Conclusion

The elliptic function transformation (1). is used here for the purpose of locating zeros and poles of a low-pass filter network function. Charts of the type shown in Figs. 7 to 12 may be prepared for any range of application whenever desired. The compactness of the expressions that give the tolerance and other characteristic quantities makes the preparation of these charts which represent a whole group of network functions with many singularities a matter of evaluating only a few terms together with a few rational operations. These charts, after they are prepared, will be very helpful for design purposes. For instance, if a required attenuation beyond twice the cut-off frequency must be greater than 13 db, Fig. 10 indicates that a filter function with the charge arrangement of Fig. 3(b) and values of a and c of 0.810 and 0.673 respectively will satisfy the requirement. The locations of all poles and zeros of this filter are determined in the z-plane. The locations of zeros and poles in the s-plane may be found by applying the inverse transformation.

Feedback Theory—Further Properties of Signal Flow Graphs*

SAMUEL J. MASON†

Summary—**A way to enhance**
Writing gain at a glance.
Dr. Tustin extended
Proof appended.
Examples illustrative
Pray not frustrative.

Background

THERE ARE many different paths to the solution of a set of linear equations. The formal method involves inversion of a matrix. We know, however, that there are many different ways of inverting a matrix: determinantal expansion in minors, systematic reduction of a matrix to diagonal form, partitioning into submatrices, and so forth, each of which has its particular interpretation as a sequence of algebraic manipulations within the original equations. A determinantal expansion of special interest is

$$D = \sum a_{1i}a_{2j}a_{3k} \cdots a_{nx} \tag{1}$$

where a_{mp} is the element in the mth row and pth column of a determinant having n rows, and the summation is taken over all possible permutations of the column subscripts. (The sign of each term is positive or negative in accord with an even or odd number of successive adjacent column-subscript interchanges required to produce a given permutation.) Since the solution of a set of linear equations involves ratios of determinantal quantities, (1) suggests the general idea that a linear system analysis problem should be interpretable as a search for all possible combinations of something or other, and that the solution should take the form of a sum of products of the somethings, whatever they are, divided by another such sum of products. Hence, instead of undertaking a sequence of operations, we can find the solution by looking for certain combinations of things. The method will be especially useful if these combinations have a simple interpretation in the context of the problem.

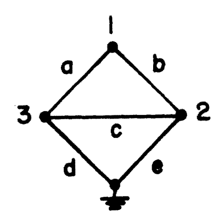

Fig. 1—An electrical network graph.

As a concrete illustration of the idea, consider the electrical network graph shown in Fig. 1. For simplicity, let the branch admittances be denoted by letters a, b, c, d, and e. This particular graph has *three* independent node-pairs. First locate all possible sets of *three* branches which *do not contain closed loops* and write the sum of their branch admittance products as the denominator of (2).

$$Z_{12} = \frac{ab + ac + bc + bd}{abd + abe + acd + ace + ade + bcd + bce + bde} \tag{2}$$

Now locate all sets of *two* branches which *do not form*

* Original manuscript received by the IRE, August 16, 1955; revised manuscript received, February 27, 1956. This work was supported in part by the Army (Signal Corps), the Air Force (Office of Scientific Research, Air Research and Development Command), and the Navy (Office of Naval Research).
† Dept. of Elect. Engrg. and Res. Lab. of Electronics, M.I.T., Cambridge, Mass.

closed loops and which also *do not contain any paths from node 1 to ground or from node 2 to ground*. Write the sum of their branch admittance products as the numerator of expression (2). The result is the transfer impedance between nodes 1 and 2, that is, the voltage at node 2 when a unit current is injected at node 1. Any impedance of a branch network can be found by this process.[1]

So much for electrical network graphs. Our main concern in this paper is with signal flow graphs,[2] whose branches are directed. Tustin[3] has suggested that the feedback factor for a flow graph of the form shown in Fig. 2 can be formulated by combining the feedback loop

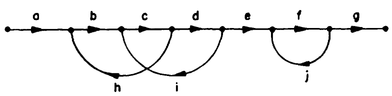

Fig. 2—The flow graph of an automatic control system.

