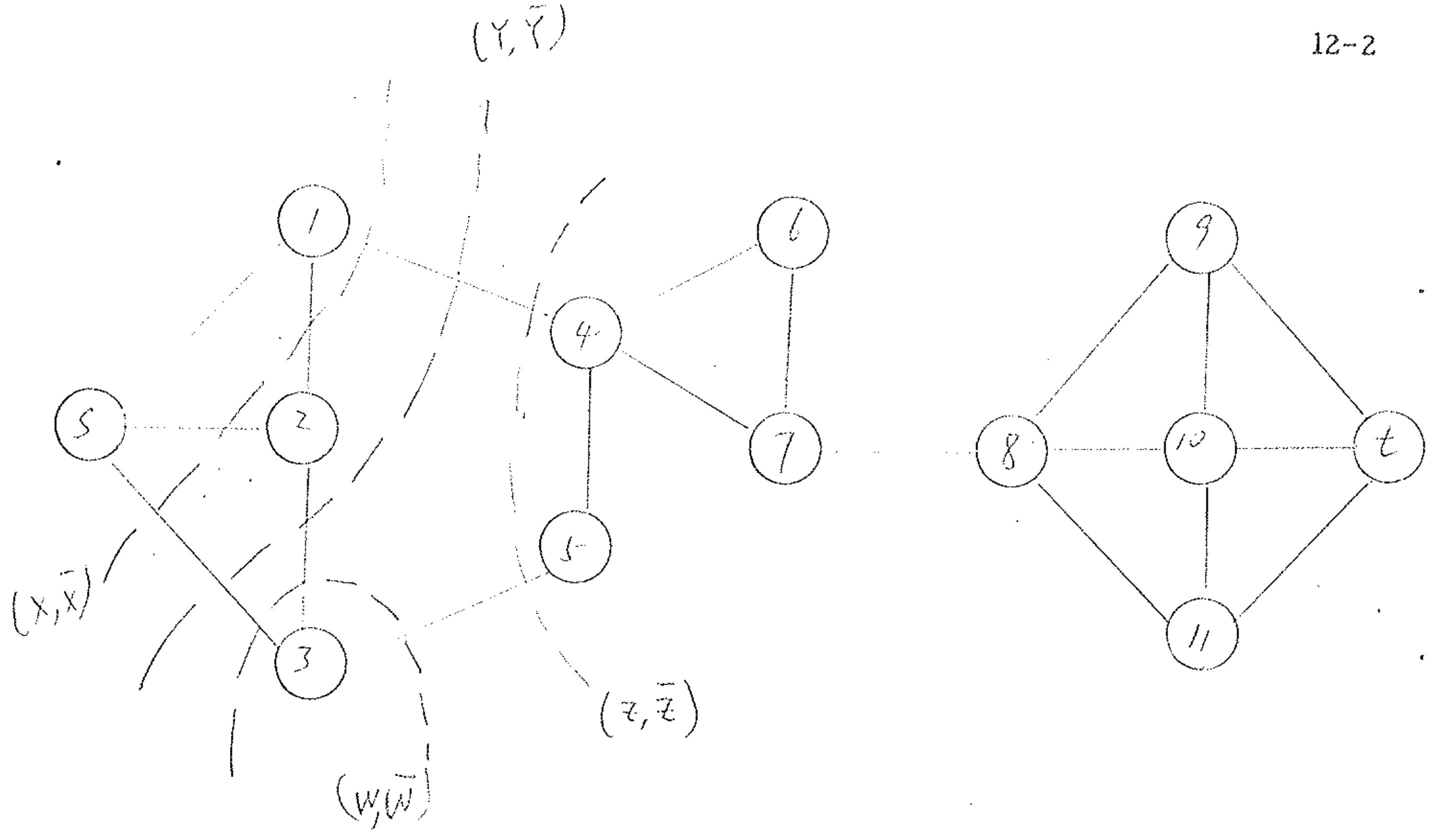



Fig. 12.1

$$(X, \bar{X}) = \{A_{14}, A_{12}, A_{s2}, A_{s3}\}$$

$$(Y, \bar{Y}) = \{A_{14}, A_{23}, A_{s3}\}$$

$$(Z, \bar{Z}) = \{A_{14}, A_{35}\}$$

$$(W, \bar{W}) = \{A_{23}, A_{s3}, A_{35}\}$$

We shall use $\sim$ to indicate that two arcs are neighbors.

$(X, \bar{X})$ and $(Y, \bar{Y})$ are neighbors because

1. $A_{14}, A_{s3}$ are in both cuts
2. $A_{12} \sim A_{23}$, $A_{s2} \sim A_{23}$ where $A_{12}, A_{s2}$ belong to $(X, \bar{X})$ and $A_{23}$ belongs to $(Y, \bar{Y})$.

$(Y, \bar{Y})$ and $(W, \bar{W})$ are not neighbors because

$A_{14}$ of $(Y, \overline{Y})$ is not a neighbor of $A_{35}$ which belongs to $(W, \overline{W})$. Similarly, the reader can verify that $(Y, \overline{Y})$ and $(Z, \overline{Z})$ are neigbhors, $(X, \overline{X})$ and $(Z, \overline{Z})$ are not neighbors, $(X, \overline{X})$ and $(W, \overline{W})$ are not neighbors, and $(Z, \overline{Z})$ and $(W, \overline{W})$ are not neighbors.

In the following, we shall discuss cuts that separate $N_s$ and $N_t$. A cut such as $(W, \overline{W})$ will not be of interest. Therefore, we shall use the word "cut" to mean "cut separating $N_s$ and $N_t$."

