THE SYNTHESIS OF VOLTAGE TRANSFER FUNCTIONS

PHILIP M. LEWIS II

TECHNICAL REPORT 314

JUNE 7, 1956

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the U. S. Army (Signal Corps), the U. S. Navy (Office of Naval Research), and the U. S. Air Force (Office of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039-sc-64637, Department of the Army Task No. 3-99-06-108 and Project No. 3-99-00-100.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS

Technical Report 314 June 7, 1956

THE SYNTHESIS OF VOLTAGE TRANSFER FUNCTIONS

Philip M. Lewis II

This report is based on a thesis submitted to the Department of Electrical Engineering, M.I.T., May 14, 1956, in partial fulfillment of the requirements for the degree of Doctor of Science.

Abstract

The synthesis of voltage transfer functions (E_2/E_1), in the form of linear, lumped, finite, passive, bilateral networks containing no ideal transformers or mutual coupling, is considered. The basic realizability conditions are derived and realization procedures are developed based on these conditions, showing them to be both necessary and sufficient.

This particular class of networks places constraints on the allowable values of the constant multiplier in the voltage transfer function and on the positions of the transmission zeros in the case of grounded networks. In order to study these constraints, a new concept – the concept of the one – is introduced. This concept gives a certain physical significance to the constant multiplier, which allows the basic realizability conditions to be derived in a simple fashion. These conditions are:

1. The maximum gain obtainable from a given pole-zero plot independent of configuration.
2. The maximum gain obtainable from a given pole-zero plot for a given complexity of the network configuration.
3. The maximum gain obtainable from a given pole-zero plot for a realization in terms of a symmetrical network.
4. The general conditions under which a grounded (three-terminal) realization of any transfer function is possible.

Various synthesis procedures are developed based on these realizability conditions:

1. A general lattice synthesis in the two-element and three-element cases.
2. A general method of realizing grounded symmetrical networks, by first realizing a lattice and then unbalancing it.
3. A general realization procedure yielding any allowable gain in both the two-element and the three-element cases.

Table of Contents

INTRODUCTION

1. THE SYNTHESIS APPROACH

At one time, among engineers, the word synthesis meant the opposite of the word analysis. Analysis referred to finding the response of a system to a given excitation, while synthesis referred to finding the system that corresponds to a given input and response. In a restricted sense, these definitions are still valid, but in recent years the scope of the synthesis field has broadened and the methods of reasoning and the philosophy of network synthesis have come to be used in other branches of engineering, and a new approach to engineering problems has evolved.

One of the main attributes of the synthesis approach to engineering problems is the separation of the approximation part of the problem from the realization part. Every practical engineering problem is first stated in ideal form: we should like a filter that would be perfectly flat out to one megacycle and exactly zero thereafter; we should like a filter that would perfectly separate signal from noise; we should like an antenna that would radiate exactly in one direction only; we should like a guided missile that would be one hundred per cent accurate. The man who states the problems knows that these ideal specifications cannot be met, but he would like to know how close he can come to them and for what price.

In order to answer this question some sort of approximation or compromise must be made. But before the approximation problem can be solved, the engineer must have at his disposal detailed knowledge of the exact types of behavior that are possible for the physical systems involved and he must also know which mathematical functions characterize this behavior. The establishment of these realizability conditions belongs properly, not to the approximation part of the problem at all, but to the realization part. However, the conditions themselves are needed to perform the approximations.

By using these realizability conditions, a suitable approximation can then be made to the ideal specifications for the system. Very often, as an aid to the approximation, certain error criteria are applied, such as maximally flat, equiripple, minimum mean-square, and so forth; in other cases a cut-and-try procedure is employed. The approximation problem is solved when a mathematical function is found that is both a close enough approximation to the desired specifications and also a representation of the behavior of a physically realizable system. The physical system represented by this function is then realized exactly, using the methods pertinent to the realization part of the problem. This approach to engineering problems is a significant departure from conventional engineering methods, which attempt to lump the approximation and realization parts of the problem under the one category—design.

That which we have called the realization part of the problem is really the portion of synthesis that is the opposite of analysis. However, it is common practice not to make this distinction; indeed the words "realization" and "synthesis" are very often used synonymously and the approximation is considered to be a separate entity. In the subsequent parts of this report, we shall not be concerned with the approximation problem; therefore we shall use the common term, "synthesis", for the realization part of the problem.

Over the years, a classic approach to the synthesis problem has been developed. This approach is patterned after the work of Foster (1) on lossless driving-point impedance synthesis. It consists of the following steps.

1. A class of systems is defined.

2. A study of this class of systems is made to determine exactly what types of behavior are possible and impossible for such a system. Mathematicians refer to this step as finding necessary conditions for the realizability of the class.

3. Realization methods are developed that depend on some of the conditions derived in step 2, thus showing these conditions to be sufficient as well as necessary.

2. VOLTAGE TRANSFER FUNCTIONS

The class of systems considered in this report consists of linear, lumped, finite, passive, bilateral, electrical networks containing no ideal transformers or mutual reactance. (An additional restriction will be added, namely, that RLC networks have no axis poles. This case is considered in detail by Fialkow and Gerst (15, 16); it appears to add little but complication to the discussion.) The type of input-output relation that will be studied is the voltage transfer function, E_2/E_1. These particular physical constraints were chosen, even though a voltage transfer function may be impracticable to realize and use unless a series resistor is available, because they appear to be the only set of conditions that allow a simple physical interpretation to be given to the constant multiplier. Such an interpretation will be developed in Section I, after which it will be possible (in Section II) to derive some theorems concerning the ultimate limits of performance that can be obtained when a particular configuration (of the class considered) is used to realize a given voltage transfer function and to develop some realization methods based on these theorems (Section III–VI).

3. HISTORICAL BACKGROUND

The problem of the design of a network with prescribed input and output functions is probably as old as the field of Electrical Engineering, but the set of procedures called "Network Synthesis" has only been developed in the past twenty-five or thirty years. The basic procedures for driving-point synthesis were published between 1924 and 1939 by Foster (1), Cauer (2), Brune (3), Darlington (4), and later by Bott and Duffin (5). The first work on transfer synthesis was done by Gewertz (6) in 1932. He derived some basic theorems on transfer impedance synthesis and gave some synthesis procedures that were, in many cases, unwieldy and difficult to use.

In 1939, Darlington (4) proposed a method of transfer synthesis in terms of a lossless network, loaded at one or both ends with a resistor. Also in 1939, Cauer (7) showed an alternative procedure for the design of lossless networks terminated in a resistor. These methods are still among the basic procedures for transfer synthesis, although their use, in practice, is usually limited to the synthesis of functions that have all their zeros on the j-axis.

In the years after 1939, many papers were written on various aspects of the problem. In 1944, Guillemin (8) showed that any RC transfer function can be realized as a lattice. Bower and Ordung (9), in 1950, gave other RC synthesis procedures in terms of lattices; they mention a maximum level of transmission which is related to the maximum gain of a lattice. In 1949 Guillemin (10) showed that any RC transfer function with positive coefficients can be realized as the parallel combination of ladders. Weinberg (11), in 1951, derived various synthesis procedures for RC and RLC networks. Also in 1951, Dasher (12) proposed a method of RC synthesis in terms of zero sections; this method was recently improved by Guillemin (13) and extended to the LC case.

Prior to 1952, much of the work on voltage transfer functions consisted of the realization of transfer impedance or admittance terminated at one or both ends with resistors. In 1952, Fialkow and Gerst (14) published a paper that dealt with voltage transfer synthesis as a separate entity, and derived some conditions for maximum gain of RC networks, as well as procedures for their realization. About a year later, the work which led to this report was begun and, since then, many of the basic realizability theorems have been derived almost simultaneously (14–20), although independently and in slightly different forms. In addition, realization methods based on these theorems have also been derived almost simultaneously, although here the duplication is not so apparent, because, since the realizability theorems were obtained in slightly different forms, the realizations that were attempted were also different. However, the realization methods derived by Fialkow and Gerst have been included, whenever they are necessary for the sake of continuity. In several places, their methods are presented in slightly different forms that simplify the realization or require less elements.

It is not the intention of this work to claim credit or dispute about any of the material that was also derived by Failkow and Gerst. Here a new approach to the problem is presented, which has a certain physical significance and yields the basic theorems and realizations in a simplified form.

I. THE CONCEPT OF THE ONE

1. INTRODUCTION

The voltage transfer function of a two terminal-pair network is defined as the ratio of the output voltage that appears across terminal pair two to the input voltage applied at terminal pair one. If the network is linear and the input is of the form

$$e_1 = E_1 e^{st}$$

the output voltage will be of the form

$$e_2 = E_2 e^{st}$$

and the voltage transfer function is defined as the ratio of the vector E_2 to the vector E_1. If, in addition, the network is lumped, finite, and bilateral, the voltage transfer ratio is a rational function of the complex frequency s:

$$A = \frac{E_2}{E_1} = K \frac{s^n + a_{n-1} s^{n-1} + -- a_1 s + a_o}{s^m + b_{m-1} s^{m-1} + -- b_1 s + b_o}$$

$$= \frac{KN}{D} = K \frac{(s - s_1)(s - s_2) -- (s - s_n)}{(s - s_I)(s - s_{II}) -- (s - s_m)}$$

Such a rational function is described completely by three sets of quantities, its poles, its zeros, and its constant multiplier K. If additional constraints are placed on the network, then restrictions appear on the allowable values for these poles, zeros, and constant multiplier. If, for instance, the network is passive, certain restrictions (described below) appear on the allowable positions of the poles.

These conditions—that the network be linear, lumped, finite, passive, and bilateral—define the class of networks usually studied in network theory. For such networks, the voltage transfer function can be expressed in terms of the impedance and admittance functions:

$$A = \frac{y_{12}}{y_{22}} = \frac{z_{12}}{z_{11}}$$

and the properties of the voltage transfer function can be derived from the known properties of impedances and admittances. These voltage transfer properties are well known and are repeated here for convenience and reference.

2. PROPERTIES OF VOLTAGE TRANSFER FUNCTIONS

a. RC and RL Networks:

1. All poles are simple and on the negative real axis.
2. All residues are real but can be positive or negative.
3. Zeros can be anywhere in the s-plane, but complex zeros must be in conjugate pairs.
4. Zero and infinite frequency cannot be poles.

b. LC Networks:

1. All poles are simple and on the j-axis.

2. Residues are pure imaginary but can be positive or negative. The proof of this last condition follows from the fact that A is the ratio of an LC transfer impedance to an LC driving-point impedance and is, therefore, an even polynomial over an even polynomial or an odd polynomial over an odd polynomial.

3. Zeros can be anywhere in the s-plane, but complex zeros off the j-axis must occur in quadruplets, and real zeros must occur in left-right half-plane pairs.

4. Zero and infinite frequency cannot be poles.

c. RLC Networks:

1. All the poles are in the left half-plane.

2. J-axis poles must be simple with pure imaginary residues. The fact that the residues must be imaginary can be seen from the following reasoning. Since $A = z_{12}/z_{11}$, it can only have a j-axis pole when z_{11} has a j-axis zero. However, the realizability conditions for the impedance functions state that on the j-axis, $R_{11}R_{22} - R_{12}^2 \geqslant 0$. Thus when z_{11} has a j-axis zero, the real part of z_{12} must be zero, or z_{12} must be imaginary; the above statement follows directly. However, we are not interested here in RLC functions with j-axis poles, so that this condition will not be used in the following discussion.

3. Zeros can be anywhere in the plane, but complex zeros must be in conjugate pairs.

4. Zero and infinite frequency cannot be poles.

If these restrictions—that the network be linear, lumped, finite, passive, and bilateral—are the only ones applied to the problem of voltage transfer functions, then nothing more can be said about the allowable values of the zeros and the constant multiplier. In fact, the constant multiplier can be adjusted to any value with an ideal transformer and the transmission zeros can lie anywhere in the s-plane. If, however, an additional constraint is placed on the problem—that the networks have no ideal transformers or mutual coupling—then the network places additional constraints on the values of the constant multiplier and on the positions of the transmission zeros in the s-plane.

3. DEFINITION OF THE ONE

In order to study these constraints in more detail, it is convenient to introduce the concept of the "one". A function is said to have a one at some particular value of the complex frequency if it equals exactly one at angle zero. (A function can also have a minus one if it equals one at an angle of 180°.) To solve for the ones of a particular function it is only necessary to find the zeros of the expression

$$1 - A = 0$$

$$1 - \frac{KN}{D} = 0$$

$$D - KN = 0$$

This equation can be solved analytically or by using the methods of root locus. The latter methods are particularly helpful in obtaining general properties of the solution.

Since the degree of N cannot be greater than the degree of D, this equation has a number of roots just equal to the degree of the polynomial D. Thus the number of ones of a function is just equal to the number of its poles. The positions of these ones in the s-plane are a function of the parameter K. As K is increased from zero, the ones move from the poles toward the zeros along lines of zero phase angle (since the above equation can only have a root when A is pure real).