gains in a certain way. The three loop gains are

$$T_1 = bch \tag{3a}$$

$$T_2 = cdi \tag{3b}$$

$$T_3 = fj \tag{3c}$$

and the forward path gain is

$$G_0 = abcdefg. \tag{4}$$

The gain of the complete system is found to be

$$G = \frac{G_0}{[1 - (T_1 + T_2)](1 - T_3)} \tag{5}$$

and expansion of the denominator yields

$$G = \frac{G_0}{1 - (T_1 + T_2 + T_3) + (T_1T_3 + T_2T_3)} \cdot \tag{6}$$

Tustin recognized the denominator as unity plus the sum of all possible products of loop gains taken one at a time $(T_1+T_2+T_3)$, two at a time, $(T_1T_3+T_2T_3)$, three at a time, and so forth, excluding products of loops that touch or partially coincide. The products T_1T_2 and $T_1T_2T_3$ are properly and accordingly missing in this particular example. The algebraic sign alternates, as shown, with each succeeding group of products.

Tustin did not take up the general case but gave a hint that a graph having several different forward paths could be handled by considering each path separately and superposing the effects. Detailed examination of the general problem shows, in fact, that the form of (6) must be modified to include possible feedback factors in the numerator. Otherwise (6) applies only to those graphs in which each loop touches all forward paths.

[1] Y. H. Ku, "Resume of Maxewll's and Kirchoff's rules for network analysis," *J. Frank. Inst.*, vol. 253, pp. 211–224; March, 1952.

[2] S. J. Mason, "Feedback theory—some properties of signal flow graphs," Proc. IRE, vol. 41, pp. 1144–1156; September, 1953.

[3] A. Tustin, "Direct Current Machines for Control Systems," The Macmillan Company, New York, pp. 45–46, 1952.

The purposes of this paper are: to extend the method to a general form applicable to any flow graph; to present a proof of the general result; and to illustrate the usefulness of such flow graph techniques by application to practical linear analysis problems. The proof will be given last. It is tempting to add, at this point, that a better understanding of linear analysis is a great aid in problems of nonlinear analysis and linear or nonlinear design.

A Brief Statement of Some Elementary Properties of Linear Signal Flow Graphs

A signal flow graph is a network of directed branches which connect at nodes. Branch jk originates at node j and terminates upon node k, the direction from j to k being indicated by an arrowhead on the branch. Each branch jk has associated with it a quantity called the branch gain g_{jk} and each node j has an associated quantity called the node signal x_j. The various node signals are related by the associated equations

$$\sum_j x_j g_{jk} = x_k, \qquad k = 1, 2, 3, \cdots \tag{7}$$

The graph shown in Fig. 3, for example, has equations

$$ax_1 + dx_3 = x_2 \tag{8a}$$

$$bx_2 + fx_4 = x_3 \tag{8b}$$

$$ex_2 + cx_3 = x_4 \tag{8c}$$

$$gx_3 + hx_4 = x_5. \tag{8d}$$

We shall need certain definitions. A *source* is a node having only outgoing branches (node 1 in Fig. 3). A *sink* is a node having only incoming branches. A *path* is any continuous succession of branches traversed in the indicated branch directions. A *forward path* is a path from source to sink along which no node is encountered more than once (*abch*, *aeh*, *aefg*, *abg*, in Fig. 3).

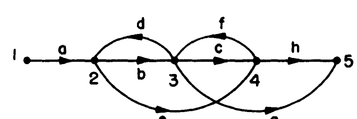

Fig. 3—A simple signal flow graph.

A *feedback loop* is a path that forms a closed cycle along which each node is encountered once per cycle (*bd*, *cf*, *def*, but not *bcfd*, in Fig. 3). A *path gain* is the product of the branch gains along that path. The *loop gain* of a feedback loop is the product of the gains of the branches forming that loop. The *gain of a flow graph* is the signal appearing at the sink per unit signal applied at the source. Only one source and one sink need be considered, since sources are superposable and sinks are independent of each other.

Additional terminology will be introduced as needed.

