LINEAR GRAPHS AND ELECTRICAL NETWORKS

by

SUNDARAM SESHU

Department of Electrical Engineering
Syracuse University

and

MYRIL B. REED

Department of Electrical Engineering
Michigan State University

ADDISON-WESLEY PUBLISHING COMPANY, INC.

READING, MASSACHUSETTS, U.S.A.

LONDON, ENGLAND

Printed in the United States of America

Library of Congress Catalog Card No. 61-5309

To

The "book widows"

Lily and Georgia

Our graduate students

on whom our foil techniques have
been developed.

Our world today

It has allowed two different minds,
trained in two different ways,
in two different generations
and in two different hemispheres
to cooperate in producing this book.

PREFACE

This text has grown out of a graduate course entitled "Foundations of Electric Network Theory," organized at the University of Illinois by the second author in 1949. Such a course has since been taught by the two authors regularly at Illinois, Syracuse, and Michigan State Universities. Over the period of years, the material has naturally evolved into a shape quite different from the original. However, the basic philosophy of mathematical precision, coordinated with the objective of establishing the foundation of network theory, has remained unchanged throughout.

For many years, an intensive search was made (especially by the second author) for a way to determine precisely, rather than dimly suspect, the mathematical properties of the Kirchhoff equations of electrical network theory. In themselves, these equations seemed to be infinitely varied and to fit into no detectable pattern. Darkly at first, but with accelerated clarity as linear graph concepts were brought to bear, it became evident that here was the tool for the Kirchhoff-equations problem. In retrospect, it seems obvious that since the linear graph determines the coefficient matrices of these equations, it is in the linear graph that the properties of the equations are to be found.

Theory of graphs depends on the mathematical discipline of linear algebra, which is not very familiar to electrical engineers. We have kept this fact in mind and at least tried to explain briefly such concepts as *field*, *ring*, *linear vector space*, etc., that are used. However, we assume knowledge of matrix algebra and use it without explanation. Similarly, in the applications presented, Laplace transformation and theory of functions are assumed in the network theory, and Boolean algebra in the switching theory.

The guiding light throughout has been mathematical precision. However, there are some places where we have avoided making a fetish of precision, in the interest of readability. For example, in Chapter 1 we exclude isolated vertices from graphs, but we admit them in Chapters 3 and 8. Similarly, we exclude single-edge loops (or self-loops) in Chapter 1 but admit them in parts of Chapters 3, 9, and 10. Also, the vertex is bound to the edge in Chapter 1 and divorced from it when convenient in Chapters 3 and 4. There is, of course, no need to be so inconsistent. But we feel that these are places where the cure is worse than the disease. We could admit self-loops from the beginning and insert the hypothesis "if the graph does not contain any self-loops" into every theorem. We could call the edge minus the vertices by another name, say *arc*. Instead, we prefer to treat the exceptional cases individually by reminding the reader that the word is being used in a different sense rather than complicate the whole book by burdensome additional terminology. The same remark applies to notation as well. The symbols **A** and **B** are used for matrices of incidence and **Q** for the cut-set matrix, both in nonoriented graphs and in directed graphs, even though the elements are chosen from different fields in the two cases. Since the matrices in the two cases are very closely related and have identical properties, we feel there is an advantage in using the same symbolism.

The first five chapters contain the basic theory of graphs. There is no intention that these five chapters should constitute a treatise on graph theory. On the contrary, we have carefully omitted all aspects of graph theory that are unrelated to the applications discussed here. The relevant concepts are, however, discussed in much greater detail than they would be in a general treatise on graph theory. Of particular interest in the applications (considered here) are the matrices of the graph. Therefore we devote considerable space to the matrices of a graph.

The last five chapters, constituting almost two thirds of the book, discuss the various applications. Three of these chapters are devoted to electrical network theory, which happens to be the major field of interest of the authors. In each of these chapters, we assume that the reader is familiar with the elementary aspects of the subject and devote the discussions to those aspects of the theory that are strongly dependent on the theory of graphs.

The present text is aimed primarily at the advanced graduate student who has attained some mathematical maturity and has had at least one graduate course in network theory covering approximately the material in *Linear Network Analysis* by S. Seshu and N. Balabanian (John Wiley and Sons, New York, 1959). It is our sincere hope, of course, that research workers in the field of electrical networks and others utilizing the theory of graphs will find this material useful. The segregation of the theory of graphs from applications and the collection of applications into almost self-contained chapters has been made with the research worker in mind, at least in part.

It is virtually impossible to acknowledge everyone who has contributed directly and indirectly to this book. The most significant contributions have come from the many graduate students who worked through versions of the text in the form of preliminary notes—with the sole objective of making them obsolete. Thanks are also due to Professor M. E. Van Valkenburg, who read the manuscript critically and made valuable suggestions. Finally we wish to express our thanks to Professor W. H. Huggins and the Addison-Wesley Publishing Company for the inclusion of this book in the Systems Engineering series.

S. S.
M. B. R.

CONTENTS

CHAPTER 1

BASIC CONCEPTS

1–1 Survey of applications. In this text, a detailed study of the theory of graphs is presented first, before discussing any applications of the theory. This procedure, while being very satisfactory as a logical order, is unsatisfactory in another sense. One is not always aware of the need for the various concepts that are being introduced, or of the utility of the various theorems that are being proved. The purpose of this first section is to provide a little motivation by briefly surveying a few of the many applications of the theory of graphs. In this section, we anticipate many of the results that are rigorously proved later on. No precision is attempted in this section, the purpose being mainly to show the utilitarian aspects of graph theory. Abstract graph theory has its own beauty, of course, but this can be appreciated only after a detailed study.

From the point of view adopted here, the most important application of graph theory is in the physical science for which G. Kirchhoff formulated the theory of graphs, namely electrical network theory. Let us first attempt to clarify the concept of an electrical network; the process will bring out the concept of a graph. The laboratory electrical network consists of a number of devices with terminals. When an attempt is made to represent or, better, to *model* these devices in a *network diagram* on paper, combinations of two-terminal elements are usually used. Let us draw such a diagram as a concrete example, as in Fig. 1–1. What is meant by such a diagram? We mean, first of all, that six devices are in use, each of them having two terminals. They are interconnected as shown in the figure. For instance, one terminal of each of the devices 1, 2, and 4 are connected together. The other terminal of device 2 is con-

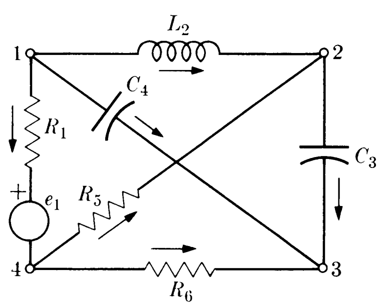


FIG. 1–1. Example of a network.

nected to one terminal of device 3 and to one terminal of device 5, etc. The various symbolic lines marked R, L, C, etc., indicate the relations between the voltages and currents associated with each of the devices. How these voltages and currents are to be measured is indicated by the arrows and plus signs.

The physical system has many other characteristics that are not shown on the network diagram, such as physical dimensions, color, weight, etc. A characteristic the omission of which is pertinent to the present discussion is the relative location of the components. The diagram of Fig. 1–1 does not imply that the six component devices are located in space relative to each other as shown. For instance, if Fig. 1–1 is redrawn as in Fig. 1–2, it is still a model of the same network. The important feature is the interconnection of the components and not their relative space location. A comparison of the underside of a broadcast receiver with its schematic will convince anyone of this fact. (Of course in the case of a high-frequency device, the relative location of components in the physical network is important, but such a consideration is not part of electrical network *theory*.) Thus a network diagram represents two (independent) aspects of an electrical network: the interconnection between components and the voltage-current relationships of each component. Network topology is primarily a study of the former aspect. Therefore, let us try to extract from Fig. 1–1 the information about the interconnection, or the network geometry, without regard to voltages, currents, and their interrelations. The interconnection of the network may be portrayed as in Fig. 1–3(a) or, more simply, as in Fig. 1–3(b). It must be remembered that the lines of Fig. 1–3(b) represent network elements and are not necessarily "short circuits." An interconnected system of line segments such as Fig. 1–3(b) is a *linear graph*, and *graph theory* is a study of such structures. The points 1, 2, 3, and 4 are the *vertices* of the graph and the line segments are the *edges* of the graph.

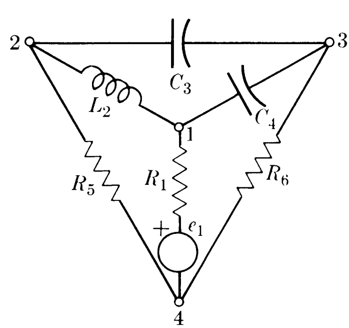


FIG. 1–2. A redrawing of Fig. 1–1.

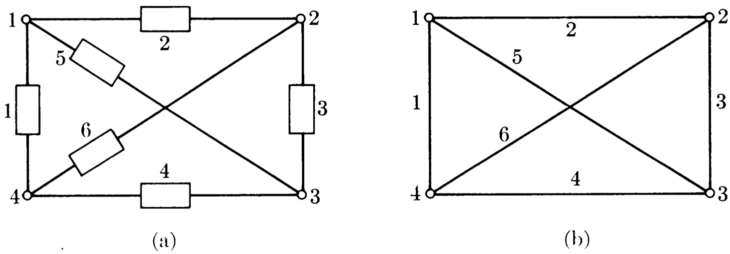


FIG. 1-3. Geometry of Fig. 1-1.

Let us ask what can be learned about electrical networks from a study of such structures. We have, first of all, Kirchhoff's original reasons for inventing graph theory. Namely, we can establish from such a study, in a mathematically rigorous fashion, the numbers of linearly independent current and voltage equations. We are acquainted with these numbers, because they have been established for a few simple cases and projected into the general case. But we cannot *prove* the validity of these statements in the general case without appealing to graph theory. Kirchhoff's laws can be written in matrix notation as

$$\mathbf{A}_a\mathbf{I}(t) = 0 \qquad \text{(Kirchhoff's current law)},$$

and

$$\mathbf{B}_a\mathbf{V}(t) = 0 \qquad \text{(Kirchhoff's voltage law)}. \tag{1-1}$$

The matrices $\mathbf{A}_a$ and $\mathbf{B}_a$ are the *matrices of incidence* of the linear graph, relating vertices to edges and edges to loops respectively. We prove later that these two matrices have ranks $v - 1$ and $e - v + 1$ respectively, where e is the number of edges and v is the number of vertices. Therefore $v - 1$ and $e - v + 1$ are also the numbers of linearly independent Kirchhoff's current and voltage equations. Notice that $\mathbf{A}_a$ and $\mathbf{B}_a$ are associated with the *graph*. Their ranks have nothing to do with currents and voltages. If, for instance, the linear graph represents a lumped mechanical system, with the vertices representing rigid bodies, exactly the same matrices $\mathbf{A}_a$ and $\mathbf{B}_a$ would arise for Newton's force equations and the displacement equations respectively (as for the electrical network). Then the same numbers, $v - 1$ and $e - v + 1$, represent the numbers of linearly independent force equations and displacement equations.

Secondly, it is possible to establish rigorously the validity of the loop and node systems of equations and find their generalizations. We can also find the conditions under which unique solutions can be found for these equations. Finally, we can justify the various duality procedures.

None of this is new; the present study merely permits a justification of familiar procedures.

But more than this is available. We discover short-cut methods of writing, by inspection from the network, the determinants and cofactors of the loop and node systems of equations, *without even writing out the equations.* For this, we need to know the structure of the coefficient matrices of the loop and node systems of equations

$$\mathsf{Z}_m = \mathsf{BZB}' \qquad \text{and} \qquad \mathsf{Y}_n = \mathsf{AYA}', \tag{1-2}$$

where A and B are again coefficient matrices of Kirchhoff's current and voltage equations, where not all the equations, but only the independent ones, are represented. The prime denotes the transpose. A and B depend only on the interconnections and hence are properties of the linear graph. Much of the material of the first part of this book is devoted to a study of their properties. The property that is used in the short-cut method is the relation between the nonsingular submatrices of A and B and the structure of the graph. The nonsingular submatrices of A correspond to the *trees* of the graph and those of B to complements of trees. A *tree* is a connected subgraph containing all the vertices (nodes) and not containing any loops, and the *complement* is the set of all elements not in the tree. Knowing this, we prove that the node determinant (det Y_n) is the sum of tree-admittance products and that the loop determinant (det Z_m) is the sum of impedance products of complements of trees. We also extend these formulas to the various cofactors.

More interesting than the short-cut procedures are the applications of graph theory to network synthesis. Several relationships are established in Chapter 8 between the topology of the network and the analytic behavior of the network functions. For instance, it is proved that a trans-

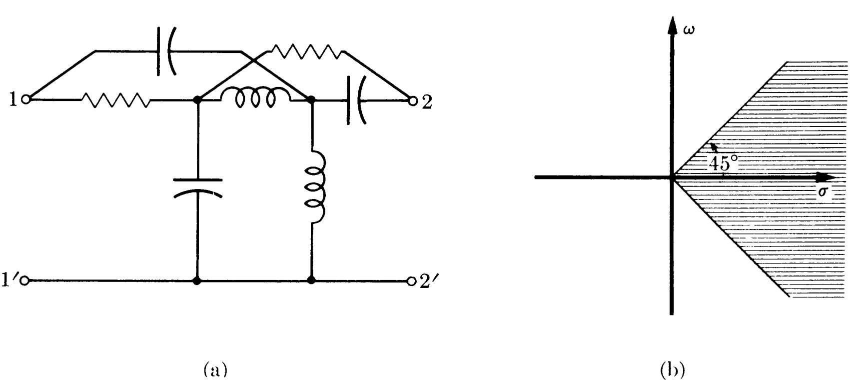


FIG. 1–4. Example in network synthesis.

formerless realization of a minimum positive real function cannot contain any paths between the input terminals, or cut-sets separating them, consisting only of one type of element (R, L, or C). As another example we can say, *by inspection* (that is without any computations), that all the zeros of transmission of the network of Fig. 1-4(a) are *outside* the shaded region shown in Fig. 1-4(b).

Another major application of graph theory, which is of interest to the electrical engineer, occurs in the theory of switching. Here graph theory finds significant applications in both combinational and sequential network theory. A combinational contact network has an obvious interpretation as a linear graph. For instance, the contact network of Fig. 1-5(a) has the linear graph representation shown in Fig. 1-5(b). Many graph-theoretic ideas are therefore applicable to contact network theory. For example, the switching function (in its so-called "admittance representation") is expressible as the sum of path products over all the paths between the input terminals. The zeros of the switching function correspond to the cut-sets separating the input terminals. Concepts such as duality carry over immediately. It is also possible to relate conventional networks and contact networks. The known methods of proving the minimality of a contact network depend very strongly on graph theory. We discuss the three known methods of proof of minimality. The first of these is due to C. Cardot, who showed that the parity function of n variables,

$$F(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \cdots \oplus x_n \tag{1-3}$$

(of which $n = 3$ is illustrated in Fig. 1-5), requires $4n - 4$ contacts for its realization. The second is a very elegant graph-theoretic argument due to C. E. Shannon, proving that the 18-contact realization of the 16 switching functions of two variables is minimal. The third is a matrix technique due to R. Gould.

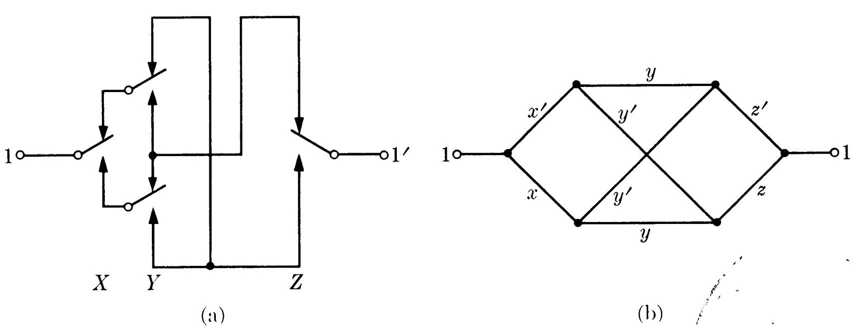


FIG. 1-5. (a) A contact network and (b) its graph.

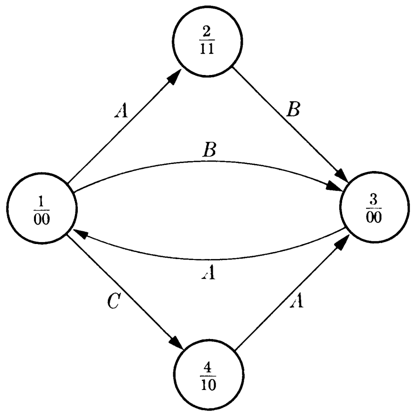


FIG. 1–6. A state diagram.

Graph theory (in the form of theory of nets) is also applicable to (electronic) logic-circuit representations of Boolean functions. However, not much work has been done on them, and so the subject is mentioned only briefly.

The best-known representation of a sequential switching system takes the form of a directed graph, known as a *state diagram*. An example of a state diagram is shown in Fig. 1–6. The vertices, which are now drawn as circles, represent the "states" of the machine (memory states, for instance). The edges represent transitions between states, with the input symbol causing the transition associated with the edge. The output of the state is associated with the vertex. Thus the representation becomes a *weighted directed graph* or a *net*. It is clear that the theory of directed graphs should play an important role in the theory of sequential machines.

The general concept of a net has many applications besides the theory of sequential machines. One of the most natural applications is to communication networks. The vertices represent stations and the edges represent channels of communication. "Communication network" is itself a general concept applicable to voice (or message) communication, oil or gas pipelines, railroads, highways, etc. The weights associated with the edges depend upon the particular application. They may be channel capacities (bits* per second or gallons per hour or cars per hour), probabilities of channel availability, etc. An interesting problem here is to compute the maximum rate of flow (of whatever is being communicated) from one given point of the communication network to another given

* Binary digits.

point. The solution, due to two independent groups of workers, Elias, Feinstein, and Shannon, and Ford and Fulkerson, is graph-theoretic and simple. The maximum flow is the capacity of the smallest cut-set separating the two points.

The calculus of binary relations is another natural application of the theory of nets. Here we are concerned with a set of objects (human beings, for instance) with some relations defined between pairs of objects. Typical relations are "son of," "father of," "friend of," etc. Now, each object is a vertex of the net and the relations are shown by directed edges weighted with the relation. For example, the familiar riddle, "Brothers and sisters I have none, but that man's father is my father's son," has the representation shown in Fig. 1–7. Here s_1 is the speaker, f is his father, m is "that man," and s_2 is the father of m. F and S stand for "Father of" and "Son of." The solution to the riddle may be read from the net. (There are many other mathematical games, unconnected with the calculus of relations, which depend on graph theory for their solutions.) Social sciences are very often concerned with such relations. The social structure of a group of individuals can be represented as a net, and much about the group can be learned from the net. For instance, a "clique" is a maximal complete subgraph of the net. Yet another net that is familiar to many electrical engineers is the representation of a set of equations as a *signal-flow graph.* Still another application is to neural networks.

It is clear that very similar (and often identical) methods of attack will be found useful in these various applications of the theory of nets.

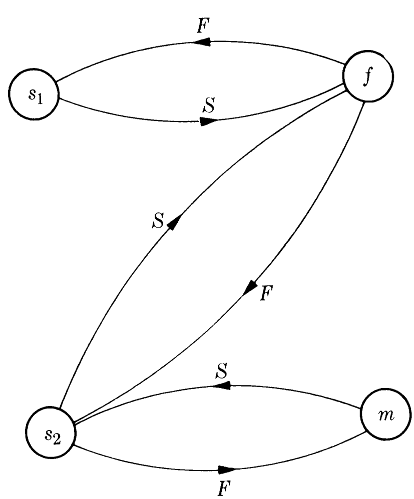


FIG. 1–7. Graph of riddle.

Matrices of the net have been found to be the most useful tools. These are considered in Chapters 9 and 10.

Enumeration problems in general find graph theory a useful tool. Enumeration problems arise in such diverse fields as chemistry, psychology, and (classical mathematical) combinatorial analysis. However, this is a topic that is not considered in this book. Interested readers are referred to excellent treatments by Riordan [148] and Harary and Norman [73].*

In fact, it is because graph theory has so many diverse applications that we should study abstract graph theory as a subject by itself and not mix it up with some specific application such as electrical network theory. Despite this separation of theory from application, the study of graphs in this book is oriented toward applications in electrical engineering.

We have seen examples of applications of both the directed graph (with orientations assigned to the edges) and the nonoriented graph. For the application considered here in the greatest detail, namely electrical network theory, we need both directed and nonoriented graphs. In network analysis, directed graphs are used. However, the orientation is rather artificial (introduced to take care of the reference systems for current, voltage, and magnetic polarity) and so disappears when system functions are computed. For these reasons, we choose to begin graph theory with nonoriented graphs even though the algebra involved (modulo 2 algebra) is unfamiliar. This fact is in itself an advantage because it prevents potential confusion with familiar concepts in electrical network theory and focuses attention on graph theory instead.

1–2 The nonoriented graph. It is unfortunate that every mathematical theory has to begin with a long list of definitions. Even more unfortunate is that nothing can be done about it. One must have a few words to talk with, and in the interest of precision these have to be formally defined. It is possible to reduce somewhat the number of definitions (as for instance in defining a *path*), but then each definition becomes much more complicated and hence nearly incomprehensible. The intuitive concepts of a *path* and a *loop* turn out to be relatively hard to define, mainly because all point-set topological concepts are avoided in order to develop a theory that is independent of the relative locations of the elements. This independence is, however, an essential part of both the theory and its application to electrical networks. The one saving feature in graph theory is that many of the terms used have nearly the same meaning as in everyday English and so very little conscious effort is required to remember

---

* Numbers in brackets refer to the bibliography at the end of the book.

them. We will also refer to several illustrative diagrams to act as a buffer in this initial barrage of definitions.

DEFINITION 1–1. *Edge.* A line segment together with its distinct endpoints is an *edge.*

In this book, *edge* and *element* are used as synonyms. While *element* is the more common engineering term, it can sometimes be confusing when one has to talk of elements of matrices or some other sets as well. *Edge* is more convenient to use then.

Definition 1–1 as stated places two requirements on the edge. First, the endpoints belong to the edge. Second, the endpoints are distinct. Clearly, both are merely conventions, and they are introduced here to simplify the statements of many theorems. On occasion, however, these conventions cease to be convenient. For instance, in certain interpretive statements in Section 2–4, it is more convenient to regard the endpoints as not belonging to the edge. If absolute precision is required, the edge without its endpoints must then be defined to be a new entity: *arc,* for instance. However, in this text, the name *edge* is used in this connotation also, with a reminder to the reader that it does not include the endpoints. Similarly, in the discussion of duality in Chapter 3, and in the discussion of the applications of nets in Chapters 8 and 10, edges with coincident endpoints (*self-loops*) are needed. Again, to avoid additional terminology, such edges are simply admitted where necessary. The definition of an edge, as given, is used unless the discussion explicitly states otherwise.

DEFINITION 1–2. *Vertex.* A *vertex* is an endpoint of an edge.

Point, 0-*cell,* and *node* are three other names commonly used for a vertex. By convention, and to simplify the statements of theorems, we do not usually consider an isolated point as a vertex. On some occasions it is convenient to regard an isolated point as a vertex (which we do, after warning the reader).

DEFINITION 1–3. *Linear graph.* A *linear graph* is a collection of edges, no two of which have a point in common that is not a vertex.

Linear complex and 1-*complex* are other words used for a graph. Some examples of linear graphs are shown in Fig. 1–8. A graph as defined here is an abstract graph and need not have any geometric significance whatsoever. It is true, however, that one can consider a linear graph as a configuration in a 3-dimensional euclidean space. The vertices can be interpreted as points and the edges as arcs. Without further specific statement, only *finite graphs* are considered here, that is, graphs containing only a finite number of edges (and hence a finite number of vertices). Infinite graphs have some interesting properties, but do not (so far) have

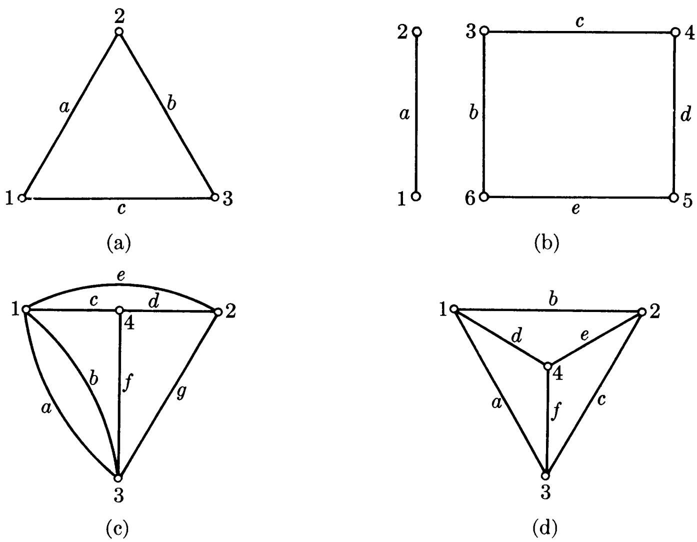


FIG. 1–8. Examples of linear graphs.

any applications. The interested reader may refer to Koenig [88] for a discussion of infinite graphs.

Graph theory derives its potency and its rich variety of applications from the single axiom of graph theory which follows.

AXIOM OF GRAPH THEORY. If M is an arbitrary finite or infinite collection of objects, and if to each (unordered) pair (a, b) of M is assigned a nonnegative integer $M_{ab} = M_{ba}$ (which may be zero), such that for each a at least one M_{ab} is nonzero, then there exists a graph G which has the elements of M for vertices and in which vertices a and b are connected by M_{ab} edges.

Despite its simplicity and intuitive "validity" from a geometric viewpoint, the axiom of graph theory is very important. In fact, it is this simplicity (and generality) that makes graph theory applicable to a very large number of situations. As an example, consider a problem that might arise in a civil defense communication network.

Suppose that there is a network of five stations and we wish to find out whether there are any "weak spots" in the network that need to be "strengthened" by additional channels of communication. That is, we wish to know whether any station or set of stations can become isolated from the rest of the group by the failure of a very small number of channels of communication. Let the five stations be a, b, c, d, and e. These are the *elements of M* for the application of the axiom. For each pair of stations, we list the number of channels through which they can communicate. The channels themselves may take on any physical form. Some of them may be radio links, others telephone lines, or even semaphore links or roads along which messengers can go. Suppose that this table takes the form of Table 1–1.

TABLE 1–1

Pairs of stations	Number of links
a, b	1
a, c	0
a, d	2
a, e	0
b, c	1
b, d	3
b, e	0
c, d	0
c, e	1
d, e	1

Let us draw the graph for such a system. The graph has five vertices corresponding to the five elements of M. The number of edges between vertices i and j is the number M_{ij}, to be found in the second column of Table 1–1. For example, there are three edges between b and d and none between b and e. The graph is shown in Fig. 1–9. The broken lines on the figure show how a subset of stations can become isolated from the rest.

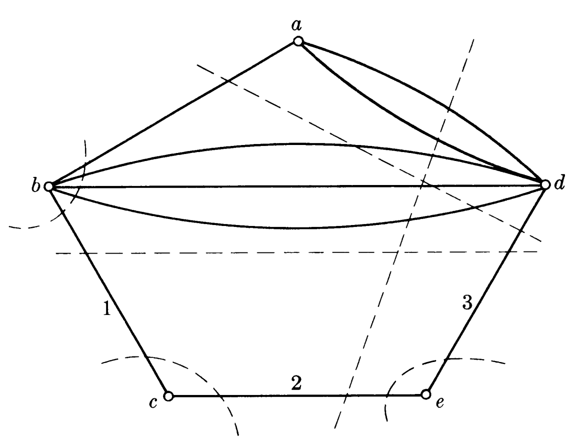


FIG. 1–9. Graph of communication network.

If all the channels of communication that correspond to the edges cut by a broken line fail, the sets of stations on the two sides of the broken line are isolated from each other. (These are called *cut-sets*, to be defined later.) We see from the graph that the failure of a single channel cannot isolate a station or set of stations. But if any two of the three channels marked 1, 2, and 3 in the figure fail, either or both of stations c and e would be isolated from the rest. To increase the strength of the system, at least two more channels are needed. By inspection of the graph it is seen that a suitable way to add two channels is to add one between a and c and another between a and e. The graph of the new system is shown in Fig. 1–10. At least three channels have to fail in the new system before any station becomes isolated. The probability of three simultaneous failures would naturally be less than the probability of two simultaneous failures.

DEFINITION 1–4. *Subgraph.* A *subgraph* is a subset of the edges of the graph (and thus is itself a graph). The subgraph is a *proper subgraph* if it does not contain all the edges of the graph.

DEFINITION 1–5. *Incidence.* A vertex and an edge are *incident* with each other if the vertex is an endpoint of the edge.

For example, in Fig. 1–8(a), edge a is incident with vertex 1 and vice versa. On the other hand, edge a is not incident with vertex 3.

In Section 1–1, it was shown that it is possible to draw an electrical network or a linear graph in different ways. In such cases, one would like a precise way of saying that the two graphs are really the same even though they are drawn differently and the vertices and edges are labeled

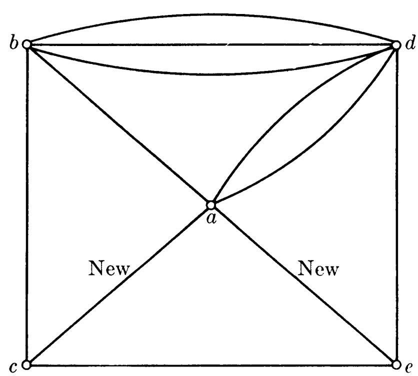


FIG. 1–10. Graph of new network.

differently. The next definition provides the terminology necessary for this purpose.

DEFINITION 1–6. *Isomorphism.* Two graphs G and G' are *isomorphic* (or *congruent*) if there is a one-to-one correspondence between the vertices of G and G' and a one-to-one correspondence between the edges of G and G' which preserves the incidence relationships.

For example, the graph of Fig. 1–8(d) is isomorphic to the graph of Fig. 1–3(b). We establish their isomorphism by means of Table 1–2.

TABLE 1–2

(a) Vertex		(b) Edge	
Fig. 1–3(b)	Fig. 1–8(d)	Fig. 1–3(b)	Fig. 1–8(d)
1	1	1	d
2	2	2	b
3	4	3	c
4	3	4	f
		5	a
		6	e

It can be readily verified that corresponding edges are incident at corresponding vertices. For those unfamiliar with the concept of an isomorphism, it should be noted that a one-to-one correspondence is not, by itself, an isomorphism. The correspondence must be *preserved* (or be *invariant*) under whatever relation happens to be of interest, in this case incidence. If, for example, edges a and e are interchanged in the above table, making a correspond to 6 and e to 5 (leaving all other entries unaltered), it is no longer an isomorphism. For, a has vertices 1 and 3 in Fig. 1–8(d). From the vertex table, the corresponding vertices in Fig. 1–3(b) are 1 and 4. But edge 6 has vertices 2 and 4 in Fig. 1–3(b), not 1 and 4.

The next sequence of definitions is directed toward the realization of a reasonable definition of a path and a loop.

DEFINITION 1–7. *Edge sequence.* If the edges of a graph or a subgraph *can be* ordered such that each edge has a vertex in common with the preceding edge (in the ordered sequence) and the other vertex in common with the succeeding edge, the subgraph is an *edge sequence.*

In this definition, note that an edge sequence is a graph. When it is expressed as a sequence of edges, each edge may appear any number of times. In fact, an edge may follow itself in the sequence. If we trace the sequence on the graph, the resulting line may intersect itself or retrace parts several times. For example, in Fig. 1–8(c),

$$abcffdecdg$$

is the ordering of an edge sequence, the edge sequence in this case being the whole graph.

DEFINITION 1–8. *Multiplicity.* The number of times an edge appears in an edge sequence is the *multiplicity* of the edge.

In the example of the edge sequence given above, the edges c, f, and d have multiplicity 2 and all others have multiplicity 1.

DEFINITION 1–9. *Edge train.* If each edge of an edge sequence has multiplicity 1, the sequence is an *edge train.*

An example of an edge train in Fig. 1–8(c) is *abcdgf.* Thus an edge train can intersect itself (that is, go through a vertex more than once), but cannot retrace parts of itself, as an edge sequence can.

DEFINITION 1–10. *Initial, final, and terminal vertices.* The vertex of the first edge of an edge sequence (or an edge train) that is not shared by the second edge is the *initial vertex.* Similarly, the vertex of the last edge that is not common to the previous edge is the *final vertex.* The initial and final vertices are the *terminal vertices* of the sequence. (*Initial* and *final* refer to the ordering of the edge sequence. The graph is nonoriented.)

It is implicitly assumed in this definition that the first and second edges are not the same and similarly that the final edge is different from the preceding edge. We also say that the edge sequence is *between* the initial and final vertices and that the terminal vertices are *connected* by the edge sequence.

DEFINITION 1–11. *Closed and open edge trains.* If the terminal vertices of an edge train coincide, it is a *closed* edge train, otherwise it is *open.*

DEFINITION 1–12. *Degree of vertex.* The *degree* of a vertex is the number of edges incident at the vertex.

DEFINITION 1–13. *Path.* If the degree of each internal (nonterminal) vertex of an edge train is 2 and the degree of each terminal vertex is 1, the edge train is a *path.* (This degree is to be counted with respect to the edge train only and not with respect to the graph in which the edge train may be situated.)

Some examples of paths in Fig. 1–8(c) are *cd*, *agd*, and *gec*.

DEFINITION 1–14. *Circuit* or *loop*. If an edge train is closed and all vertices are of degree 2, the edge train is a *circuit* or a *loop*.

This may seem like a lot of fuss just to introduce the intuitively obvious concepts of a path and a loop. The difficulty, as already indicated, is that all point-set-theoretic ideas are avoided so that we may develop a theory that is independent of coordinate systems (i.e., independent of the relative locations of vertices and edges in a 3-dimensional space) and dependent only on incidence relationships.

Also note that the word *circuit* is being defined as synonymous with the word *loop*, or closed path. In the language of electrical engineers of yesterday, *circuit* and *network* were considered synonymous. We are here restoring the original meaning of *circuit* (German *Kreis*), which was *circle*. Modern electrical engineering terminology tends to designate an electrical network as a *network* and not as a *circuit*. This may cause an initial confusion to those used to the older terminology, but it need only be an initial confusion. A more important point of conceptual importance is that we are defining a *loop* to be a *subgraph*, i.e., a collection of edges *rather than the operation of "going around" a closed path*. One immediate consequence of this definition is that the number of possible loops in a finite graph is finite—for which fact we can be grateful when confronted with Kirchhoff's voltage law.

A much more elegant way of defining a circuit proves very useful, since it gives intuitive insight into the structure of circuits. Before we can give this definition, we need the concept of connectedness.

DEFINITION 1–15. *Connected graph*. A graph G is *connected* if there exists a path between any two vertices of the graph.

Thus, intuitively speaking, a graph is connected if it is in one piece. Fig. 1–11(a) is an example of a graph which is *not* connected, and Fig. 1–11(b) shows one which is connected.

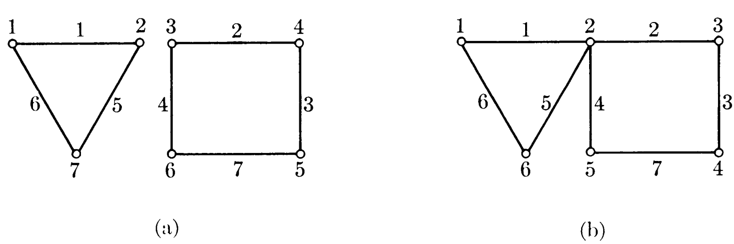


FIG. 1–11. Unconnected and connected graphs.

The alternative definition of a circuit, due to Veblen [190], is: a *circuit* is a connected graph or subgraph in which each vertex is of degree 2.

THEOREM 1–1. Veblen's definition of a circuit is equivalent to Definition 1–14. (In other words a subgraph G_s is a circuit according to Definition 1–14 if and only if G_s is connected and each vertex of G_s is of degree 2.) (See Problem 1–2.)

DEFINITION 1–16. *Noncircuit and circuit elements.* An element of a graph G which is not contained in any circuit of G is a *noncircuit element.* All other edges are *circuit elements.*

THEOREM 1–2. If G is a connected graph and one of the circuit elements of G is removed, the resultant graph is connected and contains all the vertices of G.

Proof. Let e_1 be a circuit element, and let G_s be the subgraph obtained when e_1 is removed. Since there is a circuit in G containing e_1, the vertices of e_1 are common to other edges of G (Veblen's definition of a circuit). Hence G_s contains all the vertices of G. Only the paths in G which contain e_1 are absent in G_s. Since there is a circuit in G containing e_1, there is a path P_2 in G_s between the vertices of e_1 (which therefore does not contain e_1). If in any path P_1 of G containing e_1, e_1 is replaced by the path P_2, an edge sequence is obtained, which contains a path (Problem 1–4). Hence the theorem.

This is a very useful result.

If the graph G happens to be an unconnected graph, as in Fig. 1–11(a), then it is obvious that it must consist of a number of "connected pieces." We next attempt to make this intuitive concept precise.

By Problem 1–12, the existence of a path between vertices is an equivalence relation. Any such equivalence relation defines a partition of the vertices of the graph into sets such that any two vertices in a set are connected by a path in G. Alternatively, we could also construct the sets. Beginning with any vertex v_1, consider all the vertices of G which can be connected to v_1 by a path in G. Then the elements of G incident at these vertices constitute a connected subgraph G_s. Furthermore, if any other element of G is added to this subgraph to form G_1, then G_1 is not connected. Thus G_s is a *maximal connected subgraph* of G. G may or may not have any more vertices than are contained in G_s. If G has other vertices (not in G_s), consider one of these vertices v_i. By a similar process, we can now construct a maximal connected subgraph containing v_i. The process can be repeated until there are no more vertices left, provided G is finite. The number of these maximal connected subgraphs is denoted by p.

THEOREM 1–3. The decomposition of a graph into maximal connected subgraphs is unique.

THEOREM 1–4. $p = 1$ for a graph G if and only if G is connected.

PROBLEMS

1–1. Show that the terminal vertices of an open edge train are of odd degree and that the other (internal) vertices are of even degree, where the degree is to be counted with respect to the edge train only.

1–2. Prove Theorem 1–1. [*Hint:* Problem 1–1.]

1–3. If there is an open edge train between vertices a and b, show that there is a path between vertices a and b.

1–4. If there is an edge sequence with terminal vertices a and b, show that there is a path between a and b.

1–5. Prove that a single edge is a path.

1–6. Let a path P between two vertices a and b be considered as a subgraph G_s. Prove:

(a) There exists one and only one path in G_s between vertex a and any other vertex of G_s.

(b) There exists one and only one path in G_s between any two vertices of G_s.

(c) The number of edges e and the number of vertices v of G_s are related by $e = v + 1$.

(d) Any connected subgraph of G_s is a path.

1–7. If there are two different paths P_1 and P_2 (differing in at least one edge) between two vertices a and b, show that there is a circuit consisting of some of the edges of P_1 and P_2.

1–8. If a closed edge train contains vertex a, show that there is a circuit containing vertex a.

1–9. Prove that every graph contains at least one connected subgraph.

1–10. Prove that if G_1 and G_2 are two subgraphs of a connected graph G such that G_1 and G_2 have no edges in common and together include all edges of G, then G_1 and G_2 have at least one common vertex.

1–11. Let a circuit be considered as a subgraph G_c. Prove:

(a) If G_c contains e edges and v vertices, then $e = v$.

(b) There are exactly two paths between any two vertices of G_c.

(c) G_c contains at least two edges.

(d) Any proper subgraph of G_c contains at least two vertices of degree 1.

(e) G_c contains a path.

(f) The complement of any path in G_c is also a path.

1–12. If there is a path between vertices a and b and there is a path between vertices b and c, show that there is a path between vertices a and c. Thus if we write $a \; P \; b$ to say that there is a path between a and b, show that the relation P satisfies

(a) $a \; P \; a$ by definition (reflexivity),

(b) $a \; P \; b$ implies $b \; P \; a$ (symmetry),

(c) $a \; P \; b$ and $b \; P \; c$ imply $a \; P \; c$ (transitivity).

A relation between two objects (a *binary* relation) that is reflexive, symmetric, and transitive is known as an *equivalence* relation.

1–13. Show that an equivalence relation defined on a finite set S of objects defines a partition of S into disjoint (mutually exclusive) subsets $s_1, s_2, \ldots, s_k$ which includes all elements of S.

1–14. Show that the partitioning defined by an equivalence relation (Problem 1–13) is unique.

1–15. Prove Theorem 1–3.

CHAPTER 2

CIRCUITS AND CUT-SETS

2-1 The Königsberger bridge problem. Euler [51] wrote perhaps the first paper on graph theory in 1736. Euler's interest was due to a problem which arose in Königsberg, Germany. This celebrated problem, called the Königsberger bridge problem, may be stated as follows.

The shaded areas of Fig. 2-1 denote a river, and the regions A, B, C, and D denote land. There are seven bridges across the river. The problem was to cross all seven bridges, passing over each one only once. Euler solved the problem by showing that it was impossible, and laid the foundations of graph theory. Euler's formulation is in terms of islands and bridges.

If we draw a graph for the bridge problem, with a vertex for each region and an edge for each bridge, we get the graph of Fig. 2-2. The problem now is to draw this graph as an open or closed edge train. This problem is solved after the discussion of Euler graphs. (Here, we are certainly not interested in the solution to this problem but in the fundamental ideas that have arisen from it.)

Since graphs are considered as sets of edges, a few set-theoretic terms (which are almost self-explanatory) are used in the following discussion, without formal definitions. The *union* of two subgraphs G_1 and G_2 is the set of all edges which are in G_1 or in G_2 or in both. The *intersection* of two subgraphs G_1 and G_2 is the set of all edges which are (simultaneously) in both G_1 and G_2. Two subgraphs G_1 and G_2 are *edge-disjoint* if they have no common edges. G_1 and G_2 are *vertex-disjoint* or simply

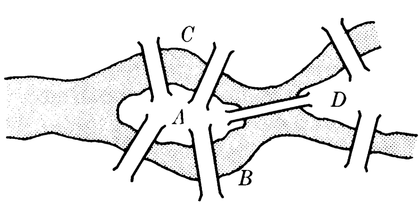


FIG. 2-1. Königsberger bridge problem.

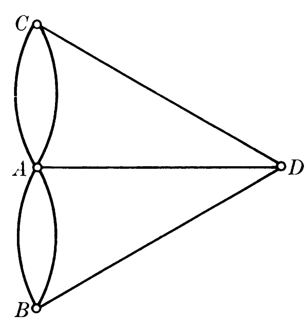


FIG. 2-2. Graph of the Königsberger bridge problem.

disjoint if they have no common vertices (and hence no common edges either). Usually the union of G_1 and G_2 is denoted by $G_1 \cup G_2$ and the intersection as $G_1 \cap G_2$. Thus if G_1 and G_2 are disjoint, $G_1 \cap G_2$ is the *null graph.*

DEFINITION 2–1. *Euler line.* If a closed edge train of the graph contains all the edges of the graph, then it is an *Euler line* of the graph.

DEFINITION 2–2. *Euler graph.* A graph in which every vertex is of even degree is an *Euler graph.*

We next relate these two seemingly different concepts, the Euler line and the Euler graph.

THEOREM 2–1 (Veblen). A graph G is an Euler graph if and only if G is a union of circuits, no two of which have an element in common.

Proof. If G is an element-disjoint union of circuits, then the degree of each vertex is even, and the graph is an Euler graph by Definition 2–2. Let G be an Euler graph. Let us begin at any vertex v_1. There are at least two elements at v_1. Let (v_1v_2) be one element. Since v_2 is of even degree, there is another element (v_2v_3). Now either $v_3 = v_1$ or there is an element (v_3v_4). Proceeding in this fashion, we must incorporate at some stage a vertex that has already been included. Then we would have formed a circuit (with some additional elements, possibly). Let this circuit be removed. The complement (remainder) in G is still an Euler graph. Thus we can repeat the procedure until no elements are left. Hence the theorem.

COROLLARY 2–1. Every vertex of an Euler graph is contained in a circuit.

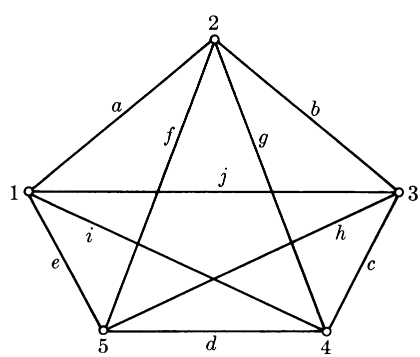


FIG. 2–3. The complete pentagon.

An example of an Euler graph is shown in Fig. 2–3. Each vertex of this graph is of degree 4. This graph may be described in a number of ways as a disjoint union of circuits. For example, it consists of the two circuits *abcde* and *fgijh*. Or it is made up of the three circuits *aef*, *cdh*, and *bjig*.

THEOREM 2–2. *A graph can be drawn as a closed edge train if and only if it is a connected Euler graph.*

Proof. If the graph can be drawn as a closed edge train, then every vertex is of even degree. The graph is therefore connected and is an Euler graph. Let G be a connected Euler graph. Let Z be a closed edge sequence of G which contains the maximum number of elements of G. If $(G - Z) = G'$ is not empty, then G' is an Euler graph and has a vertex in common with Z. Let this vertex be v_i. By Corollary 2–1, there is a circuit C in G' containing v_i. Now $Z + C$ is a closed edge train which can be described by starting at v_i, describing Z, and then describing C. This contradicts the assumption that Z was a maximal closed edge train and $(G - Z)$ is nonempty. Thus the theorem is proved.

The complete pentagon of Fig. 2–3 can be drawn as a closed edge train as follows. Starting with vertex 1, we describe the circuit *abcde* and then describe the circuit *jhfgi*.

The next theorem solves the Königsberger bridge problem. It is a special case of a theorem first stated by Listing and later proved by Lucas. We need the following lemma, which is interesting in itself and has a very neat proof.

LEMMA 2–1. *In any finite graph G, there is an even number of vertices of odd degree.*

Proof. Let ρ_n be the number of vertices of degree n, and let G contain e elements. Since each element has two vertices, we have

$$2e = \rho_1 + 2\rho_2 + 3\rho_3 + \cdots + p\rho_p. \tag{2–1}$$

Since $2e$ is even, so is

$$2e - 2\rho_2 - 2\rho_3 - 4\rho_4 - 4\rho_5 - 6\rho_6 - 6\rho_7 - \cdots = \rho_1 + \rho_3 + \rho_5 + \rho_7 + \cdots \tag{2–2}$$

Hence the lemma.

THEOREM 2–3. *A graph is an open edge train if and only if it is connected and contains exactly two vertices of odd degree.*

Proof. If the graph is an open edge train, then the two terminal vertices are of odd degree, and the internal vertices are of even degree. The graph is obviously connected. If there are two vertices of odd degree, let these be Q_1 and Q_2. Addition of an edge (Q_1Q_2) makes the graph into a connected Euler graph, which is a closed edge train by Theorem 2–1. Removal of (Q_1Q_2) makes the edge train open.

The graph of the Königsberger bridge problem contains four vertices of odd degree and thus is not an open edge train.

2–2 Circuits. Circuits are fundamental in electrical network theory. This section is devoted to an examination of the concept of a circuit.

Whitney [199], in a fundamental paper, defines circuits for a rather general class of objects called "matroids," by means of the following three postulates.

C_1. No proper subset of a circuit is a circuit.

C_2. If P_1 and P_2 are circuits, if e_1 is in both P_1 and P_2, and if e_2 is in P_1 but not in P_2, then there is a circuit in $P_1 + P_2$ containing e_2 but not e_1.

C_3. If P_1 and P_2 contain only one common element e, then $P_1 + P_2 - e$ is a union of a set of circuits.

In these postulates, $+$ stands for set-theoretic union. Thus $P_1 + P_2$ consists of all the elements which are in P_1 or in P_2 or in both. $P_1 + P_2 - e$ is the union of P_1 and P_2 with the element e removed.

Before examining Whitney's postulates, it is useful to introduce the concept of a ring and a ring sum of sets.

Definition 2–3. *Ring.* A *ring* is a collection S of objects with two (binary) operations, addition and multiplication defined, which satisfy the conditions:

(a) If α and β are in S, so is $\alpha + \beta = \beta + \alpha$.
(b) There exists a 0 in S such that $0 + \alpha = \alpha$ for all α in S.
(c) For each α in S, there is an α' in S such that $\alpha + \alpha' = 0$.
(d) $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$; α, β, and γ in S.
(e) If α and β are in S, so is $(\alpha\beta)$.
(f) If α, β, and γ are in S, $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.
(g) If α, β, and γ are in S, then $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ and $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$.

The first four conditions define an abelian group, abelian since the commutative law is obeyed.

A simple example of an abelian group under addition is the set of integers (positive, negative, and zero). This set is also a ring, as can be verified easily. Another example of a ring is the set of real (or complex) matrices of order (n, n). Real matrices of order (m, n), where $m \neq n$, constitute an abelian group under addition but do not constitute a ring since (e), (f), and (g) are not satisfied. The product of two such matrices is not defined. A more involved but more familiar example of an abelian group that is not a ring is the set of 3-dimensional vectors. Here, if the product is taken as the cross product, multiplication is defined and yields a 3-dimensional vector. It is also distributive over addition. But multiplication is not associative. That is,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C}. \tag{2–3}$$

(It is not commutative, either.)

DEFINITION 2–4. *Ring sum (mod 2 sum).* The *ring sum* $S_1 \oplus S_2$ of two sets S_1 and S_2 is the set of all elements of S_1 and S_2 which occur in S_1 or S_2 but not in both.

Thus $S_1 \oplus S_2$ is the difference between the logical sum and the logical product (for those familiar with this terminology):

$$S_1 \oplus S_2 = S_1 \cup S_2 - S_1 \cap S_2. \tag{2–4}$$

For example, the ring sum of the two sets $S_1 = \{a, b, c\}$ and $S_2 = \{b, c, d\}$ is

$$S_1 \oplus S_2 = \{a, d\}.$$

The name arises because the ring sum converts the algebra of sets (Boolean algebra) into a ring.

With the aid of these definitions, we now establish some fundamental properties of the set of circuits of a graph.

THEOREM 2–4. The ring sum of two circuits is a circuit or an edge-disjoint union of circuits (i.e., a set of circuits which contain no common edges).

The proof of this theorem is almost self-evident in the light of Theorem 2–1. For, if we consider any vertex of the ring sum of two circuits which is in both circuits, the total degree of the vertex is either 2 or 4. If one of the edges incident at this vertex is common to the two circuits, then the degree is 2, otherwise 4. All other vertices are of degree 2. In any case, the degree of each vertex of the ring sum is even. Thus the ring sum is an Euler graph, and the rest follows.

THEOREM 2–5. The set consisting of the circuits and disjoint unions of circuits of G is an abelian group under the operation $\oplus$.

If we note here the familiar convention in mathematics, that the null set is a subset of every set, we may observe that the set of circuits and disjoint unions of circuits satisfy the first four conditions of Definition 2–3. We leave it as a problem for the reader to complete the details (Problem 2–2).

THEOREM 2–6. The circuits of a graph satisfy postulates C_1, C_2, and C_3 of Whitney if we interpret "+" as union.

The proof of Theorem 2–6 is left as an instructive exercise (see Problem 2–3).

2–3 Trees and fundamental systems of circuits. The "tree" is perhaps the single most important concept in graph theory, insofar as electrical network theory is concerned. The word *tree* intuitively signifies a treelike structure, namely a structure in one piece, with branches, and branches off other branches. There are no closed paths (circuits) of branches. The term *tree* signifies a very similar concept in graph theory.

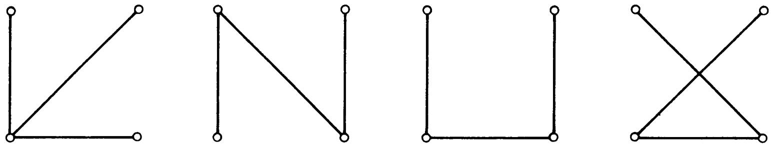
DEFINITION 2–5. *Tree.* A *tree* is a connected subgraph of a connected graph which contains all the vertices of the graph but does not contain any circuits.

This definition differs slightly from the conventional mathematical definition, in that the condition "contains all the vertices of the graph" is usually omitted. The alternative is to use the term *complete tree* for the concept we need, as in Cauer [25]. Since there is no occasion here to use an incomplete tree, we define a tree to be complete. Modern engineering terminology is in accordance with Definition 2–5. It is time to give a few examples. The graph of Fig. 1–3(b) contains sixteen trees, four similar to each of the trees shown in Fig. 2–4. As another example, the graph of Fig. 2–5(a) contains eight trees, four similar to those in Fig. 2–5(b) and two similar to each of Figs. 2–5(c) and 2–5(d). As a third example, we give two trees of a more complicated graph. The graph is shown in Fig. 2–6(a) and the trees in 2–6(b) and 2–6(c).

The tree is a very important concept because of the number of properties of the graph that can be related to the tree. The number of independent Kirchhoff equations, the methods of choosing independent equations, the structure of the coefficient matrices, and the topological formulas for network functions, are all stated in terms of the single concept of a tree. It has been considered so important that it has found its way even into several undergraduate texts on network theory.

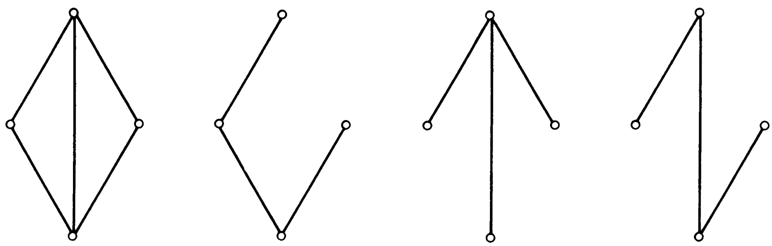
THEOREM 2–7. A finite graph is a tree if and only if there exists exactly one path between any two vertices of the graph.

Proof. If the graph is a tree, there is at least one path between any two vertices, since the tree is connected. If there are two paths P_1 and



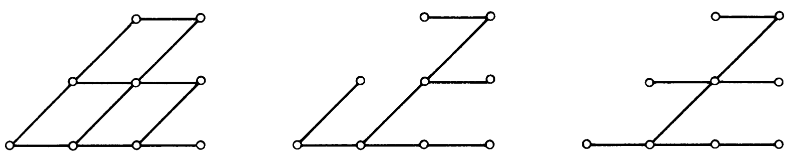
(a) (b) (c) (d)

FIG. 2-4. Trees of Fig. 1-3(b).



(a) (b) (c) (d)

FIG. 2-5. An example for trees.



(a) (b) (c)

FIG. 2-6. A graph and its trees.

P_2 between two vertices of a tree, then there is a circuit in $P_1 + P_2$. But a tree contains no circuits.

If the graph contains one and only one path between any two vertices, it is connected and contains no circuits. Hence it is a tree (of itself).

THEOREM 2-8. Every finite connected graph contains a tree.

Proof. If the graph itself is not a tree, it contains a circuit. Removal of an element of the circuit leaves the graph connected and does not remove a vertex, by Theorem 1-2. The circuit, however, is destroyed. Repeated application of this procedure yields a tree.

THEOREM 2-9. If a tree contains v vertices, it contains $v - 1$ elements.

Proof. The theorem is proved by induction on the number vertices. If $v = 2$, the tree can contain only one element, since it contains no circuits. Let the theorem be true for $v = n$. For a tree with $n + 1$ vertices, there is at least one end vertex a (that is, a vertex of degree 1) by Problem 2–6. Let (ab) be the element incident at a. If (ab) is removed from the tree, the result is a subgraph of n vertices, which is its own tree. By induction hypothesis, this subgraph contains $n - 1$ elements. Adding the element (ab), therefore, the tree of $n + 1$ vertices contains n elements.

A tree thus has four properties: connected, no circuits, v vertices, and $v - 1$ elements. It can be shown (Problem 2–7) that any three of these properties imply the fourth. This raises the question of whether any two will suffice. The answer is given next.

THEOREM 2–10. If G is a connected graph of v vertices, and G_s is a subgraph of G with $v - 1$ elements and containing no circuits, then G_s is a tree of G.

Proof. First we show that G_s is connected. For, let G_s consist of p maximal connected subgraphs. Let $s_1, s_2, \ldots, s_p$ be the subgraphs, and let v_i be the number of vertices in s_i. Since each s_i is connected and contains no circuits, s_i is its own tree. Hence s_i contains $v_i - 1$ elements. Since $s_1, s_2, \ldots, s_p$ contain no common elements or vertices, and together contain all vertices of G,

$$\sum_{i=1}^{p} v_i = v. \tag{2–5}$$

Hence the number of elements in G_s is equal to

$$\sum_{i=1}^{p} (v_i - 1) = \sum_{i=1}^{p} v_i - p = v - p = v - 1, \tag{2–6}$$

by hypothesis. Hence $p = 1$, or G_s is connected. Now G_s is its own tree and contains $v - 1$ elements. Hence G_s contains v vertices, or is a tree of G. Problem 2–8 disposes of the other pairs of conditions.

DEFINITION 2–6. *Branch.* An element of a tree is a *branch.*

DEFINITION 2–7. *Chord* or *link.* An element of the complement of a tree is a *chord* (*link*).

THEOREM 2–11. A connected graph of v vertices and e edges contains $v - 1$ branches and $e - v + 1$ chords.

When we speak of chords and branches, it is with reference to a chosen tree.

If we add one chord to a tree, the resulting graph is, of course, no longer a tree. The chord, and the path in the tree between the vertices of the chord, constitute a circuit. This, however, is a unique circuit and the only circuit of the resulting graph.

DEFINITION 2–8. *f-circuit. f-circuits* (fundamental circuits) of a connected graph G for a tree T are the $e - v + 1$ circuits formed by each chord and its unique tree path.

The concept of fundamental circuits is due to Kirchhoff [86] and is very useful. If the graph is not connected, it consists of maximal connected subgraphs. One can find a tree for each subgraph. The set of these trees is called a *forest* of G. It follows immediately that there are $v - p$ elements in a forest and $\mu = e - v + p$ elements not in the forest.

DEFINITION 2–9. *Nullity.* The *nullity* of a graph with e edges, v vertices, and p maximal connected subgraphs is $\mu = e - v + p$. Nullity is also known by the names of *cyclomatic number*, *connectivity*, and *first Betti number*.

DEFINITION 2–10. *Rank.* The *rank* of a graph with v vertices and p maximal connected subgraphs is $v - p$. (The reason for the name *rank* is seen in Chapter 4.)

The fundamental system of circuits for an unconnected graph is obtained by taking the fundamental systems for each maximal connected subgraph.

Frequently, it is of interest to know whether a subgraph can be made part of a tree. The following theorem is useful in this connection.

THEOREM 2–12. A subgraph G_s of a connected graph G can be made part of a tree if and only if G_s contains no circuits.

*Proof.** The necessity follows from the definition of a tree. Suppose that G_s contains no circuits. Let T be any tree of G (G connected by hypothesis). Consider $T + G_s = G_1$. G_1 contains all the vertices of G. G_1 may contain circuits. Any circuit C of G_1 contains at least one element not in G_s, since G_s contains no circuits. Removal of such an element destroys C without removing a vertex. Repeated application of this procedure yields the result of the theorem.

* The method of proof is due to Prof. P. W. Ketchum of the University of Illinois.

2–4 Cut-sets and fundamental systems of cut-sets. The preceding discussions have certainly indicated that a circuit is an important subgraph, and the discussions to follow add to the stature of circuits. A second class of subgraphs, the *cut-set*, closely parallels the circuit (or, to anticipate later discussion, is *dual* to it) and finds important use in electrical network theory. Because these two concepts are so closely related, circuits and cut-sets are introduced early and kept late in the presentation of the material in this text. Whitney [194] seems to have originated the concept of a cut-set during his fundamental work on the theory of graphs.

In discussing cut-sets, it is conceptually convenient to regard an edge as open, that is, as not including the endpoints, and admit isolated vertices into the graph. The definitions and theorems stated here are all formulated in such a way as to apply with or without this interpretation. In many of the explanatory statements, however, the endpoints are considered as not belonging to the edge. (We may note that rank and nullity of a graph are unaltered by the insertion or removal of isolated vertices.)

DEFINITION 2–11. *Cut-set.* A *cut-set* is a set of edges of a connected graph G such that the removal of these edges from G reduces the rank of G by one, provided that no proper subset of this set reduces the rank of G by one when it is removed from G.

Since the rank of a graph is $v - p$ (Definition 2–10), removal of a cut-set of edges without their vertices from a connected graph yields an unconnected graph (with one isolated vertex, possibly). The rank of the new graph is $v - 2$. Hence it follows that "cutting" this set of edges separates the graph into two pieces. The name *cut-set* has its origin in this interpretation. For examples of cut-sets, consider the graph of Fig. 2–7. The sets of edges *aige*, *chgfe*, and *bhige* are examples of cut-sets. The broken lines on Fig. 2–7 show how these cut-sets "cut" the graph. Drawing broken lines on the graph is a good way to find most of the

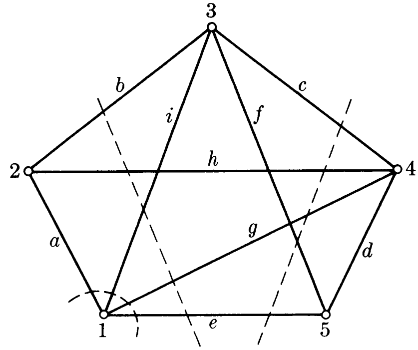


FIG. 2–7. Examples of cut-sets.

cut-sets; however, there may be other cut-sets which cannot be shown by a single straight line drawn across the graph, because of the ways in which the graph may be drawn. For instance in Fig. 2-8(a), the set of edges *abcd* is a cut-set, but we can draw a straight line through them only when the graph is redrawn as in Fig. 2-8(b).

Whitney's original definition of a cut-set was given in 1933. Although the concept has been used widely by a number of authors, including Guillemin [68] and Foster [59], very little work was done on the relationship of cut-sets to the other concepts of graph theory until quite recently. Our present discussion of cut-sets is based almost entirely on one of our own papers [154]. The orientation of the discussion here is toward showing that cut-sets bear the same relationship to circuits that circuits bear to incidence relationships. Thus we find the duals of a number of theorems proved earlier about circuits. It is also our purpose to show later (in Chapter 4) that the set of cut-sets contains the same essential information as the graph itself.

The cut-set (*aige*) of Fig. 2-7 is an interesting example of a cut-set, since these are all the edges incident at vertex 1 of the graph. It is evident, in general, that if we remove all the edges that are incident at a

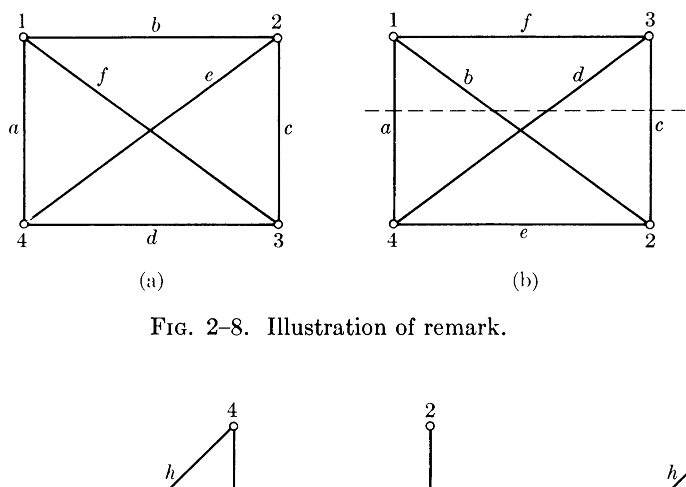


FIG. 2-8. Illustration of remark.

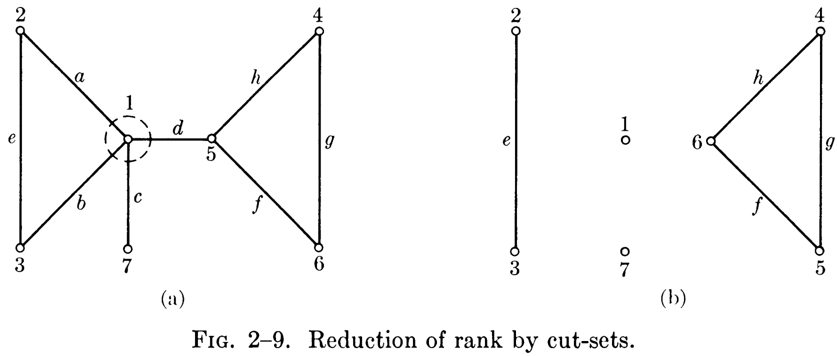


FIG. 2-9. Reduction of rank by cut-sets.

vertex, that vertex would be isolated. Thus the rank of the graph would be reduced by at least one. The "at least" in the previous sentence needs an explanation. Let us consider an example in which the removal of all the edges incident at some vertex reduces the rank of a graph by more than one.

In Fig. 2–9(a), if we remove all the edges incident at vertex 1, leading to Fig. 2–9(b), the rank of the graph is reduced from 6 to 3, a reduction of three instead of one as required. In fact, we see that the sets ab, c, and d are each cut-sets, so that the set of edges incident at vertex 1 is a disjoint union of cut-sets. Why this happened is discussed next. Vertex 1 is the only vertex common to the subgraphs $\{a, e, b, c\}$ and $\{d, h, f, g\}$. By Problem 1–10, any two subgraphs G_1 and G_2 of a connected graph G which are edge-disjoint and together include all edges of G must have a common vertex. In case two such subgraphs G_1 and G_2 have only one common vertex, that vertex is called a *cut-vertex* of the graph (also called *articulation point*). The formal definition of a cut-vertex is given in the next chapter along with the related concept of separability. We also note here that the edge c by itself is a cut-set and so is edge d. These two edges are not in any circuit of the graph. These few elementary properties are collected into the next theorem.

THEOREM 2–13. The set of edges incident at a vertex is a cut-set provided that this vertex is not a cut-vertex (articulation point) of the graph. Each noncircuit element is a cut-set (by itself). A circuit element (by itself) is not a cut-set.

THEOREM 2–14. Every cut-set C contains at least one branch of every tree T.

Proof. If we remove C, and T remains, there would be a path between any two vertices through T, so that C is not a cut-set.

A stronger version of this theorem, which is an elegant characterization of cut-sets, is possible:

THEOREM 2–15. C is a cut-set if and only if C is a minimal set of edges which contains at least one branch of every tree.

Proof. Let C be a minimal set containing at least one branch of every tree of the connected graph G. Then the complement G_c of C with respect to G does not contain any tree and so is either not connected or contains one less vertex than G. Hence the rank of G is reduced by one. Since C is a minimal such set, G becomes connected when any edge of C is returned to the graph. On the other hand, if C is a cut-set, the complement in G is not connected (counting isolated vertices), and so C contains at least one branch of every tree of G. If C is not a minimal such set, some proper subset of C becomes a cut-set by the first part of the proof, contradicting the definition of a cut-set.

Theorem 2–15 is the analog of Whitney's [199] characterization of a circuit as "a minimal set containing at least one chord of every tree" (Problem 2–22).

A cut-set can be interpreted in another useful fashion. Let G be a connected graph, and let C be a cut-set of G. Then the graph obtained by removing C is in two pieces (one of the pieces may be an isolated vertex). Let A and B be the sets of vertices in these two pieces. Then A and B are mutually exclusive and together include all the vertices of G. Further, any two vertices of A can be joined by a path not containing any vertex of B; and similarly for vertices in B. The edges of C have one vertex in A and another in B. No other edge has this property. Conversely, if the vertices of the graph G were partitioned into two sets A and B such that any two vertices in the same set can be connected by paths not containing a vertex of the other set, then the edges of G which have one vertex in A and the other in B constitute a cut-set.

This partitioning of vertices can be done by means of a tree. Let T be any tree of G, and let b_i be a branch of T. Since T contains no circuits, $\{b_i\}$ is a cut-set of T (where T is considered as a connected graph). Therefore this cut-set $\{b_i\}$ defines a partition of the vertices of T (which are all the vertices of G) into two sets A and B with the required property. Now let us consider the *cut-set of* G corresponding to this partition. This cut-set contains only one branch of T (and some chords with respect to T). Such a cut-set is called a *fundamental cut-set* for the following reason. By Theorem 2–14, each cut-set includes at least one branch of T, and the fundamental cut-set includes exactly one. In this respect, it is similar to a fundamental circuit. Each circuit includes at least one chord of T, and a fundamental circuit includes exactly one. Furthermore, fundamental cut-sets are very closely related to fundamental circuits, as we will see after the following formal definition.

Definition 2–12. *f-cut-set. The fundamental system of cut-sets* with respect to a tree T is the set of $v - 1$ cut-sets, one for each branch, in which each cut-set includes exactly one branch of T.

Theorem 2–16. If T is a tree of the connected graph G, the f-cut-set determined by branch b_i of T contains exactly those chords of G for which b_i is in each of the fundamental circuits determined by these chords.

We leave the proof of this important theorem as an instructive problem (Problem 2–23).

2–5 Cut-sets and circuits. In this section, we indicate the most important properties of cut-sets, namely, their relationship to the circuits of the graph, which form the basis of the discussion of the cut-set matrix in Chapter 4.

THEOREM 2–17. Every cut-set contains an even number of edges in common with every circuit.

Proof. Let α_i be a cut-set and c_j a circuit. Let Π_1 and Π_2 be the (necessarily disjoint) sets of vertices of the two subgraphs into which α_i separates the connected graph G. If α_i and c_j have no elements in common, the theorem is proved. If α_i and c_j have elements in common, then c_j contains vertices from both Π_1 and Π_2. Let the vertices of c_j be ordered cyclically so that any two successive vertices are endpoints of an element of c_j. Starting with a vertex in Π_1, we get to Π_2 by an edge of the cut-set. We can get back to Π_1 only by another edge of α_i. Since the circuit is a closed edge train, we have to get back to Π_1 finally. Thus the number of common elements is even.

The next theorem, which is the converse of the preceding, is the dual of Veblen's theorem on Euler graphs and characterizes the structure of cut-sets.

THEOREM 2–18. A nonempty set α of elements of a connected graph G, such that α has an even number of elements in common with every circuit, is a cut-set or an element-disjoint union of cut-sets.

Proof. Case 1. α has no elements in common with any circuit. Then every element of α is a noncircuit element, and so each element is a cut-set.

Case 2. Let the noncircuit elements be deleted from α, but retained in G. Let e_1 be an element of α (which now contains no noncircuit elements). Remove e_1 from α, and let α_1 be the remainder. Remove e_1 from G, and let G_1 be the remainder. G_1 is still connected and contains all the vertices of G. If α_1 contains a noncircuit element of G_1, let this element be e_2. Then $\{e_1, e_2\}$ is a cut-set of G. Otherwise, let another element e_2 be removed from α_1 and G_1, resulting in α_2 and G_2. This procedure of removing elements from α and G is repeated until the remainder of α contains a noncircuit element with respect to the remainder of G. This procedure cannot result in an empty set for the remainder of α, for suppose that all but one of the elements of α have been removed both from α and from G. Let this last element be e_n. If e_n is in any circuit of the remainder of G, α initially had at least two elements in common with this circuit and so the circuit has been destroyed when one of the other elements was removed from G, contradicting the hypothesis. Thus, at some stage, we are left with a noncircuit element whose removal reduces the rank of G by one. Thus α contains a cut-set of G. Let such a *cut-set* be removed from α but retained in G. (We can find this cut-set, if required, by returning the elements one by one to the graph until the rank is restored to $v - 1$.) Let α_1' be the rest of α. Now α_1' has the same characteristic as α; namely, it has an even number of elements in common with every circuit. This follows because α had an even number of elements in common with every circuit, by hypothesis, and the cut-set which was deleted from α also had an even number of elements in common with every circuit, by Theorem 2–17. Hence the same procedure may be repeated.

PROBLEMS

2–1. Extend Listing's theorem (2–3) to show that a graph with $2k$ vertices of odd degree can be drawn as k open edge trains, no two of which have a common edge (theorem of Lucas).

2–2. Prove Theorem 2–5.

2–3. Prove Theorem 2–6. [*Hint:* Theorem 2–4.]

2–4. Prove that if every vertex of a graph G is of degree 2, then G is a circuit or a vertex-disjoint union of circuits.

2–5. For a given graph, find a reasonable way of computing the number of trees of the graph. [*Hint:* Theorem 2–12.]

2–6. Show that every tree contains *at least* one vertex of degree 1 (an end vertex).

2–7. Show that any three of the following four conditions imply the fourth.

(a) G_s contains all v vertices of G.

(b) G_s contains $v - 1$ edges.

(c) G_s is connected.

(d) G_s contains no circuits.

2–8. Show by means of counterexamples that except for the pair (b) and (d), no other pair of conditions of Problem 2–7 implies the other two conditions.

2–9. If μ stands for the nullity of the graph $(e - v + p)$, show that for any graph, $\mu \geq 0$.

2–10. Show that a graph is a forest (a collection of trees) if and only if the nullity μ of the graph is zero.

2–11. Show that μ is invariant under insertion or removal of vertices of degree 2 (either by splitting an edge into two edges in series or by merging two edges in series into one).

2–12. Prove that any *end element* (an element with a vertex of degree 1) of a connected graph is contained in every tree of that graph.

2–13. Prove that any element of a connected graph is a branch of some tree.

2–14. Either prove or give a counterexample: any element of a connected graph is a chord for some tree.

2–15. Prove that a path is its own tree.

2–16. Prove that there are at most $e + 1$ vertices in a connected graph of e elements.

2–17. Prove that every connected graph contains a cut-set.

2–18. Prove that the complement of a tree does not contain a cut-set.

2–19. Prove that the complement of a cut-set does not contain a tree.

2–20. Prove the following dual of Theorem 2–12. A subgraph G_s of a connected graph G can be included in the complement of a tree if and only if G_s contains no cut-sets of G. [*Hint:* Consider $G - G_s$. Also Problems 2–17 and 2–18.]

2–21. Write out in detail the proof of Theorem 2–13.

2–22. Prove the analogue of Theorem 2–15; that is, prove that a subgraph G_s of a connected graph G is a circuit if and only if G_s is a minimal set of edges containing at least one chord of every tree of G.

2–23. Prove Theorem 2–16. [*Hint:* Method of proof of Theorem 2–17.]

2–24. Prove that the set of cut-sets and disjoint union of cut-sets is an abelian group under the ring sum ($\oplus$).

2–25. The set of all trees of a connected graph G is (b, c, e), (a, d, e), (a, c, d), (b, c, d), (a, b, c), (a, b, d), (b, d, e), (a, c, e). Find the fundamental system of cut-sets by using Theorem 2–15, and find a fundamental system of circuits by using Problem 2–22, both for the same tree. Verify Theorem 2–16 for this example.

CHAPTER 3

NONSEPARABLE, PLANAR, AND DUAL GRAPHS

In the theory of electrical networks without transformers and in the theory of combinational contact networks, we assume most of the time that the network is not separable. A network is separable if it consists of two subnetworks that are joined at only one node. In such a case, we know from experience that the two subnetworks can be treated as distinct subnetworks, independent of each other. The graph corresponding to a nonseparable network is a nonseparable graph. Another concept that is useful in network theory is that of duality. In this chapter, the graph-theoretic concepts of separability and duality are presented. This chapter is based almost entirely on two classical papers by H. Whitney, "Non-Separable and Planar Graphs" and "2-Isomorphic Graphs" [195, 198]. However, only such parts of these two papers as are of particular interest in network theory are introduced.

In Section 3–3, we introduce several inconsistencies in terminology to avoid the encumbrance of new words. The dual of a single edge (with distinct endpoints) happens to be a loop consisting of a single edge, i.e., an edge with coincident endpoints or a "self-loop." Such edges were excluded in Chapter 1 but are admitted in Section 3–3 to avoid complicating all the theorems about duality. (All cases of interest in the applications concern nonseparable graphs, and in such cases self-loops do not appear.) Also, in some of the proofs of this section it is convenient to consider the vertices as not belonging to the edge, and we admit isolated vertices. We adopt this procedure in preference to the introduction of new terminology. (The definitions and theorems still hold under the original definitions.) In the earlier sections of the chapter, however, isolated vertices and single-edge loops are not admitted.

3–1 Nonseparable graphs.

DEFINITION 3–1. *Nonseparable.* A graph G is *nonseparable* if every subgraph of G has at least two vertices in common with its complement. All other graphs are separable.

Thus a graph that is not connected is a trivial example of a separable graph. Several examples of separable graphs are shown in Fig. 3–1. Examples of nonseparable graphs are shown in Fig. 3–2. It follows from the definition that a connected separable graph G must contain at least

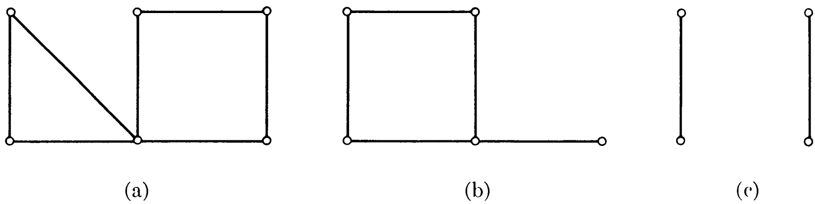


FIG. 3–1. Separable graphs.

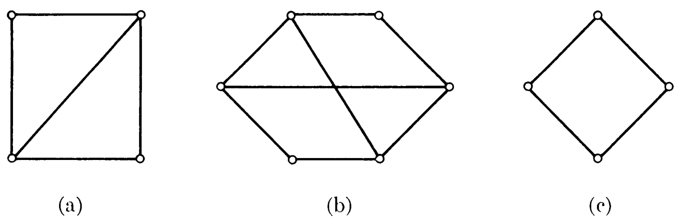


FIG. 3–2. Nonseparable graphs.

one subgraph which has only one vertex in common with its complement. In Section 2–4 such a vertex was named a cut-vertex. Formally, a cut-vertex is defined as follows.

DEFINITION 3–2. *Cut-vertex.* Let G be a connected separable graph, and let v_c be the single vertex in common between a subgraph G_s and its complement. Then v_c is a *cut-vertex* (*articulation point*) of G.

THEOREM 3–1. A necessary and sufficient condition that a connected graph be nonseparable is that it contain no cut-vertex. (This theorem is merely a restatement of the definition.)

THEOREM 3–2. A necessary and sufficient condition that v_c be a cut-vertex of a graph G is that there exist two vertices v_a and v_b (other than v_c) in G such that every path from v_a to v_b contains v_c.

Proof. Suppose that every path from v_a to v_b contains v_c. Let S be the set of vertices which can be connected to v_a by a path not containing v_c. Let v_a and v_c be added to S to make S_1. Then v_b is not in S_1. Consider the subgraph G_s consisting of the edges which have both vertices in S_1. This subgraph has only the vertex v_c in common with its complement. For if v_d is any other common vertex, and e_d is an edge of the complement incident at v_d, we see that the other vertex v_f of e_d can be connected to v_a without passing through v_c by a path from v_a to v_d together with e_d. Thus v_f belongs to S_1 and so e_d belongs to G_s, contrary to assumption.

Thus G is separable and v_c is a cut-vertex. The necessary part of the condition of the theorem is evident (from a sketch containing a cut-vertex).

A different way of looking at a nonseparable graph is given by Definition 3-3 and Theorem 3-3, which follow.

DEFINITION 3-3. *Cyclically connected.* A graph is *cyclically connected* if any two vertices in the graph can be placed in a circuit.

THEOREM 3-3. A necessary and sufficient condition that a graph containing at least two edges be cyclically connected is that it be nonseparable.

Proof. Without loss of generality, we may assume that the graph is connected, as the theorem is trivial otherwise. If G is separable, by Theorem 3-2 there exist two vertices v_a and v_b which cannot be placed in any circuit. Suppose that there are two vertices v_a and v_b which are not in any circuit. If there is an edge (v_av_b) in G, then there is no other path from v_a to v_b. Now we see that G is separable by a proof similar to the proof of Theorem 3-2. Otherwise, let $v_a, v_d, \ldots, v_b$ be the vertices of a path from v_a to v_b. If there is no circuit containing v_a and v_d, the first proof applies. Otherwise, let v_c be the last vertex of the path that can be placed in the same circuit as v_a. Let v_f be the next vertex of the path. Then every path from v_a to v_f passes through v_c. For suppose that there is a path from v_a to v_f not containing v_c. Let this path be p. Then we can construct a circuit containing v_a and v_f (thus contradicting the hypothesis that v_c is the last vertex of the path that can be placed in the same circuit as v_a) as follows. Let C be the circuit containing v_a and v_c. Starting with v_f, follow p until a vertex of C is reached. If v_a is not the first vertex reached, follow C to v_a. Continue along C to v_c and complete the circuit by the edge (v_cv_f), which will complete the proof.

THEOREM 3-4. A nonseparable graph containing at least two edges is of nullity $\mu > 0$. Each vertex is at least of degree 2.

The proof is omitted. (See Problems 2-9 and 2-10.)

THEOREM 3-5. A nonseparable graph of nullity 1 is a circuit and conversely.

The proof is left as a problem (Problem 3-3).

If the connected graph G is separable, we may separate the graph into two new connected graphs by splitting the cut-vertex in two. Insofar as electrical networks are concerned, this splitting is an operation that does not change the current or voltage of any network element. This process may be continued with another cut-vertex (if one exists) until every maximal connected subgraph is nonseparable. This process is known as *decomposition* of a separable graph into its *components* (German *Glieder*). We may now observe:

THEOREM 3–6. Every nonseparable subgraph of G is contained wholly in one of its components.

A more important theoretical result is:

THEOREM 3–7. The decomposition of a graph into its components is unique.

Proof. The theorem follows immediately from Theorem 3–6. For suppose that $G_1, G_2, \ldots, G_k$ and $G'_1, G'_2, \ldots, G'_n$ are two decompositions of G. Since G_i is a nonseparable subgraph of G, it is contained wholly in one of the components $G'_1, G'_2, \ldots, G'_n$, say G'_j. But G'_j is a nonseparable subgraph of G and so is contained in a G_k. Hence G_i is contained in G_k or they are identical. Hence also G_i and G'_j are identical. Repeated application of the argument yields the result.

In Chapter 1 we defined isomorphism for two graphs. The concept of decomposition of a separable graph gives a different type of equivalence. Consider a separable graph with a cut-vertex v_1. It may be decomposed by replacing v_1 by two vertices v'_1 and v''_1. Now, if we like, we can reconnect the two parts by coalescing some vertex v_i of one part with some vertex v_j of the other. If the graph represents an electrical network, the new network has the same currents and voltages as the old one. Two such graphs are shown in Fig. 3–3. A more general type of equivalence that is of interest in electrical network theory is the interchange of series-connected elements or subnetworks. This whole class of equivalences was investigated by Whitney [198] who named the equivalence a 2-isomorphism.

DEFINITION 3–4. *2-isomorphism.* Two graphs G_1 and G_2 are 2-*isomorphic* if they become isomorphic under (repeated applications of) either or both of the following operations:

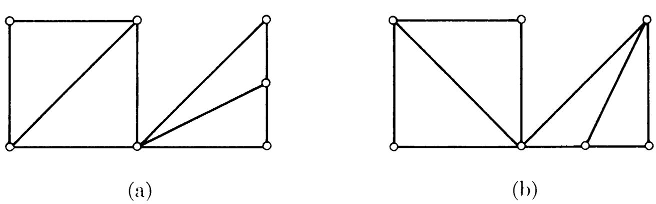


FIG. 3–3. Equivalent graphs.

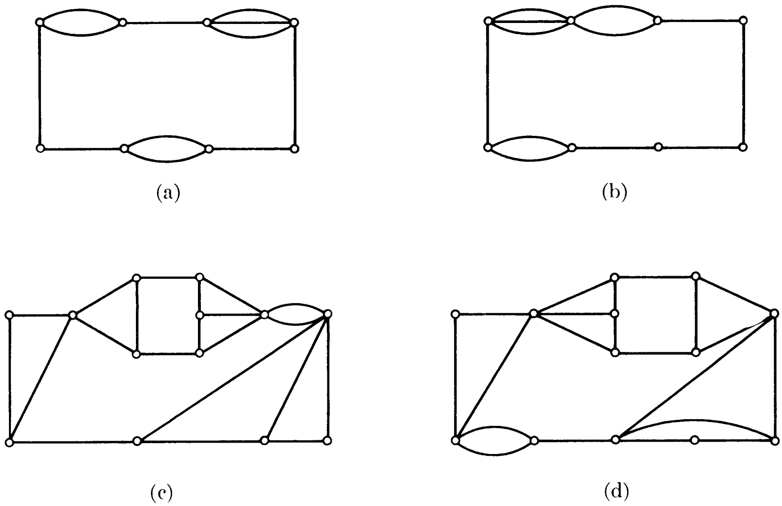


FIG. 3–4. Pairs of 2-isomorphic graphs.

(a) Separation into components.

(b) If the graph consists of two subgraphs H_1 and H_2, which have only two vertices a and b in common, the interchange of their names in one of the subgraphs.

Geometrically, the subgraph is "turned around" at these vertices, under operation 2. The most important result about 2-isomorphic graphs is the following result of Whitney.

THEOREM 3–8. If there is a one-to-one correspondence between the edges of two graphs G_1 and G_2 such that circuits correspond to circuits, then the two graphs are 2-isomorphic. Conversely, if G_1 and G_2 are 2-isomorphic (and hence have a one-to-one correspondence between their edges), then circuits in either graph correspond to circuits in the other.

The second half of the theorem is fairly evident. The proof of the first part of the theorem is too long to be given here. We therefore refer the reader to the original paper by Whitney [198]. Two pairs of 2-isomorphic graphs are shown in Fig. 3–4.

3–2 Planar graphs. The discussions up to this point have been entirely in terms of the abstract graph; the diagrams have served merely as illustrations of the theory. In this section, the problem considered is that of *mapping* a graph on a plane. Naturally, only a geometric graph can be

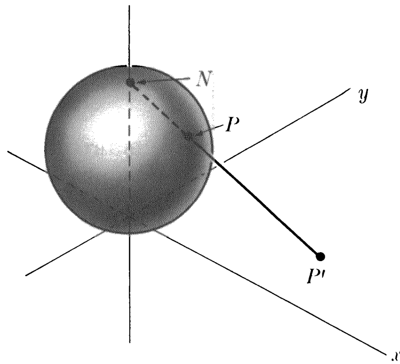


FIG. 3–5. Stereographic projection.

mapped. It was remarked in Section 1–2 that in a 3-dimensional euclidean space, a geometric structure can be associated with each abstract graph. This geometric structure is referred to as the *topological* graph. The distinction between the abstract graph and the topological graph must be carefully kept in mind. This section and parts of the next are concerned with the topological graph.

DEFINITION 3–5. *Planar graph.* A topological graph is *planar* if it can be mapped onto a plane such that no two edges have a point in common that is not a vertex. An abstract graph is planar if the corresponding topological graph is planar.

THEOREM 3–9. If a graph can be mapped (as in Definition 3–5) onto a sphere, it can be mapped onto a plane, and conversely.

Proof. To prove the theorem, we use the familiar stereographic projection of the sphere onto a plane. (This is the mapping of the complex plane onto the Riemann sphere.) The sphere is kept on the plane. The coordinate system in the plane is such that the point of contact is the origin, as in Fig. 3–5. The topmost point of the sphere is N (the north pole). Joining N to any point P of the sphere by a straight line and extending the line to meet the plane at P' establishes a one-to-one correspondence between points on the plane and points on the sphere. This procedure is referred to as mapping the plane onto the sphere, and conversely. Suppose that we have the graph mapped onto a sphere. Place the sphere on the plane so that the north pole is not a point of the graph (that is, it is not a vertex and is not on any edge of the graph). The stereographic projection now maps the graph onto the plane. The converse is proved similarly.

Definition 3–6. *Region.* The *regions of a planar graph* are the regions into which the graph divides the plane or the sphere when mapped onto the plane or the sphere.

In network theory, regions of a planar graph are usually referred to as *windows* or sometimes as *meshes.* A given region of the graph is characterized by the edges on the boundary of the region. When the graph is mapped onto a plane, the unbounded region is also referred to as the *outside region.*

Theorem 3–10. A planar graph may be mapped onto a plane such that any given region is the outside region.

Proof. Map the graph onto the sphere by Theorem 3–9. Rotate the sphere so that the north pole is inside the given region. Map the graph back onto the plane.

The most important application of planar graphs is in connection with duality, to be considered in the next section.

3–3 Dual graphs.

Following Whitney, we first give an algebraic definition of duality and later show that it agrees with the familiar geometrical definition. In a discussion of dual graphs, it is necessary to admit a "self-loop," i.e., an edge with coincident endpoints.

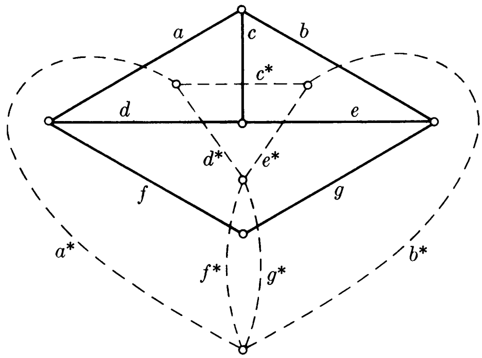
Definition 3–7. *Dual.* G_2 is a *dual* of G_1 if there is a one-to-one correspondence between the edges of two graphs G_1 and G_2 such that if H_1 is any subgraph of G_1 and H_2 is the *complement of the corresponding subgraph* of G_2, then

$$r_2 = R_2 - n_1, \tag{3–1}$$

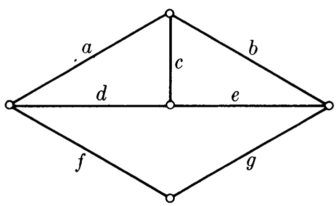
where r_2 and R_2 are ranks of H_2 and G_2, respectively, and n_1 is the nullity of H_1.