Similar reasoning shows that as K is increased from zero, the minus ones move from the poles toward the zeros along lines of 180° phase angle, and their positions for any particular value of K are given by the equations,

$$1 + A = 0$$

$$1 + \frac{KN}{D} = 0$$

$$D + KN = 0$$

EXAMPLE.

$$A = \frac{KN}{D} = \frac{Ks}{(s+1)(s+3)} = \frac{Ks}{s^2 + 4s + 3}$$

The ones of this function are the roots of the equation

$$D - KN = 0$$

$$s^2 + (4 - K)\, s + 3 = 0$$

while the minus ones are the roots of the equation

$$D + KN = 0$$

$$s^2 + (4 + K)\, s + 3 = 0$$

Since these polynomials are quadratic, the loci of the roots can be drawn by inspection; they are plotted in Fig. 1. The dashed line is the locus of ones and the dot-and-dash line is the locus

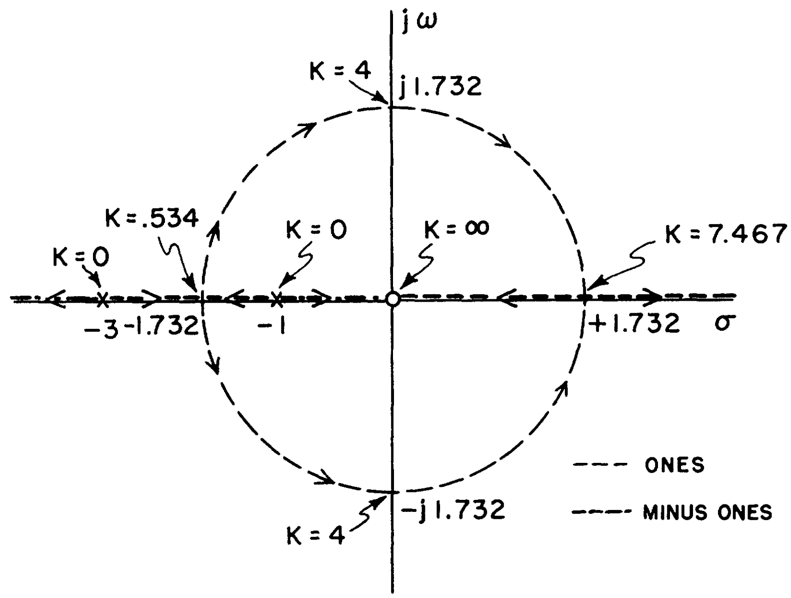

Fig. 1. Root locus for ones and minus ones.

of minus ones. The appropriate values of K are added in several places to the locus of ones.

4. ADDITIONAL PROPERTIES OF THE ROOT LOCUS

Some additional properties of the locus of the ones and minus ones are given below. Many of these properties are already apparent from the preceding example.

a. As K is increased from zero, the ones leave the (simple) poles at an angle (measured from a line parallel to the positive real axis) just equal to the angle of the residue of A in that pole, while the minus ones leave at 180° plus that angle.

COROLLARY 1. Since the residue in any simple, negative real axis pole must be pure real, the one leaving that pole (as K is increased from zero) will stay on the axis and move around until it meets another one, after which the two ones will leave the axis and become a pair of complex ones. Thus, in the case of RC and RL networks, which have all their poles on the negative real axis, there is a value of K below which all the ones and minus ones are on the negative real axis and above which some of the ones or minus ones are off the axis.

Corollaries 1 and 2 are stated for the ones of the function, but the same results apply to the the minus ones.

COROLLARY 2. A similar statement can be made concerning the ones of LC networks. (If A has any j-axis poles this statement may still be true under certain conditions, if the residues are pure imaginary. These conditions are given in (15, 16) and will not be needed in the present discussion.) Since these networks have all simple zeros on the j-axis with all imaginary residues, an exactly analogous situation exists. As K is increased from zero, each one stays on the j-axis until it meets another one, then they separate and move into the left and right half-plane as a quadruplet (together with the conjugate pair). Thus, for LC networks, there is a value of K below which all the ones and minus ones are on the j-axis and above which some of the ones or minus ones are off the axis.

COROLLARY 3. A somewhat similar statement can also be made concerning RLC networks that contain no j-axis poles. (See Corollary 2.) There is a value of K below which all the ones and minus ones are in the left half-plane and above which some of the ones or minus ones are in the right half-plane. The ones or minus ones can cross into the right half-plane in three ways: by passing through zero frequency; by passing through infinite frequency; and by crossing over the j-axis at some finite nonzero frequency.

b. Since A is pure real on the whole positive real axis of the s-plane, the following statements can be made concerning the behavior of the locus near this axis.

1. The whole positive real axis is part of the one or the minus one locus.

2. There is some value of K below which no ones or minus ones are on the axis and above which some of the ones or minus ones are on the axis.

3. As K is increased, the ones (or minus ones) can approach the positive real axis in three ways: by passing through zero frequency; by passing through infinite frequency; and by a pair of conjugate ones or minus ones crossing into the right half-plane at some finite nonzero j-axis point, and later (at some higher value of K) approaching the positive real axis.

5. METHODS OF PRODUCING ONES

In general, then, these loci extend over the entire s-plane. However, if we apply the restriction that the given function is the voltage transfer ratio of a network that contains no ideal transformers or mutual coupling, then certain constraints exist on the allowable positions of the ones. In order to study these constraints, we must study the mechanisms by which ones can be produced in such networks. There are only two ways in which ones can be produced: if a certain circuit element or combination of circuit elements has a pole or a zero that effectively connects the corresponding input and output terminals or, if there is cancellation, that is, if the transmission just happens to be one, without any circuit elements that have poles or zeros. (The word "cancellation" is used to describe the condition in which ones or minus ones are not produced by elements that have poles or zeros in order to bring out the analogy between ones and transmission zeros, which is discussed in a later section.) Examples will now be given of networks that have both types of one.

EXAMPLE 1. Consider the ELL-network of Fig. 2. The voltage transfer function of this network can have a one only if Z_a is a short circuit (has a zero) or Z_b is an open circuit (has a pole). Thus all the ones of this network arise from elements having zeros or poles, and these ones must therefore be in the left half-plane for RLC networks, on the j-axis for LC networks and on the negative real axis for RC or RL networks. This determines the maximum value of K for this type of network.

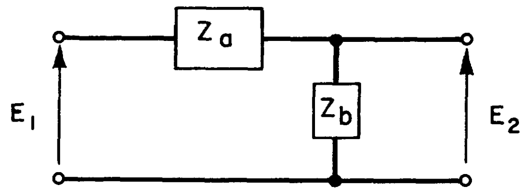

Fig. 2. ELL-network.

EXAMPLE 2. In the lattice network of Fig. 3, all the ones and minus ones are caused by elements that have zeros or poles. The ones can occur only when Z_a has a zero or Z_b has a pole, and the minus ones can occur only when Z_a has a pole or Z_b has a zero. This time, all the ones and minus ones must be in the left half-plane for RLC networks, on the j-axis for LC networks and on the negative real axis for RC or RL networks; again this determines the maximum gain for the lattice (and incidentally for all symmetrical networks).

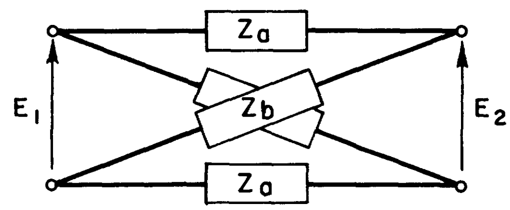

Fig. 3. A lattice.

EXAMPLE 3. Consider the ladder network of Fig. 4, which has the voltage transfer function

$$A = \frac{5/2\,s}{(s + 1)\,(s + 3)}$$

Since the denominator is quadratic, there are two ones, and these are just the roots of the polynomial

$$D - KN = 0$$

$$s^2 + (4 - 2.5)\,s + 3 = 0$$

$$(s + 0.75 - 22.9)\,(s + 0.75 + 22.9) = 0$$

Since these ones are complex, they plainly are not caused by RC circuit elements that have poles or zeros and are, therefore, (by definition) caused by cancellation.

It is not obvious at this point that the network puts any constraints on the positions of ones caused by cancellation; however, these restrictions do exist and are derived in Section II.

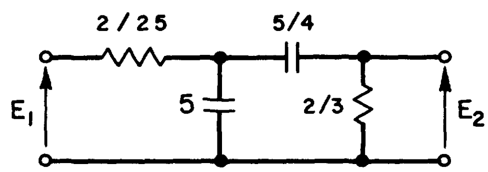

Fig. 4. A ladder.

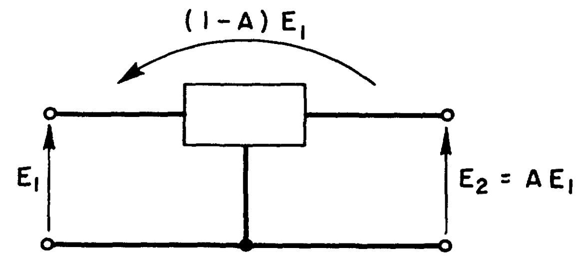

Fig. 5. Grounded configuration.

6. RELATION OF ONES TO TRANSMISSION ZEROS

There is an interesting analogy between the mechanisms by which ones and transmission zeros are produced. It is well known that there are two ways in which transmission zeros can occur: if a certain circuit element or combination of circuit elements has a pole or zero which effectively isolates the input of the network from the output or, if there is cancellation caused by the transmission through two or more parallel paths just happening to add up to zero. The analogy with ones is apparent. It is also well known that, because of the reasons given above, certain network configurations can have zeros of transmission only in certain portions of the s-plane. For example, the zeros of the ELL-network of Fig. 2 occur only where Z_a has a pole or Z_b has a zero and are, therefore, constrained to certain parts of the s-plane by the type of element (RLC, LC, RC, RL) used in the synthesis, while the zeros of the lattice of Fig. 3 arise from cancellation caused by the balancing of a bridge ($Z_a = Z_b$) and can be anywhere in the s-plane. These constraints on the positions of zeros are well known and are clearly analogous to those for the ones.

In the special case of a grounded network, this analogy becomes a much closer relationship. The ones and transmission zeros of a grounded network can be interchanged by a change of reference terminals. (See Fig. 5.) Thus the general grounded network can have transmission zeros only where it can have ones and vice versa. Any restriction on the positions of two ones is also a restriction on the positions of the zeros.

A word of caution is in order; the above statements refer to general grounded networks. Particular network configurations may add additional constraints, which may restrict the ones and/or transmission zeros to particular parts of the s-plane. For example, the configuration of the RC ladder used in the previous example restricts the transmission zeros to the negative real axis, while no such restriction exists for the ones. The above statements refer to general grounded networks with unspecified configurations.

II. BASIC LIMITATIONS ON NETWORK BEHAVIOR

1. INTRODUCTION

The concept of the one places in evidence various constraints on network performance. (Hereafter, the word "network" refers to linear, lumped, finite, passive, bilateral networks that contain no ideal transformers or mutual coupling.) Certain of these constraints apply to particular network configurations, as demonstrated in the examples of Section I, while others apply to all networks independently of their configuration. This second type is, in a sense, more fundamental in that it gives the basic limitations on the behavior that can be obtained from physical networks. Such limitations are of importance in determining exactly which transfer functions can be realized and which cannot.

In this section these limitations are derived in the form of four theorems that state necessary properties of physical networks. In subsequent sections synthesis procedures are derived based on these theorems, thus showing them to be sufficient as well as necessary.

THEOREM 1. No network can have a one or minus one on the positive real axis except at zero or infinite frequency where the gain can be one but no greater (17). (An obvious exception to this theorem is any network that has a gain identically equal to one.)

PROOF. On the positive real axis, all the circuit elements behave like positive resistors whose values depend on σ. Clearly, such a "resistive" network cannot have a gain greater in magnitude than one. Moreover, the gain can only equal one or minus one if some of the resistors have zero or infinite resistance, but the "resistors" considered here can only have zero or infinite resistance at zero or infinite frequency.

This theorem gives the maximum value of the constant multiplier that can be obtained from a given pole-zero constellation no matter what network configuration is used in the synthesis. As shown in the previous section, there is some value of K below which no ones or minus ones are on the positive real axis and above which some of the ones or minus ones are on this axis; this is the maximum value of K.

2. TEST FOR MAXIMUM K

A graphical statement of the above theorems can be given as follows:

> The curve of A versus σ must lie between the lines $A = 1$ and $A = -1$ everywhere on the positive real axis; the curve can only touch one of the lines at zero or infinite frequency.

At the maximum value of K, there is some frequency at which the curve just touches one of the lines. This can occur at zero frequency, infinite frequency or at any finite nonzero value of σ. If it occurs at zero or infinite frequency, the curve may cross through the $A = 1$ or $A = -1$ line, but if it occurs at any other value of σ, the curve must not cross the line and must be passing through one of its maxima or minima. Thus the frequency at which the curve touches the line is either zero or infinity or one of the frequencies at which A has one of its extrema.