A cut is called a relative minimum cut if its capacity is less than or equal to the capacity of all of its neighboring cuts. For example, let every arc in Figure 12.1 have the same arc capacity, then $(Z, \overline{Z})$ is a relative minimum cut, so is the cut which consists of the single arc $A_{78}$.

A minimum cut as defined in Chapter 8 is certainly a relative minimum cut but the converse is not true. For example, $A_{78}$ is clearly the minimum cut separating $N_s$ and $N_t$. $(Z, \overline{Z})$ is a relative minimum cut but not a minimum cut. Therefore, a minimum cut in our terminology is really an absolute minimum cut in a sense. For a cut to be a minimum cut it is necessary but not sufficient that it be a relative minimum cut. Consider all the cuts separating $N_s$ and $N_t$ in a network. The neighboring relation between cuts is analogous to the distance between points in a plane. For a given point "a" there are points within distance $\epsilon$ to that given point "a"; these points are said to be in the $\epsilon$-neighborhood of "a". Similarly, for a given cut there are cuts which are neighboring cuts of the given cut. In order for a function $f(x)$ to be an absolute minimum of a point $a$, it is necessary that $f(x)-f(a) \geq 0$ for $|x-a| < \epsilon$

say. Here, in order for a cut to be a minimum cut, it is necessary that it be a relative minimum cut.

In calculus, or in functional analysis, a local minimum is obtained. If a globle minimum of the function is desired, we have to compare all the local minima. In network flow theory we are interested in minimum cuts separating the source and the sink. Take the network in Figure 12. 1 for example, $(Z, \overline{Z})$ and $A_{78}$ are not neighboring cuts, and we want an absolute minimum cut among all cuts separating $N_s$ and $N_t$. Thus <u>the labelling method for getting the maximal flow (hence locating the minimum cut) is therefore a technique which locates an absolute minimum which is not implied by a local minimum condition.</u> We shall explore more on this aspect in later sections.

## §2. Node-constraint networks

In Chapter 8 of the book, we associated with every arc of the network an arc capacity which indicates the maximum amount of flow that can pass through the arc. There is no limitation on the amount of flow that can pass a node except that flow must conserved at every node. Now we shall let the nodes have capacity restrictions and the arcs have no capacity restrictions. We develop this model mostly for the use in the section 3.

We consider a network consisting of nodes $N_i$ and arcs $A_{ij}$ connecting $N_i$ and $N_j$. Each node $N_i$ has associated with it a node capacity $w_i$, which indicates the maximum amount of flow that can pass through the node. Let $x_{ij}$ be the flow from $N_i$ to $N_j$ in the arc $A_{ij}$. Since the flow is conserved at every node, the amount of flow passing a node $N_j$ is therefore $\frac{1}{2}\sum_i |x_{ij}|$, this is denoted by $x_j$. The maximal flow problem for a node-constraint network is therefore

$$\max \quad v$$

subject to

$$\sum_i x_{ij} - \sum_j x_{jk} = \begin{cases} -v & j = s \\ 0 & j \neq s, t \\ v & j = t \end{cases}$$

$$x_{ij} \geq 0 \tag{1}$$

$$0 \leq x_j \leq w_j \qquad \text{for all } i, j$$

$$\frac{1}{2} \sum_i |x_{ij}| = x_j \ .$$

Note that we put absolute value sign on $x_{ij}$ because some of the arc flow are going into the node $N_j$ and some of the arc flows are leaving the node $N_j$.

A cut separating $N_s$ and $N_t$ in the node-constraint network will be a set of nodes, the removal of which will disconnect the network into <u>two or more parts</u>, one part containing $N_s$, another part containing $N_t$, and no proper subset of which should have this property.

Two nodes are said to be neighbors if there is an arc connecting them. Although the removal of the cut will separate the network into more than two parts in general, we can still use the notation $(X, \overline{X})$ to denote a cut. $X$ then denotes all nodes in the part of the network containing $N_s$ and $(X, \overline{X})$ will be all the neighboring nodes of $X$. A cut with the sum of the node capacities a minimum is called a minimum cut. Just like the Max Flow Min Cut Theorem in Chapter 1, we can prove an analogue theorem for the node-constraint network. It is true that a node-constraint network can be converted into an arc-constraint network (see Ford and Fulkerson [ ]), but this greatly increases the number of arcs. We shall deal with the node-constraint network directly.

Max Flow Min Cut Theorem (Node-constraint case). For a node-constraint network, the maximal flow value from $N_s$ to $N_t$ is equal to the minimum cut capacity of all cuts separating $N_s$ and $N_t$.

Proof: Assume that all node capacities $w_i$ are positive integers and let $f_{st}$ be any flow, not necessarily maximum. If the value of the flow is equal to the capacity of a cut, the theorem is proved. Otherwise, we shall search for a chain along which the flow value can be increased. This is done by defining a subset $X$ such that flow can be sent to any node in $X$.

Based on the current flow, we shall define the set $X$ recursively as follows: (let $w_{ij} = \min(w_i, w_j)$)

Rule 0. $N_s \in X$

Rule 1. If $N_i \in X$ and $x_i < w_i$, $x_{ij} < w_{ij}$, $x_j < w_j$ then $N_j \in X$

Rule 2. If $N_i \in X$ $x_{ij} > 0$ then $N_j \in X$

Rule 3. If $N_i \in X$ $x_i < w_i$, $x_{ij} < w_{ij}$, $x_j = w_j$, $x_{kj} > 0$ then $N_k \in X$

Based on this recursive definition, either $N_t$ is in $X$ or in $\bar{X}$.

