Topology of Thin Film RC Circuits

By F. W. SINDEN

(Manuscript received August 31, 1966)

Integrated RC circuits can be made by depositing exceedingly thin metallic and dielectric films in suitable patterns on an insulating substrate. Resistors are strips of conductor; capacitors are patches on which conducting, dielectric, and conducting layers are superimposed. Since conductors can cross at capacitor patches, RC networks need not be strictly planar to be realizable in thin film.

Determining which RC circuits are realizable poses new problems in topology which are remarkably simple to state but are as yet unsolved. The results reported here are fragmentary, but they do cover some cases of small order that may be of practical interest.

I. INTRODUCTION

Integrated RC circuits can be made by depositing exceedingly thin metallic and dielectric films in suitable patterns on an insulating substrate. A resistor is made by depositing a long, narrow strip of conductor (usually in a zag-zag for compactness); a capacitor is made by superimposing conducting, dielectric, and conducting layers. Because the dielectric is thin, the capacitance per unit area is high. Fig. 1 shows a typical thin film pattern.

Ordinarily printed circuits are strictly planar; crossovers are made only by leading one of the conductors entirely out of the plane of the circuit. In the thin film technique, however, conductors can be separated by thin insulating layers within the plane of the circuit. Thus, crossovers can be permitted provided a nonzero capacitance between the crossing conductors is acceptable. If an RC circuit can be laid out so that conductors cross only if the circuit requires a nonzero capacitance between them, we will say the circuit is realizable in thin film or just *realizable.*

An example of a realizable nonplanar circuit is shown in Fig. 2. In this case, the schematic thin film layout brings out intrinsic symmetries not displayed by the circuit diagram.

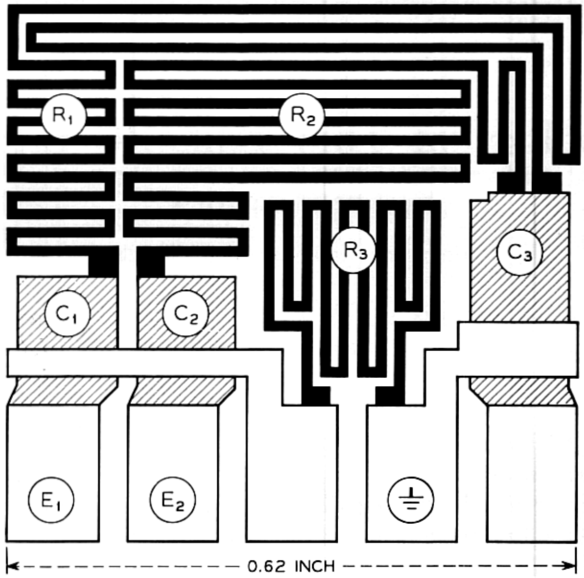


Fig. 1 — Thin film layout for a notch filter (courtesy W. H. Orr). Black region is bottom conductor; shaded region is dielectric; white region is top conductor.

Finding feasible layouts, or even determining when they exist, leads to unsolved problems in topology. The results presented here give answers only in special cases. Moreover, these results concern only the topological side of the problem; *electrical* equivalences are not taken into account. It is assumed that the network is given topologically and that

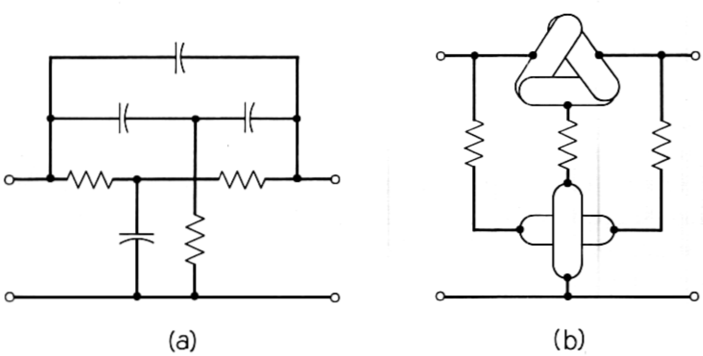


Fig. 2 — (a) Nonplanar circuit ("twin-tee", Ref. 3, p. 309); (b) schematic thin film layout for the circuit in (a).

terminals to the outside are located in given fixed positions on the periphery of the board.

II. SEPARATION OF THE RESISTIVE AND CAPACITIVE PARTS

Given an RC network N, let R_N be the purely resistive network obtained by replacing every capacitor by a direct connection. Clearly N is not realizable in thin film unless R_N is. R_N is realizable only if its graph (a vertex for each conductor, an edge for each resistor) is planar under the restrictions imposed by the locations of the terminals to the outside (see Fig. 3). This observation provides a first check: if R_N is not planar, there is no need to proceed further.

Each vertex in the graph of R_N replaces a purely capacitive network. In Fig. 3, for example, the vertex V in R_N replaces the network shown in Fig. 4.

One way to construct a realization of N is to construct realizations for the individual vertex-networks, and then to fit these into the planar layout of R_N. Since the layout of R_N may not be unique (there may be more than one ordering of edges about a vertex) the conditions on the vertex-networks may not be unique.

Another approach, discussed briefly in the final section, is to modify algorithms for purely capacitive networks to take account of resistors. In either case, one needs to study the purely capacitive networks first.

III. PURE C NETWORKS

A pure C network is a set of zero-resistance conductors $c_1, \cdots, c_r$ some pairs of which are connected by capacitors. The problem of finding a feasible layout for such a network is the following:

For each conductor c_i find a connected region R_i in the plane such that

(*i*) R_i and R_j have common points if and only if c_i and c_j are connected by a capacitor, and

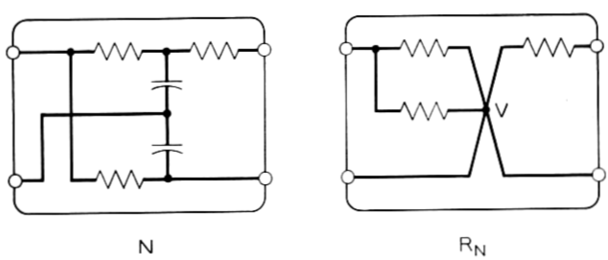


Fig. 3 — Nonplanar RC network N and reduced purely resistive network R_N .

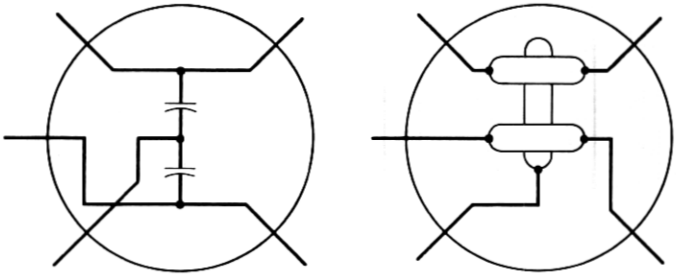


Fig. 4 — Capacitive network for vertex V of Fig. 3 and realization of this network.