These properties form the basis of the test for maximum K. First the frequencies at which all the maxima and minima occur are found by setting the first derivative of A equal to zero. Then the transfer function is evaluated at each of these frequencies, as well as at zero and infinite frequencies, and the value of K is chosen that makes the magnitude of A equal to one at one of these frequencies and less than one at all of the others.

To determine the frequencies at which the function has its maxima and minima, the derivative is set equal to zero.

$$\frac{d}{ds}\left(\frac{KN}{D}\right) = K\left(\frac{DN' - ND'}{D^2}\right) = 0$$

$$DN' - ND' = 0$$

where the prime denotes differentiation with respect to s. The frequencies of interest are the positive real roots of this equation. The actual factoring of the polynomial must be done by using Newton's method or some other approximate method. But before beginning one of these lengthy procedures, various tests can be used to determine the number of positive real roots:

a. Descartes' rule of signs states that the number of positive real roots of a polynomial is not greater than the number of sign variations in its coefficients. A particular case of interest is one in which there are no sign variations. In this case, no positive real roots are present and the transfer function goes through one at zero or infinite frequency.
b. Sturm's test can be used to determine the exact number of positive real roots and to help localize those roots.

EXAMPLE. Consider the voltage transfer function

$$A = K\,\frac{s\,(s-1)}{(s+1)\,(s^2+1)}$$

This function has positive real axis zeros at zero frequency, infinite frequency, and at the frequency s = 1. A sketch of the function (see Fig. 6) shows that there are at least two extrema

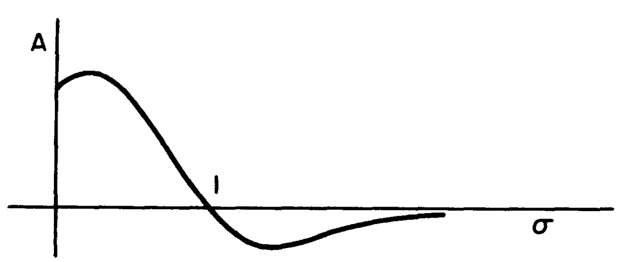

Fig. 6. Sketch of a transfer function on the positive real axis.

on the positive real axis, one between s = 0 and s = 1, and one above s = 1. The polynomial DN′ − ND′ is found to be:

$$DN' - ND' = s^4 - 2s^3 - 2s^2 + 1 = 0$$

Descartes' rule of signs shows that there are no more than two positive real roots. Thus there must be exactly two such roots. The factors of the polynomial are found to be:

$$DN' - ND' = (s - 2.65)\,(s - 0.598)\,(s^2 + 1.255 + 0.621)$$

The voltage transfer function is then evaluated at each of the positive real extrema:

At s = 0.548	A = 0.119 K
At s = 2.65	A = −0.15 K

The maximum K is then:

$$K < \frac{1}{0.15} = 6.67$$

The strict less than sign is used, because, as shown later, this maximum can be approached but never attained.

THEOREM 2. A grounded network cannot have a transmission zero on the positive real axis except at zero or infinite frequency (17).

PROOF. This can be proved by using the relationship between ones and transmission zeros of grounded networks or by using the same reasoning as in the last proof. On the positive real axis all the circuit elements look like positive resistors. Such a grounded resistive network cannot have transmission zeros unless some of the resistors have zero or infinite resistance, but, again, these resistors can only have zero or infinite resistance at zero or infinite frequency.

COROLLARY. Since grounded networks have no zeros on the positive real axis, the transfer function must be either everywhere positive or everywhere negative on this axis. Using the same kind of reasoning as above, we can easily show that the voltage transfer function of a grounded network must be positive everywhere on the positive real axis.

Theorem 2 states the difference between the functions that can be realized as grounded networks and those that must be realized as ungrounded networks.

EXAMPLE. The voltage transfer function

$$A = K\,\frac{s\,(s-1)}{(s+1)\,(s^2+1)}$$

which was considered in the previous example, cannot be realized as a grounded network, but can be realized as an ungrounded network (as will be shown later).

In a more complicated case, Sturm's theorem can be used to determine whether or not there are any positive real zeros. Of course, this test is only necessary if the numerator has any negative coefficients. In particular, if the first or last coefficient of the numerator is negative, then the function cannot be positive everywhere on the positive real axis and is excluded as a grounded network.

3. NETWORK CONSTRAINTS CAUSED BY COMPLEXITY

Theorems 1 and 2 raise some interesting questions. If a grounded network cannot have a zero on the positive real axis, how close to this axis can it have a zero? If no network can have a one on the positive real axis, how close can the ones come? The answers to those questions involve some physical constraints imposed by the network configuration on the allowable positions of the ones (or the zeros in the grounded case). The complexity of the network (the actual number of resistors, inductors, and capacitors in the network) determines how close the ones (or the zeros in the grounded case) can come to the positive real axis. As the complexity increases, the ones (or zeros in the grounded case) can approach closer and closer to the axis, finally reaching it only if the network has an infinite number of elements. In order to demonstrate this several theorems will be derived.

THEOREM 3. The transfer function (voltage, impedance or admittance) of a grounded network has all positive coefficients. By transfer function is meant the actual transfer function of the network before common factors in the numerator and denominator are canceled. If the transfer function is expressed as the ratio of determinants Δ_{1k}/Δ, then both Δ_{1k} and Δ have all positive coefficients.

PROOF. Consider making a star-delta transformation at each of the internal nodes of the network so that the resultant network has only two nodes, one at the input and one at the output. Such a transformation has the property that, although it may yield nonrealizable elements, it always yields positive coefficients. The theorem follow immediately. (This simple proof of Theorem 3 was suggested by Professor S. J. Mason.)

This theorem states the constraints on the positions of right half-plane zeros as a function of the complexity of the network. If a given transfer function has some zeros in the right half-plane, some of the coefficients in the numerator may be negative. Then, in order to realize the function as the transfer function of a grounded network, the numerator and denominator must be multiplied by some Hurwitz polynomial (or some negative real root polynomial in the RC or RL case, or j-axis polynomial in the LC case) so that the resultant numerator has all positive coefficients. This in turn will increase the complexity of the network.

The degree of the required augmenting polynomial is a function of the angle that the zero makes with the positive real axis (or more correctly the angle made by the line joining the origin with the zero). As the zero gets closer to the positive real axis, the angle gets smaller, the degree of the required polynomial, and thus of the required network, increases rapidly, approaching infinity as the angle approaches zero. Some quantitative results that relate the degree of the required polynomial to the angle of the zero will be derived in a later section.

EXAMPLE. The all-pass voltage transfer function

$$A = \frac{s^2 - s + 1}{s^2 + s + 1}$$

which can be readily synthesized as a lattice or as a grounded network that contains mutual reactance, cannot be realized as a grounded network that contains no mutual reactance or ideal transformers in its present form, because of the negative coefficient. However, if the numerator and denominator are augmented with the polynomial s + 1, the resultant functions become:

$$A = \frac{(s^2 - s + 1)(s + 1)}{(s^2 + s + 1)(s + 1)} = \frac{s^3 + 1}{s^3 + 2s^2 + 2s + 1}$$

which can be realized as a grounded network.

THEOREM 4. The expressions 1 + A and 1 − A have all positive coefficients.

PROOF. 1. For grounded networks.

a. Since A has all positive coefficients, 1 + A has all positive coefficients.

b. As demonstrated in Section I, 1 − A also represents the voltage transfer function of a grounded network and, therefore, has all positive coefficients.

2. For ungrounded networks.

Any ungrounded two terminal-pair voltage transfer function can be written as the difference between two grounded voltage transfer functions (see Fig. 7), each of which must have all positive coefficients. Now 1 − A and 1 + A are given by

$$1 - A = (1 - A_{14}) + A_{13}$$

$$1 + A = A_{14} + (1 - A_{13})$$

Both of these expressions have all positive coefficients, since they are the sum of two expressions, each of which has all positive coefficients.

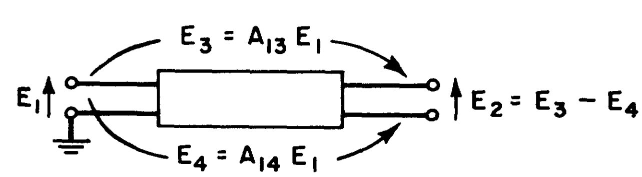

Fig. 7. Two terminal-pair network.

This theorem places the limits on the positions of ones and minus ones in the right half-plane as a function of the complexity of the network. Given a voltage transfer function, there is a minimum number of elements that can be used to synthesize that function, determined by the number of its poles and zeros, and there is a maximum gain that can be achieved by using this number of elements, determined by the greatest value of K for which $D - KN$ and $D + KN$ have positive coefficients. This value of K may or may not be the highest possible K as given by Theorem 1.

a. If, for this value of K, $D - KN$ or $D + KN$ has a zero at zero or infinite frequency, then the network has a one, in this limiting fashion, on the positive real axis, and no increase in gain is possible.
b. In all other cases, the negative coefficients that would result from an increase in K are caused by ones or minus ones in the right half-plane, and it is possible to obtain a higher gain at the expense of increasing the number of circuit elements in the network. This is done by multiplying the numerator and denominator by some Hurwitz polynomial (or negative real root polynomial in the RC or RL case or j-axis polynomial in the LC case), which enables the ones and minus ones to move closer to the positive real axis before the expressions $D - KN$ or $D + KN$ have negative coefficients. Thus the gain can be increased somewhat; however, there is a maximum gain, as determined by Theorem 1, which can never be exceeded even if the network is infinitely complicated.

EXAMPLE. Consider the voltage transfer function

$$A = \frac{Ks\,(s-1)}{(s+1)\,(s^2+1)} = \frac{K\,(s^2-s)}{s^3+s^2+s+1}$$

which was studied in a previous example. It was shown that the maximum gain obtainable from this function for any network configuration is 6.67. For this function,

$$D - KN = s^3 + (1-K)\,s^2 + (1+K)\,s + 1$$

$$D + KN = s^3 + (1+K)\,s^2 + (1-K)\,s + 1$$

Thus the maximum gain that can be obtained without augmentation is $K = 1$.

If, now, the numerator and denominator are augmented with the polynomial $(s + 1)$

$$A = \frac{K\,(s^2-s)\,(s+1)}{(s^3+s^2+s+1)\,(s+1)}$$

$$= \frac{K\,(s^3-s)}{(s^4+2s^3+2s^2+2s+1)}$$

then, for this augmented voltage transfer function,

$$D - KN = s^4 + (2 - K)\ s^3 + 2s^2 + (2 + K)\ s + 1$$

$$D + KN = s^4 + (2 + K)\ s^3 + 2s^2 + (2 - K)\ s + 1$$

Now a K of 2 can be obtained with no further augmentation.

EXAMPLE.

$$A = \frac{1}{s^2 + s + 1}$$

For this function,

$$D - KN = s^2 + s + (1 - K)$$

The following statements can be made about determining the maximum gain.

a. For a value of K greater than 1, D − KN has negative coefficients.

b. For a value of K equal to 1, D − KN has a zero on the positive real axis.

Therefore, K = 1 is the maximum possible gain, and it is obtainable without any augmentation.

The function

$$A = \frac{1}{s^2 + s + 1} = \frac{\frac{1}{s}}{s + 1 + \frac{1}{s}}$$

can be realized by inspection as an ELL-network, as in Fig. 8. No other network can be built

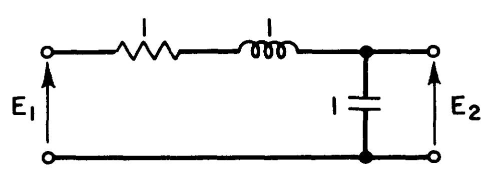

Fig. 8. ELL-network.

that possesses this transfer function with a higher constant multiplier K.

Clearly, this type of reasoning applies to all functions of the form

$$A = \frac{K}{P\,(s)}$$

and also to functions of the form

$$A = \frac{Ks^m}{s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_o}$$

All functions of this type can be realized as networks with maximum gain without any augmentation.

4. COMPLEXITY OF POLYNOMIALS

Now some results will be derived concerning the relationship between the degree of a polynomial that has all positive coefficients and the angle of its right half-plane zeros.

THEOREM 5. A polynomial of degree n which has all positive coefficients can have a zero no closer to the positive real axis than an angle of $\pm \pi/n$ radians. (The following proof of Theorem 5 was given by Professor S.J. Mason.)

PROOF. Consider making a Nyquist plot around the pie-shaped wedge in the s-plane formed by the rays at an angle $\pm \pi/n$ radians. Along the two rays the given polynomial,

$$P(s) = \sum_0^n a_k s^k$$

can be expressed as a sum of vectors,

$$P(\rho e^{\pi/n}) = \sum_0^n a_k \rho^k e^{jk(\pi/n)}$$

Since each of these vectors has an angle less than π/n radians (except the n^{th} which just equals π/n radians), and since all the coefficients are positive, the locus cannot encircle the origin and the polynomial cannot have a zero inside the pie-shaped wedge. If all the coefficients except a_o and a_n are zero, the locus just touches the origin and there is a zero on the boundary.

Thus the Butterworth polynomials $s^n + 1$ (here normalized) are the polynomials of smallest degree that have zeros closest to the positive real axis.