General Formulation of Flow Graph Gain

To begin with an example, consider the graph shown in Fig. 4. This graph exhibits three feedback loops, whose gains are

$$T_1 = h \tag{9a}$$

$$T_2 = fg \tag{9b}$$

$$T_3 = de \tag{9c}$$

and two forward paths, whose gains are

$$G_1 = ab \tag{10a}$$

$$G_2 = ceb. \tag{10b}$$

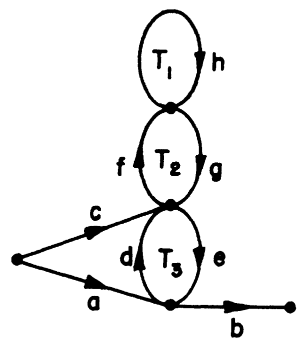

Fig. 4—A flow graph with three feedback loops.

To find the graph gain, first locate all possible *sets of nontouching loops* and write the algebraic sum of their gain products as the denominator of (11).

$$G = \frac{G_1(1 - T_1 - T_2) + G_2(1 - T_1)}{1 - T_1 - T_2 - T_3 + T_1T_3} \tag{11}$$

Each term of the denominator is the gain product of a set of nontouching loops. The algebraic sign of the term is plus (or minus) for an even (or odd) number of loops in the set. The graph of Fig. 4 has no sets of three or more nontouching loops. Taking the loops two at a time we find only one permissible set, T_1T_3. When the loops are taken one at a time the question of touching does not arise, so that each loop in the graph is itself an admissible "set." For completeness of form we may also consider the set of loops taken "none at a time" and, by analogy with the zeroth power of a number, interpret its gain product as the unity term in the denominator of (11). The numerator contains the sum of all forward path gains, each multiplied by a factor. The factor for a given forward path is made up of all possible sets of loops which *do not touch each other* and which also do *not touch that forward path.* The first forward path ($G_1 = ab$) touches the third loop, and T_3 is therefore absent from the first numerator factor. Since the second path ($G_2 = ceb$) touches both T_2 and T_3, only T_1 enters the second factor.

The general expression for graph gain may be written as

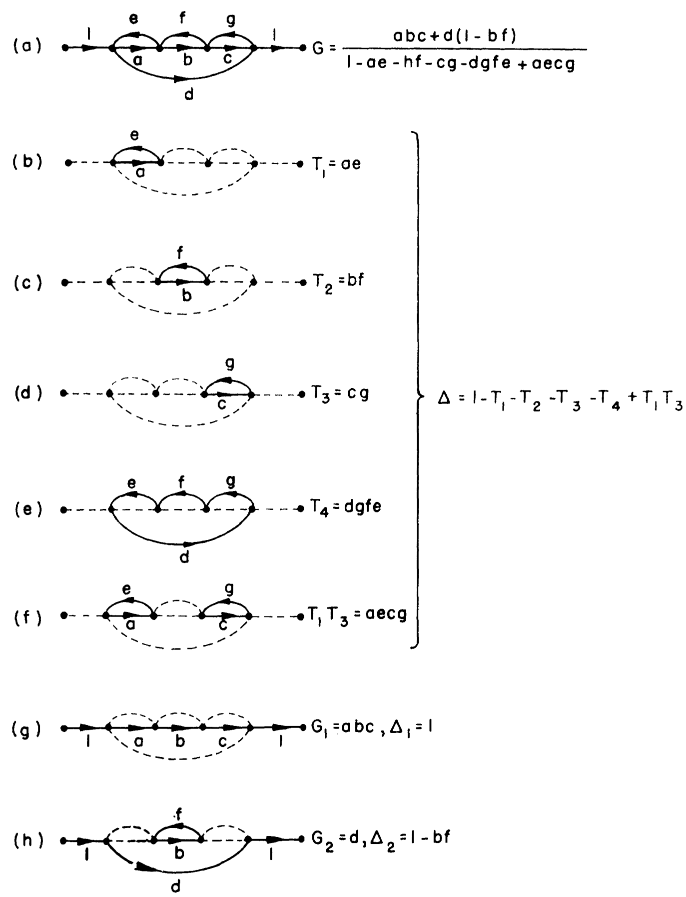

Fig. 5—Identification of paths and loop sets.