(*ii*) no point belongs to more than two regions.

Condition (*ii*) says that no more than two conductors (separated by dielectric) may be superimposed. If, contrary to condition (*ii*), conducting and dielectric layers can be stacked up indefinitely, then *every* connected C network has a feasible layout. (The network is *connected* if any conductor can be reached from any other through a sequence of capacitors.) This is not quite immediately obvious; a proof is given in Appendix A.1.

Indefinite stacking offers other advantages as well.[1] Unfortunately it also presents technical difficulties. To date most thin film circuits have been limited to two conducting layers.

It does not change the problem to replace the connected regions R_i by curves C_i of finite length, since a connected region can be nearly filled by a curve of finite length, and a curve of finite length can be approximated by a narrow region. When convenient, the curves can have branches, although this is not necessary, since a branch can be approximated by letting the curve double back. In some cases, a pair of curves, whether branched or not, have to cross more than once (examples later). Such multiple crossings will be permitted on the assumption that a capacitance, if need be, can be distributed over several crossovers. Sometimes the curves are more convenient and sometimes the regions. I will use both.

In addition to satisfying conditions (*i*) and (*ii*) the regions (or curves) may have to satisfy constraints associated with the terminals to the outside. More specifically, $R_1, \cdots, R_r$ may be required to lie within a given region R and certain of the R_i may be required to contain specified points P_i on the boundary of R. I will consider mainly the two extreme cases where (*a*) there are no such terminal constraints and (*b*) every region R_i satisfies a terminal constraint.

IV. UNCONSTRAINED CASE

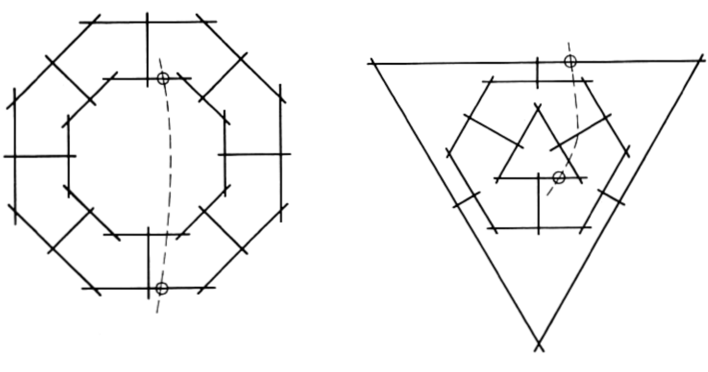
The problem is simply stated: It is specified which pairs of a set of curves (or connected regions) in the plane cross and which pairs do not. When are such specifications consistent?

To get a feeling for the problem, the reader may wish to try the examples in Fig. 5.

The crossings are conveniently specified by means of a graph G. Associate a vertex with each curve, and let two vertices be joined by an edge if and only if the corresponding curves are required to cross. If a set of curves satisfying the crossing specifications exists, we will say that the graph G is *realizable*.

If G is planar, then it is realizable. In a planar representation of G one has merely to replace each vertex v_i by a star-shaped region R whose points extend out along the edges emanating from v_i far enough to overlap the points of neighboring regions.

The converse is not true; some nonplanar graphs are realizable. For instance, any complete graph (nonplanar if the order is greater than four) is realizable, for in this case every curve C_i crosses every other. (Let the C_i be straight lines in general position; i.e., no two parallel, no three through a point.)



(a) (b)

Fig. 5 — Examples of unconstrained case. With the exception of the dashed curve, a pair of curves must cross if and only if they cross in the figure. The dashed curve must make only the encircled crossings. One of these examples has a solution; the other does not. Answers are given in Appendix A.2.

Although nonrealizable is different from nonplanar there is a class of nonrealizable graphs that is related to nonplanar graphs. If G is nonplanar, then the graph G^* obtained by inserting a new vertex into each edge of G is nonrealizable (see Fig. 6). If G^* were realizable, one could construct a planar representation of G as follows. In a realization of G^* let each of the curves C_i corresponding to an original vertex of G shrink to a point in such a way that no new crossings are generated. This is always possible. Since by assumption the remaining curves (corresponding to edges of G) do not cross each other, the resulting figure is a planar representation of G.

A theorem of Kuratowski[2] states that any nonplanar graph can be reduced to one of two minimal nonplanar graphs G_1 or G_2 (Fig. 7) by (*i*) deleting edges and (*ii*) combining adjacent vertices.†

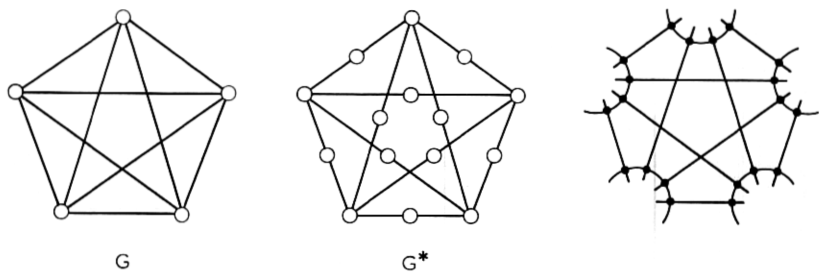


Fig. 6 — G is nonplanar; G^* is nonrealizable. On the right is a nonrealization of G^*; crossings marked with dots are required, no others are permitted.

The two operations (*i*) and (*ii*) clearly preserve planarity. Operation (*ii*) also preserves realizability, but (*i*) does not. (If it did, all graphs would be realizable, since any graph can be constructed by deleting edges of a complete graph, which is realizable.) To preserve realizability it is necessary to replace (*i*) by the weaker operation (*i'*): deleting *vertices* (together with attached edges). To see that (*i'*) and (*ii*) do indeed preserve realizability one has only to interpret them as operations on the curves C_i.

Using operations (*i'*) and (*ii*) and Kuratowski's theorem we can identify a class of nonrealizable graphs as follows.

Let G_1^* and G_2^* be the graphs obtained by inserting a new vertex

† G_1 is the graph involved in the familiar problem of connecting three utilities (e.g., the gas, water, and electric plants) to three houses without crossing lines. Since G_1 is nonplanar there is no solution. In Fig. 7 vertices 1, 3, and 5 can be taken as the utilities and 2, 4, and 6 as the houses.

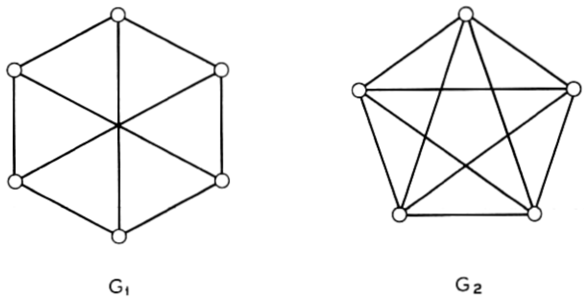


Fig. 7 — Kuratowski graphs.

into each edge of the Kuratowski graphs G_1 and G_2. A graph is nonrealizable if it can be reduced to G_1^* or G_2^* by application of (*i'*) and (*ii*). G_1^* and G_2^* are themselves irreducible. In Appendix A.2 one of the examples in Fig. 5 is shown to be reducible to G_1^*, hence nonrealizable.