If the right half-plane factor is given by

$$s^2 - 2\cos\theta\, s + 1$$

and if θ is an exact divisor of π (or even if $\theta = m\pi/n$ with m and n integers), the above theorem immediately gives a polynomial with all positive coefficients that contain that factor. If θ is not an exact divisor of π, the following method can be used to obtain a positive coefficient polynomial.

a. Multiply $s^2 - 2\cos\theta\, s + 1$

by $s^2 + 2\cos\theta\, s + 1$

and obtain $s^4 - 2(2\cos^2\theta - 1)s^2 + 1 = s^4 - 2\cos 2\theta\, s^2 + 1$

b. Continue multiplying

$$s^4 - 2 \cos 2\,\theta\, s^2 + 1$$

by $$s^4 + 2 \cos 2\,\theta\, s^2 + 1$$

and obtain $$s^8 - 2 \cos 4\,\theta\, s^4 + 1$$

c. Repeat this process for p steps, and obtain

$$s^{2^{P-1}} - \cos 2^P\,\theta\, s^{2P} + 1$$

or

$$s^n - \cos \frac{n\theta}{2}\, s^{n/2} + 1$$

where $\cos \frac{n\theta}{2}$ is negative or where

$$\frac{n\theta}{2} \geqslant \frac{\pi}{2}$$

$$n \geqslant \frac{\pi}{\theta}$$

If we compare this value of n with that specified in the theorem, it is apparent that, for $n \geqslant 4$, this method yields the polynomial of smallest degree that contains a given right half-plane zero. For $n \leqslant 4$, the polynomial $s+1$ is an optimum augmenting polynomial.

A similar result was derived by Poincaré (23), who showed that a polynomial of the form

$$s^2 - 2 \cos \theta\, s + 1$$

can be augmented with the polynomial

$$s^{n-2} \sin (n-1)\,\theta + s^{n-3} \sin (n-2)\,\theta + --- + s \sin 2\,\theta + \sin \theta$$

where n is the smallest integer

$$n \geqslant \frac{\pi}{\theta}$$

to obtain

$$\sin (n-1)\,\theta\, s^n - \sin n\,\theta\, s^{n/2} + \sin \theta$$

Poincaré's augmenting polynomial can be shown to be the product of all the augmenting polynomials used in the previous method.

Unfortunately, for n greater than 4 or θ less than 45°, the polynomial that is to be used to multiply the numerator and denominator is no longer a Hurwitz polynomial, but this theorem does put a lower bound on the complexity of RLC networks.

For RC and RL networks, a similar lower bound is easily derived in the form of the statement that

$$(s^2 - 2\cos\theta\, s + 1)(s + 1)^n$$

has all positive coefficients for

$$n \geqslant \frac{2\cos\theta}{1 - \cos\theta} \text{ for n even}$$

$$n + 1 \geqslant \frac{2\cos\theta}{1 - \cos\theta} \text{ for n odd}$$

which can easily be verified by using the binomial theorem.

For LC networks, an exactly similar relationship exists

$$(s^4 - 2\cos\theta\, s^2 + 1)(s^2 + 1)^n$$

has all positive coefficients for

$$n \geqslant \frac{2\cos\theta}{1 - \cos\theta} \text{ for n even}$$

$$n + 1 \geqslant \frac{2\cos\theta}{1 - \cos\theta} \text{ for n odd}$$

EXAMPLE. Consider the voltage transfer function

$$A = \frac{Ks}{s^2 + s + 1}$$

A. Maximum possible gain. To compute the maximum gain, the function $DN' - ND'$ is calculated.

$$DN' - ND' = (s^2 + s + 1) - (2s + 1)s$$

$$= (s + 1)(s - 1)$$

The function has a positive real zero at $s = 1$. At this frequency, the transfer function is evaluated at

$$s = 1, \qquad A = K/3$$

Therefore, the maximum gain is $K \leqslant 3$.

B. Maximum gain without augmentation. To compute the maximum gain without augmentation, the function D – KN is formed.

$$D - KN = s^2 + (1 - K)\, s + 1$$

The maximum K without augmentation is K = 1.

C. Augmentation. If numerator and denominator are multiplied by (s + 1), then

$$A = \frac{Ks\,(s + 1)}{(s^2 + s + 1)\,(s + 1)} = \frac{K\,(s^2 + s)}{s^3 + 2s^2 + 2s + 1}$$

Now

$$D - KN = s^3 + (2 - K)\, s^2 + (2 - K)\, s + 1$$

A value of K = 2 can be obtained without further augmentation.

If the lower limits derived in the previous section are applied, the following table is obtained according to the relation

$$n \geqslant \frac{\pi}{\cos^{-1} \frac{K - 1}{2}}$$

Table I

Obtainable Value of K	Lower Limit on Degree of Resultant Denominator
1	2
2	3
2.5	5
2.9	10
2.99	32
2.999	96
3.0	∞

III. NETWORK REALIZATION

1. THE REALIZATION PROBLEM

The synthesis approach to the design of any physical system is logically divided in two parts. The first part is concerned with a study of a given system, to establish exactly the possible and impossible types of behavior for such a system. To speak mathematically, this part consists of establishing the necessary conditions for the realizabilty of the system. The second part is concerned with the determination of realization techniques based on these properties that yield systems which have any allowable behavior that is determined by the first part. Mathematically, the second part determines those conditions that are both necessary and sufficient for the realization of that function. There may still be a third part, which consists of looking for practical realization techniques; that is, techniques that have certain desirable properties, such as easy alignment, input and terminating capacitors, and so forth, but for our purposes, the synthesis is complete if any network realization method is found.

The synthesis problem has been completely solved for driving-point and transfer impedances and admittances. The results derived in the first two sections represent the first part of an attempted solution to the synthesis of voltage transfer functions. The succeeding sections deal with the second half of the problem—namely, realization techniques.

Very often the realization problem is difficult in the general case and the solutions obtained are unwieldy and difficult to use. Therefore, it is customary to approach the problem in small steps. The general case is subdivided into simpler cases that can more easily be solved; and the general case is then built up as a combination of these simpler cases. Thus, in the driving-point impedance problem, the two-element-kind case (RC, LC and RL) is first solved. The solution of the general case (RLC) is an extension of these simple techniques. In the synthesis of transfer impedances and admittances, the two-element-kind case is further subdivided into:

1. Those transfer functions whose transmission zeros are either on the j-axis if the function is LC or on the negative real axis if the function is RC or RL.

2. Those transfer functions whose transmission zeros are elsewhere in the s-plane.

Functions of the first kind can be realized in such a way that the transmission zeros are produced by circuit elements that have poles or zeros, such as those in a ladder configuration, while in the second case the problem is more difficult because the transmission zeros cannot be produced by circuit elements that have poles or zeros but must be produced by cancellation. However, the problem can be solved by a simple extension of the methods used in the first part; for example, a parallel combination of ladders can be used.

For the synthesis of voltage transfer functions, we propose to make still another subdivision into:

A. Those voltage transfer functions whose ones and minus ones are all: in the left half-plane for RLC functions; on the negative real axis for RC or RL functions; and on the j-axis for LC functions, so that, in a realization of such a function, these ones and minus ones can be produced by network elements that have poles and zeros.

B. Those voltage transfer functions that have ones or minus ones elsewhere in the s-plane, so that these ones and minus ones must be produced by cancellation.

It will be shown later that the first case can always be realized as a symmetrical network, while the second case requires more general (and more complicated) methods.

This type of subdivision is slightly different from the ones considered previously because the same transfer function that has the same poles and zeros, may belong to either group, depending on the value of the constant K. As shown in Section I, for K below a certain critical value, all voltage transfer functions (except possibly for RLC functions with j-axis poles) have their ones and minus ones in positions in the s-plane that correspond to the first group; that is, in the left half-plane for RLC functions; on the negative real axis for RC and RL functions; and on the j-axis for LC functions. For larger values of K, the ones and minus ones move to positions that correspond to the second group.

To clarify the relationship between the critical value of K considered here and the various other critical values of K discussed earlier, the important values of K will be briefly reviewed:

A. The highest value of K for which the voltage transfer function has no ones or minus ones on the positive real axis – This value of K yields the maximum gain obtainable from any network that has the given poles and zeros.

B. The highest value of K for which the functions $1 - A$ and $1 + A$ have all positive coefficients – This value of K yields the greatest gain obtainable from the given poles and zeros without augmentation.

C. The highest value of K for which the ones and minus ones are all: in the left half-plane for RLC functions; on the negative real axis for RC functions; and on the j-axis for LC functions. This is the greatest value of K for which the function belongs to the first group and can be realized simply as a symmetrical network.

Any two or all three of these values of K may be equal. For example, all three of the critical values of K for the function

$$A = \frac{K}{s + 1}$$

occur for $K = 1$. This function can be realized with one resistor and one capacitor (or one inductor) in the form of an ELL-network (which is a degenerate symmetrical network) and no other network can be built that possesses this voltage transfer function and has a greater gain.

2. THE IMPEDANCE POINT OF VIEW

The subdivision we have been considering can also be discussed in terms of the impedance and admittance functions of the network. If the voltage transfer ratio is expressed in terms of these functions,

$$A = \frac{z_{12}}{z_{11}} = \frac{y_{12}}{y_{22}}$$

then a little thought shows that there are only three ways in which ones can occur.

1. If at some particular frequency, z_{12} and z_{11} have a pole with the same residue, at that frequency, y_{12} and y_{22} will, of course, be equal.

2. If at some particular frequency, y_{12} and y_{22} have a pole with the same residue, at that frequency, z_{12} will just equal z_{22}.

3. If at some particular frequency, z_{11} just happens to equal z_{12} and y_{22} just happens to equal y_{12}, neither the impedances nor the admittances will have a pole or a zero at that frequency. (The case of z_{12} and z_{11} both having a zero with the same slope is not a separate case, because if z_{12} and z_{11} both have zeros, then y_{12} and y_{22} both have poles, which corresponds to case two.)

Again, the discussion includes only ones; the discussion for minus ones is very similar except for several minus signs in residues.

The first two cases correspond to ones and minus ones, because the elements have zeros and poles, and the third case corresponds to cancellation. (For two-element-kind symmetrical networks, cases one and two correspond to the residue condition, being satisfied with the equality sign.) This way of looking at the subdivision throws some light on one of the apparent inconsistencies of voltage transfer synthesis when it is considered from the impedance point of view. For a given voltage transfer function,

$$A = \frac{y_{12}}{y_{22}} = \frac{z_{12}}{z_{11}}$$

there appears to be insufficient information for determining the Z and Y functions of the network. The numerator and denominator of the transfer function serve to determine the zeros of z_{12} and z_{11} (y_{12} and y_{22}) but what of the poles of the impedances (and admittances)? They appear to be arbitrary, subject only to certain realizability conditions. Indeed, in the general case, it will be shown that these poles are arbitrary, but in the particular case considered here, the poles of the Z and Y functions are directly related to the ones and minus ones of the function, as is shown by the following reasoning. If

$$A = \frac{y_{12}}{y_{22}}$$

then

$$1 - A^2 = 1 - \frac{y_{12}^2}{y_{22}^2}$$

or

$$\frac{(D - KN)(D + KN)}{D^2} = \frac{y_{22}^2 - y_{12}^2}{y_{22}^2}$$

At this point in the discussion, the possibility of a symmetrical network realization for functions of group one first becomes apparent. For symmetrical networks,

$$z_{11} = \frac{y_{22}}{y_{22}^2 - y_{12}^2}$$

then

$$\frac{(D - KN)(D + KN)}{D^2} = \frac{1}{z_{11} y_{22}}$$

Thus, for symmetrical networks, all the ones and minus ones are included in the poles of z_{11} or y_{22}.

3. TEST FOR CRITICAL VALUE OF K

Before discussing the realization techniques appropriate to group one, it is necessary to develop a test for the value of K below which the ones and minus ones are in the positions in the s-plane that correspond to this group. This will enable us not only to associate an obtainable gain with a given function but also to ascertain how much of the maximum possible gain is being sacrificed in order to obtain a practical structure. Naturally, the tests will differ in the two-element-kind and the three-element-kind cases.

4. TWO-ELEMENT CASE

The two-element case consists of RC, RL, and LC networks. Since the test is very similar for these three types of network, the details will be presented for the RC case; the extension to the other two cases is straightforward. For RC networks, the object of this test is to determine the greatest value of K for which both D – KN and D + KN have all their roots on the negative real axis. The most straightforward way to perform such a test is to actually form the functions D – KN and D + KN, keeping K as a parameter, and apply a test to the resultant polynomials to determine for what values of K they have all their zeros on the negative real axis. Such a test is described below.

If P(s) is a polynomial that has all its zeros on the negative real axis, then $P(s)/\frac{dP(s)}{ds}$ is an RC impedance, since it has all simple poles on the negative real axis with positive real residues (in this case all equal to one). The test consists of expanding $P(s)/\frac{dP(s)}{ds}$ in a continued fraction expansion and demanding that all the quotients be positive. This is a general test to determine whether a given polynomial has all its zeros on the negative real axis. Unfortunately, in the application considered here, the determination of the values of K for which all the quotients are positive is computationally difficult. Each quotient in the continued fraction expansion is a rational function of K, the first of first degree, the second of second degree and so on. The computation becomes prohibitive.