$$G = \frac{\sum_k G_k\Delta_k}{\Delta} \tag{12a}$$

wherein

$$G_k = \text{gain of the } k\text{th forward path} \tag{12b}$$

$$\Delta = 1 - \sum_m P_{m1} + \sum_m P_{m2} - \sum_m P_{m3} + \cdots \tag{12c}$$

P_{mr} = gain product of the mth possible combination of r nontouching loops (12d)

Δ_k = the value of Δ for that part of the graph not touching the kth forward path. (12e)

The form of (12a) suggests that we call Δ the *determinant* of the graph, and call Δ_k the *cofactor* of forward path k.

A subsidiary result of some interest has to do with graphs whose feedback loops form nontouching subgraphs. To find the *loop subgraph* of any flow graph, simply remove all of those branches *not* lying in feedback loops, leaving all of the feedback loops, and nothing but the feedback loops. In general, the loop subgraph may have a number of nontouching parts. The useful fact is that the *determinant of a complete flow graph is equal to the product of the determinants of each of the nontouching parts in its loop subgraph.*

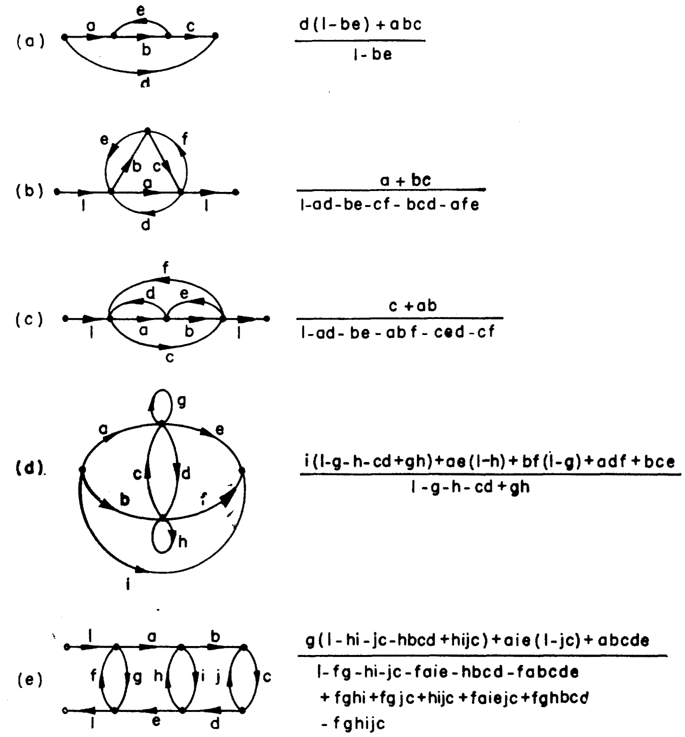

Fig. 6—Sample flow graphs and their gain expressions.

Illustrative Examples of Gain Evaluation by Inspection of Paths and Loop Sets

Eq. (12) is formidable at first sight but the idea is simple. More examples will help illustrate its simplicity. Fig. 5 (on the previous page) shows the first of these displayed in minute detail: (a) the graph to be solved; (b)–(f) the loop sets contributing to Δ; (g) and (h) the forward paths and their cofactors. Fig. 6 gives several additional examples on which you may wish to practice evaluating gains by inspection.

Illustrative Applications of Flow Graph Techniques to Practical Analysis Problems

The study of flow graphs is a fascinating topological game and therefore, from one viewpoint, worthwhile in its own right. Since the associated equations of a linear flow graph are in cause-and-effect form, each variable expressed explicitly in terms of others, and since physical problems are often very conveniently formulated in just this form, the study of flow graphs assumes practical significance.