In Definition 3–7, duality is defined for abstract graphs. Since the definition is likely to be somewhat confusing at first sight, let us consider an example of duality in the usual geometric sense and illustrate the definition. In Fig. 3–6(a), two dual graphs are mapped together, the individual graphs being shown in Figs. 3–6(b) and 3–6(c). The edges are labeled such that a corresponds to a^*, b to b^*, etc. Let H_1 be the subgraph of G_1 consisting of edges a, c, d, and g. Then the corresponding subgraph of G_2 is $\{a^*, c^*, d^*, g^*\}$, so that the complement H_2 consists of edges b^*, e^*, and f^*. These two graphs are shown in Fig. 3–7. By inspection of Figs. 3–6(c) and 3–7, we see that

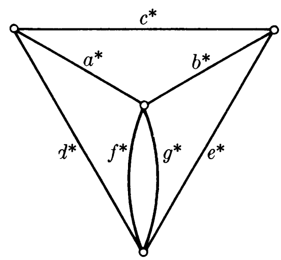
$$r_2 = 2, \qquad R_2 = 3, \qquad \text{and} \qquad n_1 = 1. \tag{3–2}$$



(a)

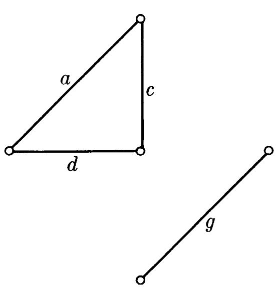


(b) G_1

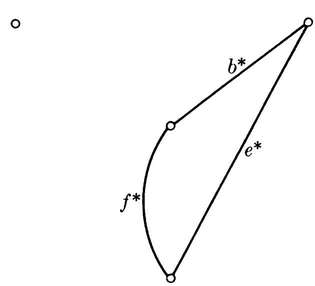


(c) G_2

FIG. 3–6. Illustration of Definition 3–7.



(a) H_1



(b) H_2

FIG. 3–7. H_1 and H_2 of Example.

Thus
$$r_2 = 2 = R_2 - n_1 = 3 - 1. \tag{3-3}$$

For another example, choose $H_1 = \{a, b, c, d, e\}$, so that $n_1 = 2$. Then $H_2 = \{f^*, g^*\}$, with $r_2 = 1$. Again this checks with the definition, since

$$R_2 - n_1 = 3 - 2 = 1. \tag{3-4}$$

Throughout this section, the notation established by Definition 3-7 is followed. Capital R and N are used for rank and nullity of a graph, and lower case letters are used for rank and nullity of subgraphs, with subscripts corresponding to graphs.

THEOREM 3-11. Let G_2 be a dual of G_1. Then

$$R_1 = N_2 \qquad \text{and} \qquad R_2 = N_1. \tag{3-5}$$

Proof. Let H_1 be a subgraph of G_1 consisting of G_1 itself. Then the corresponding subgraph of G_2 is G_2 itself, so that the complement H_2 is the null graph. Hence, $r_2 = 0$ and $n_1 = N_1$. Substituting in the definition, we find that

$$R_2 = N_1. \tag{3-6}$$

The other equation follows immediately, since the two graphs contain the same number of edges and $R + N =$ number of edges, for any graph.

THEOREM 3-12. If G_2 is a dual of G_1, then G_1 is a dual of G_2.

Proof. Let H_2 be any subgraph of G_2 and H_1 the complement of the corresponding subgraph of G_1. Since G_2 is a dual of G_1,

$$r_2 = R_2 - n_1. \tag{3-7a}$$

By Theorem 3-11,

$$R_2 = N_1. \tag{3-7b}$$

If e_1 and e_2 stand for the numbers of edges in H_1 and H_2, respectively, and E stands for the number of edges in G_1 (or G_2), we note that

$$e_1 + e_2 = E. \tag{3-8}$$

These equations give

$$\begin{aligned} r_1 = e_1 - n_1 = e_1 - (R_2 - r_2) &= e_1 - N_1 + (e_2 - n_2) \\ = E - N_1 - n_2 &= R_1 - n_2. \end{aligned} \tag{3-9}$$

Thus G_1 is a dual of G_2, by Definition 3-7.

Hence it is not necessary to say that G_1 is a dual of G_2. It suffices to say that G_1 and G_2 are dual graphs. We need the next two theorems for the proof of the main results of this section. The proof of Theorem 3–13 is not given, since it depends on several results which have not been considered here. The proof may be found in Whitney [195].

THEOREM 3–13. The dual of a nonseparable graph is nonseparable.

THEOREM 3–14. Let G_1 and G_2 be dual graphs, and let $\alpha_1(a_1, b_1)$ and $\alpha_2(a_2, b_2)$ be corresponding edges. (In this notation, a_1 and b_1 are vertices of α_1, and similarly for α_2.) Form G_1' from G_1 by deleting the edge $\alpha_1(a_1, b_1)$. Form G_2' by deleting the edge $\alpha_2(a_2, b_2)$ and letting the vertices a_2 and b_2 coalesce. Then G_1' and G_2' are duals, the correspondence between their edges being the same as in G_1 and G_2.

Proof. Let H_1' be any subgraph of G_1', and let H_2' be the complement of the corresponding subgraph of G_2'. Let H_1 be the subgraph of G_1 identical to H_1'. Then the nullities of H_1 and H_1' are the same:

$$n_1 = n_1'. \tag{3–10}$$

Let H_2 be the complement in G_2 of the subgraph corresponding to H_1. Then

$$r_2 = R_2 - n_1, \tag{3–11}$$

since G_1 and G_2 are duals. Now H_2 is the subgraph in G_2 corresponding to H_2' in G_2', except that H_2 contains $\alpha_2(a_2, b_2)$ and these vertices are distinct. If we delete $\alpha_2(a_2, b_2)$ from H_2 and let the vertices a_2 and b_2 coalesce, we form H_2'. In this operation, the number of maximal connected subgraphs remains unchanged and the number of vertices is decreased by one. Hence,

$$r_2' = r_2 - 1. \tag{3–12}$$

As a special case of this equation, if H_2 contains all the edges of G_2, then

$$R_2' = R_2 - 1. \tag{3–13}$$

Combining these equations, we find that

$$r_2' = R_2' - n_1'. \tag{3–14}$$

Hence G_2' is a dual of G_1'.

The most important result on duality is the next theorem, on the existence of a dual. This theorem and the next also relate the algebraic definition of duality with the geometrical definition.*

---

* The proofs of Theorems 3–15 and 3–16 are very involved and may be omitted if desired, without loss of continuity. They are included here because of the importance of these two theorems.

THEOREM 3–15. A graph has a dual if and only if it is planar.

Proof. The proof is considerably simplified by considering the vertices as not belonging to the edges. Since rank and nullity are unaltered by the insertion or removal of isolated vertices, the definitions apply under this convention as well.

Let G_1 be a planar graph. Let G_1 be mapped onto a sphere. If N_1 is the nullity of G_1, we observe first that G_1 divides the sphere into $N_1 + 1$ regions. To see this, construct G_1 edge by edge, starting with all the vertices in place. Initially, the graph contains v vertices and no edges, and is in v separate pieces. The rank and nullity are both zero. Every time we add an edge joining two separate pieces, the nullity and the number of regions remain the same, but the rank increases by one. Every time an edge is added, joining two vertices in the same connected subgraph, the nullity and the number of regions both increase by one. Initially, the nullity is zero and the number of regions is one. Hence after G_1 is constructed, the number of regions is $N_1 + 1$.

The graph G_2 is next constructed as follows. In each region of the graph G_1, place a vertex of the graph G_2. G_2 therefore contains $N_1 + 1$ vertices. Each edge of G_2 crosses exactly one edge of G_1, the vertices of the edge of G_2 lying in the two regions separated by the edge of G_1. Each edge of G_1 is crossed by exactly one edge of G_2. The edges of G_1 and G_2 are now in a one-to-one correspondence, as defined by the crossing relationship. In the usual geometrical (or combinatorial) sense, G_1 and G_2 are duals. It remains to prove that they are duals in the algebraic sense of Definition 3–7 as well.

Let H_1 be a subgraph of G_1, and let H_2 be the complement of the corresponding subgraph of G_2. To establish the result, we must show that $r_2 = R_2 - n_1$. For this purpose, a constructional scheme is used, simultaneously constructing H_1 and H_2. To this end, begin with G_2 on the sphere and all the vertices of G_1 in place. H_1 is now constructed by adding its edges one by one. Each time an edge of H_1 is added, delete the corresponding edge of G_2 (leaving the vertices behind). Hence when H_1 is completely constructed, H_2 is also formed. To establish the required relationship between ranks and nullities, we prove:

(1) Each time the nullity of the subgraph of G_1 is increased by one (on adding an edge), the number of connected pieces in the subgraph of G_2 is increased by one (on deleting the corresponding edge of G_2).

(2) Each time the nullity of the subgraph of G_1 is unaltered, the number of connected pieces in the subgraph of G_2 is also unaltered.

In (1), we should remember that some of the connected pieces may be isolated vertices. To prove (1), note that the nullity of the G_1-subgraph is increased only when an edge is added between two vertices in the same connected piece. Let (a_1, b_1) be such an edge, with vertices a_1 and b_1. As a_1 and b_1 were already connected by a path, this path together with the edge (a_1, b_1) forms a circuit C. Let (a_2, b_2) be the edge of G_2 corresponding to (a_1, b_1). When (a_2, b_2) is removed, the vertices a_2 and b_2 are no longer connected. For suppose that there is still a path p_2 connecting them. Since a_2 and b_2 are on opposite sides of the circuit C, p_2 must cross C. (Strictly speaking, we must appeal to the Jordan curve theorem to prove this fact.) Thus an edge of p_2 crosses an edge of C. But such an edge of p_2 was removed when the corresponding edge of C was added. Thus proposition (1) is established.

To prove (2), consider constructing the whole of the graph G_1 by this process. The total increase in nullity is then N_1. Therefore the increase in the number of connected pieces in G_2 is *at least* N_1, by (1). G_2 was initially in at least one piece, and so is finally in *at least* $N_1 + 1$ pieces. But what is left of G_2, once G_1 is constructed, is a set of v_2 isolated vertices. Since $v_2 = N_1 + 1$, by earlier construction, "at least" in the two sentences above must be replaced by "exactly." Thus G_2 is connected and the number of connected pieces increases only when the nullity of the G_1-subgraph increases.

Returning to H_1 and H_2, if we let H_2 have all the vertices of G_2, the increase in the number of connected pieces when H_2 is formed from G_2 is exactly n_1, the nullity of H_1. Since G_2 is connected, by previous argument, the rank of G_2 is $v_2 - 1$. Hence the rank of H_2 is given by

$$r_2 = v_2 - 1 - n_1 = R_2 - n_1. \tag{3-15}$$

Thus G_1 and G_2 are duals in the algebraic sense as well.

To prove the other half of the theorem, we must show that if a graph has a dual, then it is planar. Whitney [195] has shown that if the components of a graph are planar, so is the graph. (This result is fairly obvious if one considers the topological graph.) Thus it is sufficient to consider nonseparable graphs. The second part of the theorem is therefore a consequence of the following theorem.

THEOREM 3–16. Let a nonseparable graph G_1 have a dual G_2. Then G_1 and G_2 can be mapped on a sphere such that

(a) corresponding edges of G_1 and G_2 cross each other, and no other edges cross, and

(b) inside each region of one graph, there is just one vertex of the other graph.

Since the proof of the theorem is somewhat involved, we first give an outline of the proof, using diagrams, before we undertake the formal proof. The proof proceeds by induction on the number of edges in G_1

and G_2. If the graphs contain only one edge each, they can be mapped as shown in Fig. 3–8. Next, we assume that the theorem is true for all graphs with less than e edges. Consider dual graphs G_1 and G_2 with e edges. The idea of the proof is to drop one edge of G_1 and the corresponding edge of G_2, coalescing the vertices of one of the two edges. By Theorem 3–14, the new graphs G_1' and G_2' are duals and preserve the correspondence between edges. Since they have only $e - 1$ edges each, they can be mapped, as required, by the induction hypothesis. The problem now is to restore the original graphs, maintaining conditions (a) and (b). The proof is broken up into two cases for this purpose. The first (and simpler) case is that in which one of the two graphs G_1 and G_2 has a vertex of degree 2. In this case, consider the graph which has a vertex of degree 2 to be G_1. We drop one of the two edges at this vertex and coalesce the vertices of the dropped edge. The restoration merely consists of inserting a vertex on the remaining edge of G_1 and adding an edge to G_2, as in Fig. 3–9. That (a) and (b) are maintained is evident.

The other case is that in which every vertex is of degree 3 or more. Here again, an edge (a_1, b_1) of G_1 and the corresponding edge (a_2, b_2) of G_2 are dropped, and the vertices a_2 and b_2 are coalesced to form a new vertex a_2'. The new graphs G_1' and G_2' are mapped as required by the induction hypothesis. Now we must separate a_2' into two vertices and insert the dropped edges to restore the original graphs. This is more complicated than inserting a vertex on an edge as in Fig. 3–9. We must

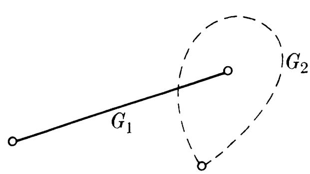


FIG. 3–8. Mapping for one edge.

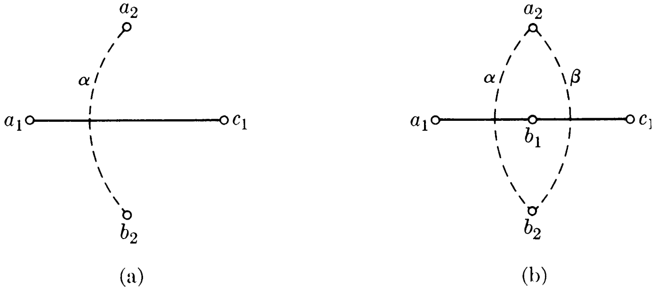


FIG. 3–9. Restoration for vertex of degree 2.

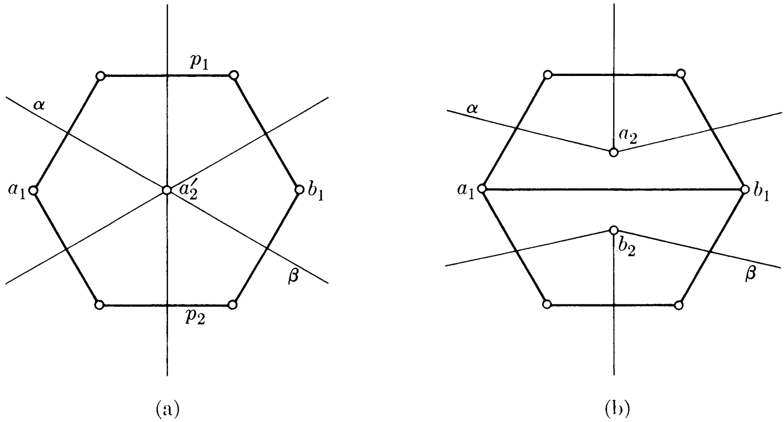


FIG. 3–10. Restoration of general case.

establish that as a_2' is separated into two vertices, no two edges of G_2 will cross each other, and when the edge (a_2, b_2) is added, it will not cross any other edge of G_2. This is established as follows.

We first show that the edges of G_1 corresponding to the edges incident at a_2 of G_2 constitute a circuit. Since we do not know that the graphs G_1 and G_2 can be mapped onto a sphere, we require an algebraic proof. Similar results hold for edges corresponding to those incident at vertices a_1, b_1, and b_2. Next, we establish that the circuit corresponding to the edges incident at a_2' consists of two paths p_1 and p_2 between the vertices a_1 and b_1, which came from the two circuits corresponding to a_2 and b_2 when (a_1, b_1) was removed. Hence when a_2' is separated into two vertices, there will be no crossing of edges of G_2. This is illustrated in Fig. 3–10. Now (a_2, b_2) is restored, and then (a_1, b_1) is restored. The need for establishing the crossing of p_1 and p_2 may be appreciated by noting that if α in Fig. 3–10(b) is incident at b_2, and β at a_2, the edge (a_1, b_1) would cross α and β and (a_2, b_2). Thus condition (a) would be violated. It is also necessary to establish that there is no "extraneous" part of G_2 inside the circuit formed by p_1 and p_2, which might be crossed by (a_1, b_1). We turn now to the formal proof.

Proof of Theorem 3–16. The theorem is easily seen to be true if the graph contains a single edge. We assume it to be true for graphs containing less than e edges and prove it for graphs with e edges. Since every edge of a nonseparable graph is in a circuit, each vertex is at least of degree 2.

Case 1. G_1 contains a vertex b_1 which is incident only with two edges (a_1, b_1) and (b_1, c_1). Since G_1 is nonseparable, there is a circuit containing these edges. Thus deleting one of them will not alter the rank, but deleting both of them reduces the rank by one. From the definition of duality, each of the two corresponding edges in G_2 is of nullity zero, and the two edges taken together are of nullity one. Thus they are of the form $\alpha(a_2, b_2)$ and $\beta(a_2, b_2)$, the first corresponding to (a_1, b_1) and the second to (b_1, c_1). Delete the edge (b_1, c_1), and let the vertices b_1 and c_1 coalesce, thus forming G_1'. Since G_1 is nonseparable, so is G_1'. Delete $\beta(a_2, b_2)$ from G_2 to form G_2'. By Theorem 3–11, G_1' and G_2' are duals and preserve the correspondence between their edges. Since these graphs contain fewer than e edges, they can be mapped together onto the sphere so that (a) and (b) hold. In particular, $\alpha(a_2b_2)$ crosses (a_1, c_1). Mark a point on the edge (a_1, c_1) of G_1' between the vertex c_1 and the point at which the edge $\alpha(a_2, b_2)$ of G_2' crosses it. Let this be the vertex b_1, dividing the edge (a_1, c_1) into the two edges (a_1, b_1) and (b_1, c_1). Draw the edge $\beta(a_2, b_2)$ crossing the edge (b_1, c_1). Now G_1 and G_2 are reconstructed and are mapped onto the sphere, as required.

Case 2. Each vertex of G_1 is at least of degree 3. Then G_1 is not a circuit and so is of nullity greater than 1. Under these conditions, it is possible to drop an edge (a_1, b_1) of G_1 (not any edge, but only a suitably chosen edge) such that the rest of the graph is still nonseparable. (The proof of this result is to be found in Whitney [195].) G_2 is nonseparable and contains more than one edge, and so the edge (a_2, b_2) corresponding to (a_1, b_1) of G_1 is not a self-loop; that is, it has distinct vertices. Delete (a_2, b_2), and let its vertices coalesce into a new vertex a_2', thus forming G_2'. By Theorem 3–14, G_1' and G_2' are duals and preserve the correspondence between their edges. Since G_1' is nonseparable, so is G_2'. Consider the edges of G_2 incident at a_2. Since a_2 is not a cut-vertex (G_2 nonseparable), these edges constitute a cut-set of edges (Theorem 2–14). Hence by Problem 3–15, the edges of G_1 corresponding to these edges of G_2 constitute a circuit C_1. One of these edges is the edge (a_1, b_1). The remaining edges constitute a path p_1. Similarly, the edges of G_1 corresponding to the edges of G_2 incident at b_2 constitute a circuit C_2, and this circuit without the edge (a_1, b_1) is a path p_2. The paths p_1 and p_2 have a_1 and b_1 for their terminal vertices. By the same argument, the edges of G_1' corresponding to the edges of G_2' incident at a_2' constitute a circuit C'. But these are the edges corresponding to the edges of G_2 incident at a_2 and b_2, except the edge (a_2, b_2), which was deleted. Hence the edges of G_1' constituting the circuit C' are the edges of the paths p_1 and p_2.

Since G_1' and G_2' have less than e edges, they can be mapped onto the sphere such that (a) and (b) hold. The vertex a_2' lies on one side of the circuit C', which we shall call the "inside." Each edge of C' is crossed by an edge incident at a_2'. Hence there are no other edges of G_2' crossing C'.

There is no part of G_2' lying inside C' other than the vertex a_2', for such a part can have only the vertex a_2' in common with its complement, whereas G_2' is nonseparable. Also, there is no part of G_1' lying inside C', for any edge must be crossed by an edge of G_2' and any vertex must be joined to the rest of G_1' by an edge, since G_1' is nonseparable.

Now replace a_2' by two vertices a_2 and b_2, restoring the original incidence for the edges at a_2' [the edge (a_2, b_2) has not yet been restored]. Since the set of edges incident at a_2 all cross the path p_1, and the set of edges incident at b_2 cross the path p_2, we can separate a_2' into a_2 and b_2 in such a way that no two edges of G_2 cross each other. We may now join a_1 and b_1 by the edge (a_1, b_1), crossing none of the other edges. This divides the circuit C' into two parts, with a_2 in one part and b_2 in the other. We may therefore join a_2 and b_2 by the edge (a_2, b_2), crossing (a_1, b_1). G_1 and G_2 are now mapped onto the sphere, as required.

A problem is suggested at the end of this chapter (Problem 3–22) to show the need for care in this argument.

Theorem 3–15 gives one characterization of planar graphs, namely that they have duals. A different characterization of planar graphs, in terms of their structure, is given by the following celebrated theorem of Kuratowski.

THEOREM 3–17 (Kuratowski). A necessary and sufficient condition that a graph be planar is that it contain neither of the following two graphs as subgraphs:

G_1. This graph is formed by taking five vertices, a, b, c, d, and e, and connecting each pair of vertices by an edge or a series connection of edges.

G_2. This graph is formed by taking two sets of three vertices, a, b, c and d, e, f, and joining each vertex of one set to each vertex of the other by an edge or a series connection of edges.

These two basic nonplanar graphs are illustrated in Fig. 3–11. The word *series* is used in the familiar sense of electrical network theory. Formally, the definitions of series and parallel are:

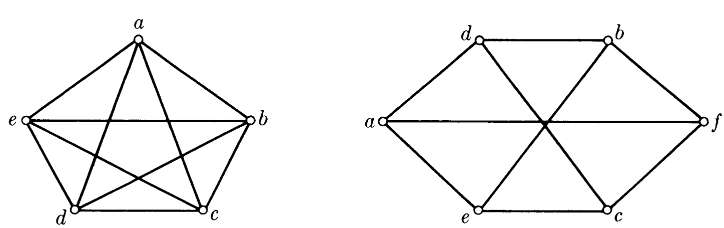


FIG. 3–11. Basic nonplanar graphs.

DEFINITION 3-8. *Series.* Two edges α and β are in *series* if they have exactly one common vertex and this vertex is of degree 2.

DEFINITION 3-9. *Parallel.* Two edges α and β are in *parallel* if they are incident at the same pair of vertices.

Kuratowski's theorem has been stated here because of its fundamental character. However, we are unable to give a proof of the theorem, since the proof depends on many point-set topological ideas that have not been developed here. The original paper of Kuratowski [94] is referred to for a proof. A proof is also given by Whitney [197]. The result itself is very useful for constructing counterexamples. A matrix method of proving that neither of the two graphs of Fig. 3-11 has a dual is suggested as a problem in Chapter 4. It follows then, from Theorem 3-15, that neither graph is planar.

In Definition 3-7 and the various theorems following it, the phrase "G_2 is *a dual* of G_1" was used instead of ". . . *the dual* . . ." One may ask whether a graph can have more than one dual and, if so, how the duals are related. We can conceive of a simple way in which two different graphs (nonisomorphic graphs) can have the same dual. Suppose that we begin with an unconnected graph G_1 and find its geometrical dual G_2 by the procedure of Theorem 3-15. From the proof of Theorem 3-15, G_2 is connected. Let us next take G_2 and construct its geometrical dual G_1' by the same procedure. Then G_1' is also connected. Hence G_1 and G_1' are both duals of G_2 and they are not isomorphic, since one is connected and the other is not. More complicated situations can occur also. The general question is answered by the next theorem.

THEOREM 3-18. The dual of a graph, when it exists, is unique within a 2-isomorphism. That is, if G_2 and G_2' are both duals of the same graph G_1, then G_2 is 2-isomorphic to G_2'.

Proof. Take any circuit of G_2 (or of G_2'). By Problem 3-15, the corresponding edges of G_1 constitute a cut-set. Again by Problem 3-15, the edges of G_2' (or of G_2) corresponding to this cut-set constitute a circuit. Thus circuits of G_2 and G_2' correspond and, by Theorem 3-8, the graphs are 2-isomorphic.

3-4 One terminal-pair graphs. In both conventional electrical networks and combinational contact networks, the concepts of separability and duality are used in connection with one terminal-pair networks in a sense that is slightly different from the definitions given in Sections 3-1 and 3-3. The following definitions serve to clarify this concept.

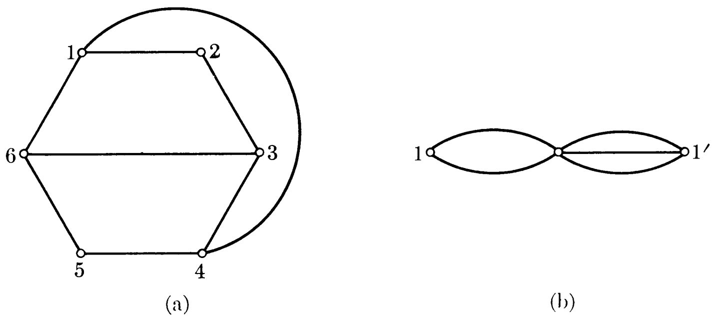


FIG. 3–12. One terminal-pair examples.

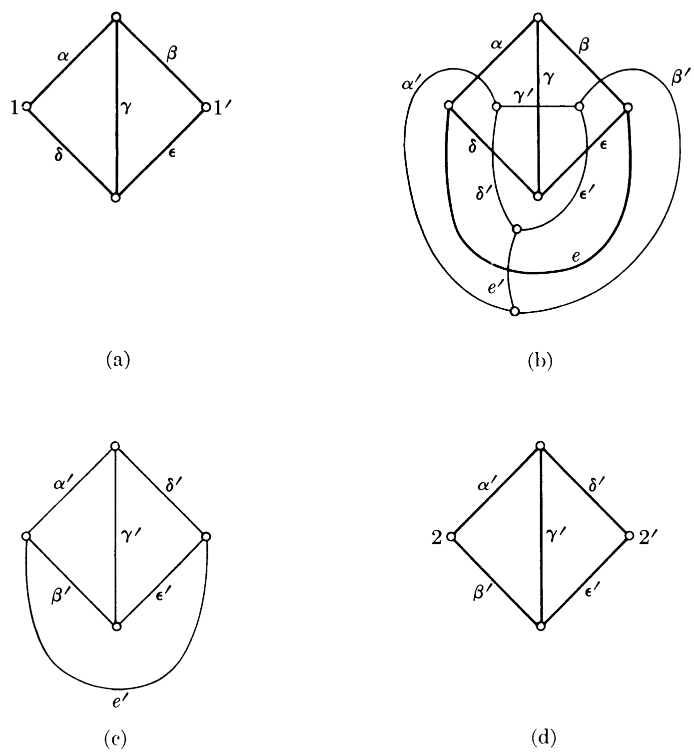


FIG. 3–13. Dual of a one terminal-pair.

DEFINITION 3–10. *One terminal-pair graph.* A *one terminal-pair graph* is a graph with two vertices (conventionally denoted by 1 and 1′) specially designated as *terminals* of the graph.

DEFINITION 3–11. *Planar one terminal-pair graph.* A one terminal-pair graph is *planar* if the graph remains planar when an edge is added between the terminals (1, 1′).

DEFINITION 3–12. *Dual of one terminal-pair graph.* The one terminal-pair graphs G_1 and G_2 with terminals (1, 1′) and (2, 2′) respectively, are *duals in the one terminal-pair sense* if the graphs obtained by adding the edges (1, 1′) and (2, 2′) are duals in the sense of Definition 3–7, with the added edges corresponding to each other.

As an example, the graph of Fig. 3–12(a) is planar. If, however, it is considered as a one terminal-pair graph, it may or may not have a dual, depending on the terminal-pair chosen. If 2 and 5 are chosen as terminals, it has no dual.

An example illustrating the process implied in Definition 3–12 is given in Fig. 3–13. Part (a) of the figure shows G_1, and part (d) shows G_2.

DEFINITION 3–13. *Nonseparable one terminal-pair graph.* A one terminal-pair graph G is nonseparable if the graph obtained by adding an edge between the terminals is nonseparable.

For example, Fig. 3–12(b) is a nonseparable one terminal-pair with terminals (1, 1′).

The concept of a dual is also useful with two (or more) terminal-pair electrical networks, but the problem is a little more involved due to the desire to have certain relationships between the network functions of the two networks. Hence the discussion is postponed to Chapter 6.

PROBLEMS

3–1. Let a graph G be a tree (of itself). Show that every vertex of G is a cut-vertex.

3–2. Prove that a connected graph G is separable if and only if G contains two edges with a common vertex v_a such that no circuit of G contains both edges.

3–3. Prove Theorem 3–5.

3–4. Prove that every nonseparable graph contains at least one circuit.

3–5. Prove that if a nonseparable graph G contains e edges and v vertices, $e \geq v$.

3–6. Prove that any edge of a nonseparable graph can be made a chord of a tree.

3–7. Prove that every cut-set of a nonseparable graph contains at least two edges.

3–8. Any two edges of a nonseparable graph can be contained in some f-circuit. True or false?

3–9. Prove that every vertex of a nonseparable graph is incident to at least two edges.

3–10. Under what conditions can any two elements of a connected graph be made chords of a tree?

3–11. Prove that the rank and nullity of a graph are invariant under the decomposition of a graph into its components.

3–12. Show that the circuits of a graph are invariant under operation (b) of Definition 3–4.

3–13. Either prove or give a counterexample: a graph is specified to within a 2-isomorphism by its rank and nullity.

3–14. Draw a few examples of planar graphs and find their duals.

3–15. Prove that if G_1 and G_2 are dual graphs, circuits in either graph correspond one-to-one with cut-sets in the other. [*Hint:* Whitney's postulates C_1, C_2, and C_3 and definitions of cut-sets and duals.]

3–16. Let G_1 and G_2 be one terminal-pair dual graphs with $(1, 1')$ and $(2, 2')$ as their terminals. Show that paths between the terminals of either graph correspond to cut-sets in the other graph, with the terminals being placed in different parts by the cut-set, and conversely.

3–17. Let G be a one terminal-pair graph with terminals $(1, 1')$. Show that every path between these terminals has an odd number of edges in common with every cut-set separating these terminals. State and prove an appropriate converse.

3–18. A one terminal-pair graph is defined to be *series-parallel* as follows:

> A single edge is series-parallel. A series or parallel combination of series-parallel graphs is series-parallel.

Show that the dual of a series-parallel graph is series-parallel. Hence show that the dual of any *non-series-parallel* (or *bridge*) graph is another bridge graph.

3–19. With dual graphs, show that disjoint unions of circuits in either graph correspond to disjoint unions of cut-sets in the other.

3–20. If G_1 and G_2 are dual graphs, show that trees in G_1 correspond to tree complements in G_2 and conversely.

3–21. If G_1 and G_2 are dual graphs, show that the f-circuits of either graph correspond to f-cut-sets in the other. [*Hint:* Problems 3–15 and 3–20.]

3–22. To illustrate the need for care in the proof of Theorem 3–16, attempt the following:

Delete one of the edges of the nonplanar graph of Fig. 3–11(a) or (b), thus making it planar. Let this be the graph G_1'. Find its dual, and let it be G_2'. Now attempt to restore the deleted edge as in the proof of Theorem 3–14 and suitably restore G_2. (Of course this is impossible, since the original graph G_1 is nonplanar; but the point at which the procedure breaks down illustrates the need for the argument of Theorem 3–16.)

CHAPTER 4

MATRICES OF A NONORIENTED GRAPH

4–1 The field modulo 2. The most convenient algebra to use in the study of nonoriented graphs is the algebra of the residue class modulo 2. This was first observed by Veblen [190]. The algebra modulo 2 consists of two elements, 0 and 1. These two symbols are used as convenient symbols and are not to be confused with the real numbers zero and one. Any two symbols, a and b for instance, might be used for the two elements; but 0 and 1 are the standard symbols. Two operations, *addition* and *multiplication*, are defined in this algebra by the rules

$$\begin{aligned} &0+0=0, \qquad 0+1=1+0=1, \qquad 1+1=0; \\ &0\cdot 0=0, \qquad 0\cdot 1=1\cdot 0=0, \qquad 1\cdot 1=1. \end{aligned} \tag{4–1}$$

Except for the addition rule $1+1=0$, the others are also the rules for the real numbers zero and one; this one rule makes the algebra distinct. To understand this algebra (and incidentally to see why the symbols 0 and 1 are used), let us list the postulates satisfied by the real number system, letting R stand for the set of all real numbers.

Postulates of Real Numbers

I. *Addition postulates:*

(i) If a and b are in R, so is $a+b$. (Closure.)
(ii) If a and b are in R, then $a+b=b+a$. (Commutative.)
(iii) If a, b, and c are in R, then $a+(b+c)=(a+b)+c$. (Associative.)
(iv) There exists a real number 0 such that $0+a=a$ for all a in R. (Identity.)
(v) For each a in R there exists b in R such that $a+b=0$. (Inverse.)

II. *Multiplication postulates:*

(i) If a and b are in R, so is $a\cdot b$. (Closure.)
(ii) If a and b are in R, then $a\cdot b=b\cdot a$. (Commutative.)
(iii) If a, b, and c are in R, then $a\cdot(b\cdot c)=(a\cdot b)\cdot c$. (Associative.)
(iv) There exists a real number 1 such that $1\cdot a=a$ for all a in R. (Identity.)
(v) For each $a\neq 0$, there exists a b such that $a\cdot b=1$. (Inverse.)

III. *Distributive law:*

If a, b, and c are in R, then $a \cdot (b + c) = a \cdot b + a \cdot c$.

IV. *Order postulates:*

(i) For each a in R, exactly one of the following three statements is true:

(1) a is positive.

(2) a is negative.

(3) $a = 0$.

(ii) If a and b are positive, so are $a \cdot b$ and $a + b$.

V. *Completeness postulate:*

Every nonempty bounded set of real numbers has a least upper bound. (Dedekind's axiom.)

Let us compare these postulates with the definition of a ring (Definition 2–1) and with some of the other algebraic systems. Note first of all that the real numbers constitute an abelian group under addition, as well as under multiplication. They are also a ring. Or looking at it the other way, we see that the various algebraic systems result if we *relax* some of the conditions imposed on real numbers. If we demand that the system satisfy only postulates (i), (iii), (iv), and (v) of addition (or of multiplication) we get an *additive* (or *multiplicative*) *group*. If we also demand that the system satisfy postulate (ii), the group becomes *abelian*. If all the postulates of addition, postulates (i) and (iii) of multiplication, and the distributive law are satisfied, the system is a *ring*. If the commutative law of multiplication is satisfied in a ring, it is a *commutative ring*. The addition of postulate (iv) makes it a *commutative ring with a unit*. A Boolean ring is an example of a commutative ring with a unit. (In fact, a Boolean ring can be defined as a commutative ring with a unit in which every element is *idempotent;* that is, $a \cdot a = a$.) The algebra of all $(n \times n)$-matrices of real or complex numbers is a ring with a unit that is not commutative.

Finally, if we add multiplication postulate (v) to a commutative ring with a unit (thus demanding the first eleven postulates), the result is a *commutative division ring*, more commonly known as a *field*. The first eleven postulates are the *algebraic postulates* of real numbers.

An example of a field that satisfies the order postulates but *not the completeness requirement* is the set of all rational numbers (real or complex). An example of a system satisfying the first eleven postulates (hence a field) and the completeness postulate (in a slightly different form) but *not the order postulates* is the set of complex numbers. Finally, any system satisfying all fourteen postulates is isomorphic to (or is indistinguishable from) the real number system.

A field is the "strongest" algebraic system. All the algebraic properties of real numbers hold in any field. The residue class modulo any prime number, defined below, is a field. Let p be the prime number. The set consists of p elements $0, 1, 2, \ldots, p - 1$. The sum of two elements a and b is found by the following procedure. First treat a and b as real integers, and find their sum:

$$a + b = q. \tag{4–2}$$

Now divide q by p to get a quotient m and remainder r:

$$q = mp + r, \qquad r < p. \tag{4–3}$$

In the algebra modulo p, we define

$$a + b = r \qquad (\text{mod } p). \tag{4–4}$$

Multiplication is performed similarly. Fields modulo a prime number are named *Galois fields* after the famous French mathematician who first formulated them.

The only Boolean ring which is also a field is the 2-element Boolean ring. This ring is isomorphic to the field modulo 2.

The familiar algebraic concepts such as linear dependence (of equations or vectors), rank of a matrix, inverse of a matrix, etc., are valid in any field. In particular, they are applicable to the field mod 2. The determinant of a matrix in the field mod 2 is found exactly as in real arithmetic, except that there are no minus signs. [Minus signs come from inverse elements for addition; but in mod 2 algebra, 1 is its own inverse (or negative) since $1 + 1 = 0$.] A few examples are now given to illustrate mod 2 algebra.

To illustrate several operations, let us find the inverse, in mod 2 algebra, of the matrix

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\cdot \tag{4–5}$$

Expanding by elements of the first column, we find that the determinant is

$$\begin{aligned} \det \mathbf{P} &= 1 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 1 \cdot (1 \cdot 0 + 0 \cdot 0) + 1 \cdot (0 \cdot 0 + 1 \cdot 0) + 1 \cdot (0 \cdot 0 + 1 \cdot 1) \\ &= 1 \cdot (0 + 0) + 1 \cdot (0 + 0) + 1 \cdot (0 + 1) \\ &= 1 \cdot 0 + 1 \cdot 0 + 1 \cdot 1 = 0 + 0 + 1 = 1. \end{aligned} \tag{4–6}$$

Note that all signs are "+" but that all other rules are the same as in ordinary arithmetic. The cofactors are

$$
\begin{aligned}
\Delta_{11} &= \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} = 0 + 0 = 0, \\
\Delta_{12} &= \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} = 0 + 0 = 0, \\
\Delta_{13} &= \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = 0 + 1 = 1, \\
\Delta_{21} &= \begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} = 0 + 0 = 0, \\
\Delta_{22} &= \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = 0 + 1 = 1, \\
\Delta_{23} &= \begin{vmatrix} 1 & 0 \\ 1 & 0 \end{vmatrix} = 0 + 0 = 0, \\
\Delta_{31} &= \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 0 + 1 = 1, \\
\Delta_{32} &= \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = 0 + 1 = 1, \\
\Delta_{33} &= \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1 + 0 = 1.
\end{aligned}
\tag{4-7}
$$

Hence the inverse matrix is

$$
\mathsf{P}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}. \tag{4-8}
$$

Note that the determinant can be only 1 or 0, for there are no other elements in this algebra. So, the inverse matrix also can have no entries other than 1 and 0.

As another example of mod 2 algebra and as an illustration of the procedure of Theorem 4-4 to follow, let us reduce the following matrix by elementary operations and determine its rank:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}. \tag{4–9}$$

Add the first row separately to the third row and to the fourth row, to reduce the first column to zeros below row 1. The result is $(1 + 1 = 0)$

$$\mathbf{P}_1 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}. \tag{4–10}$$

There is no 1 in the (2, 2)-position, and so interchange column 3 with column 2. The result is

$$\mathbf{P}_2 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{4–11}$$

(Note the change in the symbol for the matrix to emphasize that these matrices are different from $\mathbf{P}$ and from each other. All of them have the same rank, however.)

Next add row 2 to row 4, to get

$$\mathbf{P}_3 = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4–12}$$

Since there is a 1 in the (3, 3)-position and a 0 in the (3, 4)-position, move next to the fourth row. Since the (4, 4)-element is 0, we interchange the fourth column and the sixth column, getting finally

$$\mathbf{P}_4 = \left[\begin{array}{cccc|cc} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right]. \tag{4–13}$$

The submatrix consisting of the first four columns is

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4–14}$$

The zeros below the main diagonal show that det $\mathbf{Q}$ is simply the product of the diagonal entries. For, on expanding det $\mathbf{Q}$ by the first column, we find that

$$\det \mathbf{Q} = 1 \cdot \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix}, \tag{4–15}$$

since the other elements of the first column are zero. Repeat the process and expand this (3×3)-determinant by the first column:

$$\det \mathbf{Q} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1(1 + 0) = 1. \tag{4–16}$$

Hence $\mathbf{P}_4$ contains a submatrix of order 4×4 which is nonsingular. Since $\mathbf{P}_4$ cannot contain a square submatrix of a larger order, $\mathbf{P}_4$ has a rank of 4. Hence the rank of $\mathbf{P}$ is also 4.

As a third example, let us find the general solution of the system of equations

$$x_1 + x_2 + x_3 = 1, \tag{4–17a}$$

$$x_2 + x_3 = 0. \tag{4–17b}$$

Adding the two equations yields

$$x_1 + (1 + 1)x_2 + (1 + 1)x_3 = 1 + 0 = 1 \tag{4–18a}$$

or

$$x_1 = 1. \tag{4–18b}$$

Adding x_3 to both sides of Eq. (4–17b) yields

$$x_2 + (1 + 1)x_3 = 0 + x_3 \tag{4–19a}$$

or

$$x_2 = x_3. \tag{4–19b}$$

Hence the general solution is

$$x_1 = 1, \qquad x_2 = t, \qquad x_3 = t, \tag{4–20}$$

where t is a parameter (equal to 0 or 1).

4–2 The vertex or incidence matrix. We have already observed, in the axiom of graph theory, that the most fundamental characteristic of a graph is the interconnection between edges and vertices. The graph is completely specified as soon as we specify which edges are incident at which vertices. Such a specification is most conveniently done by means of a matrix. We make each row of the matrix correspond to a vertex, and each column to an edge. If the edge is connected at a vertex, we write 1; otherwise we write 0. Precisely, the definition is as follows.

DEFINITION 4–1. *Vertex*, or *incidence matrix*, A_a.
$\mathsf{A}_a = [a_{ij}]$ is a matrix of v rows and e columns for a graph of v vertices and e edges, where
$a_{ij} = 1$ if the edge j is incident at vertex i,
$a_{ij} = 0$ if the edge j is not incident at vertex i.

The subscript a denotes that all the vertices of the graph are represented. We will be dealing mostly with $v - 1$ rows of the matrix A_a, and so it will be convenient to reserve the simpler symbol A for this purpose. For example, the incidence matrix A_a of Fig. 4–1 is

$$\mathsf{A}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array}. \tag{4–21}$$

From inspection of this matrix, the following theorem is obvious.

THEOREM 4–1. Every column of A_a contains exactly two nonzero elements.

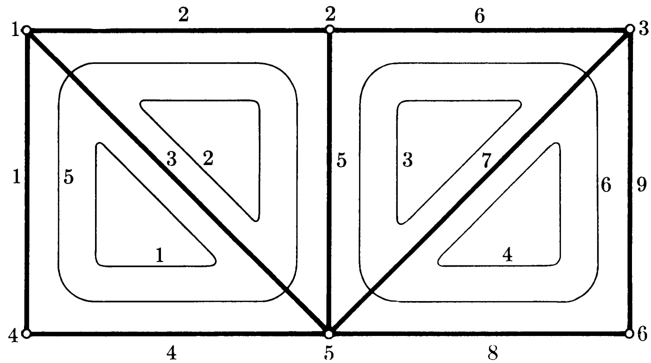


FIG. 4–1. Example.

This is the fundamental characterization of vertex matrices. Notice also, in passing, that the incidence matrix is equivalent to the graph in the sense that each is determined completely by the other. This leads us to the next theorem.

THEOREM 4–2. If graphs G_1 and G_2 have incidence matrices which differ only by a permutation of rows and columns, then G_1 and G_2 are isomorphic; and conversely.

Thus, all the information about the graph is contained in the incidence matrix. It will require the remainder of this chapter to demonstrate the full significance of this remark. The first property of interest about a matrix is its rank. Since the algebra involved in this chapter is entirely modulo 2, our discussion of rank is with respect to modulo 2 algebra.

THEOREM 4–3. The rank of the vertex matrix A_a of a connected graph is at most $v - 1$, where v is the number of vertices.

Proof. Add all the rows to the last row (which may be any row). This is an elementary operation which does not change the rank. Since each column contains exactly two nonzero elements (1's), the last row becomes a row of zeros ($1 + 1 = 0$ in modulo 2 algebra). Since the matrix has only $v - 1$ nonzero rows, the rank cannot exceed $v - 1$.

LEMMA 4–4. For a connected graph G, the sum of any r rows of A_a, with $r < v$, contains at least one nonzero element.

Proof. By contradiction, let $r < v$, and let r rows of A_a add to a row of zeros. Let the rows of A_a be rearranged such that these r rows are at the top. Since these r rows add to a row of zeros, it follows that each column of these r rows contains either two or no nonzero elements. Let the columns of A_a be permuted so that the columns with no nonzero elements in the first r rows are the last columns. There must be some columns like this, or the last $v - r$ rows contain only zeros, which is impossible since G contains no isolated vertices. Now in the first set of columns where the first r rows contain both 1's, the last $v - r$ rows must contain only zeros. Partitioned in this manner, the matrix $\mathbf{A}_a$ becomes

$$\mathbf{A}_a = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}. \tag{4–22}$$

It is seen that the first r vertices have no common elements with the last $v - r$. Hence the graph is not connected, contradicting the hypothesis.

THEOREM 4–4. The rank of the vertex matrix $\mathbf{A}_a$ of a connected graph is $v - 1$, where v is the number of vertices of the graph.

Proof. Two proofs of this important theorem are given under (a) and (b). The first, and the more elegant, makes use of concepts of linear dependence. The second is a direct proof.

(a) Let $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_v$ be the rows of $\mathbf{A}_a$. Let c_j be scalars from field modulo 2; that is, let $c_j = 0$ or 1. Then the equation

$$\sum_{j=1}^{v} c_j \mathbf{A}_j = \mathbf{0}$$

has only one nonzero solution for c_j's, namely

$$c_1 = c_2 = \cdots = c_v = 1,$$

by Lemma 4–4 and Theorem 4–3. Thus there is only one independent relation among the rows of $\mathbf{A}_a$. Since $\mathbf{A}_a$ has v rows, the rank of $\mathbf{A}_a$ is $v - 1$. (This was Kirchhoff's original proof.)

(b) Let the first $v - 1$ rows of $\mathbf{A}_a$ be added to the last row. The last row is thereby reduced to zeros. The first row contains a nonzero element. By permutation of columns, let this be brought to the (1, 1)-position. If there is any other nonzero element in column 1, say in the kth row, $k > 1$, let the first row be added to the kth row, to reduce this element to 0. By Lemma 4–4, the kth row still contains nonzero elements. There is a nonzero element in the second row, and it is not in the first column, since (after the preceding addition) the first column contains only zeros after the first row. By interchange of columns, let this nonzero element be brought to the (2, 2)-position. If the second column contains any nonzero elements below the second row, let the second row be used to reduce it to zero. Again, less than v rows have been added, and so no zero row is produced thereby. We see by repeated application of this procedure that the vertex matrix is reduced to the form

$$\begin{bmatrix} 1 & - & - & - & \cdots & - & - & - & \cdots & - \\ 0 & 1 & - & - & \cdots & - & - & - & \cdots & - \\ 0 & 0 & 1 & - & \cdots & - & - & - & \cdots & - \\ 0 & 0 & 0 & 1 & \cdots & - & - & - & \cdots & - \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & - & \cdots & - \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where the dashes may be 0 or 1. The leading square submatrix of order $v - 1$ is triangular with nonzero elements on the main diagonal and so is nonsingular. Hence the rank of A_a is $v - 1$.

The last row in the proof of Theorem 4–4 is arbitrary, so we may state the following corollary:

COROLLARY 4–4. If any row of the matrix A_a of a connected graph is omitted, the resulting matrix A has a rank of $v - 1$.

The symbol A is always used in this text to denote the incidence matrix of $v - 1$ rows of a connected graph. A is also referred to as the *vertex matrix*.

4–3 The circuit matrix. Just as we describe the relation between vertices and edges by a matrix, we also define a matrix relating edges and circuits.

DEFINITION 4–2. *The circuit matrix* B_a.

$\mathsf{B}_a = [b_{ij}]$ contains one row for each circuit of G and contains e columns; and

$b_{ij} = 1$ if element j is in circuit i

$b_{ij} = 0$ if element j is not in circuit i.

Under the definition of a circuit, a finite graph contains only a finite number of circuits. Hence B_a is finite.

As an example, let us construct the matrix B_a for the graph of Fig. 4–1. Six loops are shown in Fig. 4–1. The graph contains four more loops that have not been shown because the figure would become too confusing. These latter loops are

loop 7: (1, 2, 6, 9, 8, 4), loop 8: (2, 6, 7, 3),

loop 9: (1, 2, 6, 7, 4), loop 10: (2, 6, 9, 8, 3).

The matrix $\mathbf{B}_a$ of the set of all circuits is

$$\mathbf{B}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \end{array} \tag{4-23}$$

The fundamental set of circuits defined earlier (Definition 2-8) have an interesting circuit matrix. To have a fundamental system, we must choose a tree T of the graph. Let the f-circuits be numbered in some arbitrary manner as $1, 2, \ldots, e - v + 1$. Let the chord that appears in circuit i be numbered as edge i, $1 \leq i \leq e - v + 1$. Let the branches be numbered in some arbitrary manner as $e - v + 2, e - v + 3, \ldots, e$. If we arrange the rows and columns of the matrix in the order in which the circuits and the elements have been numbered, the matrix of f-circuits appears as

$$\mathbf{B}_f = [\mathbf{U} \quad \mathbf{B}_{12}], \tag{4-24}$$

where $\mathbf{U}$ is the unit or identity matrix of order $e - v + 1$. As an example, consider the graph of Fig. 4-2. If we choose the tree consisting of edges 1, 2, and 3, the f-circuits are as shown in the figure. The matrix of these

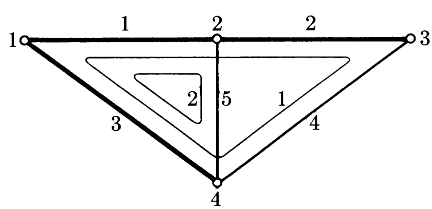


FIG. 4-2. Example for fundamental systems.

circuits is

$$\mathsf{B}_f = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right] \end{array}. \tag{4–25}$$

The matrix B_f obviously has the rank of $e - v + 1$. Further, since the f-circuits are always part of the set of all circuits, B_f is a submatrix of B_a. This leads to our next theorem.

THEOREM 4–5. The rank of the circuit matrix B_a is at least $e - v + 1$ for a connected graph G of v vertices and e elements.

To establish that $e - v + 1$ is also an upper bound for the rank of B_a, we need the following theorem. This result is of very fundamental importance even apart from establishing the rank of B_a, as is evident from the rest of this chapter.

THEOREM 4–6. If the columns of the matrices A_a and B_a are arranged in the same element order,

$$\mathsf{A}_a\mathsf{B}_a' = \mathbf{0} \quad \text{and} \quad \mathsf{B}_a\mathsf{A}_a' = \mathbf{0}, \tag{4–26}$$

where the prime indicates the transpose.

Proof. Consider the ith row of A_a and the rth column of B_a', that is, the rth row of B_a. There are nonzero elements in the corresponding positions in the ith row of A_a and the rth row of B_a if and only if the element is incident at vertex i and is in circuit r. If the vertex i is not in circuit r, there is no such element and the product is zero. If the vertex i is in circuit r, then by Veblen's definition of a circuit, exactly two elements are at vertex i and in circuit r, and so the product of the ith row of A_a by the rth column of B_a' is

$$1 + 1 = 0 \bmod 2. \tag{4–27}$$

Hence the theorem.

Let us verify Theorem 4–6 for the graph of Fig. 4–2 as an illustration. The graph has one more circuit that is not shown in the figure, consisting of edges 2, 4, and 5. The matrices A_a and B_a for Fig. 4–2 are (the order of edges is kept as in B_f above for illustrative purposes)

$$\mathsf{A}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \left[\begin{array}{ccccc} 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right] \end{array} \tag{4–28}$$

and

$$\mathsf{B}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array} . \tag{4–29}$$

Hence,

$$\mathsf{A}_a\mathsf{B}_a' = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \tag{4–30}$$

Theorem 4–6 immediately establishes an upper bound for the rank of B_a if we make use of the theorem known as *Sylvester's law of nullity.* This theorem and its proof are given below. The elements of the matrices are assumed to be from a field, so that rank, reciprocal, etc., are meaningful.

THEOREM (Sylvester's law of nullity). If

$$\mathsf{P} = [p_{ij}]_{m,n} \qquad \text{and} \qquad \mathsf{Q} = [q_{ij}]_{n,p}$$

are matrices of elements from a field, and if

$$\mathsf{PQ} = \mathbf{0},$$

then

$$(\text{rank of } \mathsf{P}) + (\text{rank of } \mathsf{Q}) \leq n.$$

Proof. Let the rank of P be r. Let the rows and columns of P be rearranged to bring a nonsingular submatrix of order r to the leading position, and let the resulting matrix P_1 be partitioned so that

$$\mathsf{P}_1 = \begin{bmatrix} \mathsf{P}_{11} & \mathsf{P}_{12} \\ \mathsf{P}_{21} & \mathsf{P}_{22} \end{bmatrix} \tag{4–31}$$

and P_{11} is of order r and nonsingular. Then P_{12} contains $n - r$ columns. Let the rows of Q be rearranged to correspond to the rearrangement of the columns of P, and let the rearranged Q be partitioned after r rows, so that

$$\begin{bmatrix} \mathsf{P}_{11} & \mathsf{P}_{12} \\ \mathsf{P}_{21} & \mathsf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathsf{Q}_{11} \\ \mathsf{Q}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} . \tag{4–32}$$

From the first row,

$$\mathsf{P}_{11}\mathsf{Q}_{11} + \mathsf{P}_{12}\mathsf{Q}_{21} = \mathbf{0} \qquad \text{or} \qquad \mathsf{Q}_{11} + \mathsf{P}_{11}^{-1}\mathsf{P}_{12}\mathsf{Q}_{21} = \mathbf{0}. \tag{4–33}$$

Premultiply the rearranged Q by the nonsingular matrix

$$\begin{bmatrix} \mathsf{U} & \mathsf{P}_{11}^{-1}\mathsf{P}_{12} \\ \mathbf{0} & \mathsf{U} \end{bmatrix}.$$

The result is

$$\begin{bmatrix} \mathsf{U} & \mathsf{P}_{11}^{-1}\mathsf{P}_{12} \\ \mathbf{0} & \mathsf{U} \end{bmatrix} \begin{bmatrix} \mathsf{Q}_{11} \\ \mathsf{Q}_{21} \end{bmatrix} = \begin{bmatrix} \mathsf{Q}_{11} + \mathsf{P}_{11}^{-1}\mathsf{P}_{12}\mathsf{Q}_{21} \\ \mathsf{Q}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathsf{Q}_{21} \end{bmatrix}. \tag{4–34}$$

Premultiplication by a nonsingular matrix does not change the rank. The matrix on the right contains only $n - r$ nonzero rows. Hence (rank of Q) $\leq n - r = n -$ (rank of P), or (rank of P) $+$ (rank of Q) $\leq n$.

Using Theorem 4–6 and Sylvester's law of nullity, we immediately have the following two theorems.

THEOREM 4-7. For any graph G,

$$(\text{rank of } \mathsf{A}_a) + (\text{rank of } \mathsf{B}_a) \leq (\text{number of edges}).$$

THEOREM 4–8. For a connected graph G,

$$(\text{rank of } \mathsf{B}_a) \leqq e - v + 1.$$

Finally, from Theorems 4–5 and 4–8, we have Theorem 4–9:

THEOREM 4–9. For a connected graph G,

$$(\text{rank of } \mathsf{B}_a) = e - v + 1.$$

The symbol B is reserved in this text for a circuit matrix of a connected graph with $e - v + 1$ rows and rank $e - v + 1$.

4–4 Nonsingular submatrices of A and B and formula for B_f. In this section, we shall see further evidence of the importance of the concept of a tree. The nonsingular submatrices of A and B are very closely related to the topology of the graph, and it is the purpose of this section to investigate this relationship.

LEMMA 4–10. There exists a linear relationship among the columns of the vertex matrix A which correspond to the edges of a circuit.

Proof.

$$\mathsf{B}_a\mathsf{A}_a' = \mathbf{0}. \tag{4–35}$$

Consider circuit r. On partitioning $\mathbf{A}_a$ into columns and multiplying by the rth row of $\mathbf{B}_a$, we find that

$$[b_{r1} \quad b_{r2} \quad \cdots \quad b_{re}]\begin{bmatrix} \mathbf{A}'_1 \\ \mathbf{A}'_2 \\ \vdots \\ \mathbf{A}'_e \end{bmatrix} = \mathbf{0} \quad \text{or} \quad b_{r1}\mathbf{A}_1 + b_{r2}\mathbf{A}_2 + \cdots + b_{re}\mathbf{A}_e = \mathbf{0}. \tag{4–36}$$

If elements $i_1, i_2, \ldots, i_k$ are in this circuit,

$$b_{ri_1} = b_{ri_2} = \cdots = b_{ri_k} = 1 \tag{4–37}$$

and all other $b_{rs} = 0$. Hence

$$\mathbf{A}_{i_1} + \mathbf{A}_{i_2} + \cdots + \mathbf{A}_{i_k} = \mathbf{0}, \tag{4–38}$$

which is a linear relationship among the columns of $\mathbf{A}$ corresponding to the edges of a circuit.

THEOREM 4–10. A square submatrix of $\mathbf{A}$ of order $v - 1$ is nonsingular if and only if the elements corresponding to these columns of $\mathbf{A}$ constitute a tree of the graph.

Proof. One half of the proof is listed as a problem (Problem 4–2). Let $v - 1$ columns constitute a nonsingular submatrix of $\mathbf{A}$. The columns are therefore linearly independent. Hence the corresponding subgraph contains $v - 1$ elements and contains no circuits. Hence by Theorem 2–10, the subgraph is a tree.

Thus the trees of the graph are in one-to-one correspondence with the nonsingular submatrices of $\mathbf{A}$. This is a very fundamental relationship and points out the importance of a tree.

A dual relationship exists between the nonsingular submatrices of $\mathbf{B}$ and the chord sets.

THEOREM 4–11. Let $\mathbf{B}$ be a submatrix of $\mathbf{B}_a$ with $e - v + 1$ rows and of rank $e - v + 1$ for a connected graph G. Then a square submatrix of $\mathbf{B}$ of order $e - v + 1$ is nonsingular if and only if the columns of this submatrix correspond to a set of chords.

Proof. (a) Let the columns correspond to a set of chords, and let the columns of $\mathbf{B}$ be arranged so that this submatrix appears in the leading position. Partition $\mathbf{B}$ as

$$\mathbf{B} = [\mathbf{B}_{11} \quad \mathbf{B}_{12}], \tag{4–39}$$

where $\mathbf{B}_{11}$ corresponds to a set of chords and $\mathbf{B}_{12}$ corresponds to a tree T. There is a fundamental set of circuits for the tree T, with a matrix $\mathbf{B}_f$ which, partitioned similar to $\mathbf{B}$, is

$$\mathbf{B}_f = [\mathbf{U} \quad \mathbf{B}_{f_{12}}]. \tag{4–40}$$

Since $\mathbf{B}_f$ is a basis for the set of all circuits, there is a matrix $\mathbf{D}$ such that

$$\mathbf{B} = \mathbf{D}\mathbf{B}_f. \tag{4–41}$$

Further, since the circuits of $\mathbf{B}$ are independent (i.e., since the rank of $\mathbf{B}$ is $e - v + 1$), $\mathbf{D}$ is nonsingular. Now

$$[\mathbf{B}_{11} \quad \mathbf{B}_{12}] = \mathbf{D}[\mathbf{U} \quad \mathbf{B}_{f_{12}}], \tag{4–42}$$

from which

$$\mathbf{B}_{11} = \mathbf{D}\mathbf{U} = \mathbf{D}, \tag{4–43}$$

so that $\mathbf{B}_{11}$ is nonsingular.

(b) Let the $e - v + 1$ columns constitute a nonsingular submatrix. Let $\mathbf{B}$ be arranged as

$$\mathbf{B} = [\mathbf{B}_{11} \quad \mathbf{B}_{12}], \tag{4–44}$$

and let $\mathbf{B}_{11}$ be nonsingular. There are $v - 1$ columns in $\mathbf{B}_{12}$, and so it is sufficient to prove that there is no circuit consisting only of elements corresponding to columns in $\mathbf{B}_{12}$. If there is such a circuit, let the row $\mathbf{B}_i$ corresponding to this circuit be added to $\mathbf{B}$. Then

$$\begin{bmatrix} \mathbf{B} \\ \mathbf{B}_i \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{0} & \mathbf{B}_{i_2} \end{bmatrix}. \tag{4–45}$$

There is at least one nonzero element in $\mathbf{B}_{i_2}$. By arranging the last $v - 1$ columns, we can bring this nonzero element to the $(e - v + 2, e - v + 2)$-position. The leading square submatrix of order $e - v + 2$ is now

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{0} & 1 \end{bmatrix},$$

which is nonsingular. Hence rank of

$$\begin{bmatrix} \mathbf{B} \\ \mathbf{B}_i \end{bmatrix}$$

is $e - v + 2$. But

$$\begin{bmatrix} \mathsf{B} \\ \mathsf{B}_i \end{bmatrix}$$

is a submatrix of B_a which is of rank $e - v + 1$. Hence this is impossible.

We could also have deduced Theorem 4–11 from Theorem 4–10 and the fundamental orthogonality relation, Theorem 4–6. Whitney [199] shows in his fundamental paper on linear dependence that with dual matroids such as are defined by A and B, *bases in one correspond to base complements in the other*.

We stated in Section 4–2 that the incidence matrix contains all the information about the graph. As an illustration, we next give an explicit formula for the f-circuit matrix B_f in terms of the vertex matrix A.

THEOREM 4–12. The vertex matrix can always be partitioned as

$$\mathsf{A}_a = \begin{bmatrix} \mathsf{A}_{11} & \mathsf{A}_{12} \\ \mathsf{A}_{21} & \mathsf{A}_{22} \end{bmatrix}, \tag{4–46}$$

where A_{12} is a square nonsingular matrix of order $v - 1$. Then the matrix of f-circuits for the tree corresponding to the columns of A_{12} is given by

$$\mathsf{B}_f = [\mathsf{U} \quad \mathsf{A}'_{11}\mathsf{A}_{12}^{-1\prime}]. \tag{4–47}$$

Proof. Since A_a is of rank $v - 1$, the indicated partitioning is always possible. Then of course the columns of A_{12} correspond to branches of a tree. Let B_f be the matrix of f-circuits for this tree. Arranged in the usual order, B_f may be partitioned as

$$\mathsf{B}_f = [\mathsf{U} \quad \mathsf{B}_{f_{12}}]. \tag{4–48}$$

Since

$$\mathsf{A}\mathsf{B}'_f = \mathbf{0}, \tag{4–49}$$

we have

$$\mathsf{A}_{11} + \mathsf{A}_{12}\mathsf{B}'_{f_{12}} = \mathbf{0} \quad \text{or} \quad \mathsf{B}'_{f_{12}} = \mathsf{A}_{12}^{-1}\mathsf{A}_{11} \quad \text{or} \quad \mathsf{B}_{f_{12}} = \mathsf{A}'_{11}\mathsf{A}_{12}^{-1\prime}. \tag{4–50}$$

As an example, let us verify Theorem 4–12 for the graph of Fig. 4–2. Deleting the row corresponding to vertex 4, we find that the incidence matrix of Fig. 4–2 is

$$\mathsf{A} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}. \tag{4–51}$$

Let us rearrange the incidence matrix to be in the same column order as B_f given earlier:

$$\mathsf{A}_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 4 & 5 & 1 & 2 & 3 \end{array} \\ \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{array}\right] \end{array}. \tag{4-52}$$

Now to verify Theorem 4–12:

$$\mathsf{A}_{12}^{-1}\mathsf{A}_{11} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{4-53}$$

Hence

$$[\mathsf{U} \quad (\mathsf{A}_{12}^{-1}\mathsf{A}_{11})'] = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}, \tag{4-54}$$

which is indeed the same matrix as B_f.

4–5 The cut-set matrix. The concept of a cut-set was defined in Chapter 2. In this section, a matrix formulation of the concept is used to tie together the cut-set properties developed earlier. It is the purpose of this section to show that the cut-set matrix and the incidence matrix contain essentially the same information.

DEFINITION 4–3. *Cut-set matrix.* The *cut-set matrix* $\mathsf{Q} = [q_{ij}]$ has one row for *each possible cut-set* and one column for each edge and is defined by

$q_{ij} = 1$ if edge number j is in cut set i

$q_{ij} = 0$ if edge number j is not in cut set i.

For example, the cut-set matrix of the seven possible cut-sets of the

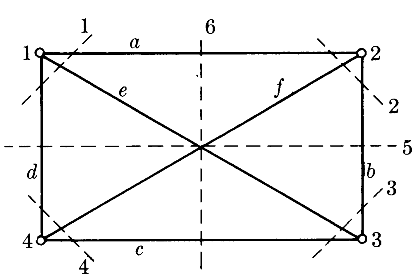


FIG. 4–3. Example for cut-set matrix.

graph of Fig. 4–3 is given by

$$\mathbf{Q}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \end{array}. \tag{4–55}$$

The last cut-set cannot be shown very easily on the diagram as drawn (see Fig. 2–8).

We are interested in answering the following questions about the matrix $\mathbf{Q}_a$.

(a) What is the rank of $\mathbf{Q}_a$?
(b) How is $\mathbf{Q}_a$ related to the incidence matrix $\mathbf{A}$ and to the circuit matrix $\mathbf{B}$?
(c) What are the nonsingular submatrices of $\mathbf{Q}_a$?
(d) How is the matrix of fundamental cut-sets related to the matrix of fundamental circuits (both formed with respect to the same tree)?

We answer these questions more or less in the order in which they are stated except that (a) and (b) are taken up together. Theorem 4–13 is obvious:

THEOREM 4–13. For a nonseparable graph G, the matrix $\mathbf{Q}_a$ contains the matrix $\mathbf{A}_a$. For any connected graph G, the rows of $\mathbf{A}_a$ are expressible as linear combinations (sums mod 2) of the rows of $\mathbf{Q}_a$.

Theorem 4–13 follows immediately from Theorem 2–13 and the elementary observations about cut-vertices.

COROLLARY 4–13. If G is a connected graph of v vertices, the rank of $\mathbf{Q}_a$ is at least $v - 1$.

THEOREM 4–14. If the columns of the matrices $\mathbf{Q}_a$ and $\mathbf{B}_a$ are arranged in the same element order,

$$\mathbf{B}_a\mathbf{Q}_a' = \mathbf{0} \bmod 2. \tag{4–56}$$

Theorem 4–14 is simply an elegant restatement of Theorem 2–17 since the sum of an even number of 1's is 0 in mod 2 algebra. The details of the proof are left as a problem.

COROLLARY 4–14. The rank of $\mathbf{Q}_a$ is at most $v - 1$ for a connected graph G of v vertices.

This corollary is an immediate consequence of Theorem 4–14 and Sylvester's law of nullity (Section 4–3). Since the upper and lower bounds for the rank of $\mathbf{Q}_a$ are equal, we have the next theorem.

THEOREM 4–15. The rank of the cut-set matrix $\mathbf{Q}_a$ of a connected graph G of v vertices is $v - 1$.

We now have the following situation. The rows of the incidence matrix $\mathbf{A}$ are linear combinations of the rows of $\mathbf{Q}_a$; or if we were to select a submatrix $\mathbf{Q}$ of $v - 1$ rows and rank $v - 1$ from $\mathbf{Q}_a$, all rows of $\mathbf{A}$ are expressible as linear combinations of the rows of $\mathbf{Q}$; that is, there exists a matrix $\mathbf{D}$ of order $(v - 1) \times (v - 1)$ such that

$$\mathbf{A} = \mathbf{DQ}. \tag{4–57}$$

But both $\mathbf{A}$ and $\mathbf{Q}$ are of rank $v - 1$ and contain $v - 1$ rows. This is possible if and only if the matrix $\mathbf{D}$ is nonsingular. But if $\mathbf{D}$ is nonsingular, we can write

$$\mathbf{Q} = \mathbf{D}^{-1}\mathbf{A}, \tag{4–58}$$

where $\mathbf{D}^{-1}$ is also nonsingular. Thus, not only can the rows of $\mathbf{A}$ be expressed in terms of the rows of $\mathbf{Q}$, but the rows of $\mathbf{Q}$ can also be expressed in terms of rows of $\mathbf{A}$. More generally, we have, by using Problem 2–24, the following theorem:

THEOREM 4–16. Each row of a matrix $\mathbf{F}$ of order $(v - 1) \times e$ and rank $v - 1$ corresponds to a cut-set or element-disjoint union of cut-sets if and only if

$$\mathbf{F} = \mathbf{DA}, \tag{4–59}$$

where $\mathbf{D}$ is nonsingular.

We may restate the same result with the use of the circuit matrix:

THEOREM 4–17. Let $\mathbf{F}$ be a matrix of order $(v - 1) \times e$ and rank $v - 1$ such that

$$\mathbf{BF}' = \mathbf{0} \bmod 2, \tag{4–60}$$

where $\mathbf{B}$ is the circuit matrix of G. Then each row of $\mathbf{F}$ corresponds to a cut-set or element-disioint union of cut-sets.

This result is very important because it gives us a means of constructing the cut-set matrix and hence (by using Theorem 4–16) a means of constructing the incidence matrix from the circuit matrix. However, the result should be considered to be obvious in the light of Theorem 2–18.

Knowing the relationship of the cut-set matrix to the incidence matrix (Theorem 4–16) and the structure of the nonsingular submatrices of the incidence matrix (Theorem 4–10), we can immediately answer the third question raised above:

THEOREM 4–18. *If* $\mathbf{Q}$ *is a cut-set matrix of* $v - 1$ *rows and rank* $v - 1$ *of a connected graph* G, *the nonsingular submatrices of* $\mathbf{Q}$ *of order* $v - 1$ *are in one-to-one correspondence with the trees of* G.

Finally, let us turn our attention to the matrix of fundamental cut-sets. To have a fundamental system, we have to choose a tree T; so it is also natural to arrange the columns of all matrices in the order of chords and branches of T. Let us therefore number the edges so that $1, 2, \ldots, e - v + 1$ are chords and $e - v + 2, \ldots, e$ are branches of T, and arrange the columns in this order. Further, let us also number the cut-sets of the fundamental system in such a fashion that a unit matrix results. Let the first cut-set of the fundamental system contain the first branch, namely edge number $e - v + 2$. Let the second cut-set contain edge $e - v + 3$, etc. Then the matrix $\mathbf{Q}_f$ of the fundamental system of cut-sets has the form

$$\mathbf{Q}_f = [\mathbf{Q}_{f_{11}} \quad \mathbf{U}], \tag{4–61}$$

where $\mathbf{U}$ is a unit matrix of order $v - 1$.

For example, in the graph of Fig. 4–3, let us choose the tree consisting of edges a, c, and e. Then the fundamental cut-sets are the cut-sets numbered 2, 4, and 5. The fundamental cut-set matrix is

$$\mathbf{Q}_f = \begin{array}{c} \\ 2 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} b & d & f & a & c & e \end{array} \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \end{array}. \tag{4–62}$$

We may now interrelate the matrices $\mathbf{A}$, $\mathbf{B}_f$, and $\mathbf{Q}_f$ in a number of ways. These relationships stem from Theorems 2–17, 4–14, and 4–16. We state these results as the next theorem and leave the proof as an exercise (Problem 4–11).

THEOREM 4–19. *If the columns of the matrices* $\mathbf{A}$, $\mathbf{B}_f$, *and* $\mathbf{Q}_f$ *are arranged in the order of chords and branches for the tree* T *for which the fundamental systems are formed, and partitioned as*

$$\mathsf{A} = [\mathsf{A}_{11} \quad \mathsf{A}_{12}], \qquad \mathsf{B}_f = [\mathsf{U} \quad \mathsf{B}_{f_{12}}], \qquad \text{and} \qquad \mathsf{Q}_f = [\mathsf{Q}_{f_{11}} \quad \mathsf{U}], \tag{4–63}$$

then we have the following interrelationships:

$$\mathsf{Q}_{f_{11}} = \mathsf{A}_{12}^{-1}\mathsf{A}_{11} = \mathsf{B}'_{f_{12}} \qquad \text{and} \qquad \mathsf{Q}_f = \mathsf{A}_{12}^{-1}\mathsf{A} = [\mathsf{B}'_{f_{12}} \quad \mathsf{U}]. \tag{4–64}$$

Thus, in fact, we can start with any one of the three matrices A, B_f, and Q_f and construct the others.

4–6 Linear vector spaces. We come to one more very useful algebraic concept in the theory of graphs, the last major algebraic concept to be introduced here, namely *linear vector spaces*. The concept of a linear vector space is not really new. It is mainly an extension of the set of familiar 3-dimensional vectors. It serves as a unifying concept allowing us to bring together our knowledge of vectors, matrices, and linear equations. And in this process it adds geometric intuition to many abstract algebraic concepts. The brief discussion of linear vector spaces given here is included for this purpose. Before embarking on a discussion of the general concept of a linear vector space, we begin by selecting from the theory of 3-dimensional vectors those properties which can be extended easily to more than three dimensions, and which characterize the space as *linear*.

For notational convenience, a 3-dimensional vector is denoted as a column matrix, as

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

instead of as

$$a\mathsf{i} + b\mathsf{j} + c\mathsf{k}.$$

Let X, Y, Z, ... stand for 3-dimensional vectors, and let a, b, c, ... stand for real numbers. Then the following properties of 3-dimensional vectors are familiar.

I. (i) If X and Y are vectors, so is $\mathsf{X} + \mathsf{Y}$.
(ii) $\mathsf{X} + \mathsf{Y} = \mathsf{Y} + \mathsf{X}$.
(iii) $\mathsf{X} + (\mathsf{Y} + \mathsf{Z}) = (\mathsf{X} + \mathsf{Y}) + \mathsf{Z}$.
(iv) $\mathbf{0} + \mathsf{X} = \mathsf{X}$ for all X, where $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.
(v) For each X, there is a Y such that $\mathsf{X} + \mathsf{Y} = \mathbf{0}$.

II. (i) $a \cdot \mathsf{X}$ is a vector.
(ii) $a \cdot (b \cdot \mathsf{X}) = (a \cdot b) \cdot \mathsf{X}$.
(iii) $(a + b) \cdot \mathsf{X} = a \cdot \mathsf{X} + b \cdot \mathsf{X}$.
(iv) $a \cdot (\mathsf{X} + \mathsf{Y}) = a \cdot \mathsf{X} + a \cdot \mathsf{Y}$.

III. Every vector $\mathbf{X}$ can be expressed as $\mathbf{X} = a_1\mathbf{D}_1 + a_2\mathbf{D}_2 + a_3\mathbf{D}_3$, where $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ are any three fixed (independent of $\mathbf{X}$) noncoplanar vectors, and a_1, a_2, and a_3 are real numbers.

We have grouped these properties into three sets in a natural fashion. The first set characterizes the set as an abelian group under addition. The second set gives the properties of scalar multiplication. Finally, set III states the 3-dimensional character.

A familiar concept with 3-dimensional vectors is *orthogonality.* Two vectors are defined as orthogonal if their scalar product (dot product) is zero. In matrix notation, the vectors

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \text{and} \qquad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

are *orthogonal* if $\mathbf{X}' \cdot \mathbf{Y} = 0$. The product $\mathbf{X}' \cdot \mathbf{Y}$ is clearly the same as the dot product. Let us exploit this concept a little further by using it to construct the set $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{D}_3$ of III. Choose $\mathbf{D}_1$ and $\mathbf{D}_2$ to be any two vectors which are not collinear. Two vectors $\mathbf{X}$ and $\mathbf{Y}$ are collinear if and only if

$$\mathbf{X} = a\mathbf{Y} \qquad \text{or} \qquad \mathbf{Y} = a\mathbf{X}, \tag{4–65}$$

where a is a scalar. In the language of linear dependence, therefore, two vectors are collinear if and only if they are linearly dependent. Let us first consider the class of vectors that are linearly dependent on the noncollinear vectors $\mathbf{D}_1$ and $\mathbf{D}_2$, that is, the set of vectors that can be expressed as

$$\mathbf{C} = c_1\mathbf{D}_1 + c_2\mathbf{D}_2, \tag{4–66}$$

where c_1 and c_2 are scalars. We recognize these as the vectors that are in the *plane* defined by $\mathbf{D}_1$ and $\mathbf{D}_2$. Conversely, every vector in the plane can be expressed as a linear combination of $\mathbf{D}_1$ and $\mathbf{D}_2$. We can also observe that the set of vectors in this plane satisfies all the conditions laid in I, II, and III, except that there are only two vectors, $\mathbf{D}_1$ and $\mathbf{D}_2$, in III. We say that the plane is a *2-dimensional subspace* of the 3-dimensional space.

Returning to the mainstream, we need a vector $\mathbf{D}_3$ that is not in the plane defined by $\mathbf{D}_1$ and $\mathbf{D}_2$. An interesting vector to choose is a nonzero vector $\mathbf{D}_3$ that is orthogonal to the plane. $\mathbf{D}_3$ can be chosen by the following argument. $\mathbf{D}_3$ must be orthogonal to every vector in the plane. In particular

$$\mathbf{D}_3'\mathbf{D}_1 = 0 \qquad \text{and} \qquad \mathbf{D}_3'\mathbf{D}_2 = 0. \tag{4–67}$$

Conversely, if $\mathbf{D}_3$ satisfies these two equations, it is orthogonal to every vector in the plane. For, if $\mathbf{C}$ is any vector in the plane, we have

$$\mathbf{C} = c_1\mathbf{D}_1 + c_2\mathbf{D}_2, \tag{4–68a}$$

and so

$$\mathbf{D}_3'\mathbf{C} = c_1\mathbf{D}_3'\mathbf{D}_1 + c_2\mathbf{D}_3'\mathbf{D}_2 = 0. \tag{4–68b}$$

Thus to find $\mathbf{D}_3$, two simultaneous equations have to be solved. To make the notation look familiar, we write

$$\mathbf{D}_1 = \begin{bmatrix} d_{11} \\ d_{12} \\ d_{13} \end{bmatrix}, \qquad \mathbf{D}_2 = \begin{bmatrix} d_{21} \\ d_{22} \\ d_{23} \end{bmatrix}, \qquad \mathbf{D}_3 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{4–69}$$

Then the equations to be satisfied by $\mathbf{D}_3$ are, in scalar notation,

$$d_{11}x_1 + d_{12}x_2 + d_{13}x_3 = 0 \tag{4–70a}$$

and

$$d_{21}x_1 + d_{22}x_2 + d_{23}x_3 = 0, \tag{4–70b}$$

or

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{4–71}$$

The nonzero solution to these equations can be obtained by the usual procedure. But this argument gives us a new point of view about homogeneous linear algebraic equations. The solution to a homogeneous system of linear algebraic equations is a *vector orthogonal to the vectors defined by the rows of the coefficient matrix.*

Since $\mathbf{D}_3$ is orthogonal to the plane defined by $\mathbf{D}_1$ and $\mathbf{D}_2$, it is certainly not in the plane. Therefore $\mathbf{D}_3$ *cannot* be expressed as

$$\mathbf{D}_3 = c_1\mathbf{D}_1 + c_2\mathbf{D}_2. \tag{4–72}$$

That is, $\mathbf{D}_3$ is not linearly dependent on $\mathbf{D}_1$ and $\mathbf{D}_2$. Thus *orthogonality implies independence* (but not conversely). Thus the matrix

$$[\mathbf{D}_1 \quad \mathbf{D}_2 \quad \mathbf{D}_3]$$

is nonsingular.

Let us also consider the vectors that are dependent on $\mathbf{D}_3$, the vectors expressible as

$$\mathbf{C} = c\mathbf{D}_3, \tag{4–73}$$

where c is a real number. These are evidently the vectors collinear with $\mathbf{D}_3$. This set of vectors again satisfies all the conditions in I, II, and III, with III containing only one vector. We say that the set of vectors $\mathbf{C} = c\mathbf{D}_3$, where $\mathbf{D}_3$ is nonzero, is a *1-dimensional subspace* of the 3-dimensional space. The two subspaces that we have (the 1-dimensional subspace defined by $\mathbf{D}_3$ and the plane defined by $\mathbf{D}_1$ and $\mathbf{D}_2$) are said to be *orthogonal* to each other, since every vector in one subspace is orthogonal to every vector in the other subspace.

$\mathbf{D}_1$ and $\mathbf{D}_2$ constitute the *basis vectors* of the plane, and $\mathbf{D}_3$ is the *basis vector* of the line. It is clear that if the basis vectors of one subspace are orthogonal to the basis vectors in the other subspace, the two subspaces are orthogonal.

One final observation may be made before leaving the special case. With the vectors as constructed, any vector in the 3-dimensional space can be expressed as a linear combination of $\mathbf{D}_1$, $\mathbf{D}_2$, and $\mathbf{D}_3$ as

$$\mathbf{C} = c_1\mathbf{D}_1 + c_2\mathbf{D}_2 + c_3\mathbf{D}_3, \tag{4–74}$$

where the scalars c_1, c_2, and c_3 are *uniquely* determined by $\mathbf{C}$. This is stated usually as: the 3-dimensional space is the *direct sum* of the two subspaces defined by $\{\mathbf{D}_1, \mathbf{D}_2\}$ and $\{\mathbf{D}_3\}$. Since they are orthogonal by construction, the subspaces are also called *orthogonal complements* of the 3-dimensional space. We now leave the special case and turn to the general concept.

DEFINITION 4–4. *Linear vector space.* Let $\mathcal{G} = \{\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \ldots\}$ be an additive abelian group, and let $\mathfrak{F} = \{a, b, c, \ldots\}$ be a field. Let there be defined a multiplication of elements of $\mathcal{G}$ by elements of $\mathfrak{F}$. Then the set of such products $\mathcal{V} = \{a\mathbf{X}, b\mathbf{Y}, a\mathbf{Z}, \ldots\}$ is an *n-dimensional linear vector space* if for all $a, b, \ldots$ in $\mathfrak{F}$ and all $\mathbf{X}, \mathbf{Y}, \ldots$ in $\mathcal{G}$,

(a) $a \cdot (b \cdot \mathbf{X}) = (a \cdot b) \cdot \mathbf{X}$,
(b) $(a + b) \cdot \mathbf{X} = a \cdot \mathbf{X} + b \cdot \mathbf{X}$,
(c) $a \cdot (\mathbf{X} + \mathbf{Y}) = a \cdot \mathbf{X} + a \cdot \mathbf{Y}$,
(d) every element of $\mathcal{V}$ is expressible as a linear combination of a fixed set of n basis vectors $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_n$ with coefficients from $\mathfrak{F}$ as $\mathbf{X} = \sum_{i=1}^{n} a_i\mathbf{D}_i$, and
(e) $1 \cdot \mathbf{X} = \mathbf{X}$, where 1 is the unit element of $\mathfrak{F}$.

It is possible to extend this definition to the case in which the scalars are chosen from a ring (instead of from a field) by adding the stipulation that the expression in (d) be *unique*. Such a generalization is not needed here. By implication, n in (d) is a finite integer. The vector space is then referred to as a *finite-dimensional* vector space. There are also infinitedimensional vector spaces. An example is the set of all real continuous functions $f(x)$ on the interval $0 \leq x \leq 1$. A basis for this set is $\{\sin 2\pi nx, \cos 2\pi nx\}_{n=0}^{\infty}$, with (d) being the Fourier representation.

In modern algebra, it is customary to study linear vector spaces without introducing any coordinate systems at all. However, for the present purposes, it is more convenient to consider a fixed-basis set of vectors as defining a coordinate system. Then each vector in the space can be described as a column matrix and we will be in the familiar domain of matrix algebra. Thus if $\mathsf{D}_1, \mathsf{D}_2, \ldots, \mathsf{D}_n$ is the fixed basis, and a vector X has the representation

$$\mathsf{X} = \sum_{i=1}^{n} x_i \mathsf{D}_i, \tag{4–75a}$$

we write

$$\mathsf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \tag{4–75b}$$

The vectors $\mathsf{D}_1, \mathsf{D}_2, \ldots, \mathsf{D}_n$ can also be represented in this way, becoming merely the columns of a unit matrix of order n. The most interesting vector spaces associated with a graph are the row spaces of the matrices A and B. In these row spaces, there is a natural way of representing vectors as matrices, and so we avoid some difficulties that might otherwise result from this unconventional procedure. The linear vector space of interest is the set of all subgraphs of a given graph. The field is the field modulo 2, so that addition becomes "ring sum." The fixed-basis set of vectors defining the coordinate system are the "elementary" or "atomic" subgraphs, each consisting of a single edge of the graph.

Every one of the properties of 3-dimensional space that were discussed earlier holds for n-dimensional space as well. (It is more correct to say that we discussed only such properties of 3-dimensional space as are true of n-dimensional space as well.) A few of the more important properties are discussed below for the general space, and the generalizations of the others are left for the reader to complete.

It is neither possible nor desirable to include a complete and rigorous discussion of linear vector spaces here. Therefore, we must be content with stating a few results that seem plausible and include only a semiformal discussion of the others. The following two results are assumed in the later discussion.

Every basis of an n-dimensional vector space contains exactly n elements.

More than n vectors chosen from an n-dimensional linear vector space are linearly dependent.

Linear dependence for vectors is defined as for equations. Vectors $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k$ are *linearly dependent* if there exist scalars $a_1, a_2, \ldots, a_k$ in $\mathfrak{F}$, *not all zero,* such that

$$\sum_{j=1}^{k} a_j \mathbf{Y}_j = \mathbf{0}. \tag{4–76}$$

It follows then that at least one of these vectors can be expressed as a linear combination of the others.

First, let us investigate the condition under which a given set of n vectors $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ is a basis for the space. The coordinate system is assumed to be defined by the basis $\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_n$. Since each $\mathbf{Y}_j$ is a linear combination of the $\mathbf{D}$'s,

$$\mathbf{Y}_j = \sum_{i=1}^{n} a_{ji} \mathbf{D}_i \tag{4–77}$$

or, expressed in matrix notation,

$$\begin{bmatrix} \mathbf{Y}_1' \\ \mathbf{Y}_2' \\ \vdots \\ \mathbf{Y}_n' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1' \\ \mathbf{D}_2' \\ \vdots \\ \mathbf{D}_n' \end{bmatrix}. \tag{4–78}$$

(The transpose notation is used partly to make the a_{kj}'s appear in natural order and partly because the vectors we deal with in graph theory are mostly expressed as row matrices.) This equation can be written more concisely as

$$\mathbf{Y}' = \mathbf{A}\mathbf{D}'. \tag{4–79}$$

Now suppose $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ is a basis. Then the vectors are clearly independent. (Otherwise we do not need all of them.) But this implies that the rows of the matrix $\mathbf{A}$ are independent and so $\mathbf{A}$ is of rank n; hence $\mathbf{A}$ is nonsingular. Conversely, if the vectors $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ are linearly independent and hence the matrix $\mathbf{A}$ is nonsingular, we can invert the equation for the $\mathbf{D}$'s and write

$$\mathbf{D}' = \mathbf{A}^{-1}\mathbf{Y}'. \tag{4–80}$$

Thus the basis vectors $\mathbf{D}_1, \ldots, \mathbf{D}_n$ can be expressed in terms of the $\mathbf{Y}$'s, and hence any vector in the space can be expressed in terms of the $\mathbf{Y}$'s. Thus:

Any set of n linearly independent vectors is a basis for the n-dimensional linear vector space.

By comparison, in three dimensions, the test for finding out whether three vectors A, B, and C are noncoplanar (and hence a basis) is to compute the volume of the parallelepiped that they enclose, which is

$$\pm\mathsf{A} \times \mathsf{B} \cdot \mathsf{C} = \pm \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \cdot \qquad (4\text{–}81)$$

An important theorem on bases, which is very closely related to Whitney's postulate B_2 to be given shortly, is the *Steinitz replacement theorem*, which states:

If $\mathsf{Y}_1, \mathsf{Y}_2, \ldots, \mathsf{Y}_s$ *are linearly independent, and if* $\mathsf{D}_1, \mathsf{D}_2, \ldots, \mathsf{D}_n$ *is a basis of the n-dimensional linear vector space (thus* $s \leq n$*), then there exists a subset of* $\mathsf{D}_1, \mathsf{D}_2, \ldots, \mathsf{D}_n$ *of* $n - s$ *elements which together with* $\mathsf{Y}_1, \mathsf{Y}_2, \ldots, \mathsf{Y}_s$ *constitute a basis.*

This result is the exact analogue of the method of proof used for Theorem 2–12. The analogue of Theorem 2–12 itself is:

Any set of linearly independent vectors can be included in a basis.

These two results are proved in exactly the same fashion as was Theorem 2–12. To prove the first of these two results, for instance, we consider the set of vectors $\{\mathsf{Y}_1, \mathsf{Y}_2, \ldots, \mathsf{Y}_s; \mathsf{D}_1, \mathsf{D}_2, \ldots, \mathsf{D}_n\}$. Every vector of the space can certainly be expressed in terms of these. This set (assuming $s > 0$) must be dependent, since it contains more than n vectors, so that

$$\sum_{i=1}^{s} a_i\mathsf{Y}_i + \sum_{j=1}^{n} b_j\mathsf{D}_j = \mathbf{0}. \qquad (4\text{–}82)$$

Since the Y_i's are linearly independent, at least one b_j is nonzero. Hence this particular D_j can be expressed in terms of the others and so can be deleted from the set. We repeat the procedure, deleting dependent D_j's until the set becomes linearly independent, and therefore a basis. Since every basis contains n elements, $n - s$ of the D_j's must have been included in the final set.

We conclude the discussion of linear vector spaces with the statements of a few useful definitions and theorems, leaving it for the reader to make the necessary extensions from three dimensions. (The theorems are italicized.)

If $\mathfrak{F}$ (in Definition 4–4) is the real field, two vectors X and Y are *orthogonal* (with respect to a given basis) if

$$\mathsf{X}'\mathsf{Y} = 0, \qquad (4\text{–}83)$$

where X and Y are assumed to be expressed as column matrices in terms of the basis.

A vector X *is orthogonal to every vector in a subspace* $\mathcal{V}_s$ *(or, briefly, orthogonal to the subspace) if and only if* X *is orthogonal to the basis vectors of* $\mathcal{V}_s$.

Two subspaces are *orthogonal* if every vector in one subspace is orthogonal to every vector in the other subspace.

Two subspaces are orthogonal if and only if the basis vectors of one subspace are orthogonal to the basis vectors of the other.

Given two subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ of an n-dimensional space $\mathcal{V}$, the *direct sum* $\mathcal{V}_1 \oplus \mathcal{V}_2$ of the two subspaces is the set of all vectors $\mathsf{X}_1 + \mathsf{X}_2$, where X_1 is in $\mathcal{V}_1$ and X_2 is in $\mathcal{V}_2$.

If $\mathcal{V}_1$ *and* $\mathcal{V}_2$ *are orthogonal subspaces of an n-dimensional space* $\mathcal{V}$ *such that the sum of their dimensions is n, then* $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$.

Thus every vector in $\mathcal{V}$ is a linear combination of the basis vectors of $\mathcal{V}_1$ and $\mathcal{V}_2$. It follows that the basis vectors of $\mathcal{V}_1$ together with the basis vectors of $\mathcal{V}_2$ constitute a basis for $\mathcal{V}$; in particular, they constitute a linearly independent set of vectors. In such a case, the subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ are called *orthogonal complements* of the space $\mathcal{V}$.

An abstract algebraic discussion of linear vector spaces may be found in any text on modern algebra (Birkhoff and MacLane [11] or Van der Waerden [187], for instance). A detailed discussion from the point of view adopted here may be found in Hohn [78]. The application to linear graphs has been discussed by Gould [67] and Doyle [46].