A much more satisfactory test involves, first, determining the frequencies at which the ones or minus ones leave the axis and then determining the corresponding value of K. The test is quite similar to the one described earlier for the maximum possible gain, in which we determined the frequencies at which the ones and minus ones left the positive real axis. By similar reasoning, it becomes apparent that ones and minus ones can only leave the negative real axis, through zero frequency, infinite frequency or at some finite nonzero frequency at which two roots come together and then leave the axis as a complex pair. As described in Section II, this latter condition can only occur at one of the maxima or minima of the function. The test then consists of the following two steps:

1. Find the frequencies at which the function has its extrema on the negative real axis. These are just the negative real zeros of the first derivative of the function, which are the negative real roots of the polynomial $DN' - ND'$.

2. At each of these frequencies, as well as at zero and infinite frequency, evaluate the function. Pick the value of K for which the function is equal to or less than one at each of these frequencies.

EXAMPLE. Consider the voltage transfer function

$$A = \frac{K\,(s^2 + 2)}{(s + 1)\,(s + 2)\,(s + 3)}$$

For this function,

$$DN' - ND' = -(s^4 - 5s^2 + 12s + 22)$$

$$= -(s + 2.50)\,(s + 1.35)\,(s^2 - 3.85s + 6.51)$$

The function is then evaluated at each of the negative real roots and at zero and infinite frequency:

$$A(0) = 1/3K$$

$$A(\infty) = 0$$

$A(2.5) = 22\,K$

$A(1.35) = -7.5\,K$

The maximum K for this test is then K = 1/22 = 0.045. Compare this value with the K of 3, which is the highest obtainable for this zero-pole configuration, and which, incidentally, is obtainable without augmentation.

5. RLC CASE

For the three-element case, the object of the test is to determine the greatest value of K for which D – KN and D + KN are Hurwitz polynomials. Again, there are two possibilities:

1. The Hurwitz test can be applied directly to the polynomials D – KN and D + KN. However, this method leads to computational difficulties.

2. The frequencies at which the ones cross the j-axis can be determined and then the corresponding values of K can be calculated. This is the test that will be discussed.

The ones and minus ones can only cross the j-axis at zero frequency, infinite frequency or at some other j-axis point at which the function is pure real. These latter frequencies are just the j-axis zeros of the imaginary part of the voltage transfer function. If the function is

$$A = \frac{m_1 + n_1}{m_2 + n_2}$$

in which m and n refer to even and odd parts and $m_2 + n_2$ is a Hurwitz polynomial, but $m_1 + n_1$ is not necessarily a Hurwitz polynomial, the imaginary part of this function on the j-axis is given by the function

$$M\left\{A(j\omega)\right\} = \frac{m_2 n_1 - m_1 n_2}{m_2^2 - n_2^2}$$

The frequencies of interest are the j-axis zeros of the polynomial

$$m_2 n_1 - m_1 n_2 = 0$$

The voltage transfer function is evaluated at each of these frequencies, as well as at zero and infinite frequencies, and the constant K is so chosen that the value of the function is less than one at each of these frequencies.

EXAMPLE. As an example of this test, consider the function

$$A = \frac{K\,(s^2 + 2)}{(s + 1)\,(s + 2)\,(s + 3)}$$

which was considered earlier. This function has all its poles on the negative real axis and can be realized as an RC function. As shown in the preceding example, the maximum gain for this function when realized as an RC symmetrical network is K = 0.045, while K = 3 is the maximum possible for this pole-zero configuration. Now we investigate the matter of what K is obtainable if an RLC symmetrical network realization is to be used.

For this function,

$$m_2n_1 = m_1n_2 = s\,(s^2 + 2)\,(s^2 + 11)$$

The function is evaluated at zero frequency, infinite frequency, and at each of the j-axis roots of this polynomial.

$$A(0) = 1/3K$$

$$A(\infty) = 0$$

$$A(j2) = 0$$

$$A = (j11) = 3/20K$$

The maximum K for this test is then K = 3. Since for this value the one crosses the j-axis at zero frequency and then finds itself on the positive real axis, K = 3 is the maximum obtainable for any configuration.

IV. LATTICE REALIZATION

1. INTRODUCTION

For many years, symmetrical networks in the form of lattices have been one of the basic building blocks used by the network designer. Lattices were first used in the image parameter techniques of conventional filter theory and later incorporated into many of the synthesis procedures of what we like to call "modern network theory." The simplicity of the realization techniques for the lattice structure, as well as its generality, have led to its wide use as a starting point from which to derive other symmetrical, as well as nonsymmetrical, networks.

A similar approach will be used here. In this section we wish to show that any voltage transfer function (A possible exception is RLC functions with j-axis poles, which will not be considered; for a discussion of this case see refs. 15 and 16) which has a value of K small enough so that all the ones and minus ones are: (a) in the left half-plane for RLC functions; (b) on the negative real axis for RC and RL functions; (c) on the j-axis for LC functions can be realized as a symmetrical lattice. The next section will consider the problem of converting the lattice to unbalanced form; it will be shown that any voltage transfer function can first be realized as a lattice, and then the lattice can always be unbalanced, provided only that the given transfer function has all positive coefficients, a condition which was shown to be necessary in Section II.

2. THE LATTICE NETWORK

It is well known that the lattice is the most general symmetrical network; that is, any symmetrical network that can be realized at all can be realized as a lattice. This configuration is, therefore, an appropriate starting point for a study of symmetrical networks.

It is readily apparent either physically or mathematically that, for the lattice, all the ones and minus ones are attributable to elements having poles or zeros. From Fig. 9, it can be seen that

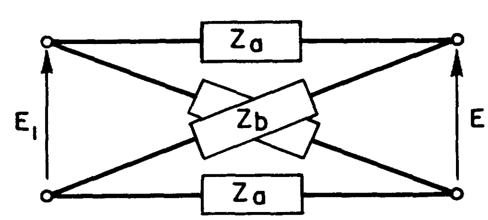

Fig. 9. A lattice.

the lattice can only have a one when Z_a is a short circuit (has a zero) or when Z_b is an open circuit (has a pole) and it can only have a minus one when Z_a has a pole or Z_b has a zero.

This can also be seen mathematically by writing

$$A = \frac{z_{12}}{z_{11}} = \frac{Z_b - Z_a}{Z_b + Z_a}$$

$$= \frac{y_{12}}{y_{22}} = \frac{Y_a - Y_b}{Y_a + Y_b}$$

A can have a one in two ways:

1. If Z_b has a pole not contained in Z_a, in which case z_{11} and z_{12} have a pole with the same residue.

2. If Z_a has a zero not contained in Z_b or in other words Y_a has a pole not contained in Y_b, in which case y_{12} and y_{22} have a pole with the same residue.

Similarly A can have a minus one in two ways:

1. If Z_a has a pole not contained in Z_b, in which case z_{12} and z_{11} have poles whose residues are equal in magnitude but opposite in sign.

2. If Z_b has a zero not contained in Z_a or in other words Y_b has a pole not contained in Y_A, in which case y_{12} and y_{11} have poles whose residues are equal in magnitude but opposite in sign.

For two-element-kind networks, these conditions correspond to the case in which the residue condition is satisfied by the equality sign.

Thus all the ones and minus ones are attributable to circuit elements that have poles or zeros, causing either z_{11} and z_{12} or y_{22} and y_{12} to have poles with residues equal in magnitude and either the same or opposite sign. The lattice then definitely belongs to group 1, as defined in Section III; it remains to be proved, however, that any voltage transfer function belonging to group 1 can be realized as a lattice.

If the expression for A is written in the form

$$A = \frac{1 - \frac{Z_a}{Z_b}}{1 + \frac{Z_a}{Z_b}}$$

and this expression is solved for Z_a/Z_b, the resultant equation reads,

$$\frac{Z_a}{Z_b} = \frac{Y_b}{Y_a} = \frac{1 - A}{1 + A} = \frac{D - KN}{D + KN}$$

It is again readily apparent that all the ones and minus ones are caused by Z_a or Z_b with poles or zeros.

3. LATTICE SYNTHESIS

Now that the properties of lattices are known, the realization can proceed in a straightforward fashion. Starting with a given voltage transfer function,

$$A = \frac{KN}{D}$$

for which an appropriate value of K has been determined by using the methods of Section III, the following expression is formed.

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN}$$

The problem is then to split Z_a/Z_b into Z_a and Z_b so that both are realizable impedances. The methods for doing this differ in the two-element and three-element cases; therefore these two procedures will be considered separately.

4. TWO-ELEMENT CASE

The method given here for the two-element lattices was first derived by Fialkow and Gerst (14). The methods for RC, RL, and LC functions are essentially similar; therefore the details of the procedure will be given only for the RC case, but examples will be given for both the RC and LC cases.

For RC functions, the problem can be stated as follows. Given the expression,

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN}$$

all of whose critical frequencies are on the negative real axis, split up Z_a/Z_b into Z_a and Z_b, both of which are realizable RC impedances.

Clearly, the critical frequencies of the quotient of two RC impedances Z_a/Z_b must obey certain order relationships on the negative real axis. Similarly, because of the principle of the continuity of the roots of an algebraic equation with a parameter K, the critical frequencies of the expression $(D - KN)/(D + KN)$ must also obey certain order relationships on the negative real axis, provided that K is small enough so that all the zeros of both polynomials are on this axis. We shall show that these order relationships are identical and, moreover, that a knowledge that such a relationship exists allows an expression of the form Z_a/Z_b to be readily split into Z_a and Z_b.

The relationship involved is one of double alternation, which can be defined as follows with the help of Fig. 10. The poles and zeros of a function exhibit the property of double alternation on the negative real axis if, starting from zero frequency, each pair of critical frequencies contains one pole and one zero in either order. The highest critical frequency may be a pole or a zero and may or may not belong to a pair.

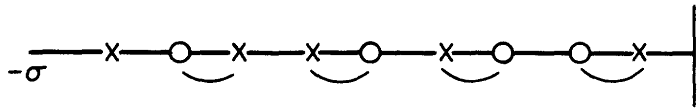

Fig. 10. Double alternation.

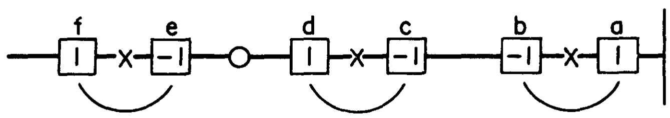

Fig. 11. An RC voltage transfer function.

a. The Critical Frequencies of D – KN/D – KN Exhibit Double Alternation

PROOF. The zeros and poles of the function D – KN/D + KN are the ones and minus ones of the given voltage transfer function. The proof then consists of showing that the ones and minus ones exhibit double alternation. To clarify the discussion, Fig. 11 shows the critical frequencies that lie on the negative real axis of a typical RC voltage transfer function. All of the poles of the function are on the negative real axis; some of the zeros are on the real axis and some are off, but only those that are on the axis contribute to the discussion; therefore only these are shown. Corresponding to each pole of the transfer function is a one and a minus one. The one is to the right of the pole and the minus one to the left of the pole if the residue in that pole is positive, and the one is to the left and the minus one to the right if the residue is negative. Whether the residue is positive or negative depends only on the critical frequencies that are on the real axis. If there are none or if there are an even number of zeros to the right of the first pole, the residue will be positive; if there are an odd number, the residue will be negative. Similarly, if there are none or if there are an even number of zeros between the first and second poles, the residue in the second pole will be of opposite sign from the residue in the first pole; if there are an odd number, the residues will be of the same sign. By the use of this reasoning, the ones and minus ones can be drawn approximately as in Fig. 11. Since the figure is drawn for K that is sufficiently small, the property of double alternation is clearly satisfied. The question remains, however, as to whether or not this condition will be true for all values of K for which the ones and minus ones are all on the negative real axis. As K is increased, three things might happen to upset the situation:

1. b and c might come together. This condition specifies the maximum K for this configuration.

For all K below this value, the double alternation property exists, and for K above this value, the function no longer has all its ones and minus ones on the axis; therefore it is no longer of interest. For K that is equal to this critical value, the double alternation property still exists in a limiting form.

2. a might go through zero frequency, or f might go through infinite frequency. Again, this specifies the maximum K for this configuration.

3. d and e might cross, thus ruining the double alternation property. This condition can never occur, since d and e can only reach the zero that is between them for an infinite value of K. Since all cases have been covered, it becomes apparent that the ones and minus ones exhibit the property of double alternation for all values of K for which all the ones and minus ones are on the negative real axis.

b. The Quotient of Two RC Impedances Exhibits Double Alternation

PROOF. The easiest way to become convinced of the truth of this statement is by considering a simple example. Fig. 12 shows a pole-zero plot of two arbitrarily drawn RC impedances and the

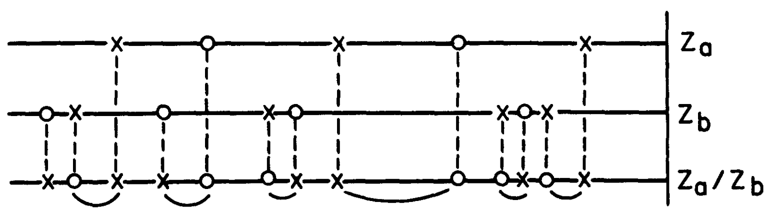

Fig. 12. Double alternation of the quotient of two RC impedances.

pole-zero plot of their quotient. Clearly, double alternation will always exist. The steps in the realization procedure are now clear.