The analogue of Kuratowski's theorem which would say that every nonrealizable graph can be reduced to G_1^* or G_2^* is false. An example of a nonrealizable graph that cannot be so reduced is given in Appendix A.3.

V. CONSTRAINED CASE

In addition to satisfying the conditions (*i*) and (*ii*) in Section III, the curves C_i (or the regions R_i) will now be required to lie within a simply-connected region R (which we shall take to be a disk) and each C_i will be required to contain a specified point P_i on the boundary of R. (This covers the case where a single conductor is required to join two or more separate terminals. One has only to require that the corresponding curves cross each other; their union represents the conductor.)

Before proceeding further, the reader may wish to try the examples in Fig. 8.

In passing, we observe that any constrained problem can be imbedded in an unconstrained problem. The constraints can be simulated by means of a ring structure containing $2r$ curves, where r is the number of curves in the constrained problem. This is proved in connection with the example discussed in Appendix A.3. Unfortunately, this observation is of little use in the absence of more information about the unconstrained case.

We will regard the vertices $v_1, \cdots, v_r$ of graph G as residing at the terminal points $P_1, \cdots, P_r$. We will often make use of the complement

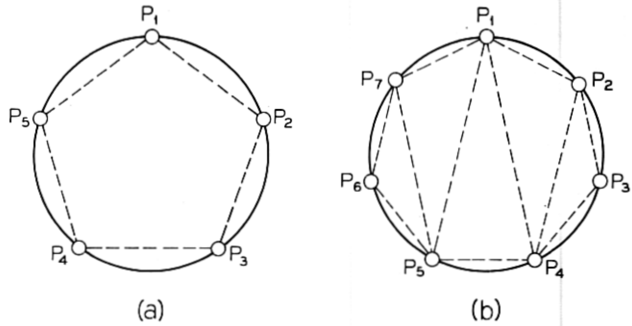


Fig. 8 — Examples of constrained case. Curve C_i must contain point P_i and lie otherwise within the circle. The dashed lines show the edges *not* in G, i.e., if P_i and P_j are connected by a dashed line then curves C_i and C_j may *not* cross; otherwise C_i and C_j *must* cross. One example has a solution, the other does not. Answers in Appendix A.4.

$\bar{G}$ of G, where $\bar{G}$ consists of all edges *not* in G. Edges in G will be shown as solid lines, edges in $\bar{G}$ as dashed lines.

A subset of vertex points $P_{i_1}, \cdots, P_{i_n}$ such that $i_1 < i_2 < \cdots < i_n$ will be called a *cycle* if all the pairs

$$(P_{i_1}, P_{i_2}),(P_{i_2}, P_{i_3}), \cdots, (P_{i_n}, P_{i_1})$$

are joined by edges. A cycle will be called empty if no other pairs are joined by edges. We will be primarily concerned with empty cycles in the complementary graph $\bar{G}$. (See Fig. 9)

Theorem 1: A necessary condition for a constrained graph G to be realizable is that $\bar{G}$ contain no empty cycles of order four or more.

Proof: (*i*) If G is an empty cycle of order four, then G is not realizable. This is easily verified by inspection. If, therefore, $\bar{G}$ *contains* an empty cycle of order four, then G is not realizable.

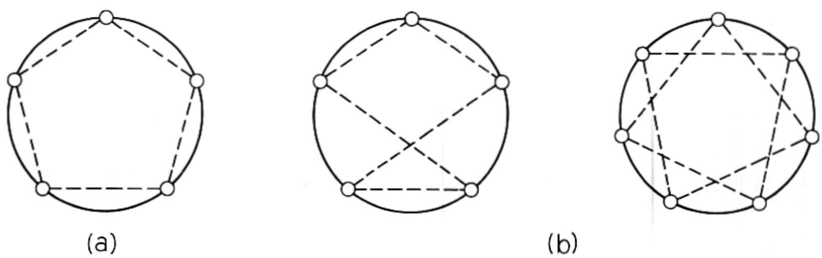


Fig. 9 — (a) Empty cycle in $\bar{G}$, (b) non-cycles. Dashed edges belong to $\bar{G}$; edges not shown belong to G.

(*ii*) Suppose the theorem is known to be true for cycles of order 4, $\cdots$, $m - 1$ and suppose, contrary to the theorem, that $\bar{G}$ contains an empty cycle of order m and that G is realizable. The realization of G can be generated in the following way: let curve C_1 grow continuously out of point P_1 until it reaches its full length, then let curve C_2 grow out of point P_2 until *it* reaches *its* full length, and so on until all curves are complete.

Let $\bar{G}(t)$ be the corresponding complementary graph at time t. At the beginning, $\bar{G}(t)$ is the complete graph (no crossings); as the crossings are generated one by one, edges are deleted from $\bar{G}(t)$. At some stage the postulated empty cycle of order m, which is contained in the final form of $\bar{G}$, must have just one internal edge left. But this last internal edge forms two empty cycles inside the final cycle, at least one of which must be of order four or more (since $m > 4$) and less than order m. Therefore, by the induction hypothesis, there can be no realization at this intermediate stage. Contradiction.

For some time it appeared to me that the empty cycle condition was not only necessary for the realizability of a constrained graph, but sufficient as well. Recently, though, I found a counterexample of order eight. This example is discussed in Appendix A.5.

Following are a number of results that help to identify and construct special classes of realizable constrained graphs. Taken together these seem to cover most cases of small order.

If no two edges of G cross, then clearly G is realizable. Less obvious is a similar result for $\bar{G}$:

Theorem 2: A sufficient condition for a constrained graph G to be realizable is that $\bar{G}$ contain no empty cycles of order four or more and that no two edges of $\bar{G}$ cross.

An example of such a $\bar{G}$ is the triangulated polygon of Fig. 8(b). This example, typical of the genre, has a complicated solution with unavoidable multiple crossings.

Theorem 2 is proved in Appendix B. A more general result, also proved in Appendix B, is the following:

Theorem 3: (*i*) *If (P_1, P_k) is an edge of G that crosses no other edges of G, and if the subgraphs G' with vertices $P_1, P_2, \cdots, P_k$, and G'' with vertices $P_k, \cdots, P_r, P_1$ are both realizable, then G is realizable.*

(*ii*) *If (P_1, P_k) is an edge of $\bar{G}$ that crosses no other edges of $\bar{G}$, and if subgraphs G' with vertices $P_1, P_2, \cdots, P_k$, and G'' with vertices $P_k, \cdots, P_r, P_1$ are both realizable, then G is realizable.*

The following two theorems describe circumstances under which a new curve C_{r+1} can be added to an existing solution. In many cases the entire solution can be generated by adding curves one at a time.