4–7 Vector spaces associated with a graph. We now reinterpret the properties of the matrices, cut-sets, and circuits of a linear graph in the language of linear vector spaces. In the first (and more important) interpretation, the vector space $\mathcal{V}_G$ consists of the set of all subgraphs of the given linear graph G. The graph G is assumed to be connected, but the assumption is not necessary. The extension to unconnected graphs is not difficult and so is omitted from the present discussion.

The field $\mathfrak{F}$ over which the subgraphs of G constitute a linear vector space is the field mod 2, and addition of vectors is the ring-sum operation. It can be verified directly from Definition 4–4 that the set of all subgraphs of G constitutes a linear vector space of dimension e. The coordinate system is that defined by the "atomic" elements, each of which consists of a single edge of G. If the edges are numbered $1, 2, \ldots, e$, any subgraph can be expressed as an e-tuple $(g_1 g_2 \cdots g_e)$ of 1's and 0's. In particular, the rows of the matrices A, B, and Q are vectors of the space $\mathcal{V}_G$.

DEFINITION 4–5. *Subspaces $\mathcal{V}_Q$ and $\mathcal{V}_B$.* $\mathcal{V}_Q$ is the set of all linear combinations of the rows of the matrix $\mathbf{A}$ over the field mod 2; $\mathcal{V}_B$ is the set of all linear combinations of the rows of $\mathbf{B}$ over the field mod 2, where $\mathbf{A}$ and $\mathbf{B}$ are the incidence and circuit matrices of the graph G.

THEOREM 4–20. There are 2^{v-1} vectors in $\mathcal{V}_Q$, and each of these is a cut-set or disjoint union of cut-sets.

Proof. Since the rank of $\mathbf{A}$ is $v - 1$, there are $v - 1$ vectors in a basis. Each vector in $\mathcal{V}_Q$ can therefore be expressed as

$$\sum_{i=1}^{v-1} a_i\mathbf{A}_i,$$

where the $\mathbf{A}_i$ are the basis vectors of $\mathcal{V}_Q$. Since there are two choices, 0 or 1, for each a_i, there are 2^{v-1} vectors in $\mathcal{V}_Q$, including the vector $\mathbf{0}$. The rest of the theorem follows from Theorems 4–13 and 4–17 since all vectors in $\mathcal{V}_Q$ are orthogonal to the rows of $\mathbf{B}$. The analogous result for $\mathcal{V}_B$ is stated in the next theorem:

THEOREM 4–21. There are 2^{μ} vectors (including $\mathbf{0}$) in $\mathcal{V}_B$, where μ is the nullity of the graph G, and each of these is a circuit or disjoint union of circuits of G.

Thus every graph G of e edges defines two subspaces $\mathcal{V}_Q$ and $\mathcal{V}_B$ of the linear vector space $\mathcal{V}_G$ of dimension e. In the case of the directed graphs considered in Chapter 5, $\mathcal{V}_Q$ and $\mathcal{V}_B$ become orthogonal complements of $\mathcal{V}_G$. In the field mod 2, however, orthogonality cannot be meaningfully defined. There are cases in which the same vector can belong to both $\mathcal{V}_Q$ and $\mathcal{V}_B$ (that is, a circuit is also a cut-set), as for example in the graph consisting of two parallel edges. We can now interpret 2-isomorphism and duality as follows.

THEOREM 4–22. 2-isomorphic graphs G_1 and G_2 define the same Q- and B-subspaces of an e-dimensional space, with $\mathcal{V}_{Q1} = \mathcal{V}_{Q2}$ and $\mathcal{V}_{B1} = \mathcal{V}_{B2}$. Conversely, any two graphs with $\mathcal{V}_{Q1} = \mathcal{V}_{Q2}$ or $\mathcal{V}_{B1} = \mathcal{V}_{B2}$ are 2-isomorphic.

A one-to-one correspondence between the edges of the two graphs is implicitly assumed. The result follows from Theorems 3–8 and 4–17.

Since the Q-subspaces of 2-isomorphic graphs agree, the basis vectors defined by the incidence matrix of either graph are also basis vectors for the Q-subspace of the other graph. Thus if $\mathbf{A}_1$ and $\mathbf{A}_2$ are incidence matrices of the two graphs, the rows of $\mathbf{A}_1$ are a basis set of vectors for $\mathcal{V}_Q$, and so are the rows of $\mathbf{A}_2$. Hence the next theorem.

THEOREM 4–23. Two graphs G_1 and G_2 are 2-isomorphic if and only if their incidence matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are related by

$$\mathbf{A}_2 = \mathbf{D}\mathbf{A}_1, \tag{4–84}$$

where $\mathbf{D}$ is a nonsingular matrix of integers mod 2.

THEOREM 4–24. If G_1 and G_2 are dual graphs, they define the same subspaces of the e-dimensional space, with

$$\mathcal{V}_{Q1} = \mathcal{V}_{B2} \qquad \text{and} \qquad \mathcal{V}_{Q2} = \mathcal{V}_{B1}. \tag{4–85}$$

Theorem 4–24 follows from Problem 3–15.

COROLLARY 4–24. If G_1 and G_2 are dual graphs, the incidence matrix of either graph is a circuit matrix of the other (with the proper rank, and each row representing a circuit); that is,

$$\mathbf{A}_1 = \mathbf{B}_2 \qquad \text{and} \qquad \mathbf{A}_2 = \mathbf{B}_1. \tag{4–86}$$

The inverse problem of *synthesis* of a graph from a given decomposition of a vector space is much more involved, and consideration of this question is postponed to Chapter 5.

Vector spaces defined by the columns of the matrices $\mathbf{A}$ and $\mathbf{B}$ can also be considered, but these are not particularly interesting; however, they do relate to the work of Whitney [199]. For example, Whitney calls a tree a *basis for the graph* because the columns of $\mathbf{A}$ corresponding to the branches of a tree constitute a basis for the vector space defined by the columns of $\mathbf{A}$. Whitney defines a basis as a set of elements with the properties that

B_1. no proper subset of a basis is a basis and
B_2. if D and D' are bases, and if e is in D, there exists an e' in D' such that $D - e + e'$ is a basis.

We recognize B_2 as essentially the Steinitz replacement theorem. In the space defined by the columns of $\mathbf{B}$, it is the chord sets that constitute bases. This complementary relationship is characterized by Whitney in the statement that the two spaces are *dual matroids*. One of Whitney's theorems on dual matroids is that "Bases in one matroid correspond to basis complements in the dual matroid." We do not pursue this topic further here.

PROBLEMS

4–1. Show that the rank of the vertex matrix of a graph with v vertices and p maximal connected subgraphs is $v - p$. [*Hint:* Arrange the rows and columns of $\mathbf{A}_a$ according to subgraphs, and partition similarly.]

4–2. Prove that if T is a tree of the connected graph G, then the $v - 1$ columns of the incidence matrix $\mathbf{A}$ of G, corresponding to the edges of T, constitute a nonsingular submatrix of $\mathbf{A}$. [*Hint:* Find the incidence matrix $\mathbf{A}_T$ of T, and find its rank.]

4–3. Show that in general the rank of $\mathbf{B}_a$ is $e - v + p$.

4–4. Given the matrix $\mathbf{A}_a$, how can you find out whether the graph is connected?

4–5. Let $\mathbf{A}$ be the vertex matrix of a connected graph G. Let $\mathbf{R} = [r_{ij}]$ be a matrix of elements 0 and 1, such that

(a) $\mathbf{R}$ is of order $(e - v + 1) \times e$, where $\mathbf{A}$ is of order $(v - 1) \times e$, and

(b) $\mathbf{RA}' = \mathbf{0}$.

Show that each row of $\mathbf{R}$ represents a circuit or disjoint union of circuits. If $\mathbf{R}$ contains a unit of matrix, how can the conclusion be strengthened?

4–6. Determine the rank of the matrix $\mathbf{B}_a$ of Fig. 4–1 by the procedure used in the example of Section 4–1. (It should be 4, of course.)

4–7. Find all the trees of Fig. 4–2 and verify that the corresponding submatrices of $\mathbf{A}$ are nonsingular. See if there are any other nonsingular submatrices in $\mathbf{A}$.

4–8. Form another circuit matrix $\mathbf{B}$ (of $e - v + 1$ rows) for Fig. 4–2. Find all the nonsingular submatrices of this matrix, and check against Problem 4–7.

4–9. Find the rank of the matrix $\mathbf{Q}_a$ of the graph of Fig. 4–3 by reducing the matrix, using elementary operations.

4–10. Write out the details of the proof of Theorem 4–14.

4–11. Prove Theorem 4–19.

4–12. Outline the procedure for obtaining a graph, given a basis for the cut-sets of the graph. Apply this procedure and obtain a graph which has the following cut-sets for a basis. Sets of edges: (abd), (cef), (cfg), (dfg). Can you obtain more than one graph with these cut-sets? How are the graphs that you obtain related to each other?

4–13. A graph has the following sets of edges as the basis for the set of all circuits: (abc), (cde), (bdf). Find (in order) (a) a fundamental system of circuits, (b) a fundamental system of cut-sets, and (c) the graph.

4–14. Prove that a graph is determined to within a 2-isomorphism by the set of all trees. Given that the trees of a graph are (ace), (bcd), (abd), (abe), (ade), (bde), (bce), and (acd), find the fundamental system of cut-sets with the aid of Theorem 2–15, and hence find the graph.

4–15. How will parallel edges manifest themselves in the matrix $\mathbf{Q}_a$? And series edges?

4–16. Show that the fundamental system of circuits is a basis for the set of all circuits and disjoint unions of circuits of a graph, and hence show that this set is a linear vector space of dimension $e - v + p$ over the field mod 2.

4–17. If a linear vector space over the field mod 2 is of dimension r, show that the number of bases for the space is

$$(2^r - 2^0) \cdot (2^r - 2^1) \cdot (2^r - 2^2) \cdots (2^r - 2^{r-1}).$$

[*Hint:* There are 2^r vectors in the space. Since any nonzero vector is independent, it can be included in a basis. Hence, the first vector can be chosen in $2^r - 1 = 2^r - 2^0$ ways. Any k vectors define a subspace consisting of 2^k vectors. So any one of the other $2^r - 2^k$ vectors is independent of these k vectors.]

4–18. Show that the trees of the graph satisfy Whitney's postulates B_1 and B_2.

4–19. Repeat Problem 4–18 for chord sets.

4–20. A tree T_1 is *adjacent* to a tree T_2 if T_1 and T_2 contain the same branches, with one exception. Given any two trees T_1 and T_n of a connected graph G, show that there is a sequence of trees $T_1, T_2, \ldots, T_{n-1}, T_n$ such that any two successive trees are adjacent. That is, T_1 can be transformed into T_n by replacing edges of T_1 one at a time, the structure remaining a tree throughout the transformation. [*Hint:* Whitney's postulate B_2.]

4–21. Let 1 and 1′ be any two vertices of a nonseparable graph G. By a *cut-set (1, 1′)* is meant a cut-set which places vertices 1 and 1′ into two different connected parts. Show that the cut-sets (1, 1′) contain a basis for the set of all cut-sets (that is, $\mathcal{V}_Q$). [*Hint:* In the incidence matrix $\mathbf{A}$, let 1′ be the omitted vertex. Derive the matrix $\mathbf{Q}$ by adding rows in such a fashion that row 1 has been added to each of the others. Equivalently, add an edge (1, 1′) to the graph. Now modify $\mathbf{A}$ (with 1′ still omitted) by row additions such that there is a 1 in every row in the column corresponding to edge (1, 1′), and then eliminate disjoint unions of cut-sets.] The dual of this result is stated at the end of Chapter 5.

4–22. Prove that a cut-set is a minimal (nonempty) set of edges such that the columns of $\mathbf{B}$ corresponding to these edges are linearly dependent.

4–23. Prove that a circuit is a minimal set of edges which has an even number of edges in common with each cut-set.

4–24. State and prove the analogue of Lemma 4–10 for cut-sets.

CHAPTER 5

DIRECTED GRAPHS

5–1 The vertex matrix. Most applications require linear graphs in which each edge is oriented, rather than the nonoriented graphs discussed so far. Such graphs are called *directed* graphs instead of the (perhaps) more natural *oriented* graphs because, by long-standing convention, the name *oriented graph* is applied to graphs in which there is at most one directed line segment between any two vertices. Parallel edges are allowed in the present discussion. In some applications, the orientation of the edges is a "true" orientation in the sense that the system represented by the graph exhibits some unilateral property, as for example in signal-flow graphs, information theory, or sequential machines. In electrical network theory on the other hand, the orientation used is a "pseudo"-orientation, used in lieu of an elaborate reference system. The edges, in electrical network theory, are assigned arbitrary orientations.

DEFINITION 5–1. *Oriented edge.* An *oriented edge* is an edge with an orientation assigned by ordering its vertices.

In a diagram, the orientation is shown by an arrowhead on the edge pointing toward the second vertex of the ordered pair. For example, Fig. 5–1 shows an oriented edge (a, b). The edge is said to be oriented *away from* the first vertex and *toward* the second vertex of the ordered pair.

FIG. 5–1. Oriented edge.

DEFINITION 5–2. *Directed graph.* A graph in which every edge has been assigned an orientation is a *directed graph.*

DEFINITION 5–3. *Connected.* A directed graph is *connected* if the corresponding nonoriented graph is connected.

This appears to be an unnatural concept of connectedness for a directed graph. It is more natural to require the orientations of the edges of the path to be all alike. However, Definition 5–3 is used in electrical networks (as the orientation itself is "unnatural" in this case). In the theory of sequential machines, the other concept is useful, and is called *strong connectedness.* This concept is introduced in Chapter 9.

DEFINITION 5–4. *Vertex matrix.* The *vertex matrix* A_a of a directed graph is defined by

$\mathsf{A}_a = [a_{ij}]$ is of order $v \times e$ for a graph with v vertices and e edges,

$a_{ij} = 1$ if edge j is incident at vertex i and is oriented away from vertex i,

$a_{ij} = -1$ if edge j is incident at vertex i and is oriented toward vertex i, and

$a_{ij} = 0$ if edge j is not incident at vertex i.

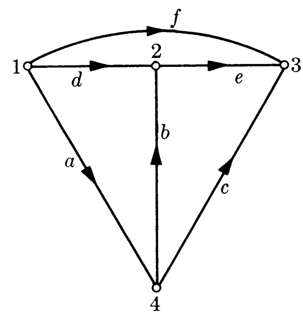


FIG. 5–2. Example for vertex matrix.

This vertex matrix A_a is the coefficient matrix of Kirchhoff's current equations, to be discussed in Chapter 6. Consequently the properties of this matrix are of considerable interest. As an example, the vertex matrix of the directed graph of Fig. 5–2 is

$$\mathsf{A}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 & -1 & -1 \\ -1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}. \tag{5–1}$$

The symbol A_a was used earlier for the incidence matrix of a nonoriented graph and is now used for the vertex matrix of a directed graph. The entries are now treated as real integers. The choice of symbolism here (as in the case of the circuit and cut-set matrices to follow) is guided by the fact that the essential structure of the matrix A_a is the same for directed and nonoriented graphs. There will normally be no confusion. Occasionally (as in Section 5–5) both the matrices (for the directed and nonoriented graphs) may be needed in the same development. In such cases, the superscript (2) will denote the mod 2 (nonoriented) matrix.

The properties of the matrices of a directed graph, and to a large extent the methods of proving them, are identical to those in the nonoriented case. Hence the proofs in this and the following three sections are given in outline form only and often omitted and suggested as problems.

LEMMA 5–1(a). The rank of the vertex matrix A_a of a directed graph of v vertices is at most $v - 1$.

Proof. Each column contains a 1 and a -1. Hence the sum of all the rows is a row of zeros.

LEMMA 5–1(b). For a connected directed graph of v vertices, the sum of any r rows of the incidence matrix, where $r < v$, is nonzero.

THEOREM 5–1. The rank of the incidence matrix $\mathbf{A}_a$ of a connected directed graph of v vertices is $v - 1$.

The proofs of Lemma 5–1(b) and Theorem 5–1 are identical to the proofs in the nonoriented case and so are suggested as a problem. As before, $\mathbf{A}$ denotes a submatrix of $\mathbf{A}_a$ of a connected directed graph obtained by deleting an arbitrary row of $\mathbf{A}_a$.

THEOREM 5–2. If T is a tree of a connected directed graph G, the $v - 1$ columns of the matrix $\mathbf{A}$ corresponding to the branches of T constitute a nonsingular matrix.

Proof. The $v - 1$ columns in question constitute the incidence matrix $\mathbf{A}_T$ of T. Since T is a connected graph of v vertices, the rank of $\mathbf{A}_T$ is $v - 1$, by Theorem 5–1. Since $\mathbf{A}_T$ is of order $(v - 1, v - 1)$, it is nonsingular.

5–2 The circuit matrix. Since the graph is directed, it is natural to consider the circuits and cut-sets also as oriented.

DEFINITION 5–5. *Oriented circuit.* A circuit with an orientation assigned by a cyclic ordering of vertices is an *oriented circuit.*

For example, in Fig. 5–2 the circuit $\{d, e, f\}$ can be oriented as $(1, 2, 3, 1)$ or as $(1, 3, 2, 1)$. Again, one can represent the orientation pictorially by an arrowhead. For the purposes of the following definition, the orientations of an edge of a circuit and the circuit "coincide" if the vertices of the edge appear in the same order both in the ordered-pair representation of the edge and in the ordered-vertex representation of the circuit. Otherwise, they are "opposite." Pictorially, the meaning is obvious.

DEFINITION 5–6. *Circuit matrix* $\mathbf{B}_a$. The *circuit matrix* $\mathbf{B}_a = [b_{ij}]$ with a finite number of rows and e columns is defined by

$b_{ij} = 1$ if edge j is in circuit i and the orientations of the circuit and the edge coincide,

$b_{ij} = -1$ if the edge j is in circuit i and the orientations do not coincide, and

$b_{ij} = 0$ if the edge j is not in circuit i.

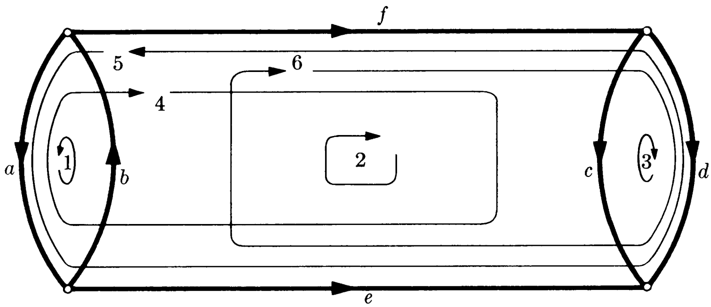


FIG. 5–3. Example for circuit matrix.

In the next chapter, B_a is shown to be the coefficient matrix of Kirchhoff's voltage equations. As an example of a circuit matrix, let us consider the set of all circuits of the graph of Fig. 5–3:

$$\mathsf{B}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 1 & -1 & 1 \end{bmatrix} \end{array}. \tag{5–2}$$

The rank of the circuit matrix is established exactly as in the nonoriented case, by making use of a fundamental system and the orthogonality of A_a and B_a.

DEFINITION 5–7. *Fundamental circuits* (*f-circuits*). The *f-circuits* of a connected directed graph with respect to a tree T are the $e - v + 1$ circuits formed by each chord and the single path in the tree between the vertices of the chord. The f-circuit orientation is chosen to agree with that of the defining chord.

The matrix B_f of these circuits arranged in the order of chords and branches of T again has the form

$$\mathsf{B}_f = [\mathsf{U} \quad \mathsf{B}_{fT}]. \tag{5–3}$$

The unit matrix fixes the rank of B_f as $e - v + 1$. Since B_f is a submatrix of B_a, we have the next theorem.

THEOREM 5–3. The rank of the circuit matrix $\mathbf{B}_a$ of a connected directed graph is at least $e - v + 1$.

THEOREM 5–4. If the columns of the matrices $\mathbf{A}_a$ and $\mathbf{B}_a$ are arranged in the same edge order,

$$\mathbf{A}_a\mathbf{B}'_a = \mathbf{0} \qquad \text{and} \qquad \mathbf{B}_a\mathbf{A}'_a = \mathbf{0}.$$

The proof is left as an interesting problem (Problem 5–7).

THEOREM 5–5. The rank of the circuit matrix $\mathbf{B}_a$ of a connected directed graph is $e - v + 1$.

The proof follows from Sylvester's law of nullity, as in the nonoriented case.

As before, $\mathbf{B}$ denotes a circuit matrix of $e - v + 1$ rows and rank $e - v + 1$ of a connected directed graph.

5–3 Nonsingular submatrices of A and B and formula for $\mathbf{B}_f$.

LEMMA 5–6. There exists a linear relationship among the columns of $\mathbf{A}$ corresponding to the edges of a circuit.

THEOREM 5–6. A square submatrix of $\mathbf{A}$ of order $v - 1$ is nonsingular if and only if the columns of this submatrix correspond to the branches of a tree.

The proofs of Lemma 5–6 and Theorem 5–6 are identical to the proofs of Lemma 4–10 and Theorem 4–10.

THEOREM 5–7. The determinant of a nonsingular submatrix of $\mathbf{A}$ is ± 1.

Proof. This important result has a very simple proof. Consider any nonsingular submatrix of $\mathbf{A}$. Each column of this submatrix has at most two nonzero elements, a $+1$ and a -1. Not every column can have both a $+1$ and a -1, for then the matrix is singular. Also, there is no zero column. Hence there is at least one column with only one nonzero element, a ± 1. Expanding by this column, we find the determinant to be

$$\Delta = \pm 1 \cdot \Delta_{ij}, \tag{5–4}$$

where (i, j) is the position of the nonzero entry. The cofactor Δ_{ij} again has, by the same reasoning, a column with a single nonzero element. Expand Δ_{ij} by this column. Repeated application of the procedure yields

$$\Delta = \pm 1. \tag{5–5}$$

THEOREM 5–8. Let $\mathbf{B}$ be a matrix of $e - v + 1$ rows and rank $e - v + 1$ for a connected graph G. A square submatrix of $\mathbf{B}$ of order $e - v + 1$ is nonsingular if and only if the columns of this submatrix correspond to the set of chords for some tree of G.

The proof of Theorem 5–8 is identical to the proof of Theorem 4–11. Thus, once again:

Nonsingular submatrices of $\mathbf{A}$ *are in one-to-one correspondence with the trees of the graph.*

Nonsingular submatrices of $\mathbf{B}$ *are in one-to-one correspondence with complements of trees of the graph.*

THEOREM 5–9. Let the vertex matrix $\mathbf{A}$ be partitioned in terms of chords and branches for a tree as

$$\mathbf{A} = [\mathbf{A}_{11} \quad \mathbf{A}_{12}]. \tag{5–6}$$

Then the matrix $\mathbf{B}_f$ of f-circuits for the tree corresponding to $\mathbf{A}_{12}$ is given by

$$\mathbf{B}_f = [\mathbf{U} \quad -\mathbf{A}'_{11} \cdot \mathbf{A}_{12}^{-1'}]. \tag{5–7}$$

A closer examination of Theorem 5–9 (which is proved as in Chapter 4) suggests an even deeper result about $\mathbf{A}_{12}$. Instead of starting with a graph, we can start with an arbitrary nonsingular matrix $\mathbf{A}_{12}$ with at most a 1 and a -1 per column. Having obtained the inverse $\mathbf{A}_{12}^{-1}$, we can write any other matrix $\mathbf{A}_{11}$, with at most a 1 and a -1 per column and the same number of rows as $\mathbf{A}_{12}$. Apart from this condition, $\mathbf{A}_{11}$ can be written completely independently of $\mathbf{A}_{12}$. We know that a graph can be drawn for the matrix

$$[\mathbf{A}_{11} \quad \mathbf{A}_{12}].$$

For the graph, by Theorem 5–9, it is necessary that

$$\mathbf{B}'_{f12} = -\mathbf{A}_{12}^{-1}\mathbf{A}_{11}. \tag{5–8}$$

But $\mathbf{B}_f$ has only elements $+1$, -1, and 0. Hence, regardless of the way in which the matrix $\mathbf{A}_{11}$ is written, the elements of $\mathbf{A}_{12}^{-1} \cdot \mathbf{A}_{11}$ must be 1, -1, and 0. Thus, $\mathbf{A}_{12}^{-1}$ must have some very special characteristic. If, for example, it were possible to have

$$\mathbf{A}_{12}^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}, \tag{5–9}$$

we could choose A_{11} as

$$\mathsf{A}_{11} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \tag{5-10}$$

and have

$$\mathsf{A}_{12}^{-1}\mathsf{A}_{11} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}, \tag{5-11}$$

which we know is impossible. Thus we could not have had a 1 and a -1 in the same row. This suggests the following result.

THEOREM 5–10. Let A_{12} be a nonsingular submatrix of the vertex matrix A. Then the nonzero elements of any row of A_{12}^{-1} are either all 1 or all -1.

Proof. We always have

$$\mathsf{A}_{12}^{-1} \cdot \mathsf{A}_{12} = \mathsf{U}. \tag{5-12}$$

Let

$$\mathsf{A}_{12}^{-1} = \mathsf{C}. \tag{5-13}$$

Suppose that there exists row i of C containing both positive and negative elements. Let the columns of C be arranged so that the first r columns of row i are $+1$, the next s columns are -1 and the rest are zeros. Let the rows of A_{12} be arranged similarly, and let A_{12} be partititioned into rows as

$$\mathsf{A}_{12} = \begin{bmatrix} \mathsf{A}_1 \\ \mathsf{A}_2 \\ \vdots \\ \mathsf{A}_{v-1} \end{bmatrix}. \tag{5-14}$$

Then the product of the ith row of C by A_{12} is

$$\sum_{k=1}^{r} \mathsf{A}_k - \sum_{k=r+1}^{r+s} \mathsf{A}_k = \mathsf{R}, \tag{5-15}$$

where R is a row matrix, and is the ith row of the unit matrix U of order $v - 1$. Thus R has 0's everywhere except in the ith column. Now A_{12} has at most a $+1$ and a -1 per column. Also from Eq. (5–15), if any column, except the ith column, has a nonzero entry in the first $r + s$ rows, then this column has both a $+1$ and a -1, both in the first r rows or both in the next s rows. The ith column, however, has only one nonzero entry in the first $r + s$ rows, since the one nonzero entry in $\mathbf{R}$ is 1. Let this nonzero entry be in row j.

Case 1. If $j \leq r$, then the ith column in rows $r + 1$ to $r + s$ contains only zeros. Thus for every 1 in these rows, there is a -1 in the same column, so that

$$\sum_{k=r+1}^{r+s} \mathbf{A}_k = \mathbf{0}; \tag{5-16}$$

or these rows are linearly dependent and $\mathbf{A}_{12}$ is singular, contrary to hypothesis.

Case 2. If $j > r$, we see by a similar argument that

$$\sum_{k=1}^{r} \mathbf{A}_k = \mathbf{0}, \tag{5-17}$$

contrary to the hypothesis that $\mathbf{A}_{12}$ is nonsingular.

Theorem 5-10 is not of any evident theoretical importance. It could be used as a computational aid, for checking computations.

5-4 Cut-sets of directed graphs. A cut-set of a directed graph is merely a cut-set of the corresponding nonoriented graph. However, since the graph is oriented, it is more natural to consider the cut-set also as oriented. To orient the cut-set, we use the interpretation of a cut-set given in Section 2-4, namely, that a cut-set defines a partition of the vertices of the graph.

DEFINITION 5-8. *Cut-set orientation.* A *cut-set* is *oriented* by ordering the sets of vertices α and β of G separated by the cut-set as (α, β) or (β, α). An edge e_i of the cut-set (α, β) has the *same orientation* as the cut-set if e_i is oriented away from its vertex in α and toward its vertex in β.

Pictorially, one can show the orientation by means of an arrow placed near the broken line defining the cut-set. An example of a directed graph

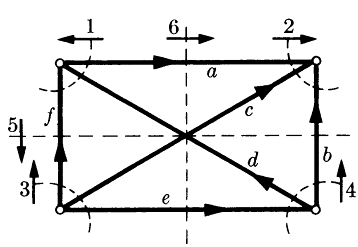


FIG. 5-4. Example for oriented cut-sets.

with oriented cut-sets is shown in Fig. 5–4. In cut-set 1, for example, edges f and d have the same orientation as the cut-set and edge a is oriented opposite to the cut-set.

Once again we can discuss cut-sets most conveniently by means of a cut-set matrix.

DEFINITION 5–9. *Cut-set matrix.* The *cut-set matrix* $\mathbf{Q}_a = [q_{ij}]$ has one row for *each possible* cut-set of the graph and one column for each edge, and is defined by

$q_{ij} = 1$ if edge j is in cut-set i and the orientations agree,
$q_{ij} = -1$ if edge j is in cut-set i and the orientations are opposite, and
$q_{ij} = 0$ if edge j is not in cut-set i.

It should be noted for emphasis that every possible cut-set is in $\mathbf{Q}_a$, except that we do not include cut-sets that are obtained by merely reversing orientations. For the graph of Fig. 5–4, there are seven cut-sets (one of which, consisting of edges a, b, e, and f, is not shown), and so the matrix appears as

$$\mathbf{Q}_a = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \\ 1 & 0 & 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 0 & -1 & -1 \end{bmatrix} \end{array}. \tag{5–18}$$

As in the nonoriented case, we have

THEOREM 5–11. The cut-set matrix $\mathbf{Q}_a$ contains the incidence matrix $\mathbf{A}$ (with some rows possibly multiplied by -1) as a submatrix if G is nonseparable. In any case, the rows of $\mathbf{A}$ are expressible as linear combinations of rows of $\mathbf{Q}_a$.

COROLLARY 5–11. For a connected graph G, of v vertices, the rank of the cut-set matrix $\mathbf{Q}_a$ is at least $(v - 1)$.

To set the upper bound, we once again attack the problem by relating cut-sets to circuits, adding the effect of orientation to Theorem 4–14.

THEOREM 5–12. The number of edges common to a cut-set and a circuit is always even. If a cut-set i has $2k$ edges in common with a circuit j, then k of these edges have the same relative orientation in the cut-set

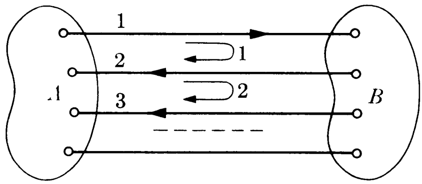


FIG. 5-5. Illustration of Theorem 5-12.

and in the circuit, and the other k have one orientation in the cut-set and the opposite orientation in the circuit.

The proof of Theorem 5-12 consists simply in formalizing the intuitive ideas presented in Fig. 5-5. Since we have already established the evenness of the number of common elements, only the orientation part needs to be proved, and this we omit as evident. (See Problem 5-13.)

THEOREM 5-13. If the columns of the circuit matrix $\mathbf{B}_a$ and the cut-set matrix $\mathbf{Q}_a$ of a directed graph G are arranged in the same edge order,

$$\mathbf{B}_a\mathbf{Q}_a' = \mathbf{0}. \tag{5-19}$$

Theorem 5-13 is merely a restatement of Theorem 5-12 in matrix notation and requires no proof. Combining Theorem 5-13 with Sylvester's law of nullity, we get the result we are after, namely Theorem 5-14.

THEOREM 5-14. The rank of the cut-set matrix $\mathbf{Q}_a$ of a directed graph G of v vertices is $v - 1$.

Since the relationships between the cut-set matrix and the circuit matrix are the same as in nonoriented graphs, we have the same results as in the nonoriented case, except that a few negative signs appear, as we now have the field of real numbers to work with. There is hardly any point in going over the same ground once again in detail, and so we shall blandly state the results, leaving the details as obvious.

DEFINITION 5-10. *f-cut-sets.* If T is a tree of a connected directed graph G, the *fundamental system of cut-sets* with respect to T is the set of $v - 1$ cut-sets in which each cut-set includes only one branch of T. The fundamental cut-set orientation is to agree with the orientation of the defining branch.

Again, if we order the columns as chords and branches and arrange the cut-sets suitably, the matrix of the fundamental system of cut-sets has the form

$$\mathbf{Q}_f = [\mathbf{Q}_{f_{11}} \quad \mathbf{U}]. \tag{5-20}$$

We then have the familiar results stated in Theorems 5–15 and 5–16:

THEOREM 5–15. If Q is a cut-set matrix of $v - 1$ rows and rank $v - 1$ of a connected directed graph G of v vertices, and A is the incidence matrix of G, then

$$\mathsf{Q} = \mathsf{DA}, \tag{5–21}$$

where D is nonsingular.

THEOREM 5–16. If the columns of A, B_f, and Q_f of a directed graph G are arranged in order of chords and branches for the tree T defining the fundamental systems of cut-sets and circuits, so that

$$\mathsf{A} = [\mathsf{A}_{11} \quad \mathsf{A}_{12}], \qquad \mathsf{B}_f = [\mathsf{U} \quad \mathsf{B}_{f_{12}}], \qquad \text{and} \qquad \mathsf{Q}_f = [\mathsf{Q}_{f_{11}} \quad \mathsf{U}], \tag{5–22}$$

we have the relations

$$\mathsf{Q}_f = \mathsf{A}_{12}^{-1}\mathsf{A} \qquad \text{and} \qquad \mathsf{Q}_{f_{11}} = -\mathsf{B}'_{f_{12}} = \mathsf{A}_{12}^{-1}\mathsf{A}_{11}. \tag{5–23}$$

THEOREM 5–17. If Q is a cut-set matrix of a connected directed graph G of v vertices, with $v - 1$ rows and rank $v - 1$, the nonsingular submatrices of Q of order $v - 1$ correspond one-to-one to the trees of G.

THEOREM 5–18. If F is any matrix of elements 1, -1, and 0, such that

$$\mathsf{BF}' = 0, \tag{5–24}$$

where B is the circuit matrix of $e - v + 1$ rows and rank $e - v + 1$ of a connected directed graph G of e edges and v vertices, then each row of F represents a cut-set or disjoint union of cut-sets.

5–5 Existence of graphs for given matrices. In recent years, there has been a considerable amount of work done on the synthesis of conventional electrical networks and combinational switching networks, by algebraic methods. This relatively new field, dating back only to 1954, is known as *topological synthesis*. Some early contributions are due to Okada [124], Traktenbrot [175], and Seshu [152, 153]. An important fundamental problem came into focus immediately [153]. In all of these topological syntheses, one arrives at a matrix of integers mod 2 which, desirably, should be the cut-set or circuit matrix of a graph. It was pointed out as early as 1935 by Whitney [199] that there exist matrices of integers mod 2 that are not cut-set matrices or circuit matrices of graphs. If a given matrix is a cut-set matrix, it can be reduced by elementary row operations to an incidence matrix; that is, a matrix with at most two 1's per column. An example of a matrix that cannot be so reduced is

$$\mathsf{F}_u = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}. \tag{5-25}$$

The complete theoretical solution of the problem was given recently by Tutte [186]. Cederbaum [32] and Gould [66] considered this problem in some detail before the solution was given by Tutte. This section is devoted to a discussion of the general problem. Tutte's general solution depends on many algebraic topological concepts that have not been developed in this book, aside from being extremely long. Therefore we are unable to give the proof of Tutte's general theorem and have to be satisfied with a statement of the result. Cederbaum's contributions are, however, considered in some detail.

Before considering specific results, let us discuss the general problem. Suppose that $\mathsf{F} = [f_{ij}]$ is a matrix of integers mod 2, which is a cut-set matrix (or a circuit matrix) of maximum rank of a linear graph G. Let us now assign arbitrary orientations to the edges of G and consider it as a directed graph G_d. Construct the cut-set matrix (or circuit matrix) F_d of G_d for the same cut-sets (circuits) as in F, retaining the column ordering as well. As we know, F and F_d have many properties in common. First of all, they have nonzero elements in the same positions. They have the same rank. Since nonsingular submatrices of F and F_d of maximum order correspond to trees (chord sets) of G, nonsingular submatrices of F and F_d correspond. Thus F is a special kind of matrix. Its 1's may be replaced

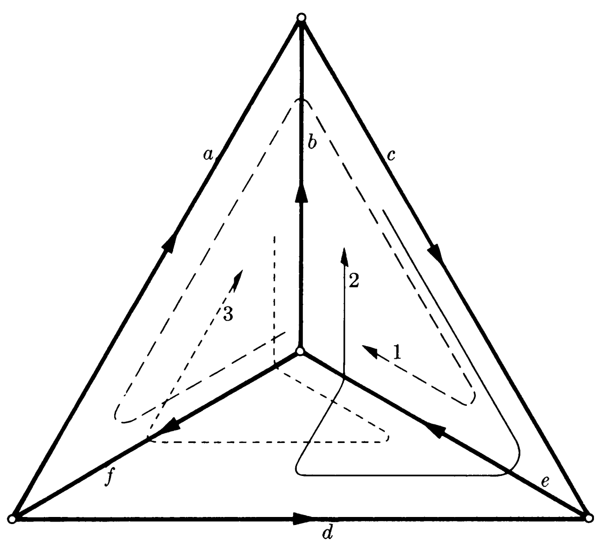


FIG. 5-6. Example illustrating remark.

appropriately by 1's and -1's in such a way that the ranks of submatrices are unaltered. The matrix $\mathbf{F}_u$ given in Eq. (5–25) is an example for which this is not possible. $\mathbf{F}_u$ cannot be replaced by a matrix of 1's, -1's, and 0's such that ranks of submatrices remain invariant.

The discussion above requires a slight qualification, since the field mod 2 and the real field are different. If we start with a set of circuits or cut-sets for a directed graph G_d which are independent over the real field, these circuits or cut-sets may not be independent over the field mod 2 when orientations are removed. Such a circumstance is uncommon, but it can occur. An example is the set of circuits shown in Fig. 5–6. (The orientations of the edges and circuits in this figure may be altered without affecting the argument.) The matrix of these circuits is

$$\mathbf{B}_d = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & -1 & -1 & 0 \end{bmatrix} \end{array}. \tag{5–26}$$

The submatrix consisting of columns a, b, and d, in that order, has a determinant -2 and so is nonsingular. $\mathbf{B}_d$ is therefore of rank 3. If graph and circuits are considered as nondirected, the matrix of these circuits is obtained by removing all the negative signs in $\mathbf{B}_d$:

$$\mathbf{B}_n = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}. \tag{5–27}$$

The mod 2 rank of $\mathbf{B}_n$ is only 2, since the sum of its rows (mod 2) is zero. Similar examples with cut-sets may be found in Cederbaum [28].

Degeneracies of this type can be avoided by requiring that the matrix under discussion contain a unit matrix, and thus correspond to f-circuits or f-cut-sets. This is a convenient assumption to make from other points of view as well, and so is made here. Thus,

$$\mathbf{F} = [\mathbf{U} \quad \mathbf{F}_{12}] \qquad \text{or} \qquad \mathbf{F} = [\mathbf{F}_{11} \quad \mathbf{U}], \tag{5–28}$$

depending on whether $\mathbf{F}$ is desired as a circuit matrix or as a cut-set matrix. This is no loss of generality, since any given matrix can be brought to this form by premultiplication by a suitable matrix. The problem may now be stated in two stages as:

(a) Under what conditions can $\mathbf{F}$ be replaced by a matrix of $+1$, -1, and 0 and keep its rank and nonsingular submatrices invariant?

(b) Under what conditions is $\mathbf{F}$ the circuit matrix or cut-set matrix of a graph?

Before considering these general questions, we discuss the contribution of Cederbaum on E-matrices and show that E-matrices belong to the class of matrices with property (a) above. Later discussion also establishes that if the matrix contains a unit matrix, then the cut-set and circuit matrices of the graph are, in fact, E-matrices. Thus the properties of E-matrices and their characterization are of fundamental importance in the theory of graphs.

Since we wish to discuss cut-set and circuit matrices simultaneously (and at this stage we do not even know whether the matrix is either a cut-set or a circuit matrix) the neutral symbol F is used for mod 2 matrices, with subscript d for real matrices.

DEFINITION 5–11. *E-matrix.* A matrix F_d of real elements is an *E-matrix* if the determinant of *every* square submatrix of F_d is 1, -1, or 0.

THEOREM 5–19. If $\mathsf{F}_d = [f_{ij}]$ is an E-matrix, then $f_{ij} = 1, -1$, or 0.

THEOREM 5–20. If F_d is an E-matrix, then so are

(a) F_d',
(b) matrices obtained by a permutation of rows or columns of F_d,
(c) all submatrices of F_d, and
(d) matrices obtained by multiplying rows or columns of F_d by -1.

Theorems 5–19 and 5–20 are more or less obvious. The next theorem depends on a theorem on matrices due to Jacobi which has not been discussed here, and so its proof is not included. Theorem 5–21 is not required for further development of the subject, and its statement is included purely for completeness.

THEOREM 5–21. If F_d is a (square) nonsingular E-matrix, so is F_d^{-1}.

The most important results on E-matrices are Theorem 5–22, which characterizes E-matrices, and the results that follow from it, namely Theorems 5–25 and 5–26.

THEOREM 5–22. Consider the equation

$$\mathsf{F}_d\mathsf{X} = \mathsf{Y}, \tag{5–29}$$

where F_d is a real matrix of order (n, n) and where $\mathsf{X} = [x_i]$ and $\mathsf{Y} = [y_i]$ are $(n \times 1)$-column matrices (vectors). A necessary and sufficient condition that F_d be an E-matrix is: on assuming *any* $n - 1$ of the $2n$ variables x_i, y_i to be zero, there exists a vector pair X, Y with $\mathsf{X} \not\equiv 0$ satisfying Eq. (5–29) and in which all the remaining $n + 1$ unspecified variables are 1, -1, or 0.

Proof. Let F_d be an E-matrix. Let an arbitrary set of $n - 1$ variables be taken as zero. Let r of these be y's ($0 \leq r \leq n - 1$) and the other $n - r - 1$ be x's. By a suitable permutation of rows and colums of F_d, we may make the vanishing y's occupy the first r positions ($y_1 = y_2 = \cdots = y_r = 0$) and the vanishing x's occupy the last $n - r - 1$ positions ($x_{r+2} = x_{r+3} = \cdots = x_n = 0$), without changing the E-character of F_d. Then the first r equations of the system are

$$\begin{aligned} f_{11}x_1 + f_{12}x_2 + \cdots + f_{1,r+1}x_{r+1} &= 0, \\ f_{21}x_1 + f_{22}x_2 + \cdots + f_{2,r+1}x_{r+1} &= 0, \\ \vdots \\ f_{r1}x_1 + f_{r2}x_2 + \cdots + f_{r,r+1}x_{r+1} &= 0. \end{aligned} \tag{5-30}$$

If every f_{ij} in Eq. (5-30) is zero, take $x_1 = 1$ and $x_2 = x_3 = \cdots = x_{r+1} = 0$. Then from the other equations of the system, $y_i = f_{i1}$ for $r + 1 \leq i \leq n$, which proves the result since $f_{i1} = \pm 1$ or 0. Otherwise, let the coefficient matrix of the system (5-30) be of rank s, $0 < s \leq r$. By permutation of rows and columns, we may assume that the top left submatrix of order s is nonsingular. Then the system can be solved for $x_1, x_2, \ldots, x_s$ in terms of $x_{s+1}, x_{s+2}, \ldots, x_{r+1}$. Take $x_{s+1} = 1$ and $x_{s+2} = x_{s+3} = \cdots = x_{r+1} = 0$. Then, solving the first s equations of Eq. (5-30) by Cramer's rule, we find

$$x_i = -\frac{\Delta_i^{(s)}}{\Delta^{(s)}}, \qquad i = 1, 2, \ldots, s, \tag{5-31}$$

where $\Delta^{(s)}$ is the determinant of the top left submatrix of order s in Eq. (5-30), and $\Delta_i^{(s)}$ is obtained by replacing column i of $\Delta^{(s)}$ by column $(s + 1)$ of the first s rows of Eq. (5-30). Since F_d is an E-matrix, $\Delta_i^{(s)} = 1$, -1, or 0, and $\Delta^{(s)} = 1$ or -1. Thus, $x_i = 1, -1$, or 0 for $i = 1, 2, \ldots, s$. To find y_p, $p = r + 1, r + 2, \ldots, n$, augment the first s equations of (5-30) by the pth equation of the original system (Eq. 5-29), to get

$$\begin{aligned} f_{11}x_1 + f_{12}x_2 + \cdots + f_{1,s+1}x_{s+1} &= 0, \\ f_{21}x_1 + f_{22}x_2 + \cdots + f_{2,s+1}x_{s+1} &= 0, \\ \vdots \\ f_{s1}x_1 + f_{s2}x_2 + \cdots + f_{s,s+1}x_{s+1} &= 0, \\ f_{p1}x_1 + f_{p2}x_2 + \cdots + f_{p,s+1}x_{s+1} &= y_p, \end{aligned} \tag{5-32}$$

since all other x_i are zero. If the coefficient matrix of the system (5-32) is singular, take $y_p = 0$. Otherwise, let $\Delta_p^{(s+1)}$ be the determinant of the coefficient matrix. Solving Eq. (5-32) for x_{s+1}, we find

$$x_{s+1} = \frac{\Delta^{(s)}}{\Delta_p^{(s+1)}}\, y_p, \tag{5-33}$$

where $\Delta^{(s)}$ is the same as the $\Delta^{(s)}$ of Eq. (5-31). Since $x_{s+1} = 1$, this yields

$$y_p = \frac{\Delta_p^{(s+1)}}{\Delta^{(s)}}. \tag{5–34}$$

Since F_d is an E-matrix, $y_p = \pm 1$. Thus, X and Y satisfying the required conditions have been found. The sufficiency of the condition is established by induction on the order of the submatrices of F_d. First choose $x_1 = x_2 = \cdots = x_{i-1} = x_{i+1} = \cdots = x_n = 0$. Since there exists an $\mathsf{X} \neq 0$ satisfying Eq. (5–29), we may assume that $x_i = 1$ (by multiplying all equations by -1 if necessary). Then from Eq. (5–29),

$$y_k = f_{ki}, \qquad k = 1, 2, \ldots, n. \tag{5–35}$$

But $y_k = 1, -1$, or 0. Hence the submatrices of order 1 of F_d have determinants (the elements of F_d) 1, -1, or 0. Next suppose that all square submatrices of order r and less have determinants 1, -1, 0, where $r \geq 2$. Consider a nonsingular submatrix of order $r + 1$. Without loss of generality, let this be the top left submatrix of F_d, and let the determinant of this submatrix be $\Delta^{(r+1)}$. Choose

$$y_1 = y_2 = \cdots = y_r = 0 \quad \text{and} \quad x_{r+2} = x_{r+3} = \cdots = x_n = 0. \tag{5–36}$$

The first $r + 1$ equations of the system (5–29) are then

$$\begin{aligned}
f_{11}x_1 + f_{12}x_2 + \cdots + f_{1,r+1}x_{r+1} &= 0,\\
f_{21}x_1 + f_{22}x_2 + \cdots + f_{2,r+1}x_{r+1} &= 0,\\
&\vdots\\
f_{r1}x_1 + f_{r2}x_2 + \cdots + f_{r,r+1}x_{r+1} &= 0,\\
f_{r+1,1}x_1 + f_{r+1,2}x_2 + \cdots + f_{r+1,r+1}x_{r+1} &= y_{r+1}.
\end{aligned} \tag{5–37}$$

Since there exists an $\mathsf{X} \neq 0$ and some Y satisfying Eq. (5–29) by assumption, and since the determinant of the system (5–37) is nonzero, $y_{r+1} \neq 0$ and so $y_{r+1} = \pm 1$. Solving Eq. (5–37) for the nonzero x's, we have

$$x_s = \frac{\Delta_s^{(r)}}{\Delta^{(r+1)}}\, y_{r+1} = \pm \frac{\Delta_s^{(r)}}{\Delta^{(r+1)}}. \tag{5–38}$$

Since $x_s \neq 0$, so is $\Delta_s^{(r)}$. Hence by the induction hypothesis,

$$\Delta_s^{(r)} = \pm 1.$$

Since $x_s = \pm 1$, by hypothesis, we have from Eq. (5–38) that

$$\Delta^{(r+1)} = \pm 1. \tag{5–39}$$

Hence the result is established by induction, and F_d is an E-matrix.

THEOREM 5–23. Any matrix which contains (as a submatrix) the matrix

$$\mathbf{N} = \begin{bmatrix} \times & 0 & \times & \times \\ \times & \times & 0 & \times \\ \times & \times & \times & 0 \end{bmatrix}, \tag{5–40}$$

where the crosses indicate nonzero entries, cannot be an E-matrix.

Proof. By Theorem 5–20, it suffices to prove that $\mathbf{N}$ is not an E-matrix. By multiplying columns by -1 where necessary, we may assume that all the nonzero entries of the first row are 1. This operation does not change the E-character, by Theorem 5–20. Similarly, by multiplying either row 2 or row 3, or both, by -1 if necessary, we make all the entries of the first column 1's. Now none of the other nonzero entries can be -1 if $\mathbf{N}$ is an E-matrix, as can be seen easily. For instance, if $n_{33} = -1$, then

$$\det \begin{bmatrix} n_{11} & n_{13} \\ n_{31} & n_{33} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2, \tag{5–41}$$

which is impossible. Hence we need only consider the matrix

$$\mathbf{N} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}. \tag{5–42}$$

The determinant of the last three columns of $\mathbf{N}$ is 2, and so $\mathbf{N}$ is not an E-matrix.

All the research workers—Whitney [199], Gould [66], Cederbaum [32], and Tutte [185]—who have attempted to solve the general problem appear to have encountered this particular matrix.

Cederbaum [32] gives several other interesting structural properties of E-matrices, and we refer to his original paper for these. Attention is next directed here to the solution of the general problem. The definitions and major theorems that follow are modifications of Tutte's [185] results. Theorems 5–25 and 5–26 relating E-matrices and regular matrices are due to Seshu [158].

Let

$$\mathbf{F}_d = [f_{ij}] \tag{5–43}$$

denote a matrix of real integers (positive, negative, and zero), and let

$$\mathbf{F} = [f_{ij}^{(2)}] \tag{5–44}$$

denote a matrix of integers mod 2 (that is, $f_{ij}^{(2)} = 1$ or 0). We assume that these matrices are of order (m, n), and rank m ($m \leq n$, naturally). Consider the set of all linear combinations of the rows of F_d with real integral coefficients, and the set of all linear combinations of rows of F with coefficients 1 or 0 (of the mod 2 algebra). Then each of these is a linear vector space of dimension m. (Strictly speaking, in standard mathematical terminology the rows of F_d generate a 0-*module* as the set of integers in a ring, not as the set of integers in a field [187].) We fix the coordinate system by admitting no column operations on these matrices other than permutations.

Definition 5–12. *Elementary vector.* The vector R_1 of either of the spaces under consideration is *elementary* if it is nonzero and there is no other vector R_2 in the space which has nonzero elements only at a proper subset of the positions in which R_1 has nonzero elements.

Definition 5–13. *Primitive vector.* A vector R of the linear vector space defined by the rows of F_d (of real integral elements) is *primitive* if it is elementary and all of its entries are 1, -1, or 0.

Definition 5–14. *Real regular matrix.* The matrix F_d of real integral elements is *regular* if, to every elementary vector in the linear vector space defined by the rows of F_d, there corresponds a primitive vector in the linear vector space, with nonzero entries in the same positions.

For convenience, we refer to a matrix containing a unit matrix,

$$\mathsf{F}_d = [\mathsf{U} \quad \mathsf{F}_{12}], \tag{5–45}$$

as a matrix in *normal form.* For matrices in normal form, Definition 5–14 can be rephrased as follows:

A matrix F_d of real integers in normal form (Eq. 5–45) is a regular matrix if for every linear combination R_1 of the rows of F_d with coefficients 1, -1, and 0, we have that (a) the elements of R_1 are 1, -1, and 0, or (b) there exists another such linear combination R_2 (with coefficients 1, -1, and 0) which has 1 and -1 for nonzero elements and these are at a (not necessarily proper) subset of the positions in which R_1 has nonzero elements.

Let us first establish the reason for considering regular matrices.

Theorem 5–24. The fundamental cut-set matrix Q_f, the incidence matrix A, and the fundamental circuit matrix B_f of a directed graph are all regular matrices.

Proof. Since Q_f and A generate the same space, it suffices to consider Q_f and B_f:

$$\mathsf{Q}_f = [\mathsf{Q}_{11} \quad \mathsf{U}] \qquad \text{and} \qquad \mathsf{B}_f = [\mathsf{U} \quad \mathsf{B}_{12}]. \tag{5–46}$$

Consider any linear combination $\mathbf{Q}_1$ of rows of $\mathbf{Q}_f$ with coefficients 1, -1, and 0. Let $i_1, i_2, \ldots, i_k$ be the cut-sets with coefficients 1 and $j_1, j_2, \ldots, j_p$ be the cut-sets with coefficients -1, so that

$$\mathbf{Q}_1 = \sum_{r=1}^{k} \mathbf{Q}_{i_r} - \sum_{r=1}^{p} \mathbf{Q}_{j_r}. \tag{5–47}$$

Now consider the graph as nonoriented, with fundamental cut-set matrix $\mathbf{Q}_f^{(2)}$, with rows and columns arranged in the same order as in $\mathbf{Q}_f$. Let

$$\mathbf{Q}_1^{(2)} = \sum_{r=1}^{k} \mathbf{Q}_{i_r}^{(2)} + \sum_{r=1}^{p} \mathbf{Q}_{j_r}, \tag{5–48}$$

where the sums are sums mod 2. By Theorem 4–17, $\mathbf{Q}_1^{(2)}$ represents a cut-set or disjoint union of cut-sets. Now we observe that wherever $\mathbf{Q}_1$ has zeros, $\mathbf{Q}_1^{(2)}$ also has zeros. [$\mathbf{Q}_1^{(2)}$ may have more zeros than $\mathbf{Q}_1$.] This result is immediate since each of the vectors on the right side of Eq. (5–47) has elements 1, -1, or 0, and so $\mathbf{Q}_1$ can have a zero only if an even number of nonzero entries has been added. Under these conditions, $\mathbf{Q}_1^{(2)}$ also has a zero since the sum of an even number of 1's is zero in mod 2 algebra. Also, since $\mathbf{Q}_f^{(2)}$ has independent rows, $\mathbf{Q}_1^{(2)} \neq \mathbf{0}$ if $\mathbf{Q}_1 \neq \mathbf{0}$. Hence we can always find another cut-set $\mathbf{Q}_2^{(2)}$ which is contained in the disjoint union of cut-sets $\mathbf{Q}_1^{(2)}$. If $\mathbf{Q}_1^{(2)}$ is a cut-set itself, take $\mathbf{Q}_2^{(2)} = \mathbf{Q}_1^{(2)}$. The directed cut-set $\mathbf{Q}_2$ corresponding to $\mathbf{Q}_2^{(2)}$ then has entries 1, -1, and 0 and has zeros wherever $\mathbf{Q}_1$ has zeros. $\mathbf{Q}_f$ is therefore regular. A similar proof shows that $\mathbf{B}_f$ is also regular.

Theorem 5–24 is actually true of most of the cut-set matrices and circuit matrices, and is not restricted to $\mathbf{Q}_f$ and $\mathbf{B}_f$. The restriction to $\mathbf{Q}_f$ and $\mathbf{B}_f$ is made merely to ensure that the corresponding mod 2 matrices have linearly independent rows so that $\mathbf{Q}_1^{(2)}$ cannot become the empty cut-set. The exceptional cases that have to be excluded are the "pathological" cases similar to the matrix $\mathbf{B}_d$ of Eq. (5–26).

Returning to general regular matrices (which may or may not be cut-set or circuit matrices of graphs), we first prove that every regular matrix in normal form is an E-matrix. The assumption "normal form" is not a serious restriction, as shown by Lemma 5–25(a).

LEMMA 5–25(a). To every regular matrix $\mathbf{F}_R$ there corresponds a regular matrix $\mathbf{F}_d$ in normal form which generates the same linear vector space with at most a permutation of coordinates (corresponding to a permutation of columns of $\mathbf{F}_R$).

Proof. Let $\mathbf{F}_R$ be of order (m, n) and rank m. We may assume that the first m columns of $\mathbf{F}_R$ constitute a nonsingular submatrix (by permutation of columns if necessary):

$$\mathsf{F}_R = [\mathsf{F}_{11} \quad \mathsf{F}_{12}]. \tag{5–49}$$

Now F_{11} can be made diagonal by using only the following row operations (see Problem 5–31):

(a) multiplication of a row by a nonzero integer,
(b) addition to one row of an integral multiple of another row, and
(c) permutation of rows.

Then F_R becomes transformed into the matrix

$$\mathsf{F}_{R1} = [\mathsf{D} \quad \mathsf{F}_{R2}], \tag{5–50}$$

where D is a diagonal matrix with nonzero (integral) diagonal entries. Since each row of F_{R1} is a linear combination of rows of F_R with integral coefficients, it belongs to the linear vector space defined by the rows of the regular matrix F_R. Hence the rows of F_{R1} can be replaced by primitive vectors preserving all the zeros in F_{R1} by the definition of regular matrix. None of the primitive vectors can have zeros in all the first m positions since F_{11} is nonsingular and therefore has independent rows. Thus D is replaced by a unit matrix. (If any of the diagonal entries is -1, multiply this vector by -1, which does not change its primitive character.) Thus the matrix becomes

$$\mathsf{F}_d = [\mathsf{U} \quad \mathsf{F}_{d2}]. \tag{5–51}$$

F_d clearly generates the same linear vector space and hence is also regular.

LEMMA 5–25(b). The rows of the regular matrix F_d in normal form are primitive vectors.

The proof is omitted and left as a problem (Problem 5–28).

LEMMA 5–25(c). Any subset of rows of a regular matrix F_d in normal form constitutes another regular matrix.

Proof. Let $r_1, r_2, \ldots, r_k$ be the rows in question. Consider any linear combination of $r_1, r_2, \ldots, r_k$ with coefficients 1, -1, and 0. In the linear vector space defined by F_d, there is a corresponding primitive vector. Because of the unit matrix, this primitive vector cannot depend on any other row r_p, for then it has a ± 1 in column r_p, where the first vector has only 0. Hence the lemma.

THEOREM 5–25. Every regular matrix F_d in normal form is an E-matrix.

Proof. By Lemma 5–25(c) (where the subset is a single row), the square submatrices of order 1 of F_d have determinants 1, -1, and 0 (these are the elements of F_d). Suppose that all square submatrices of orders up to (and including) r, $r \geq 2$, have determinants 1, -1, and 0. Consider a nonsingular submatrix of order $r + 1$, which for notational convenience we write as

$$\mathsf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} & p_{1,r+1} \\ p_{21} & p_{22} & \cdots & p_{2r} & p_{2,r+1} \\ \vdots & & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} & p_{r,r+1} \\ p_{r+1,1} & p_{r+1,2} & \cdots & p_{r+1,r} & p_{r+1,r+1} \end{bmatrix}. \tag{5–52}$$

By permuting columns, we may consider the leading $(r \times r)$-submatrix of P to be nonsingular. Consider the row vector

$$\mathsf{R} = [\Delta_{1,r+1} \quad \Delta_{2,r+1} \quad \cdots \quad \Delta_{rr} \quad \Delta_{r+1,r+1}], \tag{5–53}$$

where the elements are cofactors of the elements of the last columns of P. Clearly R is the last row of P^{-1} multiplied by $\Delta = \det \mathsf{P}$. By the induction hypothesis, $\Delta_{i,r+1} = 1, -1$, or 0, and $\Delta_{r+1,r+1} \neq 0$. Replace the last row of P by RP. Clearly $|\det \mathsf{P}|$ is unaltered by this process since the new matrix has a determinant equal to $\Delta \cdot \Delta_{r+1,r+1}$. The matrix now becomes

$$\mathsf{P}_1 = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} & p_{1,r+1} \\ p_{21} & p_{22} & \cdots & p_{2r} & p_{2,r+1} \\ \vdots & & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} & p_{r,r+1} \\ 0 & 0 & \cdots & 0 & \Delta \end{bmatrix}. \tag{5–54}$$

The last row of P_1 is a linear combination of the rows of P with coefficients $1, -1$, and 0. Consider the same linear combination of the rows of F_d corresponding to rows of P. If $\Delta \neq \pm 1$, there exists a primitive vector with no additional nonzero entries, which by Lemma 5–25(c) can be expressed as a linear combination of the same rows, with $1, -1$, and 0 as coefficients. This primitive vector cannot have a 0 in the position occupied by Δ since P is nonsingular. Use the coefficients of this linear combination (which gives the primitive vector) on P; that is, replace the last row of P by the linear combination of rows of P as defined by the primitive vector. Since all the coefficients are $1, -1$, or 0, the determinant changes at most in sign. The last row now becomes $[0 \; 0 \; 0 \; \cdots \; 0 \; \pm 1]$. Expanding the new determinant by the last row, we arrive at

$$\det \mathsf{P} = \pm\Delta_{r+1,r+1} = \pm 1. \tag{5–55}$$

Corollary 5–25. *Every submatrix of a regular matrix in normal form is itself regular (but may not be in normal form) and is an E-matrix.*

The corollary follows by the method of proof used in Theorem 5–25 and from Theorem 5–20.

The converse of Theorem 5–25 is also true, as stated by the next theorem.

THEOREM 5–26. Every E-matrix with linearly independent rows is regular.

Proof. Let F_d be an E-matrix of order (m, n), with $m \leq n$. We may assume that F_d contains no zero-row. Let any linear combination of the rows of F_d be represented as $\mathsf{X}'\mathsf{F}_d$, where X is an m-vector of elements 1, -1, and 0. Let

$$\mathsf{F}_d'\mathsf{X} = \mathsf{Y},$$
$$\mathsf{X} = [x_i], \qquad \mathsf{Y} = [y_i]. \tag{5–56}$$

If Y does not contain elements 1, -1, and 0, we have to prove that there exists another $\mathsf{X}_1 \neq \mathbf{0}$, with elements 1, -1, and 0, such that

$$\mathsf{F}_d'\mathsf{X}_1 = \mathsf{Y}_1 \neq \mathbf{0},$$
$$\mathsf{X}_1 = [x_i^1], \qquad \mathsf{Y}_1 = [y_i^1], \tag{5–57}$$

has elements 1, -1, and 0, and has zeros wherever Y has zeros. If Y contains no zeros at all, let $x_1^1 = 1$ and $x_2^1 = x_3^1 = \cdots = x_m^1 = 0$. Then Y_1 becomes the first column of F_d which has elements 1, -1, and 0. In the general case, we slightly modify the method of proof used for Theorem 5–22. Let Y contain r zeros. By permuting rows of F_d, we may consider the first r y's to be zero. Then the first r equations of the system (5–56) become

$$\begin{aligned} f_{11}x_1 + f_{21}x_2 + \cdots + f_{m1}x_m &= 0, \\ f_{12}x_1 + f_{22}x_2 + \cdots + f_{m2}x_m &= 0, \\ &\vdots \\ f_{1r}x_1 + f_{2r}x_2 + \cdots + f_{mr}x_m &= 0. \end{aligned} \tag{5–58}$$

These equations *are satisfied* by the given linear combination X. Therefore the columns of the coefficient matrix of Eq. (5–58) are linearly dependent. Therefore the rank of the coefficient matrix is $s < m$. Again by permutation of rows and columns, let the leading square submatrix of order s be nonsingular. Set $x_{s+1}^1 = 1$ and $x_{s+2}^1 = x_{s+3}^1 = \cdots = x_m^1 = 0$. The existence of x_{s+1}^1 is ensured because $s < m$. This ensures that $\mathsf{X}_1 \neq \mathbf{0}$. The new vector X_1 is found exactly as in Theorem 5–22 by solving the first s equations of (5–58) as

$$x_i^1 = -\frac{\Delta_i^{(s)}}{\Delta^{(s)}}, \qquad i = 1, 2, \ldots, s, \tag{5–59}$$

as in Eq. (5–31). The computation of y_p, $p = r + 1, r + 2, \ldots, n$, is performed exactly as in Theorem 5–22. However, we must establish that $\mathsf{Y}_1 \neq \mathbf{0}$. But this follows because F_d has linearly independent rows by hypothesis, and so F_d' has linearly independent columns.

Finally, we turn to matrices of integers mod 2 and the statements of the answers to the questions raised at the beginning of this section.

DEFINITION 5–15. *Binary matrix.* A matrix of integers 1, −1, and 0 is *binary* if the replacement of −1's by 1's leaves the ranks of submatrices unaltered, where the rank of the derived matrix is with respect to modulo 2 algebra.

THEOREM 5–27. Every E-matrix is binary.

Proof. The theorem follows from the general expansion formula for a determinant [78]. If $\mathsf{P} = [p_{ij}]_{n,n}$, then

$$\det \mathsf{P} = \sum_{(j_1 j_2 \cdots j_n)} \epsilon_{j_1 j_2 \cdots j_n} p_{1j_1} p_{2j_2} \cdots p_{nj_n}, \tag{5–60}$$

where the sum is over all permutations $(j_1 j_2 \cdots j_n)$ of $(1, 2, \ldots, n)$, and $\epsilon_{j_1 j_2 \cdots j_n}$ is $+1$ or -1 depending on whether the permutation is even or odd. For mod 2 algebra, the coefficient is always 1. The details are left as a problem (Problem 5–30).

COROLLARY 5–27. Every regular matrix in normal form is binary.

DEFINITION 5–16. *Regular matrix mod* 2. A matrix of integers mod 2 is *regular* if the replacement of a suitable set of 1's by −1's makes it regular.

Referring back to the remarks at the beginning of this section, the first problem is to characterize regular matrices mod 2. If the given matrix F contains the matrix N of Eq. (5–40), with the crosses replaced by 1's, or its transpose, it is clearly not regular. However if F does not contain N or N', one cannot immediately conclude that F is regular. For, some linear combinations of rows of F may produce N or N'. It may appear at first sight that all linear combinations of rows of F must be tried to examine this possibility. However, they are not all necessary. It is sufficient to premultiply F by the inverses of its nonsingular submatrices. Each of the resulting matrices is referred to as a *normal form* of F, since these contain unit matrices. The answers to the fundamental questions are given next.

THEOREM 5–28. A matrix F of integers mod 2 is regular if and only if no normal form of F contains either of the matrices

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

as a submatrix.

THEOREM 5-29. A matrix **F** of integers mod 2 is the cut-set matrix (circuit matrix) of a graph if and only if it is regular and no normal form of **F** contains a circuit matrix (cut-set matrix) of either of the two basic nonplanar graphs of Kuratowski (Fig. 3-9).

The necessity parts of these two theorems have been established. For Theorem 5-28, the necessity is given by Theorem 5-23. The necessity part of Theorem 5-29 has been established in Theorems 5-24 and 3-14. The sufficiency occupies 57 pages of *Transactions of the American Mathematical Society* [185, 186] and so cannot be given here. The theorems, however, are of fundamental importance.

5-6 Summary of important results on graphs. Having completed the general discussion of graphs, and before proceeding to their applications, it is useful to collect the most important results that have been obtained. In a logical development, the important results are indistinguishable from the many auxiliary results that are needed for the proofs of the main results and the minor results included for the sake of completeness. Also, many of these have been listed as problems and are likely to have been overlooked. Let us therefore spotlight the significant results. To keep the summary compact, we shall use the same termininology and notation used earlier, without further explanation. To avoid verbosity, we assume that all graphs are connected and omit words like *non-empty*. A theorem or problem number indicates where the proof of the result may be found in the text. Where significant, we include the name of the person who first proved the result, and the year of that proof. Because of the classification, there are many repetitions.

I. *Circuit:*

(a) Connected graph with every vertex of degree 2. (Veblen, 1911; Theorem 1-1.)
(b) Minimal set of edges not contained in any tree. (Whitney, 1935; Problem 2-22.)
(c) Minimal dependent set of columns of **A** or **Q**. (Whitney, 1935; Lemma 4-10.)
(d) Minimal set with at least one chord of each tree. (Whitney, 1935; Problem 2-22.)
(e) Minimal set with an even number of edges from each cut-set. (Problem 4-23.)

II. *Cut-set:*

(a) Dual of a circuit.
(b) Minimal set of edges not contained in any tree complement. (Theorem 2-15.)

(c) Minimal dependent set of columns of $\mathbf{B}$. (Problem 4–22.)
(d) Minimal set with at least one branch of every tree. (Theorem 2–15.)
(e) Minimal set with an even number of edges from each circuit. (Theorem 2–18.)

III. *Trees:*

(a) G_s contained in a tree if and only if G_s contains no circuits. (Theorem 2–12.)
(b) G_s contained in a tree complement if and only if G_s contains no cut-sets of G. (Problem 2–20.)
(c) Maximal independent set of columns of $\mathbf{A}$. (Theorems 4–10 and 5–6.)
(d) Complement of maximal independent set of columns of $\mathbf{B}$. (Theorems 4–11 and 5–8.)

IV. *Duality.* (*Hypothesis: G_1 is a dual of G_2*):

(a) G_2 is a dual of G_1. (Theorem 3–12.)
(b) $R_1 = N_2$ and $R_2 = N_1$. (Theorem 3–11.)
(c) $\mathbf{A}_1 = \mathbf{B}_2$ and $\mathbf{A}_2 = \mathbf{B}_1$. (Corollary 4–25.)
(d) $\mathbf{B}_1 = \mathbf{Q}_2$ and $\mathbf{B}_2 = \mathbf{Q}_1$. (Theorem 4–25.)
(e) If G_3 is a dual of G_2, then G_3 is 2-isomorphic to G_1. (Theorem 3–17.)

V. *Determination of a graph to within a 2-isomorphism:*

(a) Matrix $\mathbf{A}$. (Isomorphism.)
(b) Matrix $\mathbf{B}$. (Whitney, 1933; Theorem 4–19.)
(c) Matrix $\mathbf{Q}$. (Theorem 4–19.)
(d) Set of all trees. (Whitney, 1935; Problem 4–14.)
(e) Set of all chord sets. (Whitney, 1935; Problem 4–14.)
(f) Set of all cut-sets separating any two vertices of a nonseparable graph. (Problem 4–21.)
(g) Set of all paths between any two vertices of a nonseparable graph. (Ashenhurst, 1954; Theorem 9–5.)

VI. $\mathbf{A}$ *and* $\mathbf{Q}$:

(a) Rank of $v - 1$. (Kirchhoff, 1847; Theorems 4–4 and 5–1.)
(b) Nonsingular submatrices $\overset{1:1}{\leftrightarrow}$ trees. (Theorems 4–10 and 5–6.)
(c) For directed graphs ($\mathbf{Q}$ assumed $\mathbf{Q}_f$), determinant of a nonsingular submatrix is 1 or -1. (Theorems 5–7 and 5–25.)
(d) Regular matrices. (Theorem 5–24.)
(e) $\mathbf{Q} = \mathbf{TA}$, $\mathbf{T}$ nonsingular. (Theorems 4–19 and 5–15.)
(f) $\mathbf{Q}_f = \mathbf{A}_{12}^{-1}\mathbf{A}$. (Theorems 4–19 and 5–16.)
(g) If G_1 and G_2 are 2-isomorphic, then $\mathbf{A}_1 = \mathbf{Q}_2$, $\mathbf{A}_2 = \mathbf{Q}_1$, and $\mathbf{A}_2 = \mathbf{TA}_1$, $\mathbf{T}$ nonsingular. (Theorem 4–24.)

VII. $\mathbf{B}$:

(a) Rank of $e - v + 1$. (Kirchhoff, 1847; Theorems 4-9 and 5-5.)
(b) Nonsingular submatrices $\overset{1:1}{\leftrightarrow}$ tree complements. (Theorems 4-11 and 5-8.)
(c) For $\mathbf{B}_f$ of directed graphs and "windows" of planar directed graphs, determinant of nonsingular submatrix is 1 or -1. (Theorem 5-24 and Problem 5-26.)
(d) Regular matrix ($\mathbf{B}_f$). (Theorem 5-24.)
(e) $\mathbf{AB}' = \mathbf{0}$, $\mathbf{QB}' = \mathbf{0}$. (Theorems 4-6 and 4-14.)
(f) $\mathbf{B}_f = [\mathbf{U} \;\; \mathbf{A}'_{11} \cdot \mathbf{A}_{12}^{-1'}]$ nonoriented. (Theorem 4-19.)
$\mathbf{B}_f = [\mathbf{U} \;\; -\mathbf{A}'_{11} \cdot \mathbf{A}_{12}^{-1'}]$ directed. (Theorem 5-16.)
(g) G_1 and G_2 2-isomorphic if and only if $\mathbf{B}_1 = \mathbf{B}_2$.

VIII. *Vector spaces of nonoriented graphs* (dim $\mathcal{V}$ *stands for dimension of* $\mathcal{V}$):

(a) dim $\mathcal{V}_Q = v - 1$ and dim $\mathcal{V}_B = e - v + 1$.
(d) $\mathcal{V}_Q \leftrightarrow S_1 = \{$cut-sets and edge-disjoint unions of cut-sets$\}$.
(e) $\mathcal{V}_B \leftrightarrow S_2 = \{$circuits and edge-disjoint unions of circuits$\}$.
(f) G_1 and G_2 are 2-isomorphic if and only if $\mathcal{V}_{Q1} = \mathcal{V}_{Q2}$ and $\mathcal{V}_{B1} = \mathcal{V}_{B2}$.
(g) G_1 and G_2 are duals if and only if $\mathcal{V}_{Q1} = \mathcal{V}_{B2}$ and $\mathcal{V}_{B1} = \mathcal{V}_{Q2}$.
(h) $\mathcal{V}_Q$ contains 2^{v-1} vectors, and $\mathcal{V}_B$ contains 2^{e-v+1} vectors.

IX. *Vector spaces of directed graphs:*

(a), (f), (g) Same as for nonoriented graphs.
(b) $\mathcal{V}_Q$ orthogonal to $\mathcal{V}_B$.
(c) $\mathcal{V}_Q \oplus \mathcal{V}_B = \mathcal{V}_G$.
(d) $\{$Vectors in $\mathcal{V}_Q$ with coordinates 1, -1, and 0$\}$ $\leftrightarrow$ $S_1 = \{$cut-sets and edge-disjoint unions of cut-sets$\}$.
(e) $\{$Vectors in $\mathcal{V}_B$ with coordinates 1, -1, and 0$\}$ $\leftrightarrow$ $S_2 = \{$circuits and edge-disjoint unions of circuits$\}$.
(h) Both $\mathcal{V}_Q$ and $\mathcal{V}_B$ contain an infinite number of vectors.

X. *Matrices mod 2 and graphs:*

(a) Given matrix $\mathbf{F}$ of integers mod 2, $\mathbf{F}$ can be replaced by a matrix of 1, -1, and 0, keeping ranks of all submatrices invariant if and only if no normal form of $\mathbf{F}$ contains either

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

as a submatrix. (Theorem 5-28.)

(b) Given **F** satisfying conditions in (a), **F** is the cut-set matrix of a graph if and only if no normal form of **F** contains the circuit matrix of a nonplanar graph. (Theorem 5–29.)

(c) Given **F** satisfying conditions in (a), **F** is the circuit matrix of a graph if and only if no normal form of **F** contains the cut-set matrix of a nonplanar graph. (Theorem 5–29.)

XI. *Planar graphs:*

(a) A graph is planar if and only if it does not contain a Kuratowski graph. (Theorem 3–16.)

(b) A graph is planar if and only if it has a dual. (Theorem 3–14.)

Problems

5–1. Prove Theorem 5–1.

5–2. Orient the graph of Fig. 1–10. Construct the matrix **A**. Choose a tree and show, by elementary operations, that the submatrix of **A** corresponding to this tree is nonsingular.

5–3. It is always possible to arrange the rows and columns of **A** such that

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

and $\mathbf{A}_{11}$ is square. However, the order of $\mathbf{A}_{11}$ changes, depending on various factors. Examine these factors and state the criterion for making $\mathbf{A}_{11}$ as large as possible, with no row of $\mathbf{A}_{11}$ being zero.

5–4. Prove that given any subgraph containing no circuits, the incidence matrix **A** of the graph can be arranged as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix},$$

where the columns of $\mathbf{A}_{11}$ correspond to the edges of the subgraph and $\mathbf{A}_{11}$ is nonsingular.

5–5. Find the rank of the matrix $\mathbf{B}_a$ of Fig. 5–3 by reducing the matrix, using elementary operations.

5–6. Construct the matrix for a fundamental system of circuits of Figs. 3–9(a) and (b).

5–7. Prove Theorem 5–4. [*Hint:* If a vertex is in a circuit, consider all possible orientations.]

5–8. Show that any set of $e - v + 1$ circuits of a connected graph G such that the matrix of these circuits has a rank $e - v + 1$ includes every circuit element. Thus show that the sophomore law "Be sure to include every network element in at least one loop" is superfluous. [*Hint:* Follow proof of the second part of Theorem 4–11.]

5–9. Show that any incidence matrix **A** "covers" the graph as indicated for the circuit matrix in Problem 5–8.

5–10. The graph of Fig. 3–9(a) can be "covered" (all edges included) with two, three, four, five, or six circuits in such a way that each circuit contains (at least) one edge which is in no other circuit. Show this. Can the graph be so covered with seven circuits? What general conclusion can be drawn from this example?

5–11. State the analogue of Problem 4–5, in directed graphs, and prove it.

5–12. Construct a fundamental system of cut-sets for Fig. 5–4 and the corresponding matrix.

5–13. Complete the formal proof of Theorem 5–12.

5–14. Assign edge orientations to Fig. 1–8(c) in any arbitrary fashion. Then, (a) establish $\mathbf{A}_a$, (b) establish $\mathbf{Q}_a$, and (c) show that $\mathbf{A}_a$ is contained in $\mathbf{Q}_a$.

5–15. Repeat Problem 5–2 by using $\mathbf{Q}_f$.

5–16. Repeat Problem 5–9 for any cut-set matrix $\mathbf{Q}_f$.

5–17. Prove that any row in any matrix $\mathbf{Q}_f$ is a linear combination of some set of rows of $\mathbf{A}_a$.

5–18. Take any cut-set matrix of Fig. 5–4 of maximal rank (3) and find the transformation $\mathbf{D}$ of Theorem 5–15, relating the cut-set matrix to the incidence matrix.

5–19. Prove Theorem 5–17.

5–20. Prove Theorem 5–18.

5–21. Orient Fig. 2–2. Establish $\mathbf{A}$, $\mathbf{Q}_f$, and $\mathbf{B}_f$ corresponding to some tree. Show by calculation that $\mathbf{A}\mathbf{B}_f' = \mathbf{0}$ and $\mathbf{Q}_f\mathbf{B}_f' = \mathbf{0}$.

5–22. Orient Fig. 2–2. Determine $\mathbf{Q}_f$ for four different trees. Show that in each of the four matrices, the submatrices corresponding to the four trees are nonsingular.

5–23. Show that $\mathbf{Q}_f$ and $\mathbf{B}_f$ can be calculated from any given $\mathbf{A}$.

5–24. Orient the graph of Fig. 3–9(a). Calculate the number of trees in this graph from the formula

$$(\text{number of trees}) = \det \mathbf{A}\mathbf{A}'.$$

(This formula is established in Chapter 7.)

(a) There are evidently many trees and therefore many matrices $\mathbf{B}_f$ and $\mathbf{Q}_f$. How many matrices $\mathbf{A}$ exist?
(b) Try to find the number of rows in $\mathbf{B}_a$.
(c) Repeat part (b) for cut-set matrix $\mathbf{Q}_a$.

5–25. Let G be a planar directed graph. Describe the procedure for constructing the geometrical dual of G (including the orientations of edges) which should be performed so that the incidence matrix of either graph is the circuit matrix of the other.

5–26. Show that with planar graphs G, if the "windows" are chosen for loops, Theorem 5–7 is applicable to the circuit matrix also. (Cederbaum [26].)

5–27. Do the statements of Problems 3–8 and 3–9 remain true for directed graphs G and G^* (oriented as in Problem 5–25) even when orientations of paths, cut-sets, and circuits are taken into account? Justify your answer.

5–28. Prove Lemma 5–25(b). [*Hint:* Argument of Lemma 5–25(c).]

5–29. It follows from Theorems 5–24 and 5–25 that the determinants of a square submatrix of order $v - 1$ of $\mathbf{Q}_f$ and order $e - v + 1$ of $\mathbf{B}_f$ are 1, -1, or 0. Use this to show that if $\mathbf{B}$ is any circuit matrix (not necessarily fundamental or even regular), the determinants of all nonsingular submatrices of $\mathbf{B}$ have the same magnitude. [*Hint:* Express $\mathbf{B}$ in terms of $\mathbf{B}_f$ and use: the determinant of a product of two square matrices is the product of the determinants.]

5–30. Complete the detailed proof of Theorem 5–27.

5–31. Justify the reduction procedure used in Lemma 5–25(a) to reduce $\mathbf{F}_R$ to the form $[\mathbf{D}\ \mathbf{F}_{R2}]$.

5–32. Show, without using Theorem 5–25, that the matrix $\mathbf{B}_d$ of Eq. (5–26) is not regular.

CHAPTER 6

APPLICATIONS TO NETWORK ANALYSIS

As mentioned in Section 1–1, Kirchhoff founded the theory of graphs in its present form (as opposed to Euler's discussions) in 1847, specifically for its application to electrical networks. (Kirchhoff's contributions are distributed throughout this book; Chapters 5, 6, and 7 contain most of Kirchhoff's contributions to electrical network theory.) The present chapter is concerned with those aspects of electrical network analysis which depend on the theory of graphs. Much of the discussion is sufficiently general to be applicable to general linear systems, as is well recognized in the engineering profession.

The main purpose of this chapter is to provide a rigorous mathematical foundation for the discipline of electrical network theory, justifying many of the familiar statements and procedures of network analysis. A general familiarity with network analysis, including the Laplace transformation technique, is assumed in this chapter. Therefore no time is devoted to the "physical aspects" or to the relationship to other disciplines, e.g., the equations of Maxwell and those of Lagrange.

6–1 Kirchhoff's laws. Since the purpose here is to "prove" some properties of Kirchhoff's current and voltage equations, it is necessary to begin with a precise (postulational) formulation of Kirchhoff's laws. A very brief discussion of the concept of a reference is given first, to allow for the correlation of the present formulation with the conventional ones.

Electrical network theory is formulated in terms of two variables, current and voltage, associated with each network element (*branch,* in conventional terminology). As in any other physical science, these quantities are correlated with the readings of certain instruments, which in this case are called *ammeters* and *voltmeters.* Since our discussion here is concerned with current and voltage as *real functions of time,* $i(t)$ and $v(t)$, the meters should be of the "instantaneous-value" kind. They might be center-zero D'Arsonval meters or oscilloscopes, for instance. As is well known, the sign of the reading depends on the way in which the instrument is connected in the network; reversing the terminals changes the reading from positive to negative or vice versa. Hence, for unique correlation of theory with experiment, it is necessary to specify, on the network diagram, how these quantities are to be measured. Such a specification is done by means of current and voltage *references.* Figures 6–1 and 6–2 show the common references used and the meter connections implied by them.

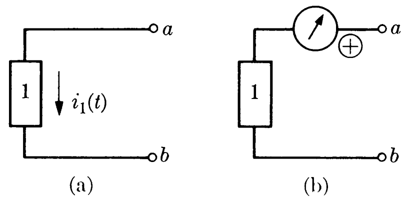


FIG. 6–1. Current reference convention.

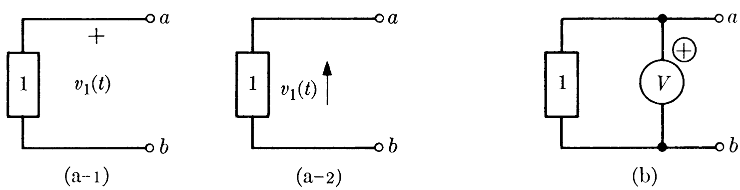


FIG. 6–2. Voltage reference convention.

In each figure, part (a) shows the reference, and part (b) shows the meter connection. The ⊕ on the meter stands for the +-terminal of the meter or the red terminal of the oscilloscope. Since current and voltage references can be arbitrarily assigned to a network element, there is no need to carry two sets of references. Hence, in this book, they are combined into one reference (which is used later as the magnetic polarity reference as well). The combined reference is identified as the edge orientation in the directed graph. The convention adopted here is shown in Fig. 6–3. Thus, all the voltage +-references are assumed to be at the tails of the current-reference arrows. (Those accustomed to the "rise" convention of Fig. 6–2(a–2) may find this a little confusing initially.)

Since the present formulation of Kirchhoff's laws may be unfamiliar, an example is given first, before the formal statement. In Fig. 6–4(a), a network is shown in familiar form, with all the current and voltage references shown and the voltage + being kept at the tail of the current-reference arrow. The three loops in the network and the loop references are also shown. From earlier experience, Kirchhoff's current and voltage equations for this network may be written as

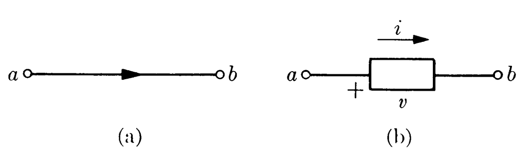


FIG. 6–3. Combined reference.

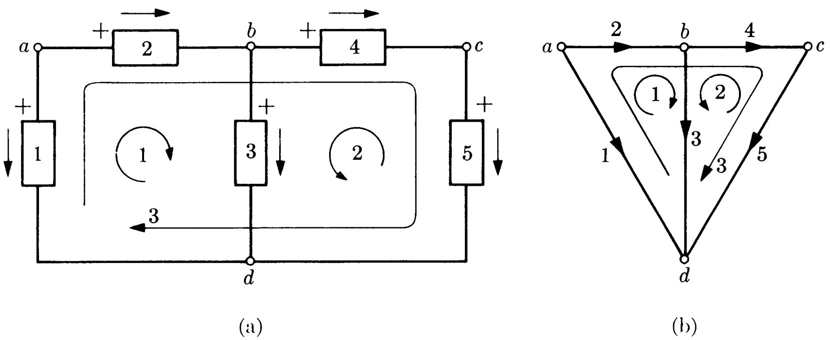


Fig. 6–4. Example for Kirchhoff's laws.

Current equations:

$$\begin{array}{lrcl}
\text{Node } a: & i_1(t) + i_2(t) & = 0, & \\
\text{Node } b: & -\, i_2(t) + i_3(t) + i_4(t) & = 0, & \\
\text{Node } c: & -\, i_4(t) + i_5(t) & = 0, & (6\text{–}1) \\
\text{Node } d: & -i_1(t) \quad - i_3(t) \quad - i_5(t) & = 0; &
\end{array}$$

Voltage equations:

$$\begin{array}{lrcl}
\text{Loop } 1: & -v_1(t) + v_2(t) + v_3(t) & = 0, & \\
\text{Loop } 2: & v_3(t) - v_4(t) - v_5(t) & = 0, & (6\text{–}2) \\
\text{Loop } 3: & -v_1(t) + v_2(t) \quad + v_4(t) + v_5(t) & = 0. &
\end{array}$$

Collecting these two systems of equations in matrix notation results in

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ -1 & 0 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} i_1(t) \\ i_2(t) \\ i_3(t) \\ i_4(t) \\ i_5(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{6–3}$$

and

$$\begin{bmatrix} -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \\ -1 & 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ v_4(t) \\ v_5(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \cdot \tag{6–4}$$

The coefficient matrices of Eqs. (6–3) and (6–4) are now recognized as the vertex and circuit matrices, respectively, of the directed graph of Fig. 6–4(b). It is not difficult to see that this observation is perfectly general, and is not peculiar to the example. The postulational forms of Kirchhoff's laws now follow.

DEFINITION 6–1. *Electrical network.* An *electrical network* is a directed linear graph with two real-valued functions v and i of the real variable t, which are of bounded variation, associated with each edge, satisfying the three postulates N_1, N_2, and N_3 below.*

POSTULATE N_1. *Kirchhoff's current law:*

$$\mathsf{A}_a \mathsf{i}(t) = \mathbf{0}, \tag{6–5}$$

where A_a *is the vertex matrix of the directed graph*

$$\mathsf{i}(t) = \begin{bmatrix} i_1(t) \\ i_2(t) \\ \vdots \\ i_e(t) \end{bmatrix} \tag{6–6}$$

and $i_j(t)$ *is associated with edge* j.

POSTULATE N_2. *Kirchhoff's voltage law:*

$$\mathsf{B}_a \mathsf{v}(t) = \mathbf{0}, \tag{6–7}$$

where B_a *is the circuit matrix of the directed graph*

$$\mathsf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_e(t) \end{bmatrix} \tag{6–8}$$

and $v_j(t)$ *is associated with edge* j.

Since the properties of incidence and circuit matrices are known, it suffices to restate these results as properties of Kirchhoff's current and voltage equations.

THEOREM 6–1. For a connected network, exactly $v - 1$ of Kirchhoff's current equations are linearly independent. In general, if the network is in p connected pieces, there are $v - p$ linearly independent Kirchhoff's current equations. In both cases, v is the number of vertices.

---

* The statement of Postulate N_3 is postponed to Section 6–3 to avoid confusion in the following discussion of Kirchhoff's laws.

Theorem 6–1 follows immediately from Theorem 5–1, since the linear dependence of a system of equations is decided by the rank of the coefficient matrix. Similarly, Theorem 6–2 follows from Theorem 5–5.

THEOREM 6–2. There are exactly $e - v + 1$ linearly independent Kirchhoff's voltage equations for a connected network with e edges and v vertices.

More interesting results are obtained on translating Theorems 5–6 and 5–8.

THEOREM 6–3. If T is any tree of a connected network, the voltage functions of the chords of T are expressible as linear combinations of the voltage functions of the branches of T, and the current functions of the branches of T are expressible as linear combinations of the current functions of the chords of T.

The proof of Theorem 6–3 is straightforward and is left as a problem (Problem 6–2).

Since the incidence matrix and the cut-set matrix differ only by a nonsingular transformation, it is possible to state Kirchhoff's current law by using the cut-set matrix.

THEOREM 6–4. If Q is a cut-set matrix of $v - 1$ cut-sets and rank $v - 1$, Kirchhoff's current equations

$$\mathsf{Ai}(t) = \mathbf{0} \tag{6–9}$$

are equivalent* to the system of equations

$$\mathsf{Qi}(t) = \mathbf{0}. \tag{6–10}$$

This result, which is easily proved (Problem 6–4) is familiar in a different form. If a subnetwork N_s is connected to the rest of the network by means of k wires, it is a familiar fact that the sum of the currents in the k wires (with references taken into account) is zero. Theorem 6–4 makes precisely this statement, since the k wires constitute a cut-set. Essentially, the matrices A and Q should be considered to be interchangeable. Almost any statement about A is also true about Q and conversely (except for A's property of having one 1 and one -1 per column, and the related result of Theorem 5–7). For this reason, many authors prefer to consider the cut-set matrix rather than the vertex matrix (see, for instance, Foster [59] or Guillemin [68]).

---

* Two systems of linear equations are *equivalent* if they have the same solution.