Case 1. $N_t$ is in $X$. We have to show that the flow value can be increased. Since $N_t$ is in $X$, there must be a chain from $N_s$ to $N_t$ along which one of the above three rules must be true. If Rule 1 and Rule 2 holds, then we can clearly send flow along the chain. If the third rule hold, let $\epsilon = \min(w_{ij} - x_{ij}, x_{kj})$. Then we can add $\epsilon$ to $x_{ij}$ and subtract $\epsilon$ from $x_{kj}$. This is equivalent to sending flow to $N_k$ and keeping $N_j$ saturated with flow. Since $w_j$ are integers, $\epsilon$ will be an integer. As the maximal flow is bounded from above, this case can not be repeated indifinitely.

Case 2. $N_t$ in $\bar{X}$. We have to show that the value of the present flow is equal to the capacity of a cut. Let the neighboring nodes of $X$ be called $\gamma$ nodes. We have to show the following four things.

1. There is no arc flow from a $\gamma$ node to a node in $X$.

2. There is no arc flow from $\overline{X} - \gamma$ to $\gamma$.

3. There is no arc flow from one $\gamma$ node to another.

4. All $\gamma$ nodes are saturated.

Each of the above can be proved as follows.

1. There is no $x_{ji} > 0$ with $N_i \epsilon X$ and $N_j \epsilon \gamma$ because Rule 2 would then label $N_j$ to be in $X$.

2. There is no $x_{kj} > 0$ with $N_j \epsilon \gamma$ and $N_k$ in $\overline{X} - \gamma$, because Rule 3 would then label $N_k$ to be in $X$.

3. There is no $x_{jk} > 0$ with $N_j$, $N_k \epsilon \gamma$, because Rule 3 would then label $N_j$ to be in $X$.

4. Let $N_i$ be in $X$ and $N_j$ be a $\gamma$ node and a neighbor of $N_i$.

   If $N_i$ is saturated and all the arc flows out of $N_i$ are to $\gamma$ nodes, then $N_i$ could not be in $X$.

   If $N_i$ is saturated and there is an $N_k \epsilon X$ and $x_{ik} > 0$, then the node $N_j \epsilon \gamma$ would be labeled as an $X$ node from $N_k$ by Rule 3.

   If $N_i$ is not saturated and if $x_{ij} < w_{ij}$, then $N_j$ will be in $X$ unless $N_j$ is saturated. Therefore all $\gamma$ nodes are saturated.

The value of the flow is therefore $\Sigma x_{ij}$, $N_i \epsilon X$, $N_j \epsilon \gamma$, which is equal to $\Sigma w_j$ ($N_j \epsilon \gamma$). Because $\gamma$ nodes are all the neighboring nodes of $X$, it is a cut. This completes the theorem.

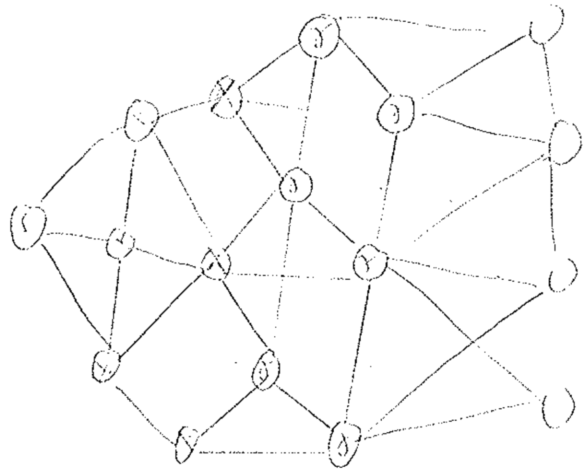


Fig. 12.2

We can clearly develop a labeling method, analogue to the one in chapter 8 and based on the three rules. However, the third rule requires that not only do we have to look at all neighboring nodes of a node but also all neighboring nodes of the neighboring nodes. This greatly increases the amount of computation. The following is a labeling procedure, which uses only first and second rules plus minor modifications. Nodes will be classified into five types:

L nodes: Labeled nodes

LS nodes: Labeled and scanned nodes

R nodes: Rejected nodes

RS nodes: Rejected and scanned nodes

U nodes: Unlabeled nodes

<u>L node:</u> A node $N_i$ is a L node if using the first and second rules, we label it to be in X.

<u>LS node</u>: A node $N_i$ is a LS node if it is a L node and we have looked all neighboring nodes $N_j$ of $N_i$.

<u>R node:</u> A node $N_j$ is a R node if it is looked at and found to be at its capacity, i. e., $x_j = w_j$.

<u>RS node:</u> A node $N_j$ is a RS node if it is a R node and we have looked all neighboring R nodes $N_k$ (of $N_j$) and found that there is no arc flow $x_{kj} > 0$.

<u>U node;</u> A node $N_j$ is a U node if it is not any of the above types.

The labeling procedure proceeds as follows: We first label a node to be in X only if it belongs to X due to the first and second rule. A node is labeled a R node if it is saturated. If break through occurs, then we increase the flow, erase all labels, and start a new cycle. If a non-break through occurs, then we look at all R nodes $N_j$ one by one. For a given R node $N_j$, if there is no $x_{kj} > 0$ with $N_k$ a R node, then $N_j$ is labeled an RS node. If there is a R node $N_k$ with $x_{kj} > 0$, then that R node $N_k$ is changed into a L node and we search all neighboring nodes of $N_k$ to see if any more nodes can be labeled. If there is no L node and an R node and the non-breakthrough occurs, then the present flow is already maximum. In this labeling procedure a node $N_j$ may be first a R node because $x_j = w_j$ and second become a RS node due to non-existence of an R node $N_k$ with $x_{kj} > 0$. Let us assume that $x_{kj} > 0$ for a R node $N_k$. Later the R node $N_k$ may become an L node and that node $N_j$ will be changed into a L node also as $x_{jk} > 0$. Therefore, the longest sequence of changing labels that can happen to a node is indicated below

$$R \rightarrow RS \rightarrow L \rightarrow LS.$$

Therefore, a node together with its neighbors is looked at, at most twice during this labeling method.