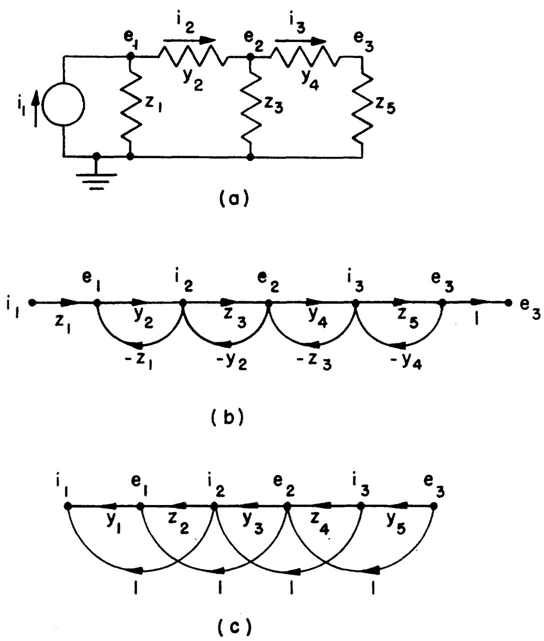

Fig. 7—The transfer impedance of a ladder.

Consider the ladder network shown in Fig. 7(a). The problem is to find the transfer impedance e_3/i_1. One possible formulation of the problem is indicated by the flow graph Fig. 7(b). The associated equations state that $e_1 = z_1(i_1 - i_2)$, $i_2 = y_2(e_1 - e_2)$, and so forth. By inspection of the graph,

$$\frac{e_3}{i_1} = \frac{z_1 y_2 z_3 y_4 z_5}{1 + z_1 y_2 + y_2 z_3 + z_3 y_4 + y_4 z_5 + z_1 y_2 z_3 y_4 + z_1 y_2 y_4 z_5 + y_2 z_3 y_4 z_5} \tag{13a}$$

or, with numerator and denominator multiplied by $y_1 y_3 y_5 = 1/z_1 z_3 z_5$,

$$\frac{e_3}{i_1} = \frac{y_2 y_4}{y_1 y_3 y_5 + y_2 y_3 y_5 + y_1 y_2 y_5 + y_1 y_4 y_5 + y_1 y_3 y_4 + y_2 y_4 y_5 + y_2 y_3 y_4 + y_1 y_2 y_4}. \tag{13b}$$

This result can be checked by the branch-combination method mentioned at the beginning of this paper.

A different formulation of the problem is indicated by the graph of Fig. 7(c), whose equations state that $i_3 = y_5 e_3$, $e_2 = e_3 + z_4 i_3$, $i_2 = i_3 + y_3 e_2$, and so forth. In the physical problem i_1 is the primary cause and e_3 the final effect. We may, however, *choose* a value of e_3 and then calculate the value of i_1 *required* to produce that e_3. The resulting equations will, from the analysis viewpoint, treat e_3 as a primary cause (source) and i_1 as the final effect (sink) *produced* by the chain of calculations. *This does not in any way alter the physical role of* i_1. The new graph (c) may appear simpler to solve than that of (b). Since graph (c) contains no feedback loops, the determinant and path cofactors are all equal to unity. There are many forward paths, however, and careful inspection is required to identify the sum of their gains as

$$\frac{i_1}{e_3} = y_1z_2y_3z_4y_5 + y_1z_2y_3 + y_1z_2y_5 + y_1z_4y_5 + y_3z_4y_5 + y_1 + y_3 + y_5 \tag{13c}$$

which proves to be, as it should, the reciprocal of (13b). Incidentally, graph (c) is obtainable directly from graph (b), as are all other possible cause-and-effect formulations involving the same variables, by the process of path inversion discussed in a previous paper.[2] This example points out the two very important facts: 1) the primary *physical* source does not *necessarily* appear as a source node in the graph, and 2) of two possible flow graph formulations of a problem, the one having fewer feedback loops is not necessarily simpler to solve by inspection, since it may also have a much more complicated set of forward paths.

Fig. 8(a) offers another sample analysis problem, determination of the voltage gain of a feedback amplifier. One possible chain of cause-and-effect reasoning, which leads from the circuit model, Fig. 8(b), to the flow graph formulation, Fig. 8(c), is the following. First notice that e_{g1} is the difference of e_1 and e_k. Next express i_1 as an effect due to causes e_{g1} and e_k, using superposition to write the gains of the two branches entering node i_1. The dependency of e_{g2} upon i_1 follows directly. Now, e_2 would be easy to evaluate in terms of either e_{g2} or i_f if the other were zero, so superpose the two effects as indicated by the two branches entering node e_2. At this point in the formulation e_k and i_f are as yet not explicitly specified in terms of other variables. It is a simple matter, however, to visualize e_k as the superposition of the voltages in R_k caused by i_1 and i_f, and to identify i_f as the superposition of two currents in R_f caused by e_k and e_2. This completes the graph.