1. Choose K so that D − KN and D + KN have all their zeros on the negative real axis.
2. Form

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN}$$

This function will exhibit double alternation.

3. Split Z_a/Z_b into Z_a and Z_b. This splitting can be done in a large number of ways, either with or without surplus factors. One way that will always work is as follows. If

$$\frac{Z_a}{Z_b} = \frac{K\,(s+\sigma_1)\,(s+\sigma_2)\,---}{(s+\sigma_a)\,(s+\sigma_b)\,---}$$

where the numerical and literal order refer to the order on the negative real axis starting from zero frequency, then for RC networks,

$$Z_a = \frac{K_1 (s + \sigma_2)(s + \sigma_4) - -}{(s + \sigma_a)(s + \sigma_c) - -}$$

$$Z_b = \frac{K_2 (s + \sigma_b)(s + \sigma_d) - -}{(s + \sigma_1)(s + \sigma_3) - -}$$

where

$$K = \frac{K_1}{K_2}$$

EXAMPLE 1. The voltage transfer function

$$A = \frac{K (s^2 + 2)}{(s + 1)(s + 2)(s + 3)}$$

was considered in previous problems. It has been shown that K = 0.045 is the largest K for which this function can be realized as an RC lattice. Then, let us realize

$$A = \frac{0.045 (s^2 + 2)}{(s + 1)(s + 2)(s + 3)}$$

For this function,

$$D - KN = s^3 + 5.955s^2 + 11s + 5.91 = (s + 2.5)^2 (s + 0.955)$$

$$D + KN = s^3 + 6.045s^2 + 11s + 6.09 = (s + 1.05)(s + 3.19)(s + 1.81)$$

The factoring of D − KN was particularly easy because the double-order one at $s = -2.5$ was known beforehand from the calculation of maximum K. Then

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{(s + 2.5)^2 (s + 0.955)}{(s + 1.05)(s + 3.19)(s + 1.81)}$$

One way to form Z_a and Z_b is as follows:

$$Z_a = \frac{s + 2.5}{(s + 1.05)(s + 3.19)}$$

$$Z_b = \frac{s + 1.81}{(s + 0.955)(s + 2.5)}$$

EXAMPLE 2. Consider the LC voltage transfer function

$$A = \frac{K(s^4 + 1)}{(s^2 + 1)(s^2 + 2)}$$

We first determine the maximum K, by finding the extrema of the function on the j-axis.

$$DN' - ND' = 6s\left(s^4 + \frac{2}{3}s^2 - 1\right) = 6s(s^2 + 1.39)(s^2 + 0.73)$$

Evaluating the voltage transfer function at the j-axis zeros of this polynomial, as well as at zero and infinite frequency, yields

$$A(\infty) = K$$

$$A(0) = \frac{1}{2}K$$

$$A(j\sqrt{1.39}) = 46.5$$

Then the maximum K is

$$K = \frac{1}{46.5} = 0.0215$$

Just to be different, we shall realize the function with a K that is less than the maximum

$$A = \frac{0.02(s^4 + 1)}{(s^2 + 1)(s^2 + 2)}$$

For this function,

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{0.96(s^2 + 0.96)(s^2 + 2.10)}{(s^2 + 1.05)(s^2 + 1.89)}$$

Z_a and Z_b can be formed in a number of ways. Three ways will be shown: two without surplus factors, and one with surplus factors.

$$Z_a = \frac{0.96(s^2 + 0.96)}{s(s^2 + 1.89)} \qquad Z_b = \frac{s^2 + 1.05}{s(s^2 + 2.10)} \qquad (1)$$

$$Z_a = \frac{0.96\,(s^2 + 0.96)\,(s^2 + 2.10)}{s\,(s^2 + 1.05)} \qquad Z_b = \frac{s^2 + 1.89}{s} \tag{2}$$

$$Z_a = \frac{0.96\,(s^2 + 0.96)\,(s^2 + 2.10)}{s\,(s^2 + 1.5)\,(s^2 + 3)} \qquad Z_b = \frac{(s^2 + 1.05)\,(s^2 + 1.89)}{s\,(s^2 + 1.5)\,(s^2 + 3)} \tag{3}$$

5. THREE-ELEMENT CASE

For RLC networks, the problem can be stated as follows. Given the function

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN}$$

the numerator and denominator of which are both Hurwitz polynomials, split this function into Z_a and Z_b both of which are positive real impedances. This problem is inherently more difficult than the two-element case, because

1. The realizable conditions for RLC functions do not easily lend themselves to the methods of root locus, which we have been employing.

2. The realization techniques of RLC networks are inherently more complex than those of the two-element case, so that a general RLC network may be difficult to build.

For these reasons, methods that are not completely general but yield networks with simple realizations are often more useful than the general solution. In this section, six solutions will be presented, which can be divided into three general categories.

1. Methods that work in particular cases, and yield simple realizable networks with any realizable gain.

2. Methods that will always work and yield simple realizable networks but sometimes give less than the maximum gain.

3. A general solution that will always yield a network with any allowable gain but uses surplus factors and does not allow a simple realization.

These methods will now be discussed.

a. A Particular Method – Lossless Impedances in Series or Parallel with Resistors

In certain cases, a realization in terms of lossless impedances in series or parallel with resistors is possible. For such a realization, either the odd or the even parts of the numerator and denominator of the given voltage transfer function $A = KN/D$ must be proportional. The details will be presented for the odd parts proportional; the extension to the other case is obvious. If

$$A = \frac{KN}{D} = \frac{m_N + k\,n_D}{m_D + k\,n_D}$$

then

$$\frac{Z_a}{Z_b} = \frac{m_D - m_N + (1 - k)\, n_D}{m_D + m_N + (1 + k)\, n_D}$$

which can be written

$$\frac{Z_a}{Z_b} = \frac{m_1 + C\, n_2}{m_2 + n_2}$$

where

$$m_1 = m_D - m_N \qquad n_2 = (1 + k)\, n_D$$

$$m_2 = m_D + m_N \qquad C = \frac{1 - k}{1 + k}$$

Then, since the odd parts of the numerator and denominator of the function Z_a/Z_b are proportional, we can write

$$\frac{Z_a}{Z_b} = \frac{C + \dfrac{m_1}{n_2}}{1 + \dfrac{m_2}{n_2}}$$

The function can then be split into Z_a and Z_b, either

$$Z_a = C + \frac{m_1}{n_2} \qquad Z_b = 1 + \frac{m_2}{n_2}$$

or

$$Z_a = \frac{1}{1 + \dfrac{m_2}{n_2}} \qquad Z_b = \frac{1}{C + \dfrac{m_1}{n_2}}$$

Then the lattice impedances are lossless networks in series or parallel with resistors. Part of this resistance can be removed from the lattice either as a series or as a shunt termination.

If, in the given voltage transfer function, k is equal to zero, that is, if the numerator is even (or odd), then in the function Z_a/Z_b, $C = 1$ and the resultant impedances Z_a and Z_b are both lossless networks in series or parallel with one-ohm resistors, which can be completely removed from the lattice, leaving a lossless network terminated in a one-ohm resistor. This is a familiar realization for transfer functions that have even or odd numerators. However, this method is slightly more general than that, in that it applies when the numerator has both odd and even coefficients, as long as the odd or even parts of the numerator are proportional.

EXAMPLE 1. Consider again the voltage transfer function

$$A = \frac{3(s^2+2)}{(s+1)(s+2)(s+3)}$$

Since the numerator is even, we can write

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^3 + 3s^2 + 11s}{s^3 + 9s^2 + 11s + 12} = \frac{1 + \dfrac{3s}{s^2 + 11}}{1 + \dfrac{9s^2 + 11}{s(s^2 + 11)}}$$

Then

$$Z_a = 1 + \frac{3s}{s^2 + 11}$$

$$Z_b = 1 + \frac{9s^2 + 12}{s(s^2 + 11)}$$

EXAMPLE 2. Consider the voltage transfer function

$$A = \frac{s^4 + 5s^2 + s + 4}{3s^4 + 2s^3 + 15s^2 + 5s + 12}$$

For this function,

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^4 + s^3 + 5s^2 + 2s + 4}{2s^4 + s^3 + 10s^2 + 3s + 8}$$

$$\frac{Z_a}{Z_b} = \frac{1 + \dfrac{s(s^2+2)}{(s^2+1)(s^2+4)}}{2 + \dfrac{s(s^2+3)}{(s^2+1)(s^2+4)}}$$

b. Methods That Always Work but Sometimes Yield Less Than the Maximum Gain

1. For $|A(j\omega)| \leqslant 1$, if $|A(j\omega)| \leqslant 1$, it is well known that the function $(1 - A)/(1 + A)$ is positive real, in which case the expression

$$\frac{Z_a}{Z_b} = \frac{1 - A}{1 + A} = \frac{D - KN}{D + KN}$$

can be split into

$$Z_a = \frac{1 - A}{1 + A}$$

$$Z_b = 1$$

The proof that $1 - A/1 + A$ is positive real for $|A(j\omega)| \leqslant 1$ can be carried out in a somewhat novel fashion by using the methods developed here.

A. Since $|A(j\omega)| \leqslant 1$, then $|D(j\omega)| \geqslant |KN(j\omega)|$ and, by Rouché's theorem, D, D − KN, and D + KN all have the same number of zeros in the left plane. This proves that D − KN and D + KN are both Hurwitz polynomials.

B. If

$$A = \frac{m_1 + n_1}{m_2 + n_2}$$

then

$$\frac{1 - A}{1 + A} = \frac{m_2 - m_1 + n_2 - n_1}{m_2 + m_1 + n_2 + n_1}$$

The real part of this function on the j-axis is given by

$$\operatorname{Re}\left\{\frac{1 - A(j\omega)}{1 + A(j\omega)}\right\} = \frac{m_2^2 - n_2^2 - m_1^2 - n_1^2}{(m_2 + m_1)^2 - (n_2 + n_1)^2}$$

$$= \frac{\left[1 - \dfrac{m_1^2 - n_1^2}{m_2^2 - n_2^2}\right](m_2^2 - n_2^2)}{(m_2 + m_1)^2 - (n_2 + n_1)^2}$$

$$= \frac{\left[1 - |A(j\omega)|^2\right](m_2^2 - n_2^2)}{(m_2 + n_1)^2 - (n_2 + n_1)^2}$$

Thus the real part of $1 - A/1 + A$ is positive on the whole j-axis when the magnitude of A is greater than one on the whole j-axis, and, for this condition, $1 - A/1 + A$ is a positive real impedance.

EXAMPLE. Consider again the function

$$A = \frac{K(s^2 + 2)}{(s + 1)(s + 2)(s + 3)}$$

We must first determine the appropriate value of K for which this procedure is possible. This can be done by forming A (jω) and then differentiating to determine its extrema, after which the function is evaluated at each of these frequencies and an appropriate value of K is picked so that the magnitude of A is less than one at each of the extrema. However, in this simple case, a glance at the pole-zero plot shows that the maximum value of the magnitude occurs at zero frequency. Thus a K of 3 is possible without the magnitude of the function becoming greater than one anywhere on the j-axis. In this case, no decrease in gain below the maximum is required for this method. For this value of K,

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^3 + 3s^2 + 11s}{s^3 + 9s^2 + 11s + 12}$$

This is a positive real function, as can be readily verified. Then let

$$Z_a = \frac{s^3 + 3s^2 + 11s}{s^3 + 9s^2 + 11s + 12}$$

$$Z_b = 1$$

2. For $|\text{Re}[A(j\omega)]| \leqslant 1$, the requirement that the magnitude of the real part on the j-axis be less than one is clearly less of a restriction on gain than that the magnitude be less than one. For this case, two realization procedures are possible.

Method 1. If

$$A = \frac{m_1 + n_1}{m_2 + n_2}$$

form

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{m_2 - m_1 + n_2 - n_1}{m_2 + m_1 + n_2 + n_1}$$

Let

$$Z_a = \frac{m_2 - m_1 + n_2 - n_1}{m_2 + n_2} \qquad Z_b = \frac{m_2 - m_1 + n_2 - n_1}{m_2 + n_2}$$

To prove that these two functions are both positive real for $|\mathrm{Re}\,[A\,(j\omega)]| \leqslant 1$, form their real part on the j-axis.

$$Z_{a,b}(j\omega) = \frac{(m_2^2 - n_2^2) \mp (m_1 m_2 - n_1 n_2)}{m_2^2 - n_2^2}$$

$$= 1 \mp \frac{m_1 m_2 - n_1 n_2}{m_2^2 - n_2^2}$$

$$= 1 \mp A\,(j\omega)$$

It is clear that for $|\mathrm{Re}\,[A\,(j\omega)]| \leqslant 1$, the real parts of both Z_a and Z_b are positive on the whole imaginary axis.