6–2 Mesh (loop) and node transformations. The discussion up to this point has been in terms of the so-called *branch variables*, namely the currents and voltages associated with the network elements. Although these quantities are more "basic" in the sense of being directly measurable, the loop and node variables are used more often in network analysis. This section is devoted to a justification of their use and discussion of the consequences. We decompose the vector space associated with the graph into orthogonal complements for this purpose. A justification based on the theory of equations is also possible and may be found elsewhere [156].

As observed earlier, the orthogonality relation

$$\mathsf{AB}' = \mathbf{0} \tag{6–11}$$

shows that the vector subspaces $\mathcal{V}_Q$ and $\mathcal{V}_B$ are orthogonal complements of the e-dimensional linear vector space $\mathcal{V}_G$. From the discussion of linear algebraic equations in Section 4–6, we recognize that the vector $\mathsf{i}(t)$ satisfying Kirchhoff's current equation must belong to $\mathcal{V}_B$, and similarly $\mathsf{v}(t)$ must belong to $\mathcal{V}_Q$. Since the rows of the matrices B and A are respectively bases for these two subspaces, $\mathsf{i}(t)$ must be a linear combination of the columns of B', and $\mathsf{v}(t)$ must be a linear combination of the columns of A'. Such a linear combination can be written in matrix notation as in Section 5–5. For instance, if B is partitioned into rows,

$$\mathsf{B} = \begin{bmatrix} \mathsf{B}_1 \\ \mathsf{B}_2 \\ \vdots \\ \mathsf{B}_\mu \end{bmatrix}, \qquad \mu = e - v + 1, \tag{6–12}$$

the expression for $\mathsf{i}(t)$ can be written as

$$\mathsf{i}(t) = [\mathsf{B}_1' \quad \mathsf{B}_2' \quad \cdots \quad \mathsf{B}_\mu'] \begin{bmatrix} i_{m1}(t) \\ i_{m2}(t) \\ \vdots \\ i_{m\mu}(t) \end{bmatrix} \tag{6–13}$$

where $i_{m1}, i_{m2}, \ldots, i_{m\mu}$ are the coefficients of the linear combination. Since B' is a matrix of constants and $\mathsf{i}(t)$ is a matrix of functions of t, the coefficients of the linear combination must be functions of t. Similar remarks apply to $\mathsf{v}(t)$.

THEOREM 6–5. The column matrix $\mathsf{i}(t)$ of element-current functions of a connected network satisfies Kirchhoff's current equation,

$$\mathsf{Ai}(t) = \mathbf{0}, \tag{6–14}$$

if and only if there exists a set of $e - v + 1$ functions $i_{mj}(t)$ such that

$$\mathsf{i}(t) = \mathsf{B}'\mathsf{i}_m(t), \tag{6–15}$$

where

$$\mathsf{i}_m(t) = \begin{bmatrix} i_{m1}(t) \\ i_{m2}(t) \\ \vdots \\ i_{m\mu}(t) \end{bmatrix}, \qquad \mu = e - v + 1, \tag{6–16}$$

and B is a circuit matrix of the network of $e - v + 1$ rows and rank $e - v + 1$.

The sufficiency of Eq. (6–15) is observed immediately, since

$$\mathsf{A}\mathsf{i}(t) = \mathsf{A}[\mathsf{B}'\mathsf{i}_m(t)] = (\mathsf{A}\mathsf{B}')\mathsf{i}_m(t) = \mathbf{0}. \tag{6–17}$$

Equation (6–15) is the *mesh,* or *loop, transformation.* The corresponding theorem for $\mathsf{v}(t)$ is given next and follows from similar arguments.

THEOREM 6–6. The column matrix of voltage functions $\mathsf{v}(t)$ of a connected network satisfies Kirchhoff's voltage equation,

$$\mathsf{B}\mathsf{v}(t) = \mathbf{0}, \tag{6–18}$$

if and only if there exists a set of $v - 1$ functions $v_{nj}(t)$ such that

$$\mathsf{v}(t) = \mathsf{A}'\mathsf{v}_n(t), \tag{6–19}$$

where

$$\mathsf{v}_n(t) = \begin{bmatrix} v_{n1}(t) \\ v_{n2}(t) \\ \vdots \\ v_{n\rho}(t) \end{bmatrix}, \qquad \rho = v - 1, \tag{6–20}$$

and A is an incidence matrix of $v - 1$ rows of the network.

Equation (6–19) is the *node transformation.* The variables $i_{mj}(t)$ are known as *mesh,* or *loop, currents,* and $v_{nj}(t)$ are the *node voltages* (also known as *node-to-datum voltages*).

In Theorem 6–6, the incidence matrix A can evidently be replaced by a cut-set matrix Q of $v - 1$ rows and rank $v - 1$. Then the transformation becomes

$$\mathsf{v}(t) = \mathsf{Q}'\mathsf{v}_p(t). \tag{6–21}$$

With most (but not all) cut-set matrices, the variables in $\mathsf{v}_p(t)$ can be identified as *node-pair* voltages, i.e., voltages between some pairs of nodes in the network. If Q becomes Q_f, the variables are the voltages of the branches of the corresponding tree (see Section 6–4). In these cases (where the interpretation *node-pair voltages* is possible) Eq. (6–21) is referred to as the *node-pair transformation.* However, for any cut-set matrix $\mathbf{Q}$, Eq. (6–21) is valid, provided only that $\mathbf{Q}$ has $v - 1$ rows and rank $v - 1$.

Although the concepts of loop currents and node voltages are extremely familiar, let us illustrate the transformations by means of an example, to show that the matrix equations agree with the familiar conceptions of loop currents and node voltages; only the notation is new.

For the network of Fig. 6–5, the incidence and circuit matrices are

$$\mathbf{A} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{bmatrix} -1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{bmatrix} \end{array} \tag{6–22}$$

and

$$\mathbf{B} = \begin{array}{c} 1 \\ 2 \\ 3 \end{array}\!\!\begin{bmatrix} 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{bmatrix}. \tag{6–23}$$

The mesh transformation is therefore

$$\mathbf{i}(t) = \mathbf{B}'\mathbf{i}_m(t),$$

$$\begin{bmatrix} i_1(t) \\ i_2(t) \\ i_3(t) \\ i_4(t) \\ i_5(t) \\ i_6(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_{m1}(t) \\ i_{m2}(t) \\ i_{m3}(t) \end{bmatrix}, \tag{6–24}$$

which becomes, on multiplying out,

$$\begin{aligned} i_1(t) &= i_{m1}(t), \\ i_2(t) &= -i_{m1}(t) + i_{m2}(t), \\ i_3(t) &= i_{m1}(t) - i_{m3}(t), \\ i_4(t) &= -i_{m2}(t) + i_{m3}(t), \\ i_5(t) &= -i_{m2}(t), \\ i_6(t) &= i_{m3}(t). \end{aligned} \tag{6–25}$$

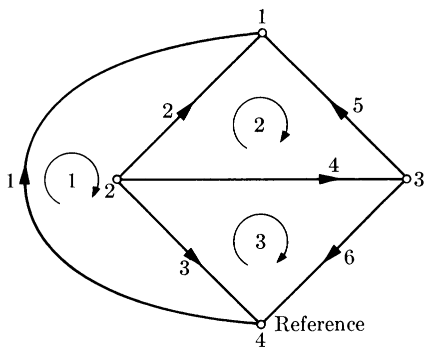


FIG. 6-5. Loop and node transformations.

On examining Fig. 6-5, it is seen that these are indeed the correct expressions for the element currents in terms of the loop currents.

The node transformation is

$$\mathsf{v}(t) = \mathsf{A}'\mathsf{v}_n(t),$$

$$\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ v_4(t) \\ v_5(t) \\ v_6(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{n1}(t) \\ v_{n2}(t) \\ v_{n3}(t) \end{bmatrix} \tag{6-26}$$

or, in scalar form,

$$\begin{aligned} v_1(t) &= -v_{n1}(t), \\ v_2(t) &= -v_{n1}(t) + v_{n2}(t), \\ v_3(t) &= \quad v_{n2}(t), \\ v_4(t) &= \quad v_{n2}(t) - v_{n3}(t), \\ v_5(t) &= -v_{n1}(t) + v_{n3}(t), \\ v_6(t) &= \quad v_{n3}(t). \end{aligned} \tag{6-27}$$

Since all the element-voltage references are at the tails of the current reference arrows, these are indeed the correct expressions for the element voltages in terms of the voltages of nodes 1, 2, and 3 with respect to the reference (or datum) node 4.

Thus the results of Theorems 6-5 and 6-6 are not particularly new. However they do justify the use of loop and node variables in network analysis, and provide an elegant notation for writing these expressions. The next theorem makes use of this elegance of notation.

THEOREM 6–7 (power relation). If the currents and voltages in a network satisfy Kirchhoff's laws, then

$$\sum_{j=1}^{e} v_j(t) \cdot i_j(t) = \mathbf{i}'(t)\mathbf{v}(t) = 0. \tag{6–28}$$

Proof. By Theorems 6–5 and 6–6,

$$\mathbf{i}(t) = \mathbf{B}'\mathbf{i}_m(t) \quad \text{and} \quad \mathbf{v}(t) = \mathbf{A}'\mathbf{v}_n(t). \tag{6–29}$$

Hence

$$\mathbf{i}'(t) \cdot \mathbf{v}(t) = \mathbf{i}'_m(t)\mathbf{B}\mathbf{A}'\mathbf{v}_n(t) = \mathbf{i}'_m(t)(\mathbf{B}\mathbf{A}')\mathbf{v}_n(t) = 0 \tag{6–30}$$

since

$$\mathbf{B}\mathbf{A}' = \mathbf{0}. \tag{6–31}$$

For the reference convention adopted, the expression for "power absorbed in element j" is

$$p_j(t) = v_j(t)i_j(t). \tag{6–32}$$

From this, Theorem 6–7 can be interpreted as stating that

$$\sum_{j=1}^{e} p_j(t) = 0. \tag{6–33}$$

On integrating Eq. (6–33) between any two limits t_1 and t_2, the result is Theorem 6–8.

THEOREM 6–8 (conservation of energy). If the energy function is absolutely continuous, so that

$$p_j(t) = \frac{dw_j(t)}{dt} \quad \text{and} \quad w_j(t) = \int_a^t p_j(x)\,dx, \tag{6–34}$$

then Kirchhoff's laws imply conservation of energy:

$$\sum_{j=1}^{e} w_j(t) \text{ is constant.}$$

Stated differently, Theorem 6–8 implies that conservation of energy need not be added as a postulate of the discipline of network theory. It is already included in the theory, so long as energy is a well-behaved function.

The mesh and node transformations are of interest from a mathematical point of view because they are singular transformations (defined by singular matrices). It turns out to be an interesting problem to investigate when a given singular transformation can be used and when it cannot be used. The interested reader is referred to an original article [151] for such a discussion.

6-3 The third postulate. Postulates N_1 and N_2 are concerned only with the way in which the network elements are interconnected. The character of the network elements (resistor, inductor, capacitor, generator, etc.) does not enter the discussion of Kirchhoff's laws in any way. Kirchhoff's laws are associated purely with the topology of the network. On the other hand, the character of the individual network element (whether it is a resistor or inductor or generator) is clearly independent of where in the network the element happens to be located. The network element is characterized by the relationship between voltage and current. Postulate N_3 concerns this relationship. The independence of the two aspects of a network (the geometry, or interconnection aspect, and the character of the network elements) must be clearly borne in mind.

The functions associated with each element of a network, in addition to satisfying Kirchhoff's laws, are required to satisfy a system of integro-differential equations. These element equations have the general form

$$\mathsf{v}(t) = \mathsf{L}\frac{d\mathsf{i}(t)}{dt} + \mathsf{R}\mathsf{i}(t) + \mathsf{D}\int_0^t \mathsf{i}(x)\,dx + \mathsf{e}(t) + \mathsf{v}_C(0+). \quad (6\text{–}35)$$

The entries in the matrices L, R, and D characterize the equations and the network.

(a) If the matrices are symmetric, the network is *bilateral,* or *reciprocal;* otherwise it is *nonreciprocal.*

(b) If the entries in the matrices L, R, and D are independent of the dependent variables v_j and i_j, then the network is *linear;* otherwise it is *nonlinear.*

(c) If the entries in the matrices L, R, and D are functions of the independent variable t, but not of i_j and v_j, the network is a *linear time-variable* network.

(d) If the matrices L, R, and D are positive semidefinite or definite, and if $\mathsf{e}(t) = 0$, the network is *passive.*

(e) If the matrices R, L, and D contain only constants, the network is *linear time-invariant.*

The general principles of the discussions in this chapter are applicable to all linear time-invariant networks. The discussions up to this point are applicable to all lumped networks. However, for the major theorems in the rest of this chapter, a linear, reciprocal, time-invariant network with positive semidefinite matrices is assumed. The reason for this restriction is given by means of an example at the end of Section 6–4. The type of network to be considered is characterized by the third postulate.

POSTULATE N_3. *The functions* $\mathsf{v}(t)$ *and* $\mathsf{i}(t)$ *are related by*

$$\mathsf{v}(t) = \mathsf{L}\frac{d}{dt}\,\mathsf{i}(t) + \mathsf{R}\mathsf{i}(t) + \mathsf{D}\int_0^t \mathsf{i}(x)\,dx + \mathsf{v}_C(0+) + \mathsf{e}(t), \qquad (6\text{–}36)$$

where R *and* D *are real diagonal matrices with nonnegative entries on the main diagonal, and* L *is real symmetric, with the nonzero rows and columns of* L *constituting a positive definite submatrix.*

A detailed discussion of the concept of positive definiteness may be found in Hohn [78]. The definition and some important properties are given here for the continuity of the discussion. A real symmetric matrix F of order n is positive definite if for every real vector $\mathsf{X} \neq \mathsf{0}$ of order n,

$$\mathsf{X}'\mathsf{F}\mathsf{X} > 0,$$

where $\mathsf{X}'\mathsf{F}\mathsf{X}$ is a quadratic form. It can be expanded as

$$\mathsf{X}'\mathsf{F}\mathsf{X} = \sum_{r=1}^{n}\sum_{s=1}^{n} x_r f_{rs} x_s.$$

The definition can also be formulated in terms of complex vectors X, in which case the transpose of X must be replaced by the transposed conjugate. It is left as a problem (Problem 6–14) to show that the two are equivalent for real matrices F. It is a trivial consequence of the definition that every diagonal matrix with positive diagonal entries is positive definite, for then the quadratic form is merely

$$\sum_{i=1}^{n} x_i^2 f_{ii}.$$

A useful test for positive definiteness is the following. A *leading principal* minor of order r is the determinant of the submatrix consisting of the first r rows and the first r columns. It can be shown that a symmetric matrix is positive definite if and only if all the leading principal minors of order r are positive for $r = 1, 2, 3, \ldots, n$.

Positive semidefiniteness is defined as follows. The real symmetric matrix F is positive semidefinite if for all real vectors $\mathsf{X} \neq \mathsf{0}$ of order n,

$$\mathsf{X}'\mathsf{F}\mathsf{X} \geq 0,$$

provided there is at least one $\mathsf{X} \neq \mathsf{0}$ for which the equality sign applies. If the matrix F is positive semidefinite, then all the leading principal minors are nonnegative. The converse, however, is not true in this case. A counterexample is the matrix

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

The leading principal minors of $\mathbf{F}$ are all nonnegative, but $\mathbf{F}$ is neither positive definite nor semidefinite. Define a *principal minor* of order r to consist of rows $i_1, i_2, \ldots, i_r$ and columns $i_1, i_2, \ldots, i_r$ (that is, chosen symmetrically) of the matrix $\mathbf{F}$. Then the matrix is positive semidefinite if and only if all the principal minors are nonnegative.

The condition that the matrix $\mathbf{L}$ be positive semidefinite is equivalent to requiring that a passive system be stable. In Section 6–5 it is shown that a passive network satisfying N_3 is stable. Conversely, the following statement is proved in Section 6–5. *If a set of inductors can be found such that the matrix* $\mathbf{L}$ *of these inductors is neither positive definite nor semidefinite, then we can build a passive network consisting of these inductors and some positive resistors which is unstable.* Postulate N_3 makes a stronger requirement, namely that the nonzero rows and columns of $\mathbf{L}$ must constitute a positive definite submatrix. This condition is equivalent to prohibiting "perfectly coupled" transformers. The uniqueness theorems established in Section 6–4 are not true for networks containing perfectly coupled transformers. It is also possible to justify the positive semidefiniteness of the matrix $\mathbf{L}$ by showing that the quadratic form $\mathbf{i}'\mathbf{L}\mathbf{i}$ is the energy stored in the magnetic field of the inductors, by using a rather complicated field-theoretic argument. However, we prefer to base the justification on stability.

The matrix $\mathbf{e}(t)$ corresponds to the so-called *driving functions* or *generators.* These are the elements of the network for which either $v(t)$ or $i(t)$ is a specified function. If $v(t)$ is specified, it is referred to as a *voltage driver* or *voltage generator,* and if $i(t)$ is specified, it is a *current driver* or *current generator.*

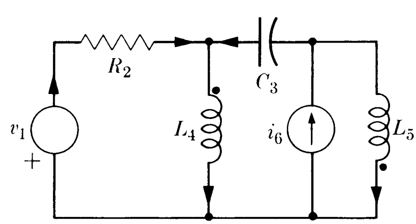


FIG. 6–6. Example for Postulate N_3.

Since these equations have been written in somewhat unfamiliar notation, let us consider a simple example and write out the equations for the example. For the network of Fig. 6–6, the element equations of Postulate N_3 are

$$
\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ v_4(t) \\ v_5(t) \\ v_6(t) \end{bmatrix}
=
\begin{bmatrix} v_1(t) \\ 0 \\ 0 \\ 0 \\ 0 \\ v_6(t) \end{bmatrix}
+
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & L_{44} & L_{45} & 0 \\
0 & 0 & 0 & L_{45} & L_{55} & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} \frac{d}{dt} i_1(t) \\ \frac{d}{dt} i_2(t) \\ \frac{d}{dt} i_3(t) \\ \frac{d}{dt} i_4(t) \\ \frac{d}{dt} i_5(t) \\ \frac{d}{dt} i_6(t) \end{bmatrix}
$$

$$
+
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & R_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} i_1(t) \\ i_2(t) \\ i_3(t) \\ i_4(t) \\ i_5(t) \\ i_6(t) \end{bmatrix}
$$

$$
+
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & D_3 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} \int_0^t i_1(x)\,dx \\ \int_0^t i_2(x)\,dx \\ \int_0^t i_3(x)\,dx \\ \int_0^t i_4(x)\,dx \\ \int_0^t i_5(x)\,dx \\ \int_0^t i_6(x)\,dx \end{bmatrix}
+
\begin{bmatrix} 0 \\ 0 \\ v_3(0+) \\ 0 \\ 0 \\ 0 \end{bmatrix}
. \tag{6–37}
$$

These equations can be rewritten more concisely (and with fewer zeros) as

$$\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ v_4(t) \\ v_5(t) \\ v_6(t) \end{bmatrix} = \begin{bmatrix} v_1(t) \\ 0 \\ 0 \\ 0 \\ 0 \\ v_6(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{p} D_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{44}p & L_{45}p & 0 \\ 0 & 0 & 0 & L_{45}p & L_{55}p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_1(t) \\ i_2(t) \\ i_3(t) \\ i_4(t) \\ i_5(t) \\ i_6(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ v_3(0+) \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{6-38}$$

where the operational notation

$$pi(t) = \frac{di(t)}{dt} \qquad \text{and} \qquad \frac{1}{p}\, i(t) = \int_0^t i(x)\, dx \tag{6-39}$$

has been used.

A few remarks on these equations might eliminate some possible confusion. First, the rows and columns in the coefficient matrix which correspond to the drivers are zero. Second, each R, L, C, and generator has been considered to be a separate element. In network analysis, such an assumption is unnecessary. But we find this convention useful later, and so a uniform convention is adopted. A possible source of confusion is the use of L_{45} in the matrix, whereas the polarity marks in Fig. 6–6 seem to indicate $-L_{45}$ as the appropriate entry. In this case, L_{45} has been taken to be a negative number. It is theoretically more convenient to write

$$\mathsf{L} = [L_{ij}] \tag{6-40}$$

and let L_{ij} be positive or negative rather than to write

$$\mathsf{L} = [\pm L_{ij}], \tag{6-41}$$

with the choice of signs in the matrix depending on the particular network. The following convention is associated with Eq. (6–40). If the two polarity marks on windings i and j are similarly situated with respect to the edge orientation in the directed graph of the network (i.e., both at the tail or both at the head of the orientation arrow), L_{ij} is a positive number; otherwise it is negative. If for any particular reason it is desired that mutual inductance should be nonnegative, mutual inductance can be defined as $M_{ij} = |L_{ij}|$.

Postulate N_3 can be written concisely in operational notation as

$$\mathsf{v}(t) = \mathsf{e}(t) + \mathsf{Z}(p)\mathsf{i}(t) + \mathsf{v}_C(0+), \tag{6-42}$$

where, it should be emphasized, $\mathbf{Z}(p)$ contains operators and so must be handled with care.

6–4 Loop and node systems of equations. In this section, the loop and node systems of equations are established on a firm foundation for rather general networks. We also examine the conditions under which these equations have unique solutions, i.e., the conditions under which the coefficient matrices are nonsingular. The three fundamental systems of equations of network theory constitute the starting point of the development. These are

$$\begin{aligned} \mathbf{Ai}(t) &= \mathbf{0}, \qquad \mathbf{Bv}(t) = \mathbf{0}, \\ \mathbf{v}(t) &= \mathbf{e}(t) + \mathbf{Z}(p)\mathbf{i}(t) + \mathbf{v}_C(0+). \end{aligned} \tag{6–43}$$

Since Postulate N_3 leads to a system of ordinary linear integrodifferential equations with constant coefficients, the Laplace transform method of solution is the most convenient. In this text, a uniform convention is adopted for Laplace transforms. Capital letters always stand for Laplace transforms of the corresponding lower-case symbol. Thus, for instance,

$$\mathcal{L}\{f(t)\} = \int_0^\infty \epsilon^{-st} f(t)\,dt = F(s).$$

In Laplace transforms, the three fundamental systems of equations (6–43) become:

$$\begin{aligned} \mathbf{AI}(s) &= \mathbf{0}, \qquad \mathbf{BV}(s) = \mathbf{0}, \\ \mathbf{V}(s) &= \mathbf{E}(s) + \mathbf{Z}(s)\mathbf{I}(s) - \mathbf{Li}(0+) + \frac{1}{s}\,\mathbf{v}_C(0+). \end{aligned} \tag{6–44}$$

As implied by the notation, $\mathbf{Z}(s)$ is obtained by replacing p by s in $\mathbf{Z}(p)$ of Eq. (6–43). The last two terms in Eq. (6–44) correspond to the initial values. Clearly, the matrix $\mathbf{i}(0+)$ can be replaced by one containing only the inductor currents $\mathbf{i}_L(0+)$, since all other entries of $\mathbf{L}$ are zero.

In Laplace transforms, the mesh and node transformations are

$$\mathbf{I}(s) = \mathbf{B}'\mathbf{I}_m(s) \qquad \text{and} \qquad \mathbf{V}(s) = \mathbf{A}'\mathbf{V}_n(s), \tag{6–45}$$

which are respectively equivalent to Kirchhoff's current and voltage equations, as observed earlier.

The systems of loop and node equations to be derived here are for very general networks with arbitrary distributions of current and voltage generators and with all initial conditions taken into account; consequently, they are "complicated." Therefore it is worth while to draw a "flow chart," which incidentally shows the derivation of simplified systems of loop and node equations. The systems illustrated in the flow chart of Fig. 6–7 are for networks satisfying the conditions:

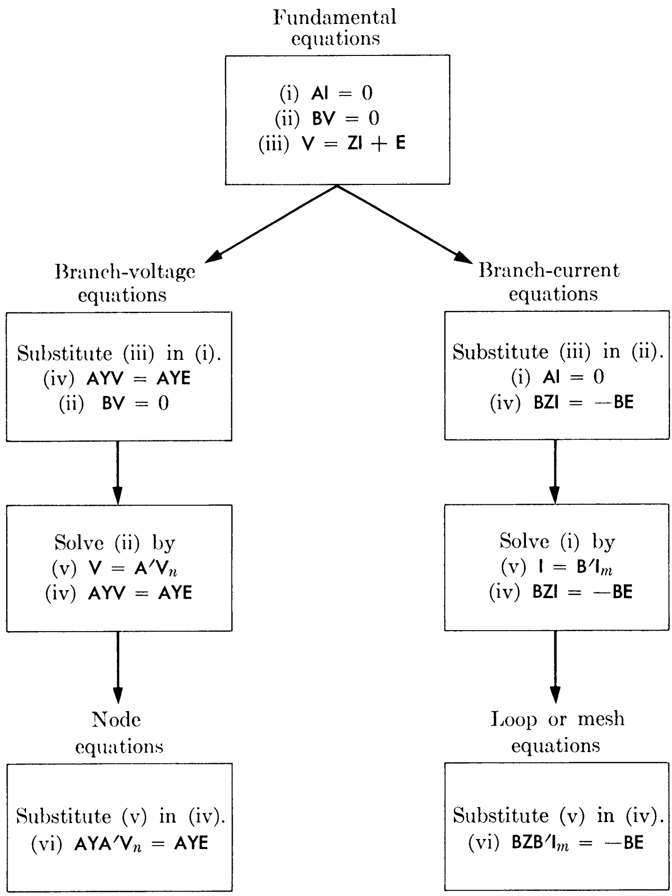


FIG. 6–7. Flow chart for loop and node equations.

(a) All initial conditions are zero.

(b) The network contains no current generators.

(c) Each voltage generator has an R, L, or C in series, and the two together are considered as a network element.

In the chart, $\mathbf{Y} = \mathbf{Z}^{-1}$, which exists by conditions (b) and (c). The branch-current equations in column 2 of the chart are of historical importance because many of the early workers, including Kirchhoff himself, used the branch-current system of equations. The general system, to be derived next, follows the same pattern as Fig. 6–7 but is more involved, since the matrices have to be partitioned in various ways. For most practical purposes, the simplified systems shown in the chart of Fig. 6–7 suffice. The main purpose in deriving the generalized systems of equations is to establish rigorously the conditions under which unique solutions can be obtained for an electrical network.

In the general case, a network satisfying Postulates N_1, N_2, and N_3 is considered. For the principle of the derivation, the restriction of N_3 to reciprocal networks is unnecessary, but is vital to Theorems 6–11 and 6–12. In the general derivation, each R, L, C, and generator is assumed to be a separate network element. Two general assumptions are made as follows:

(a) Whenever a row (and column) of $\mathsf{Z}(s)$ is zero, the corresponding element is a driver; and either $v(t)$ or $i(t)$ is specified for this element.

(b) There exists a tree T_1 of the network such that all the current generators are chords for this tree, and there exists a tree T_2 (which may or may not be the same as T_1) such that all the voltage generators are branches of T_2. (T_1 and T_2 usually have other chords and branches, respectively, as well.)

Assumption (a) is meaningful, for otherwise we have an element in the network about which nothing is known. Assumption (b), although it appears artificial, is not a restriction. In the interest of logical order, it is shown before the general derivation is undertaken that assumption (b) is a necessary condition for the unique solvability of the network equations.

THEOREM 6–9. If for a connected network the equations

$$\begin{aligned} \mathsf{AI}(s) &= \mathbf{0}, \qquad \mathsf{BV}(s) = \mathbf{0}, \\ \mathsf{V}(s) &= \mathsf{E}(s) + \mathsf{Z}(s)\mathsf{I}(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{Li}_L(0+) \end{aligned} \tag{6–46}$$

have a unique solution for $\mathsf{I}(s)$ and $\mathsf{V}(s)$, then there exists a tree such that the current drivers are chords for this tree and the voltage drivers are branches for this tree.

Proof. First write the three systems of equations together so that the known quantities can be separated from the unknowns and the coefficient matrix examined:

$$\begin{bmatrix} \mathsf{A} & \mathbf{0} \\ \mathbf{0} & \mathsf{B} \\ -\mathsf{Z}(s) & \mathsf{U} \end{bmatrix} \begin{bmatrix} \mathsf{I}(s) \\ \mathsf{V}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathsf{E}(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{Li}_L(0+) \end{bmatrix}. \tag{6–47}$$

Now the matrices must be partitioned so that the known quantities can be transposed to the right and the unknowns to the left. [In Eq. (6–47) the currents of the current drivers and voltages of the voltage drivers are the known quantities on the left, and the voltages of the current drivers are unknowns appearing on the right.] To this end, arrange the variables such that the current drivers appear first and the voltage drivers last. Rearrange the rows and columns of $\mathbf{Z}$, and the columns of $\mathbf{A}$ and $\mathbf{B}$, such that the equations are unaltered. Let a subscript 1 denote the current drivers, a subscript 3 the voltage drivers, and a subscript 2 the others. Then, in partitioned form, the network equations are

$$\begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{B}_1 & \mathbf{B}_2 & \mathbf{B}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Z}_{22} & \mathbf{0} & \mathbf{0} & \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_1(s) \\ \mathbf{I}_2(s) \\ \mathbf{I}_3(s) \\ \mathbf{V}_1(s) \\ \mathbf{V}_2(s) \\ \mathbf{V}_3(s) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{E}_1(s) \\ \frac{1}{s}\,\mathbf{v}_C(0+) - \mathbf{L}\mathbf{i}_L(0+) \\ \mathbf{E}_3(s) \end{bmatrix}. \tag{6–48}$$

The known quantities in these equations are $\mathbf{I}_1(s)$, $\mathbf{E}_3(s)$, $\mathbf{v}_C(0+)$, and $\mathbf{i}_L(0+)$. The others, including the voltages of the current drivers and the currents of the voltage drivers, are unknowns. Transposing the known $\mathbf{I}_1(s)$ to the right and the unknown $\mathbf{E}_1(s)$ to the left, the third and fifth equations become trivial since $\mathbf{E}_1(s)$ and $\mathbf{E}_3(s)$ are merely alternative symbols for $\mathbf{V}_1(s)$ and $\mathbf{V}_3(s)$. Deleting these trivial equations, we find

$$\begin{bmatrix} \mathbf{A}_2 & \mathbf{A}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_1 & \mathbf{B}_2 \\ -\mathbf{Z}_{22} & \mathbf{0} & \mathbf{0} & \mathbf{U} \end{bmatrix} \begin{bmatrix} \mathbf{I}_2(s) \\ \mathbf{I}_3(s) \\ \mathbf{V}_1(s) \\ \mathbf{V}_2(s) \end{bmatrix} = \begin{bmatrix} -\mathbf{A}_1\mathbf{I}_1(s) \\ -\mathbf{B}_3\mathbf{V}_3(s) \\ \frac{1}{s}\,\mathbf{v}_C(0+) - \mathbf{L}\mathbf{i}_L(0+) \end{bmatrix}. \tag{6–49}$$

By the hypotheses of the theorem, these equations have a unique solution. The coefficient matrix is therefore nonsingular. The rows of the coefficient matrix are therefore linearly independent and hence so is any subset of rows. Consider the first $v - 1$ rows. Since these are linearly independent, the matrix $[\mathbf{A}_2\ \mathbf{A}_3]$ contains a nonsingular submatrix of order $v - 1$. By Theorem 5–6, the columns of this submatrix correspond to a tree T_1 of the network. For this tree, the current drivers are evidently chords, since $\mathbf{A}_1$ corresponds to the set of current drivers. Similarly, the second set of $e - v + 1$ rows is linearly independent, and so the matrix $[\mathbf{B}_1\ \mathbf{B}_2]$ contains a nonsingular submatrix of order $e - v + 1$. By Theorem 5–8, the columns of this submatrix correspond to the set of all chords of a tree T_2. Since the voltage drivers correspond to the columns of $\mathbf{B}_3$, the voltage drivers are branches of T_2. To complete the proof, we must show that T_1 and T_2 *can be chosen* to be the same tree. This result is established as a separate theorem.

THEOREM 6–10. Let G be a connected graph. Let S_1 and S_2 be edge disjoint subsets of G, such that (a) there exists a tree T_1 such that the edges of S_1 are chords of T_1 (not necessarily the whole of the chord set of T_1) and (b) there exists a tree T_2 such that the edges of S_2 are branches of T_2 (again T_2 may have other branches besides edges of S_2); then there exists a tree T for which the elements of S_2 are branches and the elements of S_1 are chords.

Proof. Delete the elements of S_1 from G. Let the rest of G be denoted by G_1. Since G_1 contains the tree T_1 of G, G_1 is connected and contains all the vertices of G. S_2 is contained in G_1 since S_1 and S_2 have no common edges. Since, by hypothesis, S_2 is contained in a tree of G, S_2 contains no circuits, and hence can be made part of a tree of G_1 (Theorem 2–12). Let this tree be T. Then T is also a tree of G, and satisfies the conditions imposed in the theorem.

The proof of Theorem 6–9 is now complete and the general assumption (b) on the distribution of generators is justified. It is possible to show that the hypotheses of Theorem 6–9 are also *sufficient* for the unique solvability of the network equations (6–49) (see Problem 6–15). However, this result is established with greater ease with the help of mesh and node systems of equations. The mesh equations are considered next.

Let the variables be arranged as in the proof of Theorem 6–9. Let T be a tree of the network which contains all the voltage drivers and none of the current drivers. For the current drivers, choose fundamental circuits. The circuits for the other chords need not be f-circuits, but are chosen so that they do not contain the current drivers. Thus the current drivers are placed in exactly one circuit each. Then, partitioning the columns of $\mathbf{B}$ as in Theorem 6–9, and partitioning the rows after the rows corresponding to the f-circuits for the current drivers, we find that Kirchhoff's voltage equations become

$$\begin{bmatrix} \mathbf{U} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{0} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1(s) \\ \mathbf{V}_2(s) \\ \mathbf{V}_3(s) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{6–50}$$

(The unit matrix and the zeros below it arise because the current drivers are in exactly one circuit each.)

Of these two equations (6–50), set aside the first equation,

$$\mathbf{V}_1(s) = -\mathbf{B}_{12}\mathbf{V}_2(s) - \mathbf{B}_{13}\mathbf{V}_3(s), \tag{6–51}$$

for the present. It is used later to find $\mathbf{V}_1(s)$ after the voltage transforms $\mathbf{V}_3(s)$ are found.

In this development, the same circuit matrix is also used for the mesh transformation, although it is clear from Theorem 6–5 that it is not necessary to do so. Any circuit matrix of the network containing $e - v + 1$ rows and of rank $e - v + 1$ can be used. If this procedure is followed [i.e., if different circuit matrices are used for (1) Kirchhoff's voltage equations and (2) the mesh transformation], the coefficient matrix of the loop system of equations will be asymmetrical even for a reciprocal (or bilateral) network. However, there is no generality achieved by following such a procedure. The symmetry, on the other hand, is convenient. Furthermore, for the positive-definiteness arguments to follow, the symmetry is essential. Therefore, if we use the same circuit matrix as in Eq. (6–50), the mesh transformation becomes

$$\begin{bmatrix} \mathsf{I}_1(s) \\ \mathsf{I}_2(s) \\ \mathsf{I}_3(s) \end{bmatrix} = \begin{bmatrix} \mathsf{U} & \mathsf{0} \\ \mathsf{B}'_{12} & \mathsf{B}'_{22} \\ \mathsf{B}'_{13} & \mathsf{B}'_{23} \end{bmatrix} \begin{bmatrix} \mathsf{I}_{m1}(s) \\ \mathsf{I}_{m2}(s) \end{bmatrix}. \tag{6–52}$$

We see from the first row that

$$\mathsf{I}_{m1}(s) = \mathsf{I}_1(s) \tag{6–53}$$

and hence is known. (This should be obvious from earlier experience, since the current drivers are in exactly one loop each and so the loop current is equal to the generator current.) We see from the third row of the mesh transformation that

$$\mathsf{I}_3(s) = \mathsf{B}'_{13}\mathsf{I}_1(s) + \mathsf{B}'_{23}\mathsf{I}_{m2}(s). \tag{6–54}$$

This equation is also set aside for computing the functions $\mathsf{I}_3(s)$ after $\mathsf{I}_{m2}(s)$ have been found. From Theorem 6–5, Kirchhoff's current law need not be considered after the mesh transformation has been used.

The voltage-current relations written in the present partitioned form are

$$\begin{bmatrix} \mathsf{V}_1(s) \\ \mathsf{V}_2(s) \\ \mathsf{V}_3(s) \end{bmatrix} = \begin{bmatrix} \mathsf{0} & \mathsf{0} & \mathsf{0} \\ \mathsf{0} & \mathsf{Z}_{22}(s) & \mathsf{0} \\ \mathsf{0} & \mathsf{0} & \mathsf{0} \end{bmatrix} \begin{bmatrix} \mathsf{I}_1(s) \\ \mathsf{I}_2(s) \\ \mathsf{I}_3(s) \end{bmatrix} + \begin{bmatrix} \mathsf{E}_1(s) \\ \mathsf{0} \\ \mathsf{E}_3(s) \end{bmatrix} + \begin{bmatrix} \mathsf{0} \\ \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+) \\ \mathsf{0} \end{bmatrix}. \tag{6–55}$$

As in Theorem 6–9, the first and third equations are trivial. The second equation is

$$\mathsf{V}_2(s) = \mathsf{Z}_{22}(s)\mathsf{I}_2(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+). \tag{6–56}$$

Now perform the combinations indicated in the chart of Fig. 6–7. The equations to be considered are

$$\begin{aligned} \mathsf{B}_{22}\mathsf{V}_2(s) + \mathsf{B}_{23}\mathsf{V}_3(s) &= \mathbf{0}, \\ \mathsf{V}_2(s) &= \mathsf{Z}_{22}\mathsf{I}_2(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+), \qquad (6\text{–}57) \\ \mathsf{I}_2(s) &= \mathsf{B}'_{12}\mathsf{I}_1(s) + \mathsf{B}'_{22}\mathsf{I}_{m2}(s). \end{aligned}$$

Performing the operations indicated in the chart of Fig. 6–7, we find that the final mesh or loop system of equations is

$$\begin{aligned} &\mathsf{B}_{22}\mathsf{Z}_{22}(s)\mathsf{B}'_{22}\mathsf{I}_{m2}(s) \\ &\quad = -\mathsf{B}_{23}\mathsf{V}_3(s) - \mathsf{B}_{22}\mathsf{Z}_{22}\mathsf{B}'_{12}\mathsf{I}_1(s) - \frac{1}{s}\,\mathsf{B}_{22}\mathsf{v}_C(0+) + \mathsf{B}_{22}\mathsf{L}_{22}\mathsf{i}_{L2}(0+). \qquad (6\text{–}58) \end{aligned}$$

By anticipating the results of Theorem 6–12, we can solve these equations for $\mathsf{I}_{m2}(s)$ in terms of the known generators and initial values. Substitute the solution for $\mathsf{I}_{m2}(s)$ in

$$\mathsf{I}_2(s) = \mathsf{B}'_{12}\mathsf{I}_1(s) + \mathsf{B}'_{22}\mathsf{I}_{m2}(s) \qquad (6\text{–}59)$$

to find $\mathsf{I}_2(s)$ and in

$$\mathsf{I}_3(s) = \mathsf{B}'_{13}\mathsf{I}_1(s) + \mathsf{B}'_{23}\mathsf{I}_{m2}(s) \qquad (6\text{–}60)$$

to find $\mathsf{I}_3(s)$. Substitute $\mathsf{I}_2(s)$ in

$$\mathsf{V}_2(s) = \mathsf{Z}_{22}\mathsf{I}_2(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+) \qquad (6\text{–}61)$$

to find $\mathsf{V}_2(s)$. Finally, substitute $\mathsf{V}_2(s)$ in

$$\mathsf{V}_1(s) = -\mathsf{B}_{12}\mathsf{V}_2(s) - \mathsf{B}_{13}\mathsf{V}_3(s) \qquad (6\text{–}62)$$

to find $\mathsf{V}_1(s)$. The time functions are found by inverting the Laplace transforms. Then all the currents and voltages are found, and the analysis of the network is complete. Observe that the loop method of analysis is an organized procedure for reducing the number of equations to be solved simultaneously, from $2e$ to $e - v + 1 -$ (number of current generators), which latter number is the number of equations in the system (6–58). The rest of the analysis consists of substitution.

The coefficient matrix of the loop system of equations

$$\mathsf{Z}_m(s) = \mathsf{B}_{22}\mathsf{Z}_{22}\mathsf{B}'_{22} \qquad (6\text{–}63)$$

is generally known as the *loop-impedance matrix*. Before establishing the nonsingularity of the loop-impedance matrix, the generalized node system of equations is derived. The generalized system to be derived is not the conventional node system shown in Fig. 6–7 but is the node-pair system of equations in which the variables are the branch voltages of a suitable tree of the network. In this, we follow an earlier paper [151].

Let the variables be arranged as before, with a subscript 1 denoting current drivers, a subscript 3 denoting voltage drivers, and a subscript 2 denoting the others. By general assumption (b), there exists a tree T for which the voltage drivers are branches and the current drivers are chords. Consider the current equations of the fundamental system of cut-sets for this tree T. By Theorem 6–4, these equations are equivalent to Kirchhoff's current equations. If we arrange the cut-set equations suitably, and partition after the cut-sets defined by the voltage drivers, these equations are

$$\begin{bmatrix} \mathsf{Q}_{11} & \mathsf{Q}_{12} & \mathsf{U} \\ \mathsf{Q}_{21} & \mathsf{Q}_{22} & \mathsf{0} \end{bmatrix} \begin{bmatrix} \mathsf{I}_1(s) \\ \mathsf{I}_2(s) \\ \mathsf{I}_3(s) \end{bmatrix} = \begin{bmatrix} \mathsf{0} \\ \mathsf{0} \end{bmatrix}. \tag{6–64}$$

As in the loop system, set aside the equations for the currents of the voltage drivers,

$$\mathsf{I}_3(s) = -\mathsf{Q}_{11}\mathsf{I}_1(s) - \mathsf{Q}_{12}\mathsf{I}_2(s), \tag{6–65}$$

and consider the second set

$$\mathsf{Q}_{21}\mathsf{I}_1(s) + \mathsf{Q}_{22}\mathsf{I}_2(s) = \mathsf{0}. \tag{6–66}$$

The voltage-current relations are, as before,

$$\begin{bmatrix} \mathsf{V}_1(s) \\ \mathsf{V}_2(s) \\ \mathsf{V}_3(s) \end{bmatrix} = \begin{bmatrix} \mathsf{0} & \mathsf{0} & \mathsf{0} \\ \mathsf{0} & \mathsf{Z}_{22}(s) & \mathsf{0} \\ \mathsf{0} & \mathsf{0} & \mathsf{0} \end{bmatrix} \begin{bmatrix} \mathsf{I}_1(s) \\ \mathsf{I}_2(s) \\ \mathsf{I}_3(s) \end{bmatrix} + \begin{bmatrix} \mathsf{E}_1(s) \\ \mathsf{0} \\ \mathsf{E}_3(s) \end{bmatrix} + \begin{bmatrix} \mathsf{0} \\ \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+) \\ \mathsf{0} \end{bmatrix}. \tag{6–67}$$

As before, the first and third equations in Eq. (6–67) are trivial, and the second is

$$\mathsf{V}_2(s) = \mathsf{Z}_{22}\mathsf{I}_2(s) + \frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+). \tag{6–68}$$

Following the chart of Fig. 6–7, these equations must be solved for $\mathsf{I}_2(s)$ and the solution substituted in Eq. (6–66). To this end, we must state and prove Theorem 6–11.

THEOREM 6–11. If the network satisfies Postulate N_3, and if for every element with a zero row and column in $\mathsf{Z}(s)$, either $V(s)$ or $I(s)$ is specified, then the matrix $\mathsf{Z}_{22}(s)$ of Eq. (6–68) is positive definite for positive real s and so is nonsingular. Hence det $\mathsf{Z}_{22}(s)$* is not identically zero in s.

* The determinant of $\mathsf{Z}_{22}(s)$ is a polynomial in s divided by some power of s and so will have some zeros; in fact, if each R, L, and C is taken as a separate element as is done here, the zeros are at $s = 0$ or $s = \infty$.

Proof. By suitable permutation of rows and columns, $\mathbf{Z}_{22}(s)$ can be brought to the form

$$\mathbf{Z}_{22}(s) = \begin{bmatrix} \mathbf{R} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & s\mathbf{L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{s}\mathbf{D} \end{bmatrix}. \tag{6–69}$$

By Postulate N_3, $\mathbf{R}$ is diagonal with positive diagonal entries, and so is $\mathbf{D}$. Also by Postulate N_3, $\mathbf{L}$ is positive definite. Hence:

$$\det \mathbf{Z}_{22}(s) = (\det \mathbf{R})\,(\det \mathbf{D})\,(\det \mathbf{L}) s^k, \tag{6–70}$$

where k is an integer (positive, negative, or zero). Since each of the terms of the product on the right side of Eq. (6–70) is nonzero, $\det \mathbf{Z}_{22}(s) \not\equiv 0$. The positive definiteness is immediately observed.

We remind the reader that $\mathbf{L}$ need not be diagonal; the network may contain magnetic coupling; only perfectly coupled transformers are prohibited. Since $\mathbf{Z}_{22}(s)$ is nonsingular, the inverse is defined as

$$\mathbf{Y}_{22}(s) = \mathbf{Z}_{22}^{-1}(s). \tag{6–71}$$

(In the absence of mutual coupling, $\mathbf{Y}_{22}$ is easily found by taking reciprocals of diagonal elements of $\mathbf{Z}_{22}$. In the general case, the submatrix corresponding to the coupled coils is inverted, the other elements being still the reciprocals of corresponding diagonal elements of $\mathbf{Z}_{22}$.) Now solving the voltage-current relations of Eq. (6–68) for $\mathbf{I}_2(s)$, we find that

$$\mathbf{I}_2(s) = \mathbf{Y}_{22}(s)\left[\mathbf{V}_2(s) - \frac{1}{s}\,\mathbf{v}_C(0+) + \mathbf{L}_{22}\mathbf{i}_{L2}(0+)\right]. \tag{6–72}$$

If we substitute this expression for $\mathbf{I}_2(s)$ in the cut-set current equations (6–66), the result is

$$\mathbf{Q}_{21}\mathbf{I}_1(s) + \mathbf{Q}_{22}\mathbf{Y}_{22}(s)\left[\mathbf{V}_2(s) - \frac{1}{s}\,\mathbf{v}_C(0+) + \mathbf{L}_{22}\mathbf{i}_{L2}(0+)\right] = \mathbf{0}. \tag{6–73}$$

If we separate known quantities from unknowns in this equation, we find

$$\mathbf{Q}_{22}\mathbf{Y}_{22}(s)\mathbf{V}_2(s) = -\mathbf{Q}_{21}\mathbf{I}_1(s) + \mathbf{Q}_{22}\mathbf{Y}_{22}(s)\left[\frac{1}{s}\,\mathbf{v}_C(0+) - \mathbf{L}_{22}\mathbf{i}_{L2}(0+)\right]. \tag{6–74}$$

The next step is to use the node-pair voltage transformation (see Fig. 6–7). Again, since the use of a different cut-set matrix in this transformation results in no additional generality, the same cut-set matrix as in Eq. (6–64) is used for the node-pair transformation, which is

$$\begin{bmatrix} \mathsf{V}_1(s) \\ \mathsf{V}_2(s) \\ \mathsf{V}_3(s) \end{bmatrix} = \begin{bmatrix} \mathsf{Q}'_{11} & \mathsf{Q}'_{21} \\ \mathsf{Q}'_{12} & \mathsf{Q}'_{22} \\ \mathsf{U} & \mathsf{0} \end{bmatrix} \begin{bmatrix} \mathsf{V}_{p1}(s) \\ \mathsf{V}_{p2}(s) \end{bmatrix}. \tag{6–75}$$

From the third row, we see that

$$\mathsf{V}_{p1}(s) = \mathsf{V}_3(s), \tag{6–76}$$

which again should be obvious from the choice of fundamental cut-sets. Again, the first equation,

$$\mathsf{V}_1(s) = \mathsf{Q}'_{11}\mathsf{V}_3(s) + \mathsf{Q}'_{21}\mathsf{V}_{p2}(s), \tag{6–77}$$

is reserved for later use; the second is used for the node-pair equations

$$\mathsf{V}_2(s) = \mathsf{Q}'_{12}\mathsf{V}_3(s) + \mathsf{Q}'_{22}\mathsf{V}_{p2}(s). \tag{6–78}$$

Since the node-pair transformation has been used, Kirchhoff's voltage equations need not be considered any further. Now on substituting the expression for $\mathsf{V}_2(s)$ given in Eq. (6–78) into the fundamental cut-set equations (6–74) and transposing the known $\mathsf{V}_3(s)$ to the right, we find the node-pair system of equations to be

$$\begin{aligned} &\mathsf{Q}_{22}\mathsf{Y}_{22}(s)\mathsf{Q}'_{22}\mathsf{V}_{p2}(s) \\ &\quad = -\mathsf{Q}_{22}\mathsf{Y}_{22}\mathsf{Q}'_{12}\mathsf{V}_3(s) - \mathsf{Q}_{21}\mathsf{I}_1(s) + \mathsf{Q}_{22}\mathsf{Y}_{22}\left[\frac{1}{s}\,\mathsf{v}_C(0+) - \mathsf{L}_{22}\mathsf{i}_{L2}(0+)\right]. \end{aligned} \tag{6–79}$$

With some additional computation (see Problem 6–16), we can simplify the right side by showing that

$$\mathsf{Y}_{22}\,\frac{1}{s}\,\mathsf{v}_C(0+) = \mathsf{C}_{22}\mathsf{v}_C(0+) \quad \text{and} \quad \mathsf{Y}_{22}\mathsf{L}_{22}\mathsf{i}_{L2}(0+) = \frac{1}{s}\,\mathsf{i}_{L2}(0+), \tag{6–80}$$

so that the node-pair equations can be written as

$$\begin{aligned} &\mathsf{Q}_{22}\mathsf{Y}_{22}\mathsf{Q}'_{22}\mathsf{V}_{p2}(s) \\ &\quad = -\mathsf{Q}_{22}\mathsf{Y}_{22}\mathsf{Q}'_{12}\mathsf{V}_3(s) - \mathsf{Q}_{21}\mathsf{I}_1(s) + \mathsf{Q}_{22}\left[\mathsf{C}_{22}\mathsf{v}_C(0+) - \frac{1}{s}\,\mathsf{i}_{L2}(0+)\right], \end{aligned} \tag{6–81}$$

where, as before, $\mathsf{i}_{L2}(0+)$ are the initial values of the functions $i(t)$ of the inductors only.

It would be pointless, in manual computations, to use these general loop and node-pair systems of equations developed here for solving simple, specific problems. Several simple rules (which can be derived easily from the general equations above) are known for writing these equations by inspection. One may therefore wonder whether these general systems are of any value. The derivation here has three purposes. First it is valuable to have seen, at least once, the complete and detailed justification of such common procedures as loop and node analyses. The second purpose is to establish the equations on a sufficiently firm basis to prove the nonsingularity of the coefficient matrices under the general assumptions made. The general equations are, finally, suitable for solving problems when digital computers are available.

The next lemma is needed in the proofs of the main theorem on uniqueness of solutions to the loop and node systems of equations, as well as in the later discussion of energy functions. It is a generalization of a well-known theorem [78].

LEMMA 6–12. Let $\mathbf{P}$ be a real positive definite matrix of order n. Let $\mathbf{T}$ be a real matrix of order (r, n) and rank r ($\leq n$, naturally). Then $\mathbf{TPT}'$ is positive definite. If $\mathbf{P}$ is positive semidefinite, then $\mathbf{TPT}'$ is positive definite or semidefinite.

Proof. Let $\mathbf{X}$ be a column vector of r rows, of real elements, with $\mathbf{X} \neq \mathbf{0}$. We need to show that $\mathbf{X}'\mathbf{TPT}'\mathbf{X} > 0$. To this end, define

$$\mathbf{Y} = \mathbf{T}'\mathbf{X}. \tag{6–82}$$

Then $\mathbf{Y}$ is an n-vector. Since $\mathbf{T}'$ has a rank equal to the number of columns, the equation

$$\mathbf{T}'\mathbf{X} = \mathbf{0} \tag{6–83}$$

has only the trivial solution

$$\mathbf{X} = \mathbf{0}. \tag{6–84}$$

Since $\mathbf{X} \neq \mathbf{0}$, it follows that $\mathbf{Y} \neq \mathbf{0}$. Since $\mathbf{P}$ is positive definite,

$$\mathbf{Y}'\mathbf{PY} > 0, \tag{6–85}$$

or

$$(\mathbf{T}'\mathbf{X})'\mathbf{P}(\mathbf{T}'\mathbf{X}) > 0 \qquad \text{or} \qquad \mathbf{X}'(\mathbf{TPT}')\mathbf{X} > 0. \tag{6–86}$$

Hence $\mathbf{TPT}'$ is positive definite. The rest of the theorem follows similarly.

THEOREM 6–12. If the network satisfies Postulate N_3, and the conditions

(a) there is no loop consisting only of voltage drivers,
(b) there is no cut-set consisting only of current drivers, and

(c) whenever a row and column of $\mathbf{Z}(s)$ are zeros, either $v(t)$ or $i(t)$ is specified,

then the coefficient matrices of the loop and the node-pair systems of equations are both nonsingular.

Proof. In this theorem, the driver-distribution conditions have been stated in a more natural form. It is left as a problem (Problem 6–18) to show that conditions (a) and (b) of Theorem 6–12 are equivalent to the conditions of Theorem 6–9. From Theorem 6–11 and Lemma 6–1, it is sufficient to prove that the matrices $\mathbf{B}_{22}$ and $\mathbf{Q}_{22}$ have ranks equal to the number of rows.

Remove all the current drivers from the network. Then by condition (a), the rest of the network, say N^*, is still connected and contains all the vertices of the original network. The matrix $[\mathbf{B}_{22}\ \mathbf{B}_{23}]$ is now seen to be a circuit matrix of N^*, with the proper number of rows and rank, where the submatrices $\mathbf{B}_{22}$ and $\mathbf{B}_{23}$ are the same as in Eq. (6–50). Now by condition (a) there exists a tree of N^* containing the voltage drivers, which are the elements corresponding to the columns of $\mathbf{B}_{23}$. Therefore a subset of columns of $\mathbf{B}_{22}$ corresponds to the chord set of this tree. Hence $\mathbf{B}_{22}$ contains a nonsingular submatrix of maximum order. Hence $\mathbf{B}_{22}$ has a rank equal to the number of its rows.

The cut-set matrix of N^* is obtained by simply deleting the first column of the partitioned matrix of Eq. (6–64). The cut-set matrix of N^* is therefore

$$\begin{bmatrix} \mathbf{Q}_{12} & \mathbf{U} \\ \mathbf{Q}_{22} & \mathbf{0} \end{bmatrix}.$$

Since N^* is connected and contains all the vertices of the original network, the rank of the cut-set matrix of N^* is $v - 1$. Since the matrix above has exactly $v - 1$ rows, the rows are linearly independent. In particular, the rows of $[\mathbf{Q}_{22}\ \mathbf{0}]$, which are a subset of the rows of the cut-set matrix, are also linearly independent. The matrix $[\mathbf{Q}_{22}\ \mathbf{0}]$ therefore contains a nonsingular submatrix of maximum order, which evidently has to be contained in $\mathbf{Q}_{22}$. Thus $\mathbf{Q}_{22}$ has a rank equal to the number of its rows, and so the proof is complete.

To complete the argument of network analysis, one should prove at this point that the functions $\mathbf{I}_{m2}(s)$ and $\mathbf{V}_{p2}(s)$, as well as the other functions defined in terms of these, are Laplace transforms; that is, since they are rational functions or the product of transforms ($\mathbf{I}_1$ and $\mathbf{V}_3$) by rational functions, one should prove that the degree of the numerator polynomial is lower than the degree of the denominator or, in the second case, at most equal to the degree of the denominator polynomial. Further, one should

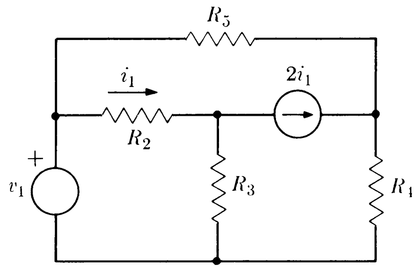


FIG. 6–8. Active network.

show that the time functions so obtained satisfy the initial conditions specified. It is well known that these existence theorems are not true unless the initial conditions satisfy certain requirements. (Otherwise the so-called *impulse* functions are present in the solution and the initial conditions are not satisfied.) Unfortunately, no such proof is available. Therefore these existence theorems are left as unsolved problems.

Before concluding the discussion of the loop and node systems of equations, it is in order to explain why the restriction "reciprocal passive" is necessary in Theorem 6–12. If the network contains dependent generators, the nonsingularity of the loop-impedance matrix (or the node-admittance matrix) is not decided, in general, purely by the topological structure; it may well depend on the values of the parameters of the network. For example, in the network of Fig. 6–8, where $2i_1$ is a dependent current generator, the loop-impedance matrix is singular if $R_2 = R_3$, and nonsingular if $R_2 \neq R_3$.

6–5 Energy functions and stability. The main purpose of this section is to show that a passive network satisfying Postulate N_3 is stable and thereby, in a sense, to justify N_3; namely, we wish to prove the statement made about the positive definiteness of $\mathbf{L}$. The energy functions of the network are defined for this purpose and for later use in Chapter 8. Many of the concepts and methods of proof used in this section are due to Bode [12].

DEFINITION 6–2. *Stable, strongly stable.* The system represented by the set of ordinary integrodifferential equations with constant coefficients

$$\mathbf{P}(p)\mathbf{X}(t) = \mathbf{Y}(t) \tag{6–87}$$

[where $p = d/dt$, $1/p = \int dt$, $\mathbf{P}$ is a matrix of polynomials in p and $1/p$, and $\mathbf{Y}(t)$ is the matrix of (known) driving functions] is *stable* if det $\mathbf{P}(s)$ has no zeros in Re $(s) > 0$, and is *strongly stable* if all the zeros of det $\mathbf{P}(s)$ are in Re $(s) < 0$.

For the conventional definition of stability (in terms of the solutions of the homogeneous equation, or the *transient* solution), one must add the stipulation that $\mathbf{P}^{-1}(s)$ has at most simple poles on the imaginary s-axis. The next theorem shows that a passive network satisfying Postulate N_3 is stable according to Definition 6–2.

THEOREM 6–13. The determinants of the loop-impedance matrix

$$\mathbf{Z}_m(s) = \mathbf{B}_{22}\mathbf{Z}_{22}(s)\mathbf{B}'_{22} \tag{6–88a}$$

and the node-pair admittance matrix

$$\mathbf{Y}_p(s) = \mathbf{Q}_{22}\mathbf{Y}_{22}(s)\mathbf{Q}'_{22} \tag{6–88b}$$

of a network satisfying Postulate N_3 have no zeros inside the right half-plane Re $(s) > 0$.

Proof. For convenience, define

$$\mathbf{R}_m = \mathbf{B}_{22}\mathbf{R}_{22}\mathbf{B}'_{22}, \qquad \mathbf{L}_m = \mathbf{B}_{22}\mathbf{L}_{22}\mathbf{B}'_{22}, \qquad \mathbf{D}_m = \mathbf{B}_{22}\mathbf{D}_{22}\mathbf{B}'_{22}. \tag{6–89}$$

By Lemma 6–1 and Theorem 6–12, the matrices $\mathbf{R}_m$, $\mathbf{L}_m$, and $\mathbf{D}_m$ are positive semidefinite or definite; the sum of the three is positive definite. Consider the homogeneous system of loop integrodifferential equations

$$\mathbf{Z}_m(p)\mathbf{i}_m(t) = \mathbf{0} \tag{6–90}$$

or

$$\left(\mathbf{L}_m p + \mathbf{R}_m + \frac{1}{p}\,\mathbf{D}_m\right)\mathbf{i}_m(t) = \mathbf{0}. \tag{6–91}$$

As is well known, if s_k is a zero of det $\mathbf{Z}_m(s)$, then

$$\mathbf{i}_m(t) = \mathbf{i}_{mk}\epsilon^{s_k t} \tag{6–92}$$

is a solution of the homogeneous system (6–90) for a suitable matrix of (complex) constants $\mathbf{i}_{mk}$. Substituting Eq. (6–92) into the homogeneous system (6–91) and performing the indicated integration and differentiation, we find that

$$\left(s_k\mathbf{L}_m + \mathbf{R}_m + \frac{1}{s_k}\,\mathbf{D}_m\right)\mathbf{i}_{mk}\epsilon^{s_k t} = \mathbf{0}; \tag{6–93}$$

or, since the exponential function is never zero, we may divide by $\epsilon^{s_k t}$ to get

$$\left(s_k\mathbf{L}_m + \mathbf{R}_m + \frac{1}{s_k}\,\mathbf{D}_m\right)\mathbf{i}_{mk} = \mathbf{0}. \tag{6–94}$$

Premultiply this equation by $\mathbf{i}^{*\prime}_{mk}$ (the transposed conjugate of $\mathbf{i}_{mk}$) to convert it into a scalar equation with quadratic forms:

$$(\mathbf{i}_{mk}^{*\prime}\mathbf{L}_m\mathbf{i}_{mk})s_k + (\mathbf{i}_{mk}^{*\prime}\mathbf{R}_m\mathbf{i}_{mk}) + (\mathbf{i}_{mk}^{*\prime}\mathbf{D}_m\mathbf{i}_{mk})\,\frac{1}{s_k} = 0. \tag{6–95}$$

The quadratic forms in this equation are positive definite or semidefinite. Hence when we multiply through by s_k, Eq. (6–95) becomes a quadratic equation in s_k with real nonnegative coefficients. Such an equation has no solutions with $\mathrm{Re}\,(s_k) > 0$. Hence the result is established. The proof for det $\mathbf{Y}_p(s)$ is similar.

The quadratic forms in Eq. (6–95) are known as energy functions for the following reasons. Taking the quadratic form of $\mathbf{R}_m$, for example, we see that

$$\mathbf{i}_{mk}^{*\prime}\mathbf{R}_m\mathbf{i}_{mk} = \mathbf{i}_{mk}^{*\prime}(\mathbf{B}_{22}\mathbf{R}_{22}\mathbf{B}_{22}^{\prime})\mathbf{i}_{mk} = (\mathbf{B}_{22}^{\prime}\mathbf{i}_{mk})^{*\prime}\mathbf{R}_{22}(\mathbf{B}_{22}^{\prime}\mathbf{i}_{mk}). \tag{6–96}$$

But

$$\mathbf{B}_{22}^{\prime}\mathbf{i}_{mk} = \mathbf{i}_2, \tag{6–97}$$

where $\mathbf{i}_2$ are the element currents for this set of loop currents, from the mesh transformation. Hence

$$\mathbf{i}_{mk}^{*\prime}\mathbf{R}_m\mathbf{i}_{mk} = \mathbf{i}_2^{*\prime}\mathbf{R}_{22}\mathbf{i}_2 = \sum_{j=1}^{e_2} R_j|i_j|^2, \tag{6–98}$$

the last step following from the fact that $\mathbf{R}_{22}$ is a diagonal matrix. The reason for the name *energy function* is now quite clear. This name is used also when the variables of the quadratic forms are not complex numbers but are Laplace transforms of the current functions. The definition of energy functions is extended by Definition 6–3 in the general case of the transformed loop equations,

$$\mathbf{Z}_m(s)\mathbf{I}_m(s) = \mathbf{E}_m(s) + \mathbf{K}(s, 0+), \tag{6–99}$$

where $\mathbf{K}$ stands for the matrix of initial values.

DEFINITION 6–3. *Energy functions.* The energy functions of an electrical network are

$$F(s) = \mathbf{I}_m^{*\prime}(s)\mathbf{R}_m\mathbf{I}_m(s), \qquad T(s) = \mathbf{I}_m^{*\prime}(s)\mathbf{L}_m\mathbf{I}_m(s), \qquad V(s) = \mathbf{I}_m^{*\prime}(s)\mathbf{D}_m\mathbf{I}_m(s). \tag{6–100}$$

We make use of this definition in the next section to get an expression for the driving-point impedance in terms of the energy functions. We turn next to the "justification" of Postulate N_3.

THEOREM 6–14. If a set of inductors can be found such that the inductance matrix $\mathbf{L}$ of these inductors is neither positive definite nor semidefinite, then an unstable passive network *can be constructed* consisting of some of these inductors and some positive resistors.

Proof. Let $L_1, L_2, \ldots, L_n$ be a set of n inductors such that the matrix of these inductances,

$$\mathsf{L} = \begin{bmatrix} \mathsf{L}_{11} & \mathsf{L}_{12} & \cdots & \mathsf{L}_{1n} \\ \mathsf{L}_{21} & \mathsf{L}_{22} & \cdots & \mathsf{L}_{2n} \\ \vdots & & & \\ \mathsf{L}_{n1} & \mathsf{L}_{n2} & \cdots & \mathsf{L}_{nn} \end{bmatrix}, \tag{6–101}$$

is neither positive definite nor semidefinite. Then there must be at least one principal minor (not necessarily a leading principal minor) which must be negative. For convenience of notation, let this be the leading principal minor of order k. The network is now constructed as follows. Connect a 1-ohm resistor across each of $L_1, L_2, \ldots, L_k$. Leave the others open-circuited as in Fig. 6–9.

Let us now write loop equations for this network, choosing loop references agreeing with the reference marks for the inductors [those used in the construction of the matrix L of Eq. (6–101)]; the loop-impedance matrix becomes

$$\mathsf{Z}_m(s) = \begin{bmatrix} sL_{11} + 1 & sL_{12} & \cdots & sL_{1k} \\ sL_{12} & sL_{22} + 1 & \cdots & sL_{2k} \\ \vdots & & & \\ sL_{1k} & sL_{2k} & \cdots & sL_{kk} + 1 \end{bmatrix}. \tag{6–102}$$

Let

$$\Delta(s) = \det Z_m(s). \tag{6–103}$$

Then $\Delta(s)$ is a polynomial of degree k in s. By the usual rule for adding determinants, $\Delta(s)$ can be expanded as

$$\Delta(s) = \begin{bmatrix} L_{11}s & L_{12}s & \cdots & L_{1k}s \\ L_{12}s & L_{22}s & \cdots & L_{2k}s \\ \vdots & & & \\ L_{1k}s & L_{2k}s & \cdots & L_{kk}s \end{bmatrix} + \begin{bmatrix} 1 & L_{12}s & L_{13}s & \cdots & L_{1k}s \\ 0 & L_{22}s & L_{23}s & \cdots & L_{2k}s \\ \vdots & & & & \\ 0 & L_{2k}s & L_{3k}s & \cdots & L_{kk}s \end{bmatrix} + \cdots$$

$$\cdots + \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{6–104}$$

It is observed from Eq. (6–104) that the coefficient of s^k is given by

$$a_k = \det \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1k} \\ L_{12} & L_{22} & \cdots & L_{2k} \\ \vdots & & & \\ L_{1k} & L_{2k} & \cdots & L_{kk} \end{bmatrix}, \tag{6–105}$$

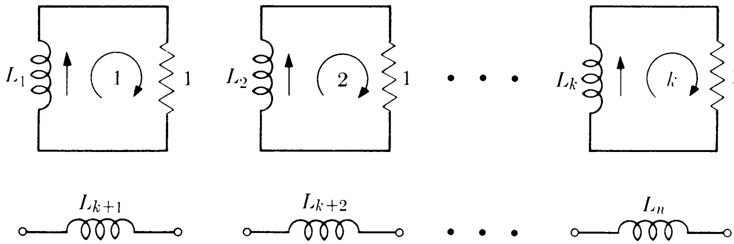


FIG. 6-9. Unstable "passive" network.

which is negative by hypothesis. The constant term is given by the last term in Eq. (6-104), which is 1. Hence

$$\Delta(s) = a_k s^k + a_{k-1} s^{k-1} + \cdots + a_1 s + 1. \tag{6-106}$$

Thus for sufficiently large real positive s,

$$\Delta(\sigma) < 0. \tag{6-107a}$$

But

$$\Delta(0) = 1 > 0. \tag{6-107b}$$

Being a polynomial and thus continuous, $\Delta(s)$ must pass through zero somewhere on the positive real axis. Thus $\Delta(s)$ has a zero in the right half-plane and the network is unstable.

Certainly, one does not expect a physical network such as that in Fig. 6-9 to go up in smoke if it contains no generators. Thus Postulate N_3 is justified.

When only two inductors are coupled, positive definiteness is the same as the condition that the coefficient of coupling be less than 1; that is,

$$L_{11}L_{22} - L_{12}^2 \geq 0. \tag{6-108}$$

But if more than two inductors are coupled, positive definiteness is a stronger requirement. For instance, the matrix

$$\mathsf{L} = \begin{bmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.2 \\ 0.9 & 0.2 & 1 \end{bmatrix} \tag{6-109}$$

is not positive definite or semidefinite, and so is unrealizable, even though all the "coefficients of coupling" are less than 1. (See Tokad and Reed [174].)

6–6 Dual networks. We conclude the discussion of network analysis with a treatment of the *dual* in both one and two terminal-pair networks. We assume a general familiarity with such networks and the derivations of the various describing functions from the loop and node systems of equations. (For the derivations, see [156].) Only the definitions of these functions are given here.

DEFINITION 6–4. *Driving-point impedance* and *driving-point admittance.* Let N be an electrical network not containing any (independent) generators, and let two vertices $(1, 1')$ of N be designated as input vertices. Then the ratio of the transform of $v_1(t)$ to the transform of $i_1(t)$, with references as shown in Fig. 6–10, under zero initial conditions, is the *driving-point impedance* at $(1, 1')$:

$$Z_d(s) = \left.\frac{V_1(s)}{I_1(s)}\right|_{\text{all initial conditions zero}} \tag{6–110}$$

The reciprocal of $Z_d(s)$ is the *driving-point admittance:*

$$Y_d(s) = \left.\frac{I_1(s)}{V_1(s)}\right|_{\text{all initial conditions zero}} \tag{6–111}$$

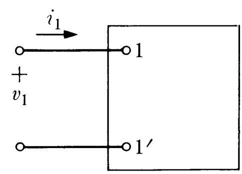


FIG. 6–10. Driving-point functions.

[It may happen that for the given $v_1(t)$ there is no solution $i_1(t)$ under zero initial conditions. The definition refers merely to the formal procedure.]

DEFINITION 6–5. *Dual of a one terminal-pair network.* The *dual* of a given planar one terminal-pair electrical network without generators or transformers is the one terminal-pair dual of the corresponding graph (Definition 3–12) with the element-impedance matrix $\mathbf{Z}(s)$ of either network being the element-admittance matrix $\mathbf{Y}(s)$ of the other network.

The requirement in the definition on the corresponding elements is the usual replacement of an inductor by a capacitor of equal value and vice versa, and the replacement of a resistor by another of reciprocal value.

THEOREM 6–15. If N and N^* are dual one terminal-pair networks, then the driving-point impedance of either network is equal to the driving-point admittance of the other; that is,

$$Z_d(s) = Y_d^*(s) \qquad \text{and} \qquad Y_d(s) = Z_d^*(s). \tag{6–112}$$

Proof. Let e_0 and e_0^* be the edges connected across the input vertices of N and N^* respectively. Then by the corollary to Theorem 4–25, the incidence matrix of $N + e_0$ is the circuit matrix of $N^* + e_0^*$ and conversely; that is,

$$\mathbf{A} = \mathbf{B}^* \qquad \text{and} \qquad \mathbf{B} = \mathbf{A}^*. \tag{6–113}$$

Further, by Definition 6–4,

$$\mathbf{Z} = \mathbf{Y}^* \qquad \text{and} \qquad \mathbf{Y} = \mathbf{Z}^*. \tag{6–114}$$

Hence the loop-impedance matrix of either network is the node-admittance matrix of the other:

$$\mathbf{Z}_m = \mathbf{Y}_n^* \qquad \text{and} \qquad \mathbf{Y}_n = \mathbf{Z}_m^*. \tag{6–115}$$

Since the matrices are equal, so are the determinants and cofactors. The rest follows from the usual formulas for $Z_d(s)$ and $Y_d(s)$ with the reference conventions of Fig. 6–11:

$$Z_d(s) = \left.\frac{\Delta(s)}{\Delta_{11}(s)}\right|_z \tag{6–116a}$$

and

$$Y_d^*(s) = \left.\frac{\Delta^*(s)}{\Delta_{11}^*(s)}\right|_y, \tag{6–116b}$$

where z indicates that the determinant and cofactor are chosen from the loop-impedance matrix, and y similarly refers to the node-admittance matrix.

Two networks satisfying the reciprocal relationship of Eq. (6–112) are generally referred to as *inverse networks.* By Theorem 6–15, dual one terminal-pairs are also inverse networks. The converse, however, is not true. By the well-known results of Brune [16], every passive one terminal-pair network has an inverse, whereas only planar one terminal-pair networks without transformers have duals.

For later use (in Chapter 8), the expression for the driving-point impedance $Z_d(s)$ in terms of energy functions is developed before considering the dual of a two terminal-pair. The loop equations for the network of Fig. 6–10, with v_1 in loop 1 *only* and with the initial conditions equal to zero, are written as

$$\left(\mathbf{R}_m + s\mathbf{L}_m + \frac{1}{s}\,\mathbf{D}_m\right)\mathbf{I}_m(s) = \mathbf{E}_m(s), \tag{6–117}$$

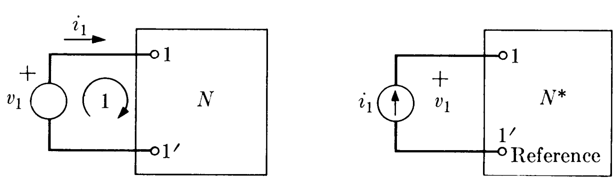


FIG. 6–11. References for duals.

where $I_{m1}(s) = I_1(s)$, $E_{m1}(s) = V_1(s)$, and all other $E_{mj}(s) = 0$, by the choice of loops. If we premultiply Eq. (6–117) by $\mathbf{I}_m^{*\prime}(s)$, the result is

$$\left[F(s) + sT(s) + \frac{1}{s}\,V(s)\right] = V_1(s) \cdot I_1^*(s), \tag{6–118}$$

with the usual notation for energy functions. Hence

$$Z_d(s) = \frac{V_1(s)}{I_1(s)} = \frac{V_1(s) \cdot I_1^*(s)}{I_1(s) \cdot I_1^*(s)} \tag{6–119}$$

can be expressed as

$$Z_d(s) = \frac{1}{|I_1|^2}\left[F(s) + sT(s) + \frac{1}{s}\,V(s)\right], \tag{6–120}$$

which is the desired expression.

Definition 6–6. *Planar two terminal-pair.* A two terminal-pair network, with (1, 1′) and (2, 2′) designated as the terminal-pairs, is *planar* if the network remains planar on adding an edge e_1 between terminals (1, 1′) and an edge e_2 between terminals (2, 2′), with e_1 and e_2 *being on the boundary of a common region,* when the network is mapped onto a plane (or onto a sphere).

Definition 6–6 is seen to be more restrictive than Definition 3–11. For example, the network of Fig. 6–12(a) is *not* a planar two terminal-pair network with the terminal-pairs (1, 1′) and (2, 2′), even though it remains planar on adding e_1 and e_2 as in Fig. 6–12(b).

Definition 6–7. *Dual of a two terminal-pair.* The dual of a planar two terminal-pair network N is obtained by finding the dual network of $N + e_1 + e_2$ (with the notation of Definition 6–6) and deleting the edges e_1^* and e_2^*, corresponding to e_1 and e_2. The terminals of e_1^* and e_2^* are the terminal-pairs of the dual N^*.

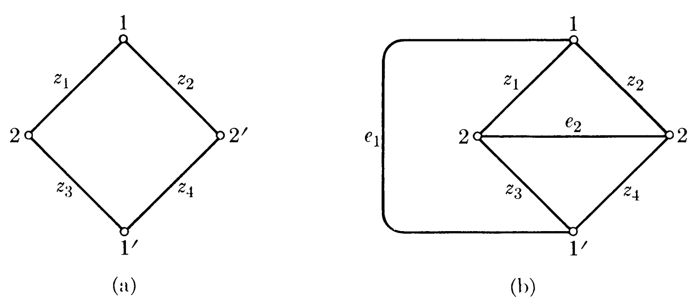


Fig. 6–12. A nonplanar two terminal-pair.

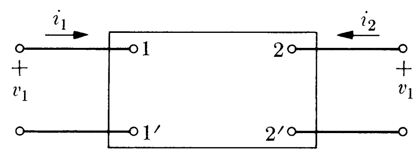


FIG. 6–13. Two terminal-pair reference convention.

A two terminal-pair network (see Fig. 6–13) is characterized by its open-circuit impedance matrix or its short-circuit admittance matrix, which are respectively the coefficient matrices in the equations

$$\begin{bmatrix} V_1(s) \\ V_2(s) \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1(s) \\ I_2(s) \end{bmatrix}$$

and (6–121)

$$\begin{bmatrix} I_1(s) \\ I_2(s) \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1(s) \\ V_2(s) \end{bmatrix},$$

with all initial conditions equal to zero, as before. The coefficient matrices are respectively denoted by $\mathbf{Z}_{oc}$ and $\mathbf{Y}_{sc}$.

THEOREM 6–16. If N and N^* are dual two terminal-pair networks, then the short-circuit admittance matrix of either network is equal to the open-circuit impedance matrix of the other.

The proof follows directly from the definition of two terminal-pair duals (Definition 6–7) on observing that the formulas for z_{ij} and y_{ij} of Eq. (6–121) are the same (in terms of the loop-impedance matrix and node-admittance matrix respectively) if the network is common-terminal, i.e., if terminals 1′ and 2′ are the same terminal. The details are left as a problem (Problem 6–24).

It is easily appreciated, by drawing the planar two terminal-pair such that e_1 and e_2 are on the boundary of the "outside" region, that the dual of a planar two terminal-pair will always be common-terminal. Since the dual of the dual of a graph G is 2-isomorphic to G, we have the next theorem.

THEOREM 6–17. A planar two terminal-pair network N is 2-isomorphic to a two terminal-pair network N_1 which is common-terminal and has the same open-circuit and short-circuit matrices.

Thus only a very restricted subclass of two terminal-pair networks have duals.

Problems

6–1. Discuss, to the fullest extent you are able, the implication of Theorem 6–2 on using less than or more than $e - v + 1$ voltage equations. Establish as many criteria as you can for choosing a set of $e - v + 1$ circuits which leads to a circuit matrix of rank $e - v + 1$.

6–2. Prove Theorem 6–3. Carry out an example as an illustration. What happens if you take a fundamental system of circuits for Kirchhoff's voltage law and a fundamental system of cut-sets for Kirchhoff's current law before solving for branch currents and link voltages?

6–3. Theorem 6–3 specifies certain restrictions on the location of driving currents and driving voltages in a network. State them.

6–4. Prove Theorem 6–4 by first showing that two homogeneous systems of algebraic equations are equivalent if and only if their matrices differ by a nonsingular transformation.

6–5. The result indicated by Eq. (6–10) implies that certain sets of currents in a network add to zero. What are these sets in terms of "looking" at a network or a graph?

6–6. Choose a fundamental system of cut-sets for an example (say Fig. 6–5). Write out the node-pair transformation using the fundamental cut-set matrix. Verify that the variables $v_{pj}(t)$ are branch voltages. Prove this result in general.

6–7. Prove Theorem 6–6.

6–8. Can we express all the element currents in terms of less than $e - v + 1$ loop currents? Justify your answer.

6–9. Prove that of the three postulates (Arsove [2])

$$\mathbf{A}\mathbf{i}(t) = \mathbf{0}, \qquad \mathbf{B}\mathbf{v}(t) = \mathbf{0}, \qquad \mathbf{i}'(t)\mathbf{v}(t) = 0,$$

any two imply the other. [*Hint:* Consider the vector subspaces $\mathcal{V}_Q$ and $\mathcal{V}_B$.]

6–10. How do we justify that $\sum p_j = 0$ in a network when we know that it takes some energy to run an electrical network?

6–11. In certain singular problems (such as two identical capacitors charged to different voltages and connected in parallel), we know that energy in the network is not conserved. What happens to Theorem 6–8 in these cases?

6–12. The Riemann integral $\int_0^t i(x)\,dx$ has certain mathematical properties just because it is an integral. Determine precisely what these are [particularly the possibility of discontinuity in $v_C(t)$]. Consider Problem 6–11 in view of these properties. In particular, show that a mathematical contradiction arises.

6–13. Find the expression for node voltages in terms of the branch voltages of a tree, and the expression for loop currents in terms of chord currents of a tree.

6–14. If $\mathbf{A}$ is a real symmetric matrix, show that

$$\mathbf{X}'\mathbf{A}\mathbf{X} > 0$$

for all nonzero real vectors $\mathbf{X}$ (that is, $\mathbf{A}$ is positive definite for real $\mathbf{X}$) if and only if

$$\mathbf{Y}^{*\prime}\mathbf{A}\mathbf{Y} > 0$$

for all nonzero complex vectors $\mathbf{Y}$.

6–15. By reducing the coefficient matrix of Eq. (6–49) to the triangular form, show that the driver conditions of Theorem 6–9 are also *sufficient* to ensure the unique solvability of the network equations (6–46). [*Hint:* Use fundamental systems of circuits and cut-sets for Kirchhoff's voltage and current equations, both for the same tree. Partition the matrix further according to chords and branches of this tree. Use Theorem 5–9. The problem is somewhat lengthy.]

6–16. Show that

$$\frac{1}{s}\,\mathbf{Y}_{22}\mathbf{v}_C(0+) = \mathbf{C}_{22}\mathbf{v}_C(0+) \quad \text{and} \quad \mathbf{Y}_{22}\mathbf{L}_{22}\mathbf{i}_{L2}(0+) = \frac{1}{s}\,\mathbf{i}_{L2}(0+). \tag{6–80}$$

6–17. Work out the details of the derivation of the simplified loop and node systems of the equations shown in Fig. 6–7.

6–18. Show that conditions (a) and (b) of Theorem 6–12 are equivalent to the driver conditions of Theorem 6–9.

6–19. What would be wrong in writing 5-loop equations for a 6-loop network and solving them? Try an example such as Fig. 6–14.

6–20. What happens if you choose more than $e - v + 1$ loop equations?

6–21. Show that the zeros of the network determinants for *LC*, *RL*, and *RC* networks are on $j\omega$-, $-\sigma$-, and $-\sigma$-axes respectively. [*Hint:* Proof of Theorem 6–13.]

6–22. Develop the expression for $Y_d(s)$ in terms of the node equations, into a form analogous to the energy-function expression for $Z_d(s)$ of Eq. (6–120).

6–23. Prove that if a network contains a tree such that all chords of the tree are resistors, then the network is strongly stable according to Definition 6–2. [*Hint:* Proof of Theorem 6–13.]

6–24. Complete the details of the proof of Theorem 6–16.

6–25. In Theorem 6–17, show that the open-circuit impedance matrices of N and N_1 are identical.

6–26. Find the one terminal-pair duals of Figs. 7–1, 7–21, and 8–5.

6–27. Examine the networks of Figs. 7–9 and 7–22 to determine whether they have two terminal-pair duals. Find the duals when they exist.

6–28. Give a simple reasoning, with an example, to show why two terminal-pair networks that merely remain planar when the source and load are added may not have two terminal-pair duals (Fig. 6–12, for example).

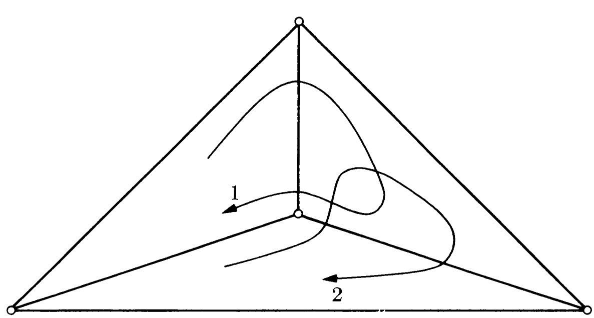


FIGURE 6–14

CHAPTER 7

TOPOLOGICAL FORMULAS

The name *topological formulas* is applied to the formulas for writing certain classes of network functions (driving-point and transfer functions) by inspection of the network diagram without actually expanding various determinants and cofactors. As such, these formulas have applications to both network analysis and network synthesis. In analysis, topological formulas provide a short-cut method of evaluating network determinants and cofactors because the usual cancellations inherent in evaluation of determinants are avoided. The recent increase in interest in topological formulas is due mainly to this fact, since the evaluation of determinants (especially with polynomial entries) by conventional procedures is a time-consuming operation when digital computers are used. The application of topological formulas to network synthesis has barely begun (as mentioned in Section 5–5). In synthesis, the main virtue of topological formulas, so far, has been to provide a certain intuition. It seems virtually certain that future applications in network synthesis will entrench topological formulas much more firmly in the field than any computer method of analysis.

Like many other topics discussed in this book, the basic concepts of topological formulas are not new; they date back to Kirchhoff (1847) and Maxwell (1892). The application to active networks is, however, very recent (1957). The discussion in this chapter is restricted to the basic formulas for network functions. For detailed discussions of all the variations and ramifications, the reader is referred to Mayeda and Seshu [109] for passive networks and to Mayeda [111] or Coates [36] for active networks. Although it is possible to obtain the general formulas for active networks directly and treat passive networks as a special case, passive networks without mutual inductances are considered first in the following discussion. The formulas for the special case (passive networks without mutual inductances) are much the simplest, and the special case is sufficiently important to be considered separately.

7–1 Node determinant and cofactors. The network under consideration in this section is assumed to be passive and without mutual inductance. The topological formulas for such networks, in terms of the admittances of the network elements, were first given by Maxwell [108]. The first formal proof for the node-determinant formula was given by Brooks, Smith, Stone, and Tutte [14]. The most recent interest in these formulas started with Percival [129].

All topological formulas depend on a theorem of matrix theory known as the *Binet-Cauchy theorem*, which may be stated as follows.

BINET-CAUCHY THEOREM. If $\mathbf{P}$ of order (m, n) and $\mathbf{Q}$ of order (n, m) are matrices of elements from a field $(m \leq n)$,

$$\det \mathbf{PQ} = \sum \begin{pmatrix} \text{products of corresponding} \\ \text{major determinants of } \mathbf{P} \text{ and } \mathbf{Q} \end{pmatrix}, \tag{7-1}$$

where the summation is over all such major determinants.

The *major determinant* (or briefly, *major*) referred to in the theorem is a determinant of order m, since $\mathbf{P}$ is of order (m, n). The word *corresponding* implies the following. If columns $j_1, j_2, \ldots, j_m$ of $\mathbf{P}$ constitute the major of $\mathbf{P}$ that is chosen, the corresponding major of $\mathbf{Q}$ consists of rows $j_1, j_2, \ldots, j_m$ of $\mathbf{Q}$. The proof of the theorem may be found in Hohn [78]. As an illustration of the theorem, let

$$\mathbf{P} = \begin{bmatrix} 1 & -1 & 3 \\ 2 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{Q} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}. \tag{7-2}$$

There are three major determinants (of order 2) to be considered. Applying the Binet-Cauchy theorem, we find that

$$\begin{aligned} \det(\mathbf{PQ}) &= \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} + \begin{vmatrix} -1 & 3 \\ 1 & 0 \end{vmatrix} \cdot \begin{vmatrix} -1 & 1 \\ 1 & 0 \end{vmatrix} \\ &= 3 \cdot 3 + (-6)(-1) + (-3)(-1) = 18, \end{aligned} \tag{7-3}$$

which result can be verified by computing $\mathbf{PQ}$ and finding its determinant directly.

DEFINITION 7-1. *Tree-admittance product.* A *tree-admittance product* is the product of the admittances of the branches of a tree.

THEOREM 7-1. The determinant Δ_n of the node-admittance matrix $\mathbf{Y}_n$ of a passive network N without mutual inductances is

$$\Delta_n = \sum_{\substack{\text{all} \\ \text{trees}}} \begin{pmatrix} \text{tree-admittance product} \\ \text{of tree } t_i \text{ of } N \end{pmatrix}. \tag{7-4}$$

Proof. Since N contains no generators, we may write

$$\mathbf{Y}_n = \mathbf{AYA}', \tag{7-5}$$

using the notation of Chapter 6. Since N contains no mutual inductances,

Y is diagonal:

$$\mathsf{Y} = \begin{bmatrix} y_1 & 0 & 0 & \cdots & 0 \\ 0 & y_2 & 0 & \cdots & 0 \\ 0 & 0 & y_3 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & y_e \end{bmatrix}. \tag{7-6}$$

Hence the product AY differs from A only in that column i is multiplied by y_i for $i = 1, 2, \ldots, e$. (The two matrices AY and A have the same structure otherwise.) By the Binet-Cauchy theorem,

$$\Delta_n = \det \mathsf{Y}_n = \sum \begin{pmatrix} \text{products of corresponding} \\ \text{majors of } \mathsf{AY} \text{ and } \mathsf{A}' \end{pmatrix}. \tag{7-7}$$

By Theorems 5–6 and 5–7, the nonzero majors of A correspond to trees and have the values (± 1). Hence the nonzero majors of AY also correspond to trees and have the values $(\pm 1)\ y_{i_1}y_{i_2}\cdots y_{i_{v-1}}$, where elements i_1, $i_2, \ldots i_{v-1}$ constitute a tree of N. Since the corresponding major of A' is the transpose of the major of A, the two majors have the same value (1 or -1). The theorem now follows immediately.

COROLLARY 7–1. The node determinant Δ_n of a connected network containing no mutual inductances is a homogeneous polynomial of degree $v - 1$ in the variables $y_1, y_2, \ldots, y_e$ and is a linear function of any one y_j.

The second part of the corollary is not true for mutual inductances.

Let us illustrate the theorem by means of an example. For the network of Fig. 7–1, there are eight trees, consisting of the edges

(123), (124), (134), (135), (145), (234), (235), (245).

By Theorem 7–1, the node determinant is therefore

$$\begin{aligned}\Delta_n &= y_1y_2y_3 + y_1y_2y_4 + y_1y_3y_4 + y_1y_3y_5 \\ &\quad + y_1y_4y_5 + y_2y_3y_4 + y_2y_3y_5 + y_2y_4y_5. \end{aligned}\tag{7-8}$$

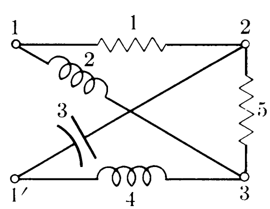


FIG. 7–1. Example for Maxwell's rule.