Theorem 4: Let G be a constrained graph with vertices $P_1, \cdots, P_r, P_{r+1}$. *G is realizable if (i) the subgraph of G with vertices* $P_1, \cdots, P_r$ *is realizable, and (ii) there do not exist three vertices* P_i, P_j, P_k, $i < j < k < r + 1$ *such that* $P_{r+1}P_i$ *and* $P_{r+1}P_k$ *are edges of* $\bar{G}$ *and* P_iP_k *and* $P_{r+1}P_j$ *are edges of G. (See Fig. 10.)*

Though cumbersome to state, this theorem is usually easy to apply. The following special cases are often useful by themselves. Let S be the set of vertices joined to P_{r+1} by edges of $\bar{G}$. Special case 1: the vertices of S are an adjacent string. Special case 2: every pair of vertices in S is joined by an edge of $\bar{G}$. Special case 2, for instance can solve examples like 8(b) in which $\bar{G}$ is a triangulated polygon. One has only to add new vertices one at a time in such a way that each additional vertex forms one new triangle in $\bar{G}$. The set S always has just two members.

Theorem 4 is proved in Appendix B. Though somewhat involved when worked out in detail, the idea of the proof is simple. In the situation of Fig. 10 the curves C_i and C_k (emanating from P_i and P_k) form a barrier which C_{r+1} cannot cross. This does not necessarily prevent C_{r+1} from intersecting C_j, for it is possible that C_j could cross the barrier. If, however, the barrier is not there, then C_{r+1} can reach C_j on its own without C_j's help. If there are no barriers of the Fig. 10 type, then C_{r+1} can reach all of the curves it is supposed to cross no matter how these may have been drawn. Thus, the new curve C_{r+1} can be added without disturbing the old ones.

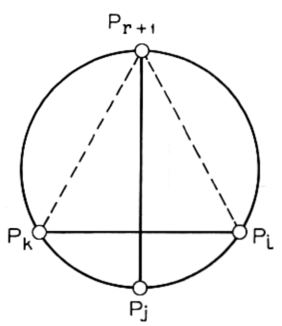


Fig. 10 — Configuration forbidden by hypothesis of Theorem 4. Dashed lines show edges of $\bar{G}$; solid lines show edges of G.

The next theorem concerns an operation which I will call an *adjacent interchange*. Given the circle R with the peripheral points $P_1, \cdots, P_r$, let R' be a slightly smaller circle concentric to R with corresponding peripheral points $P_1', \cdots, P_r'$. Let the primed points have the same order as the unprimed points except for one adjacent pair P_k', P_{k+1}', which is interchanged. The points $P_1, \cdots, P_r$, can be joined, respectively, to $P_1', \cdots, P_r'$ by curves $C_1 \cdots, C_r$ in such a way that only C_k and C_{k+1} cross. (See Fig. 11.)

If the operation is repeated by means of a new circle R'' inside R', then the curves C_i are extended inward and one new crossing is generated. A sequence of such operations can be specified by giving the pair of currently adjacent points that is to be interchanged.

Theorem 5 states the conditions under which all of the intersection requirements of a curve can be satisfied by a sequence of adjacent interchanges. These conditions involve cycles in $\bar{G}$ (not necessarily empty) as defined just before Theorem 1. Note that the order of vertices in a cycle of $\bar{G}$ is invariant under adjacent interchanges.

We will say that a member P_i of a cycle in $\bar{G}$ is *active* if it is joined to some other member of the cycle by an edge of G.

Theorem 5: The intersection requirements of a curve C_i can be satisfied entirely by a sequence of adjacent interchanges if and only if P_i is not an active member of any cycle in $\bar{G}$.

Theorems 4 and 5 tend to be complementary; where one fails, the other often works. Fig. 8(b) is an example where Theorem 5 fails (every vertex is an active member of several cycles) and Theorem 4 works.

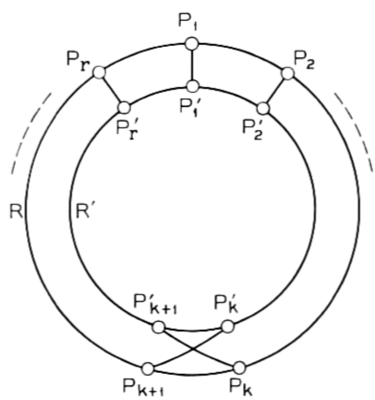


Fig. 11 — An adjacent interchange.

An example of the opposite kind is shown in Fig. 12. In this example Theorem 4 fails (every vertex has the forbidden configuration) but Theorem 5 works. The whole realization can be constructed by adjacent interchanges.

A realizable example to which neither Theorem 4 nor Theorem 5 applies is given in Appendix A.6. This is the smallest such example I have found (twelve vertices), but I doubt that it is really minimal.

VI. ORDER OF CROSSINGS

It is possible to obtain directly from the graph G information about the order in which crossings must occur along a given curve C_i. This information is contained in configurations I will call *empty chains.*

An *empty chain* is a subset of vertex points $P_{i_1}, P_{i_2}, \cdots, P_{i_n}$ in cyclic order such that the pairs (P_{i_1}, P_{i_2}), (P_{i_2}, P_{i_3}), $\cdots$, $(P_{i_{n-1}}, P_{i_n})$ are joined by edges of $\bar{G}$ and all other pairs are joined by edges of G.

An empty chain is just an empty cycle with a gap in it. Since the empty cycle is nonrealizable, it is not surprising to find that the realization of the empty chain, though not quite unique, is tightly determined. (See Fig. 13.)

Theorem 6: Let $P_1, \cdots, P_n$ be the vertices of an empty chain. Along curve C_k the first crossings with $C_1, \cdots, C_{k-2}$ must occur in that order; the first crossings with $C_{k+2}, \cdots, C_n$ must occur in reverse order.

The proof is given in Appendix B.

Every empty chain of length four or more yields ordering information. If, for instance, P_1, P_2, P_5, P_7 is an empty chain, then C_1 must cross C_7 before it crosses C_5 and C_7 must cross C_1 before C_2. Since most

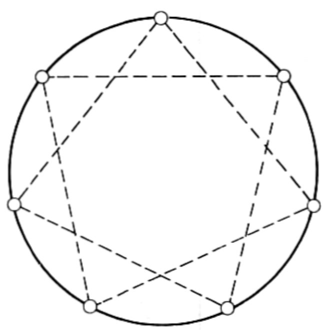


Fig. 12 — No vertex is an active member of any cycle in $\bar{G}$, therefore, a realization exists.