EXAMPLE. Consider again the function

$$A = \frac{K\,(s^2 + 1)}{(s + 1)\,(s + 2)\,(s + 3)}$$

The rigorous way to find the maximum K that is appropriate for this method is to form the real part of A on the j-axis and then differentiate to find its extrema. However, in this case, we already know that a K of 3 is the maximum gain, and since, for this value of K, the magnitude of the function on the j-axis is not greater than one, then the real part on the j-axis also cannot be greater than one; therefore a K of 3 can be realized with this method.

For K = 3,

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^3 + 3s^2 + 11s}{s^3 + 9s^2 + 11s + 12}$$

Then let

$$Z_a = \frac{s^3 + 3s^2 + 11s}{(s + 1)\,(s + 2)\,(s + 3)}$$

$$Z_b = \frac{s^3 + 9s^2 + 11s + 12}{(s + 1)\,(s + 2)\,(s + 3)}$$

both of which are positive real impedances, as can be readily checked.

Method 2. Two parallel lossless networks terminated in resistors.

Using the fact that, for two networks in parallel,

$$A = \frac{y_{12}^{(1)} + y_{12}^{(2)}}{y_{22}^{(1)} + y_{22}^{(1)}}$$

we can write

$$A = \frac{m_1 + n_1}{m_2 + n_2}$$

as

$$A = \frac{\dfrac{m_1}{m_2 + n_2} + \dfrac{n_1}{m_2 + n_2}}{\dfrac{m_2}{m_2 + n_2} + \dfrac{n_2}{m_2 + n_2}}$$

and then realize the over-all function as two networks in parallel. To realize one of them, for example, let

$$y_{12}^{(1)} = \frac{m_1}{m_2 + n_2}$$

and

$$y_{22}^{(1)} = \frac{m_2}{m_2 + n_2}$$

The voltage transfer function of this network is

$$A^{(1)} = \frac{y_{12}}{y_{22}} = \frac{m_1}{m_2}$$

which represents the voltage transfer function of a lossless network. Then, if we write

$$y_{22}^{(1)} = \frac{1}{1 + \dfrac{n_2}{m_2}} = \frac{1}{1 + z_2}$$

this amounts to a lossless network in series with a one-ohm resistor at the output end, which does not affect the voltage transfer function of that network. Considering the lossless part of the network, we wish to construct a lattice with

$$y_{22}^{(1)\prime} = \frac{m_2}{n_2}$$

$$A^{(1)} = \frac{m_1}{m_2}$$

For this lossless lattice, form

$$\frac{Y_b^{(1)}}{Y_a^{(1)}} = \frac{1 - A}{1 + A} = \frac{m_2 - m_1}{m_2 + m_1}$$

and let

$$Y_b^{(1)} = \frac{m_2 - m_1}{n_2}$$

$$Y_a^{(1)} = \frac{m_2 + n_1}{n_2}$$

These admittances are guaranteed realizable if $|\mathrm{Re}\,[A\,(j\omega)]| \leqslant 1$, because the numerator and denominator of these admittances are the odd and even parts of the impedances considered in the previous method, which were shown to be positive real for $|\mathrm{Re}\,[A\,(j\omega)]| \leqslant 1$.

The over-all network, which consists of the parallel connection of two networks each of which is a lossless lattice in series with a one-ohm resistor, is shown in Fig. 13. For this structure

$$y_{22} = \frac{m_2}{m_2 + n_2} + \frac{n_2}{m_2 + n_2} = 1$$

While this type of structure is not, strictly speaking, a lattice, it enjoys the advantages of easy

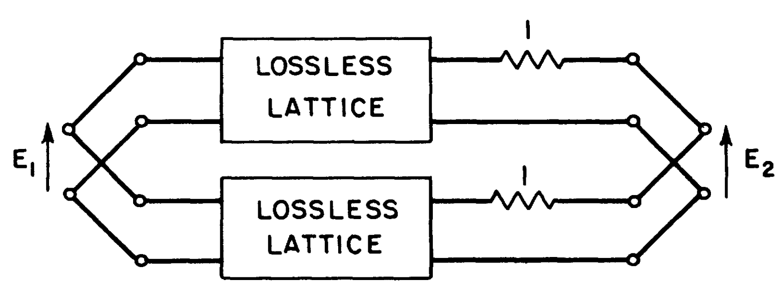

Fig. 13. A type of lattice structure.

realization and, as will be shown in Section V, it is possible to unbalance the lattices by using the methods applicable to two-element-kind networks.

EXAMPLE. Consider the voltage transfer function

$$A = \frac{K\,(s^3 - 2)}{(s + 1)\,(s^2 + s + 1)}$$

For this function,

$$\text{Re}\left\{A\,(j\omega)\right\} = \frac{K\,(\omega^6 - 2\omega^4 - 4\omega^2 - 2)}{\omega^6 + 1}$$

$$\text{Im}\left\{A\,(j\omega)\right\} = \frac{K\,2\omega\,(\omega^4 + \frac{3}{2}\omega^2 + 2)}{\omega^6 + 1}$$

A. Maximum gain for any lattice. To find the maximum gain obtainable for a lattice realization, we must evaluate the voltage transfer function at zero frequency, infinite frequency, and at each of the j-axis zeros of the imaginary part. Since the imaginary part has no finite nonzero j-axis zeros, the maximum gain obtainable is ½.

B. Maximum gain with this method. The maximum K obtainable with this method is that value of K for which the real part of K is always less than one on the j-axis. To determine this value of K, we must find the maxima of the real part. Differentiating, we obtain

$$\frac{\partial}{\partial\omega}\left[A\,(j\omega)\right] = \frac{4\omega\,(\omega^2 - 11)\,(\omega^2 - 3.16)\,(\omega^4 + 0.26\omega^2 + 1.13)}{(\omega^6 + 1)^2}$$

The j-axis extrema of the real part occur at $\omega^2 = 1.1$, $\omega^2 = 3.16$. The real part is then evaluated at these frequencies, as well as at zero and infinite frequency

$$A\,(0) = -2K$$

$$A\,(\infty) = K$$

$$A\left[j\,(1.1)^{1/2}\right] = 0.985K$$

$$A\left[j\,(3.16)^{1/2}\right] = 2.00K$$

Thus a K of ½ is obtainable with this method and no decrease from the maximum is required. For K = ½,

$$A = \frac{\frac{1}{2}(s^3 - 2)}{s^3 + 2s^2 + 2s + 1}$$

write

$$A = \frac{\dfrac{\frac{1}{2}s^3}{s^3 + 2s^2 + 2s + 1} - \dfrac{1}{s^3 + 2s^2 + 2s + 1}}{\dfrac{s^3 + 2s}{s^3 + 2s^2 + 2s + 1} + \dfrac{2s^2 + 1}{s^3 + 2s^2 + 2s + 1}}$$

Network 1.

$$y_{12}^{(1)} = \frac{\frac{1}{2}s^3}{s^3 + 2s^2 + 2s + 1} = \frac{n_1}{m_2 + n_2}$$

$$y_{22}^{(1)} = \frac{s^3 + 2s}{s^3 + 2s^2 + 2s + 1} = \frac{n_2}{m_2 + n_2}$$

For the lossless part of this network,

$$y_{22}^{(1)} = \frac{n_2}{m_2} = \frac{s^3 + 2s}{2s^2 + 1}$$

$$A^{(1)} = \frac{n_1}{n_2} = \frac{\frac{1}{2}s^3}{s^3 + 2s}$$

Then

$$Y_b^{(1)} = \frac{n_2 - n_1}{m_2} = \frac{\frac{1}{4}s(s^2 + 4)}{s^2 + \frac{1}{2}}$$

$$Y_A^{(1)} = \frac{n_2 + n_1}{m_2} = \frac{3}{4}\,\frac{s(s^2 + \frac{4}{3})}{s^2 + \frac{1}{2}}$$

Network 2.

$$y_{12}^{(2)} = \frac{-1}{s^3 + 2s^2 + 2s + 1} = \frac{m_1}{m_2 + n_2}$$

$$y_{22}^{(2)} = \frac{2s^2 + 1}{s^3 + 2s^2 + 2s + 1} = \frac{m_2}{m_2 + n_2}$$

Then

$$Y_b^{(2)} = \frac{m_2 - m_1}{n_2} = \frac{2(s^2 + 1)}{s(s^2 + 2)}$$

$$Y_a^{(2)} = \frac{m_2 + m_1}{n_2} = \frac{2s}{s^2 + 2}$$

The over-all network is then as pictured in Fig. 14.

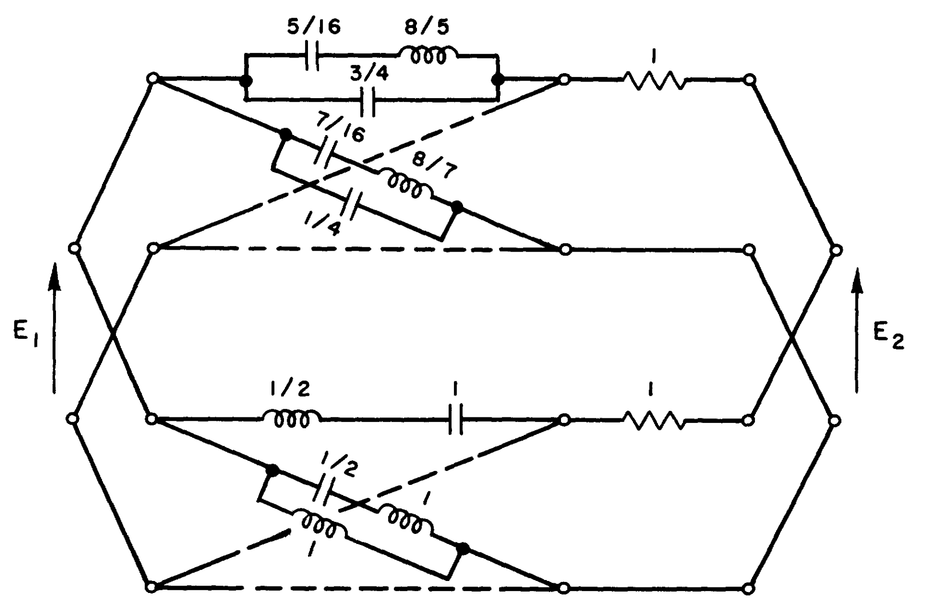

Fig. 14. Parallel lattice network.

c. A Partial-Fraction Method

If the maximum gain is not required, a synthesis in terms of partial fractions is possible. Starting with the expression,

$$\frac{Z_a}{Z_b} = \frac{1 - A}{1 + A}$$

write

$$Z_a = 1 - A$$

$$Z_b = 1 + A$$

$$\left(\text{or } \frac{1}{Z_a} = 1 + A \qquad \text{and } \frac{1}{Z_b} = 1 - A\right)$$

Now expand A in partial fractions. Some of the terms in the expansions for A and $-A$ will be realizable and some not, but those that are not can be made realizable by the addition of a suitable resistor. Consider the one in the expressions $1 - A$ and $1 + A$ as a one-ohm resistance reservoir, which is available for this purpose, and pick K small enough so that the one-ohm resistance is sufficient to make all the terms realizable. K will certainly have to be smaller than the value required to make the real part of A less than one on the j-axis, since we have shown that this value of K is necessary in order for $1 - A$ and $1 + A$ to be positive real.

The resultant RLC partial-fraction canonic forms are well known. Incidentally, the input impedance of this lattice is

$$Z_{11} = \frac{1}{2}(Z_a + Z_b)$$

$$= \frac{1}{2}(1 - A) + (1 + A)$$

$$= 1 \text{ ohm}$$

which is particularly convenient for some applications.

EXAMPLE 1. Consider the function

$$A = \frac{K\,(s^2 + 1)}{(s + 1)\,(s + 2)\,(s + 3)}$$

Expanding A in partial fractions, we obtain

$$A = \frac{\frac{3}{2}K}{s+1} - \frac{6K}{s+2} + \frac{\frac{11}{2}K}{s+3}$$

Then write

$$Z_a = 1 - A = 1 - \frac{\frac{3}{2}K}{s+1} - \frac{\frac{11}{2}K}{s+3} + \frac{6K}{s+2}$$

$$Z_b = 1 + A = 1 - \frac{6K}{s+2} + \frac{\frac{3}{2}K}{s+1} + \frac{\frac{11}{2}K}{s+3}$$

A K of 3/10 will yield realizable partial-fraction terms. Then,

$$Z_a = \frac{\frac{3}{2}s}{s+1} + \frac{\frac{11}{2}s}{s+3} + \frac{\frac{9}{5}}{s+2}$$

$$Z_b = \frac{1}{10} + \frac{\frac{9}{10}s}{s+2} + \frac{\frac{9}{20}}{s+1} + \frac{\frac{33}{20}}{s+3}$$

EXAMPLE 2.

$$A = \frac{K(s^3 - 2)}{(s+1)(s^2+s+1)}$$

Expanding this function in partial fractions, we obtain

$$A = K - \frac{3K}{s+1} + \frac{Ks}{s^2+s+1}$$

Then write

$$Z_a = 1 - A = 1 - K + \frac{3K}{s+1} - \frac{Ks}{s^2+s+1}$$

$$Z_b = 1 + A = 1 + K - \frac{3K}{s+1} + \frac{Ks}{s^2+s+1}$$

A K of ½ will yield realizable partial-fraction terms. Then

$$Z_a = \frac{\frac{3}{2}}{s+1} + \frac{s^2+1}{s^2+2+1}$$

$$Z_b = \frac{\frac{3}{2}s}{s+1} + \frac{\frac{1}{2}s}{s^2+s+1}$$

d. The General Solution

The general solution that will be described was discovered by Kahal (21,16). It is included here for the sake of completeness together with a partial proof.