By substituting the admittances, we find

$$\Delta_n = \frac{G_1C_3}{L_2} + \frac{G_1}{s^2L_2L_4} + \frac{G_1C_3}{L_4} + (G_1C_3G_5)s$$
$$+ \frac{G_1G_5}{L_4s} + \frac{C_3}{L_2L_4s} + \frac{G_5C_3}{L_2} + \frac{G_5}{L_2L_4s^2}; \tag{7-9}$$

or, by collecting coefficients,

$$\Delta_n = (G_1C_3G_5)s + \frac{G_1C_3}{L_2} + \frac{G_1C_3}{L_4} + \frac{G_5C_3}{L_2}$$
$$+ \left[\frac{G_1G_5}{L_4} + \frac{C_3}{L_2L_4}\right]\frac{1}{s} + \left[\frac{G_1 + G_5}{L_2L_4}\right]\frac{1}{s^2}. \tag{7-10}$$

With some experience, one can write the last step directly. Two important facts should be noted. First, the node-admittance matrix need not be written. Its determinant is found directly. Second, there was no cancellation, and so no unnecessary work has been done. It is necessary only to write the node-admittance matrix for Fig. 7-1 and compute its determinant to appreciate this fact.

Finally, note that the determinant of the node-admittance matrix is independent of the reference node, a fact that can also be proved directly [162].

Maxwell originated the topological formula for the node determinant in the form:

> *Δ_n is the sum of products of conductivities taken $v - 1$ at a time, omitting all those terms which contain products of the conductivities of branches which form closed circuits.*

For convenience, we use the shorthand notation

$$V(Y) = \sum \text{(tree-admittance products)}. \tag{7-11}$$

This expression (in terms of y_j's) is known as the *node discriminant* [59].

Let us next investigate the cofactor of an element on the main diagonal of $\mathbf{Y}_n(s)$. The cofactor of an element in the (i, i)-position is obtained by deleting the ith row and ith column of the matrix $\mathbf{Y}_n(s)$ and taking the determinant of the resultant matrix. Since

$$\mathbf{Y}_n(s) = \mathbf{AYA}', \tag{7-12}$$

deleting the ith row from $\mathbf{Y}_n(s)$ is equivalent to deleting the ith row from $\mathbf{A}$. Let $\mathbf{A}_{-i}$ be the matrix obtained by deleting the ith row from $\mathbf{A}$. Similarly, deleting the ith column from $\mathbf{Y}_n(s)$ is equivalent to deleting the ith column from $\mathbf{A}'$, that is, deleting the ith row from $\mathbf{A}$. Thus the cofactor of the (i, i)-element is

$$\Delta_{ii} = \det \mathbf{A}_{-i}\mathbf{Y}\mathbf{A}'_{-i}. \tag{7–13}$$

Exactly the same technique that was applied to Δ_n can be applied to Δ_{ii}. However, it is more instructive to construct the graph for which Δ_{ii} is the node determinant and apply Theorem 7–1. Let the ith vertex of the network N be shorted to the reference vertex. If this new combined vertex is used as the reference vertex, the node-admittance matrix of the new network N_1 is precisely

$$\mathbf{Y}_{n1} = \mathbf{A}_{-i}\mathbf{Y}\mathbf{A}'_{-i}. \tag{7–14}$$

Thus Δ_{ii} is simply the sum of tree products for the graph obtained by identifying the ith vertex with the reference vertex.

Let us examine the subgraphs of the network N which become the trees of the network N_1 (ith vertex and reference identified) so that we may extend the formula to the case of asymmetrical minors. N_1 contains $v - 1$ vertices, and so a tree of N_1 contains $v - 2$ elements. The subgraph of N corresponding to such a tree of N_1 will not contain any circuits. However, since it contains only $v - 2$ elements, it will not be connected; it will be in two connected parts. One of the two parts may consist of an isolated vertex. The vertex i and the reference vertex will be in two different connected parts of this subgraph, in N. (If they were in the same connected part, shorting the ith vertex with the reference vertex would produce a circuit.) Such a geometrical configuration has been named a *2-tree* by Percival [129].

Definition 7–2. *2-tree.* A *2-tree* is a pair of unconnected, circuitless subgraphs, each subgraph being connected, which together include all the vertices of the graph. One (or, in trivial graphs, both) of the subgraphs may consist of an isolated vertex.

The symbol T_2 will be used for a 2-tree. Very often, 2-trees in which certain designated vertices are required to be in different connected parts are used. Then subscripts are used to denote such 2-trees. For example, $T_{2_{ab,cde}}$ is the symbol for a 2-tree in which the vertices a and b are in one connected part and the vertices c, d, and e are in the other connected part.

Definition 7–3. *2-tree product.* A *2-tree product* is the product of the admittances of the branches of a 2-tree. Again, one of the two parts may be an isolated vertex. The product for an isolated vertex is defined to be 1. A 2-tree such as $T_{2_{i,i}}$, in which the same vertex i is required to be in different connected parts, has by definition a zero product.

A sum of 2-tree products such as occurs in the expansion of a symmetrical cofactor of the node-admittance matrix is symbolized by $W(Y)$ with subscripts denoting any special vertices which are required to be in different parts. In terms of 2-trees, the formula for the symmetrical cofactor can be expressed as in Theorem 7–2:

THEOREM 7–2. If r is the reference vertex of node equations, the cofactor of an element in the (i, i)-position is given by

$$\Delta_{ii} = \sum_{\text{all 2-trees}} (T_{2_{i,r}} \text{ products}); \tag{7–15a}$$

that is,

$$\Delta_{ii} = W_{i,r}(Y). \tag{7–15b}$$

As an example, let us find the cofactor Δ_{11} for the network of Fig. 7–1, with $1'$ as the reference vertex. For illustrative purposes, the 2-trees $T_{2_{1,1'}}$ of Fig. 7–1 are shown in Fig. 7–2. Note that in some of these 2-trees, either vertex 1 or $1'$ appears as an isolated vertex. Now

$$\begin{aligned}\Delta_{11} &= W_{1,1'}(Y) = \sum (T_{2_{1,1'}} \text{ products}) \\ &= (G_5C_3)s + G_1G_5 + \frac{C_3}{L_2} + \frac{C_3}{L_4} \\ &+ \frac{G_1/L_2 + G_1/L_4 + G_5/L_2 + G_5/L_4}{s}. \end{aligned} \tag{7–16}$$

Again note the absence of any cancellation, which leads to maximum efficiency of computation.

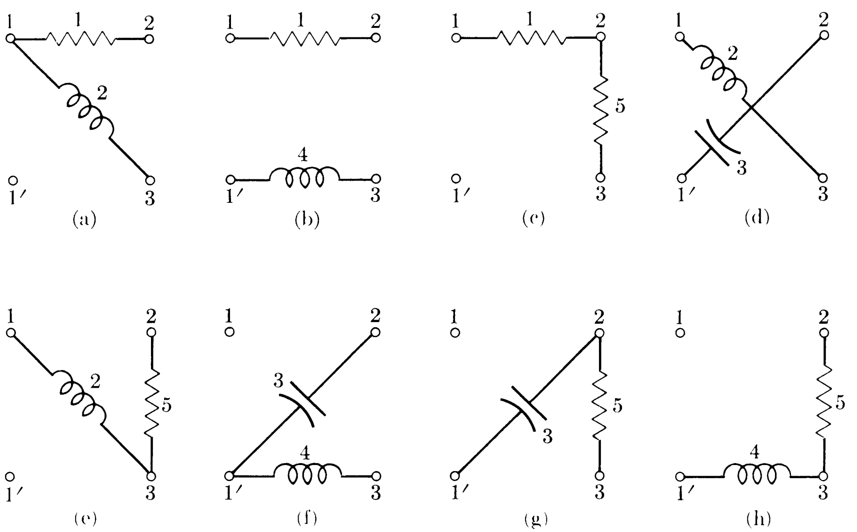


FIG. 7–2. 2-trees (1, 1′) of Fig. 7–1.

Asymmetrical cofactors of the node-admittance matrix are considered next.

THEOREM 7-3. Let $1'$ be the reference vertex of a system of node equations for a network which contains no magnetic coupling. Then the cofactor of an element in the (i, j)-position is given by

$$\Delta_{ij} = W_{ij,1'}(Y) = \sum (T_{2_{ij,1'}} \text{ products}), \tag{7-17}$$

where the summation is over all the 2-trees with vertices i and j in one connected part and vertex $1'$ in the other.

*Proof.** The cofactor of an element in the (i, j)-position is given by

$$\Delta_{ij} = (-1)^{1+j} M_{ij}, \tag{7-18}$$

where M_{ij} is the determinant of a matrix obtained by deleting the ith row and jth column from the node-admittance matrix Y_n. Hence

$$M_{ij} = \det \mathsf{A}_{-i}\mathsf{Y}\mathsf{A}'_{-j}, \tag{7-19}$$

where, as before, the subscript indicates the row which has been deleted from the incidence matrix. As in the case of the symmetrical minors, observe that the nonzero majors of the matrix A_{-i} correspond one-to-one to the 2-trees of the network which have the vertex i in one connected part and the vertex $1'$ in the other. Similarly, the nonzero majors of the matrix A_{-j} are in one-to-one correspondence with the 2-trees of the network which have the vertex j in one connected part and the vertex $1'$ in the other. Using the Binet-Cauchy theorem once again, we find that

$$M_{ij} = \sum \begin{pmatrix} \text{products of corresponding} \\ \text{majors of } \mathsf{A}_{-i}\mathsf{Y} \text{ and } \mathsf{A}'_{-j} \end{pmatrix}. \tag{7-20}$$

As before, the matrix product $\mathsf{A}_{-i}\mathsf{Y}$ differs from A_{-i} only in that the pth column is multiplied by y_p, $p = 1, 2, \ldots, e$. Thus a nonzero major of $\mathsf{A}_{-i}\mathsf{Y}$ is (except possibly for sign) a 2-tree product of a 2-tree $T_{2_{i,1'}}$. Similarly the nonzero majors of A'_{-j} correspond to the 2-trees $T_{2_{j,1'}}$. The product of the corresponding majors on the right side of Eq. (7-20) is therefore nonzero only when the set of edges (corresponding to the columns of A_{-i} and A_{-j}) constitute a 2-tree $T_{2_{i,1'}}$ as well as a 2-tree $T_{2_{j,1'}}$. Thus the nonzero products correspond to 2-trees in which both the vertices i and j are in one connected part and the vertex $1'$ is in the other, i.e., subgraphs which are 2-trees, $T_{2_{ij,1'}}$.

---

* Proof first given by W. Mayeda in a term paper at the University of Illinois 1955.

To establish the sign to be prefixed to the 2-tree products, let us select from the incidence matrix A the submatrix consisting of the columns corresponding to the elements of one of the 2-trees of the type $T_{2_{ij,1'}}$. This submatrix is of order $(v - 1, v - 2)$. If we delete row i from this matrix and take the determinant, we get the major of A_{-i}, which gives the sign of the major of $\mathsf{A}_{-i}\mathsf{Y}$. If we delete row j from this matrix and take the determinant, we similarly get the sign of the major of A_{-j} and hence of the corresponding major of A'_{-j}. Each such 2-tree necessarily contains a path between the vertices i and j. Let this path from i to j consist of the edges $e_{r_1}, e_{r_2}, e_{r_3}, \ldots, e_{r_k}$ in order. In the chosen submatrix of A, the columns corresponding to these elements will have the following structure. Column r_1 will have a nonzero entry in row i. Columns r_1 and r_2 will have nonzero entries in the same row, this row being different from row i. Columns r_2 and r_3 will have nonzero entries in another common row, etc., and column r_k has a nonzero entry in row j. Let two columns which have nonzero entries in the same row be called *adjacent*, since they correspond to adjacent elements of the graph. Then, in the sequence of columns $r_1, r_2, \ldots, r_k$, successive columns are adjacent and no others are. Using these results, we now reduce the chosen submatrix of A to one in which column r_1 has nonzero entries in rows i and j and zeros in the other rows. This reduction is achieved by means of column operations only, so that the majors of A_{-i} and A_{-j} are left invariant under these operations.

Let column r_1 have a 1 in the ith row; the case in which this entry is -1 is the same and will not be considered. Column r_1 has a -1 in another row, say p. Column r_2 has a nonzero entry in this row, by the above argument. If this entry is $+1$, add column r_2 to column r_1. If this entry is -1, subtract column r_2 from column r_1. In either case, column r_1 has, after the operation, a $+1$ in row i, a -1 in another row (the row in which column r_3 has a nonzero entry) and zeros in all other rows. Next, consider columns r_1 and r_3. If the common row entries have the same sign, subtract column r_3 from column r_1; if they have opposite signs, add. Then the -1 in column r_1 is moved to a row in which column r_4 has a nonzero entry. After repeated application of this procedure, we finally arrive at a stage when the -1 appears in a row in which column r_k has a nonzero entry not adjacent to column r_{k-1}, namely row j. Now we have a matrix in which column r_1 has a 1 in row i, a -1 in row j, and zeros in all other rows. Let this final matrix be denoted by A_d.

There are two cases to consider: $i > j$ and $i < j$. The two cases are identical, and so let $i > j$.

Consider the major of A_{-i}. This major is obtained by deleting row i from the matrix A_d obtained above and taking the determinant. Expand this major by column r_1. Column r_1 has only one nonzero entry. This entry is a -1 and is now in the jth row. (Since $i > j$, the deleted row is below row j, and so the row index of row j is unaltered.) Let the determinant of the matrix obtained by deleting rows i and j and column r_1 from A_d be denoted by D. Then

$$(\text{major of } \mathsf{A}_{-i}) = (-1)^{r_1+j}(-1)D = (-1)^{r_1+j+1}D. \tag{7-21}$$

Consider the major of A_{-j}. This major is obtained by deleting row j from the matrix A_d and taking the determinant. Column r_1 of this determinant has a 1 in row $i - 1$ and zeros in all other rows. (The row index of this row has decreased by one because row j has been deleted.) Expand the determinant by column r_1. The minor obtained by deleting column r_1 and row $i - 1$ is the same determinant D that was obtained earlier. Hence,

$$(\text{major of } \mathsf{A}_{-j}) = (-1)^{r_1+i-1}(1)D. \tag{7-22}$$

Hence the product of the two majors of $\mathsf{A}_{-i}\mathsf{Y}$ and A'_{-j} is given by

$$(-1)^{2r_1+i+j}D^2(T_{2_{ij,1'}} \text{ product}),$$

which is the same as

$$(-1)^{i+j}(T_{2_{ij,1'}} \text{ product})$$

since D is either 1 or -1 because it is selected from the incidence matrix A. Note that i and j are independent of the major selected from A_{-i} and A_{-j}. Hence

$$\begin{aligned} \det \mathsf{A}_{-i}\mathsf{Y}\mathsf{A}'_{-j} &= (-1)^{i+j} \sum (T_{2_{ij,1'}} \text{ products}) \\ &= (-1)^{i+j} W_{ij,1'}(Y). \end{aligned} \tag{7-23a}$$

Hence finally,

$$\Delta_{ij} = (-1)^{i+j} \det \mathsf{A}_{-i}\mathsf{Y}\mathsf{A}'_{-j} = W_{ij,1'}(Y), \tag{7-23b}$$

and the theorem is proved.

Before proceeding further, observe that the formula for the asymmetrical cofactor contains, as a special case, the formula for a symmetrical cofactor. For, if we let $i = j$ in Theorem 7-3,

$$\Delta_{ii} = W_{ii,1'}(Y) = W_{i,1'}(Y) \tag{7-24}$$

since the vertex i is always in the same part as vertex i.

The topological formulas require that *all* the trees or 2-trees (of the given type) be located. The following two rules, due to Percival [129], and the 2-tree identities following the rules are useful for this purpose.

Rule 1. If $V_1(Y)$, $V_2(Y)$, ..., $V_k(Y)$ are the tree-admittance polynomials for the components G_1, G_2, ..., G_k of a separable graph G, then the polynomial $V(Y)$ of G is given by

$$V(Y) = V_1(Y)V_2(Y)V_3(Y)\cdots V_k(Y). \tag{7-25}$$

Rule 2. If two subgraphs G_1 and G_2 of a connected graph G have exactly two vertices i and j in common, then for G consisting of G_1 and G_2,

$$V(Y) = V_1(Y)W_{2_{i,j}}(Y) + V_2(Y)W_{1_{i,j}}(Y). \tag{7-26}$$

These two rules are so obvious that they require no proof. Rule 2 is seen to be valid by observing that every tree must contain a path between vertices i and j, either in G_1 or in G_2, but not in both. Thus, every tree consists of a tree in G_1 and a 2-tree in G_2 or vice versa. Conversely, a tree of one of the subgraphs and a 2-tree in the other, separating vertices i and j, constitute a tree of G.

Rule 2 is very useful in computation. One can first choose an element of G as G_1. If this element is y_k, with vertices i and j, then

$$V(Y) = y_kW_{i,j}(Y) + V_2(Y), \tag{7-27}$$

where W is now simply the 2-tree sum of the graph and V_2 is the sum of tree products when y_k is removed from the graph. Next, another element may be chosen for computing V_2 similarly; the process may be repeated until the polynomial can be written by inspection.

EXAMPLE. For Fig. 7-3,

$$\begin{aligned} V(Y) &= y_1W_{1,2}(Y) + V_2(Y) \\ &= y_1[y_3y_4 + y_3y_5 + y_4y_5 + y_4y_6 + y_5y_6] + V_2(Y), \end{aligned} \tag{7-28a}$$

$$V_2(Y) = y_2[y_3y_4 + y_3y_5 + y_4y_5 + y_4y_6 + y_5y_6] + V_3(Y), \tag{7-28b}$$

and

$$V_3(Y) = y_3[y_4y_6 + y_5y_6] + y_4y_5y_6. \tag{7-28c}$$

Hence

$$\begin{aligned} V(Y) = {} & (y_1 + y_2)(y_3y_4 + y_3y_5 + y_4y_5 + y_4y_6 + y_5y_6) \\ & + y_3(y_4y_6 + y_5y_6) + y_4y_5y_6. \end{aligned} \tag{7-28d}$$

The first 2-tree identity, which is self-evident, is

$$W_{i,j} = W_{j,i}, \tag{7-29}$$

since every 2-tree with vertices i and j in different parts appears in both polynomials. The second useful identity is

$$W_{i,j} = W_{i,jk} + W_{ik,j}, \tag{7-30}$$

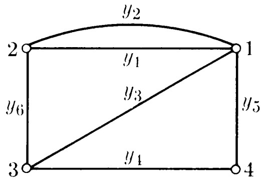


FIG. 7-3. Example for Rule 2 of Percival.

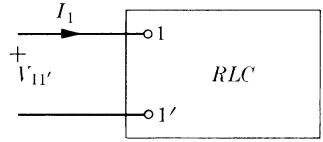


FIG. 7-4. Driving-point functions.

where k is any other vertex ($i \neq k$, $k \neq j$). This identity is seen to be true since k must be in one of the two connected parts. This equation may also be stated in the more convenient form

$$W_{i,j} - W_{ik,j} = W_{i,jk}. \tag{7-31}$$

7-2 Driving-point and transfer admittances. Figure 7-4 shows a one terminal-pair network not containing any generators. $V_{11'}$ and I_1 denote the transforms of the voltage and current respectively, with references as shown. By Definition 6-4, the driving-point admittance at terminals (1, 1′) is

$$Y_d(s) = \frac{I_1(s)}{V_{1,1'}(s)}, \tag{7-32}$$

with all initial conditions equal to zero. If node equations are written with 1′ as the reference node, then

$$Y_d(s) = \frac{\Delta}{\Delta_{11}}, \tag{7-33}$$

where Δ and Δ_{11} are the determinant and cofactor (1, 1) respectively, of the node-admittance matrix, as in Eq. (6-116b). It is important to note that the node-admittance matrix of the network of Fig. 7-4 (including I_1) is the same as the matrix $\mathsf{Y}_n(s)$ for the one terminal-pair alone, without $I_1(s)$. (If loop equations are used, the matrices with and without the driver are different.) Hence the driving-point admittance formula is obtainable directly by using Theorems 7-1 and 7-2, as we show in the next theorem.

THEOREM 7-4. For a one terminal-pair passive network which contains no magnetic coupling,

$$Y_d(s) = \frac{V(Y)}{W_{1,1'}(Y)}, \tag{7-34}$$

where 1 and 1′ are the input vertices.

Evidently, the computation of $V(Y)$ and $W_{1,1'}(Y)$ can be done without any regard to which vertex is used in writing the node equations. This is not surprising, since the driving-point admittance is certainly independent of the reference vertex chosen.

For example, the driving-point admittance of the network of Fig. 7–1 at terminals (1, 1′) is, from Eqs. (7–10) and (7–16),

$$\begin{aligned} Y_d(s) &= [(G_1C_3G_5)s^3 + (G_1C_3\Gamma_2 + G_1C_3\Gamma_4 + G_5C_3\Gamma_2)s^2 \\ &\quad + (G_1G_5\Gamma_4 + \Gamma_2\Gamma_4C_3)s + (G_1 + G_5)\Gamma_2\Gamma_4]/[G_5C_3s^3 \\ &\quad + (G_1G_5 + C_3\Gamma_2 + C_3\Gamma_4)s^2 + (G_1\Gamma_2 + G_1\Gamma_4 + G_5\Gamma_2 + G_5\Gamma_4)s], \end{aligned} \tag{7–35}$$

where $\Gamma_i = 1/L_i$.

The transfer admittance is considered next. Maxwell gave the original rule for the transfer function of a two terminal-pair network. Maxwell's rule for the current in an element between vertices r and s and oriented away from r, due to a voltage driver E with vertices p and q and with reference + at q, is

$$i_{rs} = Y_{rs}Y_{pq}\frac{\Delta_{rs,pq}}{\Delta}E. \tag{7–36}$$

In this formula, we recognize the term $\Delta_{rs,pq}$ to be the difference of cofactors selected from the node-admittance matrix. Maxwell's rule for this factor is:

> $\Delta_{rs,pq}$ is the sum of products of admittances, taken $v - 2$ at a time, omitting all the terms which contain Y_{rs} or Y_{pq} and other terms either making closed circuits with themselves or with the help of Y_{rs} and Y_{pq}. The terms which contain Y_{qr} (or which form a closed circuit with Y_{qr}) and Y_{ps} (or those forming closed circuits with Y_{ps}) are taken as positive terms, and similar terms with Y_{pr} and Y_{qs} are taken as negative terms.

First observe that, because each term contains $v - 2$ factors and does not include a circuit, each product in Maxwell's formula corresponds to a 2-tree product. Second, neither pair of vertices (p, q) or (r, s) can be in the same connected part, since the terms which form closed circuits with Y_{pq} or Y_{rs}, and those containing Y_{pq} or Y_{rs}, are to be omitted. Thus the 2-trees selected are simultaneously $T_{2_{p,q}}$ and $T_{2_{r,s}}$. Thus there are two

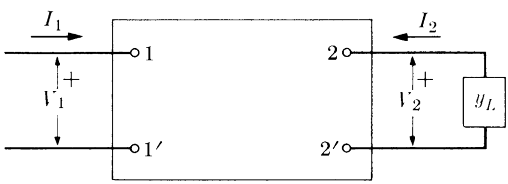


FIG. 7–5. Terminated two terminal-pair.

possible sets of 2-trees to be selected: $T_{2_{pr,qs}}$ and $T_{2_{ps,qr}}$. Maxwell affixes a positive sign to the second set of 2-trees and a negative sign to the first set. We next restate Maxwell's rule in terms of 2-trees after introducing the more common notation in the theory of two terminal-pair networks.

Let Fig. 7-5 represent a two terminal-pair network with input vertices $(1, 1')$ and output vertices $(2, 2')$. Let the references for the input current and voltage and the output current and voltage be as shown. Let y_L be a load connected across the output terminals $(2, 2')$. Let the node equations be written for this network with the vertex $1'$ as the reference vertex. These equations have the form

$$\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1,v-1} \\ y_{21} & y_{22} & \cdots & y_{2,v-1} \\ \vdots & & & \\ y_{v-1,1} & y_{v-1,2} & \cdots & y_{v-1,v-1} \end{bmatrix} \begin{bmatrix} V_{11'} \\ V_{21'} \\ \vdots \\ V_{v-1,1'} \end{bmatrix} = \begin{bmatrix} I_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{7-37}$$

when all initial conditions are zero. The output voltage $V_{22'} = V_2$ is given by

$$V_2 = \frac{\Delta_{12} - \Delta_{12'}}{\Delta} I_1. \tag{7-38}$$

Thus, Maxwell's rule above states that

$$\Delta_{12} - \Delta_{12'} = \sum (T_{2_{12,1'2'}} \text{ products}) - \sum (T_{2_{12',1'2}} \text{ products}). \tag{7-39}$$

Maxwell's formula is established by the use of the topological formula for asymmetrical cofactors (Theorem 7-3) and the 2-tree identities.

THEOREM 7-5. If $\mathbf{Y}_n$ is the node-admittance matrix of a passive network which does not contain any mutual inductances, and $1'$ is the reference node, then

$$\Delta_{12} - \Delta_{12'} = W_{12,1'2'}(Y) - W_{12',1'2}(Y). \tag{7-40}$$

Proof. The proof is immediate, on observing that

$$W_{12,1'}(Y) = W_{12,1'2'} + W_{122',1'}$$

and (7-41)

$$W_{12',1'}(Y) = W_{12',1'2} + W_{122',1'},$$

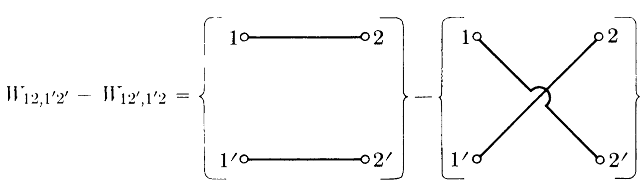


FIG. 7-6. Percival's intuitive representation.

and so the admittance products of 2-trees of the form $T_{2_{122',1'}}$ cancel on subtraction.

Percival [129] expresses this rule in the intuitive fashion shown in Fig. 7–6. The argument above illustrates the typical character of all topological formulas; namely, one does not calculate any superfluous terms in following topological formulas, as one does in evaluating determinants. Only those terms which do not cancel are included.

7–3 The short-circuit admittance functions. As remarked in Section 6–6, two terminal-pair networks are more often described independently of the load y_L by means of the coefficient matrix of the system of equations

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \tag{7–42}$$

with references as in Fig. 7–7. The functions y_{ij} of this matrix are known as *short-circuit admittance functions*, since setting the appropriate voltage equal to zero equates these functions to the current-voltage ratio.

Let node equations be written for the network of Fig. 7–7 with node 1′ as the reference node. Then on solving them as usual [156], we get the open-circuit impedance matrix $\mathbf{Z}_{oc}$ [see Eq. (6–49)] and its inverse, the short-circuit admittance matrix $\mathbf{Y}_{sc}$, as

$$\mathbf{Z}_{oc} = \frac{1}{\Delta}\begin{bmatrix} \Delta_{11} & \Delta_{12} - \Delta_{12'} \\ \Delta_{12} - \Delta_{12'} & \Delta_{22} + \Delta_{2'2'} - 2\Delta_{22'} \end{bmatrix}$$

and (7–43)

$$\mathbf{Y}_{sc} = \frac{1}{\Delta_{1122} + \Delta_{112'2'} - 2\Delta_{1122'}} \begin{bmatrix} \Delta_{22} + \Delta_{2'2'} - 2\Delta_{22'} & \Delta_{12'} - \Delta_{12} \\ \Delta_{12'} - \Delta_{12} & \Delta_{11} \end{bmatrix}.$$

All the cofactors in $\mathbf{Y}_{sc}$ (and $\mathbf{Z}_{oc}$) can be expressed in terms of 2-trees, except those in which two rows and columns have been deleted. To express these terms topologically, the 3-tree, defined below, is needed.

DEFINITION 7–4. *3-tree and 3-tree product.* A *3-tree* is a set of $v - 3$ elements which does not contain a circuit. (Thus a 3-tree is a set of three unconnected circuitless subgraphs which together include all the vertices of the graph. One or two of these subgraphs may consist of isolated vertices.) A *3-tree product* is the product of the admittances of a 3-tree; the product for an isolated vertex is 1, by definition.

Certain specified vertices may be required to be in different connected parts of the 3-tree. Such a 3-tree is denoted as $T_{3_{ab,c,def}}$, which is a 3-tree in which the vertex sets (a, b), (c), and (d, e, f) are required to be in dif-

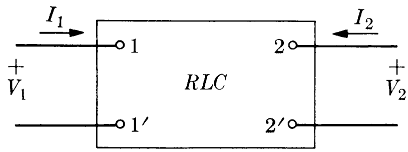


FIG. 7–7. Two terminal-pair reference convention.

ferent connected parts. The sum of 3-tree products is denoted by the symbol $U(Y)$, with subscripts on U to denote any specified distribution of vertices. We see at once, by arguments similar to those of Theorem 7–3, that

$$\Delta_{1122} = U_{1,2,1'}, \quad \Delta_{112'2'} = U_{1,2',1'}, \quad \Delta_{1122'} = U_{1,22',1'}, \tag{7–44}$$

since $1'$ is the reference vertex. 3-trees of the form $T_{3_{1,22',1'}}$ occur both in $U_{1,2',1'}$ and in $U_{1,2,1'}$. Such terms therefore cancel in the det $\mathbf{Z}_{oc}$ expansion because of the $-2\Delta_{1122'}$ term. Therefore

$$\begin{aligned}\Delta_{1122} &+ \Delta_{112'2'} - 2\Delta_{1122'} \\ &= U_{12',2,1'} + U_{1,2,1'2'} + U_{12,2',1'} + U_{1,2',1'2}.\end{aligned} \tag{7–45}$$

The other entries of Y_{sc} are

$$\begin{aligned}\Delta_{22} + \Delta_{2'2'} - 2\Delta_{22'} &= W_{2,1'} + W_{2',1'} - 2W_{22',1'} \\ &= W_{2,1'2'} + W_{2',1'2} \\ &= W_{2,2'},\end{aligned} \tag{7–46}$$

$$\begin{aligned}\Delta_{12'} - \Delta_{12} &= W_{12',1'} - W_{12,1'} \\ &= W_{12',1'2} - W_{12,1'2'},\end{aligned} \tag{7–47}$$

and

$$\Delta_{11} = W_{1,1'}. \tag{7–48}$$

In the sequel, the abbreviation $\sum U$ is used for the sum

$$\sum U = U_{12',2,1'} + U_{1,2,1'2'} + U_{12,2',1'} + U_{1,2',1'2}.$$

THEOREM 7–6. For a passive two terminal-pair network which contains no mutual inductances, the matrix of the short-circuit admittances is given by:

$$\mathbf{Y}_{sc} = \frac{1}{\sum U}\begin{bmatrix} W_{2,2'} & W_{12',1'2} - W_{12,1'2'} \\ W_{12',1'2} - W_{12,1'2'} & W_{1,1'} \end{bmatrix}. \tag{7–49}$$

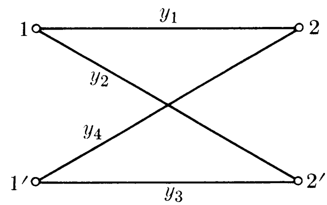


FIG. 7–8. First example for two terminal-pair.

From the computation that was performed for det $\mathbf{Z}_{oc}$, we can also write the topological formula for the determinant of the short-circuit admittance matrix, since

$$\mathbf{Y}_{sc} = \mathbf{Z}_{oc}^{-1} \qquad \text{and so} \qquad \det \mathbf{Y}_{sc} = \frac{1}{\det \mathbf{Z}_{oc}}. \tag{7–50}$$

THEOREM 7–7. For a passive two terminal-pair network which contains no mutual inductances, the determinant of the short-circuit admittance matrix is given by

$$\det \mathbf{Y}_{sc} = \frac{V(Y)}{U_{12',2,1'} + U_{1,2,1'2'} + U_{12,2',1'} + U_{1,2',1'2}} = \frac{V(Y)}{\sum U(Y)}. \tag{7–51}$$

THEOREM 7–8. For a two terminal-pair network which contains no mutual inductances, the open-circuit impedance matrix is given by

$$\mathbf{Z}_{oc} = \frac{1}{V(Y)} \begin{bmatrix} W_{1,1'}(Y) & W_{12,1'2'}(Y) - W_{12',1'2}(Y) \\ W_{12,1'2'}(Y) - W_{12',1'2}(Y) & W_{2,2'}(Y) \end{bmatrix}. \tag{7–52}$$

THEOREM 7–9. For a two terminal-pair network which contains no mutual inductances, the determinant of the open-circuit impedance matrix $\mathbf{Z}_{oc}$ is given by

$$\det \mathbf{Z}_{oc} = \frac{\sum U(Y)}{V(Y)}. \tag{7–53}$$

EXAMPLES. For the two terminal-pair network shown in Fig. 7–8, the required 3-tree and 2-tree products are

$$\begin{gathered} U_{12',2,1'} = y_2, \quad U_{1,2,1'2'} = y_3, \quad U_{12,2',1'} = y_1, \quad U_{1,2',1'2} = y_4; \\ W_{2,2'} = (y_1 + y_2)(y_3 + y_4), \quad W_{1,1'} = (y_1 + y_4)(y_2 + y_3); \\ W_{12',1'2} = y_2 y_4, \quad W_{12,1'2'} = y_1 y_3. \end{gathered} \tag{7–54}$$

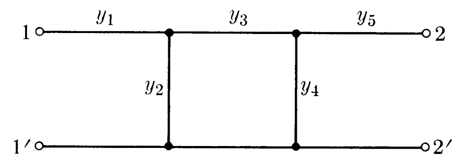


FIG. 7-9. Second example for two terminal-pair.

Hence the short-circuit admittance matrix is

$$\mathbf{Y}_{sc} = \frac{1}{y_1 + y_2 + y_3 + y_4}\begin{bmatrix} (y_1 + y_2)(y_3 + y_4) & y_2y_4 - y_1y_3 \\ y_2y_4 - y_1y_3 & (y_1 + y_4)(y_2 + y_3) \end{bmatrix}. \tag{7-55}$$

As another example, consider Fig. 7-9. The 3-tree and 2-tree products are

$$\begin{aligned} U_{12',2,1'} &= U_{12,1',2'} = U_{1,2',1'2} = 0; \\ U_{1,2,1'2'} &= (y_1 + y_2)(y_3 + y_4 + y_5) + y_3(y_4 + y_5); \\ W_{12',1'2} &= 0, \qquad W_{12,1'2'} = y_1y_3y_5; \\ W_{1,1'} &= y_1y_3y_5 + y_2y_3y_5 + y_1y_4y_5 + y_2y_4y_5 + y_3y_4y_5, \\ W_{2,2'} &= y_1y_2y_3 + y_1y_2y_4 + y_1y_3y_4 + y_1y_2y_5 + y_1y_3y_5. \end{aligned} \tag{7-56}$$

Hence

$$\mathbf{Y}_{sc} = \frac{1}{(y_1 + y_2)(y_3 + y_4 + y_5) + y_3(y_4 + y_5)} \times \begin{bmatrix} y_1y_2(y_3 + y_4 + y_5) + y_1y_3(y_4 + y_5) & -y_1y_3y_5 \\ -y_1y_3y_5 & y_3y_5(y_1 + y_2 + y_4) + y_4y_5(y_1 + y_2) \end{bmatrix}. \tag{7-57}$$

7-4 Kirchhoff's rules. Kirchhoff [86] gave a set of rules, completely dual to those of Maxwell, for the computation of network response. (Kirchhoff's rules were stated almost forty years before Maxwell's rules.) Kirchhoff gave his rules in terms of resistances. We interpret his rules in terms of impedances and loop equations, even though Kirchhoff's rules were stated in terms of the "branch current" system of equations. (Loop currents were invented by Helmholtz about thirty years after Kirchhoff's paper was written.)

It is possible to give a detailed treatment of Kirchhoff's rules, as was done in Section 7-2 for Maxwell's rules. However, since as a matter of convenience we intend to use Maxwell's rules rather than Kirchhoff's in the next chapter, we do not follow such a procedure; further, the development of the formulas in terms of element impedances is completely dual. Therefore we shall be satisfied with the basic formulas for the loop determinant and cofactors. The reader is referred to Mayeda and Seshu [109] for the detailed treatment.

For the mesh determinant, we have that

$$\Delta_m = \det \mathbf{BZB}' = \sum \begin{pmatrix} \text{products of corresponding} \\ \text{majors of } \mathbf{BZ} \text{ and } \mathbf{B}' \end{pmatrix}. \tag{7–58}$$

For a network that contains no mutual inductances, $\mathbf{Z}(s)$ is diagonal and so, as before,

$$\Delta_m = \sum_i z_{i_1} z_{i_2} \cdots z_{i_\mu} \cdot (\text{major of } \mathbf{B})^2, \tag{7–59}$$

where $\mu = e - v + 1$.

By Theorem 5–8, the nonzero majors of $\mathbf{B}$ are in one-to-one correspondence with the chord sets of the network. However, such a major does not necessarily have a value ± 1 in general. Okada [126] states that the value of a nonzero major of $\mathbf{B}$ is $\pm 2^i$, i being a nonnegative integer, fixed for a given $\mathbf{B}$. (See also Problem 5–29.) Thus if we define a *chord-set product* to be the product of the impedances of the chords of a tree of the network, we get the following topological formula for the mesh determinant.

THEOREM 7–10. For a network that contains no mutual inductances,

$$\Delta_m = \det \mathbf{BZB}' = 2^{2i} \sum (\text{chord-set products}). \tag{7–60}$$

There are two cases for which i is certainly zero; $i = 0$ for fundamental circuits, and $i = 0$ for the set of meshes ("windows") of a planar network. (See Problems 5–26 and 5–29.) A detailed discussion of this question has been given by Cederbaum [28]. Since the network functions are independent of the circuit basis chosen, we may assume that the fundamental system of circuits is chosen and so let $i = 0$. Then we have that

$$\Delta_m = \sum (\text{chord-set products}). \tag{7–61}$$

This topological formula was originally given by Kirchhoff [86] in the form:

> Δ_m is the sum of products of resistances taken $e - v + 1$ at a time, which have the common property that, when these elements are removed, no circuits remain.

The topological formulas of Theorems 7–1 and 7–10 can be combined by observing that

$$z_1 \cdot y_1 = 1$$

and so

$$(z_1 \cdot z_2 \cdots z_e)(\text{tree product}) = (\text{chord-set product of the same tree}).$$

THEOREM 7–11. For a network without mutual inductances,

$$\Delta_m = z_1 \cdot z_2 \cdot z_3 \cdots z_e \, \Delta_n, \tag{7-62}$$

where fundamental circuits are used.

The result is originally due to Tsang [181] and has been extended by Cederbaum [28] to networks containing magnetic coupling. The notation for the *mesh discriminant* is simplified by introducing the following complement convention of Percival [129].

Given the polynomial $V(Y)$, the complementary polynomial $C[V(Y)]$ is formed by replacing each product in $V(Y)$ by the product of the variables not in this product. The polynomial $C[V(Z)]$ is obtained by replacing y_i by z_i in $C[V(Y)]$. With these conventions,

$$\Delta_m = C[V(Z)]. \tag{7-63}$$

In using this complement convention, we also adopt the convention that the complement of zero is zero.

The cofactor of the element in the (i, i)-position of the matrix $\mathbf{Z}_m(s)$ is of interest only when there is at least one element in the ith circuit which is not in any other circuit. Hence it is assumed that there is an element y_j in circuit i which is in no other circuit. Let the vertices of y_j be 1 and $1'$.

Using the same notation as before, we write

$$\Delta_{ii} = \det \mathbf{B}_{-i}\mathbf{Z}\mathbf{B}'_{-i}. \tag{7-64}$$

With the assumption that y_j is in no other circuit, we find that this matrix $\mathbf{B}_{-i}\mathbf{Z}\mathbf{B}'_{-i}$ is the mesh impedance of the network obtained by deleting element y_j. Let the network obtained by deleting element y_j be denoted by N_1.

THEOREM 7–12.

$$\Delta_{ii} = \sum \text{(chord-set products of } N_1) = C[V_1(Z)]. \tag{7-65}$$

Kirchhoff gave the following rule for the computation of the current I_{rs} in an element between vertices r and s with reference from r to s, due to a generator E between vertices p and q with a reference $+$ at q:

$$I_{rs} = E_{qp} \frac{\Delta_{ab}}{\Delta}, \tag{7-66}$$

where Δ is the mesh determinant, which has already been considered. Kirchhoff's rule for Δ_{ab} is:

Δ_{ab} is the sum of (signed) products of impedances taken $e - v$ at a time, which have the common property that, after these elements have

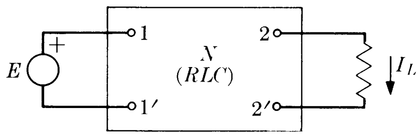


FIG. 7–10. Conventions for Kirchhoff's rule.

been removed, there is only one circuit left, and this circuit contains both the generator E and the element in which the current is being computed. The terms for which the remaining circuit goes through both E_{qp} and I_{rs} in the same relative direction are taken with a positive sign, and those for which the remaining circuit goes through E and I_{rs} in opposite directions are taken with a negative sign. (The orientation is with reference to the element orientation.)

To correlate Kirchhoff's rules with 2-trees, it is convenient to introduce the following conventions. Let N denote the two terminal-pair network of Fig. 7–10, excluding the generator E and the load Z_L. Also, N consists of R-, L-, and C-elements only. Consider one of the products in Δ_{ab} of Kirchhoff's rule. There are $e - v$ elements in this product, where e is the number of elements in the complete network, including E and Z_L. When this set of elements is removed, there are v elements remaining. This set of elements includes exactly one circuit which contains both E and Z_L. Hence if either E or Z_L (but not both) is removed, the rest is a tree of the network $N + E + Z_L$. Therefore, if both E and Z_L are removed, the rest is a 2-tree of N, which separates the vertices of E as well as the vertices of Z_L. Hence the rest is both $T_{2_{1,1'}}$ and $T_{2_{2,2'}}$. Once again the 2-tree may be either $T_{2_{12,1'2'}}$ or $T_{2_{12',1'2}}$.

The products in Δ_{ab} consist of the elements of N which are not in these 2-trees. Thus Δ_{ab} contains $C[T_{2_{12,1'2'}}]$ impedance products and $C[T_{2_{12',1'2}}]$ impedance products, where C denotes complementation with respect to N only. Kirchhoff affixes a positive sign to the products of the first type and a negative sign to the products of the second type. Thus by Kirchhoff's formula,

$$\Delta_{ab} = C[W_{12,1'2'}(Z)] - C[W_{12',1'2}(Z)]. \tag{7–67}$$

Complementation is with respect to N, and the Z in parentheses implies that the products are impedance products.

To prove* Kirchhoff's formula, we need first to observe that *a set of fundamental circuits can always be chosen for the complete network $N + E + Z_L$, such that neither E nor Z_L is in more than one circuit, although they may both be in the same fundamental circuit or in different ones.*

* Proof as given by R. Obermeyer in a term paper at the University of Illinois, 1956.

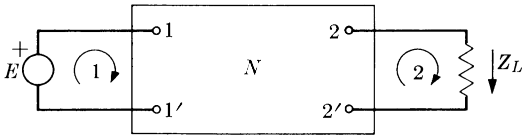


FIG. 7–11. References for loops.

If the network N is connected, we can find a tree of N. The elements E and Z_L are chords for such a tree and so are in only one circuit each and in different circuits. If N is not connected, $N + E$ is connected. Otherwise $N + E + Z_L$ would be separable (we are assuming that it is nonseparable). Choosing a tree of $N + E$, which necessarily contains E, we observe that the element E would be only in the fundamental circuit of Z_L and in no other. Thus E and Z_L are in only one circuit.

The latter case, in which E and Z_L are in the same circuit, is the driving-point case, which has already been considered. Hence it will be assumed that E and Z_L are in different fundamental circuits. Let E be in circuit 1 and Z_L be in circuit 2 for notational convenience. Let these circuits be oriented as shown in Fig. 7–11. Then, obviously,

$$I_L = \frac{\Delta_{12}}{\Delta} E, \tag{7–68}$$

with reference to the mesh equations. Since fundamental circuits were chosen,

$$\Delta = \sum \text{(chord-set products of } N + E + Z_L) \tag{7–69}$$

(without any factor 2^{2i}). Also, we find that

$$\begin{aligned} \Delta_{12} &= [\text{cofactor of the (1, 2)-element of } \mathsf{B}_f\mathsf{Z}\mathsf{B}_f'] \\ &= (-1)^{1+2} \det \mathsf{B}_{-1}\mathsf{Z}\mathsf{B}_{-2}', \end{aligned} \tag{7–70}$$

using the same notation as in Section 7–1, the subscript denoting the deleted row. Once again

$$\det \mathsf{B}_{-1}\mathsf{Z}\mathsf{B}_{-2}' = \sum \begin{pmatrix} \text{products of corresponding} \\ \text{majors of } \mathsf{B}_{-1}\mathsf{Z} \text{ and } \mathsf{B}_{-2}' \end{pmatrix}. \tag{7–71}$$

Deleting row 1 from B_f yields the circuit matrix of the network when circuit 1 is destroyed, which effect is obtained by deleting element E. Thus, *nonzero majors of* B_{-1} *are in one-to-one correspondence with chord sets of* $N + Z_L$.

Similarly, deleting row 2 of B_f yields the circuit matrix when circuit 2 is destroyed, which is the same as deleting element Z_L. Hence, *nonzero majors of* B_{-2} *are in one-to-one correspondence with the chord sets of* $N + E$.

Since $\mathbf{Z}$ is a diagonal matrix, it introduces no complications.

Thus, to get a nonzero product of the two majors, the set of elements must be a chord set of both $N + E$ and $N + Z_L$. Thus, the chord set cannot include either E or Z_L. Hence E in $N + E$ and Z_L in $N + Z_L$ must be branches of the trees for which this set is a chord set. The elements of N which are branches for these trees must therefore constitute a 2-tree of N. This 2-tree separates the vertices of both E and Z_L. Hence, it is a 2-tree of one of the two types $T_{2_{12,1'2'}}$ or $T_{2_{12',1'2}}$. Conversely, the product of the elements in the complement in N of every such 2-tree is a term in det $\mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2}$, since each such 2-tree with E is a tree of $N + E$, and with Z_L is a tree of $N + Z_L$.

It remains to establish the signs of $C[W_{12,1'2'}(Z)]$ and $C[W_{12',1'2}(Z)]$. We follow a procedure analogous to the one adopted in establishing Maxwell's rule for asymmetrical cofactors of the node-admittance matrix.

Let $e_{q_1}, e_{q_2}, \ldots, e_{q_{e-v}}$ be a set of elements corresponding to the columns of a nonzero major in $\mathbf{B}_{-1}$ and $\mathbf{B}_{-2}$. To establish the signs of these two majors, consider the complete fundamental-circuit matrix $\mathbf{B}_f$ in which the columns are rearranged in the order

$$1, 2, q_1, q_2, \ldots, q_{e-v}, \ldots, q_{e-2}.$$

Since the order of the columns $q_1, q_2, \ldots, q_{e-v}$ has not been changed, the major determinants of interest remain the same. Now the set of elements complementary to the set $q_1, q_2, \ldots, q_{e-v}$ (with respect to N) is a 2-tree of N separating the pairs of vertices $(1, 1')$ and $(2, 2')$. If we adjoin both E and Z_L to this 2-tree, the resultant graph contains one circuit K. This circuit K contains both E and Z_L. Since every circuit can be built up from fundamental circuits, so can K. Let the coefficients of the linear combination (of fundamental circuits) which produces K be $(\epsilon_1, \epsilon_2, \ldots, \epsilon_\mu)$, where each $\epsilon_j = 1, -1$, or 0, and $\mu = e - v + 1$. We may take $\epsilon_1 = 1$; then $\epsilon_2 = 1$ or -1. (Since K contains both E and Z_L, and since these appear only in the first and second circuits of the fundamental set, respectively, $\epsilon_1 \neq 0$ and $\epsilon_2 \neq 0$.) Therefore

$$[\epsilon_1\epsilon_2\epsilon_3 \cdots \epsilon_\mu]\mathbf{B}_f = \mathbf{K}, \tag{7-72}$$

where $\mathbf{K}$ stands for the row matrix of the circuit K.

Premultiply the circuit matrix $\mathbf{B}_f$ by the nonsingular matrix

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & \epsilon_3 & \epsilon_4 & \epsilon_5 & \cdots & \epsilon_{\mu-1} & \epsilon_\mu \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{7-73}$$

Since rows 1 and 2 have not been used as "tool" rows in this set of row operations, the major determinants of $\mathbf{B}_{-1}$ and $\mathbf{B}_{-2}$ are unaltered in this process. Let

$$\mathbf{M}\mathbf{B}_f = \mathbf{B}_K. \tag{7–74}$$

In the matrix $\mathbf{B}_K$, if we multiply row 2 by ϵ_2 and add to the first row, the first row becomes the circuit K. This circuit K contains E, Z_L, and elements from the 2-tree. Hence K does not contain any of the elements $q_1, q_2, \ldots, q_{e-v}$. And $\epsilon_2 = 1$ or -1. Hence (Case 1) the entries in columns $q_1, q_2, \ldots, q_{e-v}$ of the first two rows of the matrix $\mathbf{B}_K$ are either identical or (Case 2) the entries in the second row are the negatives of the entries in the first row.

Case 1. In this case, the majors of $\mathbf{B}_{-1}$ and $\mathbf{B}_{-2}$ are identical, since deleting the first row of the submatrix containing columns $q_1, \ldots, q_{e-v}$ produces the same submatrix as deleting the second row. Hence the product of the two majors is equal to one.

In this case, ϵ_2 must equal -1 to produce the desired zeros for circuit K. Hence circuit K has the form

$$\begin{array}{rcccccccccc} & E & Z_L & q_1 & q_2 & q_3 & \cdots & q_{e-v} & \cdots & q_e & \\ \mathbf{K} = & [1 & -1 & 0 & 0 & 0 & \cdots & 0 & & ---]; \end{array} \tag{7–75}$$

that is, E and Z_L appear with opposite signs. With reference to Fig. 7–11, the circuit K goes through the vertices of E and Z_L in the order $1'12'21'$. Therefore the 2-tree must be $T_{2_{12'},_{1'2}}$ to provide the required paths for circuit K of this form. The converse is also seen to be true. Thus

$$C[W_{12',1'2}(Z)]$$

has a positive sign in the expansion of det $\mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2}$.

Case 2. In this case, the major of $\mathbf{B}_{-2}$ is obtained by multiplying the first row of the major of $\mathbf{B}_{-1}$ by -1. Hence

$$(\text{major of } \mathbf{B}_{-1})(\text{major of } \mathbf{B}'_{-2}) = -1.$$

Also, in this case $\epsilon_2 = 1$, and so following the same argument as before, we see that the circuit K is of the form

$$\begin{array}{rccccccccc} & E & Z_L & q_1 & q_2 & \cdots & q_{e-v} & \cdots & q_e & \\ \mathbf{K} = & [1 & 1 & 0 & 0 & \cdots & 0 & & --- &]. \end{array} \tag{7–76}$$

Hence the 2-tree must be of the type $T_{2_{12},_{1'2'}}$ to provide the required path for the circuit K. Conversely, every such 2-tree leads to a circuit K for which $\epsilon_2 = 1$. Hence in the expansion of det $\mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2}$, the sum $C[W_{12,1'2'}(Z)]$ has a negative sign.

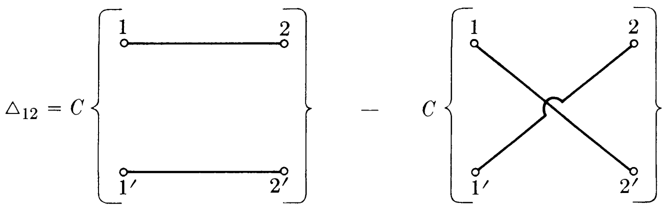


FIG. 7–12. Percival's representation of Kirchhoff's rule.

Hence,

$$\det \mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2} = C[W_{12',1'2}(Z)] - C[W_{12,1'2'}(Z)]. \tag{7–77a}$$

Finally, since

$$\Delta_{12} = (-1)^{1+2} \det \mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2} = -\det \mathbf{B}_{-1}\mathbf{Z}\mathbf{B}'_{-2}, \tag{7–77b}$$

we have that

$$\Delta_{12} = C[W_{12,1'2'}(Z)] - C[W_{12',1'2}(Z)]. \tag{7–77c}$$

This formula can be written in an intuitive fashion by following Percival, as in Fig. 7–12.

THEOREM 7–13. For a network containing no mutual inductances, if circuits 1 and 2 contain the elements $(1, 1')$ and $(2, 2')$, respectively, and these elements are in no other circuits, the cofactor $(1, 2)$ of the mesh-impedance matrix is given by

$$\Delta_{12} = C[W_{12,1'2'}(Z)] - C[W_{12',1'2}(Z)]. \tag{7–78}$$

7–5 General linear networks. The assumptions of reciprocity and no mutual inductance, made in the earlier sections of this chapter, are now dropped, and topological formulas are developed for the more general class of lumped linear networks, including nonreciprocal elements and mutual inductances. However, the network is assumed to have either an element-admittance matrix $\mathbf{Y}(s)$ or an element-impedance matrix $\mathbf{Z}(s)$. This assumption excludes the so-called "ideal" transformer. If $\mathbf{Y}(s)$ exists, admittance formulas are possible, and if $\mathbf{Z}(s)$ exists, impedance formulas are possible. Only the admittance formulas are considered here, as the two are dual developments. For the admittance formulas, the "perfectly coupled" transformers (for which the nonzero rows and columns of $\mathbf{L}$ constitute a semidefinite matrix) must also be excluded.

Three methods of developing topological formulas for such networks are known (due to Mason [107], Coates [36], and Mayeda [111]). The procedure due to Mason is conceptually very different from the methods of analysis discussed in this text; hence it is not included here. The other

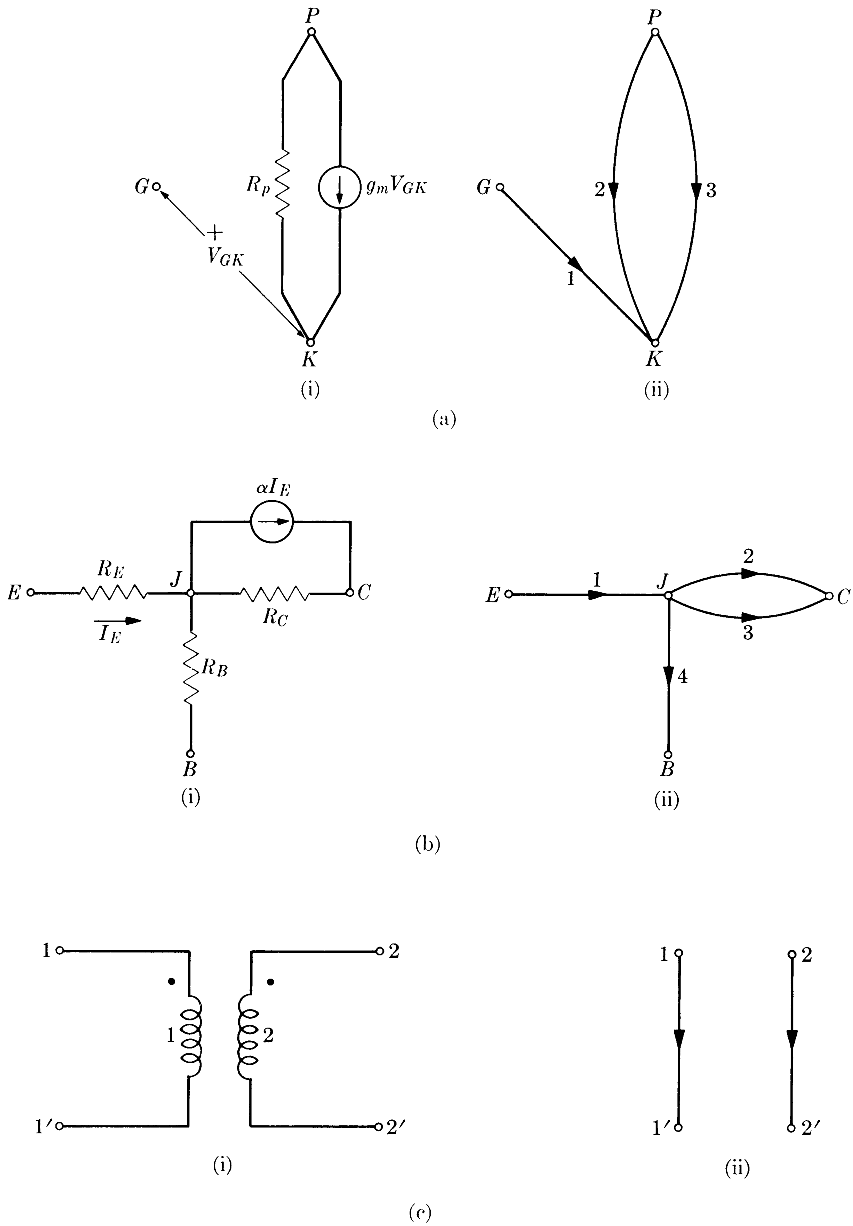


FIG. 7-13. Some common networks.

two are identical as computational schemes, when applied to most practical networks. However, from a theoretical point of view, the development due to Coates is more general. The development given in this section is a mixture of the theories of Coates and Mayeda.

It is assumed that the element-admittance matrix exists, so that the element-current transforms can be expressed in terms of the element-voltage transforms as

$$\mathsf{I}(s) = \mathsf{Y}(s)\mathsf{V}(s) + \mathsf{K}(s, 0+), \tag{7-79}$$

where $\mathsf{K}(s, 0+)$ contains the initial values. As before, the initial values are assumed to be zero, since the objective is to compute network functions. Then Eq. (7–79) becomes

$$\mathsf{I}(s) = \mathsf{Y}(s)\mathsf{V}(s). \tag{7-80}$$

The following assumption is made about $\mathsf{Y}(s)$. If

$$\mathsf{Y}(s) = [y_{ij}] \qquad \text{for all } i \text{ and } j, \tag{7-81a}$$

either

$$y_{ij} = y_{ji} \tag{7-81b}$$

or (if $y_{ij} \neq y_{ji}$)

$$\text{one of } y_{ij},\, y_{ji} \text{ is } 0. \tag{7-81c}$$

This assumption is satisfied in all practical networks where each R, L, C, and generator is considered as a separate network element. Mayeda implicitly makes such an assumption; Coates does not. The resultant generality of the Coates theory has been dropped in the present discussion, for simplicity. As an example of the significance of the assumption of Eq. (7–81c), consider the three common networks shown in Fig. 7–13. The corresponding matrices $\mathsf{Y}(s)$ are given in Eq. (7–82a, b, c), where the subscripts for the currents and voltages correspond to graphs in part (ii) of each figure:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & G_p & 0 \\ g_m & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}, \tag{7-82a}$$

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} G_E & 0 & 0 & 0 \\ \alpha G_E & 0 & 0 & 0 \\ 0 & 0 & G_C & 0 \\ 0 & 0 & 0 & G_B \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix}, \tag{7-82b}$$

$$\begin{bmatrix} I_1 \\ \\ I_2 \end{bmatrix} = \begin{bmatrix} \dfrac{L_{22}}{s\Delta} & -\dfrac{M_{12}}{s\Delta} \\ \\ -\dfrac{M_{12}}{s\Delta} & \dfrac{L_{11}}{s\Delta} \end{bmatrix} \begin{bmatrix} V_1 \\ \\ V_2 \end{bmatrix}, \tag{7–82c}$$

where $\Delta = L_{11}L_{22} - M_{12}^2$ $(\Delta \neq 0)$. In Fig. 7–13(a), if the voltage-generator equivalent is used, the admittance matrix $\mathsf{Y}(s)$ does not exist.

The assumption that $\mathsf{Y}(s)$ exists thus prohibits all dependent voltage generators (unless they are converted into current generators by the use of Norton's theorem); even current generators that depend on currents can be admitted only if the current (on which the generator depends) is in an element with a finite admittance—not a short circuit, in other words.

Much of the simplicity of the topological formulas derived for passive reciprocal networks without mutual inductances is due to the fact that $\mathsf{Y}(s)$ is a diagonal matrix. The matrices in Eq. (7–82a, b, c) are not. The Coates-Mayeda technique is to *modify* the graph in such a fashion that the node-admittance matrix $\mathsf{Y}_n(s)$ can be written as

$$\mathsf{Y}_n(s) = \mathsf{A}_i\mathsf{Y}(s)\mathsf{A}_v', \tag{7–83}$$

where $\mathsf{Y}(s)$ is the element-admittance matrix of the new graph and *is diagonal.* A_i and A_v are two new incidence matrices (to be defined shortly). The node-admittance matrix $\mathsf{Y}_n(s)$ of Eq. (7–83) is the same as the node-admittance matrix of the original graph. Thus, the problem is only slightly more complicated than the case of passive reciprocal networks without mutual inductances [complicated by the difference between A_i and A_v of Eq. (7–83)]. The modification proceeds as follows.

The procedure is simple, but its formal description is involved because of the various possibilities to be considered. Therefore an example is given first, before the formal description. Consider the example of Fig. 7–13(b) with the associated matrix of Eq. (7–82b):

$$\mathsf{Y} = \begin{bmatrix} G_E & 0 & 0 & 0 \\ \alpha G_E & 0 & 0 & 0 \\ 0 & 0 & G_C & 0 \\ 0 & 0 & 0 & G_B \end{bmatrix}. \tag{7–84}$$

Begin with the original graph of Fig. 7–13(b)(ii). When the diagonal element of Y is nonzero, associate this entry as the weight of the cor-

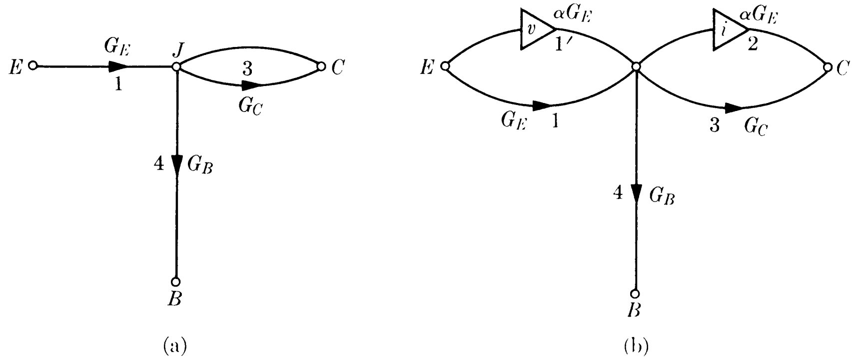


FIG. 7–14. Modified graph of transistor.

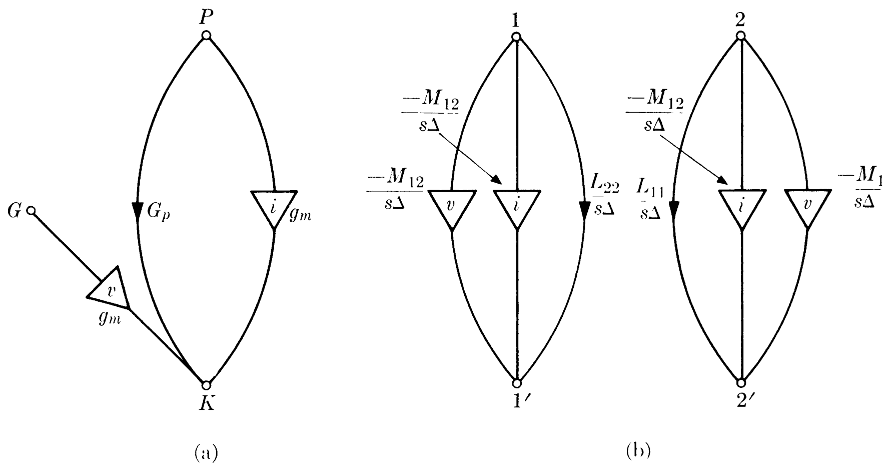


FIG. 7–15. Modified graphs of (a) vacuum tube and (b) transformer.

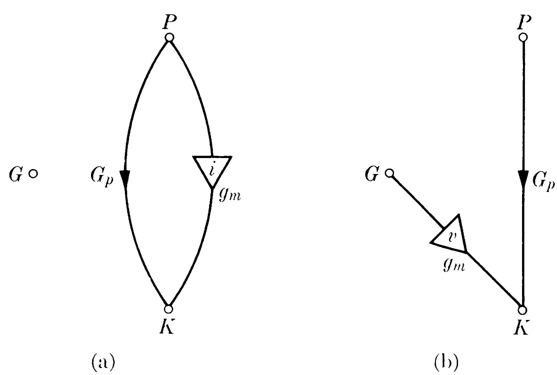


FIG. 7–16. Expansion of Fig. 7–15(a). (a) Current graph. (b) Voltage graph.

responding edge, as in Fig. 7–14(a). For the off-diagonal entry in the (2, 1)-position of Eq. (7–84), there are two associated edges in the graph, as in Fig. 7–14(b). One of these is the original current generator (edge 2), and the other is an added edge (edge 1′). This added edge merely indicates the voltage on which the current generator depends. The current of the added edge is zero. However, the weight of edge 1′ is also made αG_E, the same as the other edge (2) corresponding to the (2, 1)-entry of $\mathbf{Y}$. Edges 2 and 1′ constitute an *edge-pair* in the terminology of Coates. Edges 2 and 2′ must be distinguished from each other by some means. Edge 2 is the *current edge* and edge 1′ is the *voltage edge*. Mayeda's convention is shown in Fig. 7–14(b). All other edges are *ordinary edges* (Coates: *single edges*) and are to be treated as both current edges and voltage edges.

The procedure in the general case is the same as in the preceding example. The modified graph has the same vertices as the original graph. Whenever the diagonal entry of $\mathbf{Y}$ is nonzero, the corresponding edge is given this (diagonal-entry) weight. For each nonzero off-diagonal entry, the modified graph contains a pair of edges (one of which may be an edge of the original graph). If $y_{ij} \neq 0$, place two edges in the modified graph between the pairs of vertices at which edges i and j of the original graph were incident, with the same orientation as edges i and j, respectively. The edge with the vertices of edge i is the current edge, and the edge with the vertices of edge j is the voltage edge. With each of these is associated the weight y_{ij}. It is important to note that if also $y_{ji} \neq 0$, *another* pair of edges must be added, the current edge between vertices of edge j and voltage edges between vertices of edge i. The modified graphs of Figs. 7–13(a) and (c) are shown in Figs. 7–15(a) and (b), respectively. For Fig. 7–15(b), $\Delta = L_{11}L_{22} - M_{12}^2$ ($\Delta \neq 0$). The Coates representation of a transformer is different from Fig. 7–15(b).

In the Mayeda representation, used here, the graph of the network consists of two graphs, a current graph and a voltage graph, which are shown together purely for convenience. For instance, Fig. 7–15(a) represents the current and voltage graphs of Fig. 7–16. The current graph contains only the current elements, and the voltage graph contains only the voltage elements. The ordinary elements appear in both graphs. The edges marked g_m in Fig. 7–16, and more generally the current and voltage elements for $y_{ij} \neq 0$, are considered to be the *same edge*, occupying different positions. Coates prefers to think of these edges as different, and during the computation interchanges their positions. Although the Coates conception is logically more satisfying, the Mayeda procedure is more convenient for practical computations.

The matrices $\mathbf{A}_i$ and $\mathbf{A}_v$ are now defined to be the *incidence matrices* of the current and voltage graphs respectively, *with the same reference vertex.*

For ordinary elements, the corresponding columns of A_i and A_v are identical. For the others, if column k of A_i corresponds to a current element, column k of A_v corresponds to the corresponding voltage element. Thus the two columns k (of A_i and A_v) are unrelated. For example, for the graph of Fig. 7–14(b), the incidence matrices are (with vertex B as reference)

$$\mathsf{A}_i = \begin{array}{c} \\ E \\ J \\ C \end{array} \begin{array}{c} \begin{array}{cccc} G_E & G_C & G_B & \alpha G_E \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \\ 0 & -1 & 0 & -1 \end{bmatrix} \end{array}$$

and (7–85)

$$\mathsf{A}_v = \begin{array}{c} E \\ J \\ C \end{array} \begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & -1 & 0 & 0 \end{bmatrix}.$$

The element-admittance matrix $\mathsf{Y}(s)$ of the modified graph is defined to be a diagonal matrix:

$$\mathsf{Y}(s) = \begin{bmatrix} y_1 & 0 & 0 & \cdots & 0 \\ 0 & y_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & y_e \end{bmatrix}, \tag{7–86}$$

where y_j is the weight of edge j in the modified graph.

THEOREM 7–14. Let G_1 be the graph of a network with an element-admittance matrix Y_1, and let G_2 be the modified graph derived by the procedure above, with admittance matrix Y_2. Then if $\mathsf{Y}_n(s)$ is the node-admittance matrix of G_1, with reference vertex v,

$$\mathsf{Y}_n(s) = \mathsf{A}_i\mathsf{Y}_2\mathsf{A}_v', \tag{7–87}$$

where A_i and A_v are current- and voltage-incidence matrices of G_2 with reference vertex v, and all vertices (rows and columns of Y_n, and rows of A_i and A_v) appear in their natural order.

Proof. Let

$$\mathsf{Y}_n(s) = [Y_{ij}] = \mathsf{A}\mathsf{Y}_1\mathsf{A}', \tag{7–88}$$

where A is the incidence matrix of G_1 with reference v. Then

$$Y_{ij} = \sum_{k=1}^{n} \sum_{p=1}^{n} a_{ik}y_{kp}a_{jp}, \tag{7–89}$$

where

$$\mathsf{A} = [a_{ij}] \qquad \text{and} \qquad \mathsf{Y}_1 = [y_{kj}]. \tag{7-90}$$

On the other hand, let

$$\mathsf{Y}_2 = [y_{ij}^{(2)}]\,, \qquad \mathsf{A}_i = [a_{ij}^{(i)}]\,, \qquad \mathsf{A}_v = [a_{ij}^{(v)}]\,; \tag{7-91}$$

then

$$\mathsf{A}_i\mathsf{Y}_2\mathsf{A}_v' = [\eta_{ij}], \tag{7-92a}$$

where

$$\eta_{ij} = \sum_{k=1}^{n_2} a_{ik}^{(i)} y_{kk}^{(2)} a_{jk}^{(v)}, \tag{7-92b}$$

since Y_2 is a diagonal matrix. Consider the product $a_{ik}y_{kp}a_{jp}$. For this product to be nonzero, y_{kp} cannot equal 0, element k should be incident at vertex i, and element p should be incident at vertex j. If $k = p$, this edge has been preserved in G_2, say as edge m, and is an ordinary edge. Hence

$$a_{ik}y_{kk}a_{jk} = a_{im}^{(i)}y_{mm}a_{jm}^{(2)}. \tag{7-93}$$

If $k \neq p$, there are two edges in the modified graph, a current edge, edge m say, incident at vertex i (in the same way as edge k) and a voltage edge, also edge m, incident at vertex j (in the same way as edge p). Hence

$$a_{im}^{(i)} = a_{ik}, \qquad y_{kp} = y_{mm}^{(2)}, \qquad a_{jp}^{(v)} = a_{jk}. \tag{7-94}$$

Hence the product is preserved once again, and the theorem is established.

Since the matrices are equal, determinants and all cofactors are also equal. Thus it suffices to find topological formulas for the determinant and cofactors of $\mathsf{A}_i\mathsf{Y}(s)\mathsf{A}_v'$ for the modified graph. Since A_i and A_v are incidence matrices of graphs, they have all the familiar properties. In particular, nonsingular submatrices correspond to trees and have determinants ± 1. By referring back to the application of the Binet-Cauchy theorem to the node determinant (Section 7-1), we see that only such trees as are trees of both the current and voltage graphs contribute to det $\mathsf{A}_i\mathsf{Y}(s)\mathsf{A}_v'$.

DEFINITION 7-5. *Complete tree* and *complete-tree product.* The set of edges with weights $y_{i_1}, y_{i_2}, \ldots, y_{i_k}$ constitute a *complete tree* if the current edges with weights $y_{i_1}, y_{i_2}, \ldots, y_{i_k}$ constitute a tree of the current graph and the voltage edges with weights $y_{i_1}, y_{i_2}, \ldots, y_{i_k}$ constitute a tree of the voltage graph. (Some or all of these edges may be ordinary edges.) The *complete-tree product* is the product of the admittances $y_{i_1}y_{i_2}\cdots y_{i_k}$ of a complete tree.

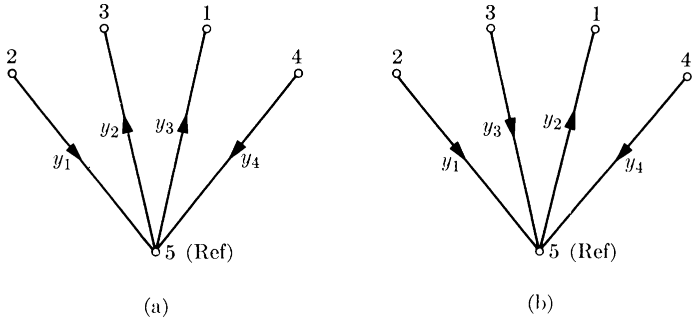


FIG. 7–17. Complete tree. (a) Current graph. (b) Voltage graph.

Thus, from the Binet-Cauchy theorem,

$$\begin{aligned}\det \mathbf{Y}_n(s) &\\ &= \sum (\text{complete-tree product}) \times (\text{major of } \mathbf{A}_i) \times (\text{major of } \mathbf{A}'_v)\\ &= \sum_j \epsilon_j \,(\text{complete-tree product of tree } j),\end{aligned} \tag{7–95}$$

where $\epsilon_j = 1$ or -1. The problem thus reduces to the computation of the relative signs of the majors chosen from $\mathbf{A}_i$ and $\mathbf{A}_v$. If all elements of the complete tree are ordinary elements, the two majors are identical and $\epsilon_j = 1$. The procedure given here for the general case is Mayeda's algorithm, which is equivalent to the procedure described by Coates.

Consider first the simplest case, in which all edges of the complete tree are incident to the reference vertex in both current and voltage graphs, as in Fig. 7–17. (Some of these edges are not ordinary edges and so appear in different places in the two parts.) Remembering that it suffices to find the relative signs of the two majors, we see that the most efficient procedure is to find the number of changes necessary to make the two trees identical and then to compute the effect on the incidence matrix of such changes. Interchanging two elements of a tree is equivalent to interchanging two columns of the incidence matrix, hence changing the sign of the major. If y_2 and y_3 are interchanged in Fig. 7–17(b), the positions of the nonzero entries in the two corresponding majors are identical, but their signs are not. To make the signs agree as well, the reference (orientation) arrow for y_3 must be reversed. Reversing a reference arrow is equivalent to multiplying the corresponding column by -1 and hence changing the sign of the major. In this case, two changes are required to make the trees identical—one interchange and one reversal of arrow. Hence the two majors have the same sign, and the coefficient ϵ_j for the complete-tree product $y_1y_2y_3y_4$ in Eq. (7–95) is 1.

TABLE 7–1

SIGN PERMUTATION FOR THE EXAMPLE OF FIG. 7–17

1	2	3	4
y_3^-	y_1	y_2^-	y_4
y_2^-	y_1	y_3	y_4

In the general algorithm (written for a computer) it is found easier to arrange all the reference arrows to point toward the reference vertex rather than to compare relative orientations in the two trees. Here again the total number of reversals of arrows in the two graphs plus the number of element interchanges required in one graph, decides the sign of ϵ_j. If this number is odd, $\epsilon_j = -1$, otherwise $\epsilon_j = +1$.

The algorithm is the following. Make a table of two rows and $v - 1$ columns, where v is the number of vertices. List the vertices in natural order as the columns. In the first row, list the edges of the current graph incident at the vertex corresponding to each column. If any edge is oriented away from the reference vertex, add a superscript minus. Repeat in the second row, for the voltage graph. This table, called a *sign permutation*, is, for the example of Fig. 7–17, shown as Table 7–1. Now $\epsilon_j = (-1)^{n+m}$, where n is the number of minus signs in the superscripts (3 in Table 7–1), and m is the number of interchanges required in the second row to make the rows identical (1 in this example).

For the general tree, which is not star-shaped like Fig. 7–17, Mayeda notes the changes that have to be made to convert it into a star-shaped tree. In practice, this conversion need not be made; it is necessary only to establish the procedure. Consider any tree T and its incidence matrix A_T, with reference vertex v. By Problem 2–6, T contains an end-vertex, a say, at which only one element y_i is incident. If the other vertex of y_i is v, no change is required. If it is not, let the other vertex be k. Row a of A_T contains only one nonzero element, in column y_i. Add row a to row k. In this process, det A_T is unaltered, but the element in row k, column y_i, is now zero. No other element of A_T is changed. The new matrix is the incidence matrix of a tree T_1 in which y_i is incident to vertices a and v and all other edges are as in T. Remove vertex a and y_i from T_1. The rest is still a tree (of $v - 1$ vertices), and so the procedure can be repeated until T is converted into a star-shaped tree, with det A_T remaining unaltered throughout. In the star-shaped tree that results, each nonreference vertex is incident with exactly one edge. To make up the sign-permutation table, it is necessary to find this edge only. It is not necessary to actually convert the tree into a star shape. A little thought about the process described above will reveal that the edge y_i is incident to vertex a in the star-shaped tree if a is a vertex of y_i in T and the (unique) path on T from a to v contains y_i.

DEFINITION 7–6. *Principal edge.* With respect to a given tree T and a given reference vertex v, edge y_i is the *principal edge* of vertex a if y_i is incident at a and the unique path in T from vertex a to the reference vertex contains y_i.

DEFINITION 7–7. *Sign permutation.* For a given complete tree τ and a given reference vertex v, the *sign permutation* is a matrix of order $(2, v - 1)$ with columns corresponding to vertices, and rows corresponding to current and voltage graphs. The (i, j)-entry is y_k if y_k is the principal edge of vertex j in graph i $(i = 1, 2)$ and is oriented away from vertex j; it is y_k^- if y_k is the principal edge of vertex j and is oriented toward vertex j.

Thus with the concept of a principal edge, the sign permutation can be formed directly and the same rule as before gives ϵ_j.

THEOREM 7–15. If $\mathsf{Y}_n(s)$ is the node-admittance matrix of a network with a Y-matrix,

$$\det \mathsf{Y}_n(s) = \sum_{\substack{\text{all}\\ \text{complete}\\ \text{trees}}} \epsilon_t \times (\text{complete-tree product}), \tag{7–96}$$

where

$$\epsilon_t = (-1)^{n+m},$$

in which

$$n = \begin{pmatrix}\text{total number of superscript minus signs in the}\\ \text{sign permutation}\end{pmatrix}$$

and

$$m = \begin{pmatrix}\text{number of interchanges required in one row of}\\ \text{sign permutation to make the rows identical}\end{pmatrix}.$$

EXAMPLE. A transformer-coupled transistor amplifier is shown in Fig. 7–18. The driving-point impedance at terminals (1, 1′) and the transfer impedance from (1, 1′) to (2, 2′) are required. The determinant is calculated at this point; the cofactors are computed later. The modified graph of the network is shown in Fig. 7–19 [cf. Figs. 7–14 and 7–15(b)]. Here,

$$y_4 = \frac{L_5}{s\Delta}, \qquad y_5 = \frac{L_4}{s\Delta}, \qquad y_8 = y_9 = \frac{-M_{45}}{s\Delta}.$$

The trees consisting only of ordinary elements are easily disposed of, as they are complete trees and $\epsilon_t = 1$. The contribution of such trees to Δ_n is

$$(y_5 + G_6)[(G_1 + G_7)(G_2G_3 + G_2y_4 + G_3y_4) + G_1G_7(G_3 + y_4)].$$

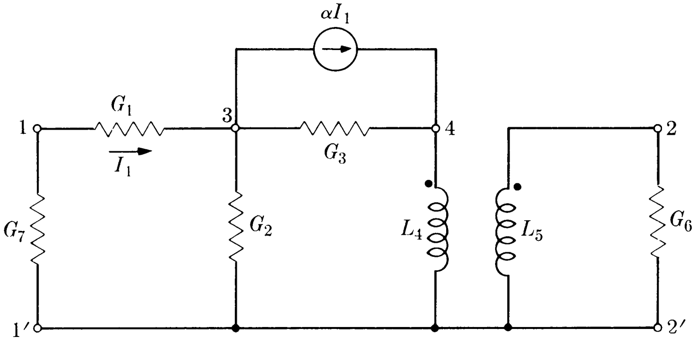


FIG. 7–18. Example for topological formulas.

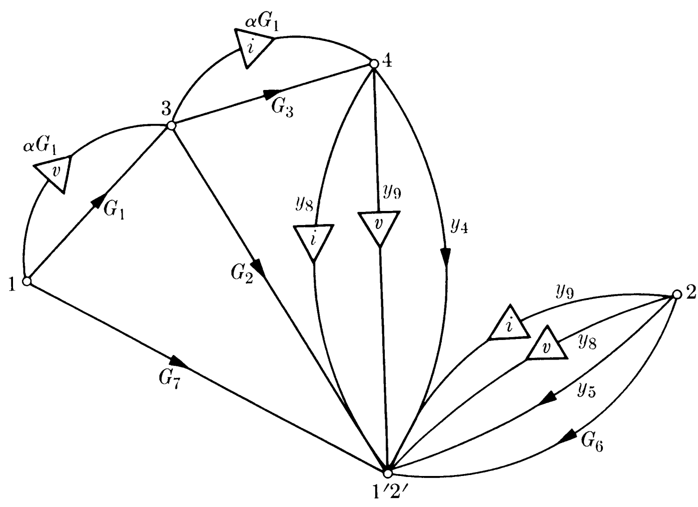


FIG. 7–19. Modified graph of Fig. 7–18.

The others are a little more difficult to find. The 2-tree concepts can be used to advantage to locate them. The complete trees containing the edge-pairs are

$\{\alpha G_1, G_7, y_4, y_5\}$, $\{\alpha G_1, G_7, y_4, G_6\}$, $\{y_8, y_9, G_2, G_1\}$, $\{y_8, y_9, G_2, G_7\}$,
$\{y_8, y_9, G_3, G_1\}$, $\{y_8, y_9, G_3, G_7\}$, $\{y_8, y_9, G_1, G_7\}$, $\{y_8, y_9, \alpha G_1, G_7\}$.

The sign permutations are determined next, as in Table 7–2. Hence the node determinant is given by

$$\begin{aligned}\Delta_n &= (y_5 + G_6)[(G_1 + G_7)(G_2G_3 + G_2y_4 + G_3y_4) + G_1G_7(G_3 + y_4)] \\ &\quad + \alpha G_1G_7y_8y_9 - \alpha G_1G_7y_4(y_5 + y_6) - y_8y_9(G_2 + G_3)(G_1 + G_7) - y_8y_9G_1G_7.\end{aligned} \tag{7–97}$$

TABLE 7–2

SIGN PERMUTATION FOR THE EXAMPLE OF FIG. 7–19

Tree	Sign permutation 1 2 3 4	ϵ_t
$\{\alpha G_1, G_7, y_4, y_5\}$	$\begin{pmatrix} G_7 & y_5 & \alpha G_1 & y_4 \\ G_7 & y_5 & \alpha G_1^- & y_4 \end{pmatrix}$	-1
$\{\alpha G_1, G_7, y_4, G_6\}$	$\begin{pmatrix} G_7 & G_6 & \alpha G_1 & y_4 \\ G_7 & G_6 & \alpha G_1^- & y_4 \end{pmatrix}$	-1
$\{y_8, y_9, G_2, G_1\}$	$\begin{pmatrix} G_1 & y_9 & G_2 & y_8 \\ G_1 & y_8 & G_2 & y_9 \end{pmatrix}$	-1
$\{y_8, y_9, G_2, G_7\}$	$\begin{pmatrix} G_7 & y_9 & G_2 & y_8 \\ G_7 & y_8 & G_2 & y_9 \end{pmatrix}$	-1
$\{y_8, y_9, G_3, G_1\}$	$\begin{pmatrix} G_1 & y_9 & G_3 & y_8 \\ G_1 & y_8 & G_3 & y_9 \end{pmatrix}$	-1
$\{y_8, y_9, G_3, G_7\}$	$\begin{pmatrix} G_7 & y_9 & G_3 & y_8 \\ G_7 & y_8 & G_3 & y_9 \end{pmatrix}$	-1
$\{y_8, y_9, G_1, G_7\}$	$\begin{pmatrix} G_7 & y_9 & G_1^- & y_8 \\ G_7 & y_8 & G_1^- & y_9 \end{pmatrix}$	-1
$\{y_8, y_9, \alpha G_1, G_7\}$	$\begin{pmatrix} G_7 & y_9 & G_1 & y_8 \\ G_7 & y_8 & G_1^- & y_9 \end{pmatrix}$	$+1$

It is seen that the topological formula for the general linear network is much more involved than the formula for a passive network without mutual inductances and consequently is more difficult to use for theoretical investigations. However, it is a useful procedure when computing machines are available.

Attention is next focused on the computation of the cofactors Δ_{kj} of the node-admittance matrix, which proceeds along similar lines. To avoid extremely complicated notation, the determination of the relative signs of the majors (which now correspond naturally to 2-trees) is here reduced to the sign permutation for trees (following Mayeda).

As before, the cofactor Δ_{kj} is given by

$$\Delta_{kj} = (-1)^{k+j} \det \mathbf{A}_{i_{-k}} \mathbf{Y} \mathbf{A}'_{v_{-j}}, \tag{7–98}$$

where $\mathbf{A}_{i_{-k}}$ is the current-incidence matrix with row k removed, and $\mathbf{A}_{v_{-j}}$ is the voltage-incidence matrix with row j removed. Both the symmetrical cofactors ($k = j$) and the asymmetrical cofactors ($k \neq j$) are included in the present discussion. From earlier discussions, nonsingular submatrices of A_{i_k} correspond one-to-one to the 2-trees of the current graph separating vertices k and (the reference) r; similarly for A_{v_j}. Thus nonzero terms of the expansion correspond to 2-trees, which are 2-trees (k, r) of the current graph and 2-trees (j, r) of the voltage graph.

DEFINITION 7-8. *Complete 2-tree* and *complete 2-tree product.* The edges $y_{p_1}, y_{p_2}, \ldots, y_{p_{v-2}}$ constitute a *complete 2-tree* $\binom{k,r}{j,r}$ of the graph if the current edges with these weights constitute a 2-tree separating vertices k and r of the current graph, and the voltage edges with these weights constitute a 2-tree separating vertices j and r of the voltage graph. The product of the edge weights $y_{p_1}y_{p_2}\cdots y_{p_{v-2}}$ of a complete 2-tree is a *complete 2-tree product.*

Thus,

$$\Delta_{kj} = (-1)^{k+j} \sum_{\substack{\text{all}\\ \text{complete}\\ \text{2-trees}}} \epsilon_t \times [\text{complete 2-tree } \tbinom{k,r}{j,r} \text{ product}], \qquad \epsilon_t = \pm 1. \tag{7-99}$$

It remains to determine when $\epsilon_t = 1$ and when $\epsilon_t = -1$. The simplest procedure is to convert the complete 2-tree into a complete tree, in such a way that the relative signs of the two majors of A_{i_k} and A_{v_j} are simply related to the relative signs of the majors of A_i and A_j. To this end, connect a current element y_0 between vertices k and r and the corresponding voltage element y_0 between vertices j and r. If $k = j$ (symmetrical-cofactor case), y_0 becomes an ordinary element. Let both elements y_0 be directed toward the reference vertex r. Given any complete 2-tree $\binom{k,r}{j,r}$, it is clear that the addition of y_0 makes this a complete tree. Conversely, if τ is any complete tree of the new graph containing y_0, it is clear that removing y_0 from τ leaves a complete 2-tree $\binom{k,r}{j,r}$. Thus complete 2-trees $\binom{k,r}{j,r}$ are in one-to-one correspondence with complete trees of the modified graph, which contains y_0. This is a useful computational procedure. Consider the majors of A_i and A_v corresponding to such a complete tree. Let y_0 occupy column p (in both matrices). Column p contains exactly one nonzero element, $+1$, in both majors (of A_i and A_v). This $+1$ is in row k of A_i and row j of A_v. Expand both majors by column p. Then,

$$\begin{aligned}
(\text{major of } \mathsf{A}_i) &= 1 \times (-1)^{k+p} \times [\text{minor } (k, p) \text{ of } \mathsf{A}_i\text{-submatrix}],\\
(\text{major of } \mathsf{A}_v) &= 1 \times (-1)^{j+p} \times [\text{minor } (j, p) \text{ of } \mathsf{A}_v\text{-submatrix}].
\end{aligned} \tag{7-100}$$

Now minor (k, p) of the A_i-submatrix is the major of A_{i_k} corresponding to the complete 2-tree in question. Similarly, minor (j, p) of A_v is the required major of A_{v_j}. Therefore, multiplying the corresponding sides of the two equations (7–100), we get

$$(\text{major of } \mathsf{A}_{i\_k}) \times (\text{major of } \mathsf{A}_{v\_j}) = (-1)^{k+j} \times (\text{major of } \mathsf{A}_i) \times (\text{major of } \mathsf{A}_v). \quad (7\text{–}101)$$

Taking into account the $(-1)^{k+j}$ in Eq. (7–98), we arrive at the next theorem.

THEOREM 7–16. If Δ_{kj} is the cofactor of the (k, j)-element of the node-admittance matrix $\mathsf{Y}_n(s)$ with r as the reference vertex,

$$\Delta_{kj} = \sum_{\substack{\text{all complete} \\ \text{2-trees } \binom{k,r}{j,r}}} \epsilon_t \,[\text{complete 2-tree } \tbinom{k,r}{j,r} \text{ product}],$$

where ϵ_t is the same as in Theorem 7–15, determined by adding a current element y_0 *from k to r* and a voltage element y_0 *from j to r*.

It is possible to carry this discussion further to cofactors of the type Δ_{1122} by defining 3-trees and 3-tree products as in the simpler case.*

EXAMPLE. Let us find Δ_{11} and Δ_{12} for the network of Fig. 7–18, with (1′, 2′) as reference, completing the computation of the driving-point and transfer impedances.

Δ_{11} is easy to find, since the added ordinary element is in parallel with G_7. Hence Δ_{11} simply contains the trees which contain G_7, with G_7 removed from them. Hence, from the expression for Δ_n of the preceding example,

$$\begin{aligned}\Delta_{11} = {} & (y_5 + G_6)[(G_1 + G_2)(G_3 + y_4) + G_3 y_4] \\ & + \alpha G_1 y_8 y_9 - \alpha G_1 y_4 (y_5 + y_6) - y_8 y_9 (G_2 + G_3) - y_8 y_9 G_1. \quad (7\text{–}102)\end{aligned}$$

For the cofactor Δ_{12}, add a current element y_0 from vertex 1 to vertex (1′, 2′) and a voltage element y_0 from vertex 2 to vertex (1′, 2′), resulting in Fig. 7–20.

It is clear that any complete tree containing y_0 must include y_9 to include vertex 2 in the current graph without having a circuit in the voltage graph. Inspection shows two complete trees containing y_0:

$$\{y_0, y_9, G_1, G_3\} \qquad \text{and} \qquad \{y_0, y_9, \alpha G_1, G_2\}.$$

The sign permutations of these two trees are, respectively,

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \qquad\qquad \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array}$$

$$\begin{pmatrix} y_0 & y_9 & G_1^- & G_3^- \\ G_1 & y_0 & G_3 & y_9 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_0 & y_9 & G_2 & \alpha G_1^- \\ \alpha G_1 & y_0 & G_2 & y_9 \end{pmatrix}.$$

---

* See Mayeda [111] or Coates [36] for further development of the subject.