## §3. Flows in a Continuous Media

Let us consider a special problem of calculus of variations. Consider the rectangular region with four sides A, S, B, T in Figure 12.3. A bounded continuous weighting function $w(x,y) > 0$ is defined on the rectangular region. We want to find a curve from the side A to the side B such that the line integral

$$\int w(x, y)\, dc \quad \text{is a minimum.}$$

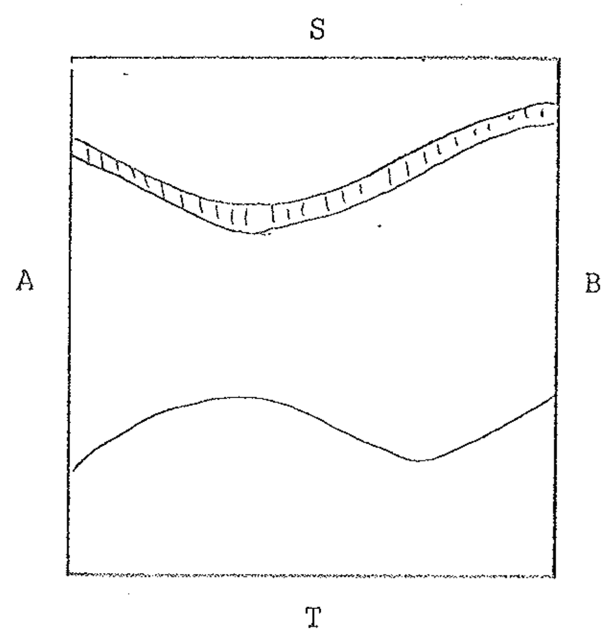



Figure 12.3

This is a problem of calculus of variations and can be solved by the usual technique if the weighting function $w(x, y)$ is sufficiently smooth. As with all problems of calculus of variations, the minimum, if obtained, is a local minimum which may not imply a global minimum.

Assume that the problem is an abstraction of a practical problem which is to locate a cheapest path for an automobile to go from the side A to the side B. The weighting function $w(x, y)$ may indicate the amount of fuel consumption at the point $w(x, y)$. There are two reasons that the curve with the line integral $\int w(x, y)\, dc$ a minimum is not what we really want. The first reason is that the weighting function may be very small on the optimum curve but may be very large near the curve. This means that if the automobile should deviate slightly from its prescribed optimum curve, it is very expensive. Since no automobile can be controlled with 100% precision, it is better to locate a strip with width $\epsilon$ in which weighting function $w(x, y)$ is small. The second reason is that the locus of an automobile unlike the locus of a point is not a curve but a strip. Thus, in practice, we would rather get a strip with width $\epsilon$ from A to B such that

$$\iint w(x, y)\, dA \text{ is a minimum.}$$

(Furthermore, we want the integral to be a global minimum compared with the integrals of other strips of width $\epsilon$ from the side A to the side B). It is very easy to construct an example in which, for a given $\epsilon$, the strip with $\iint w(x, y)\, dA$ minimum does not contain the optimum curve with $\int w(x, y)\, dA$

minimum. However, if we let $\epsilon$ go to zero, then the optimum strip will approach the optimum curve as the limit. We shall describe a technique which will give a strip of width $\epsilon$ with $\iint w(x, y)\, dA$ a global minimum, and if $\epsilon$ goes to zero, the strip will approach the optimum curve as a limit.

Note that any continuous curve $\Gamma$ from A to B will separate the rectangular region into two parts, one part containing the side S and one part containing the side T, and any curve from S to T will meet the curve $\Gamma$. In anology to a network with a finite number of nodes, we may think of the rectangular region as a network with the side S as the source and the side T as the sink. The weighting function $w(x,y)$ can be thought of as the capacity function of the media at the point $(x, y)$. Any curve from A to B corresponds to a cut, and the line integral is the capacity of a cut. If we can define flows in the continuous media and locate the minimum cut using maximum-flow minimum cut theorem, then the minimum cut will be the curve with $\int w(x, y)\, dc$ a globle minimum. The approach described below approximates the continuous media by a finite network.

The finite network consists of many nodes, each node represents a small square of the rectangular region. Each node has a capacity which equals the weight of the small square it represents. The node-capacity network will have a set of nodes as its minimum cut. If the approximation is a good one, the set of small squares corresponding to the nodes should resemble a strip of width $\epsilon$.

Thus, we have to keep three things in mind, first, the network with prescribed connection between nodes, where a minimum cut is a subset of the nodes. Second, the rectangular region is partitioned into small squares The union of the small squares, which are represented by the nodes of the minimum cut, is a subregion. Third, the subregion should approach a strip of width $\epsilon$ which separates the rectangular region.

The nodes and the small squares are in one to one correspondence. For node-capacity network there always exists a minimum cut, and hence we can always get a set of small squares which are represented by the nodes. As the size of the squares goes to zero, the subregion should approach a strip of width $\epsilon$ .

Let us first try the most straight forward way of approximating the continuous media and see what difficulties would arise. Let the rectangular region in Figure 12.3 be divided into uniform small squares of side h . At the center of every square we put a node with the capacity of the node equal to the total weight of the square. (See §2 for the definition of a network with node capacity and without arc capacity). Every node is connected by an arc to its neighboring nodes of distance h apart. Such a network is shown in Figure 12.4.