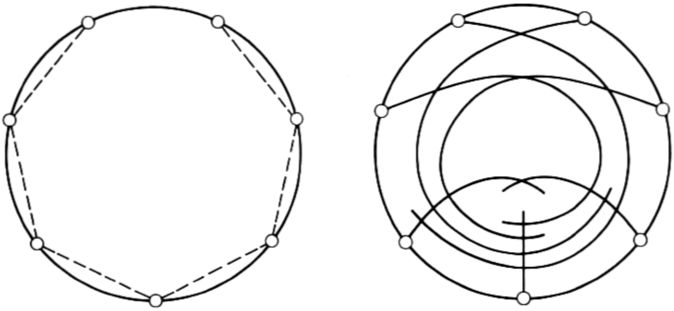


Fig. 13 — Realization of the empty chain of order seven.

examples of interest contain several such empty chains, this theorem is very generally applicable. The example of Fig. 8(b), for instance, contains six empty chains of order four and one of order five, which together give complete information about first crossings.

Searching for empty chains is tedious to do by hand, but could easily be done by machine.

A weakness of Theorem 6, evident in the example of Fig. 8(b), is that it says nothing about multiple crossings. It is clear in many examples that multiple crossings are determined by G. A way of extracting this information would be very useful.

VII. CONSTRUCTION OF SOLUTIONS — SUMMARY

The preceding results are not strong enough to define a guaranteed procedure for constructing realizations of constrained graphs. They do, however, seem to work in most cases of small order. To apply them one can proceed as follows:

(*i*) Look for empty cycles in $\bar{G}$ of order four or more. If any exist, G cannot be realized (Theorem 1).

(*ii*) Look for edges of G (or $\bar{G}$) that do not cross other edges of G (or $\bar{G}$). Such edges, if internal, permit the graph to be broken into two independent parts (Theorem 3).

(*iii*) Look for vertices that are free of the configuration shown in Fig. 10. Such vertices can be temporarily deleted, since the corresponding curves can be drawn in after the remaining curves have been drawn (Theorem 4).

(*iv*) Look for vertices that are not active members of any cycle in $\bar{G}$. These are typically on tree-like branches of $\bar{G}$. The corresponding curves can be constructed either at the beginning or the end by means of adjacent interchanges (Theorem 5).

(*v*) Find all the empty chains of order four or more and write down all the ordering relations they imply. Try to locate each crossing on both of its curves. This cannot always be done uniquely.

In a systematic procedure one could combine 1, 4, and 5 since these all involve chains and cycles.

Chains and cycles in $\bar{G}$ seem to be important in this problem; they certainly yield much information. But apparently they are not enough. To set up necessary and sufficient conditions for realizability, some other element is needed.

VIII. LOOSE ENDS

So far we have considered only completely constrained and completely unconstrained graphs, corresponding to networks where none or all of the conductors are connected to outside terminals. In general, of course, one wants the intermediate case where only some of the conductors are connected to outside terminals. This remains to be studied.

The preceding results can be used to construct realizations for the pure C networks represented by the nodes of the resistive network R_N. (See Section II.) Alternatively, one can generalize the pure C problem as follows to take account of resistors *a priori*.

The graph G can be replaced by its associated matrix A, where $a_{ij} = X$ (for "crossing") if conductors C_i and C_j are connected through a capacitor (or a short circuit) and $a_{ij} = 0$ (for "no crossing") if C_i and C_j are not so connected. To take account of resistors, we let $a_{ij} = T$ if C_i is connected to C_j through a resistor but not through a capacitor. This will mean topologically that C_i and C_j must touch without crossing.

T and X can be defined more precisely as follows. Consider instead of the curves C_i the regions R_i. We can assume that the R_i are simply connected. If $a_{ij} = T$, then the part of R_i's boundary that lies inside R_j must be connected (i.e., a single piece). If $a_{ij} = X$ then the part of R_i's boundary inside R_j may (but need not) consist of several pieces.

A. J. Goldstein has observed that in constructing an algorithm, the regions R_i have advantages over the curves C_i. (The ends of the curves have an unnecessarily special character.) He suggests that an algorithm might be constructed that would keep track of all of the pieces of the boundaries of the R_i and take, so far as possible, only steps that are topologically mandatory. Such an algorithm could easily take account of both T and X connections. This idea has not been worked out in detail and we do not know how often one would be forced to take an arbitrary step that might be wrong.

APPENDIX A

Examples and Answers

A.1 If indefinite stacking of conducting and dielectric layers is permitted, then any connected G is realizable regardless of the positions of the outside terminals. A universal realization can be constructed as follows.

Since G is connected, there is a path in G that contains every vertex at least once. In their order along this path let the vertices be $v_1, \cdots, v_n$. Over a disk D, stack n layers of conductor separated by layers of dielectric. Associate the conductors with the vertices of G according to their order along the path. This is permissible since the conductors have nonzero capacitances only with their neighbors in the stack. These capacitances correspond to the edges in the path. An extension of any conductor can be brought out of the stack radially in any direction. Thus, any pair of conductors required to have a nonzero capacitance can be brought out together and superimposed in an arbitrarily long radial strip. Similarly, any conductor can be brought out in the appropriate direction to connect to an outside terminal.

Although this construction shows the existence of a topological realization, it would hardly do as a practical layout in every case, even if indefinite stacking were permitted. Some of the metrical difficulties can be overcome by substituting an annulus for the disk D, but even so, this construction should be regarded as an existence proof, not as a practical solution.

A.2 *Answers to the Examples in Fig. 5.*

The example (a) of Fig. 5, constructed by R. L. Graham, was the first nonrealizable example found. It turns out to be of the type discussed in the text. Its graph is shown in Fig. 14(a). By deleting vertices and combining adjacent vertices it can be reduced to the graph shown in Fig. 14(b), which is a Kuratowski graph with a vertex inserted into each edge. Therefore, the example is nonrealizable. (See discussion subsequent to Fig. 5.)

Example (b) of Fig. 5 has the solution shown in Fig. 15.

A.3 Fig. 16(a) shows a nonrealizable graph which does not contain either of the augmented Kuratowski graphs G_1^* or G_2^*. The outer ring

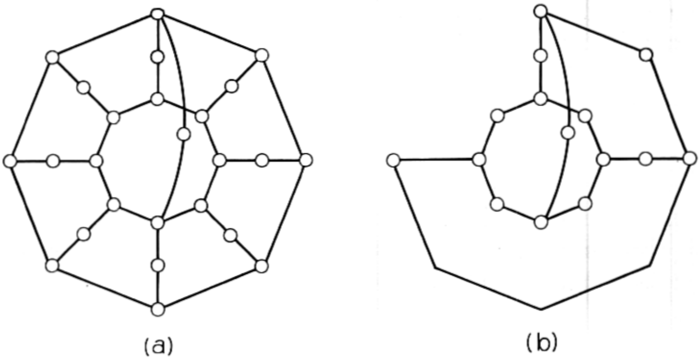


Fig. 14 — (a) Graph for example (a) of Fig. 5. (b) Reduced graph G_1^*.