The path from e_k to e_{g1} to i_1 may be lumped in parallel with the branch entering i_1 from e_k. This simplification, convenient but not necessary, yields the graph shown in Fig. 9. We could, of course, have expressed i_1 in terms of e_1 and e_k at the outset and arrived at Fig. 9 directly. All simplifications of a graph are themselves

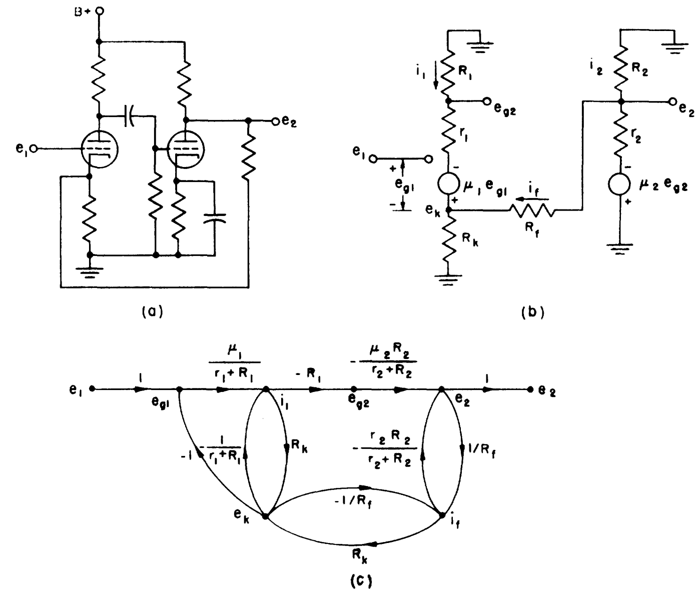

Fig. 8—Voltage gain of a feedback amplifier. (a) A feedback amplifier; (b) The midband linear incremental circuit model; (c) A possible flow graph.

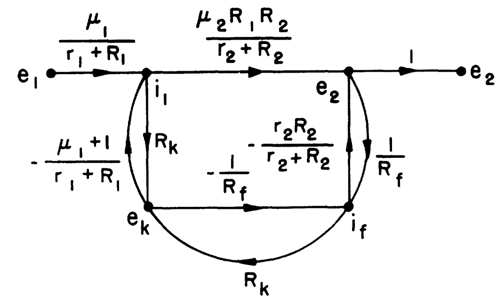

Fig. 9—Elimination of superfluous nodes e_{g1} and e_{g2}.

possible formulations. The better our perception of the workings of a circuit, the fewer variables will we need to introduce at the outset and the simpler will be the resulting flow graph structure.

In discussing the feedback amplifier of Fig. 8(a) it is common practice to neglect the loading effect of the feedback resistor R_f in parallel with R_k, the loading effect of R_f in parallel with R_2, and the leakage transmission from e_k to e_2 through R_f. Such an approximation is equivalent to the removal of the branches from e_k to i_f and i_f to e_2 in Fig. 9. It is sometimes dangerous to make early approximations, however, and in this case no appreciable labor is saved, since we can write the exact answer by inspection of Fig. 9:

$$\frac{e_2}{e_1} = \frac{\dfrac{\mu_1\mu_2R_1R_2}{(r_1+R_1)(r_2+R_2)}\left[1+\dfrac{R_k}{R_f}\right] + \dfrac{\mu_1R_kr_2R_2}{(r_1+R_1)(R_f)(r_2+R_2)}}{1+\dfrac{(\mu_1+1)R_k}{r_1+R_1}+\dfrac{R_k}{R_f}+\dfrac{r_2R_2}{R_f(r_2+R_2)}+\dfrac{(\mu_1+1)\mu_2R_1R_kR_2}{R_f(r_1+R_1)(r_2+R_2)}+\dfrac{(\mu_1+1)R_kr_2R_2}{R_f(r_1+R_1)(r_2+R_2)}} \tag{14}$$