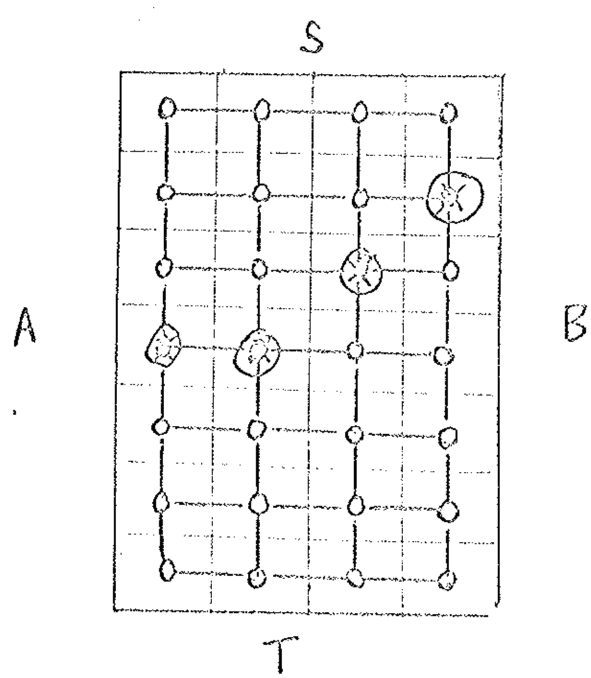


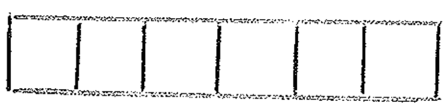

Fig 12.4

The boundry squares touching the side S are represented by sources with limited supply. The amount of supply of a source is the total weight of the square. Similarily, the boundry squares near the side T become sinks with limited demand. This many-source and many-sink network can be converted into a one-source and one sink network by the technique of §3, Chapter 8. A cut of the network in Figure 12.4 will be a set of nodes the removal of which together with its incident arcs will disconnect the network into two parts, one part containing all the sources and one part containing all the sinks. The set of squares represented by the set of nodes should look like a strip with the total weight a minimum. It seems plausible that as h approaches zero, the strip will approach the optimum curve as the limit. Unfortunately, the network described above can not do the job.

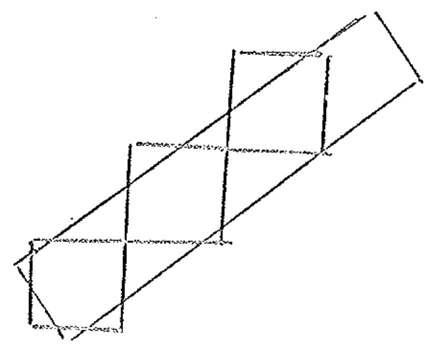
The first difficulty of the approach is if one connects the set of nodes in the minimum cut, the curve will always consist of horizontal straight lines, vertical straight lines and lines with 45° angles with the horizontal or vertical lines. One such set of nodes is marked with x in Figure 12.4. Consequently, the strip which is the union of the squares represented by the nodes will consist of horizontal strip, vertical strip, and 45° strip. This is not desireable, as we need a smooth strip which will approach a smooth optimum curve as h goes to zero.

The second difficulty is that the total weight of the set of nodes may not be equal to the strip which the set of nodes should represent.

If the set of nodes in the minimum cut all lie in a horizontal line say, then the weight of the set of nodes is equal to the weight of the horizontal strip of width $h$. (See Figure 12.5a) If the set of nodes in the minimum cut should all lie in a 45° line say, then the total weight of the set of nodes does not equal to the weight of the strip of width $h$ inclined at 45°. This is shown in Figure 12.5b.



12.5 (a)



12.5 (b)

In order to overcome the difficulties just mentioned, we shall construct a different network as follows. The rectangular region is again divided uniformly into small squares of side $h$. At the center of each square, we put a node with its capacity equal to the total weight of the square. Now we connect any two nodes with an arc if the distance between the two nodes is equal to or less than $r$ ($r >> h$). This network is called a r-connected network. The minimum cut of this network is then a set of nodes which looks like the lattice points in a strip with width $r$, and the total weight of the set of nodes is equal to the total weight of the strip.

Furthermore, if h goes to zero, r goes to zero, and h/r goes to zero, then the limit of the strip will be the curve with $\int w(x, y)\, dc$ minimum.

In this r-connecting network, removing a set of nodes in a horizontal row will not disconnect the network into two parts, since there can be two nodes one above and one below the horizontal row, which are connected. In order to block the flow from S to T, we have to remove $<\frac{r}{h}>$ rows of nodes, for example. The labelling method for a node-capacity constraint network will pick out the set of nodes of the minimum cut. Since it is a cut, the network should be separated into two parts once the minimum cut is removed. Since it is a minimum cut, no proper subset of nodes should have the separating properties.

Since the small squares are in one to one correspondence with the nodes, the subregion, which is the union of small squares, should have the analogous properties of a minimum cut, namely, it separates the rectangular region into two parts and any single small square of the subregion if removed from the subregion will destroy the separating property of the subregion. Now the region containing S as well as the region containing T both are unions of squares. The two regions are considered to be separated if there exist no two small squares, one in each region, with the distance between the two centers of the squares less than or equal to r.

The question now is, if $h \to 0$, $r \to 0$, $h/r \to 0$, what should the subregion approach as a limit? This question is studied in more detail in the appendix. Briefly, the subregion will approach a strip of width r from

A to B.

Since r is comparatively larger than h, there are many more squares completely inside the strip of width r than squares which are partially in the strip. The squares completely inside the strip are called the interior squares, and the squares partially in the strip are called the boundary squares. For the weighting function to satisfy condition (1) in the appendix, the total weight of the interior squares is always of an order of magnitude greater than that of the total weight of the boundary squares, and the total weight of the subregion or that of the nodes in the minimum cut approaches the weight of the strip as $h/r \to 0$.