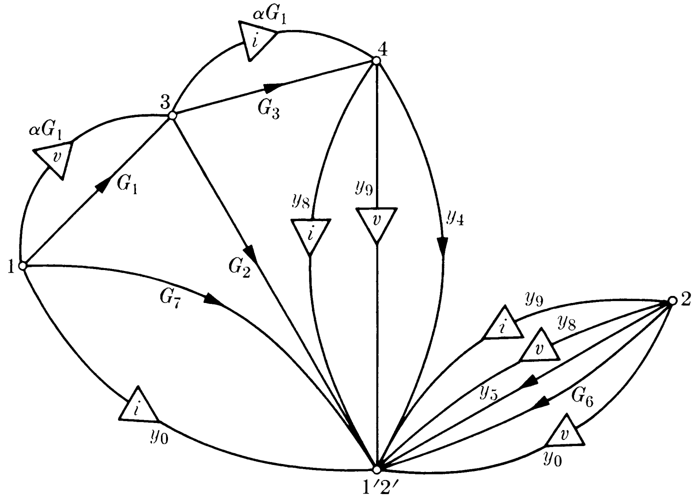


FIG. 7-20. Computation of Δ_{12}.

As contrasted with the computation of Δ_n, the ordinary elements also change positions in this tree. Thus in general, the entire sign permutation has to be examined and no short cuts are possible. The coefficients ϵ_t for the two trees are now found from the sign permutations to be -1 and -1. Hence

$$\Delta_{12} = -y_9(G_1G_3 + \alpha G_1G_2). \tag{7-103}$$

The required network functions can now be computed.

PROBLEMS

7–1. Write out the node-admittance matrix of Fig. 7–1, with 1′ as the reference vertex, and compute Δ and Δ_{11} by conventional procedures. Compare with Eqs. (7–10) and (7–16).

7–2. Find the driving-point admittance of the network in Fig. 7–21 by using Maxwell's formulas.

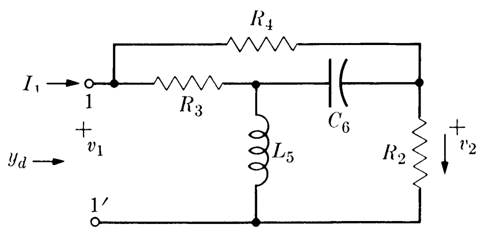


FIGURE 7–21

7–3. In Fig. 7–21, find the transfer impedance $Z_{21} = V_2/I_1$ (all initial conditions zero) by using Maxwell's rules.

7–4. Find the open-circuit impedance matrix and short-circuit admittance matrix of the networks in Fig. 7–22 by the use of topological formulas.

7–5. A technique for finding all the trees of a graph would be useful. Can you formulate one?

7–6. Calculate the inverse of the node-admittance matrix of Fig. 7–21 by the standard cofactor method. Repeat, using topological formulas.

7–7. Let N denote a passive network without mutual inductances, as in Fig. 7–23. The driving-point impedance of N at the terminals (1, 1′) is defined by

$$Z_d(s) = \frac{E_1(s)}{I_1(s)}, \qquad \text{all initial conditions zero.}$$

By using the formulas for the determinant and cofactor of the loop-impedance matrix, show that

$$Z_d(s) = \frac{C[W_{1,1'}(Z)]}{C[V(Z)]},$$

where the polynomials V and W and the complements are computed for the network N only.

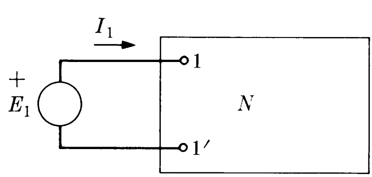


FIGURE 7–23

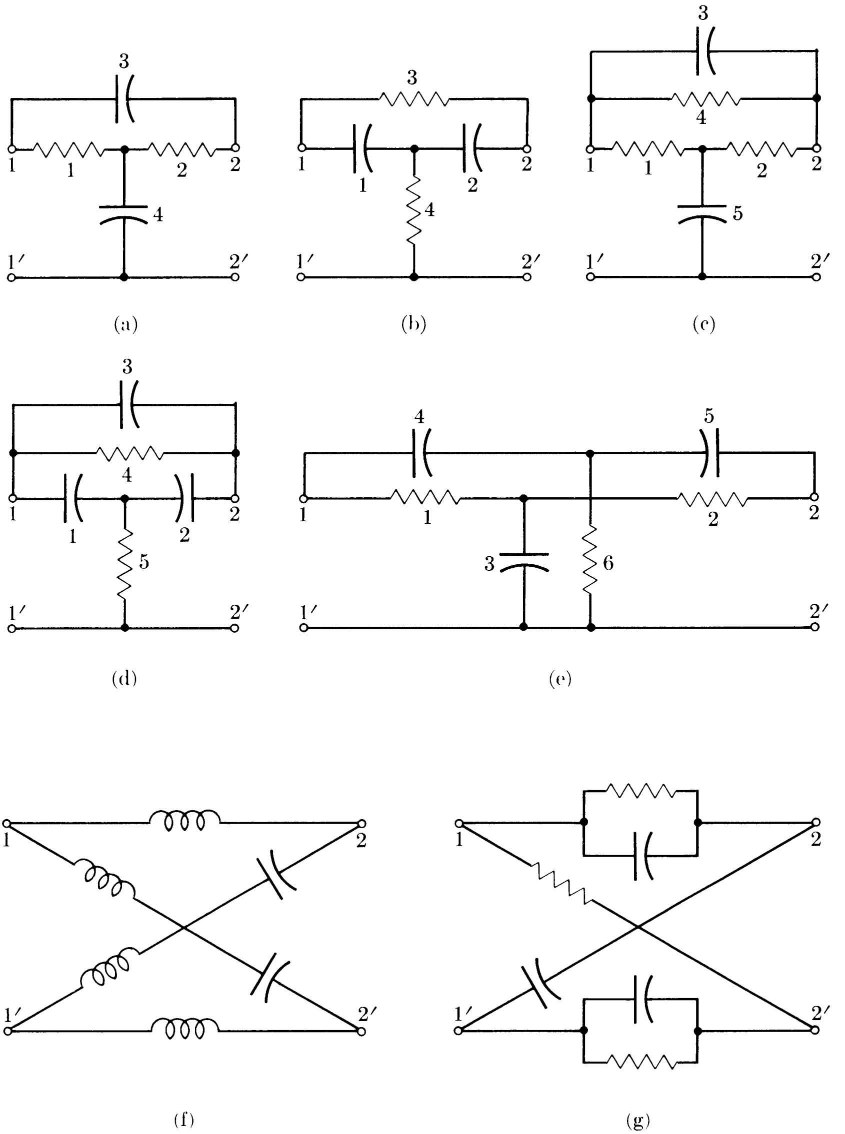


FIGURE 7–22

7–8. If G and G^* are dual graphs, show that trees of either graph correspond to tree complements in the other. Using this fact, Theorem 3–13, and Problem 7–7, obtain an alternative proof of the well-known geometrical procedure of obtaining inverse networks by duality.

7–9. How can we use Maxwell's and Kirchhoff's formulas to compute response when there are several generators in the network?

7–10. Obtain the open- and short-circuit functions of Figs. 7–22(e) and (g) by direct determinant computations and compare with the topological formulas.

7–11. Prove that the 2-trees that appear in the numerator of z_{12} (or y_{12}) are precisely those 2-trees which are common to the numerator of both z_{11} and z_{22} (or y_{11} and y_{22}). (We use this result to derive the powerful theorems of Fialkow and Gerst in the next chapter.)

7–12. From the topological formulas for $\mathbf{Z}_{oc}$ and $\mathbf{Y}_{sc}$, derive the formulas for the admittances of the T- and π-equivalents of two terminal-pair networks.

7–13. Prove that (number of trees of a graph) $=$ det $\mathbf{AA}'$. (Trent [176].)

7–14. Prove that det $\mathbf{B}_f\mathbf{B}_f' =$ (number of trees of graph).

7–15. It can be shown that the matrix

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$$

is nonsingular for nondirected and for directed graphs. Do so. Prove further that in directed graphs

$$\det \begin{bmatrix} \mathbf{A} \\ \mathbf{B}_f \end{bmatrix} = \pm \text{ (number of trees of graph).}$$

[*Hint:* Postmultiply the matrix

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B}_f \end{bmatrix}$$

by its transpose, and use Theorem 5–4 and Problems 7–13 and 7–14.]

7–16. Compute the driving-point impedance at terminals (1, 1′) and the open-circuit transfer impedance z_{21} of the network of Fig. 7–24 by topological formulas.

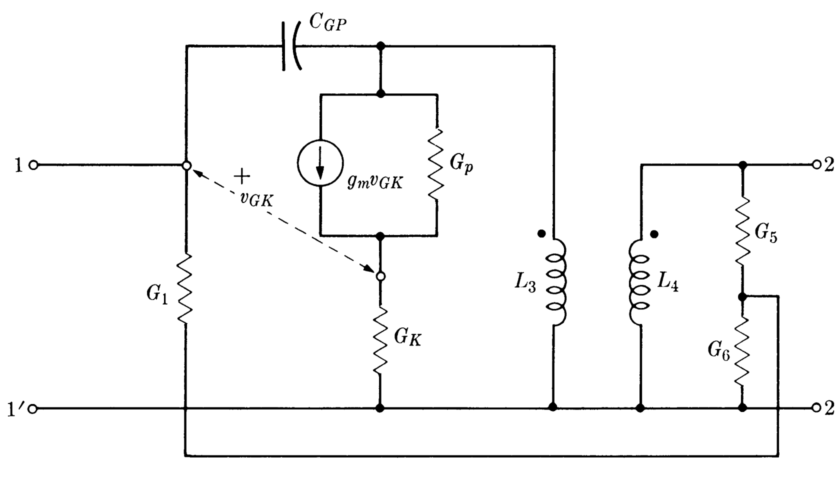


FIGURE 7–24

CHAPTER 8

APPLICATIONS TO NETWORK SYNTHESIS

As remarked earlier, the application of graph theory to problems in network synthesis is a recent development. In this chapter, most of the known applications are reviewed. Since the subject matter is of recent origin, the chapter is more in the nature of a report on the "state of the art" rather than a well-organized logical development. The entire chapter is concerned with passive reciprocal networks without mutual inductances; only R-, L-, and C-elements are admitted.

8–1 Enumeration of natural frequencies. The zeros of the network determinant (either loop or node) are referred to as the *natural frequencies* of the network, for these are the frequencies of the transient response. One of the classical problems is to count the number of natural frequencies of a network by inspection. An early solution to this problem is an algorithm due to Guillemin [69], applicable to networks which contain no all-capacitor or all-inductor loops. More recently, Reza [145] gave the solution for networks containing only two types of elements. The complete solution was obtained independently by Bryant [18], Bers [10], and Seshu (in unpublished notes, 1958).

DEFINITION 8–1. *Order of complexity.* The *order of complexity* of a network is the number of finite nonzero zeros of the determinant (loop or node), with each R, L, and C considered as a network element.

Since by Theorem 7–10, the loop and node determinants are related by

$$\Delta_m = (\det \mathbf{Z})\, \Delta_n = ks^m\, \Delta_n, \tag{8–1}$$

where m is an integer (positive, negative, or zero), the loop and node determinants have the same zeros, excluding $s = 0, \infty$. Thus Definition 8–1, which is due to Reza [145], is meaningful. Therefore also, either Δ_m or Δ_n may be chosen as the basis of the enumeration formula. In the present discussion, the node determinant Δ_n is chosen as the basis. Let

$$\Delta_n = a_k s^k + a_{k-1} s^{k-1} + \cdots + a_0 + a_{-1} s^{-1} + \cdots + a_{-p} s^{-p}, \tag{8–2}$$

where $a_k \neq 0$, $a_{-p} \neq 0$, and p and k are integers. The order of complexity is evidently $k + p$. Hence it suffices to determine k and p. Since

$$\Delta_n = \sum_{\substack{\text{all}\\ \text{trees}}} (\text{tree-admittance products}), \tag{8–3}$$

it is clear that

$$k = (\text{number of capacitors in } T_1) - (\text{number of inductors in } T_1), \tag{8–4}$$

where T_1 is the tree that maximizes this difference. Similarly,

$$p = (\text{number of inductors in } T_2) - (\text{number of capacitors in } T_2), \tag{8–5}$$

where T_2 maximizes this difference.

However, by Theorem 6–10, it is possible simultaneously to maximize the number of capacitors and minimize the number of inductors in T_1, as shown by the following argument. Let G be the graph of the network. Let α_L be the maximum number of inductors contained in any tree of the graph G; let S_1 be the subgraph (of G) of such a set of α_L inductors. S_1 thus contains no circuits. Let β_C be the smallest number of capacitors contained in any tree of G; let T be a tree containing only β_C capacitors. Let S_2 be the subgraph of all capacitors *excluding* the capacitors contained in T. S_2 thus contains no cut-sets of G (Problem 2–20). S_1 and S_2 are clearly edge-disjoint, since S_1 contains only inductors and S_2 contains only capacitors. Hence, by Theorem 6–10, there exists a tree T_1 with edges of S_1 as (some of the) branches and for which the edges of S_2 are (some of the) chords. Hence T_1 contains α_L inductors and β_C capacitors.

It remains only to establish the numbers α_L and β_C. Let n_L be the number of inductors in the graph G. Consider the subgraph G_L consisting of these inductors. If G_L contains any circuits, then G_L is not contained in any tree of G. Let μ_{Lo} be the nullity of G_L. Then at least μ_{Lo} edges must be removed from G_L to destroy all circuits; also removing a suitable set of μ_{Lo} edges destroys all circuits. (Construct a forest of G_L to appreciate this fact.) Hence

$$\alpha_L = n_L - \mu_{Lo}. \tag{8–6}$$

Thus α_L is the rank of G_L.

The number β_C is established by a dual argument, which argument is, however, spelled out in detail in the next two paragraphs.

Find the all-capacitor cut-sets of the graph G. These cut-sets are referred to in the sequel as *cut-sets of* G, *contained in* G_C, where G_C is the subgraph of all capacitors. Let δ of these be linearly independent. That is, δ is the rank of the matrix $\mathbf{Q}_{Ca}$ of these all-capacitor cut-sets of G. Then at least δ capacitors must be removed from G_C to destroy all the cut-sets of G contained in G_C. For, suppose on the contrary that $e_1, e_2, \ldots, e_r$, where $r < \delta$, are a set of edges of G_C such that the subgraph G'_C obtained by removing $e_1, e_2, \ldots, e_r$ contains no cut-sets of G. Select a submatrix $\mathbf{Q}_C$ of $\mathbf{Q}_{Ca}$ of δ rows and rank δ. Consider the submatrix of $\mathbf{Q}_C$ consisting of columns corresponding to $e_1, e_2, \ldots, e_r$. Since this submatrix contains only r columns, its rank is at most r. Since it contains δ rows, and $\delta > r$, the rows of this submatrix are linearly dependent. Hence a suitable linear combination of the rows of this submatrix (with not all of the coefficients zero) yields a row of zeros. Consider the same linear combination of the rows of $\mathbf{Q}_C$. Since $\mathbf{Q}_C$ has rank δ, the rows of $\mathbf{Q}_C$ are linearly independent, and so the linear combination is nonzero. Hence it is a cut-set or disjoint union of cut-sets of G, which does not contain any of $e_1, e_2, \ldots, e_r$. Thus, at least δ capacitors must be removed from G_C to destroy the cut-sets of G contained in G_C.

Also, it suffices to remove a suitable set of δ capacitors. For this, find a nonsingular submatrix of $\mathbf{Q}_C$ of order δ. No linear combination of the rows of this submatrix can be zero (unless all the coefficients are zero). Hence, if the δ edges corresponding to the columns of this nonsingular submatrix are removed from G_C, the remaining subgraph G'_C contains no cut-sets of G. Evidently, $\beta_C = \delta$. For later use, we restate these results in more general terminology.

LEMMA 8-1. Let G_s be a subgraph of a connected graph G. Then the maximum number α_s of edges of G_s contained in any tree of G is given by

$$\alpha_s = n_s - \mu_{so} = \rho_{so}, \tag{8-7}$$

where n_s is the number of edges in G_s, μ_{so} is the nullity of G_s, and ρ_{so} is the rank of G_s. The minimum number β_s of edges of G_s contained in any tree of G is the number of linearly independent cut-sets of G, contained in G_s.

It is convenient to use a formal method of computing the number of linearly independent cut-sets of G contained in a subgraph and thereby introduce some useful notation. Let G_s be, for instance, the L-subgraph, i.e., the subgraph consisting of all the inductors. Consider short-circuiting all elements of G which are not inductors. From the vertex-partitioning interpretation of a cut-set given in Section 2-4 (immediately preceding Definition 2-12), it is clear that every all-inductor cut-set of G remains a cut-set of the resultant graph. Since the new graph contains only inductors, the number of linearly independent L-cut-sets of G is equal to the rank of the graph obtained by short-circuiting all other elements. Let this number be denoted by ρ_{Ls}, with a similar meaning for ρ_{Cs}. Similarly, the rank of the L-subgraph [= (number of inductors) − (nullity of L-subgraph)] can be found by open-circuiting all other elements and can be denoted by ρ_{Lo}, with a similar meaning for ρ_{Co}. Similarly, μ_{Lo} is the nullity of the graph obtained by deleting (open-circuiting) all non-L-elements, etc.

THEOREM 8–1. The order of complexity of a passive network without mutual inductances is

$$N = \rho_{Co} + \rho_{Lo} - \rho_{Cs} - \rho_{Ls}, \tag{8–8}$$

with the above notation.

Proof. By Lemma 8–1, the maximum number of capacitors contained in any tree is ρ_{Co}, and the minimum number of inductors in any tree is ρ_{Ls}. Since the L-subgraph and the C-subgraph are edge-disjoint, by Theorem 6–10, there exists a tree T_1 with ρ_{Co} capacitors and ρ_{Ls} inductors. Hence

$$k = \rho_{Co} - \rho_{Ls}. \tag{8–9a}$$

Similarly,

$$p = \rho_{Lo} - \rho_{Cs}, \tag{8–9b}$$

and the theorem follows.

Since (rank) + (nullity) = (number of edges) in any graph, Eq. (8–8) can be algebraically manipulated to give other useful forms.

COROLLARY 8–1(a). With the same hypotheses and notation as in the theorem,

$$N = \mu_{Cs} + \mu_{Ls} - \mu_{Co} - \mu_{Lo}, \tag{8–10a}$$

or

$$N = (\mu_{Ls} - \mu_{Lo}) + (\rho_{Co} - \rho_{Cs}), \tag{8–10b}$$

where all inductors and capacitors are retained in the Ls- and Cs-subgraphs, and single-edge loops are counted as loops.

Guillemin's [69] algorithm gives $N = \mu_{Cs} + \mu_{Ls}$, which agrees with Eq. (8–10a) under Guillemin's assumption of no all-inductor or all-capacitor loops. Equation (8–10b) was conjectured by S. J. Mason and R. Adler (quoted by Reza [145]).

COROLLARY 8–1(b) (Reza). For an LC-network,

$$N = 2(\mu - \mu_{Co} - \mu_{Lo}). \tag{8–11}$$

The proof of Corollary 8–1(b) is left as a problem (Problem 8–1). The number $\mu - \mu_{Co} - \mu_{Lo}$ was named the number of *dynamically independent loops* by Reza [145].

8–2 One terminal-pair networks.

8–2 One terminal-pair networks. This section is devoted to the known theory of minimality in transformerless realizations of driving-point functions. The emphasis is again on the relationship between the structure of the network and the corresponding network function. Theorem 8–1 and Eq. (8–8) give the data regarding the degree of the polynomial in Δ_n. By shorting the input vertices and using the same formula, the highest and lowest powers in Δ_{11} can be determined. Hence, *if there is no cancellation* of common factors between Δ and Δ_{11}, the highest power occurring in the driving-point function is known. It is evident from Theorem 8–1 that the order of complexity cannot exceed the total number of inductors and capacitors in the network. These considerations lead to the first concept in minimality.

DEFINITION 8–2. *Minimal in reactive elements.* A transformerless realization of the positive real function

$$Z(s) = \frac{c_n s^n + c_{n-1}s^{n-1} + \cdots + c_1 s + c_0}{d_m s^m + d_{m-1}s^{m-1} + \cdots + d_1 s + d_0} \tag{8–12}$$

with $c_n \neq 0$, $d_m \neq 0$, and not both $c_0, d_0 = 0$, is *minimal in reactive elements* if either

$$n = n_L + n_C \qquad \text{or} \qquad m = n_L + n_C \tag{8–13}$$

or both, where n_L and n_C are the numbers of inductors and capacitors in the realization. [The numerator and denominator in Eq. (8–12) are assumed not to have any common factors.]

It is clear from Theorem 8–1 that all-inductor and all-capacitor loops and cut-sets must be prohibited either in the original network or in the modified network (with the input terminals shorted) if the network is minimal in reactive elements. This concept is made precise in the next theorem. The notation "cut-set (1, 1′)" signifies a cut-set which places the input vertices 1 and 1′ in different connected parts.

THEOREM 8–2. A one terminal-pair network N without mutual inductances realizing $Z(s)$ of Eq. (8–12), in which $\Delta(s)$ and $\Delta_{11}(s)$ have no common factors, is minimal in reactive elements if and only if

(a) there are no all-inductor or all-capacitor loops in N,
(b) there are no all-inductor or all-capacitor cut-sets in N other than cut-sets (1, 1′), and
(c) there are no all-inductor or all-capacitor cut-sets in N_1 (when vertices 1 and 1′ are identified).

[Condition (b) is really implied by (c), but it is more convenient to state it separately.]

Proof. The sufficiency of the condition is taken up first. The necessity then becomes evident. Let

$$\Delta(s) = a_k s^k + a_{k-1}s^{k-1} + \cdots + a_{-p}s^{-p}$$

and

$$\Delta_{11}(s) = b_q s^q + b_{q-1}s^{q-1} + \cdots + b_{-r}s^{-r}. \tag{8–14}$$

There are four cases to be considered, depending on the existence, or otherwise, of paths between the input vertices 1 and $1'$ consisting only of inductors or capacitors. Such paths are denoted by L-path $(1, 1')$ and C-path $(1, 1')$.

Case 1. There is an L-path $(1, 1')$ and a C-path $(1, 1')$. It immediately follows that there is no C-cut-set $(1, 1')$ (since there is an L-path) and no L-cut-set $(1, 1')$. Hence from Theorem 8–1,

$$k = n_C, \qquad p = n_L,$$

and

$$\Delta(s) = \frac{d_{n_L+n_C}s^{n_L+n_C} + \cdots + d_0}{s^{n_L}} \tag{8–15}$$

The proof for this case is complete, but let us compute q and r as an illustration. It is clear, from condition (a), that there is only one L-path $(1, 1')$ and only one C-path $(1, 1')$. These two become loops, on identifying vertices 1 and $1'$, so that

$$q = n_C - 1 \qquad \text{and} \qquad r = n_L - 1. \tag{8–16}$$

Thus $Z(s)$ has a zero at $s = 0$ and another zero at $s = \infty$, as we expect.

Case 2. There is an L-path $(1, 1')$ but no C-path $(1, 1')$. Then there is no C-cut-set $(1, 1')$, but there may be an L-cut-set $(1, 1')$. If there is no L-cut-set $(1, 1')$, we have immediately that

$$k = n_C, \qquad p = n_L, \qquad q = n_C, \qquad r = n_L - 1, \tag{8–17}$$

so that

$$n = n_L + n_C, \qquad m = n_L + n_C, \qquad c_0 = 0, \tag{8–18}$$

with reference to Eq. (8–12). Thus $Z(s)$ has a zero at $s = 0$, and is regular and nonzero at $s = \infty$. If there is an L-cut-set $(1, 1')$, there can be only one. For if Q_1 and Q_2 are two L-cut-sets $(1, 1')$, their mod 2 sum $Q_1 \oplus Q_2$ has an even number of edges in common with a path $(1, 1')$ since each of Q_1 and Q_2 has an odd number of edges in common with such a path. Thus $Q_1 \oplus Q_2$ is not a cut-set $(1, 1')$. However, $Q_1 \oplus Q_2$ is a cut-set or disjoint union of cut-sets. [Two of these may be cut-sets $(1, 1')$.] Thus, N contains L-cut-sets that are not cut-sets $(1, 1')$. Since there is only one L-cut-set, and one L-path $(1, 1')$,

$$k = n_C - 1, \qquad p = n_L, \qquad q = n_C, \qquad r = n_L - 1. \tag{8–19}$$

In this case, the numerator of $Z(s)$ is of degree $n_L + n_C$, which completes the proof. $Z(s)$ has a zero at $s = 0$ and a pole at $s = \infty$.

Case 3. There is a C-path $(1, 1')$ but no L-path $(1, 1')$. This case is obtained by interchanging inductors and capacitors in Case 2. Now $Z(s)$ has a zero at $s = \infty$. $Z(s)$ has a pole at $s = 0$ if there is a C-cut-set $(1, 1')$. Otherwise, $Z(s)$ is regular and nonzero at $s = 0$.

Case 4. There is neither a C-path $(1, 1')$ nor an L-path $(1, 1')$. In this case we have immediately that

$$q = n_C \qquad \text{and} \qquad r = n_L, \tag{8–20}$$

so that the maximum degree appears in the numerator of $Z(s)$. As before, we find that L- and C-cut-sets $(1, 1')$ lead to poles at $s = \infty$ and $s = 0$, respectively, the two being independent of each other. $Z(s)$ is regular and nonzero at the appropriate point in the absence of such cut-sets.

The necessity part of the theorem can now be left as evident.

We can generalize the by-products of Theorem 8–2 to nonminimal structures, leading to the intuitively obvious result of Theorem 8–3.

THEOREM 8–3. The driving-point impedance of a passive network without mutual inductances has

(a) a pole at $s = 0$ if and only if there is a C-cut-set $(1, 1')$,

(b) a pole at $s = \infty$ if and only if there is an L-cut-set $(1, 1')$,

(c) a zero at $s = 0$ if and only if there is an L-path $(1, 1')$, and

(d) a zero at $s = \infty$ if and only if there is a C-path $(1, 1')$.

As a trivial application of Theorem 8–2, one can prove the well-known result that the Foster and Cauer realizations of reactance functions are minimal; and by the Cauer transformations (see Section 8–3) extend the result to RC- and RL-networks.

There is no implication in Theorem 8–2 that every positive real function $Z(s)$ has a realization containing only the minimal number of reactive elements required. It is more difficult to establish the minimal number of resistors required. If the given positive real function is not a reactance function, resistors are certainly required. If mutual inductances are permitted, the Darlington synthesis procedure realizes any positive real function with one resistor. If mutual inductance is not permitted, one resistor is not always sufficient. However, the only known result about resistors is the following.

THEOREM 8–4. If the positive real function $Z(s)$ is regular and nonzero at $s = 0$ and $s = \infty$, and

$$Z(0) \neq Z(\infty), \tag{8–21}$$

then any transformerless realization of $Z(s)$ requires at least two resistors.

This result is intuitively obvious, since the inductors and capacitors become open or short circuits at $s = 0$ and $s = \infty$. The formal proof based on topological formulas is left as a problem (Problem 8–2).

As mentioned earlier, little is known about minimality of the various network realizations. The only RLC-network that has been proved strictly minimal is the seven-element realization of the biquadratic minimum positive real function. This proof is outlined next. The complete details may be found elsewhere [157].

A *minimum positive real function* is a positive real function which has no poles or zeros on the $j\omega$-axis and which has a zero real part at some point $j\omega_0$ ($\neq 0, \infty$) on the imaginary axis. The next theorem states an important fact about the structure of the realization of a minimum p.r. function.

THEOREM 8–5. A transformerless realization of a minimum p.r. function does not contain any paths $(1, 1')$ or cut-sets $(1, 1')$ consisting of one type of element only (i.e., all-inductor, or all-capacitor, or all-resistor).

It already has been shown that L- and C-paths $(1, 1')$ and cut-sets $(1, 1')$ cannot exist in such a network (Theorem 8–3). To show that an all-resistor path $(1, 1')$ or cut-set $(1, 1')$ cannot exist, consider the expression for $Z(s)$ in terms of energy functions (Eq. 6–120):

$$Z(j\omega) = \frac{1}{|I_1|^2}\left[F_0(j\omega) + j\omega T_0(j\omega) + \frac{1}{j\omega} V_0(j\omega)\right], \qquad (8\text{–}22)$$

where I_1 is the a-c steady-state input current (phasor) and F_0, T_0, and V_0 are the energy functions (quadratic forms associated with the R-, L-, and D-matrices). At the minimum point $j\omega_0$, $F_0(j\omega_0) = 0$, so that all the resistor currents must be zero, and hence also the voltages across the resistors. An R-path $(1, 1')$ now makes the input voltage zero so that $Z(s)$ has a zero at $j\omega_0$. On the other hand, an R-cut-set $(1, 1')$ makes the input current zero, since the input current is the sum (taking references into account) of the cut-set-$(1, 1')$ currents. Thus $Z(s)$ has a pole at $j\omega_0$. Since neither a zero nor a pole can exist on the $j\omega$-axis, the result is proved.

COROLLARY 8–5(a). A transformerless realization of a minimum p.r. function contains paths $(1, 1')$ and cut-sets $(1, 1')$ consisting of any two types of elements (RC, RL, or LC).

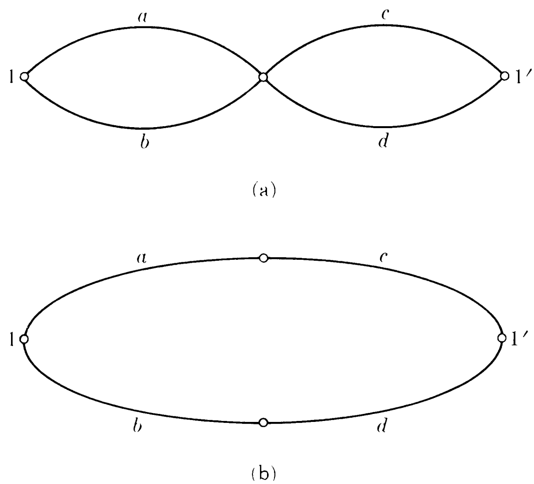


FIG. 8-1. Four-element structures.

COROLLARY 8-5(b). A transformerless realization of a minimum p.r. function contains no single-element path (1, 1′) or cut-set (1, 1′).

Thus in the terminology of Moore and Shannon [118], the *length* and *width* of such a network are at least 2. [The length is the number of elements in the shortest path (1, 1′), and the width is the number of elements in the smallest cut-set (1, 1′)]. Hence by the results of Moore and Shannon, the network contains at least four elements. However, the two four-element graphs of length 2 and width 2 shown in Fig. 8-1 cannot realize minimum p.r. functions for any assignment of R, L, and C, as shown by the next theorem.

THEOREM 8-6. A minimum p.r. function cannot be represented in a transformerless realization as a series or a parallel combination of two networks, one of which contains only two types of elements.

This result is obvious. For, if the two elements are RL or RC, the real part is nonzero at all finite nonzero frequencies; and if the two elements are LC, there is at least one pole on the $j\omega$-axis; and neither can be removed by the addition of another positive real function.

COROLLARY 8-6. A transformerless realization of a minimum p.r. function cannot be a series combination of (a) a parallel connection of two elements and (b) another network; nor can it be a parallel combination of (a) a series connection of two elements and (b) another network.

Thus the smallest network that can realize a minimum p.r. function is a five-element bridge. For example, the driving-point impedance of

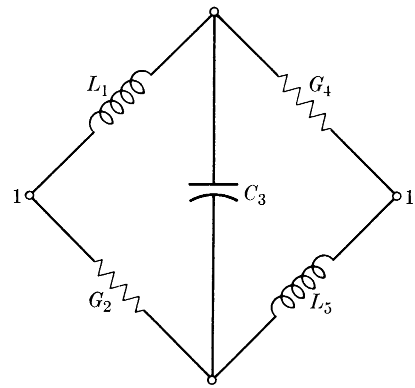


FIG. 8–2. A bridge realizing a minimum function.

the bridge network of Fig. 8–2 is a minimum p.r. function whenever $L_1 = L_5$ and $G_2 = G_4$.

The biquadratic minimum p.r. function is

$$Z(s) = R\,\frac{s^2 + a_1 s + a_0}{s^2 + b_1 s + b_0}, \tag{8–23}$$

where the a's and b's are real and positive, and satisfy

$$a_1 b_1 = (\sqrt{a_0} - \sqrt{b_0})^2. \tag{8–24}$$

Since neither a_1 nor b_1 can be zero (no poles or zeros on the $j\omega$-axis), Eq. (8–24) establishes the next theorem immediately.

THEOREM 8–7. If $Z(s)$ is a biquadratic minimum p.r. function, $Z(0) \neq Z(\infty)$, so at least two resistors are required in any transformerless realization of $Z(s)$.

Theorem 8–8 treats reactive elements.

THEOREM 8–8. At least three reactive elements are required in any transformerless realization of a biquadratic minimum p.r. function.

Proof. On the contrary, suppose that two reactive elements suffice for some biquadratic minimum function. Then these two are necessarily an L and a C. By Corollary 8–5(a), there exists an LC-path $(1, 1')$. Hence the two reactive elements must be connected as shown in Fig. 8–3, or with vertices 1 and 1′ interchanged. Now, however, there can be no LC-cut-set $(1, 1')$ (that is not an edge-disjoint union of cut-sets).

COROLLARY 8–8. If a network N without transformers realizes the biquadratic minimum function $Z(s)$, the node determinant Δ of the

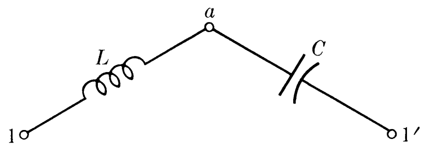


FIG. 8–3. Illustration of proof.

network (and therefore Δ_{11}) is at least a cubic polynomial divided by a power of s.

In other words, at least three of the reactive elements "contribute degree" to Δ. The proof of Corollary 8–8 follows directly from Theorem 8–1, its corollaries, and Theorem 8–8, and so is left as a problem (Problem 8–12).

Thus a common factor or factors must necessarily cancel between Δ and Δ_{11} if the network realizes a biquadratic minimum function. However, in the case of a bicubic or biquartic minimum function, this does not always happen (see Kim [84]). The known minimality argument [157] now proceeds by exhausting all possibilities, making use of the structure theorems 8–5 and 8–6. Since the four-element graphs have been disposed of by Theorem 8–6, the search begins with five-element graphs. Except for the five-element bridge (graph of Fig. 8–2), the other five-element graphs (of length 2 and width 2) are series or parallel combinations of two graphs, one of which consists of two edges. Thus, by Corollary 8–6, only the bridge graph needs to be considered. Next, all possible assignments of elements R, L, and C to the edges of the bridge must be considered. Many of these are eliminated by Corollary 8–5(a). Only six structures remain, arranged in dual pairs, so only three networks require detailed examination. The element values are computed from the behaviors at the frequencies 0, ∞, and $j\omega_0$. At $j\omega_0$, the network must be reactive (the resistors becoming shorted by series-resonant circuits or

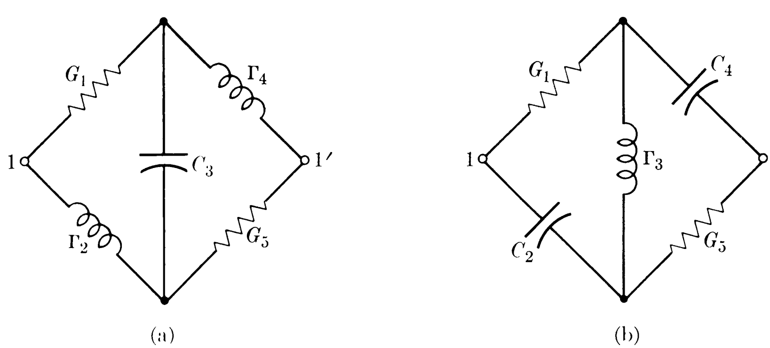


FIG. 8–4. Five-element bridges. (a) $Z(\infty) = 4Z(0)$. (b) $Z(0) = 4Z(\infty)$.

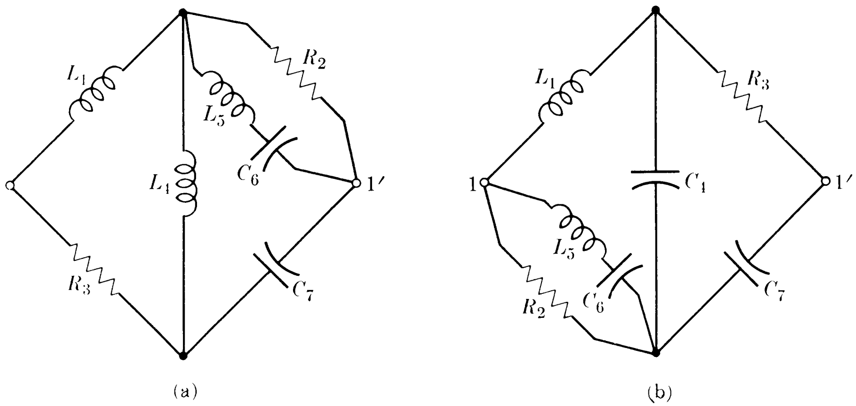


FIG. 8–5. Minimal realization of biquadratic. (a) $x_0 > 0$. (b) $x_0 < 0$.

opened by parallel-resonant circuits) with a reactance equal to $X_0 = (1/j)Z(j\omega_0)$. Finally, the two functions [the given biquadratic and $Z_d(s)$ of the bridge] are set equal, and the condition for the cancellation of a common factor in the numerator and denominator of $Z_d(s)$ is found. It is discovered that only the biquadratics which satisfy $Z(0) = 4Z(\infty)$ or $Z(\infty) = 4Z(0)$ can be realized with five elements, by the networks shown in Fig 8–4.

The search is continued now with six-element graphs. Again the series-parallel structures are eliminated by Corollary 8–6 and its extension to three elements. Four bridge graphs remain to be examined in detail, which are in dual pairs. Now R's, L's, and C's are assigned to these bridges in all ways possible without violating Theorem 8–5. Computation of element values and the check for cancellation of common factors proceed as before. It is found that no additional biquadratics are realized by six-element networks. Thus it is proved that the known seven-element realization of a biquadratic minimum function, shown in Fig. 8–5, is minimal except for the special cases shown in Fig. 8–4.

8–3 Two terminal-pair networks. In the theory of two terminal-pair structures, the adjective *minimum-phase* plays a significant role. This term is defined as follows.

DEFINITION 8–3. *Minimum phase.* A two terminal-pair network is a *minimum-phase* network if none of the zeros of y_{12} (or z_{12}) is in the open right half-s-plane. (That is, y_{12} has no zeros in $\sigma > 0$.)

Bode [12], who originated the name *minimum phase,* uses [(input)/(output)]-functions and so uses *poles* (rather than *zeros*) in its definition. The modern standard is to use (output)/(input), as is done here.

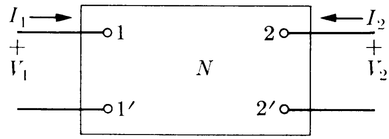


FIG. 8-6. Two terminal-pair network.

To correlate the present treatment with conventional treatments of minimum-phase networks, let us temporarily change the convention about network elements. Instead of considering each R, L, and C as a network element, let us consider each two-terminal subnetwork to be a network element. Thus the networks being considered do not have any vertices of degree 2, except possibly the input and output vertices, and no two elements are in parallel. With this convention, we define the elusive term *number of transmission paths* as follows.

DEFINITION 8-4. *Number of transmission paths.* Let N be a two terminal-pair passive network with no mutual inductances. With the convention above about network elements, the *number of paths of transmission* is the number of 2-trees in N of the types $T_{2_{12,1'2'}}$ *and* $T_{2_{12',1'2}}$. (See Fig. 8-6.)

The convention introduced does not change any of the topological formulas developed in Chapter 7. We made no assumptions in Chapter 7 about what a "network element" is. With this definition, we can prove the well-known result stated in Theorem 8-9.

THEOREM 8-9. Every two terminal-pair network without mutual inductances that contains only one path of transmission is a minimum-phase network.

The proof is sufficiently obvious to be omitted.

The best-known example of a network with a single transmission path is a ladder network. A ladder is, by definition, a network of the type shown in Fig. 8-7, in which each Z_i and Y_j may be a complex two-terminal network and there is no mutual coupling. Either or both of Z_1

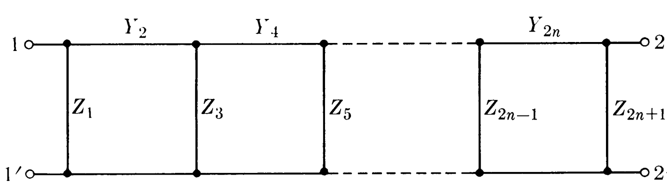


FIG. 8-7. Ladder network.

and Z_{2n+1} may be absent. For such a ladder network, we have the following theorem.

THEOREM 8–10. The zeros of transmission of a ladder network are contained among the zeros of $Y_2, Y_4, \ldots, Y_{2n}$ and $Z_1, Z_3, \ldots, Z_{2n+1}$.

Proof. For a ladder network (Fig. 8–7),

$$W_{12,1'2'} = Y_2 Y_4 \cdots Y_{2n} \qquad \text{and} \qquad W_{12',1'2} = 0. \tag{8–25}$$

The zeros of transmission are, by definition, the zeros of y_{12} (or z_{12}). Since

$$z_{12} = \frac{W_{12,1'2'}(Y)}{V(Y)}, \tag{8–26}$$

the zeros of z_{12} are the zeros of $W_{12,1'2'}(Y)$ and poles of $V(Y)$, except those that cancel. Now every tree of the network must contain the vertex $(1'2')$, and so each term of $V(Y)$ has one (or more) of $Y_1, Y_3, \ldots, Y_{2n+1}$ as a factor. Thus a pole of $Y_1, Y_3, \ldots, Y_{2n+1}$ is also a pole of $V(Y)$. (These are, of course, the zeros of $Z_1, Z_3, \ldots, Z_{2n+1}$.) Any other pole of $V(Y)$ is a pole of one of $Y_2, Y_4, \ldots, Y_{2n}$, which automatically cancels with a pole of $W_{12,1'2'}(Y)$. Thus the theorem is proved.

Theorem 8–10 can also be considered to be obvious from physical intuition, since an open circuit of one of the series arms (the even-numbered elements) or a short circuit of one of the shunt arms (the odd-numbered elements) is required for a zero of transmission. [The question of possible cancellation of zeros of $W_{12,1'2'}(Y)$ and $V(Y)$ is much more difficult and is not discussed here. It is important, however, and is used in the Cauer ladder development of two terminal-pair networks.]

Thus a nonminimum-phase structure necessarily has more than one 2-tree of the types $T_{2_{12,1'2'}}$ and $T_{2_{12',1'2}}$, or contains a transformer. The simplest examples of nonminimum-phase structures are bridged-T and lattice networks, as shown in Fig. 8–8. As can be verified, the zeros of

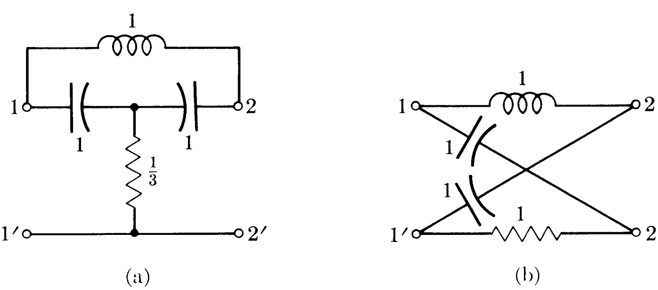


FIG. 8–8. Two nonminimum-phase networks.

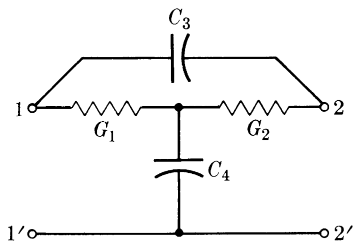


FIG. 8–9. Minimum-phase structure.

transmission of the bridged-T network are at

$$s = -1, \qquad \tfrac{1}{2}(1 \pm j\sqrt{11}),$$

and the zeros of transmission of the lattice are at

$$s = 1, \qquad -\tfrac{1}{2} \pm j\tfrac{3}{2}.$$

Thus, both of them have zeros of transmission in the right half-plane.

It is, however, quite possible for the network to have more than one path of transmission and still be a minimum-phase structure. An example of such a network is shown in Fig. 8–9. Multiplicity of transmission paths is a necessary condition for the network to be nonminimum-phase, but is not a sufficient condition.

In communication networks, where the transmission networks are operated in conjunction with active elements or coaxial cables or both, one finds it useful to require that the terminals 1′ and 2′ of the network of Fig. 8–6 be the same terminal. Such networks are variously known as *common-ground, unbalanced, common-terminal, 3-terminal,* etc., networks. They are conventionally shown as in Fig. 8–10(a) or (b). A special subclass of these networks consists of those containing only two types of elements (LC, RC, or RL) and not containing any transformers. We consider both the general common-terminal structure and the two-element types next. Let us return to the original convention of considering each R, L, and C to be an element.

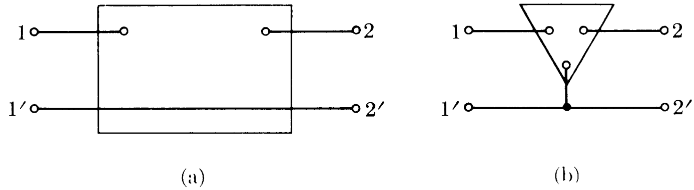


FIG. 8–10. Common-terminal network.

THEOREM 8–11 (Fialkow-Gerst). The coefficients of the numerator (and denominator) polynomials of z_{12} of a common-terminal network are all nonnegative real numbers if common factors in Δ and Δ_{12} are not cancelled.

Proof. From Theorem 7–8,

$$z_{12} = \frac{W_{12,1'2'}(Y) - W_{12',1'2}(Y)}{V(Y)}. \tag{8–27}$$

Since $1'$ and $2'$ are the same vertex, there can be no 2-tree $T_{2_{12',1'2}}$ in which $1'$ and $2'$ are in different connected parts. Therefore

$$W_{12',1'2} = 0 \tag{8–28}$$

and

$$z_{12} = \frac{W_{12,1'2'}(Y)}{V(Y)}, \tag{8–29}$$

which makes the theorem obvious.

The theorem can be restated for y_{12}, since

$$y_{12} = -\frac{W_{12',1'2}(Y) - W_{12,1'2'}(Y)}{\sum U(Y)}, \tag{8–30a}$$

which becomes

$$y_{12} = \frac{W_{12,1'2'}(Y)}{\sum U(Y)} \tag{8–30b}$$

for common-terminal structures. From this, we have the next two corollaries.

COROLLARY 8–11(a). If in a common-terminal network without transformers, common factors of Δ_{12} and Δ_{1122} are not cancelled and y_{12} is written with the leading coefficient in the denominator positive, then all the coefficients in the denominator are nonnegative real numbers, and all the coefficients in the numerator are nonpositive real numbers.

COROLLARY 8–11(b). No common-terminal transformerless network can have a zero of transmission on the positive real axis.

Corollary 8–11(b) is obvious since the sum of positive numbers cannot be zero. (The points ∞ and 0 are not considered here.)

Since there can be no zero on the positive real axis, and zeros in the right half-s-plane are possible (Fig. 8–8a), an interesting question to ask is, how close to the positive real axis can a zero be? To answer this question, we need the following result, due to H. Poincaré.

LEMMA 8-12. Let $P(s)$ be a polynomial of degree $n > 0$, with non-negative real coefficients. Then $P(s)$ has no zeros in the sector

$$|\arg s| < \frac{\pi}{n}. \tag{8-31a}$$

Further, if $P(s)$ has a zero s_0 with

$$|\arg s_0| = \frac{\pi}{n}, \tag{8-31b}$$

then

$$P(s) = a_n s^n + a_0. \tag{8-31c}$$

Proof. Let

$$P(s) = a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0, \tag{8-32}$$

with $a_k \geq 0$, $k = 0, 1, \ldots, n - 1$, and $a_n > 0$. Let

$$s_0 = r\epsilon^{j\theta}$$

be a zero of $P(s)$. Then, equating real and imaginary parts of $P(s_0)$ to zero, we have that

$$a_n r^n \cos n\theta + a_{n-1} r^{n-1} \cos (n - 1)\theta + \cdots + a_1 r \cos \theta + a_0 = 0 \tag{8-33a}$$

and

$$a_n r^n \sin n\theta + a_{n-1} r^{n-1} \sin (n - 1)\theta + \cdots + a_1 r \sin \theta = 0. \tag{8-33b}$$

Now if $|\theta| < \pi/n$, then all the terms of the second sum have the same sign. Further,

$$\begin{array}{lll} \sin k\theta > 0 & \text{if } \theta > 0 & \\ & & \text{for } 1 \leq k \leq n. \\ \sin k\theta < 0 & \text{if } \theta < 0 & \end{array}$$

Thus the second sum could not be equal to zero, so

$$|\arg s_0| \geq \frac{\pi}{n}.$$

If

$$|\arg s_0| = \frac{\pi}{n},$$

then

$$\sin n\theta = 0.$$

But if

$$|\arg s_0| = \frac{\pi}{n}$$

then

$$\sin k\theta \gtrless 0 \qquad \text{for } 1 \leq k \leq n - 1,$$

according as $\theta \gtrless 0$. Hence for $\text{Im}\, P(s_0) = 0$, we must have

$$a_k = 0 \qquad \text{for } 1 \leq k \leq n - 1 \tag{8–34a}$$

so that

$$P(s) = a_n s^n + a_0, \tag{8–34b}$$

as was to be proved.

Now we can answer the question of how close to the positive real axis the zero can be.

THEOREM 8–12. Let

$$z_{12} = \frac{s^k(a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0)}{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0} \tag{8–35}$$

for a common-terminal transformerless two terminal-pair network, where $a_n \neq 0$ and $a_0 \neq 0$. Then the zeros of z_{12} are all in the region

$$|\arg s| \geq \frac{\pi}{n}$$

(except for any zeros at $s = 0, \infty$).

Theorem 8–12 follows immediately from (Poincaré's) Lemma 8–12 and needs no comment. However, let us examine its application to two terminal-pair networks.

To apply this theorem, we must find n. Since each R, L, and C is considered as a network element, the denominator of z_{12} at most changes s^k of Theorem 8–12. To find n, we must find the 2-trees of the network, of the type $T_{2_{12,1'2'}}$; however, it is not necessary to list them. It is necessary only to find the highest and lowest powers of s that occur in $W_{12,1'2'}(Y)$. The highest power of s is contributed by the 2-tree which has the most capacitors and fewest inductors. The lowest power of s results from the most inductors and fewest capacitors. Let us consider a few examples to clarify this point. (It is certainly possible, with the use of topological formulas, to write the numerator of z_{12}, but the point of the present discussion is to find its degree n without actually writing the numerator.)

In Fig. 8–11, there are two capacitors, C_1 and C_2, which constitute a 2-tree by themselves. Hence the highest power of s obtainable is 2. The lowest power here is 0, since there are no inductors in the network. The power 0 is realized by the 2-tree of resistors G_3 and G_4. Thus $n = 2$, so the zeros of transmission are restricted to

$$|\arg s| \geq \frac{\pi}{2},$$

that is, the left half-plane, making Fig. 8–11 a minimum-phase structure.

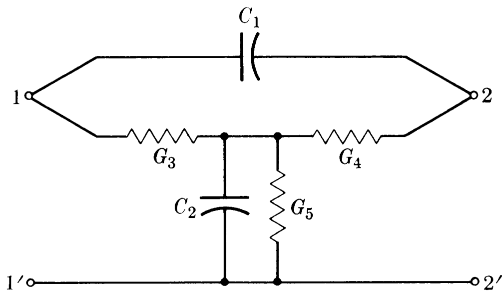


FIG. 8-11. Example for zeros of transmission.

Figure 8-12 is the familiar twin-T structure with some unfamiliar components. The highest power of s obtainable is 2, realized by the 2-tree (C_1, C_2, R_5). The lowest power is -2, realized by the 2-tree (L_3, L_4, R_6). Thus $n = 4$, and the zeros of transmission are the region

$$|\arg s| \geq \frac{\pi}{4} \cdot$$

This network *can* be a nonminimum-phase network. Whether it is a nonminimum-phase network depends on component values and should be investigated.

In Fig. 8-13, the highest power of s realizable is 3 and is actually realized by the 2-tree (C_2, C_4, C_6). The lowest possible power is -2, but there is no 2-tree containing L_7, L_3, and no capacitors. In fact, there is no 2-tree containing L_3 and L_7, since vertices 1 and 2 cannot be connected without going through one of the vertices 3 and 4. Thus the lowest power realized is only -1, realized by the 2-trees (R_1, L_3, R_5) and (R_1, R_5, L_7).

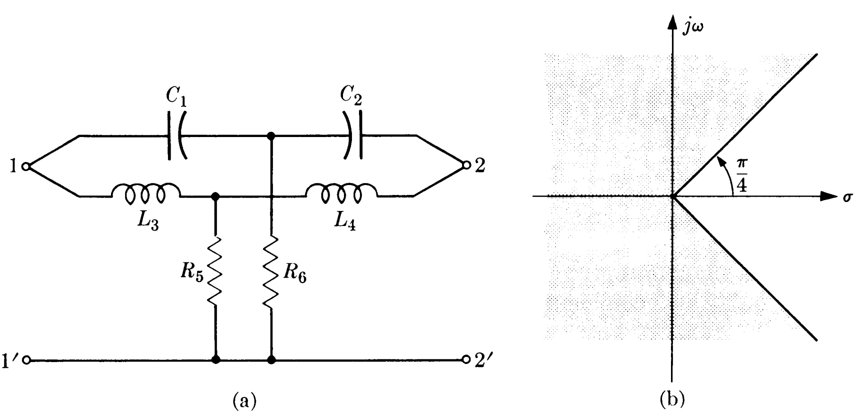


FIG. 8-12. Second example for zeros.

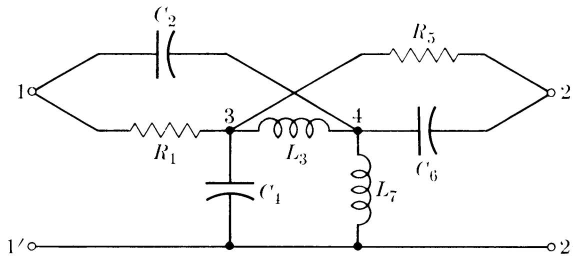


FIG. 8–13. A third example for zeros.

Hence $n = 4$, and the zeros of transmission are in the region

$$|\arg s| \geq \frac{\pi}{4}.$$

This last example brings up the question of "minimal" structures. Insofar as the location of the zero of transmission is concerned, L_7 appears superfluous. If L_7 is replaced by a resistor, the degree of the numerator of z_{12} would not be altered. The zeros of transmission would move, certainly, and function z_{12} would change. This question of minimality is very complicated and one that has not yet been satisfactorily answered.

The next result is a fundamental theorem due to Fialkow and Gerst [54].

THEOREM 8–13 (Fialkow-Gerst). Let Y_n be the node-admittance matrix of a common-terminal network with no transformers, with 1′ as the reference node. Let

$$\Delta_{11} = [\text{cofactor } (1, 1) \text{ of } \mathsf{Y}_n] = \frac{1}{s^{v-2}}[a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0]$$

and

$$\Delta_{12} = [\text{cofactor } (1, 2) \text{ of } \mathsf{Y}_n] = \frac{1}{s^{v-2}}[b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0]. \tag{8–36a}$$

(8–36b)

Then

$$0 \leq b_k \leq a_k \quad \text{for} \quad 0 \leq k \leq n. \tag{8–36c}$$

Thus, if no common factors are cancelled, and the network is common-terminal, the coefficients of the numerator polynomial of

$$\mu_{21} = \frac{z_{21}}{z_{11}}$$

are bounded by the corresponding coefficients of the denominator.

Proof. Since the network contains no magnetic coupling, topological formulas apply. Thus,

$$\Delta_{12} = W_{12,1'}(Y) \quad \text{and} \quad \Delta_{11} = W_{1,1'}(Y). \tag{8-37}$$

From the 2-tree identities of Section 7-2,

$$W_{1,1'}(Y) = W_{12,1'}(Y) + W_{1,1'2}(Y). \tag{8-38}$$

Thus every 2-tree admittance product that appears in Δ_{12} also appears in Δ_{11}:

$$\Delta_{11} = \Delta_{12} + W_{1,1'2}(Y). \tag{8-39}$$

Since the coefficients in $W_{1,1'2}(Y)$ are necessarily nonnegative,

$$0 \leq b_k \leq a_k.$$

The second half of the theorem is merely a restatement of the same result for common-terminal networks. μ_{21} has the interpretation "voltage-ratio transfer function."

We extend this theorem to general two terminal-pairs later.

THEOREM 8-14 (Fialkow-Gerst). A necessary condition for the realizability of a voltage-ratio transfer function $\mu_{21}(s)$ as a common-terminal network without transformers is

$$0 \leq \mu_{21}(\sigma) \leq 1 \qquad \text{for } 0 \leq \sigma \leq \infty, \tag{8-40}$$

where the equalities can hold only at the extremities ($\sigma = 0, \infty$) of the range unless $\mu_{21}(s) \equiv 0$ or $\mu_{21}(s) \equiv 1$.

Theorem 8-14 is actually just a restatement of Theorem 8-13. However, it is a fundamental result of two terminal-pair theory. Fialkow and Gerst [54] have actually shown that this condition, together with the stability requirement that $\mu_{21}(s)$ has no poles in the right half-plane, is sufficient for realizability as a common-terminal network. This, however, is network synthesis proper and so is not discussed here.

Theorem 8-13 is sometimes stated in the following form.

Let the numerator polynomials of z_{11} and z_{12} be respectively

$$p(s) = a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0$$

and $$q(s) = b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0. \tag{8-41}$$

Then for a common-terminal network without transformers,

$$0 \leq b_k \leq a_k \qquad \text{for } 0 \leq k \leq n.$$

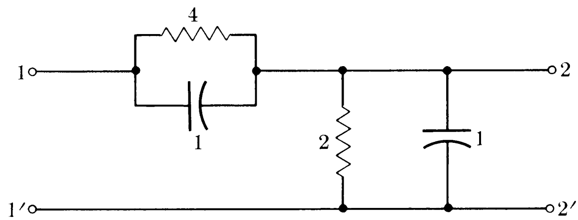


FIG. 8–14. Counterexample.

In this form, the theorem is not true if any common factors have been cancelled in z_{12} or z_{11}. A counterexample to such a statement is given in Fig. 8–14, where

$$z_{11} = \frac{16s + 6}{8s^2 + 6s + 1} \qquad \text{and} \qquad z_{12} = \frac{8}{8s + 4}.$$

As can be observed from the proofs of Theorems 8–13 and 8–14, Δ_{11} can be replaced by Δ_{22} and $\mu_{21}(s)$ by $\mu_{12}(s)$ and these theorems still remain true.

Let us conclude the discussion of common-terminal networks with a brief discussion of an interesting application of the intuition that arises from topological formulas. This application is the *generation* of two-element kinds of networks that have specified locations of zeros of transmission in the complex plane. As the model, we take the most important RC-networks. It is sufficient to consider one of the three combinations, for we can apply the results to any other combination by the following transformations, due to W. Cauer.

An RC-network with impedances $Z_j(s)$ is converted to an LC-network with impedances $\zeta_j(\lambda)$ by defining

$$\zeta_j(\lambda) = \lambda Z_j(\lambda^2). \tag{8–42}$$

An RC-network with impedances $Z_j(s)$ is converted to an RL-network with impedances $\eta_j(s)$ by defining

$$\eta_j(s) = Z_j\left(\frac{1}{s}\right). \tag{8–43}$$

Finally, an RL-network with impedances $Z_j(s)$ is converted to an LC-network with impedances $\zeta_j(\lambda)$ by defining

$$\zeta_j(\lambda) = \frac{1}{\lambda} Z_j(\lambda^2). \tag{8–44}$$

With RC-networks, the admittances are G and sC, so that $W_{12,1'2'}(Y)$ is simply a polynomial in s. Thus it is easy to determine the various degrees. As observed earlier, the degree of $W_{12,1'2'}$ (in s) is equal to the number of capacitors, if there is a 2-tree containing all these capacitors. The degree of $W_{12,1'2'}(Y)$ in y_k's is $v - 2$, where v is the number of vertices. With these few facts in mind, we can generate a number of configurations for zeros of transmission in various parts of the plane.

For example, let us try to design a *minimal* network (RC-common-terminal) which realizes any pair of complex zeros of transmission in the left half-s-plane. If the zeros are at

$$s = -\alpha \pm j\beta, \qquad \alpha, \beta \geq 0, \tag{8-45}$$

the numerator of z_{12} can be written as

$$k(s^2 + 2\alpha s + \alpha^2 + \beta^2) = k(s^2 + 2\alpha s + \omega_0^2), \tag{8-46}$$

where

$$\omega_0^2 = \alpha^2 + \beta^2. \tag{8-47}$$

Therefore, we want

$$W_{12,1'2'}(Y(s)) = k(s^2 + 2\alpha s + \omega_0^2). \tag{8-48}$$

Since the degree of the polynomial is 2, at least four vertices are required. Also, at least two capacitors are needed. Since the constant term ω_0^2 is not zero, an all-resistor 2-tree is required. Hence at least two resistors are necessary.

Let us therefore see if the network can be designed with just two resistors and two capacitors (and four vertices), which is certainly minimal.

Since four vertices are needed, one vertex besides the terminals 1, 2, and 1′ is necessary. Let this internal vertex be 3. The possible 2-trees $T_{2_{12,1'2'}}$ in such a vertex configuration are listed in Fig. 8-15.

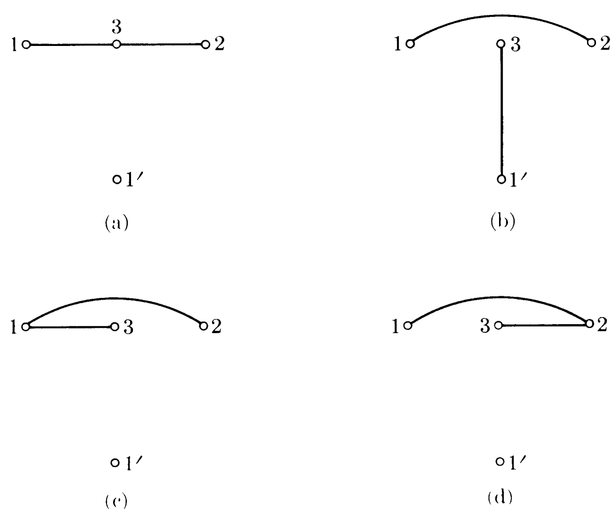


FIG. 8-15. 2-trees with four vertices.

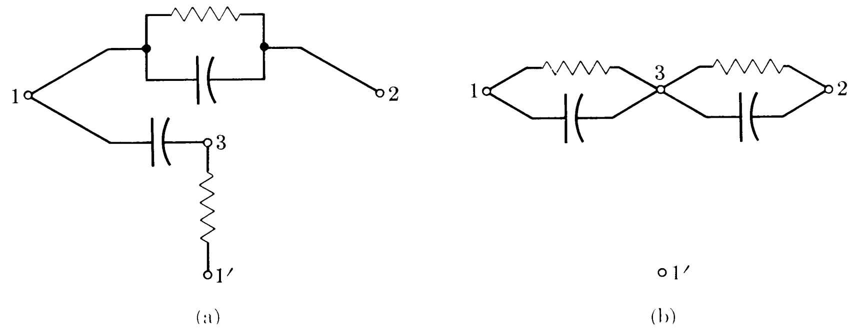


FIG. 8–16. Discarded networks.

All that remains is to "fit" two resistors and two capacitors into these patterns. Several of the possible distributions can be discarded immediately on the basis that they cannot give complex zeros of transmission because the function z_{12} becomes a driving-point function (of the RC-network). Examples are the networks of Fig. 8–16.

In a very short time, it can be seen that there are only two distributions that can possibly give complex zeros of transmission, and these are the bridged-T structures of Fig. 8–17. It has not yet been shown that these networks realize complex zeros of transmission. To show this, compute $W_{12,1'2'}$. For Fig. 8–17(a),

$$W_{12,1'2'}(Y) = s^2C_3C_4 + sC_3(G_1 + G_2) + G_1G_2. \tag{8–49}$$

This expression must equal the given polynomial. Therefore, set them equal and solve for the coefficients:

$$s^2C_3C_4 + sC_3(G_1 + G_2) + G_1G_2 = ks^2 + 2k\alpha s + k\omega_0^2, \tag{8–50}$$

which leads to the simultaneous multilinear equations

$$C_3C_4 = k, \qquad C_3(G_1 + G_2) = 2k\alpha, \qquad G_1G_2 = k\omega_0^2. \tag{8–51}$$

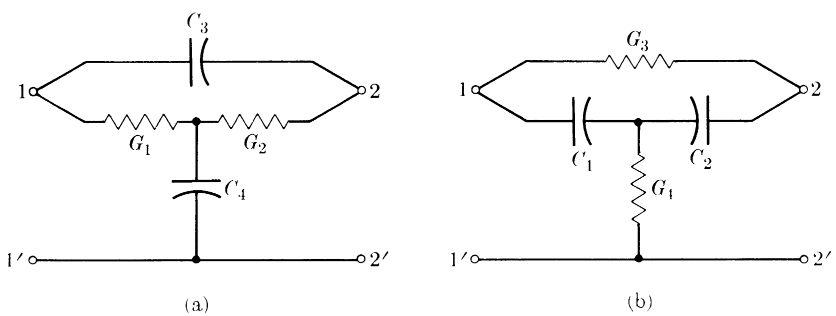


FIG. 8–17. Bridged-T networks.

It is here that the procedure can break down. Because these are multilinear equations, there is no straightforward procedure for solving them. But the problem is worse than that. The solutions (C_3, C_4, G_1, and G_2) must be *real positive numbers.* Not even existence theorems are known for such a problem.

In this particular case, however, the equations are solvable and do have real positive solutions. The solutions are

$$\begin{aligned}
C_3 &= \frac{k}{C_4}, \\
C_4 &= \text{(real positive but arbitrary)}, \\
G_1 &= \left(\frac{1}{4k^2\omega_0^4} + 2C_4\alpha\right)^{1/2} - \frac{1}{2k\omega_0^2}, \\
G_2 &= \frac{k\omega_0^2}{G_1}.
\end{aligned} \tag{8-52}$$

Thus the network realizes any pair of complex zeros in the left half-plane.

For a second example, let us try something slightly more complicated. Let us try to generate a minimal configuration that realizes a pair of zeros in the right half-plane. The sector in the right half-plane must be chosen before the number of components required is known. Suppose that the region of interest is the shaded region of Fig. 8-18.

By (Poincaré's) Lemma 8-12, the numerator has to be a cubic. Hence three capacitors are required, and these must constitute a 2-tree $T_{2_{12,1'2'}}$. Also, there must be a 2-tree of resistors; otherwise an s will factor out of the numerator of z_{12}, leaving a quadratic with positive coefficients, which has zeros in the left half-plane. Thus *at least* three capacitors, three resistors, and five vertices are necessary. Thus at least two internal

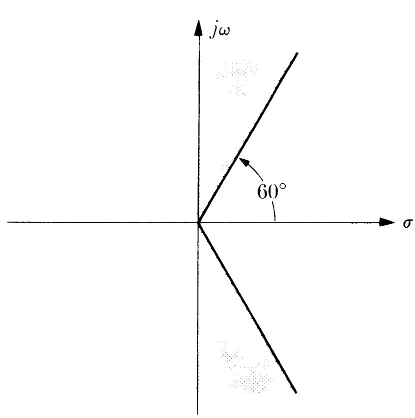


FIG. 8-18. Region in the right half-plane.

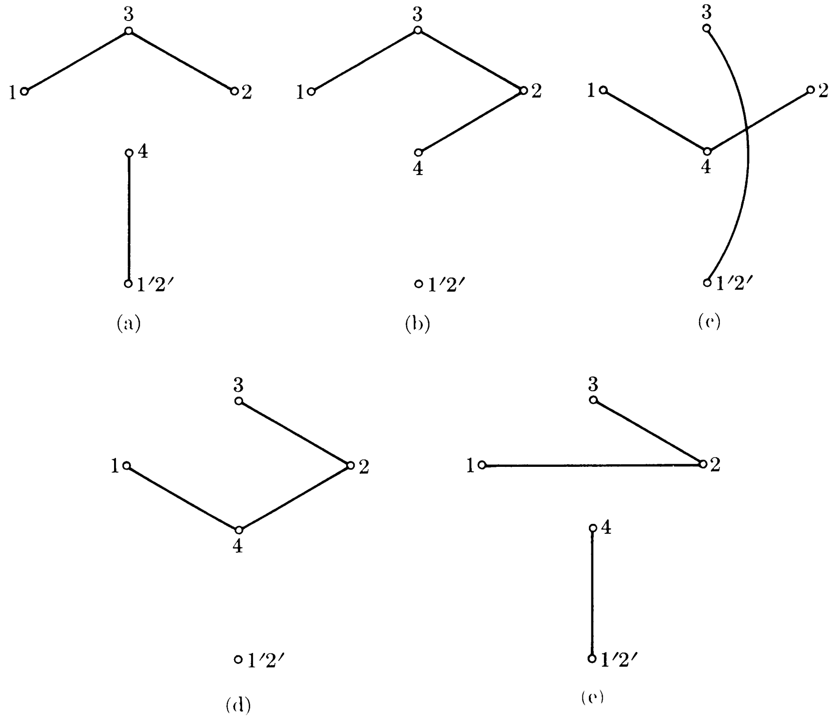


FIG. 8–19. 2-trees with five vertices.

vertices have to be added, exactly two if the minimum number of three capacitors and three resistors is used. The next problem is to fit these together. Let us look at a few 2-trees that are possible, as in Fig. 8–19. Other 2-trees are also possible, and the reader can easily construct all the possible 2-trees, starting from the ones in Fig. 8–19.

Many choices exist for assigning R's and C's to the 2-trees of Fig. 8–19 and those that can be derived from Fig. 8–19. This should be considered as encouraging rather than as discouraging. Let us look at one of the possibilities, which leads to a familiar network. This network is obtained by taking (a) and (c) of Fig. 8–19 and making one of the two 2-trees a capacitor 2-tree and the other a resistor 2-tree. Then the twin-T structure of Fig. 8–20 results.

To determine whether this twin-T will satisfy the requirements, we must follow computations similar to those in the preceding example. Only now they are more complicated, since the multilinear equations are of degree 3 instead of 2. However, the computations can be performed to show that any two complex zeros in the shaded region of Fig. 8–18 can be realized.

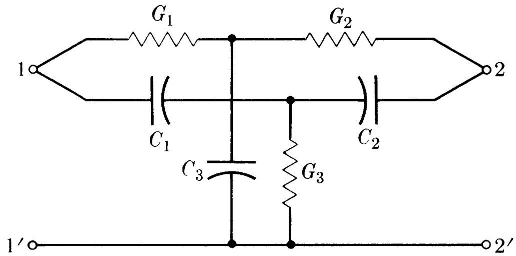


FIG. 8-20. Twin-T network.

Since a cubic is involved, it has another zero on the negative real axis. Let the zeros be at $\alpha \pm j\beta$ and $-\sigma_1$. Positiveness of the coefficients of $W_{12,1'2'}$ requires that

$$\sigma_1 > 2\alpha \quad \text{and} \quad \alpha^2 + \beta^2 > 2\alpha\sigma_1; \quad \text{that is,} \quad \frac{\alpha^2 + \beta^2}{2\alpha} > \sigma_1 > 2\alpha. \tag{8-53}$$

If, in addition, we make σ_1 satisfy

$$2(-\alpha^2 + \beta^2) < \sigma_1^2, \tag{8-54}$$

the following component values give the desired zeros:

$$\begin{aligned} C_1 &= \text{(arbitrary)}, \\ C_2 &= \frac{1}{C_1}, \\ C_3 &= 1, \\ G_3 &= \frac{\alpha^2 + \beta^2}{G_1 G_2}\,\sigma_1. \end{aligned} \tag{8-55}$$

G_1 and G_2 are the two (necessarily real positive) solutions of

$$G^2 - (\sigma_1 - 2\alpha)G + \frac{C_1}{C_1^2 + 1}(\alpha^2 + \beta^2 - 2\alpha\sigma_1) = 0. \tag{8-56}$$

Other similar examples are to be found in Hakimi and Seshu [70]. The important idea that we wish to convey here is that we now have a means of generating canonical configurations of two terminal-pair sections.

Leaving the case of the common-terminal structure, we prove the Fialkow-Gerst theorems for the general two terminal-pair network.

THEOREM 8-15 (Fialkow-Gerst). Let the node equations of the two terminal-pair network (Fig. 8-6) be written with vertex 1′ as the reference vertex. Further, let there be no magnetic coupling in the network. With Δ_{ij} representing the cofactor of the (i, j)-element of the node-admittance matrix, let

$$\begin{aligned}\Delta_{11} &= \frac{1}{s^{v-2}}\,[a_n s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0],\\ \Delta_{12} &= \frac{1}{s^{v-2}}\,[b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0],\\ \Delta_{12'} &= \frac{1}{s^{v-2}}\,[c_n s^n + c_{n-1}s^{n-1} + \cdots + c_1 s + c_0].\end{aligned} \tag{8-57}$$

Then

$$0 \le b_k \le a_k \tag{8-58a}$$

and

$$0 \le c_k \le a_k, \qquad 0 \le k \le n, \tag{8-58b}$$

so that

$$|b_k - c_k| \le a_k. \tag{8-58c}$$

Thus if no common factors are cancelled, the absolute values of the coefficients of the numerator polynomial of

$$\mu_{21} = \frac{z_{21}}{z_{11}} \tag{8-59}$$

are bounded by the corresponding coefficients of the denominator.

There is really no need to write a new proof, as the proof of Theorem 8-13 establishes the inequalities

$$0 \le b_k \le a_k \qquad \text{and} \qquad 0 \le c_k \le a_k. \tag{8-60}$$

Subtracting one from the other (both ways), we get

$$|b_k - c_k| \le a_k. \tag{8-61}$$

As before, the statement about coefficients of the numerator and denominator in the voltage-ratio transfer function is merely a restatement of the first part.

The second theorem can also be generalized, as follows.

THEOREM 8-16 (Fialkow-Gerst). A necessary condition for the realizability of a voltage-ratio transfer function $\mu_{21}(s)$ as a two terminal-pair network without transformers is

$$|\mu_{21}(\sigma)| \le 1 \qquad \text{for } 0 \le \sigma \le \infty, \tag{8-62}$$

where the equality can hold only at the extremities of the range unless it holds identically.

Theorems 8-14 and 8-16 are to be interpreted as the bounds on the "gain" or constant multiplier of z_{12} (or y_{12}) that can be realized. Once again Fialkow and Gerst [55] have also established the sufficiency of this condition (together with stability).

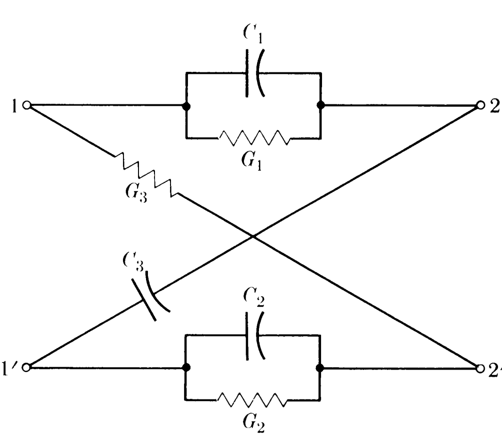


FIG. 8-21. Network realizing zeros.

Theorem 8-11 and its corollaries do not extend to general two terminal-pairs and we cannot make any general statement about the locations of the zeros of transmission. They may be located anywhere in the plane. Let us illustrate this last remark by generating a network that realizes the pair of zeros of transmission:

$$s = -\alpha \pm j\beta, \qquad \text{where} \qquad \alpha \gtreqless 0. \tag{8-63}$$

We evidently require that

$$W_{12,1'2'} - W_{12',1'2} = s^2 + 2\alpha s + (\alpha^2 + \beta^2). \tag{8-64}$$

Since α may be positive or negative,

$$W_{12',1'2} \neq 0. \tag{8-65}$$

Let us try an RC realization. We need a two-capacitor 2-tree and an all-resistor 2-tree, both in $W_{12,1'2'}$. We need one 2-tree containing one capacitor, of the type $T_{2_{12',1'2}}$. Putting these ideas together, we get, as the simplest possible structure, the network of Fig. 8-21.

For the network of Fig. 8-21,

$$W_{12,1'2'} - W_{12',1'2} = s^2C_1C_2 + (G_1C_2 + C_2G_1 - G_3C_3)s + G_1G_2. \tag{8-66}$$

We see that a possible solution of this problem is

$$\begin{aligned} C_1 = C_2 = 1, &\qquad G_1 = G_2 = (\alpha^2 + \beta^2)^{1/2}, \\ C_3 = 1, &\qquad G_3 = 2(\alpha^2 + \beta^2)^{1/2} - 2\alpha, \end{aligned} \tag{8-67}$$

and all values are nonnegative, since

$$(\alpha^2 + \beta^2)^{1/2} \geq \alpha. \tag{8-68}$$

Problems

8–1. Prove Corollaries 8–1(a) and 8–1(b).

8–2. Give a formal proof for Theorem 8–4, based on topological formulas.

8–3. Guillemin's algorithm for evaluating the order of complexity N is as follows. Order the circuits in the network as $1, 2, \ldots, \mu$, and examine them in this order. If loop 1 contains an inductor and a capacitor, the weight of the loop is 2, and one inductor and one capacitor of the loop are "assigned" to loop 1, by marking "1" next to them. If the loop contains inductors and no capacitors or vice versa, its weight is 1, and a reactive element is assigned to loop 1. If it contains only resistors, the loop has weight 0. Next examine loop 2. At each stage, only unassigned reactive elements can be counted. After all loops are examined, the weights are added; the sum is N. Show that this algorithm is equivalent to counting $\mu_{Cs} + \mu_{Ls}$.

8–4. Prove that if a planar one terminal-pair network without transformers is minimal in reactive elements, so is its dual. (Reza [145].)

8–5. Investigate the relation between the poles and the zeros of $z_{21}(s)$ for the bridged-T networks of Fig. 8–17. How can we modify this network to be able to specify the pole independently of the zeros? (Dasher [42].)

8–6. Where are the poles of $z_{21}(s)$ for the twin-T network of Fig. 8–20? Is it possible for the real zero of $z_{21}(s)$ of this network to cancel with a pole?

8–7. Design an LC common-terminal network that will realize any pair of complex zeros in the region $\pi/6 < |\arg s| \leq \pi/2$.

8–8. Design another RC minimal network to realize any pair of complex zeros in the region $\pi/3 < |\arg s| \leq \pi/2$.

8–9. Let the open-circuit voltage-ratio transfer function be

$$\mu_{21}(s) = K \frac{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}{s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0} = K \frac{N(s)}{D(s)}.$$

Show that the largest "gain" K that is realizable is such that $0 \leq K < K_0$, where $K_0 = \min K$ such that $D(s) - K\,N(s)$ has a real positive zero. (Fialkow and Gerst [54].)

8–10. Can a complex zero of transmission be realized by an RC ladder network? Why?

8–11. A network with $|\mu_{21}(j\omega)| \equiv 1$ is called an "all-pass" network. Show that a nontrivial all-pass network [i.e., one for which $\mu_{21}(s) \not\equiv 1$] is necessarily nonminimum-phase. Design the simplest all-pass networks with (a) one pole and one zero, and (b) two poles and two zeros.

8–12. Prove Corollary 8–8.

8–13. Prove that an RC common-terminal network with only two capacitors (Fig. 8–11, for example) cannot have a zero of transmission on the imaginary axis, so that the zeros are in the region $|\arg s| > \pi/2$ (thus strengthening the argument about Fig. 8–11). [*Hint:* Steinitz replacement theorem.]

8–14. Extend Problem 8–13 to show that a common-terminal network with n reactive elements has no zeros on the lines $|\arg s| = \pi/n$.

CHAPTER 9

APPLICATIONS TO THE THEORY OF SWITCHING

Like many other chapters in this text, this chapter is based on a few published papers and is a collection of the known applications of graph theory to the theory of switching. In the application to contact networks (Section 9–1), the emphasis is on the relationships between electrical networks and contact networks and on the known minimality proofs in contact network theory. The discussion of the relationship between conventional electrical networks and contact networks is based on the work of Belevitch [8], Seshu [153], and Mayeda [112]. The minimality proofs discussed are due to Cardot [22] and Shannon and Gould [66]. The application of directed graphs to the mathematical model (state diagram) of a sequential machine is the topic of Section 9–2. Although a great deal of work has been done on state diagrams, only a relatively small part of it is a direct application of graph theory. Here the connection matrix of Hohn, et al., [77] is made the basis of the discussion, primarily because it is related very closely to the flow graphs to be considered in Chapter 10. The transition matrices of Seshu, Miller, and Metze [155] are also introduced primarily because of their relationship to the relation matrices and structure matrices mentioned in Chapter 10. The application of directed graphs to logic networks has not been developed completely, and a brief outline of the work of Shelly [163] is given in Section 9–3. Familiarity with elementary theory of switching is assumed in this chapter. (See, for instance, Caldwell [21].)

9–1 Contact networks. The notation of this section is based on the so-called *admittance representation* of a contact network. Thus 1 stands for a short circuit and 0 for open circuit. The symbol $+$ stands for Boolean addition (union) and $\cdot$ stands for Boolean multiplication (intersection) with the usual convention that xy stands for $x \cdot y$. Complement of x is denoted x'. (The transpose of a Boolean matrix P is denoted P^T.) A contact network and its switching functions may be defined from the viewpoint of graph theory as follows.

DEFINITION 9–1. *Contact network.* A *contact network* is a nonoriented graph with a Boolean variable x_i (or x_i') associated with each edge.

DEFINITION 9–2. *Path product.* A *path product* π_{ij} is the product of the variables associated with the edges of a path from vertex i to vertex j of the contact network.

DEFINITION 9–3. *Switching function.* The *switching function* f_{ij} between vertices i and j of a contact network is

$$f_{ij} = \sum_k \pi_{ij}^{(k)}, \tag{9–1}$$

where the summation is Boolean addition and extends over all the paths from vertex i to vertex j.

To show the relationship between contact networks and conventional networks, let us first define the primitive connection matrix of Hohn and Schissler [76].

DEFINITION 9–4. *Primitive connection matrix.* The *primitive connection matrix* $\mathsf{P} = [p_{ij}]$ of a contact network is of order (v, v), where v is the number of vertices in the network and

$$p_{ij} = \sum_k w_{ij}^k, \qquad i \neq j, \tag{9–2}$$

where w_{ij} is the Boolean variable associated with the edge between vertices i and j and the summation is over all such edges. If there is no such edge, $p_{ij} = 0$; further,

$$p_{ii} = 1 \qquad \text{for all } i. \tag{9–3}$$

It is evident that the primitive connection matrix is closely related to the node-admittance matrix of conventional network theory. This relationship may be stated precisely (see Belevitch [8] and Seshu [153], as in Theorem 9–1.

THEOREM 9–1. Let $y_1, y_2, \ldots, y_e$ be the Boolean variables associated with the edges of the contact network, and let

$$\mathsf{Y} = \begin{bmatrix} y_1 & 0 & 0 & \cdots & 0 \\ 0 & y_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & y_e \end{bmatrix}. \tag{9–4}$$

Then the primitive connection matrix is given by

$$\mathsf{U} + \mathsf{A}_a \mathsf{Y} \mathsf{A}_a^T = \mathsf{P}, \tag{9–5}$$

where A_a is the incidence matrix (nonoriented), the superscript T denotes transpose, and U is the unit matrix of order v.

This result may be considered to be obvious in the light of experience with conventional networks. The formal proof is left as a problem (Problem 9–1).

The switching function f_{ij} between vertices i and j of a contact network is the analogue of the driving-point admittance between vertices i and j of an electrical network. To exhibit this relationship, some restrictions are necessary. In the case of the electrical network, the restriction is to networks with no magnetic coupling. In the case of the contact network, a stronger restriction is required, since the same variable x_i may appear in two or more different places in a general contact network, whereas it is impossible in an electrical network (even if two components have the same admittance, we consider them to be different). Therefore, attention is restricted to the special case of contact networks in which each Boolean variable appears only once (either primed or unprimed).

DEFINITION 9-5. *Single-contact (SC-) function.* A *single-contact (SC-)* network is a contact network in which each edge has a different Boolean variable associated with it. The switching function of such a network (between any two terminals) is an *SC-function.*

SC-functions are also referred to as *graph functions, network functions,* and *noniterated functions.* In case the one terminal-pair network under consideration is series-parallel (with respect to the terminal vertices), the relationship between the switching function and the driving-point admittance is direct. To establish the formal relationship, we define the *star product.*

DEFINITION 9-6. *Star product.* The *star product* $y_1 * y_2$ of two admittances y_1 and y_2 is defined by

$$y_1 * y_2 = \frac{y_1 y_2}{y_1 + y_2}\,; \tag{9-6}$$

$y_1 * y_2$ is thus the admittance of a series combination of y_1 and y_2.

It may be verified that the star product is commutative and associative, but is not distributive over addition. Next, let us state the formal relationship between series-parallel conventional networks and *SC*-networks.

THEOREM 9-2. Let 1 and 1′ be the terminal vertices of a series-parallel network. If the network is taken as a conventional network without mutual inductances, the driving-point admittance $Y_{11'}$ can be expressed by using only the two operations $*$ and $+$, with each y_j appearing only once in the expression, as

$$Y_{11'} = y_{i_1} * [y_{i_2} + y_{i_3} * (\cdots)]. \tag{9-7}$$

If the network is taken as a single-contact switching network, the switching function $F_{11'}$ can be expressed by using Boolean multiplica-

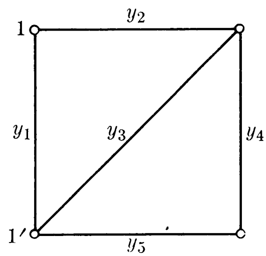


FIG. 9–1. Example for Theorem 9–2.

tion and addition only and with each variable y_j appearing only once in the expression, as

$$F_{11'} = y_{i_1}[y_{i_2} + y_{i_3}(\cdots)]. \tag{9-8}$$

The two expressions for $Y_{11'}$ and $F_{11'}$ are identical except that Boolean multiplication replaces star multiplication.

The theorem is evident, so it suffices to give an example to illustrate it. For the network of Fig. 9–1,

$$F_{11'} = y_1 + y_2(y_3 + y_4y_5)$$

and (9–9)

$$Y_{11'} = y_1 + y_2 * (y_3 + y_4 * y_5).$$

In the case of a non-series-parallel network, it is not possible to express the switching function or the driving-point admittance in such a way that each variable appears only once in the expression (using the operations $*$ and $+$ for $Y_{11'}$ and the operations $\cdot$ and $+$ for $F_{11'}$). That the star product is not distributive over addition now prevents us from correlating the two. It is, however, possible to construct $Y_{11'}$ from $F_{11'}$ and vice versa.

The driving-point admittance $Y_{11'}$ is expressible in terms of the trees and 2-trees of the network as (Eq. 7–34)

$$Y_{11'} = \frac{V(Y)}{W_{1,1'}(Y)}, \tag{9-10}$$

where $V(Y) = \sum$ (tree-admittance products, and $W_{1,1'}(Y) = \sum$ (2-tree admittance products), with vertices 1 and 1′ in different connected parts in each 2-tree. On the other hand,

$$F_{11'} = \sum_k \pi_{11'}^{(k)}, \tag{9-11}$$

where $\pi_{11'}$ is the path product of a path from 1 to 1′. From Theorem 2–12, each such path can be made part of a tree of the network; and it obviously cannot be made part of a 2-tree $(1, 1')$. Therefore, it follows that $\pi_{11'}$ is a factor of one (or more) of the tree-products in $V(Y)$ and is not a factor of any product in $W_{1,1'}(Y)$. By Problem 9–2, the converse of this statement is also true. Namely, every product of edge variables that is a factor of a product in $V(Y)$ and is not a factor of any product in $W_{1,1'}(Y)$ corresponds to a subgraph *containing* a path from 1 to 1′. Thus, in the Boolean sense, every such product is *contained* in a $\pi_{11'}^{(k)}$. Thus we have our next theorem.

THEOREM 9–3. Let the driving-point admittance between vertices 1 and 1′ be expressed as

$$Y_{11'} = \frac{V(Y)}{W_{1,1'}(Y)}. \tag{9–12}$$

Then, if this network is interpreted as an *SC*-network,

$$F_{11'} = \sum \begin{bmatrix} \text{factors of products in } V(Y) \text{ which are not} \\ \text{factors of products in } W_{1,1'}(Y) \end{bmatrix},$$

where the sum is Boolean.

In fact, we may interrelate F, W, and V in a number of ways, all of which are a consequence of the next theorem.

THEOREM 9–4. If $F_{11'}$, V, and $W_{1,1'}$ are defined as above,

$$F_{11'}\, W_{1,1'} = V \qquad \text{(Boolean equation)}. \tag{9–13}$$

The proof is simple but interesting and so is left as a problem (Problem 9–3).

An unsolved problem in this connection is to find $F + W$. Another is to find an explicit formula for F in terms of V and W. The two are, of course, related. $F_{11'}$, by itself, contains all the information about the *SC*-network. Precisely, $F_{11'}$ determines an *SC*-network to within a 2-isomorphism. This was first demonstrated by Ashenhurst [4]. His important theorem is considered next.

THEOREM 9–5 (Ashenhurst). Let G be a nonseparable graph and e_1 any edge of G with vertices 1 and 1′. Then every circuit C of G which does not contain e_1 is the ring sum of some two paths P_1 and P_2 between vertices 1 and 1′.

Proof. Let C be a circuit not containing e_1.

Case 1. The edge e_1 has two vertices in common with C. Then C consists of two disjoint paths between vertices 1 and 1′. The ring sum of these two paths is simply C.

Case 2. e_1 has one vertex in common with C. Let this vertex be 1. There is one other vertex v_3 in C which is not in e_1. Since G is nonseparable, there is a path between $1'$ and v_3 not containing 1, and hence not containing e_1. Let v_4 be the last vertex of such a path which is on C. Then there is a path between v_4 and $1'$ not containing any element of C. Let this be $p_1(v_4, 1')$. C itself consists of two paths, $p_2(1, v_4)$ and $p_3(1, v_4)$, between 1 and v_4. Now $p_1(1', v_4)$ together with $p_2(1, v_4)$ is a path between 1 and $1'$. Similarly, $p_1(v_4, 1')$, together with $p_3(1, v_4)$, is a path between 1 and $1'$. The ring sum of these two paths is C.