The two forward paths are $e_1 i_1 e_2$ and $e_1 i_1 e_k i_f e_2$, the first having a cofactor due to loop $e_k i_f$. The principal feedback loop is $i_1 e_2 i_f e_k$ and its gain is the fifth term of the denominator. Physical interpretations of the various paths and loops could be discussed but our main purpose, to illustrate the formulation of a graph and the evaluation of its gain by inspection, has been covered.

As a final example, consider the calculation of microwave reflection from a triple-layered dielectric sandwich. Fig. 10(a) shows the incident wave A, the reflection B, and the four interfaces between adjacent regions of different material. The first and fourth interfaces, of course, are those between air and solid. Let r_1 be the reflection coefficient of the first interface, relating the incident and reflected components of tangential electric field. It follows from the continuity of tangential E that the interface transmission coefficient is $1+r_1$, and from symmetry that the reflection coefficient from the opposite side of the interface is the negative of r_1. A suitable flow graph is sketched in Fig. 10(b). Node signals along the upper row are right-going waves just to the left or right of each interface, those on the lower row are left-going waves, and quantities d are exponential phase shift factors accounting for the delay in traversing each layer.

Apart from the first branch r_1, the graph has the same structure as that of Fig. 6(e). Hence the reflectivity of the triple layer will be

$$\frac{B}{A} = r_1 + (1 + r_1)(1 - r_1)G \tag{15}$$

where G is in the same form as the gain of Fig. 6(e). We shall not expand it in detail. The point is that the answer can be written by inspection of the paths and loops in the graph.

Proof of the General Gain Expression

In an earlier paper[2] a quantity Δ was defined as

$$\Delta = (1 - T_1')(1 - T_2') \cdots (1 - T_n') \tag{16}$$

for a graph having n nodes, where

T_k' = loop gain of the kth node as computed with all higher-numbered nodes split.

Splitting a node divides that node into a new source and a new sink, all branches entering that node going with the new sink and all branches leaving that node going with the new source. The loop gain of a node was defined as the gain from the new source to the new sink, when that node is split. *It was also shown that Δ, as computed according to (16), is independent of the order in which the nodes are numbered, and that consequently Δ is a linear function of each branch gain in the graph. It follows that Δ is equal to unity plus the algebraic sum of various branch-gain products.*

We shall first show that each term of Δ, other than the unity term, is a product of the gains of nontouching

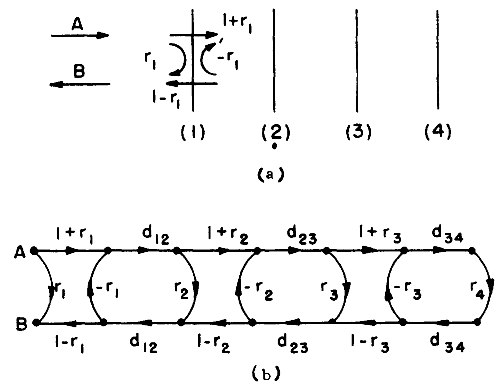

Fig. 10—A wave reflection problem. (a) Reflection of waves from a triple-layer; (b) A possible flow graph.

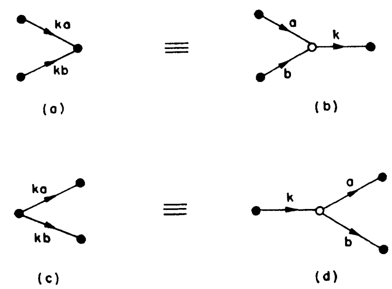

Fig. 11—Two touching paths.

feedback loops. This can be done by contradiction. Consider two branches which either enter the same node or leave the same node, as shown in Fig. 11(a) and (c). Imagine these branches imbedded in a larger graph, the remainder of which is not shown. Call the branch gains ka and kb. Now consider the equivalent replacements (b) and (d). The new node may be numbered zero, whence $T_0' = 0$, the other T' quantities in (16) are unchanged, and Δ is therefore unaltered. If both branches ka and kb appear in a term of the Δ of graph (a) then the square of k must appear in a term of the Δ of graph (b). This is impossible since Δ must be a linear function of branch gain k. Hence no term of Δ can contain the gains of two touching paths.