# APPENDIX

In this appendix, we discuss some properties of the continuous minimum cut; a full length paper by Gomory and Hu [ ] will appear elsewhere. First, we shall restrict ourselves to points in the two dimensional projective plane $R_2$. Two points a and b in a set P are said to be r-connected if there is a finite sequence of points $p_0, p_1, \dots, p_n$ with $p_i \in P$ and the distance $\rho(p_i, p_{i+1}) \leq r$, $(i = 0, \dots, n-1; p_0 = a, p_n = b)$. See for example Newman [ ]. On the other hand, two points a and b in a set P are said to be r-separated if they are not r-connected. We are interested in a set C whose removal from the plane $R_2$ will r-separate two points in $R_2 - C$. More precisely, we will say that a set $C \subset R_2$ r-separates "a" and "b" if

(i) C is closed and bounded.

(ii) $a, b \notin C$.

(iii) "a" and "b" are not r-connected in $R_2 - C$.

With these definitions, any sufficiently large set will r-separate "a" and "b". However, if we require C to be irreducible, i. e., no proper subset of C will also r-separate "a" and "b", then C is highly structured. In the following, we shall use C to denote an irreducible r-separating set. Note that an irreducible r-separating set is analogous to a cut in the continuous media and "a" and "b" are analogous to the source and the sink. We shall state the theorems and lemmas about the irreducible r-separating set. The proofs are to appear in Gomory and Hu [ ]. We shall use the following notations.

C: irreducible r-separating set, which is closed by definition

A: points r-connected to "a", which is an open set.

B: points r-connected to "b", which is an open set.

D: $R_2 - A - B - C$, which is an open set.

Theorem 1. Every r-separating set contains an irreducible r-separating set.

Theorem 2. Let p be a point of C, then $\rho(p, A) \leq r$ and $\rho(p, B) \leq r$.

Lemma 1. Let $\overline{p_1p_2}$ and $\overline{q_1q_2}$ be two line segments of length $\leq r$ that intersect at some point p. In other words, $\overline{p_1p_2}$ and $\overline{q_1q_2}$ are the diagonal lines of a quadrangle. Then there is a vertex of the quadrangle ($p_1, p_2, q_1$ or $q_2$) that has its two adjacent sides both less than or equal to r.

Let $p_1, p_2, \ldots, p_n$ be a sequence of points such that $\rho(p_i, p_{i+1}) \leq r$, then the sequence of line segments $\overline{p_1p_2}, \overline{p_2p_3} \ldots$ are said to form a r-connecting chain. An r-connecting chain is said to be simple if every vertex $p_i$ belongs to only two line segments $\overline{p_{i-1}p_i}$ and $\overline{p_ip_{i+1}}$ (except $p_1$ and $p_n$) and every other point is contained in one line segment.

Lemma 2. If there exists an r-connecting chain, there exists a simple r-connecting chain.

Theorem 3. If C is an irreducible r-separating set, then $R_2 - C = A \cup B$. In otherwords, D is empty. (Note that this Theorem is very much like the Jordan Curve Theorem which states that a simple closed curve separates

$R_2$ into two simply connected open sets. Here, an irreducible r-separating set r-separates $R_2$ into two r-connected open sets A and B.)

Since A or B may consist of many simply-connected open sets, we shall call each simply-connected set, a component of A or B.

Theorem 4. Each component of A or B is uniformly locally connected and the boundary of each component of A or B is a simple Jordan Curve.

Therefore, the boundary of C is a union of simple Jordan Curves. The structure of C can, however, be worked out in much greater detail.

Roughly speaking it can be shown that C splits into sections of two types. The first type are the tube-like sections of width r, like $C_1$ and $C_2$ in Fig. A1. The second are convex polyhedra with an even number of sides, all sides being of length exactly r. See P in Fig. A1. These polyhedra are, intuitively speaking, the areas in which the strip overlaps itself in such a ay as to maintain its irreducibility. It can be shown that any C is the union of sections of these two types.

If the weighting function $w(x, y)$ is arbitrary, then any irreducible r-separating set can be a minimum-weight separating set, simply by defining $w(x, y) = \epsilon$ in the irreducible set and $w(x, y)$ large elsewhere.

Note in Fig. A1 $\overline{p_2p_4} < r$, thus both $A_1$ and $A_2$ are r-connected to "a" and $A = A_1 \cup A_2$. The set of points B including "b" consists of points enclosed by the strip.

The strip in Fig. A1 can be a minimum-weight strip, provided the weighting function is large in A and B and small in the strip.

$R_2$ into two simply connected open sets. Here, an irreducible r-separating set r-separates $R_2$ into two r-connected open sets A and B.)

Since A or B may consist of many simply-connected open sets, we shall call each simply-connected set, a component of A or B.

Theorem 4. Each component of A or B is on uniformly locally connected and the boundary of each component of A or B is a simple Jordan Curve.

¶ insert new paragraph

~~Therefore, the boundary of C is a union of a simple Jordan Curve.~~ The structure of C can [illegible] ~~Roughly speaking,~~ it can be said that ~~C consists of tube-like sections hooked together by convex polyhedra with an even number of sides all of length r. For example, the diamond-shaped polyhedra shown in dotted lines in Fig. Al are produced by the strip overlapping itself. Nevertheless, the r-separating set is irreducible.~~ If the weighting function w(x, y) is arbitrary, then any irreducible r-separating set can be a minimum-weight separating set, simply by defining $w(x, y) = \epsilon$ in the irreducible set and w(x, y) large elsewhere.

Note in Fig. Al $\overline{p_2 p_4} < r$, thus both $A_1$ and $A_2$ are r-connected to "a" and $A = A_1 \cup A_2$. The set of points B including "b" consists of points enclosed by the strip.

The strip in Fig. Al can be a minimum-weight strip, provided the weighting function is large in A and B and small in the strip.