Case 3. e_i has no vertices in common with C. Let e_2 be some edge of C with vertices v_3 and v_4. Since G is nonseparable, there is a circuit containing e_1 and e_2. In this circuit, there is a path between 1 and one of v_3 and v_4, not containing e_1 or e_2. Suppose that this is the path between 1 and v_3. Then there is a path between $1'$ and v_4, and the two paths are disjoint. Let v_5 be the vertex at which $p_1(1, v_3)$ meets C, and v_6 the vertex at which $p_2(1', v_4)$ meets C. Then there exist distinct paths $p_3(1, v_5)$ and $p_4(1', v_6)$ joining e_1 to C. Let $p_5(v_5, v_6)$ and $p_6(v_5, v_6)$ be the paths which constitute C. Then $p_3p_5p_4$ and $p_3p_6p_4$ are two paths between 1 and $1'$, with C for the ring sum. The proof of the theorem is now complete.

LEMMA 9–6. *Let p_1 and p_2 be two paths between vertices v_1 and v_2. Then $p_1 \oplus p_2$ is an element-disjoint union of circuits.*

Proof. At each vertex of $p_1 \oplus p_2$, there is an even number of elements. Hence $p_1 \oplus p_2$ is an Euler graph, if nonempty.

THEOREM 9–6 (Ashenhurst). *The realization of an SC-function as a nonseparable SC-network is unique to within a 2-isomorphism. (The network is nonseparable in the one terminal-pair sense.)*

Proof. Let N_1 and N_2 be two networks realizing an SC-function F. Let 1 and $1'$ be the input vertices. Let an element e_0 be added to both networks, between vertices 1 and $1'$. The networks still have the same switching function. By Theorem 9–5 and Lemma 9–6, the circuits of both N_1 and N_2 are derivable from the function F. Thus they are identical. Hence N_1 and N_2 are 2-isomorphic. This completes the proof.

There is an elegant way of deriving the set of all circuits from F. Let each path and each circuit be represented as the product of the weights of elements of the path or circuit. Let the ring sum $p_i \oplus p_j$ be written also as a product. Then the set of circuits of the network is given by

$$e_0F + \sum p_i \oplus p_j, \tag{9–14}$$

where $\sum$ denotes Boolean addition.

Using Theorem 9–5, Trakhtenbrot [175], Okada [124], and Seshu [153] devised simple synthesis procedures for *SC*-functions, which have since been extended by Gould [67] to non-*SC*-functions. Let us briefly outline this procedure since it can be used as a method of establishing minimality of contact realizations. Let the *SC*-function be expressed as a sum of products as

$$F(y_1, y_2, \ldots, y_e) = \sum_j p_j(y_1, y_2, \ldots, y_e). \tag{9–15}$$

Add a "driver element" y_0 between the input vertices, and construct $y_0F = \sum y_0 p_j$. Then each $y_0 p_j$ is a circuit containing y_0. Construct the matrix $\mathbf{B}_F$ of these circuits. Since all the circuits of the graph are expressible as linear combinations of circuits containing y_0, by Ashenhurst's theorem, the mod 2 rank of $\mathbf{B}_F$ is the nullity of G. Hence, delete the superfluous rows to find $\mathbf{B}$. Now find the matrix $\mathbf{Q}$ orthogonal and complementary to $\mathbf{B}$; that is, such that

$$\mathbf{B}\mathbf{Q}^T = \mathbf{0} \qquad (\text{mod } 2)$$

and (9–16)

$$(\text{rank of } \mathbf{B}) + (\text{rank of } \mathbf{Q}) = e + 1.$$

Then $\mathbf{Q}$ is the cut-set matrix of G, from which the incidence matrix is found by elementary row operations.

Since the sum mod 2 of circuits is a circuit or disjoint union of circuits, we can state the following test (a necessary condition) for realizability.

THEOREM 9–7. If F is realizable as an *SC*-network, the sum mod 2 of an odd number of rows of $\mathbf{B}_F$ must correspond to a product contained in F (in the Boolean sense).

Gould's extension to non-*SC*-networks is to consider each contact to be a different variable.

As another consequence of Ashenhurst's theorems, Mayeda's [112] procedure for constructing the driving-point admittance $Y_{11'}$ from the switching function $F_{11'}$ is considered next. $F_{11'}$ lists all the paths between the input vertices. The ring sum of any two paths in $F_{11'}$ is a circuit or edge-disjoint union of circuits. Therefore, if the ring sum of every pair of paths in $F_{11'}$ is formed, and the minimal members chosen from the resulting sets of edges, all the circuits are found. Let $C_1, C_2, \ldots, C_k$ be the circuits that are so obtained. Now the trees of the graph can be found as follows. Find the number of linearly independent paths in $F_{11'}$ (which is the mod 2 rank of the matrix listing the paths in $F_{11'}$). This number is clearly $e - v + 2$, where e is the number of variables in $F_{11'}$ and v is the number of vertices of G; for if a column of 1's is added to the matrix, corresponding to y_0, the matrix becomes $\mathbf{B}_F$ with rank

$e' - v + 1$, where $e' = e + 1$ (since y_0 has been added). Thus v is found. Now the trees of G are sets of $v - 1$ edges such that no set contains any of the circuits $C_1, C_2, \ldots, C_k$ (see Theorem 2–10). Hence $V(Y)$ can be found. The 2-trees $(1, 1')$ are sets of $v - 2$ edges which contain none of the circuits $C_1, C_2, \ldots, C_k$ and none of the paths between vertices 1 and $1'$ (which are all in $F_{11'}$). Thus $W_{1,1'}(Y)$ can be found. Hence the driving-point admittance can be computed from $F_{11'}$ without actually constructing the graph realizing $F_{11'}$. In practice, Mayeda's procedure is not much faster than constructing the realization of $F_{11'}$; however, Mayeda's procedure is suitable for machine computation.

Another concept that can be simply extended from conventional networks to contact networks is that of duality. The following theorem is due to Shannon [159].

THEOREM 9–8. Let G be the graph of a one terminal-pair contact network which remains planar when an edge y_0 is added between the input vertices. Let $G^* + y_0^*$ be the dual of $G + y_0$, with y_0 and y_0^* being corresponding edges. Let the vertices of y_0^* be the input vertices for G^*. Let the contact variables of G and G^* be complementary; that is, let

$$y_i^* = y_i'. \tag{9–17}$$

Then the switching functions F and F^*, of G and G^* respectively, are also complementary; that is,

$$F \cdot F^* = 0 \qquad \text{and} \qquad F + F^* = 1, \tag{9–18a}$$

so that

$$F^* = F'. \tag{9–18b}$$

Proof. This theorem is a direct consequence of Whitney's result (Problem 3–16). (Shannon [159] proved that $FF^* = 0$ by a very simple argument but did not prove that $F + F^* = 1$.) To prove this result, we note from Problem 3–16 that paths in either graph between the input vertices correspond to cut-sets separating the input vertices in the dual graph. Thus, if either switching function is 1, there is a cut-set of edges in the dual with each variable of the cut-set equal to 0. On the other hand, if either switching function is 0, so that a cut-set of edges has variables equal to 0, the corresponding path in the dual has all variables equal to 1, so that the switching function of the dual is 1. Hence the result.

Minimality in contact networks, as in conventional networks, is an essentially unresolved question. Given an arbitrary contact network, there is no known method of finding out whether it is minimal; and given an arbitrary Boolean function, there is no known method of realizing a minimal contact network for the function. However, in the case of two-terminal (or one terminal-pair) networks, there are certain known methods of attack that can sometimes be used. In the case of networks with more than two terminals, not even a method of attack is available. The only known case of a multiterminal network in which minimality has been established is the special case of the network realizing all sixteen switching functions of two variables. This special case is due to Shannon.

We begin with the outline of a few "checks" that frequently suffice to establish the minimality of a realization. Then a general method of proof due to Cardot is given, which has been used extensively by E. F. Moore of Bell Laboratories in unpublished works. Finally, Shannon's proof for the special case of a multiterminal network is discussed.

Let $F(x_1, x_2, \ldots, x_n)$ be a given Boolean function of n variables $x_1, x_2, \ldots, x_n$. One of the simplest checks that can be performed on F is to find out whether F has a realization that either does not involve a variable x_i or uses only x_i or x_i' but not both. To check this, compute

$$F(x_1, x_2, \ldots, x_n)|_{x_i=0} = g_0(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

and (9-19)

$$F(x_1, x_2, \ldots, x_n)|_{x_i=1} = g_1(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

and compare g_0 and g_1. The following situations are possible:

(a) $g_0 = g_1$, (b) $g_0 \leq g_1$, (c) $g_1 \leq g_0$,
(d) g_0 and g_1 are incomparable.

Since the function F can be expressed as

$$F = x_i g_1 + x_i' g_0 \tag{9-20}$$

(with arguments as above), these four cases are easily interpreted. In the first case,

$$F = (x_i + x_i')g_0 = g_0, \tag{9-21}$$

so that x_i is not at all needed to realize F. In the second case,

$$g_1 = g_0 + g_0' g_1, \tag{9-22a}$$

so that

$$F = x_i g_0 + x_i g_0' g_1 + x_i' g_0 = g_0 + x_i g_1. \tag{9-22b}$$

Thus, no x_i'-contacts are required to realize F. Similarly, in the third case, no x_i-contact is needed. In the last case, both x_i and x_i' are needed.

If we perform such a computation for each variable, we can find both

$$m = (\text{number of variables actually involved})$$

(that is, $g_0 \neq g_1$) and

$$n = \text{(number of variables in which only one type of contact is required)}$$

(that is, $g_0 \leq g_1$ or $g_1 \leq g_0$). Then,

$$\text{(number of contacts required)} \geq 2m - n.$$

Then if we do have a realization that uses exactly $2m - n$ contacts, we immediately know that it is minimal. A simple example in which this argument is sufficient is the case of the function

$$F = x_1x'_3x_4 + x'_2x_3x_4 + x_1x'_2 + x'_1x'_2x_3 + x'_1x_2x_3 + x_1x_2x'_3x'_4. \qquad (9\text{–}23)$$

Setting the several variables equal to 1 and 0, we have

$$\begin{aligned}
x_1 &= 0: & F_{10} &= x'_2x_3x_4 + x'_2x_3 + x_2x_3 = x_3, \\
x_1 &= 1: & F_{11} &= x'_3x_4 + x'_2x_3x_4 + x'_2 + x_2x'_3x'_4 = x'_2 + x'_3, \\
x_2 &= 0: & F_{20} &= x_1x'_3x_4 + x_3x_4 + x_1 + x'_1x_3 = x_1 + x_3, \\
x_2 &= 1: & F_{21} &= x_1x'_3x_4 + x'_1x_3 + x_1x'_3x'_4 = x_1x'_3 + x'_1x_3, \\
x_3 &= 0: & F_{30} &= x_1x_4 + x_1x'_2 + x_1x_2x'_4 = x_1, \\
x_3 &= 1: & F_{31} &= x'_2x_4 + x_1x'_2 + x'_1x'_2 + x'_1x_2 = x'_1 + x'_2, \\
x_4 &= 0: & F_{40} &= x_1x'_2 + x'_1x'_2x_3 + x'_1x_2x_3 + x_1x_2x'_3 \\
& & &= x'_1x_3 + x_1(x'_2 + x'_3), \\
x_4 &= 1: & F_{41} &= x_1x'_3 + x'_2x_3 + x_1x'_2 + x'_1x'_2x_3 + x'_1x_2x_3 \\
& & &= x'_1x_3 + x_1(x'_2 + x'_3).
\end{aligned} \qquad (9\text{–}24)$$

($x'_2x_3 = x_1x'_2x_3 + x'_1x'_2x_3$, which are contained in $x_1x'_2$ and x'_1x_3, respectively.)

Since F_{10} and F_{11} are not comparable, both x_1- and x'_1-contacts are required. Since $F_{21} \leq F_{20}$, only x'_2-contact is required. Since F_{30} and F_{31} are not comparable, both x_3- and x'_3-contacts are required. Finally, since $F_{40} = F_{41}$, the variable x_4 is not involved. Using the expansion in terms of x_1, we find that

$$F = x_1F_{11} + x'_1F_{10} = x_1(x'_2 + x'_3) + x'_1x_3. \qquad (9\text{–}25)$$

The minimal realization of F is found immediately, as in Fig. 9–2.

In the cases of small numbers of variables (three or four), it is fairly easy to go one step further and examine whether the function has a realization with $2m - n$ contacts. If we can show that such a realization does not exist, and we have a $(2m - n + 1)$-contact realization, then we have a minimal network. To show that a $(2m - n)$-contact realization

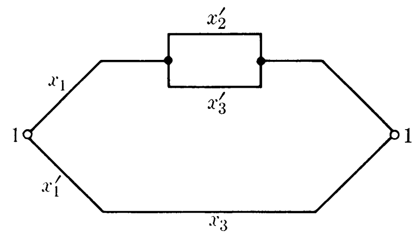


FIG. 9–2. Minimal realization of function.

does not exist, we apply the synthesis technique for *SC*-networks. However, whereas *SC*-functions have a unique "normal form" as a sum of products, the non-*SC* ones do not and it is necessary to check all possible ways of expressing the function in the $2m - n$ variables (primed and unprimed variables).

Consider, for example, the function

$$F(w, x, y, z) = wxz + wyz + xyz'. \tag{9–26}$$

The Karnaugh map of this function is shown in Fig. 9–3. By inspection of the map, it is evident that all four variables are actually involved, and so is z', so that

$$m = 4, \qquad n = 3, \qquad 2m - n = 5. \tag{9–27}$$

To test whether there is a five-contact realization of F, it is necessary to try all possible ways of writing F, involving only the variables w, x, y, z, and z'. In each of these expressions, the 1 in the y'-subcube ($wxy'z$), the 1 in the w'-subcube ($w'xyz'$), and the 1 in the x'-subcube ($wx'yz$) must be taken along with an adjacent 1, since w', x', and y' cannot appear in

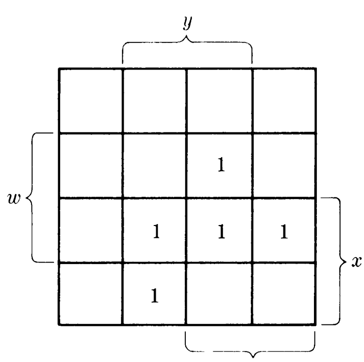


FIG. 9–3. Karnaugh map of function.

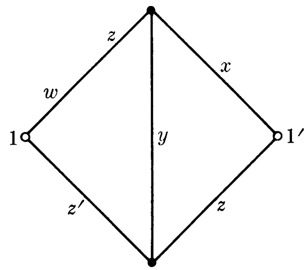


FIG. 9–4. Minimal realization of Eq. (9–26).

a five-contact realization. This condition immediately limits the realization to the two forms

$$F(w, x, y, z) = xyz' + wxz + wyz \tag{9–28a}$$

and

$$F(w, x, y, z) = xyz' + wxz + wyz + wxy, \tag{9–28b}$$

the last term in Eq. (9–28b) being redundant. We should also consider the possibility of adding either or both of the redundant terms $wxyz$ and $wxyz'$ to each of these expressions.

For Eq. (9–28a), the matrix $\mathbf{B}_F$ is

$$\mathbf{B}_F = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccccc} y_0 & w & x & y & z & z' \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array} \cdot \tag{9–29}$$

Taking the sum mod 2 of these rows, the path z' results, which is not contained in F. Therefore this realization is no good. Also, the same three rows have to be contained in every matrix $\mathbf{B}_F$ written for Eq. (9–28a) or (9–28b), including either or both of $wxyz$ and $wxyz'$. Therefore, none of these matrices, if realized, will give F. Therefore, F does not have a five-contact realization. E. F. Moore [120] has obtained the six-contact realization of this function shown in Fig. 9–4, which by the argument above is minimal.

Let us turn next to Cardot's contribution. Cardot showed that a minimal realization of the parity function of n variables,

$$F_n = x_1 \oplus x_2 \oplus \cdots \oplus x_n \tag{9–30}$$

(and its complement), requires $4n - 4$ contacts, if $n \geq 2$. To prove this result, we make use of the following properties of the parity function.

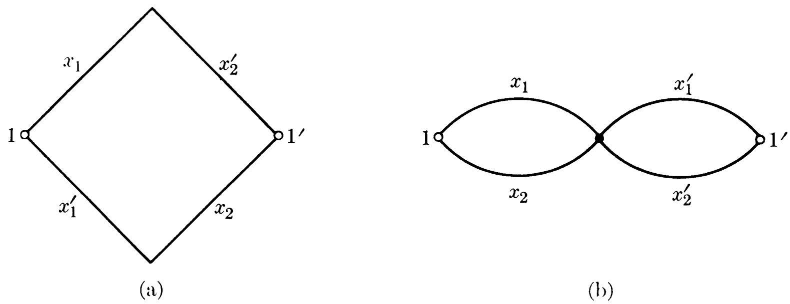


FIG. 9–5. Realization of $x_1 \oplus x_2$.

Property 1.

$$F'_n(x_1, x_2, \ldots, x_n) = F_n(x_1, \ldots, x_{i-1}, x'_i, x_{i+1}, \ldots, x_n); \tag{9–31}$$

that is, F'_n is obtained simply by replacing one of the variables in F_n by its complement.

Property 2.

$$F_n(x_1, \ldots, x_n)|_{x_i=0} = F_{n-1}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

and (9–32)

$$F_n(x_1, \ldots, x_n)|_{x_i=1} = F'_{n-1}(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Property 3. Given a set of values of $x_1, \ldots, x_n$, let F_n have the value y_1 in this state. Now if we change *one* of the values of $(x_1, \ldots, x_n)$ from 0 to 1 or vice versa, the value of the function changes to y'_1.

From Property 3, it follows that every realization of F uses both x_i and x'_i and that every path between the terminals must pass through either x_i or x'_i ($1 \leq i \leq n$). For, if there is a path not using x_i or x'_i, then by a suitable choice of values of the other variables, F_n can be made unity independently of the value of x_i, contradicting Property 3.

To prove the result, proceed by induction, starting with the parity function of two variables. For $n = 2$, it follows from Property 3 that we need four contacts, and by examining all possible ways of assigning variables to graphs with four edges, we conclude that the only two networks that realize the parity function of two variables

$$F_2(x_1, x_2) = x_1 \oplus x_2 = x_1x'_2 + x'_1x_2 \tag{9–33}$$

are those shown in Fig. 9–5.

For the parity function $F_3(x_1, x_2, x_3)$, two possible methods are available to show that it cannot be realized with less than eight contacts.

From Property 3, it follows that we need at least six. We can now exhaust all the six- and seven-contact networks and show that none of them will realize $F_3(x_1, x_2, x_3)$ under any assignment of variables. Or we can use Gould's extension of the single-contact synthesis. The parity function $F_3(x_1, x_2, x_3)$ has a unique "and, or, not" expression:

$$F_3(x_1, x_2, x_3) = x_1x_2x_3 + x_1x_2'x_3' + x_1'x_2x_3' + x_1'x_2'x_3. \tag{9-34}$$

The matrix $\mathbf{B}_F$ for this expression, using only one contact of each type, is

$$\mathbf{B}_F = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccccc} y_0 & x_1 & x_1' & x_2 & x_2' & x_3 & x_3' \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}. \tag{9-35}$$

The mod 2 sums are

rows 1, 2, 3: $x_1'x_2'x_3$,

rows 1, 2, 4: $x_1'x_2x_3'$,

rows 1, 3, 4: $x_1x_2'x_3'$,

rows 2, 3, 4: $x_1x_2x_3$,

which are all contained in F. The rank of $\mathbf{B}_F$ is 3, and the appropriate cut-set matrix is found to be

$$\mathbf{Q} = \begin{array}{c} \begin{array}{ccccccc} y_0 & x_1 & x_1' & x_2 & x_2' & x_3 & x_3' \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}. \tag{9-36}$$

Columns x_1, x_2, and x_3 of this matrix constitute one of the matrices known to be irreducible to an incidence matrix (Eq. 5–40), which can also be readily checked by trying all possibilities.

To get a seven-contact realization, an extra contact is introduced in one of the three variables or their primes. Since F_3 is symmetric, it suffices to try one variable. Let us split x_1 into two contacts x_{11} and x_{12}.

Then the matrix $\mathbf{B}_F$ becomes

$$\mathbf{B}_F = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccccccc} y_0 & x_{11} & x_{12} & x_1' & x_2 & x_2' & x_3 & x_3' \end{array} \\ \begin{bmatrix} 1 & 1 & a & 0 & 1 & 0 & 1 & 0 \\ 1 & b & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}, \tag{9–37}$$

where a and b may be either 0 or 1. Taking mod 2 sums we find

rows 1, 2, 3: $[(a' + b')x_1]x_1'x_2'x_3$,

rows 1, 2, 4: $[(a' + b')x_1]x_1'x_2x_3'$,

rows 1, 3, 4: $x_1x_2'x_3'$,

rows 2, 3, 4: $x_1x_2x_3$,

where the notation $[(a' + b')x_1]$ is used to denote that x_1 appears in these paths if $a' + b' = 1$ and not otherwise. These sums are contained in F independently of the value of a or b. The rank of $\mathbf{B}_F$, however, depends on a and b. If $a = b = 1$, the rank is 3; otherwise it is four. The cut-set matrices for the four cases are the following.

Case 1. $a = b = 1$:

$$\mathbf{Q} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccccc} y_0 & x_{11} & x_{12} & x_1' & x_2 & x_2' & x_3 & x_3' \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}. \tag{9–38a}$$

Case 2. $a = 0,\ b = 1$:

$$\mathbf{Q} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccccccc} y_0 & x_{11} & x_{12} & x_1' & x_2 & x_2' & x_3 & x_3' \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}. \tag{9–38b}$$

Case 3. $a = 1$, $b = 0$:

$$\mathsf{Q} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{cccccccc} y_0 & x_{11} & x_{12} & x'_1 & x_2 & x'_2 & x_3 & x'_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}. \tag{9–38c}$$

Case 4. $a = b = 0$:

$$\mathsf{Q} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \!\!\begin{array}{c} \begin{array}{cccccccc} y_0 & x_{11} & x_{12} & x'_1 & x_2 & x'_2 & x_3 & x'_3 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}. \tag{9–38d}$$

The matrices of cases 1 and 4 are unrealizable, since they contain the same unrealizable matrix

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

in columns x_{11}, x_2, x_3 and in columns x_{11}, x'_1, x_3, respectively. The matrices of cases 2 and 3 are unrealizable because they contain a column of zeros. We may next split x'_1 and arrive at the same conclusion as before.

Thus, at least eight contacts are required to realize $F_3(x_1, x_2, x_3)$. An eight-contact realization of F_3 is shown in Fig. 9–6. which has now been proved to be minimal.

Cardot's argument really begins at this point. Cardot defines two parameters:

$$L(n) = \begin{pmatrix} \text{number of contacts in the} \\ \text{minimal realization of } F_n \end{pmatrix},$$

$$M(n) = \begin{pmatrix} \text{Minimum number of contacts on the relay} \\ \text{having the largest number of contacts} \end{pmatrix}, \tag{9–39}$$

with the minimum $M(n)$ being computed over all realizations of

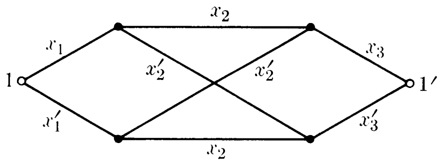


FIG. 9-6. Minimal realization of $x_1 \oplus x_2 \oplus x_3$.

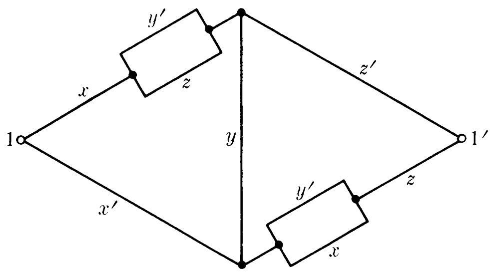


FIG. 9-7. A realization of $x_1 \oplus x_2 \oplus x_3$.

$F_n(x_1, x_2, \ldots, x_n)$. It has been shown that

$$L(3) = 8,$$

so $M(3) > 2$ [otherwise $L(3) \leq 6$]. In fact, $M(3) = 3$, as shown by the realization of Fig. 9-7. Here each relay has three contacts.

The numbers $L(n)$ and $M(n)$ satisfy the following inequalities, from Property 2 of parity functions:

$$L(n) \geq L(n-1) + M(n) \tag{9-40a}$$

and

$$M(n) \geq M(n-1), \tag{9-40b}$$

both of which are obvious consequences of the fact that we can get a realization for F_{n-1} or F'_{n-1} by setting a variable equal to 0 or 1.

From the inequalities (9-40), we need only to show the following:

(a) $M(4) = 4$.

(b) There is a realization of F_n using $4n - 4$ contacts. For, $M(n)$ is a nondecreasing function of n, by Eq. (9-40a). Therefore, for $n \geq 4$, if $M(4) = 4$,

$$\begin{aligned} L(n) &\geq L(n-1) + M(n) \geq L(n-2) + M(n-1) + M(n) \\ &\geq \cdots \geq L(3) + (n-3)M(4) \\ &= L(3) + 4n - 12 \\ &= 4n - 4. \end{aligned} \tag{9-41}$$

Let us therefore show that $M(4) = 4$, and exhibit a $(4n - 4)$-realization. Suppose, in fact, that $M(4) = 3$. Then we would have a realization of F_4 with at least eleven contacts (since F_3 requires eight). Then there would be three contacts each for three of the variables and either two or three contacts for the last variable. If we set the last variable equal to 0 we get a realization of F_3 by using three contacts per variable; and if we set the last variable equal to 1 we get a realization of F_3' by using the *same three contacts per variable.* We now show that we cannot realize both F_3 and F_3' by using the same three contacts on each of the variables.

Without any loss of generality, we may assume that there are one break (primed) and two make (unprimed) contacts per variable, and the function realized is F_3'. For, each variable has to appear both primed and unprimed; and interchanging a variable and its complement changes the function to its complement. Thus, by suitably interchanging variables with complements, we may assume that there are one break and two make contacts per variable. By interchanging y and y' in Fig. 9–7, we see that F_3' is realized with one break and two make contacts per variable. Therefore it suffices to prove that F_3 cannot be realized with the set of contacts

$$x,\ y,\ z,\ x,\ y,\ z,\ x',\ y',\ z'.$$

Since the problem has been so well defined, it is possible to use the matrix technique to show that F_3 is unrealizable with the given set of contacts. We leave this method as a problem and proceed with Cardot's proof. Cardot's argument is long and involved because a number of different cases have to be considered and the argument is geometrical. We condense it slightly by making a few observations on cut-sets before beginning Cardot's argument.

The Karnaugh map of the function

$$F_3(x, y, z) = xyz + xy'z' + x'yz' + x'y'z \tag{9–42}$$

is shown in Fig. 9–8. Any network realizing F_3 cannot have a cut-set separating terminal vertices 1 and $1'$ consisting of a single edge, since F_3 cannot be made zero by fixing one variable. Further, any cut-set $(1, 1')$ of only two edges must consist of a variable and its complement. In

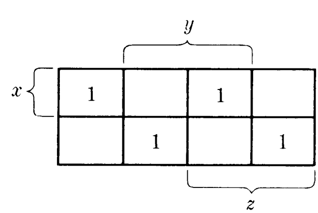


FIG. 9–8. Karnaugh map of F_3.

other words, there is no proper cut-set of width 2. Finally, there is no cut-set consisting of break contacts only.

To proceed with Cardot's proof, the following cases have to be considered.

Case 1. Neither terminal is connected to a break contact.

Case 2. One terminal, say 1, is not connected to a break contact.

Case 3. Both terminals are connected to break contacts.

Case 1. Since the set of edges incident at each terminal node is necessarily a proper cut-set (consisting only of make contacts), it follows that three make contacts are at terminal 1 and the other three at terminal 1′, making the path xyz impossible. (This path is not contained in any other.)

Case 2. It follows that each of the contacts x, y, and z is connected to terminal 1 because this is the only cut-set of make contacts. There are two subcases depending on the connection at terminal 1′. There has to be at least one make contact at terminal 1′ to make the path xyz possible. Due to the symmetry, we may assume that the break contact x' is connected to terminal 1′.

Subcase 2(a). Terminal 1′ is not connected to a make and break contact of the same variable; that is, contact x is not at terminal 1′. By symmetry, let y be at 1′. From the path $x'yz'$, it follows that z', which is single, is not connected to 1 (since 1 has no break contact connected to it) or 1′ (since the single x' is at 1′). Hence z' is between terminals 2 and 3, as in Fig. 9–9. Now to get the path $xy'z'$, either x or y' must be connected to 1′, since z' is not. Both cases are prohibited by the assumption that the make and break contacts of the same variable are not connected to 1′. We also see incidentally that neither 1 nor 1′ can be connected to two break contacts since the paths $x'y'z$, $xy'z'$, and $x'yz'$ demand that any two break contacts be in series, and each is single. (The matrix argument uses essentially this fact.)

Subcase 2(b). The terminal 1′ is connected to a make and break of the same variable, say x. By the observation above, x' is the only break

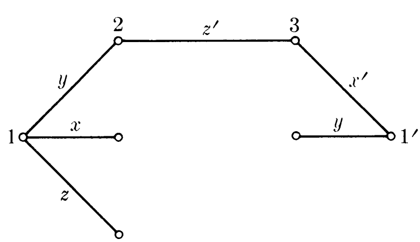


FIG. 9–9. Illustration of Subcase 2(a).

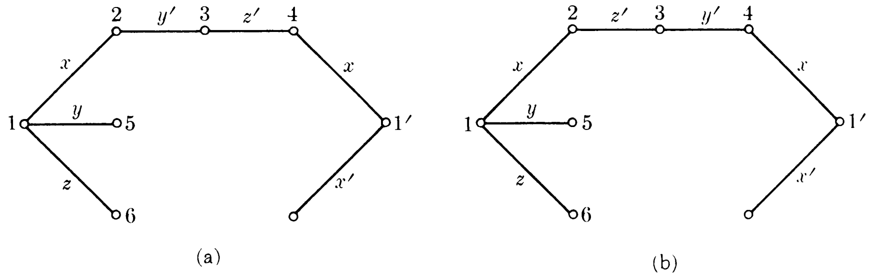


FIG. 9–10. Illustration of Subcase 2(b).

contact at 1′. Since the two x-contacts are used up, the only way to get $xy'z'$ is to connect y' and z' in series between the two x-contacts, as in Fig. 9–10(a) or (b). Since the path $x'yz'$ is required in Fig. 9–10(a), the points 5 and 3 must be connected either directly or through y, leading to a sneak path xyz' (no x' is available). In Fig. 9–10(b), the path $x'y'z$ demands a connection between 3 and 6 either directly or through z, leading to a sneak path $xy'z$.

Case 3. Let each terminal be connected to a break contact. By the observation in Subcase 2(a), each terminal can be connected to only one break contact. Since there is only one break contact of each variable, let x' be at 1 and y' be at 1′. From the path $x'yz'$, we conclude that y is at 1′. Similarly, from the path $xy'z'$, x is at 1. z' must be connected to either the y-contact at 1′ or the x-contact at 1. The two cases are identical, and so let z' be connected to x, as in Fig. 9–11. Now from the path $x'yz'$, points 3 and 4 must be connected either directly or through y, leading to sneak path xyz'.

Thus we have shown that $M(4) = 4$ and thereby proved the next theorem.

THEOREM 9–9 (Cardot). A necessary and sufficient number of contacts to realize the parity function of n variables

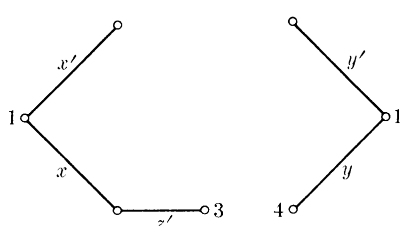


FIG. 9–11. Illustration of Case 3.

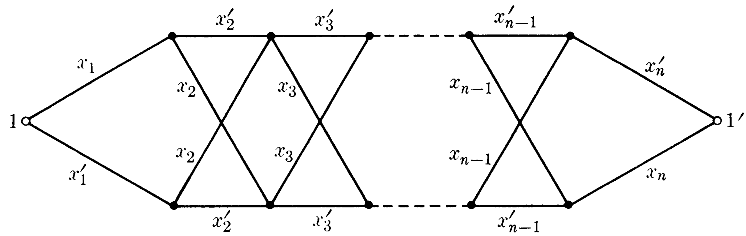


FIG. 9-12. Minimal realization of the parity function.

$$F_n(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \cdots \oplus x_n, \tag{9-43}$$

where $n \geq 2$, is $4n - 4$.

The sufficiency follows from the well-known realization of Fig. 9-12. Cardot's technique has been extended by Vasil'ev [189], who obtained five general theorems concerning the number of contacts required in a minimal realization. Using these theorems, Vasil'ev is able to find the minimal contact realizations for all functions of four variables. However, as of the time of this writing, Vasil'ev has not published the proofs of his theorems. Therefore we are unable to discuss his results.

Let us next turn to C. E. Shannon's contribution to minimality theory. As contrasted with Cardot's involved argument making use of very simple concepts, Shannon's argument is short and elegant, but makes strong use of graph theory. Mr. H. K. Bhavnani, a graduate student in the Massachusetts Institute of Technology switching circuits course in 1952, devised a network containing eighteen contacts which realizes all sixteen switching functions of two variables. Shannon, in an unpublished memorandum in 1953, proved that this network is minimal in contacts.

THEOREM 9-10 (Shannon). The necessary and sufficient number of contacts required to realize all sixteen switching functions of two variables, as switching functions between one fixed contact and each of sixteen other contacts, is eighteen.

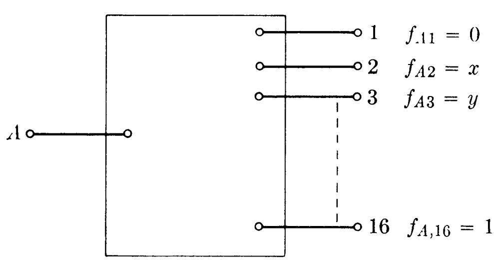


FIG. 9-13. Seventeen-terminal network.

Proof. Such a network is shown as a "black box" in Fig. 9–13. Of the sixteen terminals, terminal 1 realizing the open circuit is never connected to A; terminal 16 realizing the short circuit is always connected to A, and so it is effectively the same terminal. But the others are sometimes connected to A and sometimes not, depending on the values of x and y. Therefore, the linear graph corresponding to the contact network excluding terminal 1 is a connected linear graph with at least fifteen vertices $(A, 2, 3, 4, \ldots, 15)$. Now Shannon's proof consists of showing that nullity of the linear graph is at least 4:

$$\mu \geq 4, \tag{9–44}$$

so that the number of edges e, from

$$\mu = e - v + 1, \tag{9–45a}$$

is

$$e = \mu + v - 1 \geq 4 + 15 - 1 = 18, \tag{9–45b}$$

since

$$\mu \geq 4 \tag{9–45c}$$

and

$$v \geq 15. \tag{9–45d}$$

In fact, any network which realizes the four functions

$$\begin{aligned} f_1 &= xy + x'y', & f_2 &= xy' + x'y, \\ f_3 &= x + y, & f_4 &= x' + y' \end{aligned} \tag{9–46}$$

must have a nullity of at least 4. Let N_1, N_2, N_3, and N_4 be the nodes at which these four functions are realized. Consider N_1, realizing $xy + x'y'$. The edges incident at this node may be divided into two classes: those labeled x or y, and those labeled x' or y'. Since node A must be connected to N_1 under $xy = 1$, the first set is nonempty and, similarly, since $x'y' = 1$ also leads to a path from A to N_1, the second set is nonempty. Thus, there are at least two distinct paths from A to N_1 which contain a circuit. Now separate the node N_1 into two nodes, with the unprimed edges at one node and the primed edges at the other. This separation reduces the nullity of the network by one. We next show that this separation does not affect the realization of f_2, f_3, or f_4. No path from A to N_2 could have gone through N_1, since f_1 and f_2 are disjunctive $(f_1 \cdot f_2 = 0)$. Hence the realization of f_2 is unaffected. No previously existing path from A to N_3 has been affected. For, then, such a path should have gone through an x' and a y or an x and a y'. But if this were so, there would

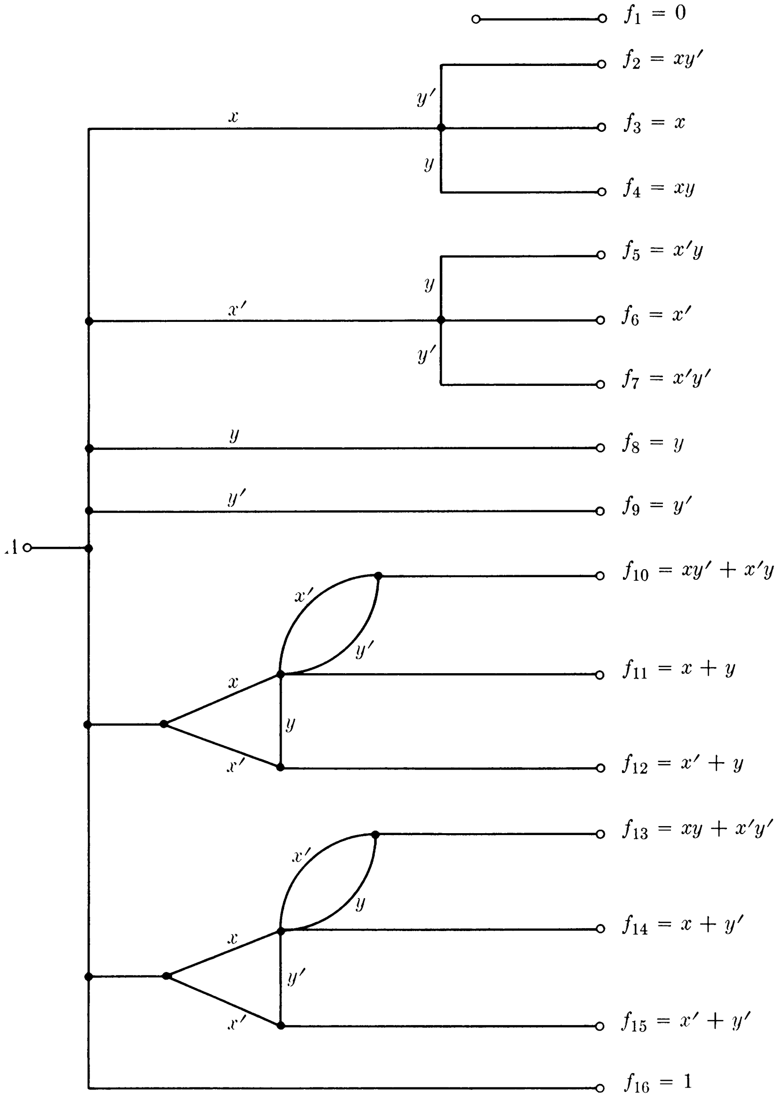


FIG. 9-14. Minimal realization of sixteen functions of two variables.

have been a path from A to N_1 under xy' or under $x'y$, which is impossible since

$$f_1 = xy + x'y'. \tag{9-47}$$

Thus f_3 is still realized. By an identical argument, f_4 is unaffected by the separation of N_1. Now separate N_2 into two nodes, with one node having edges labeled x or y' and the other having edges x' or y. Again, these two classes are nonempty, so that we destroy a circuit by this separation, decreasing the nullity of the network by one. The realizations of f_3 and f_4 are unaltered by this process since the only paths destroyed are xy and $x'y'$, which are not contained in f_2. Next, separate N_3 according to the labels x or y' and x' or y. If either class were empty, $f_3 = x + y$ could not be realized. This separation again reduces the nullity by one without destroying the realization of f_4, for the only paths destroyed are xy and $x'y'$. $f_4 \neq 1$ under xy, and $f_3 \neq 1$ under $x'y'$. Finally, separate N_4 into two classes depending on whether the incident edges are primed or unprimed.

Thus we have reduced the nullity of the graph by four. Since $\mu \geq 0$ always, the original nullity must have been at least four. Thus, the necessity of eighteen contacts has been proved. The sufficiency is established by exhibiting an eighteen-contact network realizing all sixteen functions for two variables. This network is shown in Fig. 9–14.

9–2 Sequential machines. A sequential switching system or a *sequential machine* has been abstractly defined by Moore [117] as follows.

DEFINITION 9–7. *Sequential machine.* A *sequential machine* M is a finite collection of

(a) states $q_1, q_2, \ldots, q_m$,

(b) inputs $i_1, i_2, \ldots, i_n$,

and

(c) outputs $\omega_1, \omega_2, \ldots, \omega_n$,

such that the present output is a function of the present state, and the next state is a function of the present state and the present input.

So far as the abstract Definition 9–7 is concerned, the names *state*, *input*, and *output* may be taken as undefined concepts. From a practical viewpoint, they have rather familiar interpretations. Such an abstract sequential machine has a representation as a directed graph. The vertices of the graph correspond to the states of the machine, and the edges correspond to transitions. In the pictorial representation, one draws small circles for the vertices. The name of the state and the associated output are written inside the circle. The edges are marked with the inputs causing the transition. An example of such a diagram is given in Fig. 9–15. The directed graph of Fig. 9–15 has a weight associated with each vertex (the output) and each edge (the input). A weighted directed graph of this type is called a *net* and has many applications. We refer to Hohn, Seshu, and Aufenkamp [77] for a general discussion of nets. The objective now

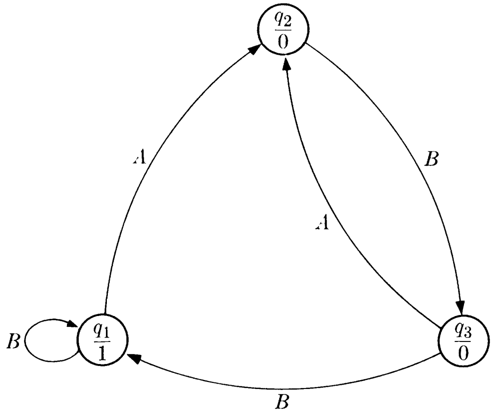


FIG. 9-15. A state diagram.

is merely to show the relationship between the connection matrix of Hohn and Aufenkamp and the incidence matrix of the graph and give the "state-removal" algorithm that finds an interesting parallel in Mason's theory of signal-flow graphs (see Chapter 10).

DEFINITION 9-8. *Connection matrix.* The *connection matrix*

$$\mathsf{C} = [c_{ij}]_{m,m} \tag{9-48}$$

of the sequential machine has one row and one column for each state of the machine and is defined by

$$c_{ij} = \sum_k w_{ij}^k, \tag{9-49}$$

where w_{ij}^k is the input associated with an edge from i to j and the sum (with the interpretation "or") is over all such edges (which go from i to j).

The connection matrix, together with the specification of vertex weights, is thus equivalent to the graph. In attempting to relate the connection matrix to the incidence matrix, we encounter the difficulty that we cannot define A_a for a graph that contains self-loops (as at the state q_1 of Fig. 9-15). To overcome this difficulty, we define two matrices of incidence as follows.

DEFINITION 9-9. *Positive incidence matrix* and *negative incidence matrix.* The *matrices of incidence*

$$\mathsf{A}_a^+ = [a_{ij}^+]_{v,e} \qquad \text{and} \qquad \mathsf{A}_a^- = [a_{ij}^-]_{v,e} \tag{9-50}$$

are defined by

$$\begin{array}{lll} a_{ij}^{+}(a_{ij}^{-}) = 1 & \text{if edge } j \text{ is incident at vertex } i \text{ and is} & \\ & \text{oriented away from (toward) vertex } i; & (9\text{–}51) \\ a_{ij}^{+}(a_{ij}^{-}) = 0 & \text{otherwise.} & \end{array}$$

Thus, if the graph contains no self-loops,

$$\mathsf{A}_a = \mathsf{A}_a^{+} - \mathsf{A}_a^{-}. \qquad (9\text{–}52)$$

The incidence matrix may be related to the connection matrix as follows [8, 37, 77].

THEOREM 9–11. Let $\mathbf{W}$ be a square matrix of order e, defined by

$$\begin{aligned} \mathbf{W} &= [w_{ij}]_{e,e}, \\ w_{ii} &= (\text{weight of edge } i), \\ w_{ij} &= 0 \qquad \text{if } i \neq j. \end{aligned} \qquad (9\text{–}53)$$

Then the connection matrix C is related to the matrices of incidence by

$$\mathsf{C} = \mathsf{A}_a^{+}\mathsf{W}(\mathsf{A}_a^{-})^{T}. \qquad (9\text{–}54)$$

The proof of this theorem is left as a problem (Problem 9–10).

We turn next to the state-removal algorithm of Aufenkamp and Hohn. Before giving the algorithm, let us briefly state the postulates satisfied by the edge weights $i_1, i_2, \ldots, i_n$. The set S of input sequences satisfies:

P_1. If $i_1, i_2 \in S$, so is $i_1 + i_2 = i_2 + i_1$. (+ is "or.")

P_2. $i_1 + (i_2 + i_3) = (i_1 + i_2) + i_3$.

P_3. $i_1, i_2 \in S$ implies $i_1 \cdot i_2 \in S$.

P_4. $i_1 \cdot (i_2 \cdot i_3) = (i_1 \cdot i_2) \cdot i_3$.

P_5. $i_1(i_2 + i_3) = i_1i_2 + i_1i_3$; $(i_2 + i_3)i_1 = i_2i_1 + i_3i_1$.

P_6. $i_1i_2 + i_1i_3i_2 = i_1i_2$.

P_7. $0 + i_1 = i_1$.

P_8. $0 \cdot i_1 = i_1 \cdot 0 = 0$.

In this algebra, $i_1 \cdot i_2$ signifies i_2 following i_1, and so in general $i_1i_2 \neq i_2i_1$.

It is possible to prove results analogous to those of Lunts [103] and Hohn and Schissler [76] for the connection matrix C. Specifically, we have the result stated in Theorem 9–12.

THEOREM 9–12. The (i, j)-entry of

$$\mathbf{C}^r = \mathbf{C} \cdot \mathbf{C} \cdot \mathbf{C} \cdots \mathbf{C} \qquad (r \text{ factors}) \tag{9–55}$$

lists exactly the paths of length r from vertex i to vertex j.

This theorem is fairly obvious. By using the postulates above, one can show that

$$\sum_{r=1}^{n} \mathbf{C}^r = \sum_{r=1}^{n+1} \mathbf{C}^r, \tag{9–56}$$

so one can get all the possible paths from vertex i to vertex j by exponentiating $\mathbf{C}$ and adding.

However, although we have elegant results in following this procedure, we lose some information by disregarding the self-loops at intermediate nodes, by virtue of the absorption law P_6. This law is what makes $\sum_r \mathbf{C}^r$ converge for $r \geq n$. But it destroys useful information in $\mathbf{C}$. In the study of signal-flow graphs, in Chapter 10, self-loops are of the utmost importance and cannot possibly be neglected. To show this analogy, let us disregard P_6 and give a modified version of the Hohn-Aufenkamp state-removal algorithm.

The state-removal algorithm is a method of finding all the possible paths of length less than n (where n is the number of nodes) between nodes i and j of a net. The net is represented by its connection matrix $\mathbf{C}$. In $\mathbf{C}$, permute rows and columns so that i and j are the first two rows and columns, and the state to be removed corresponds to the last row and column. If there is no self-loop at last node k, the state removal consists of adding the paths that go through node k, to the appropriate entries of $\mathbf{C}$, and then deleting the last row and column. Formally, replace an entry c_{pq} of $\mathbf{C}$ by

$$c'_{pq} = c_{pq} + c_{pk}c_{kq}. \tag{9–57}$$

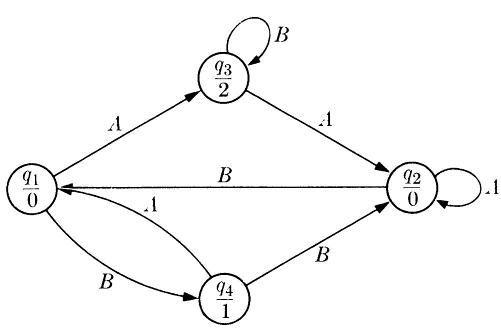


FIG. 9–16. Example for state removal.

This operation is most easily performed by taking the entry in the last column to the right of c_{pq} and postmultiplying it by the entry in the last row below c_{pq}. As an example, consider the diagram of Fig. 9–16 and the associated connection matrix

$$\mathsf{C} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & A & B \\ B & A & 0 & 0 \\ 0 & A & B & 0 \\ A & B & 0 & 0 \end{bmatrix} \end{array}. \tag{9–58}$$

Suppose that we wish to find all ways of getting from node 1 to node 2, going through at most two intermediate nodes. Let us start by pulling (removing) node 4. Since there is a B in the (1, 4)-position and zeros elsewhere in the last column, only paths from 1 will be affected. Starting with the (1, 1)-position, there is B at the right and A below. So add BA to the (1, 1)-position. Similarly, add BB to the (1, 2)-position. No other position is affected because of the zero in the (4, 3)-position. Now we may delete row 4 and column 4, leaving

$$\mathsf{C}_{(4} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} BA & BB & A \\ B & A & 0 \\ 0 & A & B \end{bmatrix} \end{array}. \tag{9–59}$$

This matrix contains all the paths among the nodes 1, 2, and 3 that are in Fig. 9–16.

If we followed the same procedure again, we would completely disregard the B in the (3, 3)-position corresponding to the self-loop at node 3. This would be in keeping with P_6 but would destroy important information. Therefore, we have to give a modified procedure (which has an analogue in signal-flow graphs). Now any path that goes through node 3 may either go directly through node 3 or may go through after several B's have occurred. Therefore, we have to *insert* this information between the incoming and outgoing edges at node 3. We may, for example, either *premultiply* all the entries of the third *row* by $1 + \sum B^k$, using 1 as identity for multiplication (but $1 + B \neq 1$), or *postmultiply* all entries of the third *column* by $1 + \sum B^k$, before proceeding. Here k is used to indicate a nonnegative integer.

Using the first procedure, we write first

$$\mathsf{C}_{(4(3s} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} BA & BB & A \\ B & A & 0 \\ 0 & (1+\sum B^k)A & 0 \end{bmatrix} \end{array} \cdot \tag{9–60}$$

Note that the (3, 3)-element is replaced by 0, thus "removing the self-loop." Now we repeat the original process and remove node 3, leaving

$$\mathsf{C}_{(4(3} = \begin{array}{c} 1 \\ 2 \end{array} \begin{bmatrix} BA & BB + A(1+\sum B^k)A \\ B & A \end{bmatrix}, \tag{9–61}$$

which lists all possible ways of getting from 1 to 2 without going through 1 or 2.

Yet another matrix that is useful in the analysis of sequential machines, and in general any system represented by a net, is a matrix which describes the distribution of edges with a given weight. In the theory of sequential machines, these matrices are known as *transition matrices;* the same matrices are *relation matrices* in the algebra of relations and *structure matrices* in the study of neuron networks (see Chapter 10). The transition matrix introduced by Seshu, Miller, and Metze [155] is defined as follows.

DEFINITION 9–10. *Transition matrix.* For each input i, the *transition matrix* T^i is an $(n \times n)$-matrix, where n is the number of states, with

$$\begin{aligned} \mathsf{T}^i &= [t^i_{kj}], \\ t^i_{kj} &= 1, \quad \text{if there is an edge with weight } i \text{ } \textit{from} \text{ state } k \text{ } \textit{to} \text{ state } j, \\ t^i_{kj} &= 0 \quad \text{otherwise.} \end{aligned} \tag{9–62}$$

Transition matrices are most directly applicable to *completely specified synchronous machines.* They can also be applied to completely specified asynchronous machines with some modification. Since the main purpose of this discussion is to introduce the basic concepts, only the completely specified synchronous case is considered. Such a machine is defined as follows.

DEFINITION 9–11. *Completely specified synchronous machine.* The state diagram G represents a *completely specified synchronous machine* if for each permissible input i and each state q_k, there is an edge *leaving* q_k with weight i (which may also be a self-loop).

For such machines, each row of the transition matrix contains precisely one 1, all other entries being zeros. Many interesting properties of the machine, in particular equivalence characteristics of states, are invariant characteristics of transition matrices—invariant under matrix multiplication. To bring out these characteristics, the state and output vectors are next defined.

DEFINITION 9–12. *State vector* and *output vector*. The *state vector* $\mathbf{Q}_0$ and the *output vector* $\boldsymbol{\Omega}_0$ are defined by

$$\mathbf{Q}_0 = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Omega}_0 = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix}, \tag{9–63}$$

where ω_j is the output in state q_j.

THEOREM 9–13. If the sequential machine M is in an initial state q_k, and receives the input sequence $i_1, i_2, \ldots, i_p$, the final state and output are given by the kth rows of the matrix products

$$\mathbf{T}^{i_1}\mathbf{T}^{i_2}\cdots\mathbf{T}^{i_p}\mathbf{Q}_0 \qquad \text{and} \qquad \mathbf{T}^{i_1}\mathbf{T}^{i_2}\cdots\mathbf{T}^{i_p}\boldsymbol{\Omega}_0,$$

respectively.

This theorem is proved by induction (see Problem 9–13).

DEFINITION 9–13. *Simple equivalence.* The states $q_{j_1}, q_{j_2}, \ldots, q_{j_k}$ are *simply equivalent* if they have the same outputs and if every input sequence applied to the machine with any one of these states as the initial state leads to the same output sequence independently of which state of the set was chosen and independently of the outputs associated with the other states.

This definition differs from that of Aufenkamp and Hohn [5], and both differ from the definition of Moore [117]. If the set of Definition 9–13 is made maximal, it agrees with that of Aufenkamp and Hohn. If, further, the machine is strongly connected, all definitions agree.

THEOREM 9–14. Let the rows and columns of the transition matrices be arranged so that $j_1, j_2, \ldots, j_k$ are the first k rows and columns, and let the matrices be partitioned after the first k rows and columns as

$$\mathbf{T}^i = \begin{bmatrix} \mathbf{T}^i_{11} & \mathbf{T}^i_{12} \\ \mathbf{T}^i_{21} & \mathbf{T}^i_{22} \end{bmatrix}. \tag{9–64}$$

Then, states $q_{j_1}, q_{j_2}, \ldots, q_{j_k}$ are simply equivalent if and only if the submatrix T^i_{12} has identical rows, for each input i. (Different transition matrices may have different rows in T^i_{12}.)

Proof. The necessity is an immediate consequence of the requirement that the next outputs be the same under any output. The sufficiency follows because this characteristic is an invariant, which is proved as the next theorem.

THEOREM 9–15. Let the transition matrices be partitioned as in Theorem 9–14. Then the property that the rows of T^i_{12} are identical in each transition matrix is an invariant under matrix multiplication.

Proof. Let

$$\mathsf{T}^i = \begin{bmatrix} \mathsf{T}^i_{11} & \mathsf{T}^i_{12} \\ \mathsf{T}^i_{21} & \mathsf{T}^i_{22} \end{bmatrix} \quad \text{and} \quad \mathsf{T}^j = \begin{bmatrix} \mathsf{T}^j_{11} & \mathsf{T}^j_{12} \\ \mathsf{T}^j_{21} & \mathsf{T}^j_{22} \end{bmatrix} \tag{9–65}$$

be two transition matrices such that T^i_{12} has identical rows and T^j_{12} has identical rows. Let

$$\mathsf{T}^i\mathsf{T}^j = \begin{bmatrix} \mathsf{T}^i_{11} & \mathsf{T}^i_{12} \\ \mathsf{T}^i_{21} & \mathsf{T}^i_{22} \end{bmatrix} \begin{bmatrix} \mathsf{T}^j_{11} & \mathsf{T}^j_{12} \\ \mathsf{T}^j_{21} & \mathsf{T}^j_{22} \end{bmatrix} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix} = \tau. \tag{9–66}$$

Then

$$\tau_{12} = \mathsf{T}^i_{11}\mathsf{T}^j_{12} + \mathsf{T}^i_{12}\mathsf{T}^j_{22}. \tag{9–67}$$

Since there is only one 1 per row and T^i_{12} has identical rows, it follows that $\mathsf{T}^i_{11} = \mathbf{0}$ or $\mathsf{T}^i_{12} = \mathbf{0}$, but not both. In the first case, $\tau_{12} = \mathsf{T}^i_{12}\mathsf{T}^j_{22}$ has identical rows since T^i_{12} has identical rows. In the second case, $\tau_{12} = \mathsf{T}^i_{11}\mathsf{T}^j_{12}$ has identical rows since they are selected from the identical rows of T^j_{12}. The rest follows by induction, since τ also has only one 1 per row and τ_{12} has identical rows.

We extend the concept of equivalence further by defining multiple equivalence.

DEFINITION 9–14. *Multiple equivalence.* Let $S_1, S_2, \ldots, S_{k+1}$ be a partition of the states of M. Let the states in each of the partitions $S_1, S_2, \ldots, S_k$ have the same output. Then the sets of states $S_1, S_2, \ldots, S_k$ are *multiply equivalent* if every input sequence applied to the machine, with the machine in any one of the states of a given S_j, $1 \leq j \leq k$, leads to the same output sequence independently of the outputs associated with the partitions $S_1, S_2, \ldots, S_k$ and the states in S_{k+1}.

This definition is meaningful because of Theorem 9–16.

THEOREM 9–16. Let $S_1, S_2, \ldots, S_{k+1}$ be a partition of the states of M such that the sets $S_1, S_2, \ldots, S_k$ are multiply equivalent. Then simultaneous identification of the states in $S_1, S_2, \ldots, S_k$ (that is, replacing each of these sets by a state) leads to an equivalent machine under Moore's [117] definition.

We next give the analogue of Theorems 9–14 and 9–15, which have been stated by Aufenkamp and Hohn [5] in a different language.

THEOREM 9–17. Let the rows and columns of the transition matrices be arranged and partitioned according to the partitioning $S_1, S_2, \ldots, S_k, S_{k+1}$, as

$$\mathbf{T}^i = \begin{bmatrix} \mathbf{T}^i_{11} & \mathbf{T}^i_{12} & \cdots & \mathbf{T}^i_{1,k} & \mathbf{T}^i_{1,k+1} \\ \mathbf{T}^i_{21} & \mathbf{T}^i_{22} & \cdots & \mathbf{T}^i_{2,k} & \mathbf{T}^i_{2,k+1} \\ \vdots & & & & \\ \mathbf{T}^i_{k,1} & \mathbf{T}^i_{k,2} & \cdots & \mathbf{T}^i_{k,k} & \mathbf{T}^i_{k,k+1} \\ \mathbf{T}^i_{k+1,1} & \mathbf{T}^i_{k+1,2} & \cdots & \mathbf{T}^i_{k+1,k} & \mathbf{T}^i_{k+1,k+1} \end{bmatrix}. \tag{9–68}$$

Then the states in each of the partitions $S_1, S_2, \ldots, S_k$ are multiply equivalent if and only if they have the same output and each of the transition matrices satisfies:

(a) All but one of the submatrices in any row j (of the partitioned matrix) are zero, $1 \leq j \leq k$.
(b) The submatrix $\mathbf{T}^i_{j,k+1}$ has identical rows, $1 \leq j \leq k$.

The necessity is obvious from the definition since the outputs after one input must agree. The sufficiency follows from Theorem 9–18.

THEOREM 9–18. Properties (a) and (b) are invariants under matrix multiplication.

Theorem 9–18 follows from a computation similar to the one performed for Theorem 9–15.

Aufenkamp and Hohn [5] have modified these results by considering the connection matrix instead of the transition matrices. (Actually, Aufenkamp and Hohn consider Mealy's model of a sequential machine to avoid the undesirable clause "independently of . . ." of Definitions 9–13 and 9–14.)

The Aufenkamp-Hohn extension is seen to be natural when we note the relationship of the transition matrices to the connection matrix, as in the next theorem.

THEOREM 9-19. The transition matrices T^{i_j} and the connection matrix C of a sequential machine are related by

$$\mathsf{C} = \sum_{j=1}^{n} i_j \mathsf{T}^{i_j}, \tag{9-69}$$

where $i_1, i_2, \ldots, i_n$ are the permissible inputs to the machine.

Aufenkamp and Hohn [5] begin by defining an *input polynomial* as a polynomial in $i_1, i_2, \ldots, i_n$, with *homogeneity* and *degree* being defined in the usual fashion. They next define an *r-matrix.*

DEFINITION 9-15. *r-matrix.* A matrix B whose elements are input polynomials is called an *r-matrix* if it has all of the following properties:
(a) All nonzero entries of B are homogeneous and of degree r.
(b) All nonzero terms in each row represent distinct input sequences.
(c) All nonzero terms which appear in any given row also appear in every other row.

Now to test for equivalence (multiple equivalence, in the present terminology), Aufenkamp and Hohn [5] partition the connection matrix C in the same fashion as the transition matrices were partitioned in Theorem 9-17:

$$\mathsf{C} = \begin{bmatrix} \mathsf{C}_{11} & \mathsf{C}_{12} & \cdots & \mathsf{C}_{1k} & \mathsf{C}_{1k+1} \\ \mathsf{C}_{21} & \mathsf{C}_{22} & \cdots & \mathsf{C}_{2k} & \mathsf{C}_{2k+1} \\ \vdots & & & & \\ \mathsf{C}_{k,1} & \mathsf{C}_{k,2} & \cdots & \mathsf{C}_{k,k} & \mathsf{C}_{k,k+1} \\ \mathsf{C}_{k+1,1} & \mathsf{C}_{k+1,2} & \cdots & \mathsf{C}_{k+1,k} & \mathsf{C}_{k+1,k+1} \end{bmatrix}. \tag{9-70}$$

Only, instead of lumping all nonequivalent states in S_{k+1}, Aufenkamp and Hohn keep the nonequivalent states in separate partitions. The matrix C^r may also be partitioned similarly. Denoting the entries of C^r as $\mathsf{C}_{ij}^{(r)}$, Aufenkamp and Hohn prove the following theorems, which are seen to be very similar to Theorems 9-17 and 9-18.

THEOREM 9-20. With C and C^r partitioned as above, if each C_{ij} is a 1-matrix then each $\mathsf{C}_{ij}^{(r)}$ is an r-matrix.

THEOREM 9-21. If a connection matrix admits of a symmetrical partitioning in which each submatrix is a 1-matrix, then the states contained in each set of the corresponding partitioning of the states are equivalent.

Aufenkamp and Hohn are unable to prove the necessity of this condition because of their weaker definition of equivalence.

A procedure can be given for the reduction of the number of states in a sequential machine, based on either the connection matrix [5] or the transition matrix [155]. The reduction algorithm for transition matrices is briefly as follows. Each equivalence class of states is replaced by a single state. The submatrices of zeros are replaced by 0's and nonzero submatrices by 1's, in the submatrices $\mathbf{T}_{11}, \ldots, T_{k,k}$. Each of the submatrices $\mathbf{T}_{1,k+1}, \ldots, \mathbf{T}_{k,k+1}$ is replaced by one of its (identical) rows. $\mathbf{T}_{k+1,k+1}$ is left unaltered. This reduction procedure (based on either the transition matrix or the connection matrix) enables us to make the conclusion stated in Theorem 9–22.

THEOREM 9–22. Given a completely specified synchronous machine M, there exists a corresponding equivalent machine N with a minimum number of states, which is unique to within an isomorphism.

E. F. Moore [117] gets this result only for strongly connected machines because of his weaker definition of equivalence.

The theory of graphs is also applicable to the problem of assignment of memory states for hazard-free operation of asynchronous sequential machines. However, this application is fairly obvious. For a detailed discussion of the problem, we refer the reader to Caldwell [21], Chapter 13.

9–3 Logic networks. The only work on the application of directed graphs to logic-network realizations of Boolean functions is a master's thesis by W. A. Shelly [163]. His contribution is reviewed briefly in this section. A formal definition of a logic network becomes too complicated to be useful and so is avoided here. A *logic unit* is a "black box" with a set of input terminals and *one* output terminal, such that the output is a Boolean function of the inputs. Such a logic unit is shown in Fig. 9–17(a)

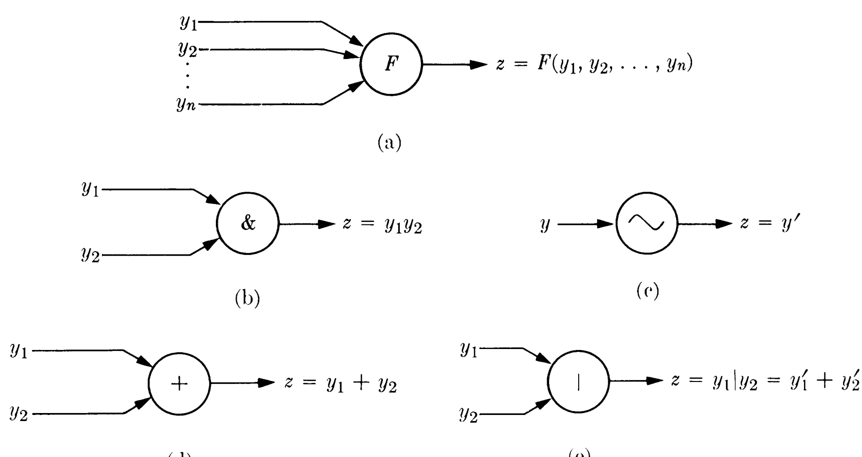


FIG. 9–17. Logic units.

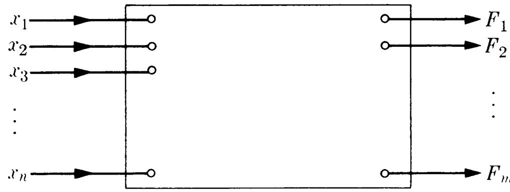


FIG. 9-18. Black-box representation of a logic network.

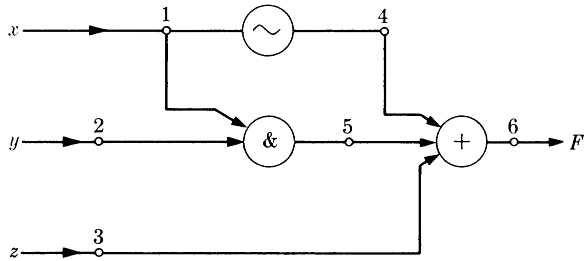
with examples in 9-17(b), (c), (d), and (e). A *logic network* is an interconnection of logic units such that no vertex is connected to more than one output and such that there are no directed circuits (feedback paths) in the network. The general logic network can be given a black-box representation as in Fig. 9-18.

In a general logic network, the functions $F_1, F_2, \ldots, F_m$ are Boolean functions of the inputs $x_1, x_2, \ldots, x_n$. In applying the theory of graphs to logic networks, one encounters the difficulty that the operations are performed in the logic units (which should normally be considered as vertices) whereas in all other applications of directed graphs, such as sequential machines, flow graphs, etc., the operations are performed by the edges of the graph.

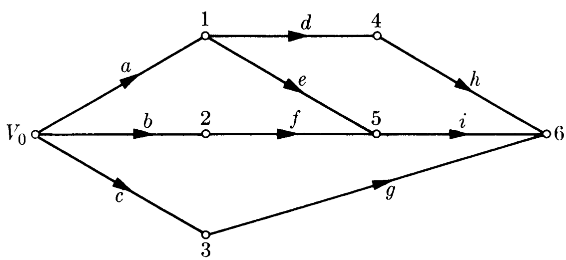
Shelly adopts two devices to overcome this difficulty. Given a logic network, the corresponding directed graph is obtained as follows. The graph contains a *reference* vertex V_0. Corresponding to each input terminal and each output terminal of the logic units is a vertex of the graph, except that if several terminals are connected together they correspond to the same vertex. The reference vertex V_0 is connected to each of the vertices corresponding to the input terminals of the logic network, by edges directed away from V_0. For the others, there is an edge from vertex j to vertex k if and only if j corresponds to the input terminal of a logic unit and k corresponds to the output terminal of the same logic unit. A simple example of a logic network is shown in Fig. 9-19(a), and its directed graph is shown in Fig. 9-19(b).

The vertices and edges of the directed graph are now given appropriate "weights." The weight of the reference vertex V_0 is 1 (the Boolean 1). The weights of the other vertices are the Boolean functions present at the corresponding terminals of the logic networks. The edge weights are defined to be *right operators* in such a fashion that the weight of the vertex at the tip of the edge is the result of the edge weight operating on the weight of the vertex at the tail. For the example of Fig. 9-19(b), if we denote vertex weights by F_0 $(=1)$, $F_1, \ldots, F_6$, we have

$$F_4 = x' = F'_1. \tag{9-71}$$



(a)



(b)

FIG. 9–19. (a) A logic network and (b) its Shelly graph.

Hence the weight of edge d is defined to be $\sim$; thus, if we multiply the weight of vertex 1 by the weight of edge d, we have

$$F_1\sim \; = F_1' = F_4. \tag{9–72}$$

Similarly, the weight of edge f is $\&F_1$, so that

$$F_5 = F_2(\&F_1) = F_2F_1 = yx. \tag{9–73}$$

The edges a, b, and c of Fig. 9–19(b) are *function generators* in Shelly's terminology. Their edge weights are defined such that $F_1 = x$, etc. Thus,

$$(\text{weight of } a) = \cdot x, \quad (\text{weight of } b) = \cdot y, \quad (\text{weight of } c) = \cdot z. \tag{9–74}$$

Hence these weights are also operators of the same kind as the rest. It is seen that the operator appears on the right. It is possible to define addition and multiplication for the operators in such a fashion that they satisfy many of the familiar postulates. Under addition, the operator algebra is closed, associative, and commutative. Under multiplication, the algebra is closed and satisfies the associative law, but is noncommutative. Identity elements exist under both operations, and the usual distributive law (multiplication over addition) holds. Under addition, the elements are idempotent ($a + a = a$).

Much of the work of Hohn and Schissler [76] can be extended to logic networks with the Shelly representation. The Shelly representation is a net, and therefore the connection matrix C is given by Definition 9–8. The entries of C are operators. The diagonal entries are defined to be 1 for convenience. When the Aufenkamp-Hohn algorithm is applied to C, and all the nodes other than the reference node and output nodes are removed, the entries in the positions $(0, m)$ where m is an output node, are the outputs of the logic network. For example, for the net of Fig. 9–19(b), the connection matrix is given by

$$\mathsf{C} = \begin{array}{cc} & \begin{array}{ccccccc} 0 & 6 & 4 & 5 & 1 & 2 & 3 \end{array} \\ \begin{array}{c} 0 \\ 6 \\ 4 \\ 5 \\ 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 & \cdot x & \cdot y & \cdot z \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & +F_5 + F_3 & 1 & 0 & 0 & 0 & 0 \\ 0 & +F_4 + F_3 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \sim & \&F_2 & 1 & 0 & 0 \\ 0 & 0 & 0 & \&F_1 & 0 & 1 & 0 \\ 0 & +F_4 + F_5 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} . \tag{9–75}$$

In practical computations, it is more convenient to begin by removing the vertices corresponding to the input terminals, in this case 1, 2, and 3; otherwise, rather complicated operators result. Whenever the function corresponding to a node is known, as F_1, F_2, and F_3 are in this example, the functions are inserted wherever they occur in the matrix. If we remove nodes 1, 2, and 3, and insert $F_1 = x$, $F_2 = y$, and $F_3 = z$, the result is

$$\mathsf{C}_{(123} = \begin{array}{cc} & \begin{array}{cccc} 0 & 6 & 4 & 5 \end{array} \\ \begin{array}{c} 0 \\ 6 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 1 & z + F_4 + F_5 & x\sim & x\&y \\ 0 & 1 & 0 & 0 \\ 0 & +F_5 + z & 1 & 0 \\ 0 & +F_4 + z & 0 & 1 \end{bmatrix} \end{array} . \tag{9–76}$$

Observe that $x\&y$ appears twice in the (0, 5)-position, which is due to the redundancy of specification in the net itself. The removal of any one input node of a logic unit specifies the output in this representation. F_4 and F_5 are now known from the first row. Writing these in more natural fashion, we have

$$\mathsf{C}_{(123} = \begin{array}{c} \\ 0 \\ 6 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 6 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & z + x' + xy & x' & xy \\ 0 & 1 & 0 & 0 \\ 0 & +xy + z & 1 & 0 \\ 0 & +x' + z & 0 & 1 \end{bmatrix} \end{array}. \tag{9–77}$$

If we remove nodes 5 and 4 in order, we arrive finally at

$$\mathsf{C}_{(123(5(4} = \begin{array}{c} \\ 0 \\ 6 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 6 \end{array} \\ \begin{bmatrix} 1 & x' + xy + z \\ 0 & 1 \end{bmatrix} \end{array}, \tag{9–78}$$

since the terms added to the (0, 6)-position are also $x' + xy + z$. Inspection of Fig. 9–19(a) verifies the conclusion.

Shelly [163] has also discussed the possibility of reversing this procedure to synthesize logic networks for given Boolean functions and a set of admissible connectives. However, the outline above suffices, since the purpose here is merely to suggest the possibility of an application.

Problems

9–1. Prove Theorem 9–1.

9–2. Prove that if a product π is a factor of a product in $V(Y)$ and is not a factor of any product in $W_{1,1'}(Y)$, then the subgraph corresponding to π contains a path from 1 to 1′.

9–3. Prove Theorem 9–4. Hence show that $V + F = F$, $V + W = W$, and $FV = WV = V$.

9–4. Let $F(x_1, x_2, \ldots, x_n)$ be an SC-switching function expressed in normal form (as a sum of products). Let

$$C_i = \{y_1, y_2, \ldots, y_k\}$$

be a *minimal* set such that C_i has an odd number of terms in common with each of the products in F. Show that C_i corresponds to a cut-set separating the input vertices. Obtain a design procedure for SC-functions based on this fact. (See Lund [102].)

9–5. Suppose that we iterate the procedure in Problem 9–3; that is, suppose that we find

$$D_p = \{z_1, z_2, \ldots, z_m\}$$

such that each D_p has an odd number of variables in common with each C_i. What can we say about D_p? (Lund [102].)

9–6. Use the procedure developed in Problem 9–4 to design the function $F = xz + xyw + vyz + vw$.

9–7. Use the simple checks to establish the minimality of the realizations of Fig. 9–20.

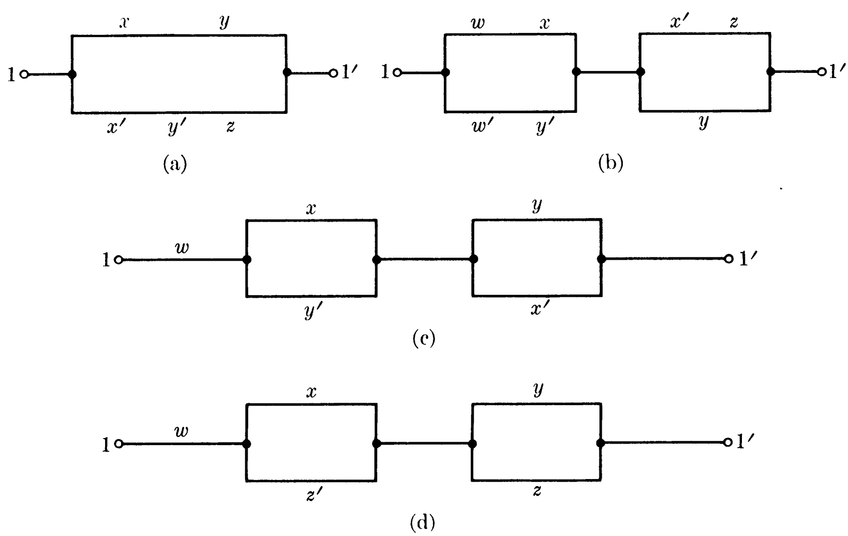


Figure 9–20

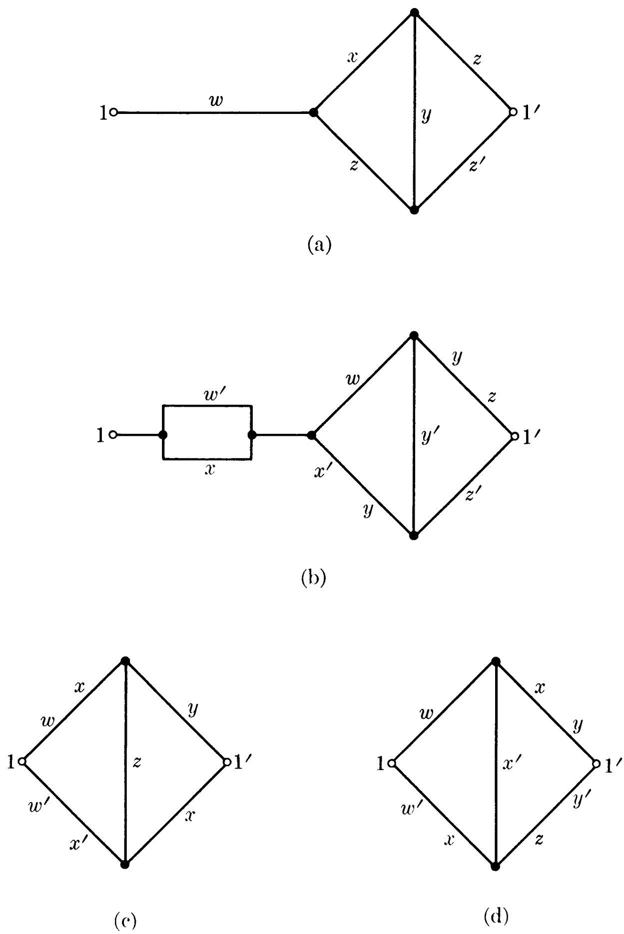


FIGURE 9–21

9–8. Use the matrix technique to establish the minimality of the realizations of Fig. 9–21.

9–9. Use the matrix technique to design minimal realizations of the functions

$$F_1 = wxy + w'x'y'z' + wx'y'z,$$
$$F_2 = wxz + wyz' + w'xy'z',$$
$$F_3 = wxz + wxy + w'x'y.$$

9–10. Prove Theorem 9–11.

9–11. Prove the following interesting analogue of Maxwell's formula. Let a *P-set of cycles* be defined as a set of oriented circuits (in which the edge orientations agree with the circuit orientation) which are vertex-disjoint and which include all the vertices of the net. Then, if we consider multiplication to be commutative, the determinant of the connection matrix C of a net is given by

$$\det \mathsf{C} = \sum (P\text{-set cycle products}),$$

the summation being over all P-sets of the net and the product being the product of edge weights. (See the discussion of Coates' flow graphs, Section 10–2.)

9–12. Use the matrix technique to show that the parity function

$$F_3(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$$

cannot be realized with the set of nine contacts

$$x, y, z, x, y, z, x', y', z'.$$

9–13. Prove Theorem 9–13.

9–14. Design logic networks realizing the same Boolean functions as are realized by the contact networks of Fig. 9–20. Analyze them by Shelly's procedure, thus checking the design.

CHAPTER 10

OTHER APPLICATIONS

This last chapter of the book is devoted to a survey of a few other applications of graph theory, particularly the theory of weighted directed graphs (or nets), which have not been discussed so far. Two of these, communication networks and flow graphs, are of great interest to electrical engineers and consequently are treated in some detail. The applications in social sciences and neural networks are reviewed only very briefly, the purpose being merely to acquaint the reader of the existence of these applications. The applications to axiomatics and the algebra of relations also find brief treatments here. This chapter, like Chapter 9, is based on a few published papers of fundamental importance, which are referred to in the appropriate sections.

10–1 Communication networks. Communication networks are "natural" applications of the theory of graphs in the sense that one thinks of the various stations as being points in a communication network and the channels of communication as being lines drawn between these points. As such, a number of investigators have applied topological ideas to problems in communication theory, either deliberately or incidentally. We consider only the most significant contributions that have been made in this connection, disregarding many obvious ones. Specifically, we consider the works of Prihar [136], Elias, Feinstein, and Shannon [50], and the independent work of Ford and Fulkerson [56].

Prihar's work closely parallels the work of Hohn et al. [77] but is not as complete. Prihar considers the problem of analyzing a communication network which may contain both one-way and two-way communication links. An example of such a system is shown in Fig. 10–1.

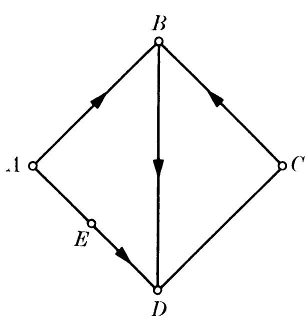


FIG. 10–1. A communication network.

For such a network, Prihar defines several matrices of real numbers, from which various characteristics of the communication system might be derived. Figure 10–1 is seen to be a mixed (directed and nondirected) graph. If we like, we can replace each two-way link by two one-way links, thus obtaining a directed graph. However, there is no reason not to admit such mixed graphs.

The first matrix that Prihar defines for the network is the *relation matrix* $\mathsf{X} = [x_{ij}]$ (which Prihar calls the matrix of *antimetries*), which has one row and column for each vertex and in which

$$\begin{aligned} x_{ij} &= 1 \quad \text{if there is a channel } \textit{from } i \textit{ to } j, \\ x_{ij} &= 0 \quad \text{if there is no channel } \textit{from } i \textit{ to } j. \end{aligned} \tag{10–1}$$

It is evident that X is the same matrix as the transition matrix of Seshu, Miller, and Metze [155] considered in Chapter 9, and we see in Section 10–3 that it is also the relation matrix of the algebra of relations. However, X may have any number of 1's in each row. For the network of Fig. 10–1,

$$\mathsf{X} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}. \tag{10–2}$$

Since there is only one type of relation being considered, X may also be considered to be the connection matrix of the net in the sense of Hohn et al. Thus we should not be surprised at the following results. (See Problem 10–1.)

THEOREM 10–1. The entries of X^n (x_{ij} considered to be real numbers) enumerate the number of paths of length n through the network, both proper and redundant (i.e., intersecting) paths being included.

Prihar fails to recognize the possibility of including intersecting paths.

THEOREM 10–2. A diagonal entry $x_{ii}^{(2)}$ of X^2 is nonzero if and only if there exists a two-way link between point i and another point of the network. More specifically, $x_{ii}^{(2)}$ is the number of points to which point i is connected by a two-way link.

THEOREM 10–3. The trace of the matrix X^2 [that is, $\sum_{i=1}^{n} x_{ii}^{(2)}$] is twice the number of two-way links.

The true connection matrix of a communication network is also defined by Prihar, who extends it to get a pair of *traffic matrices*. The *connection matrix* $\mathsf{T} = [t_{ij}]$ is defined as

$$t_{ij} = (\text{number of links from } i \text{ to } j). \tag{10–3}$$

The *incoming-traffic* matrix is defined as

$$\mathsf{P} = \mathsf{XT}, \tag{10–4}$$

and the *outgoing-traffic* matrix is defined as

$$\mathsf{Q} = \mathsf{TX}. \tag{10–5}$$

The main diagonal entries of P give the number of traffic lines entering a point, and the corresponding entries of Q give the number leaving the point. Prihar interprets the off-diagonal entries, also, as interfering links, but the interpretation is not very satisfying. He also makes use of the matrix X to analyze the minimality of an adequate communication system but fails to give the theory behind it. Therefore, we conclude the discussion of Prihar's contribution with the statement of the following result.

THEOREM 10–4. The minimum number of steps needed for complete communication (i.e., the number of steps in which any station can communicate with any other) is the smallest number m such that

$$\mathsf{A}(m) = \mathsf{X} + \mathsf{X}^2 + \cdots + \mathsf{X}^m \tag{10–6}$$

has no zeros in it.

This result is obvious and has its analogues in sequential circuit theory [155] and several other applications. (See Problem 10–2.)

Elias, Feinstein, and Shannon, as well as Ford, Fulkerson, and Dantzig, independently solved the *maximum network flow problem*, which may be stated as follows. Let a network, which may contain both directed and nondirected edges, be given, and let a real number c_{ij} be associated with each edge. The edge weight may be interpreted in different ways, depending upon the application. In a communication network, this may be the channel capacity in bits per second; in a gas pipeline or oil pipeline, this may be the capacity in cusecs*; in a highway network, this may be cars per hour; etc. In such a network, what is the maximum rate of flow (of whatever is being considered) from a given point A to a point B? In all applications, the flow must satisfy Kirchhoff's law at the nodes [(flow in) = (flow out)]. The solution must consist of three things.

* 1 cusec = 1 ft^3/sec.

First, we must state the upper bound for the flow. Second, we must state the rules for programming the flow to achieve this maximum. Third, the theorem must be translated to an algorithm for practical applicability. The first two parts of the answer are given in the proof of the main result below, which is due to Elias, Feinstein, Shannon, Ford, Fulkerson and Dantzig.

THEOREM 10–5. The maximum flow from point A to point B of a network is the minimum value A to B of all directed cut-sets (A, B), where the value of the cut-set is the sum of the weights of all the edges of the cut-set whose orientations agree with the cut-set orientation.

Proof. Since any flow from A to B must cross any cut-set (A, B) in the direction of the cut-set, it is obvious that the given flow cannot be exceeded. To show that this flow is actually achieved, we must show that the flow can be suitably programmed to achieve the maximum. We follow the procedure of Elias, Feinstein, and Shannon [50]. The procedure consists of constructing a *reduced network* with the following properties:

(a) The graph of the reduced network is the same as the graph of the original network, except possibly that some of the edges of the original network may be missing in the reduced network (i.e., capacity reduced to zero).

(b) The capacity of no edge is increased (but some may be decreased from the original network).

(c) Every edge of the reduced network is in at least one cut-set of value V, where V is the value of the minimal cut-set (A, B) of the original network.

The reduced network is constructed as follows. If any edge of the network is not in a minimum-value cut-set (A, B), reduce its capacity until either it is in a minimum-value cut-set or its capacity goes to zero. This operation cannot reduce any cut-set to below minimum. Repeat with every other such edge (taking them one by one). For a given original network, the reduced network is not unique. If we can now show that the required maximum flow can be achieved in the reduced network, then it can obviously be achieved in the original network, since no capacity is exceeded and the Kirchhoff condition is satisfied. The proof is based on induction on the number of vertices in the graph. If every path from A to B is of length 1 or 2, the network has the form shown typically in Fig. 10–2. In such a network, if each edge carries a flow equal to its capacity from left to right, it is obvious that the desired flow is achieved.

Suppose now that the theorem is true for all reduced networks with less than n vertices. We show that the theorem is true for n vertices. If the network of n vertices contains no paths of length 3 or more, the

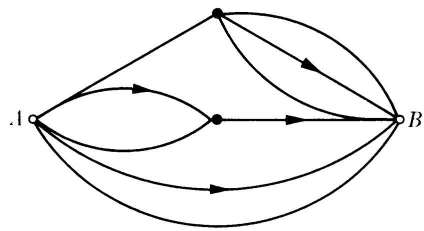


FIG. 10–2. Network of length 2.

theorem is true, by the observation above. If there is a path of length 3 or more, consider one of the edges on the path that has neither vertex A nor vertex B. This edge is in a minimal cut-set C, since the network is reduced. Replace each edge in C by two edges in series, with the same capacity as the original edge. Now identify all of the newly-formed vertices. The network then becomes a series connection of two two-terminal subnetworks. Each subnetwork has the same minimal value as the original, since it contains the cut-set corresponding to C. Each of these two subnetworks contains fewer than n vertices, by the construction. Hence, a flow program is possible in each subnetwork, by induction hypothesis. Where the common vertex is separated to give the original, the same program is seen to be satisfactory.

The theorem can be extended to cover multiterminal networks as well, by addition of channels *from* each output vertex to a common vertex and *to* each input vertex from a common vertex, as shown in Fig. 10–3.

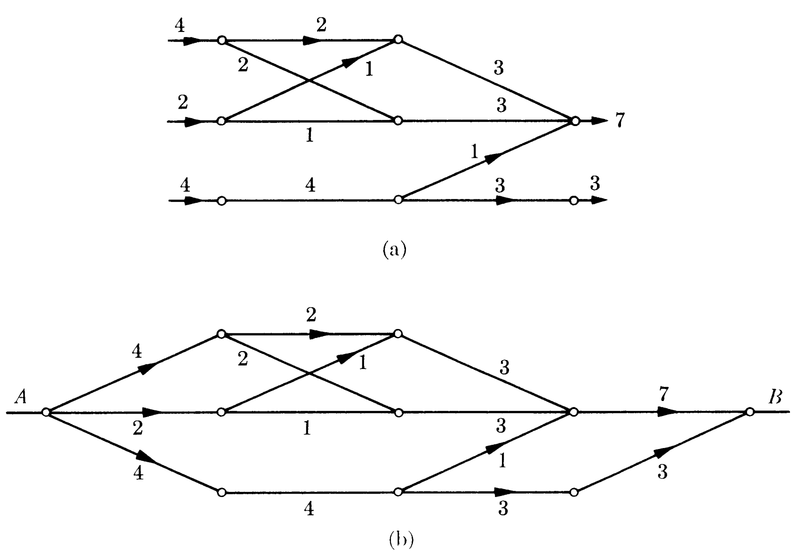


FIG. 10–3. Extension to multiterminal networks.

Although the theoretical problem is thus solved, there is still the practical question of finding the minimal cut-set. Ford and Fulkerson [57] have given a computational algorithm for finding the maximal flow, for which we refer the reader to their original paper. If the network is planar (in the two-terminal sense) we can draw its dual, assign edge weights equal to those of the corresponding edges in the original, and use Moore's [119] technique for finding the shortest path through a maze. Since this shortest path corresponds to a minimal cut-set, the problem is solved.

Another obvious application of topology to communication theory is the concept of a *probabilistic net*, i.e., a directed graph in which the weights of the edges are probabilities. Any Markov process has an obvious interpretation as a probabilistic net. For example, see Shannon [160], where a discrete source is represented as a directed graph. Although a number of workers have talked about such probabilistic nets (see several papers in *Proceedings of the Brooklyn Symposium on Information Networks*, 1954), very little concrete work has been done about them. Therefore, we list the topic as a possible avenue of exploration and close this section.

10–2 Flow graphs and signal-flow graphs. A method of solving a system of linear algebraic equations by the use of nets was first described in a paper by Mason [105] in 1953. Since then, an alternative representation of equations as a net has been described by Coates [37]. The two formulations are very closely related both in the nets that result and in the manipulations or topological formulas, as the case may be. The representation due to Mason is referred to as a *signal-flow graph.* In these pages, a signal-flow graph is also referred to as a *Mason graph,* as distinguished from the *Coates graph.* Our discussion here is mainly concerned with the application of the theory of graphs to Mason graphs and Coates graphs, and not with the application of these graphs to feedback systems. Many examples of the applications of signal-flow graphs are to be found in Truxal [179]. In this section, a general familiarity with the use of signal-flow graphs is assumed. (See, for instance, Truxal.) The emphasis is rather on the justification of the procedure and the relationship to the general theory of nets, in particular the Hohn-Aufenkamp "state-removal" algorithm given in Section 9–2. Signal-flow graphs of Mason are taken up first in the following discussion.