(B and C vertices) simulates terminal constraints; the inner part (A vertices) is a constrained graph (empty cycle of order 5) that is known to be nonrealizable.

Proof: (*i*) The graph G of Fig. 16(a) cannot be reduced to G_1^* or G_2^*. The operations (*i'*) and (*ii*) always reduce the number of vertices. But G already has the same number of vertices (fifteen) as G_1^* and G_2^*. (*ii*) G is nonrealizable. Suppose a realization exists. In this realization let $\bar{C}$ be the union of C-curves (Fig. 16(b)). No A-curve intersects $\bar{C}$. Therefore all A-curves must lie in the same mesh of $\bar{C}$. Call the interior of this mesh R. R is (or may be) partitioned into subregions by segments of B-curves. We will show that all *intersections* between pairs of A-curves lie within the same subregion of R.

The A-curves may be indexed so that in the cycle A_1, A_2, $\cdots$, A_5, A_1 each curve intersects only its neighbors. Let I be an intersection between

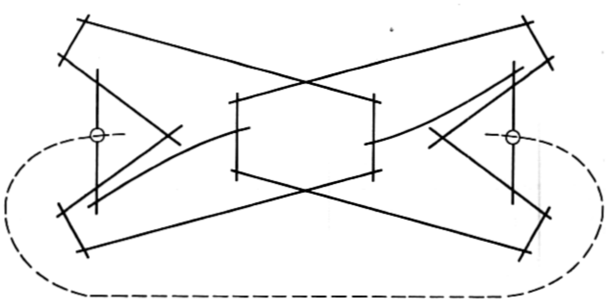


Fig. 15 — Solution to example (b) of Fig. 5. Both triangles can be drawn outside the hexagon.

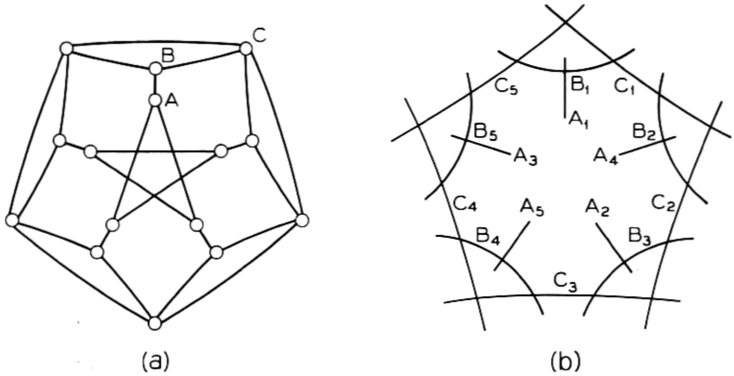


Fig. 16 — Nonrealizable graph which does not contain either of the augmented Kuratowski graphs G_1^* or G_2^* and a partial realization.

A_i and $A_{i+1(\text{mod } 5)}$ and let J be an intersection between A_j and $A_{j+1(\text{mod } 5)}$. There exist two distinct paths along A-curves joining I and J. One path P_1 traverses segments of $A_{i+1}, A_{i+2}, \cdots, A_j$ and the other path P_2 traverses $A_i, A_{i-1}, \cdots, A_{j+1}$ (indices mod 5). (In case $i = j$, P_1 traverses A_{i+1} and P_2 traverses A_i.) The sets of A-curves represented in the two paths are disjoint. Since each B-curve can cross only one A-curve and cannot cross any other B-curve it is not possible for a continuous boundary made up of B-curves to cross both P_1 and P_2. Therefore, I and J cannot belong to different subregions.

Let R^* be the subregion to which all A-intersections belong. The boundary G of R^* is made up of segments of B and C curves. (Every B-curve is represented since every A-curve must intersect its corresponding B-curve and could leave R^* only at a point belonging to this curve.) If $b_1, b_2, \cdots, b_5$ are any points on G belonging to $B_1, B_2, \cdots, B_5$, respectively (indexed according to the BC cycle), then the points $b_1, \cdots, b_5$ must lie in cyclic order around G. If not, it is possible to find a subset of four out of cyclic order, say b_1, b_3, b_2, b_4. But b_1 is joined to b_2 by a path lying within B_1, C_1, B_2, and b_3 is joined to b_4 by a path lying within B_3, C_3, B_4. These paths cannot cross, yet must be outside R^*. This is not possible under the postulated ordering b_1, b_3, b_2, b_4. All other noncyclic orderings can be similarly ruled out.

The points at which the A-curves join G must, therefore, lie in the order determined by the BC cycle and all intersections between A-curves must lie within R^*. But these are the conditions of a constrained case known (Theorem 1) to be nonrealizable.

A.4 *Answers to Examples in Fig. 8*

Example (a) of Fig. 8 has no solution (empty cycle); example (b) has the solution shown in Fig. 17.

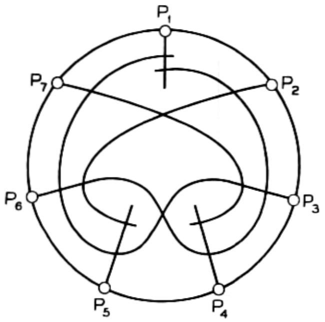


Fig. 17 — Solution to example in Fig. 8 (b). Curves C_3 and C_6 cross thrice. Multiple crossings are unavoidable in this example.

A.5 Counterexample to the conjecture that all constrained graphs free of empty cycles of order four or more are realizable. Fig. 18 shows the graphs G and $\bar{G}$ for this example and a near-realization in which only one required crossing does not occur.

The lack of empty cycles of order four or more can be verified by inspection; the nonrealizability can be shown as follows.

Consider curves 4 and 8, which do not cross. Since curve 2 crosses both of these, there exists a path from vertex 4 to vertex 8 traversing curves 4, 2, and 8. In case of multiple crossings, there may be more than one such path. We will assume that the path is chosen so that the segment of curve 2 contained in it has no crossings with curves 4 and 8 except at its endpoints. Since this path is to serve as a barrier, we will denote it by B_2.

There exists a similar path traversing curves 4, 6, and 8. We will call this one B_6.

Since curves 2 and 6 do not cross, the barriers B_2 and B_6 can have no points in common except along a single segment of curve 4 and a single segment of curve 8. Thus, the barriers must be related to each other in one of two ways shown in Fig. 19.

Curve 1 cannot cross B_2 and curve 5 cannot cross B_6. Thus, the barriers cannot be oriented as in case (a) of Fig. 19, for if they were curve 1 could not cross curve 5. By a similar argument, case (b) is eliminated by curves 3 and 7. Thus, neither case can occur; the example is nonrealizable.