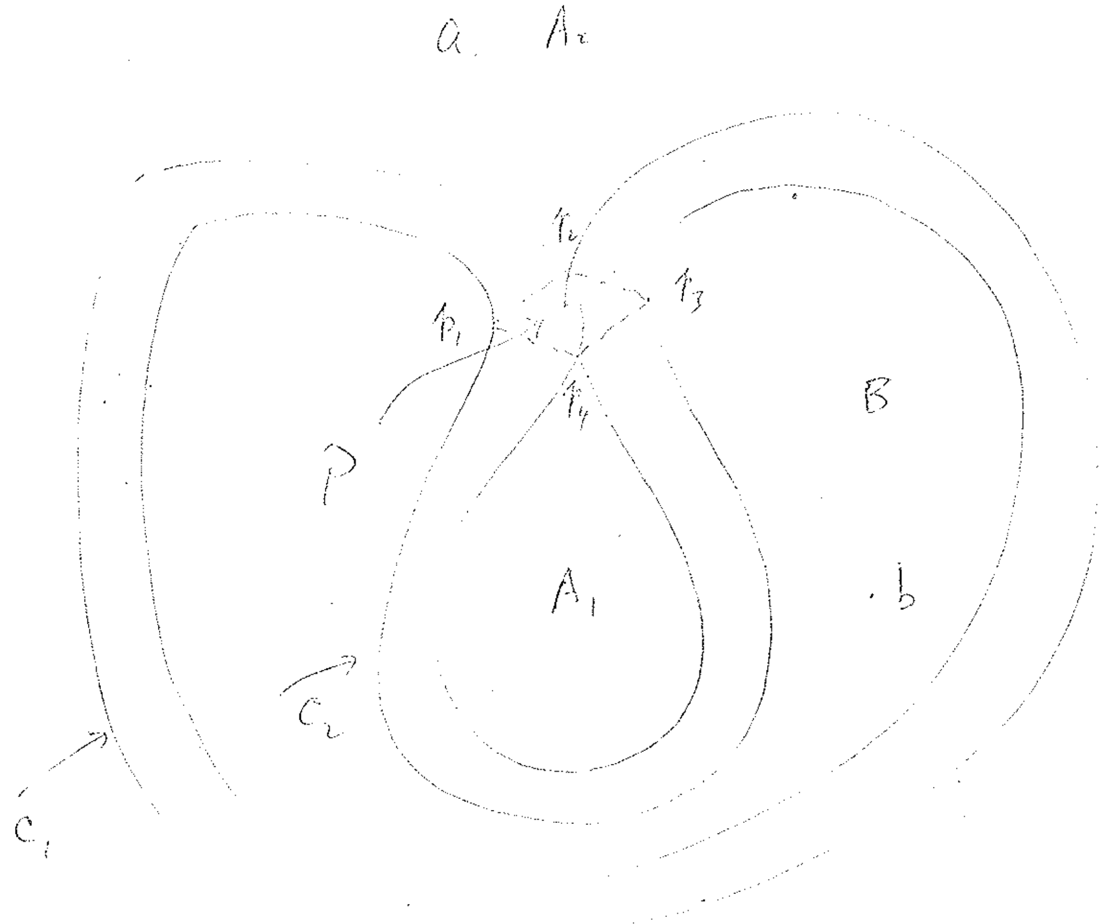



Fig. A1

Another irreducible separating strip is shown in Fig. A2.

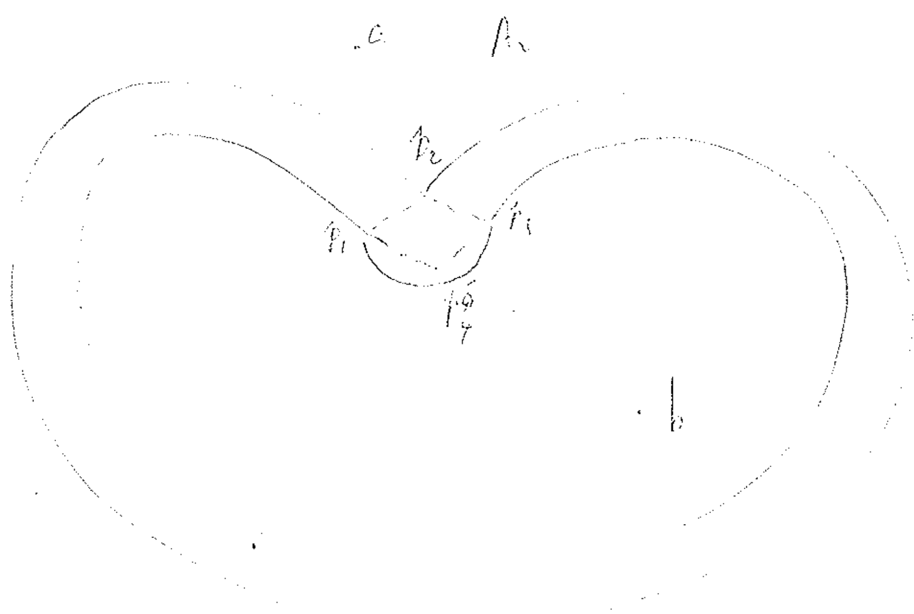



Fig. A2

Thus whether the strip in Fig. A1 or Fig. A2 is the minimum strip depends on whether the total weight of the strip from $p_1p_4$ to $p_3p_4$ in Fig. A1 or the total weight on the area in $p_1\ p_4\ p_3\ p_4'$ in Fig. A2 is smaller.