Now suppose that of the several nontouching paths appearing in a given term of Δ, some are feedback loops and some are open paths. Destruction of all other branches eliminates some terms from Δ but leaves the given term unchanged. It follows from (16) and the definitions of T_k', however, that the Δ for the subgraph containing only these nontouching paths is just

$$\Delta = (1 - T_1)(1 - T_2) \cdots (1 - T_m) \tag{17}$$

where T_k is the gain of the kth feedback loop in the subgraph. Hence the open path gains cannot appear in the given term and it follows that each term of Δ is the product of gains of nontouching feedback loops. Moreover, it is clear from the structure of Δ that a term in any subgraph Δ must also appear as a term in the Δ of the complete graph, and conversely, every term of Δ is a term of some subgraph Δ. Hence, to identify all possible terms in Δ we must look for all possible subgraphs comprising sets of nontouching loops. Eq. (17) also shows that the algebraic sign of a term is plus or minus in accord with an even or odd number of loops in that term. This verifies the form of Δ as given in (12c) and (12d).

We shall next establish the general expression for graph gain (12a). The following notation will prove convenient. Consider the graph shown schematically in Fig. 12, with node $n+1$ given special attention. Let

$\Delta' =$ the Δ for the complete graph of $n+1$ nodes.

$\Delta =$ the value of Δ with node $n+1$ split or removed.

$T =$ the loop gain of node $n+1$.

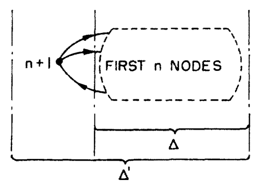

Fig. 12—A flow graph with one node placed strongly in evidence.

There will in general be several different feedback loops containing node $n+1$. Let

$T_k =$ gain of the kth feedback loop containing node $n+1$,

$\Delta_k =$ the value of Δ for that part of the graph not touching loop T_k.

With the above notation, we have from (16) that

$$1 - T = \frac{\Delta'}{\Delta} \cdot \tag{18}$$

Remembering that any Δ is the algebraic sum of gain products of nontouching loops, we find it possible to write

$$\Delta' = \Delta - \sum_k T_k \Delta_k. \tag{19}$$

Eq. (19) represents the count of all possible nontouching loop sets in Δ'. The addition of node $n+1$ creates new loops T_k but the only new loop sets of Δ' not already in Δ are the nontouching sets $T_k\Delta_k$. The negative sign in (19) suffices to preserve the sign rule, since the product of T_k and a positive term of Δ_k will contain an odd number of loops.

Substitution of (19) into (18) yields the general result:

$$T = \frac{\sum_k T_k \Delta_k}{\Delta} \cdot \tag{20}$$

With node $n+1$ permanently split, T is just the source-to-sink gain of the graph and T_k is the kth forward path. This verifies (12a).

Acknowledgment

The writer is indebted to Prof. K. Wildes and also to J. Cruz, L. Boffi, G. Amster, and other students in his 1955 summer-term class in subject 6.633, Electronic Circuit Theory, M.I.T.; for helpful suggestions.

Correction

James R. Wait, author of the Correspondence item "The Radiation Pattern of an Antenna Mounted on a Surface of Large Radius of Curvature," which appeared on page 694 of the May, 1956 issue of Proceedings, has requested that the following text, modified in editing, be reinstated in its original form.

The last sentence in the first paragraph should read:

"It is the purpose of the present note to extend and apply the Van der Pol-Bremmer theory to calculate the radiation pattern of a dipole or a slot on a conducting sphere of large radius."

The last sentence of the article should read:

"It is interesting to compare this value with the 6 db field strength reduction in the tangent plane from a slot on a flat ground plane which is abruptly truncated."

Mr. Wait has also informed the editors that in (2), the second bracketed term should be $zh_\nu^{(1)}(z)$.