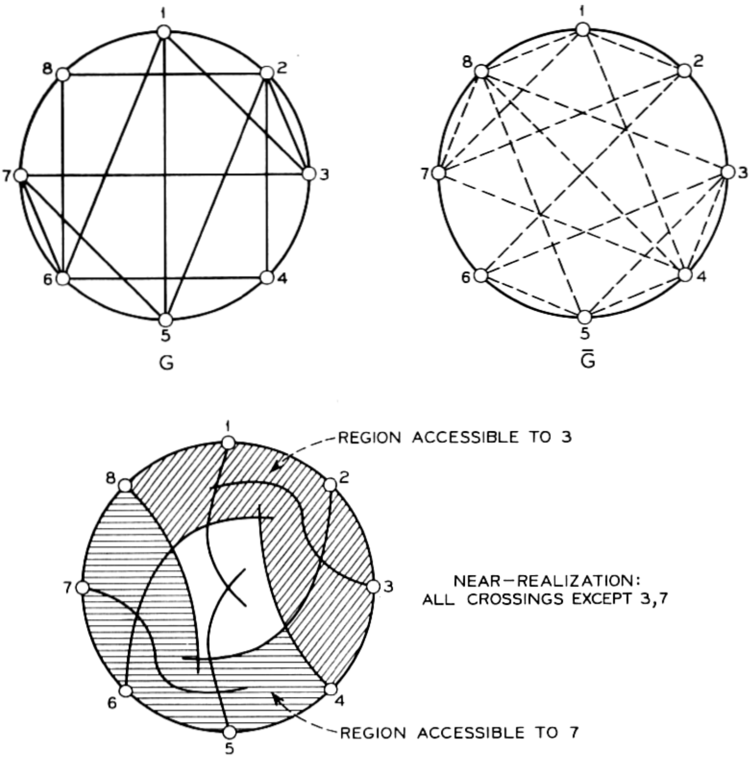


Fig. 18 — Counterexample. $\bar{G}$ contains no empty cycles of order four or more, yet G is not realizable.

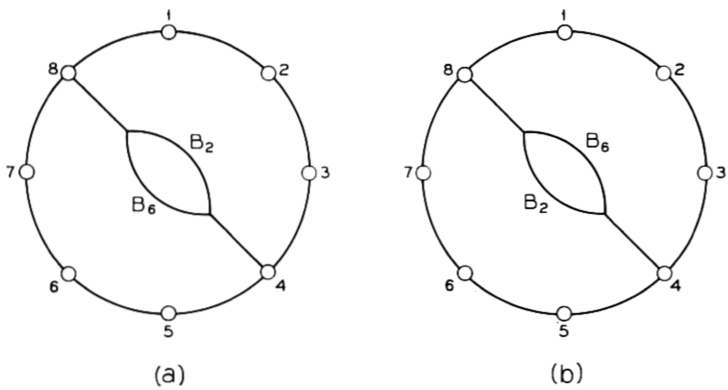


Fig. 19 — Proof of counterexample.

A.6 Fig. 20 shows the realization for an example to which neither Theorem 4 nor Theorem 5 applies. The ordering information supplied by Theorem 6 is very complete in this case. Only the order of curves 5 and 6 along curve 2 (and the symmetric counterparts) is unspecified. Indeed this could not be specified since either order is feasible. The order 6, 5 however, requires multiple crossings. The realization without multiple crossings is unique.

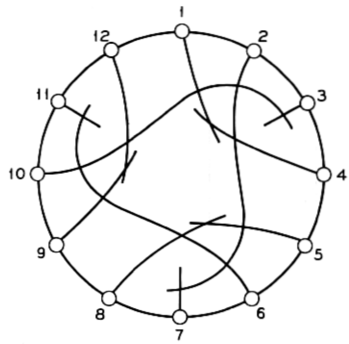


Fig. 20 — Realization for an example to which neither Theorem 4 nor 5 applies.

APPENDIX B

Proofs of Theorems

Theorem 2: A sufficient condition for a constrained graph G to be realizable is that $\bar{G}$ contain no empty cycles of order four or more and that no two edges of $\bar{G}$ cross.

Proof: The following proof depends on Theorems 3 and 5 whose proofs are independent.

The theorem is certainly true if G has three or fewer vertices. Suppose it is known to be true if G has m or fewer vertices. Consider a graph G with $m + 1$ vertices. If G satisfies the hypotheses of the theorem, then either $\bar{G}$ is an empty chain (see discussion preceding Theorem 6) or else $\bar{G}$ has an internal edge. If G is an empty chain, then by Theorem 5 it is realizable. If $\bar{G}$ has an internal edge then this edge separates $\bar{G}$ into two parts as defined in Theorem 3. Each of these parts has m or fewer vertices and is free of crossing edges and empty cycles of order four or more (by hypothesis), hence by the induction assumption is realizable. By Theorem 3, G is realizable.

Theorem 3: If (P_1, P_k) *is an edge of* G $(\bar{G})$ *that crosses no other edge of* G $(\bar{G})$, *and if the subgraphs* G' *with verticles* $P_1, P_2, \cdots, P_k$ *and* G'' *with vertices* $P_k, \cdots, P_r, P_1$ *are both realizable, then* G *is realizable.*

Proof: The method of proof is proof by picture (Fig. 21). *Case (a):* (P_1, P_k) is an edge of G crossing no other edges of G. None of the curves $C_2, \cdots, C_{k-1}$ crosses any of the curves $C_{k+1}, \cdots, C_r$. Therefore, except for C_1 and C_k the realizations of G' and G'' can be confined to separate parts of the disk R. C_1 and C_k can participate in both parts. (See Fig. 21(a).) *Case b:* (P_1, P_k) is an edge of $\bar{G}$ crossing no other edges of $\bar{G}$. Every one of the curves $C_2, \cdots, C_{k-1}$ crosses every one of the curves $C_{k+1}, \cdots, C_r$. The realizations of G' and G'' can be confined to the regions labelled with these letters in Fig. 21(b). The peripheral terminals for these realizations can be connected to the terminals on the periphery of the disk as shown in the figure. The connections to G' can cross G'''s region since this can only generate allowable crossings. The required crossings between curves of G' and curves of G'' occur in the center of the figure.

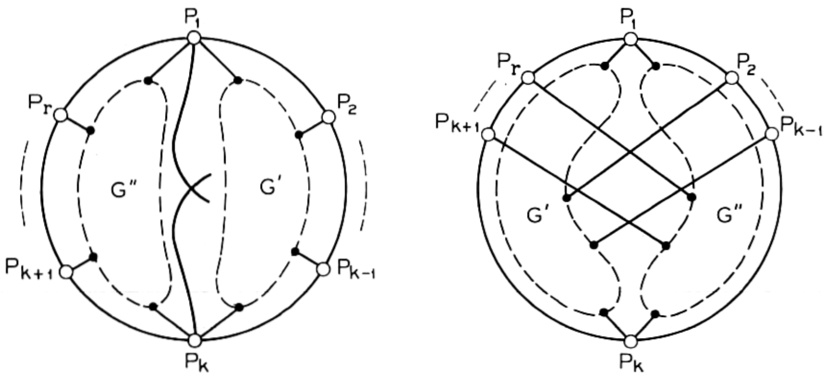


Fig. 21 — Proof of Theorem 3.