In applications, signal-flow graphs are normally drawn by inspection. However, to keep the discussion general and to prove the validity of the procedure, a system of linear algebraic equations is assumed here. Let

$$\mathbf{FX} = \mathbf{Y} \qquad \text{and} \qquad \mathbf{F} = [f_{ij}] \tag{10–7}$$

be a system of n equations in n unknowns, which is consistent and linearly independent. Thus,

$$\Delta = \det \mathsf{F} \neq 0. \tag{10–8}$$

[A homogeneous system with nontrivial solutions can be brought to the form of Eq. (10–7).] The elements of F and Y are assumed to belong to a field—real or complex numbers or analytic functions in most applications. It is assumed that the diagonal entries in F are nonzero. It is obvious from the general expansion formula [78],

$$\det \mathsf{F} = \sum_j \epsilon_{j_1 j_2 \cdots j_n} f_{1j_1} f_{2j_2} \cdots f_{nj_n}, \tag{10–9}$$

that at least one product $f_{1j_1} f_{2j_2} \cdots f_{nj_n}$ is nonzero; so, by permutation of rows of F (and Y to keep the system unaltered), the diagonal entries may be made nonzero. Thus the assumption is no restriction. To derive the Mason graph, rewrite Eq. (10–7) as

$$[(\mathsf{F} + \mathsf{U}) \quad -\mathsf{U}] \begin{bmatrix} \mathsf{X} \\ \mathsf{Y} \end{bmatrix} = \mathsf{X}. \tag{10–10}$$

However, in applications the entries in Y are frequently expressible in terms of a single function, the input, as

$$\mathsf{Y} = \mathsf{K} y_0, \tag{10–11}$$

where K is a function of the system. Then Eq. (10–10) is written as

$$[\mathsf{F} + \mathsf{U} \quad -\mathsf{K}] \begin{bmatrix} \mathsf{X} \\ y_0 \end{bmatrix} = \mathsf{X}. \tag{10–12}$$

The *signal-flow graph* corresponding to the system of equations

$$\mathsf{F}\mathsf{X} = \mathsf{K} y_0 \tag{10–13}$$

is a net with the connection matrix

$$\mathsf{C} = \begin{bmatrix} \mathsf{F} + \mathsf{U} & -\mathsf{K} \\ \mathbf{0} & 0 \end{bmatrix}' = \begin{bmatrix} \mathsf{F}' + \mathsf{U} & \mathbf{0} \\ -\mathsf{K}' & 0 \end{bmatrix} \tag{10–14}$$

(where $\mathbf{0}$ stands for a row of zeros in the middle step and a column of zeros in the extreme right of the equation), the vertex weights being (in order) $x_1, x_2, \ldots, x_n, y_0$.

The transpose notation is used here to be consistent with Chapter 9. (Coates [37] prefers to interpret c_{ij} as the weight of the edge from j to i.) The edge weights c_{ij} are referred to as *transmissions*. The last vertex y_0 is the *source node* because in applications it corresponds to a source, all other nodes being *internal nodes*.

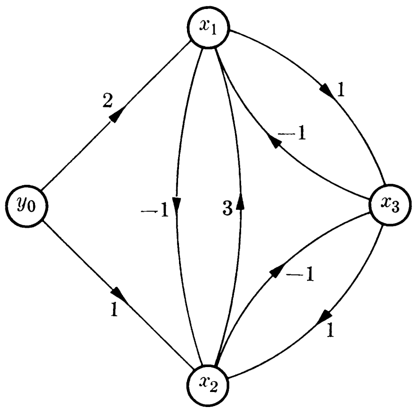


FIG. 10–4. Example of a signal-flow graph.

If there is more than one source, the column **K** is replaced by a matrix **K** with as many columns as there are sources. The basic theory remains the same, however. In the interest of simplifying the notation, a single source node is assumed here. Since each nonzero column of **C** corresponds to an equation of the system (10–12), each internal node x_j corresponds to an equation

$$\sum_{i=1}^{n} c_{ij}x_i + c_{0j}y_0 = x_j, \qquad j = 1, 2, \ldots, n. \tag{10–15}$$

Each nonzero diagonal entry in **C** corresponds to a self-loop. Since no diagonal entry of **F** is nonzero, none of the self-loops has weight 1. Usually, in the applications, the signal-flow graph is chosen to make $c_{ii} = 0$, $i = 1, 2, \ldots, n$. For example, the signal-flow graph of Fig. 10–4 stands for the system of equations

$$\begin{aligned} x_1 &= 2y_0 + 3x_2 - x_3, \\ x_2 &= y_0 - x_1 + x_3, \\ x_3 &= x_1 - x_2. \end{aligned} \tag{10–16}$$

The connection matrix for this flow graph is

$$\mathbf{C} = \begin{array}{c} \\ y_0 \\ x_1 \\ x_2 \\ x_3 \end{array}\!\!\begin{array}{c} \begin{array}{cccc} y_0 & x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 3 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \end{array} \tag{10–17}$$

(y_0 is made the first node for convenience in elimination of nodes).

The manipulations of the flow graph correspond to systematic eliminations of variables in the system of equations. Thus, implicit in the procedure is the assumption that the solution is required for only one variable x_k and not for the others. This corresponds closely to practical situations and is thus quite reasonable. In fact, in Mason's original theory the variable x_k was isolated by stipulating that no edge leave vertex x_k (the vertex is referred to by its weight to simplify notation), and the vertex was named a *sink node*. By introducing an auxiliary variable x'_k and a trivial equation $x_k = x'_k$, it is possible to ensure the existence of a sink node. However, such an assumption is unnecessary and so is not made here.

The algebraic elimination of variables corresponds to the "node-removal" process in the graph. Consider first the elimination of a variable x_i for which $c_{ii} = 0$, so that there is no self-loop at node x_i of the graph. The equation for x_i is

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ji} x_j + c_{0i} y_0. \tag{10–18}$$

Elimination of x_i consists merely of substituting this expression into all the other equations. For instance, the equation for x_k becomes ($k \neq i$)

$$\begin{aligned} x_k &= \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{jk} x_j + c_{0k} y_0 + c_{ik} x_i \\ &= \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{jk} x_j + c_{0k} y_0 + \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ji} c_{ik} x_j + c_{0i} c_{ik} y_0 \end{aligned} \tag{10–19}$$

or

$$x_k = \sum_{\substack{j=1 \\ j \neq i}}^{n} (c_{jk} + c_{ji} c_{ik}) x_j + (c_{0k} + c_{0i} c_{ik}) y_0, \tag{10–20}$$

where $k = 1, 2, \ldots, i-1, i+1, \ldots, n$.

On comparing Eq. (10–20) with Eq. (9–57), we see that the graph interpretation of the terms $c_{jk} + c_{ji}c_{ik}$ and $c_{0k} + c_{0i}c_{ik}$ is evident. The transmissions of the paths through the vertex i have been added to the direct transmission. Since the weights belong to a field, the transpose in Eq. (10–14) is immaterial and the "state-removal" algorithm can be applied directly, leading to the same result as in Eq. (10–20).

Next, consider the case in which the diagonal entry $c_{ii} \neq 0$. Then the equation for x_i becomes

$$x_i = c_{ii} x_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ji} x_j + c_{0i} y_0 \tag{10–21a}$$

or

$$(1 - c_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} c_{ji}x_j + c_{0i}y_0. \tag{10–21b}$$

If $c_{ii} = 1$, evidently the procedure breaks down, as Eq. (10–21b) must be solved for x_i and the result substituted in other equations. By the initial arrangement of equations, $f_{ii} \neq 0$, so $c_{ii} \neq 1$ when the signal-flow graph is originally drawn. However, there is no guarantee that one of the self-loops cannot become 1 somewhere in the reduction process. This circumstance demands an interchange of equations before the solution can be completed. The details are left as a problem (Problem 10–10). If $c_{ii} \neq 1$, solve Eq. (10–21b) for x_i:

$$x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{c_{ji}}{1 - c_{ii}} x_j + \frac{c_{0i}}{1 - c_{ii}} y_0 \tag{10–22}$$

Equation (10–22) is interpreted in the signal-flow graph as the operation of "removal of self-loop at x_i." Thus, to remove a self-loop, divide all the *incoming transmissions* by $1 - c_{ii}$, and then remove the self-loop c_{ii}. Now the node-removal algorithm can be applied again. When all internal nodes, and any self-loops remaining, have been removed, the graph is reduced to the form shown in Fig. 10–5, from which the solution for the desired variable x_k can be written as

$$x_k = gy_0. \tag{10–23}$$

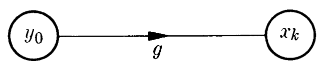


FIG. 10–5. Final graph.

The function g is referred to as the *graph gain*.

In a purely formal sense, the self-loop removal is analogous to the state-diagram procedure (Section 9–2) of multiplying the weights of all the incoming edges by $1 + \sum_k B^k$ since formally

$$1 + \sum_{k=1}^{\infty} B^k = \frac{1}{1 - B}. \tag{10–24}$$

The discussion of the flow-graph manipulation is completed with an example. The system of equations (10–16) represented by the signal-flow graph of Fig. 10–4 is now solved for x_1. The connection-matrix reduction is outlined by steps and the graph reduction is shown in Fig. 10–6, corresponding steps being identified by the same symbol.

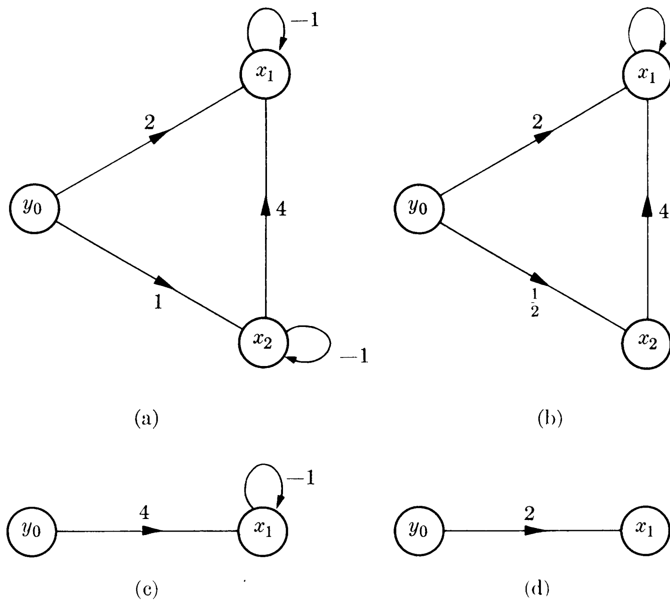


FIG. 10–6. Reduction of signal-flow graph.

(a) Remove node x_3:

$$\begin{array}{c} y_0 \\ x_1 \\ x_2 \\ \vdots \\ x_3 \end{array} \begin{bmatrix} 0 & 2 & 1 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 1 \\ 0 & 3 & 0 & \cdots & -1 \\ \vdots & & & & \\ 0 & -1 & 1 & \cdots & 0 \end{bmatrix} \sim \begin{array}{c} y_0 \\ x_1 \\ x_2 \end{array} \begin{bmatrix} 0 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 4 & -1 \end{bmatrix}.$$

(b) Remove the self-loop at x_2 by multiplying the last column (incoming transmissions) by $1/[1 - (-1)] = \frac{1}{2}$ and setting the (3, 3)-element equal to zero:

$$\begin{array}{c} y_0 \\ x_1 \\ x_2 \end{array} \begin{bmatrix} 0 & 2 & \frac{1}{2} \\ 0 & -1 & 0 \\ 0 & 4 & 0 \end{bmatrix}.$$

(c) Remove node x_2:

$$\begin{array}{c} y_0 \\ x_1 \\ \vdots \\ x_2 \end{array} \begin{bmatrix} 0 & 2 & \cdots & \frac{1}{2} \\ 0 & -1 & \cdots & 0 \\ \vdots & & & \\ 0 & 4 & \cdots & 0 \end{bmatrix} \sim \begin{array}{c} y_0 \\ x_1 \end{array} \begin{bmatrix} 0 & 4 \\ 0 & -1 \end{bmatrix}.$$

(d) Remove the self-loop at x_1 by multiplying the last column by $1/[1 - (-1)] = \frac{1}{2}$ and setting the (2, 2)-element equal to zero:

$$\begin{array}{c} y_0 \\ x_1 \end{array} \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}.$$

The solution is therefore

$$x_1 = 2y_0. \tag{10–25}$$

It is clear that the operations can be performed on either the connection matrix or the graph, as desired.

Mason [106] has given a topological formula for obtaining the graph gain by inspection of the original graph. Let y_0 be the only source node, and let x_1 be the variable for which the solution is desired. A directed path from node y_0 to node x_1 (in which all edge orientations agree with the path orientation) is referred to as a *forward path.* A directed circuit (a circuit in which edge orientations agree with the circuit orientation) is referred to as a *feedback loop.* Mason's formula is the following:

$$g = \frac{1}{\Delta} \sum_{\substack{\text{all} \\ \text{forward} \\ \text{paths}}} g_k \, \Delta_k, \tag{10–26}$$

where g_k is the transmission (product of edge weights) for the kth forward path, and where

$$\Delta = 1 - \sum_m P_{m1} + \sum_m P_{m2} - \sum_m P_{m3} + \cdots,$$

where P_{m1} is the loop transmission (product of edge weights) of the mth feedback loop, P_{mr} is the product of loop transmissions for the mth set of r vertex-disjoint feedback loops, and Δ_k is the value of Δ for the part of the graph having no vertices in common with the kth forward path.

The original argument of Mason [106] is heuristic. The formal proof becomes involved because of the modification of the equations, and is left as an unsolved problem. (For examples, see Mason [106] or Seshu and Balabanian [156].)

The Coates graph is a more natural representation of the system of equations in the sense that no modification is required. Consequently, the topological formulas are derivable by methods analogous to those of Chapter 7. Begin again with the system of equations

$$\mathbf{FX} = \mathbf{Y} = \mathbf{K}y_0, \tag{10–27}$$

with det $\mathbf{F} \neq 0$. As before, let the equations be rearranged to make the diagonal entries in F nonzero. Coates [37] rewrites Eq. (10–27) as

$$[\mathsf{F} \quad -\mathsf{K}]\begin{bmatrix} \mathsf{X} \\ y_0 \end{bmatrix} = \mathbf{0}. \tag{10–28}$$

Now the *Coates graph* of the system (10–28) is the net with the connection matrix

$$\mathsf{C} = \begin{bmatrix} \mathsf{F} & -\mathsf{K} \\ 0 & 0 \end{bmatrix}' = \begin{bmatrix} \mathsf{F}' & 0 \\ -\mathsf{K}' & 0 \end{bmatrix}, \tag{10–29}$$

where the zero column added to C is for the purpose of making it square. [Coates does not use the transpose in Eq. (10–29), as remarked earlier.] The weights of the vertices are once again $x_1, x_2, \ldots, x_n, y_0$ in that order, and the vertices are here referred to by their weights. However, the equation for vertex x_k is now

$$c_{kk}x_k + \sum_{\substack{j=1 \\ j\neq k}}^{n} c_{jk}x_j + c_{0k}y_0 = 0, \qquad \text{where} \quad \mathsf{C} = [c_{ij}]. \tag{10–30}$$

A slight modification of the graph-reduction procedure and the Hohn-Aufenkamp procedure can now be applied to the Coates graph as outlined below. The details are left as a sequence of problems (Problems 10–6 through 10–9).

Divide all incoming transmissions at node x_k by $-c_{kk}$ before pulling (removing) node x_k. If there is no self-loop at node x_k and it is desired to remove node x_k, interchange the labels of x_k and x_j, where $c_{jk} \neq 0$, and modify the (jk)-, (k, j)- and (jj)-transmissions to keep the equations unaltered (Problem 10–8). Now remove the new node x_k.

Since the new equations are closely related to the original system of equations, it is easiest to establish the topological formula for the solution. Since the relation between the connection matrix of a net and the positive and negative incidence matrices is known (Theorem 9–11), procedures analogous to those of Chapter 7 can be employed (with the use of the Binet-Cauchy theorem). For topological formulas derived on this basis, we refer the reader to Coates [37]. A simpler formula is developed here, based on Problem 9–11.

Given the system of equations

$$\mathsf{FX} = \mathsf{Y} = \mathsf{K}y_0, \qquad \text{with } \det \mathsf{F} = \Delta \neq 0, \tag{10–31}$$

the solution for a variable x_i is given by

$$x_i = \frac{1}{\Delta}\sum_{j=1}^{n} (\Delta_{ji}k_j)y_0, \qquad \text{where} \quad \mathsf{K} = [k_j]. \tag{10–32}$$

As in Cramer's rule, the sum $\sum_{j=1}^{n} \Delta_{ji} k_j$ is the determinant of the matrix F_k obtained by replacing column i of F by column K. Referring back to Eq. (10–29), we find that F' is the connection matrix of the graph obtained by deleting vertex y_0 and all edges leaving y_0 (there is no edge entering y_0). The matrix F'_K can be interpreted as the connection matrix of the following graph. Delete all edges *leaving* x_i. Now insert an edge from x_i to x_j if there is originally an edge y_0 to x_j, and give it the weight $-c_{0j}$. Repeat for every edge leaving y_0. Finally, remove y_0 and all edges leaving y_0. (Essentially, remove all edges leaving x_i and identify vertices y_0 and x_i, changing signs of the weights of edges leaving y_0.) Thus, since $\det \mathsf{F} = \det \mathsf{F}'$, it suffices to find a topological formula for the determinant of a connection matrix.

Given a net N of n vertices, let C be the connection matrix of N and let $\det \mathsf{C} \neq 0$. (Thus N has neither sources nor sinks.) The elements c_{ij} of C are assumed to be real or complex numbers or rational functions of a complex variable (in general, elements from a field). From the general expansion formula for a determinant [78], we have

$$\det \mathsf{C} = \sum_{j} \epsilon_{j_1 j_2 \cdots j_n} c_{1j_1} c_{2j_2} \cdots c_{nj_n}, \tag{10–33}$$

where $j_1, j_2, \ldots, j_n$ is a permutation of $1, 2, \ldots, n$, the sum is over all such permutations, and $\epsilon_{j_1 j_2 \cdots j_n}$ is 1 for an even permutation and -1 for an odd permutation. It clearly suffices to locate the nonzero terms in the sum. Let

$$c_{1j_1} c_{2j_2} \cdots c_{nj_n} \neq 0 \tag{10–34}$$

for some permutation. Consider the subgraph N_s of N consisting of edges $c_{1j_1}, c_{2j_2}, \ldots, c_{nj_n}$. Since each integer k appears exactly twice, once as a first subscript and once as a second subscript, it is clear that each vertex k of the subgraph N_s is incident to exactly two edges of the subgraph, one of which (c_{kj_k}) is oriented away from k and the other $(c_{ij_i}, j_i = k)$ is oriented toward k; except when $j_k = k$, in which case the two edges degenerate into a self-loop at k.

Since each vertex of N_s is of degree 2, N_s is a circuit or a vertex-disjoint union of circuits (Problem 2–4), some of which may be self-loops. Each of these circuits is a *directed circuit* by the observation regarding the orientations of edges. Also, N_s includes all vertices of N. A set of vertex-disjoint unions of directed circuits containing all vertices of N was named a *P-set of cycles* in Problem 9–11. *Thus, nonzero terms in* $\Delta = \det \mathsf{C}$ *correspond one-to-one to P-sets of cycles of N.* The converse of this statement is evident. The sign of $\epsilon_{j_1 j_2 \cdots j_n}$ can be computed by a scheme similar to the sign permutation of Section 7–4.

For each P-set of cycles, set up a $(2 \times n)$-array as follows. In the first row, list the vertices in natural order, $1, 2, \ldots, n$. In the second row below each vertex i, list the vertex that *follows* i in the directed circuit containing i. Now count the number of interchanges required to make the second row $1, 2, \ldots, n$. If this number is even, $\epsilon_j = 1$; if it is odd, $\epsilon_j = -1$. This ϵ_j is referred to as the *sign coefficient* of the P-set of cycles. The product of the edge weights of a P-set of cycles is referred to as a *P-set cycle product.* This leads us to our next theorem.

THEOREM 10–6. If $\mathbf{C}$ is the connection matrix of a net, then

$$\Delta = \det \mathbf{C} = \sum_j \epsilon_j \,(P\text{-set cycle product of set } j), \tag{10–35}$$

where the notation is that given above and the sum is over all P-sets of cycles.

On removing edges leaving x_i, identifying vertices x_i and y_0, changing signs of c_{0j}, and recomputing the determinant Δ_i, we find the solution for x_i to be

$$x_i = \frac{\Delta_i}{\Delta}\, y_0. \tag{10–36}$$

The analysis is thus complete.

As an example, consider the system of equations

$$\begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -y_0 \\ 0 \\ 0 \end{bmatrix}, \tag{10–37}$$

for which the flow graph is given in Fig. 10–7 (cf. Coates' [37] Fig. 16b). To evaluate Δ, delete y_0 and the edge (y_0x_1). There are only three types of P-sets of cycles possible: three self-loops, one self-loop, and a loop

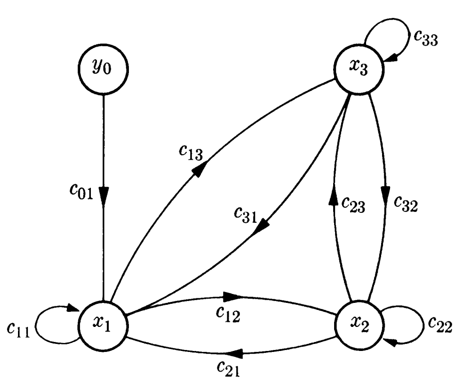


FIG. 10–7. Coates graph of Eq. (10–37).

TABLE 10–1

THE SIX POSSIBLE P-SETS OF CYCLES FOR FIG. 10–7.

P-set	Permutation	Coefficient
c_{11}, c_{22}, c_{33}	(1, 2, 3)	1
c_{11}, c_{23}, c_{32}	(1, 3, 2)	−1
c_{22}, c_{13}, c_{31}	(3, 2, 1)	−1
c_{33}, c_{12}, c_{21}	(2, 1, 3)	−1
c_{12}, c_{23}, c_{31}	(2, 3, 1)	1
c_{13}, c_{32}, c_{21}	(3, 1, 2)	1

containing the other two vertices or a single loop of all three vertices. Therefore, the six sets are as shown in Table 10–1. Therefore,

$$\Delta = c_{11}c_{22}c_{33} - c_{11}c_{23}c_{32} - c_{22}c_{13}c_{31} - c_{33}c_{12}c_{21} + c_{12}c_{23}c_{31} + c_{13}c_{32}c_{21}. \tag{10–38}$$

The example is poor, since C is completely filled and so the expression is no different from the usual expansion of det C. Only the second rows of the permutations are shown.

Suppose that the solution for x_1 is desired. Remove all edges leaving x_1 (including the self-loop c_{11}), and identify y_0 and x_1; the (y_0x_1)-edge becomes a self-loop in the process. Change signs of all edges leaving x_i. The result is shown in Fig. 10–8. The sets for this figure are listed in Table 10–2. Hence,

$$\Delta_1 = (-c_{01})c_{22}c_{33} - (-c_{01})c_{23}c_{32} = -c_{22}c_{33} + c_{23}c_{32} \tag{10–39a}$$

since

$$c_{01} = 1. \tag{10–39b}$$

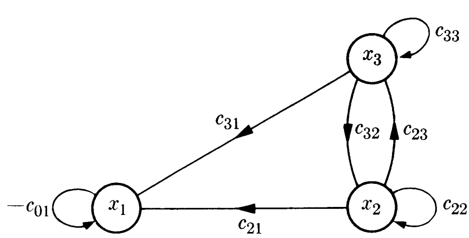


FIG. 10–8. Modified graph of Fig. 10–7.

TABLE 10–2

P-SETS OF CYCLES FOR FIG. 10–8.

P-set	Permutation	Coefficient
c_{01}, c_{22}, c_{33}	(1, 2, 3)	1
c_{01}, c_{23}, c_{32}	(1, 3, 2)	−1

For another example, in which the formula results in some reduction of effort, consider the system of equations

$$\begin{bmatrix} c_{11} & c_{21} & c_{31} \\ 0 & c_{22} & 0 \\ c_{13} & c_{23} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -c_{01} \\ -c_{02} \\ 0 \end{bmatrix} y_0. \tag{10–40}$$

Note that $c_{33} = 0$. The assumption $c_{33} \neq 0$ is necessary only if the graph-reduction scheme is used. The Coates flow graph is shown in Fig. 10–9 (cf. Coates' [37] Fig. 9). Table 10–3 is the table for Δ (after deleting y_0, c_{01}, and c_{02}). Hence,

$$\Delta = -c_{22}c_{13}c_{31}. \tag{10–41}$$

If the solution to x_2 is required, delete all edges leaving x_2, identify y_0 and x_2, and change signs, which leads to Fig. 10–10. Again there is only one P-set, $(-c_{02}, c_{13}, c_{31})$, with the coefficient -1. Hence,

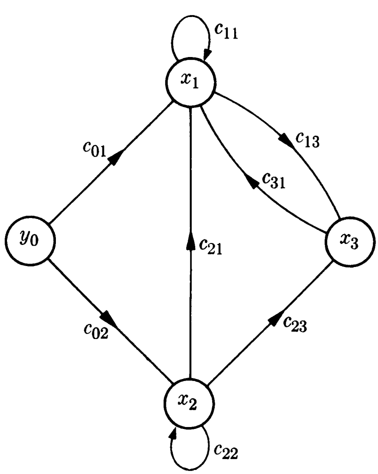


FIG. 10–9. Coates flow graph of Eq. (10–40).

TABLE 10–3

P-SETS OF CYCLES FOR FIG. 10–9.

P-set	Permutation	Coefficient
c_{22}, c_{13}, c_{31}	(3, 2, 1)	-1

and

$$\Delta_2 = +c_{02}c_{13}c_{31},$$

$$x_2 = -\frac{c_{02}c_{13}c_{31}}{c_{22}c_{13}c_{31}}\, y_0 = -\frac{c_{02}}{c_{22}}\, y_0. \tag{10–42}$$

These examples suggest the following alternate method for computing the coefficient ϵ_j, which result can be established rigorously (Problem 10–14).

THEOREM 10–7. If the circuits $C_1, C_2, \ldots, C_k$ of P-set number j consist of $r_1, r_2, \ldots, r_k$ edges, the coefficient ϵ_j of Theorem 10–6 is given by

$$\epsilon_j = (-1)^{r_1+r_2+\cdots+r_k-k}. \tag{10–43}$$

Since a P-set of cycles is a graph in which the number of edges is equal to the number of vertices, Theorem 10–7 has two corollaries.

COROLLARY 10–7(a). If a P-set of cycles consists of k circuits, the coefficient is

$$\epsilon_j = (-1)^{v-k}, \tag{10–44}$$

where v is the number of vertices in the net.

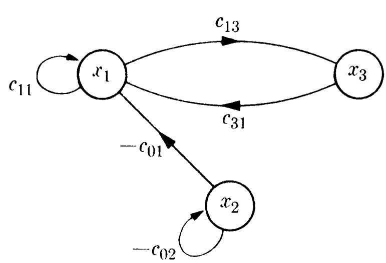


FIG. 10–10. Modified graph of Fig. 10–9.

COROLLARY 10–7(b). If C is the connection matrix of a net of v vertices, then

$$\Delta = \det \mathsf{C} = (-1)^v \sum_j (-1)^{k_j} \ (P\text{-set cycle product } j), \tag{10–45}$$

where k_j is the number of circuits in P-set j and the sum is over all P-sets of cycles.

The result is to be compared with Theorem 2 of Coates [37].

10–3 Calculus of binary relations. This section is a graph-theoretic adaptation of a matrix treatment (due to Copi [38]) of the calculus of binary relations. We, however, avoid the intricacies of logic considered by Copi, since they are not the primary objectives.

Consider a group of objects 1, 2, ..., n, where n is a finite number. The group might, for instance, consist of n persons. Consider a set of binary relations defined on this set. If R is a binary relation, then the statement $i\,R\,j$ is true if i has the relation R to j. In case the set consists of persons, examples of relations are

i is the son of j,
j is the wife of k,
j is the husband of k,
etc.

We now define the *relation matrix* T^R as follows:

$$\mathsf{T}^R = [r_{ij}]_{n,n},$$
$$r_{ij} = 1 \quad \text{if } i\,R\,j \text{ is true}, \qquad r_{ij} = 0 \quad \text{if } i\,R\,j \text{ is false}. \tag{10–46}$$

Evidently there is one such matrix for each relation. It follows that this relation matrix is the same as the *transition matrix* defined in Chapter 9. Therefore, it also follows that we may represent the system consisting of the n objects and the binary relations defined on them by means of a weighted directed graph or a *net*, with a connection matrix $\sum_R R\mathsf{T}^R$. This net contains one vertex for each object. If the object i bears the relation R to object j, that is, if $i\,R\,j$ is true, there is a directed edge from vertex i to vertex j with a weight R.

The structure of the set can be obtained directly by observation of the net. We can also state theorems in terms of the relation matrices analogous to those stated for transition matrices in Chapter 9. For instance, suppose that we are interested in the *relative product* of two relations. The relative product is defined as follows. If R and S are two relations, then $i\,RS\,k$

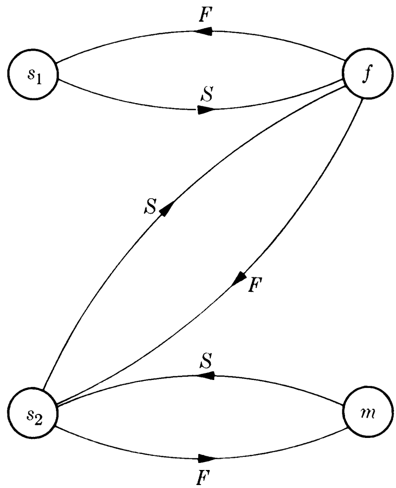


FIG. 10–11. Net for riddle.

is true if for some j, $i\ R\ j$ and $j\ S\ k$ are true. For example, a is the paternal grandmother of b if, for some c, a is the mother of c and c is the father of b. Here the relation "paternal grandmother" is the relative product of the relations "mother of" and "father of." It follows immediately that if we consider the elements of the relation matrices as Boolean ($1 \cdot 1 = 1 + 1 = 1, 0 \cdot 0 = 0 + 0 = 0, 1 + 0 = 0 + 1 = 1$), the relation matrix for the relative product RS is simply $\mathsf{T}^R\mathsf{T}^S$. In terms of the net, $i\ RS\ j$ is true if and only if there is a path of length 2 from i to j in which the first edge has weight R and the second has weight S. Thus, many of the results in the theory of binary relations have interpretations in the theory of nets, and vice versa. We leave it to the reader to look up the details of the calculus of binary relations in Copi [38] and references therein, and conclude this discussion with an example.

As an example of a net, let us construct the net for the familiar riddle "Brothers and sisters I have none, but that man's father is my father's son." Consider four persons, s_1 the speaker, m the second man, s_2 the father of m, and f the father of s_1. There are two relations, "son of" and "father of." Construct the partial net that can be constructed from the given data, as in Fig. 10–11. Immediately, the paths SF from s_1 to s_2 and vice versa are noted. Thus s_1 is the son of the father of s_2, and vice versa. However, since s_1 has no brothers, $s_1\ SF\ s_2$ is the same statement as $s_1 = s_2$, which solves the riddle.

10–4 Logic: axiomatics. In this section, we present a very brief review of the work of Herz [75], who applied the theory of directed graphs to logics consisting of simple theorems. A detailed treatment of Herz's contribution is to be found in König [88].

The system considered here is a collection of statements in which the binary relation "implies" is defined such that this relation is transitive. Thus if $\{A, B, C, \dots\}$ is the collection, whenever $A \to B$ and $B \to C$, also $A \to C$. The theorems of the logic are those statements $i \to j$ which are true. Only such simple theorems as have one hypothesis and one conclusion are admitted. Theorems with compound statements such as

$$A \;\&\; (B \text{ or } C) \to D \text{ or } F$$

are not simple theorems and are not considered in this application.

Such a logical system can be represented as a directed graph. The vertices of the graph correspond to the statements, and the edges correspond to the theorems. If $A \to B$ is a theorem, draw a directed edge from vertex A to vertex B. The graph will reflect the transitivity of the relation $\to$, and in fact we call the graph itself a *transitive graph*. The problem considered by Herz was to choose a set of axioms for the logical system. The set of axioms would consist of some of the theorems of the system with the properties that (a) all theorems can be derived from the set of axioms and (b) no axiom is derivable from the others. Herz [75] gave this problem a geometrical interpretation, as follows.

A collection B of edges of a directed graph G is an *edge basis* of G

(a) if for every edge PQ of G, not in B, there exists an oriented path in B from P to Q, with all edge orientations agreeing with the path orientation, and

(b) if PQ is any edge of B, there is no other oriented path in B from P to Q, with all edge orientations agreeing with the path orientation.

The name *oriented path* (German *Bahn*) is always used in this sense (i.e., all edge orientations agree with the path orientation). In this application, parallel edges with the same orientation evidently have no significance. If in a graph G, for any two vertices P and Q, there is at most one edge (PQ or QP) between them, we say that the graph consists of *simple edges*. The most interesting of Herz's theorems is the following uniqueness theorem.

THEOREM 10–8. A finite directed transitive graph with simple edges has a unique edge basis.

Proof. Let B and B' be two different edge bases of the graph. Then there exists at least one edge in B, not contained in B', say PQ. Since B' is a basis, we can find an oriented path $\overrightarrow{P \cdots Q}$ consisting only of edges of B'. Not all the edges of this oriented path can belong to B, and the path contains at least two edges. Let us replace each of the edges of this path

$\overrightarrow{P \cdots Q}$ (that does not belong to B) with an oriented path in B. Then we have an oriented edge sequence

$$PR_1,\ R_1R_2, \ldots, R_nQ$$

consisting only of edges of B. This sequence also must contain at least two edges, one of which has to be PQ since B is a basis. Thus one of the internal vertices is either P or Q. If $P = R_m$, $m \neq 1$, then from the transitivity of the graph and existence of the path

$$\overrightarrow{R_1R_2 \cdots P},$$

it follows that G contains an edge R_1P. Thus G contains PR_1 and R_1P, contradicting the assumption of simple edges. The case $Q = R_m$ is similar.

Herz [75] gives a method of reducing a given graph to one in which each component is a transitive graph with simple edges, thus solving the problem of choosing axioms for a system of simple theorems. For the details, we refer the reader to Herz [75] or König [88]. König also considers the analogous concept of a *vertex basis* for a graph.

10-5 Brief survey of other applications. The general concept of a net finds applications in many fields besides the ones mentioned here. Since it is not possible to develop the background material here, these applications are mentioned only briefly. The main purpose is to acquaint the reader with the existence of the various applications and provide references wherein detailed treatments are available.

In an early paper, Cayley [26] represented an abstract (mathematical) group by a net. The vertices of the net are in one-to-one correspondence with the elements $(x_1, x_2, \ldots, x_k)$ of the group, and each vertex is weighted by the corresponding element. Every pair of vertices is joined by two directed edges, one each way. The edge (x_i, x_j) is weighted by x_k if $x_i \cdot x_j = x_k$ in the group. The net obtained is the *Cayley group diagram*, and many properties of the group can be given simple graph interpretations. Details are to be found in Cayley [26] or König [88].

Another relationship between certain groups of symmetries and graphs was given by Pólya [134] in a classic paper. Pólya's formulation has applications to group theory, theory of isomers (in chemistry), and to Boolean function theory. The problem in Boolean functions is to count the number of *symmetry types* of Boolean functions of n variables. Two Boolean functions are of the same symmetry type if one can be transformed into the other by either permutation or complementation of some variables, or by both. Since any two functions of the same symmetry type have the same contact realization, the problem is of interest in the theory of switching. A Boolean function of n variables can be represented on an n-cube by marking (in some fashion) the vertices where the function is 1. Then the problem of finding the number of functions of this symmetry type reduces to that of computing the number of mappings of the graph of the n-cube onto itself (automorphisms) that retain the structure of the marked vertices. The problem was solved by Slepian [166]. Other similar enumeration problems have been treated by Harary [73], Riordan [148], Gilbert [63], and others.

A *neural network* is a collection of neurons in which some neurons have the ability to act on others (called *synapse*). Each neuron is supposed to have two stable states ("all-or-none" activity). The state of the neuron may be changed only by those acting on it, except for the initial (*afferent*) neurons. The final (*efferent*) neurons do not act on any others. The action may be *excitory* or *inhibitory*. The system is considered to be linear, so the total action on a neuron is the sum of the actions:

$$\sum \text{(excitory)} - \sum \text{(inhibitory)}.$$

Such a net was postulated by McCulloch and Pitts [114]. Landahal and Runge [95] outlined a matrix method of describing such a neuron network. The neurons are numbered arbitrarily as $1, 2, \ldots, k$. If

$$e_{ij} = \text{number of excitory inputs from neuron } i \text{ to neuron } j,$$
$$i_{ij} = \text{number of inhibitory inputs from neuron } i \text{ to neuron } j,$$

and

$$\theta_j = \text{threshold level of neuron } j,$$

the matrix is defined as

$$\mathsf{F} = [f_{ij}]_{k,k}, \qquad f_{ij} = \frac{1}{\theta_j}\{e_{ij} - i_{ij}\}. \tag{10–47}$$

Evidently we can interpret this matrix as the connection matrix of a net, with vertices corresponding to neurons, and edges corresponding to synapses. The weight of the edge (i, j) is precisely the total effect of neuron i on neuron j, namely f_{ij}. Landahal and Runge [95] and Telson-Wei [172] have studied this matrix in detail, with reference to neural nets. Shimbel [164, 165] carries this idea further by defining the *structure matrix* of a net and using powers of this matrix to give the information state of the net. The structure matrix of Shimbel is the transpose of the transition matrix defined in Chapter 9, and the results obtained by Shimbel are also analogous.

The application of graph and matrix theories to problems in the study of social groups is due to Luce and Perry [99, 100]. The fundamental concept here is again that of a relation. Relations such as "friend of," "superior of," etc., are of interest in psychology. Any group of individuals with such relations defined on them can be represented as a net. The connection matrix of the net and matrices related to the connection matrix can then be used to analyze the structure of the set. A detailed treatment of the applications of graph theory to problems in social sciences is given by Harary and Norman [73].

The theory of graphs also has important applications to the theory of games and problems in economics. These applications are treated in detail by Berge [9].

Problems

10–1. Prove (Prihar's) Theorems 10–1, 10–2, and 10–3.

10–2. Prove Theorem 10–4. Relate the integer m of Theorem 10–4 to the *diameter* of the network, which is the maximum number of edges in the shortest path between any two nodes.

10–3. Illustrate the use of Theorem 10–5 on Fig. 10–3.

10–4. Solve the system of equations

$$\begin{bmatrix} 2 & 4 & 3 & 7 \\ 1 & 4 & 6 & 3 \\ 2 & 3 & 0 & 4 \\ 5 & 11 & 7 & 14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$$

for x_1 (a) by Cramer's rule, (b) by Mason's signal-flow graphs, using reduction techniques, and (c) by writing the solution by inspection from the graph.

10–5. Repeat Problem 10–4, but use the connection matrix only, without reference to the graph.

10–6. Given a consistent system of equations, let the equations be arranged such that the diagonal entries are nonzero. Let this system be represented as a Coates flow graph. Prove that if all the incoming transmissions at each non-source node x_i are divided by $-c_{ii}$, where c_{ii} is the self-loop at x_i, and then the self-loops are removed, the resulting graph is a Mason flow graph of the same system of equations.

10–7. Justify the following procedure for removal of a node in the Coates flow graph. Let x_i be a node with a self-loop c_{ii}. Divide all incoming transmissions at x_i by $-c_{ii}$. Delete the self-loop. Now remove node x_i as in Mason flow graphs.

10–8. Justify the following operation on Coates flow graphs. Let x_i and x_j be two nodes such that there is an edge from x_i to x_j. Change the weights of the (i, i)-, (j, i)-, (i, j)-, and (j, j)-edges as in Table 10–4. Now interchange the labels x_i and x_j. This operation leaves the equations unaltered.

TABLE 10–4

Edge	Old weight	New weight
(i, i)	A	C
(i, j)	B	D
(j, i)	C	A
(j, j)	D	B

10–9. Show that by using the procedures of Problems 10–7 and 10–8, the Coates flow graph of a consistent system of equations with one source can be reduced to the form of Fig. 10–12, from which the solution can be written.

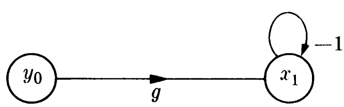


FIGURE 10–12

10–10. Check that the system of equations

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} y_0$$

is solvable and has a unique solution. Without rearranging equations (which is unnecessary since all diagonal entries are nonzero), draw a Mason flow graph. Attempt to solve for x_1 by eliminating, *in order*, the nodes x_3 and x_2.

Devise a node-interchange procedure for Mason flow graphs, similar to the procedure of Problem 10–8, to circumvent this difficulty.

10–11. Solve Problem 7–16 by using (a) Mason graphs and (b) Coates graphs.

10–12. Find the voltage ratio $V_{22'}/V_{11'}$ of Fig. 7–18 by using (a) Mason graphs and (b) Coates graphs.

10–13. Prove that if $c_{i_1j_1}, c_{i_2j_2}, \ldots, c_{i_rj_r}$ is a directed circuit, the corresponding submatrix of $\mathbf{C}$ can be diagonalized by $r - 1$ interchanges of columns. [*Hint:* Induction on r.]

10–14. Prove Theorem 10–7 and its corollaries. [*Hint:* Problem 10–13.]

APPENDIX

APPENDIX

RESEARCH PROBLEMS

This collection of research problems is divided into three groups. The first set consists of fairly simple problems, which require some original thought but for which the solutions are either known to exist or appear to be simply obtainable. These are suitable for term papers or masters' theses, depending on the problem and the depth to which the problem is pursued. Group 2 consists of true research problems which, however, appear to be solvable. Some of these may be suitable for Ph.D. theses. The last set consists of problems which are known to be difficult.

Group 1

1. Devise a simple proof of Theorem 3–8 either directly or by showing that the conditions of Theorem 4–24 imply 2-isomorphism (without using Theorem 3–8).

2. Generalize the notions of a tree, cut-set, duality, etc., to arbitrary matrices of elements from a field, and prove some of the statements at the end of Chapter 5 for such matrices.

3. Use a lattice representation of Tutte's [185] results to get a computer program for checking a matrix of integers mod 2 for regularity.

4. Attempt to prove the relation between regular and E-matrices of Chapter 5 (Theorems 5–25 and 5–26) without the additional hypotheses of "normal form" and "independent rows."

5. With the notation of Chapter 6, show that if

$$\Delta = \det\left[\mathbf{B}_{22}\left(s\mathbf{L}_{22} + \mathbf{R}_{22} + \frac{1}{s}\,\mathbf{D}_{22}\right)\mathbf{B}'_{22}\right]$$

and Δ_{ij} is the cofactor of the (i, j)-element, then Δ_{ij}/Δ has at most a simple pole at infinity.

6. Using the result of Problem 5 and the results of Dolezal [44], find the condition that the solutions (in the s-domain) for the loop currents and node voltages be proper Laplace transforms (so that no impulses occur in the time domain).

7. With the notation of Problem 5, show that the poles of Δ_{ij}/Δ on the imaginary axis are all simple.

8. With the same notation as in Problem 5, show that the residue of Δ_{ii}/Δ at a pole on the imaginary axis is real and positive (without using the theory of positive real functions). Is this result valid even if there is no edge in loop i which is in no other circuit?

9. Complete the existence theorem for the solvability of network equations. That is, determine whether the functions $\mathsf{I}_{m2}(s)$ and $\mathsf{V}_{n2}(s)$ are proper Laplace transform functions under conditions (a), (b), and (c) of Theorem 6–12, suitable requirements on the driving functions (such as continuity), and the restrictions

$$\mathsf{Q}[\mathsf{i}_L(0+) + \mathsf{i}_d(0+)] = 0, \qquad \mathsf{B}[\mathsf{v}_C(0+) + \mathsf{v}_d(0+)] = 0$$

on the initial conditions, where Q is a matrix of cut-sets containing only inductors and current drivers, B is a matrix of circuits containing only capacitors and voltage generators, $\mathsf{i}_d(t)$ are current drivers, and $\mathsf{v}_d(t)$ are voltage drivers; and determine whether $\mathsf{i}(t)$ and $\mathsf{v}(t)$ thus obtained satisfy the initial conditions. (See Dolezal [44].)

10. Find the necessary and sufficient condition that a cut-set matrix Q must satisfy if, in

$$\mathsf{V}(s) = \mathsf{Q}'\mathsf{V}_p(s),$$

the variables in $\mathsf{V}_p(s)$ are to be node-pair voltages. (That is, when is a cut-set matrix also a node-pair transformation matrix?)

11. Relate Problem 10 to the theory of regular and E-matrices.

12. Explore the possibility of using transformation matrices other than B', A', and Q' [151].

13. Make a study of other significant classes of subgraphs (collections of paths, for example) and their associated matrices.

14. Show that any lumped network (in general, most lumped systems) containing a single nonlinear element can always be described by a system of differential equations, only one of which is nonlinear and of the first degree in the nonlinear term, as

$$\begin{bmatrix} a_{11}(p) & \mathsf{A}_{12}(p) \\ 0 & \mathsf{A}_{22}(p) \end{bmatrix} \begin{bmatrix} x_1 \\ \mathsf{X}_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ \mathsf{Y}_2 \end{bmatrix},$$

where $a_{11}(p)$ contains the nonlinear term cp (c depends on some x's) and the matrices $\mathsf{A}_{12}(p)$ (a row matrix) and $\mathsf{A}_{22}(p)$ contain only linear terms. Hence obtain a general method of solving such networks.

15. Extend the method of Problem 14 to networks with two nonlinear elements.

16. Given an arbitrary reactance network, by how much can the degree of Δ (and Δ_{ij}) be increased by inserting a resistor, either in series with a reactance element or by splitting a node in two?

17. Derive general topological formulas based on impedances, dual to the Coates-Mayeda formulas.

18. Obtain a simpler method than sign permutation for evaluating the sign coefficient ϵ_j of a tree product.

19. Give a representation of general linear networks (different from Coates [36]) in which current and voltage elements are required only for unilateral devices ($y_{ij} \neq y_{ji}$) and not for a transformer. Modify the topological formulas accordingly.

20. Given $W_{1,1'}$ and $W_{2,2'}$, it is easy to show that $W_{12,1'2'}$ and $W_{12',1'2}$ contain all the 2-tree products common to $W_{1,1'}$ and $W_{2,2'}$. Find out how to separate $W_{12,1'2'}$ from $W_{12',1'2}$, given $W_{1,1'}$ and $W_{2,2'}$. (It is not possible to say which set of 2-trees belongs to $W_{12,1'2'}$; it is required only to separate the common 2-trees into two groups.)

21. Give a nontrivial sufficient condition for an active network to be minimum-phase.

22. Give a precise proof of Mason's formula for finding the gain of a signal-flow graph by inspection.

23. Make a critical study of Bashkow's [6] equations and rewrite them in matrix notation. Derive simple rules for writing the step before the final equations (before the resistor equations are eliminated) by inspection from the network.

24. Clarify the concept of an essential node in a signal-flow graph. Find the conditions on the choice of variables under which the number of essential nodes is a true measure of system complexity, that is, the conditions under which the number of essential nodes is an invariant characteristic of the system, being the same for all signal-flow graphs of the system (satisfying the conditions).

25. Relate the coefficient matrix of Bashkow's [6] equations to the topology of the network in such a way that the stability of the system can be investigated algebraically.

26. Extend the analogy between conventional networks and contact networks to transfer functions and multiterminal contact networks.

27. Derive the rules for combining multiterminal contact networks in tandem (as an extension of Shannon's disjunctivity theorem) similar to the $ABCD$-matrix technique of conventional network theory.

28. Extend Problem 27 to get a matrix-factoring technique of synthesis.

29. Discuss the classification of graphs by nullity (Foster [52] and Whitney [196]), and look for any possible application in contact network theory.

30. Mayeda [112] has given a simple method of finding all the trees of an SC-network from the switching function. Extend Mayeda's procedure to get an algorithm for finding Y_d directly from the switching function F.

31. In a sequential machine that is not completely specified, reduction procedures based on connection matrices, flow tables, or transition matrices do not necessarily lead to the minimum-state machine. (See Ginsburg [65] or Miller [116].) Find a suitable (nontrivial) condition (on the outputs, possibly) for the matrix or flow-table procedure to lead to the minimum-state machine.

Group 2

1. The edge-to-vertex dual G^* of a graph G is obtained as follows. G^* contains one vertex for each edge of G. Two vertices of G^* are connected by an edge if and only if the corresponding edges of G have a common vertex. (There are two such edges in G^* if the corresponding edges of G are in parallel.) Thus adjacency is preserved. Every G has an edge-to-vertex dual. Under what conditions does a graph have a vertex-to-edge dual which preserves adjacency? (See Kotzig [91].)

2. Classify the cut-set and circuit matrices of graphs which are also E-matrices. (See Okada [126] and Cederbaum [28].)

3. State and prove the necessary and sufficient conditions for the solvability of network equations when dependent sources of all types are admitted.

4. A familiar criterion for finding out whether a transistor switch will "turn over" is to compute the driving-point impedance at some point in the network. If the driving-point impedance has a negative real part for some frequency ω_0, the switch will "turn over." Either justify this procedure or give a counterexample.

5. Let an amplifier network be described by a set of first-order differential equations (Bashkow [6] equations, for example) as

$$\mathbf{A} \frac{d}{dt} \mathbf{X} + \mathbf{B}\mathbf{X} = \mathbf{Y},$$

where $\mathbf{B}$ is a positive definite real symmetric matrix. Then a sufficient condition for stability is that $\mathbf{A}_s = (\mathbf{A} + \mathbf{A}')/2$ be positive definite. Characterize the topological structure of amplifiers satisfying this condition. How is the gain (vs. ω) characteristic affected by such a requirement?

6. Characterize driving-point impedances of ladder networks in which each arm is a simple R, L, or C (by analytic behavior in the complex plane, not as a continued fraction). Extend to series-parallel arms.

7. Translate (Tutte's) Theorems 5–28 and 5–29 to get direct characterizations of node and mesh discriminants, $V(Y)$ and $C[V(Z)]$, without going through the matrices [152].

8. Derive more general topological formulas for networks containing (a) gyrators, (b) voltage-dependent voltage generators, and (c) current-dependent current generators, in addition to (d) R-, L-, C-, and M-elements.

9. Set lower bounds on the number of resistors required to realize positive real functions without the use of transformers.

10. Investigate the conjectures of Seshu [157] regarding the structure of the minimal realization of a minimum p.r. function.

11. Consider minimum p.r. functions which have $\operatorname{Re} Z(j\omega) = 0$ at more than one pair of points on the $j\omega$-axis. Do such functions have

realizations that are more complicated or less complicated than other minimum p.r. functions of the same degree?

12. Extend the theory of order of complexity to linear active networks.

13. Interpret Dasher synthesis as "overremoval" of poles [70] in analogy with the Cauer ladder development, and thereby obtain a simplification of the Dasher synthesis. (Dasher synthesis is complicated by the necessity of shifting a zero of the driving-point function into the complex plane to agree with a zero of transmission, by the overremoval of a single pole. Consider manipulating two or three poles simultaneously.)

14. (Suggested by Dr. R. A. Johnson, Syracuse University.) There is no known precise definition of feedback that agrees with our intuitive concept. Under any known definition, a passive network contains feedback. Attempt to define *feedback* in such a way as to agree with our intuitive concept by distinguishing between "power" and "information" transfer (where both words in quotes are used intuitively) in a network. The information feedback is used to control the power in a "true" feedback amplifier.

15. (Extension of Problem 14.) "Energy" and "information" are respectively the main concerns of the power and communication branches of electrical engineering. Yet, much of communication engineering is based on a concept of energy, whereas it is the information that is much more important. Explore the possibility of introducing information-theoretic concepts into the foundations of network theory.

16. Adapt Cardot's minimality argument to conventional networks (perhaps by trying to find the effect on the network function of opening or shorting some element).

17. Investigate the theory and applications of probabilistic nets. (See Gillespie and Aufenkamp [64].)

18. The topological formulas for the graph gain of a (Mason or Coates) flow graph are not very efficient. In particular, they can locate only the nonzero terms in the expansion of the system determinant and not the cancellations that occur. Cancellations cannot be located in the general case (for arbitrary systems of equations) since the coefficients c_{ij} (or f_{ij}) are unrelated. By suitably specializing the system of equations for a network, obtain more efficient formulas for the Mason or Coates flow graph.

19. A linear graph can be associated with a lumped (rotational or translational) mechanical system in exactly the same fashion as in electrical networks. Make a study of the techniques for establishing graphs (not signal-flow graphs) corresponding to any system which can be described by either algebraic or differential equations or both. Consider partial-differential equations as well. (See Trent [177].)

20. If F, V, and W are as defined in Chapter 9, what is $F + W$ (Boolean)?

21. Extend Shannon's minimality proof to one terminal-pair contact networks by setting lower bounds on the number of internal vertices and using König's [88] results on cut-sets of vertices.

22. Find a method of getting all the minimal realizations of a Boolean function as a contact network, given one of them.

23. Characterize single-contact switching functions directly, without using the matrix $\mathbf{B}_F$.

24. Organize Gould's matrix-synthesis technique for use on a digital computer, using Theorems 5–28 and 5–29.

25. Obtain an alternative representation of a logic network as a net which is not so redundant as Shelly's representation (Section 9–3).

26. Use Problem 1 to solve Problem 25.

27. Extend Vasil'ev's [189] theory to general contact networks.

28. Investigate the usefulness of the concept of a dual of a sequential machine, as defined by the dual graph. Also consider the edge-to-vertex dual of Problem 1.

29. The information content of a finite state diagram can be defined in the sense of discrete information theory. If there are n states, m inputs, and p outputs, the information content is

$$n(m \log_2 n + \log_2 p) \text{ bits.}$$

Given n and the set of permissible inputs ($p \leq n$ in Moore's model), devise a test for finding the structure of the machine under suitable assumptions (strong connectedness, for example). Hence find the analogue of the convolution representation (impulse-response representation) of conventional network theory.

30. Investigate the application of Trucco's [178] concept of information content of a graph to sequential machines and logic networks.

Group 3

(Excludes classical problems in graph theory. For these, see [73].)

1. Find a characterization of the incidence matrix $\mathbf{A}$ that is invariant under elementary row operations. In other words, characterize cut-set matrices of linear graphs by methods different from Theorems 5–28 and 5–29.

2. By using concepts from the theory of coding (the Hamming code and its extensions), and by using n-cubes, translate Theorem 5–28 into a form that will permit a computer to determine directly whether $\mathbf{F}$ is a regular matrix, without trying all normal forms of $\mathbf{F}$. Repeat for Theorem 5–29.

3. (Algebraic stability criterion.) Let

$$\mathsf{A}\frac{d^2}{dt^2}\mathsf{X} + \mathsf{B}\frac{d}{dt}\mathsf{X} + \mathsf{CX} = \mathsf{Y}$$

be a system of equations with real constant coefficients. If the matrices A, B, and C are symmetric, positive semidefiniteness or definiteness of A, B, and C is a necessary and sufficient condition for stability. If A, B, and C are not symmetric, but their symmetric parts [$\mathsf{A}_s = (\mathsf{A} + \mathsf{A}')/2$, etc.] are positive definite, the system is again stable. If the matrices are *normal* ($\mathsf{AA}' = \mathsf{A}'\mathsf{A}$) the positive semidefiniteness of symmetric parts is again necessary as well. Find the necessary and sufficient conditions for non-normal matrices A, B, and C.

4. Find bounds on the number of 2-trees $(1, 1')$ of a graph of e edges, v vertices, and k trees.

5. Characterize positive real functions by the coefficients of the numerator and denominator polynomials.

6. Given a positive real function $Y(s) = p(s)/q(s)$, find a procedure for choosing the element functions $y_1, y_2, \ldots, y_e$ such that $Y(s) = V[Y(s)]/W[Y(s)]$, and V and W are realizable [152].

7. In Chapter 8, it was shown that the Reza-Pantell-Fialkow-Gerst realization is strictly minimal (except for two special cases) for biquadratic minimum functions. Examine the validity of the conjecture that the Reza-Pantell-Fialkow-Gerst realization is minimal for "almost all" minimum p.r. functions, for any degree. (In this connection, see the existence theorem of Shannon [161] on realizations of Boolean functions, based on additive-number theory.)

8. State general conditions on the topology of the network and element values for the cancellation of common factors between Δ and Δ_{ij}.

9. In all applications of the theory of nets (sequential machines, flow graphs, etc.), the edge appears to perform the "operation." For instance, in sequential machines, it is the edge weight that "takes" the machine from one state to another. In signal-flow graphs, the edge "transmits" the signal from one point to another. Derive a formalism in which the "operation" is performed by the vertex. Such a system is useful in logic networks, representation of "testing sequences," etc.

10. Examine whether the vertex-to-edge dual (Problem 1, Group 2), when it exists, can be used to solve the previous problem, 9.

11. (Suggested by Dr. J. P. Runyan, Bell Telephone Laboratories.) A *feedback cut-set* of a directed graph is a set of edges which, if removed, destroys all directed circuits. The feedback cut-set containing the smallest number of edges is a *minimal feedback cut-set*. Give an efficient algorithm for finding the minimal feedback cut-set of a directed graph. Or, find a test for determining whether a given cut-set is minimal. (The algorithm

should be such that the number of operations increases linearly with the number of vertices and not exponentially. Thus, it should not require the listing of all directed circuits.)

12. The fundamental difficulty in the previous problem is that the set of directed circuits is not a group (or even a semigroup) under any of the familiar operations; much less is it a linear vector space. By contrast, the nonoriented circuits are destroyed on removing exactly μ (=nullity) edges. Find the algebraic structure of the set of all directed circuits of a directed graph.

13. Derive a graph representation for nonsimple theorems (with more than one hypothesis and/or conclusion), perhaps patterned after Shelly's [163] representation of logic networks.

14. Find an upper bound for the difference between the numbers of contacts in a minimal realization and a minimal series-parallel realization of a Boolean function of n variables.

15. Menger's theorem states that the number of edge-disjoint paths between two vertices a and b is the number of edges in the smallest cut-set (a, b). Apply this theorem to minimality in both conventional and contact networks.

16. In the theory of redundant relay-contact networks (Moore and Shannon [118]), one terminal-pair networks in which (number of edges) = (length) $\times$ (width) are useful. Such networks are called *rectangular*. It is generally desired that the number of cut-sets of the smallest width and the number of paths of the smallest length should both be minimum. The *hammock* networks of Moore and Shannon [118] are nearly optimum in this sense, but are not always so. Find the optimum network for given length and width.

17. Give a suitable definition of a *linear sequential machine* such that the network-theoretic concepts of driving-point and transfer functions can be extended to sequential machines.

18. Give the analogue of return difference and sensitivity in combinational-contact or sequential networks, or both, in such a way as to agree with the Moore-Shannon theory [118].

19. In the synthesis of sequential machines, one is interested in getting certain oriented paths in the state diagram. In these paths, the order of edges is important ($i_1i_2 \neq i_2i_1$). Therefore, the matrix technique of Gould is not applicable. Derive a formalism that is suitable for this purpose, which can be manipulated as easily as the matrices of the graph.

20. Derive a suitable measure of "complexity" in a sequential switching system, which takes into account both the memory and the logic requirements.

BIBLIOGRAPHY

BIBLIOGRAPHY

To make this bibliography more useful to the research worker, most of the significant contributions related to the subject matter of the text have been included regardless of whether any direct reference has been made in the text. In fundamental graph theory, only those papers which relate to the topics discussed in the text are included; papers on the coloring problems and enumeration problems are thus omitted. Also, the coverage is not thorough for the period 1847–1930. An attempt at completeness has been made in the application to electric network theory and to the theory of switching.

Other extensive bibliographies are to be found in (numbers refer to present bibliography) Berge [9], Cauer [25, English version], Coxeter [39], Foster [59], Hohn [78], Harary [72], König [88], Mayeda and Seshu [109], Quade [137], Reza [146], St. Lague [150] (very good coverage of early papers), Seshu [152], and Van Valkenburg [188].

1. AHRENS, W., Über das Gleichungsystem einer Kirchhoffschen Galvanishen Stromzweigung, *Math. Ann.* **49,** 311–324 (1897).

2. ARSOVE, M. G., Note on Network Postulates, *J. Math. and Phys.* **32** 203–206 (1953).

3. ASH, R. B., and KIM, W. H., On the Realizability of a Circuit Matrix, *Trans. Inst. Radio Engrs.* **CT–6,** 219–223 (June 1959).

4. ASHENHURST, R. L., *A Uniqueness Theorem for Abstract Two Terminal Switching Networks,* Bell Lab. Progr. Rpt. no. BL–10, Harvard Computation Lab. (1954).

5. AUFENKAMP, D. D., and HOHN, F. E., Analysis of Sequential Machines, *Trans. Inst. Radio Engrs.* **EC–6,** 276–285 (Dec. 1957).

6. BASHKOW, T. R., The *A* Matrix, New Network Description, *Trans. Inst. Radio Engrs.* **CT–4,** 117–119 (Sept. 1957).

7. BAYARD, M., Théorie des Réseaux de Kirchhoff, *Rev. optique* **33,** (1954).

8. BELEVITCH, V., Some Relations between the Theory of Contact Networks and Conventional Network Theory, *Proc. Intl. Symposium on Switching Circuits,* Harvard U. (1957).

9. BERGE, C., *Théorie des Graphes et ses Applications.* Paris: Dunod, 1958.

10. BERS, A., The Degrees of Freedom in *RLC* Networks, *Trans. Inst. Radio Engrs.* **CT–6,** 91–95 (March 1959).

11. BIRKHOFF, G., and MACLANE, S., *Survey of Modern Algebra.* New York: Macmillan, 1953.

12. BODE, H. W., *Network Analysis and Feedback Amplifier Design.* New York: Van Nostrand, 1945.

13. BOWER, R. E., KOSOWSKY, L. H., and ORDNUNG, P. F., Functional Characteristics of a Node Determinant, *J. Franklin Inst.* **265,** 395–406 (1958).

14. BROOKS, R. L., SMITH, C. A. B., STONE, A. H., and TUTTE, W. T., Dissection of a Rectangle into Squares, *Duke Math. J.* **7,** 312–340 (1940).

15. Bruijn, van A. de, and Ehrenfest, T. van A., *Circuits and Trees in Oriented Graphs*, Simon Stevin, Groningen, Netherlands, **28,** 203–217 (1951).

16. Brune, O., Synthesis of Two Terminal Networks, *J. Math. and Phys.* **10,** 191–236 (1931).

17. Bryant, P. R., A Topological Investigation of Network Determinants, Inst. Elec. Engrs. (London) Monograph no. 312R (Sept. 1958); *Proc. Inst. Elec. Engrs.* (London), **106** (part C), 16–22 (March 1959).

18. Bryant, P. R., *Order of Complexity of Electrical Networks*, Inst. Elec. Engrs. (London) Monograph no. 335E (June 1959).

19. Calabi, L., *Algebraic Topology of Networks I, II.* Concord, Mass.: Parke Math. Lab., 1956.

20. Calabi, L., *Algebraic Topology and Graph Theory.* Concord, Mass.: Parke Math. Lab., 1959.

21. Caldwell, S. H., *Switching Circuits and Logical Design.* New York: Wiley, 1958.

22. Cardot, C., Some Results of the Application of Boolean Algebra to Relay Switching Circuits, *Ann. telecom.* **7,** 80–82 (1952). (French.)

23. Carlitz, L., and Riordan, J., The Number of Labelled Two-Terminal Series-Parallel Networks, *Duke Math. J.* **23,** 435–445 (1956).

24. Cauer, W., Über Kirchhoffsche Regeln zur anschaulichen Ermittlung der Eigenschaften von Netzwerken, *Mix u. Genest Tech. Nachr.* **10,** 23–30 (1938).

25. Cauer, W., *Théorie der Linearen Wechselstromschaltungen.* Berlin: Akademie Verlag, 1954; English translation, New York: McGraw-Hill, 1958.

26. Cayley, A., On the Theory of Analytical Forms Called Trees, *Phil. Mag.* **13,** 172–176 (1857).

27. Cayley, A., On the Mathematical Theory of Isomers, *Phil. Mag.* **47,** 444–446 (1874).

28. Cederbaum, I., Invariance and Mutual Relations of Network Determinants, *J. Math. and Phys.* **34,** 236–244 (Jan. 1956).

29. Cederbaum, I., Some Properties of the Transfer Function of Unbalanced *RC* Networks, *Proc. Inst. Elec. Engrs.* (part C) **103,** 400–406 (Sept. 1956).

30. Cederbaum, I., On Networks without Ideal Transformers, *Trans. Inst. Radio Engrs.* **CT–3,** 179–182 (Sept. 1956).

31. Cederbaum, I., *Conditions for the Impedance and Admittance Matrices of n-Ports without Ideal Transformers*, Inst. Elec. Engrs. (London) Monograph no. 276R (Jan. 1958).

32. Cederbaum, I., Matrices all of whose Elements and Subdeterminants are 1, –1 or 0, *J. Math. and Phys.* **36,** 351–361 (1958).

33. Cederbaum, I., Applications of Matrix Algebra to Network Theory, *Trans. Inst. Radio Engrs.* **CT–6** (special supplement), 152–157 (May 1959).

34. Chow, W. L., On Electrical Networks, *J. Chinese Math. Soc.* **2,** 321–339 (1940).

35. Chuard, J., Questions d'Analysis Situs, *Rend. Circ. Math. Palermo* **46,** 185–224 (1922).

36. Coates, C. L., General Topological Formulas for Linear Network Functions, *Trans. Inst. Radio Engrs.* **CT–5,** 30–42 (March 1958).

37. Coates, C. L., Flow Graph Solutions of Linear Algebraic Equations, *Trans. Inst. Radio Engrs.* **CT–6,** 170–187 (June 1959).

38. Copi, I. M., Matrix Development of the Calculus of Relations, *J. Symbolic Logic* **13,** 193–203 (1948).

39. Coxeter, H. S. M., Self Dual Configurations and Regular Graphs, *Bull. Am. Math. Soc.* **56,** 413–455 (1950).

40. Cruickshank, A. J. O., Kron's Solution of Orthogonal Networks, *Matrix and Tensor Quart.* **7,** 51–55 (1956).

41. Dantzig, G. B., and Fulkerson, D. R., On Max-Flow Min-Cut Theorem of Networks, in Linear Inequalities, *Ann. Math. Studies,* no. 38, 215–221 (1956).

42. Dasher, B. J., Synthesis of *RC* Transfer Functions as Unbalanced Two Terminal-Pairs, *Trans. Inst. Radio Engrs.* **CT–1,** 20–34 (1952).

43. Dehn, M., and Heegaard, P., Analysis Situs, *Encyklopaedie der Mathematischen Wissenschaften* **III:** 11, 153–220 (1907).

44. Doležal, V., Systems of Ordinary Linear Integro-Differential Equations, *Appl. mat.* **4,** 1–17 (1959). (Czech.)

45. Doležal, V., and Kurzweil, J., On Certain Properties of Linear Differential Equations, *Appl. mat.* **4,** 163–174 (1959). (Czech.)

46. Doyle, T. C., Topological and Dynamical Invariant Theory of an Electrical Network, *J. Math. and Phys.* **34,** 81–94 (July 1955).

47. Effertz, H. F., Bounded Functions, Frequency Characteristics of Electric Networks and Algebraic Stability Criteria, *Z. Angew. Math. u. Mech.* **33,** 281–283 (1953).

48. Elgot, C. C., and Wright, J. B., Series-Parallel Graphs and Lattices, *Duke Math. J.* **26,** 325–338 (June 1959).

49. Elias, G. J., and Tellegan, B. D. H., *Théorie der Wisselströmen.* Groningen, Djakarta: P. Noordhoff N. V., 1952.

50. Elias, P., Feinstein, A., and Shannon, C. E., A Note on the Maximum Flow Through a Network, *Trans. Inst. Radio Engrs.* **IT–2,** 117–119 (Dec. 1956).

51. Euler, L., Solutio Problematis ad Geometriam Situs Pertinantis, *Academimae Petropolitanae* **8,** 128–140 (1736).

52. Feussner, W., Über Stromverzweigung in Netzförmigen Leitern, *Ann. Physik,* **9** (4th ser.), 1304–1329 (1902).

53. Feussner, W., Zur Berechnung der Stromstärke in Netzförmigen Leitern, *Ann. Physik* **15** (4th ser.), 385–394 (1904).

54. Fialkow, A., and Gerst, I., The Transfer Function of General Two Terminal Pair *RC* Networks, *Quart. Appl. Math.* **10,** 113–127 (1952).

55. Fialkow, A., and Gerst, I., The Transfer Function of Networks without Mutual Reactance, *Quart. Appl. Math.* **12**:2, 117–131 (July 1954).

56. Ford, L. R., and Fulkerson, D. R., Maximal Flow through a Network, *Can. J. Math.* **8,** 399–404 (1956).

57. Ford, L. R., and Fulkerson, D. R., A Simple Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem, *Can. J. Math.* **9**:2, 210–218 (1957).

58. Foster, R. M., Geometric Circuits of Electrical Networks, Bell System Monograph no. B-653; *Trans. Am. Inst. Elec. Engrs.* **51,** 309–317 (1932).

59. Foster, R. M., Topologic and Algebraic Considerations in Network Synthesis, *Proc. Polytech. Inst. Brooklyn Symposium on Modern Network Synthesis* **I,** 8–18 (April 1952).

60. Foster, R. M., The Number of Series Parallel Networks, *Proc. Int. Congr. Math.* **I,** 646 (1952).

61. Foster, R. M., Passive Network Synthesis, *Proc. Polytech. Inst. Brooklyn Symposium on Modern Network Synthesis* **5,** 3–9 (1955).

62. Gale, D. A., A Theorem on Flows in Networks, *Pac. J. Math.* **7,** 1073–1082 (1957).

63. Gilbert, E. N., Enumeration of Labelled Graphs, *Can. J. Math.* **8,** 405–411 (1956).

64. Gillespie, R. G., and Aufenkamp, D. D., On the Analysis of Sequential Machines, *Trans. Inst. Radio Engrs.* **EC–7,** 119–122 (June 1958).

65. Ginsburg, S., A Synthesis Technique for Minimal State Sequential Machines, *Trans. Inst. Radio Engrs.* **EC–8,** 13–24 (March 1959).

66. Gould, R. L., Application of Graph Theory to the Synthesis of Contact Networks, *Proc. Int. Symposium on Switching Circuits,* Harvard U., April 1957.

67. Gould, R. L., Graphs and Vector Spaces, *J. Math. and Phys.* **38,** 193–214 (1958).

68. Guillemin, E. A., *Introductory Circuit Theory.* New York: Wiley, 1953.

69. Guillemin, E. A., *Communication Networks,* vol. **1.** New York: Wiley, 1936.

70. Hakimi, S., and Seshu, S., Realization of Complex Zeros of Transmission by Means of *RC* Networks, *Proc. Natl. Electronics Conf.* **13,** 1013–1025 (1957).

71. Harary, F., and Norman, Z., *Graph Theory as a Mathematical Model in Social Science,* U. of Mich. Monograph (1953).

72. Harary, F., Graph Theory and Electric Networks, *Trans. Inst. Radio Engrs.* **CT–6** (special supplement), 95–109 (May 1959).

73. Harary, F., and Norman, Z., *The Theory of Graphs.* Reading, Mass.: Addison-Wesley (to be published).

74. Havel, V., Eine Bemerkung über die Existenz der Endlichen Graphen, *Čas. Pest. mat.* **80,** 477–480 (1955).

75. Herz, P., Über Axiomensysteme für beliebige Satzsysteme, *Math. Ann.* **87,** 246–269 (1922).

76. Hohn, F. E., and Schissler, L. R., Boolean Matrices in the Design of Combinational Switching Circuits, *Bell System Tech. J.* **34,** 177–202 (Jan. 1955).

77. Hohn, F. E., Seshu, S., and Aufenkamp, D. D., Theory of Nets, *Trans. Inst. Radio Engrs.* **EC–6,** 154–161 (Sept. 1957).

78. Hohn, F. E., *Elementary Matrix Algebra.* New York: Macmillan, 1958.

79. Howitt, N., Group Theory and the Electric Circuit, *Phys. Rev.* **37,** 1583–1595 (1931).

80. Howitt, N., Equivalent Electric Circuits, *Proc. Inst. Radio Engrs.* **20,** 1042–1051 (1932).

81. Huggins, W. H., Signal Theory, *Trans. Inst. Radio Engrs.* **CT–3,** 210–216 (Dec. 1956).

82. Huggins, W. H., Signal Flow Graphs and Random Signals, *Proc. Inst. Radio Engrs.* **45,** 74–86 (Jan. 1957).

83. Ingram, W. H., and Cramlet, C. M., On the Foundations of Electrical Network Theory, *J. Math. and Phys.* **23,** 134–155 (1944).

84. Kim, W. H., *A New Method of Synthesis of Driving Point Functions,* Interim Tech. Rpt. no. 1, U.S. Army Contract no. DA–11–022-ORD–1983, U. of Ill. (April 1956).

85. Kim, W. H., Role of Network Topology in Network Synthesis, *Proc. Midwest Symposium on Circuit Theory* **3,** 6.1–6.18 (1958).

86. Kirchhoff, G., Über die Auflösung der Gleichungen, auf welche man bei der Untersuchungen der Linearen Verteilung Galvanisher Ströme geführt wird, *Poggendorf Ann. Physik* **72,** 497–508 (1847); English translation, *Trans. Inst. Radio Engrs.* **CT–5,** 4–7 (March 1958).

87. Knichal, V., On Kirchhoff's Laws, *Mat. fys. sbornik slov. akad. vied. umeni* **2,** 13–28 (1952). (Czech.)

88. König, D., *Théorie der Endlichen und Unendlichen Graphen.* New York: Chelsea, 1950.

89. Koenig, H. E., and Reed, M. B., Linear Graph Representation of Multi-Terminal Components, *Proc. Natl. Electronics Conf.* **14,** 661–674 (1958).

90. Kotzig, A., The Significance of the Skeleton of a Graph for the Construction of Composition Bases of Some Subgraphs, *Mat. fys. čas. slov. akad. vied.* **6,** 68–77 (1956). (Slovak.)

91. Kotzig, A., Aus der Théorie der Endlichen Regulären Graphen Dritten und Vierten Grades, *Čas. pest. mat.* **82,** 76–92 (1957).

92. Ku, Y. H., Resume of Maxwell's and Kirchhoff's Rules, *J. Franklin Inst.* **253,** 211–224 (1952).

93. Kudryavcev, L. D., On Some Mathematical Problems in Electric Circuit Theory, *Uspekhi Mat. Nauk* (*N.S.*) **3,** 80–118 (1948). (Russian.)

94. Kuratowski, C., Sur le Problème des Courbes Gauches en Topologie, *Fund. Math.* **15,** 271–283 (1930).

95. Landahal, H. D., and Runge, R., Outline of a Matrix Calculus for Neural Nets, *Bull. Math. Biophys.* **8,** 75–81 (1946).

96. LeCorbeiller, P., *Matrix Analysis of Electrical Networks.* Cambridge, Mass.: Harvard U. Press, 1950.

97. Lefschetz, S., *Introduction to Topology.* Princeton, N. J.: Princeton U. Press, 1949.

98. Listing, J. B., *Vorstudien über Topologie.* Göttingen: 1848.

99. Luce, R. D., and Perry, A Matrix Method of Analysis of Group Structure, *Psychometrika* **14**:1, 95–116 (1949).

100. Luce, R. D., Connectivity and Generalized Cliques in Sciometric Group Structure, *Psychometrika* **15,** 169–190 (Jan. 1950).

101. Luce, R. D., Two Decomposition Theorems for a Class of Finite Oriented Graphs, *Am. J. Math.*, **74,** 701–722 (1952).

102. Lund, F., Connection Possibilities, *Norsk Mat. Tidsskr.* **31,** 9–31 (1949). (Norwegian.)

103. Lunts, A. G., The Application of Boolean Matrix Algebra to the Analysis and Synthesis of Relay Contact Networks, *Doklady Akad. Nauk S.S.S.R.* **70,** 421–423 (1950). (Russian.)

104. MacLane, S., A Combinatorial Condition for Planar Graphs, *Fund. Math.* **28,** 22–32 (1937).

105. Mason, S. J., Feedback Theory—Some Properties of Signal Flow Graphs, *Proc. Inst. Radio Engrs.* **41,** 1144–1156 (Sept. 1953).

106. Mason, S. J., Feedback Theory—Further Properties of Signal Flow Graphs, *Proc. Inst. Radio Engrs.* **44,** 920–926 (July 1956).

107. Mason, S. J., Topological Analysis of Linear Non-Reciprocal Networks, *Proc. Inst. Radio Engrs.* **45,** 829–838 (June 1957).

108. Maxwell, J. C., *Electricity and Magnetism,* **1,** Chap. VI and Appendix, pp. 403–410. Oxford: Clarendon Press, 1892.

109. Mayeda, W., and Seshu, S., *Topological Formulas for Network Functions,* Bull. no. 446, U. of Ill. Engineering Experiment Station (1957).

110. Mayeda, W., and Van Valkenburg, M. E., Network Analysis and Synthesis by Digital Computer, *Inst. Radio Engrs. Convention Record* (part 2) **5,** 137–144 (1957).

111. Mayeda, W., *Topological Formulas for Active Networks,* Int. Tech. Rpt. no. 8, U.S. Army Contract no. DA–11–022–ORD–1983, U. of Ill. (Jan. 1958).

112. Mayeda, W., *Application of Mathematical Logic to Network Theory,* Ph.D. thesis, U. of Ill. (1958).

113. Mayeda, W., and Van Valkenburg, M. E., Analysis of Non-Reciprocal Networks by Digital Computers, *Inst. Radio Engrs. Natl. Convention Record* (part 2) **6,** 70–75 (1958).

114. McCulloch, W. S., and Pitts, A Logical Calculus of Ideas Immanent in Nervous Activity, *Bull. Math. Biophys.* **5,** 115–133 (1943).

115. Mealy, G. H., A Method of Synthesizing Sequential Circuits, *Bell System Tech. J.* **34,** 1045–1079 (Sept. 1955).

116. Miller, R. E., State Reduction for Sequential Machines, *I.B.M. J. Res.* **3** (1959).

117. Moore, E. F., Gedanken-Experiments on Sequential Machines, Automata Studies, *Ann. Math. Study* **34,** 129–153 (1956).

118. Moore, E. F., and Shannon, C. E., Reliable Circuits Using Unreliable Relays, *J. Franklin Inst.* **262,** 191–297 *passim* (1956).

119. Moore, E. F., Shortest Path Through a Maze, *Proc. Intl. Symposium on Switching Circuits,* Harvard U. (April 1957).

120. Moore, E. F., "A Table of Two Terminal Circuits for Functions of Four Variables," in Hoggonnet, R. A., and Grea, R. A., *Logical Design of Electric Circuits.* New York: McGraw-Hill, 1958.

121. Nakagawa, N., On the Evaluation of Graph Trees and Driving Point Admittance, *Trans. Inst. Radio Engrs.* **CT–5,** 122–127 (June 1958).

122. Okada, S., and Onerda, R., On Network Topology I, *Bull. Yamagata U.* **2,** 89–117 (1952).

123. Okada, S., and Onerda, R., On Network Topology II, *Bull. Yamagata U.* **2,** 191–206 (1953).

124. Okada, S., Topology Applied to Switching Circuits, *Proc. Polytech. Inst. Brooklyn Symposium on Information Networks* **3,** 267–290 (1954).

125. Okada, S., Topologic and Algebraic Foundations of Network Synthesis, *Proc. Polytech. Inst. Brooklyn Symposium on Modern Network Synthesis* **5,** 283–322 (April 1955).

126. Okada, S., On Node and Mesh Determinants, *Proc. Inst. Radio Engrs.* **43,** 1527 (Oct. 1955).

127. Otter, R., The Number of Trees, *Ann. Math.* **49,** 583–599 (1948).

128. Pantell, R. H., A New Method of Driving Point Impedance Synthesis, *Proc. Inst. Radio Engrs.* **42,** 861 (1954).

129. Percival, W. S., Solution of Passive Electrical Networks by Means of Mathematical Trees, *J. Inst. Elect. Engrs.* (*London*) **100** (part III), 143–150 (1953).

130. Percival, W. S., Improved Matrix and Determinant Methods of Solving Networks, *J. Inst. Elec. Engrs.* (*London*) **101** (part IV), 258–265 (1954).

131. Percival, W. S., *Graphs of Active Networks,* Inst. Elec. Engrs. (London) Monograph no. 129 (1955).

132. Percus, J. K., Matrix Analysis of Oriented Graphs with Irreducible Feedback Loops, *Trans. Inst. Radio Engrs.* **CT–2,** 117–127 (June 1955).

133. Poincaré, H., Analysis Situs, *J. École Polytech.* **1** (2nd ser.), 1–121 (1895).

134. Pólya, G., Kombinatorische Anzahlbestimmungen von Gruppen, Graphen und Chemischen Verbindungen, *Acta Math.* **58,** 145–254 (1937).

135. Povarov, G. N., Matric Analysis of Connections in Partially Oriented Graphs, *Uspekhi Mat. Nauk,* 195–202 (1956). (Russian.)

136. Prihar, A., Topological Properties of Telecommunication Networks, *Proc. Inst. Radio Engrs.* **44,** 927–933 (July 1956).

137. Quade, W., Matrizenrechnung und Elektrische Netze, *Arch. Elektrotech.* **34,** 545–567 (1940).

138. Rapaport, A., Outline of a Probabilistic Approach to Animal Sociology, *Bull. Math. Biophys.* **11,** 183–281 *passim* (1945).

139. Redei, L., Über die Kantenbasen für Endliche Vollständige Gerichtete Graphen, *Acta Math. Akad. Sc. Hung.* **5,** 17–25 (1954).

140. Reed, G. B., and Reed, M. B., Patterns of Driving Elements such as appear in Tube and Transistor Networks, *Proc. 2nd Midwest Symposium on Circuit Theory,* Mich. State U., 3.1–3.16 (Dec. 1956).

141. Reed, M. B., and Seshu, S., On Topology and Network Theory, *Proc. U. of Ill. Symposium on Circuit Analysis,* 2.1–2.16 (May 1955).

142. Reed, M. B., Generalized Mesh and Node Equations, *Trans. Inst. Radio Engrs.* **CT–2,** 162–168 (June 1955).

143. Reidemeister, K., *Einführung in die Kombinatorische Topologie.* Braunschweig: Vieweg, 1932.

144. Reza, F. M., A Supplement to Brune Sythesis, *Comm. and Electronics,* no. 17, 85–90 (March 1955).

145. Reza, F. M., Order of Complexity and Minimal Structures in Network Analysis, *U. of Ill. Symposium on Circuit Analysis,* 7.1–7.33 (1955).

146. Reza, F. M., Some Topological Considerations in Network Theory, *Trans. Inst. Radio Engrs.* **CT–5,** 42–54 (March 1958).

147. RIORDAN, J., and SHANNON, C. E., On the Number of Series-Parallel Two-Terminal Circuits, *J. Math. and Phys.* **21,** 83–93 (1942).

148. RIORDAN, J., *Introduction to Combinatorial Analysis.* New York: Wiley, 1958.

149. SALTZER, C., The Second Fundamental Theorem of Electrical Networks, *Quart. Appl. Math.* **XI**:1, 119–123 (1953).

150. SAINTE LAGUE, Les Réseaux (ou Graphes), *Mémorial sci. math.* no. 18, (1926).

151. SESHU, S., and REED, M. R., Singular Transformations in Network Theory, *Proc. Natl. Electronics Conf.* **11,** 531–543 (1955).

152. SESHU, S., Topological Considerations in the Design of Driving Point Functions, *Trans. Inst. Radio Engrs.* **CT–2,** 356–367 (Dec. 1955).

153. SESHU, S., On Electric Circuits and Switching Circuits, *Trans. Inst. Radio Engrs.* **CT–3,** 172–178 (Sept. 1956).

154. SESHU, S., and REED, M. B., On Cut Sets of Electrical Networks, *Proc. 2nd Midwest Symposium on Circuit Theory,* Mich. State U., 1.1–1.13 (1956).

155. SESHU, S., MILLER, R. E., and METZE, G., Transition Matrices of Sequential Machines, *Trans. Inst. Radio Engrs.* **CT–6,** 5–12 (March 1959).

156. SESHU, S., and BALABANIAN, N., *Linear Network Analysis.* New York: Wiley, 1959.

157. SESHU, S., Minimal Realizations of the Biquadratic Minimum Function, *Trans. Inst. Radio Engrs.* **CT–6,** no. 4, 345–350 (1959).

158. SESHU, S., On E-Matrices, Regular Matrices and Graphs (unpublished).

159. SHANNON, C. E., A Symbolic Analysis of Relay and Switching Circuits, *Trans. Am. Inst. Elec. Engrs.* **57,** 713–723 (1938).

160. SHANNON, C. E., Mathematical Theory of Communication, *Bell System Tech. J.* **27,** 379–656 *passim* (1948).

161. SHANNON, C. E., The Synthesis of Two Terminal Switching Circuits, *Bell System Tech. J.* **28,** 59–98 (1949).

162. SHEKEL, J., Two Theorems Concerning the Change of Reference Voltage Terminal, *Proc. Inst. Radio Engrs.* **42,** 1125 (July 1954).

163. SHELLEY, W. A., *A Matrix Method for the Analysis and Synthesis of Logic Circuits,* Master's thesis, Syracuse U. (1958).

164. SHIMBEL, A., Application of Matrix Algebra to Communication Nets, *Bull. Math. Biophys.* **13**:3, 165–173 (1951).

165. SHIMBEL, A., Structural Parameters of Communication Networks, *Bull. Math. Biophys.* **15**:6, 501–507 (1953).

166. SLEPIAN, D., On the Number of Symmetry Types of Boolean Functions of n Variables, *Can. J. Math.* **5,** 185–193 (1954).

167. STORER, J. E., Relationship between the Bott-Duffin and Pantel Synthesis, *Proc. Inst. Radio Engrs.* **42,** 1451 (1954).

168. SYNGE, J. L., The Fundamental Theorem of Electrical Networks, *Quart. Appl. Math.* **9,** 113–127 (1951).

169. SZELE, T., Kombinatorische Untersuchungen über die Gerichteten Vollständigen Graphen, *Math. Fiz. Lapok.* **50,** 223–256 (1943).

170. Talbot, A., Some Fundamental Properties of Networks without Mutual Inductance, *Proc. Inst. Elec. Engrs.* **102,** 168–175 (Sept. 1955).

171. Tellegan, B. D. H., Geometrical Considerations and the Duality of Networks, *Phillips Tech. Rev.* **5,** 324–330 (1940).

172. Telsen-Wei, Matrices of Neural Nets, *Bull. Math. Biophys.* **10,** 63–81 (1948).

173. Ting, S. L., On the General Properties of Electrical Network Determinants, *Chinese J. Phys.* **1,** 18–40 (1935).

174. Tokad, Y., and Reed, M. B., Criteria and Tests for the Realizability on the Inductance Matrix, *Trans. Am. Inst. Elec. Engrs.*, (part I) **78,** 924–926 (Jan. 1960).

175. Trakhtenbrot, B. A., Synthesis of Non-iterated Circuits, *Doklady Akad. Nauk S.S.S.R.* **103,** 973–976 (1955). (Russian.)

176. Trent, H. M., Note on the Enumeration and Listing of All Possible Trees in a Connected Linear Graph, *Proc. Natl. Acad. Sci. U.S.*, **40,** 1004–1007 (Oct. 1954).

177. Trent, H. M., Isomorphisms between Oriented Linear Graphs and Lumped Physical Systems, *J. Acoust. Soc. Am.* **27,** 500–527 (1955).

178. Trucco, E., On the Information Content of Graphs, *Bull. Math. Biophys.* **18,** 237–253 (1956).

179. Truxal, J. G., *Control System Synthesis.* New York: McGraw-Hill, 1955.

180. Tsai, C. T., Short Cut Methods for Expanding Determinants Involved in Network Problems, *Chinese J. Phys.* **3,** 148–181 (1939).

181. Tsang, N. F., On Electrical Network Determinants, *J. Math. and Phys.* **33,** 185–193 (July 1954).

182. Tutte, W. T., A Ring in Graph Theory, *Proc. Cambridge Phil. Soc.* **43,** 26–40 (1947).

183. Tutte, W. T., A Theorem on Planar Graphs, *Trans. Am. Math. Soc.* **82,** 99–116 (1956).

184. Tutte, W. T., A Class of Abelian Groups, *Can. J. Math.* **8,** 13–28 (1956).

185. Tutte, W. T., A Homotopy Theorem for Matroids, I, II, *Trans. Am. Math. Soc.* **88,** 144–174 (May 1958).

186. Tutte, W. T., Matroids and Graphs, *Trans. Am. Math. Soc.* **90,** 527–552 (March 1959).

187. Van der Waerden, B. L., *Modern Algebra.* New York: Frederick Ungar, 1950.

188. Van Valkenburg, M. E., Topological Synthesis, *Inst. Radio Engrs. Wescon Record* (part 2) **2,** 3–9 (1958).

189. Vasil'ev, Iu. L., Minimal Contact Networks for Boolean Functions of Four Variables, *Doklady Akad. Nauk S.S.S.R.* **127**:2, 242–245 (1959). (Russian.)

190. Veblen, O., *Analysis Situs,* Am. Math. Soc. Cambridge Colloquium Publications, 1931.

191. Wang, K. T., *On a New Method for the Analysis of Networks,* Natl. Res. Inst. for Engineering, Academia Sinica, Memoir no. 2, 1–11 (1934).

192. Weinberg, L., Kirchhoff's Third and Fourth Laws, *Trans. Inst. Radio Engrs.* **CT–5,** 8–30 (March 1958).

193. Weinberg, L., *Linear Network Synthesis.* New York: McGraw-Hill (to be published).

194. Whitney, H., Congruent Graphs and Connectivity of Graphs, *Am. J. Math.* **54,** 150–168 (1932).

195. Whitney, H., Non-separable and Planar Graphs, *Trans. Am. Math. Soc.* **34,** 339–362 (1932).

196. Whitney, H., On the Classification of Graphs, *Am. J. Math.* **55,** 236–244 (1933).

197. Whitney, H., Planar Graphs, *Fund. Math.* **21,** 73–84 (1933).

198. Whitney, H., 2-Isomorphic Graphs, *Am. J. Math.* **55,** 245–254 (1933).

199. Whitney, H., On the Abstract Properties of Linear Dependence, *Am. J. Math.* **57,** 509–533 (1935).

200. Yoeli, M., The Theory of Switching Nets, *Trans. Inst. Radio Engrs.* **CT–6** (special supplement), 152–157 (May 1959).

INDEX

INDEX