Let the area of the strip from $p_1p_4$ to $p_3p_4$ in Fig. A1 be $Q_1^*$ and let the area of $p_1p_4p_3p_4'$ in Fig. A2 be $Q_2^*$, then the ratio $Q_2^*/Q_1^*$ is bounded from above since it takes certain area to bend the strip from $p_1p_4$ to $p_3p_4$. Let this upper bound be $k_1$. In general, the area of the irreducible strip overlapping itself is always larger than the irreducible strip not overlapping itself, although the total weight of the former may not be larger than that of the latter. For certain weighting functions, say a constant, any minimum weight separating strip will not overlap itself. Therefore, we are interested in finding out the restriction on the weighting function, such that the minimum weight separating strip will not overlap itself.

Let $Q_1$ and $Q_2$ be the two areas that have nonempty intersections. Let $\bar{w}(Q_1)$ and $\bar{w}(Q_2)$ denote the total weights on the area $Q_1$ and $Q_2$, respectively. If $Q_2/Q_1 \leq k_1$ implies $\bar{w}(Q_2) \leq \bar{w}(Q_1)$, then no minimum weight strip will overlap itself. In order that $\bar{w}(Q_2) \leq \bar{w}(Q_1)$, it is necessary to require $w(x, y) \geq c$, where $c$ is a positive constant. Otherwise, $w(x, y) = 0$ in $Q_1$ and $w(x, y) \neq 0$ in $Q_2$ will violate the inequality $\bar{w}(Q_2) \leq \bar{w}(Q_1)$. Furthermore, it is necessary that the total weight on any area be bounded from above so that $\bar{w}(Q_2)$ cannot be infinity. The following is a set of sufficient conditions that will assure non-overlapping of the minimum weight strip.

(i) $w(x, y) \geq c$, $c$ is a positive constant.

(ii) The weighting function satisfies the Lipschitz condition with constant $K$.

(iii) The constant $K$ satisfies $K \leq 6c/r$, where $c$ is the positive constant in (i) and $r$ is the width of the minimum separating strip. (Note that if $r \to 0$, $K$ can be arbitrarily large).

In Fig. A3, we show the enlarged picture of the portion of an overlapping strip superposed together with the nonoverlapping one.

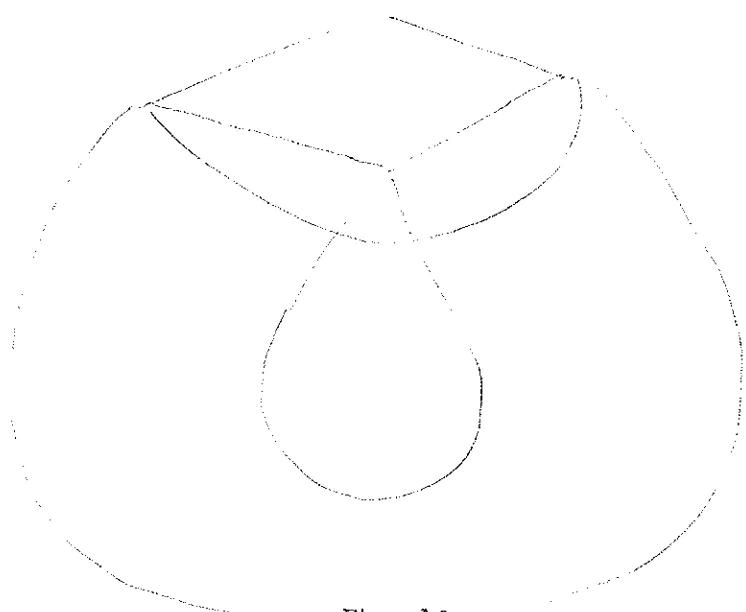


Fig. A3

The lighted shaded area is $Q_2^*$ and the heavy shaded area is $Q_1^*$. The lower bound on the total weight of $Q_1^*$ is obtained by letting

$w(x,y) = c$ throughout $Q_1^*$. The upper bound on the total weight on $Q_2^*$ is obtained by letting the weighting function grow as large as possible from the intersection point of $Q_1^*$ and $Q_2^*$. Note that the area in $Q_1$ is of order $O(r^2)$ and the area of $Q_2$ is of order $O(r^2)$. The total weight of $Q_1$ is $O(r^2)$ and the total weight of $Q_2$ is $O(r^2) + O(r^3)$.

The idea of a separating strip can be generalized to the case where we want to separate two sets of points instead of two points. In the rectangular region in §12.1, we want to separate the line S and the line T. This is equivalent to the r-separating in $R_2$ if we define $w(x,y)$ to be zero outside the rectangular region.

If we trust our intuition in three-dimensional space, the corresponding r-separating set in $R_3$ is a surface of thickness r and this leads us to Plateau's problem [ ].

Plateau's problem [ ] is to find the surface of least area spanned in a given closed Jordan curve $\gamma$. Related problems are to find minimal surfaces of least area when the whole boundary or part of it is not prescribed but left free on a given manifold. If we define a weighting function at every point of the space, we can ask the minimum-weight surface spanned in a given closed Jordan curve $\gamma$. The weight of a surface is defined to be the integral of the function w on the surface.

We shall first consider the problem of minimum-weight surface with free boundary. Consider a surface P which is topologically equivalent to a sphere. The inside open region is denoted by R where the open region plus the boundary surface $P$ is denoted by $\bar{R}$. On the surface $P$, we have two special subsurfaces $\Gamma_s$ and $\Gamma_t$. If a weighting function is defined on $\bar{R}$ and say zero elsewhere, then we can ask for the minimum weight surface separating $\Gamma_s$ and $\Gamma_t$. If $P - \Gamma_s - \Gamma_t$ is a curve $\gamma$ and the weighting function is a constant in $\bar{R}$, the problem of finding a separating surface of minimum-weight becomes Plateau's problem. In the approximation by a finite network, we first find a minimum weight surface with thickness $r$ and then let $r$ go to zero.