Theorem 4: Let G *be a constrained graph with vertices* $P_1, \cdots, P_r, P_{r+1}$. G *is realizable if (i) the subgraph of* G *with vertices* $P_1, \cdots, P_r$ *is realizable, and (ii) there do not exist three vertices* P_i, P_j, P_k, $i < j < k < r + 1$ *such that* $P_{r+1}P_i$ *and* $P_{r+1}P_k$ *are edges of* $\bar{G}$ *and* P_iP_k *and* $P_{r+1}P_j$ *are edges of* G. (See Fig. 10.)

Proof: We suppose that a realization for the subgraph with vertices $P_1, \cdots, P_r$ is at hand. It will be convenient to think of this realization as made up of regions $R_1, \cdots, R_r$ instead of curves. To simplify the notation later on we will designate the disk R to which the realization is confined by the indexed name R_0. We may assume that R_i intersects the boundary of R_0 only in the vertex P_i.

Let R^* be that connected piece of R_0 which contains P_{r+1} but is exterior to the regions that R_{r+1} may not intersect. R^* is the set of points that can be reached by R_{r+1}. We will show that the boundary of R^* contains all the vertices corresponding to regions R_{r+1} must intersect, i.e., all vertices joined to P_{r+1} by edges of G.

The boundary of R^* can be partitioned into a sequence of segments $S_1, \cdots, S_n$ where S_i belongs to the boundary of $R_{k(i)}$ and $k(i) \neq k(i+1)$. The segments S_1 and S_n adjacent to P_{r+1} belong to R_0, hence $k(1) = k(n) = 0$. If $k(i) = 0, 1 < i < n$, then S_i is that segment of the boundary of the disk R_0 which runs from $P_{k(i-1)}$ to $P_{k(i+1)}$. (The end cases $i = 1$ and $i = n$ can be included by defining $k(0) = k(n+1) = r+1$.)

Now suppose $P_{r+1} P_j$ is an edge of G (i.e., R_{r+1} must intersect R_j). We will show that P_j belongs to the boundary of R^*.

If $j > k(i), i = 1, \cdots, n$, then P_j belongs to S_n, hence to the boundary of R^*. If not, let i be the first index such that $j < k(i+1)$. It is not possible that $k(i) = j$ because R_j as a region that intersects R_{r+1} is not involved in the boundary of R^*. It is also not possible that $0 < k(i) < j$ for this would violate hypothesis (*ii*). ($R_{k(i)}$ intersects $R_{k(i+1)}$ because segments of their boundaries are adjacent.) Therefore, $k(i) = 0$. Hence, S_i runs from $P_{k(i-1)}$ to $P_{k(i+1)}$. Since $k(i-1) < j < k(i+1)$, S_i must contain P_j. Therefore, P_j is on the boundary of R^*, which was to be proved.

Theorem 5: The intersection requirements of a curve C_i can be satisfied entirely by a sequence of adjacent interchanges if and only if P_i is not an active member of any cycle in $\bar{G}$.

Proof: If: A *chain* is a sequence of vertices $P_{i_1}, P_{i_2}, \cdots, P_{i_n}$ in cyclic order such that (P_{i_1}, P_{i_2}), (P_{i_2}, P_{i_3}), $\cdots$, $(P_{i_{n-1}}, P_{i_n})$ are edges of $\bar{G}$. For the duration of this proof a chain must have at least three vertices.

Let the vertices be numbered in clockwise order and suppose P_1 is not an active member of any cycle in $\bar{G}$. Let S be the set of vertices joined to P_1 by edges of G. We will show that by a sequence of adjacent interchanges the members of S can be moved around the circle and finally interchanged with P_1.

S can be divided into three subsets:

(*i*) The clockwise set S_c : $P_k \varepsilon S_c$ if P_1 is joined to P_k by a chain whose intermediate members have indices between 1 and k.

(*ii*) The counterclockwise set S_{cc} : $P_k \in S_{cc}$ if P_1 is joined to P_k by a chain whose intermediate members have indices greater than k.

(*iii*) The rest S_R .

S_c and S_{cc} must be disjoint because otherwise P_1 would be an active member of a cycle. Let P_i be that member of S_c with highest index. P_i can be interchanged with all vertices with higher indices. Thus, it can be moved clockwise around the circle past P_1 . With P_i out of the way, the member of S_c with next highest index can also be moved clockwise past P_1 . The process can continue until all members of S_c have been interchanged with P_1 . Similarly, the members of S_{cc} can be moved counterclockwise past P_1 . The members of S_R can be moved either way. Hence, every member of S can be interchanged with P_1 , which was to be proved.

Only if: Suppose P_1 is an active member of a cycle. Then it is joined to another member P_k by an edge of G. P_1 cannot be brought adjacent to P_k because the order of vertices in a cycle is invariant under adjacent interchanges. Hence, P_1 cannot be interchanged with P_k .

Theorem 6: Let P_1 , $\cdots$, P_n be the vertices of an empty chain. Along curve C_k the first crossings with C_1 , $\cdots$, C_{k-2} must occur in that order; the first crossings with C_{k+2} , $\cdots$, C_n must occur in reverse order.

Proof: It is only necessary to prove the first part of the statement (concerning C_1 , $\cdots$, C_{k-2}) since the second part follows from the first by symmetry. The first part is trivially true if $k \leqq 3$. We assume then that $k > 3$.

The region bounded by C_{k-1} and C_{k-3} encloses C_{k-2} . (See Fig. 22.) Since C_k cannot cross C_{k-1} it must cross C_{k-3} before it can cross C_{k-2} .

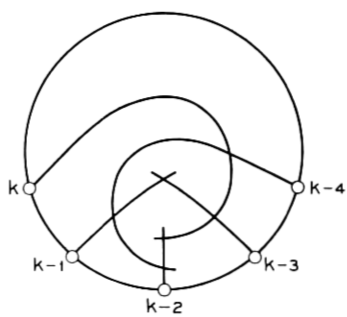


Fig. 22 — Proof of Theorem 6.

If $k > 4$, then there is a curve C_{k-4}. The region bounded by C_{k-2} and C_{k-4} encloses C_{k-3}. Therefore, C_k must cross either C_{k-2} or C_{k-4} before it can cross C_{k-3}. But by the previous argument it cannot cross C_{k-2} before C_{k-3}. Therefore, it must cross C_{k-4} before C_{k-3}. Since this argument can be iterated indefinitely, the theorem holds for arbitrary k.

REFERENCES

1. Feldman, D., private communication.
2. Berge, C., *Theory of Graphs and its Applications*, John Wiley & Sons, New York, 1962.
3. Balbanian, N., *Network Synthesis*, Prentice-Hall, 1958.