Given a function

$$\frac{D - KN}{D + KN} = \frac{P(s)}{Q(s)}$$

whose numerator and denominator are both Hurwitz polynomials, and which it is desired to split into the quotient of two positive real impedances, we wish to construct an auxiliary lossless impedance $Z(s)$ so that

$Z(s)$ is positive real

$Z_2(s) = Z(s)\frac{P(s)}{Q(s)}$ is positive real

These are the two desired impedances. In order to establish the conditions for realizability of the second impedance, we must study the behavior of its real part on the j-axis. On this axis, we can write

$$\left.\frac{P(s)}{Q(s)}\right|_{s=j\omega} = R_1(\omega) + jx_1(\omega)$$

$$Z(s) = jx(\omega)$$

Then the real part of their product is

$$Z_2(j\omega) = -x_1(\omega)\,x(\omega)$$

If we are given $\frac{P(s)}{Q(s)}$, this specifies $x_1(\omega)$ and the problem is to pick $x(\omega)$ so that the function $-x_1(\omega)\,x(\omega)$ is always positive on the j-axis. The auxiliary impedance $x(\omega)$ is only needed if $-x_1(\omega)$ should go negative; this can only happen at the odd-order j-axis zeros of $x_1(\omega)$. If $-x_1(\omega)$ is written so as to place these odd-order zeros in evidence,

$$-x_1(\omega) = -j\omega\,(\omega_1{}^2 - \omega^2)\,(\omega_2{}^2 - \omega^2)\ \_\ \_\ x_1'(\omega)$$

where the numerical values refer to the order on the j-axis and $-x_1'(\omega)$ contains no odd-order j-axis zeros, then the required impedance Z(s) can be constructed according to the following rules:

1. If $-x(\omega)$ has a positive sign,

$$-x_1(\omega) = j\omega\,(\omega_1^2 - \omega^2)\,(\omega_2^2 - \omega^2)\;--\;x'(\omega)$$

then pick

$$x(\omega) = \frac{(\omega_1^2 - \omega^2)\,(\omega_3^2 - \omega^2)\;--}{j\omega(\omega_2^2 - \omega^2)\;--}$$

so that

$$-x_1(\omega)\,x(\omega) = (\omega_1^2 - \omega^2)^2\,(\omega_3^2 - \omega^2)^2\;--\;x'(\omega)$$

2. If $-x_1(\omega)$ has a negative sign,

$$-x_1(\omega) = -j\omega\,(\omega_1^2 - \omega^2)\,(\omega_2^2 - \omega^2)\;--\;x'(\omega)$$

then pick

$$x(\omega) = \frac{j\omega(\omega_2^2 - \omega^2)\;--}{(\omega_1^2 - \omega^2)\,(\omega_3^2 - \omega^2)\;--}$$

so that

$$-x_1(\omega)\,x(\omega) = \omega^2\,(\omega_2^2 - \omega^2)^2\,(\omega_4^2 - \omega^2)^2\;--\;x'(\omega)$$

In both cases the real part of $Z_2(s)$ is positive. Then the impedances Z(s) and $Z_2(s)$ are known to be positive real, provided that it can be proved that their j-axis poles are simple with positive real residues. Kahal and Fialkow and Gerst give proofs of this last point, but we shall not include the details of their proofs.

EXAMPLE. The function

$$A = \frac{K(s^2 + 2)}{(s + 1)\,(s + 2)\,(s + 3)}$$

was considered in many of the previous examples, in which it was shown that K = 3 is the greatest obtainable gain. For

$$A = \frac{3\,(s^2 + 2)}{(s + 1)\,(s + 2)\,(s + 3)}$$

we can write

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^3 + 3s^2 + 11s}{s^3 + 9s^2 + 11s + 12}$$

To split this into Z_a and Z_b, we first form the imaginary part on the j-axis

$$\mathrm{Im}\left\{\frac{Z_a}{Z_b}(j\omega)\right\} = -\frac{16\omega\,(2 - \omega^2)\,(11 - \omega^2)}{(12 - 9\omega^2)^2 + \omega^2\,(11 - \omega^2)^2}$$

$$= -j\omega\,(2 - \omega^2)\,(11 - \omega^2)\,x'(\omega)$$

Then the reactive impedance Z(s) is

$$x_2(\omega) = \frac{j\omega\,(11 - \omega^2)}{2 - \omega^2}$$

$$Z(s) = \frac{s\,(s^2 + 11)}{s^2 + 2}$$

and we can form Z_a and Z_b as follows

$$Z_a = \frac{s\,(s^2 + 11)}{s^2 + 2} = s + \frac{9s}{s^2 + 2}$$

$$Z_b = \frac{(s^3 + 9s^2 + 11s + 12)\,(s^2 + 11)}{(s^2 + 3s + 11)\,(s^2 + 2)} = s + \frac{9s}{s^2 + 2} + \cfrac{1}{\cfrac{1}{6} + \cfrac{\frac{1}{2}s}{s^2 + 11}}$$

V. UNBALANCING LATTICES

1. INTRODUCTION

When a grounded symmetrical network is desired, the lattice is still used as a starting point in the synthesis procedure; the only additional problem involves converting the lattice to unbalanced form. Three well-known steps are available for this purpose.

1. Series impedances, common to Z_a and Z_b can be removed from the lattice as series elements (Fig. 15a).

2. Shunt admittances common to Z_a and Z_b can be removed from the lattice as shunt elements (Fig. 15b).

3. The lattice can be split into two parallel lattices and each one separately converted to unbalanced form, after which the two grounded networks are connected in parallel without the need of an ideal transformer (Fig. 15c).

The purpose of this section is to show that for the two-element-kind case a combination of these steps will always succeed in unbalancing the lattice provided that the transfer function has positive coefficients, a condition which was proved necessary in Section II. If a given two-element-kind voltage transfer function has all positive coefficients (in the general case, some decrease in gain may be necessary), a synthesis procedure is always possible in which the function is first realized as a lattice and then converted to unbalanced form. The methods are also applicable, with only slight alteration, to the synthesis of transfer impedances and admittances.

It is strongly believed that the statement of the two-element-kind case is also true for the three-element case, but no proof is available at present. However, such a conversion is always possible in the RLC case if the maximum gain is not required. As shown in Section IV, an RLC function can always be realized (with possibly a slight decrease in gain) as a parallel combination of two LC lattices each of which is in series with a resistor. The methods of this section can then be used to unbalance the lossless lattices.

The proof of the two-element case will consist of considering certain particular types of voltage transfer function in which it can easily be shown that the lattice can be unbalanced. Then the general case will be considered as a combination of these particular cases and a combination of the methods used in the simple cases will be shown to work always, as long as the given function has positive coefficients. As usual, the details will be given for RC functions, but the results will apply to all two-element-kind functions.

2. CASE 1 - THE TRANSFER FUNCTION HAS ALL ITS ZEROS AT ZERO OR INFINITE FREQUENCY

If the transfer function is of the form

$$A = \frac{K\,s^P}{s^n + b_{n-1}s^{n-1} + b_1 s + b_o}$$

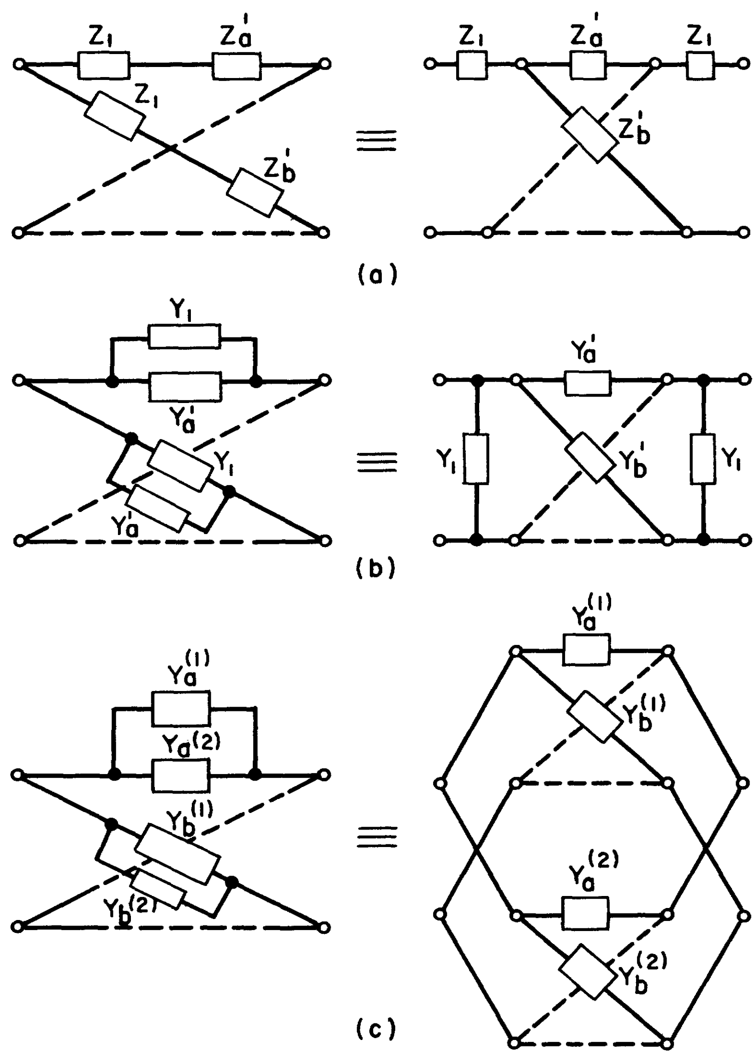

Fig. 15. Unbalancing lattices.

where $P \leqslant m$, a particularly easy procedure is possible. Let us first consider the zeros at infinite frequency.

If the transfer function has a zero at infinite frequency, then, at that frequency, $Z_a = Z_b$ and $Y_a = Y_b$, since it is the bridge action that causes the zeros. Now, in the RC case, at infinite frequency, the impedances can only behave like a resistor or a capacitor:

$$Z_a(\infty) = Z_b(\infty) = R$$

$$Z_a(\infty) = Z_b(\infty) = \frac{1}{sC(\infty)}$$

If the impedances are finite at infinite frequency, then the resistor R_∞ can be removed from Z_a and Z_b, leaving both of them positive real and RC. If the impedances have a zero at infinity, then the admittances have a pole; this pole can be removed from Y_a and Y_b as a shunt capacitor, C_∞, leaving both the admittances positive real and RC.

Now, if the original transfer function had a double-order zero at infinite frequency, the remaining lattice still has a zero at that frequency and the impedances and admittances of this remaining lattice are still equal at infinite frequency ($Z_a' = Z_b'$). Then the process can be repeated; if the first time a series resistor was removed from the impedances, this time a shunt capacitor will be removed from the admittances and vice versa. These steps can be repeated until all the zeros at infinite frequency have been developed.

The same reasoning shows that if the transfer function has a zero at zero frequency, then, since $Z_a(O) = Z_b(O)$, either a series capacitor or a shunt conductance can be removed from the lattice. Then, if the original function has a double-order zero at zero frequency, the remaining impedances are still equal at zero frequency and another element can be removed. The process can be repeated until all the zeros at zero frequency have been developed.

If the transfer function has all its zeros at either zero or infinite frequency, then, at any point in the procedure, a zero at either infinite frequency or zero frequency can be developed by removing the appropriate elements, and the process can always be continued until all the zeros have been developed and the lattice is then unbalanced. Because of the choice in the order of removal of zeros, a variety of configurations is frequently possible.

The process will always unbalance the lattice if all the transmission zeros are at zero or infinite frequency. If only some of the zeros are at zero or infinite frequency, these methods can still be used to develop these zeros and thus simplify the given lattice.

EXAMPLE. Consider the voltage transfer function

$$A = \frac{K_1 s}{(s+1)(s+2)(s+3)}$$

It can be readily calculated, by using the methods of Section III, that the maximum gain for a lattice realization is K = 0.06. For

$$A = \frac{0.06\,s^2}{(s+1)(s+2)(s+3)}$$

we have

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{(s+0.96)(s+2.49)^2}{(s+1.06)(s+1.76)(s+3.24)}$$

This function can be split into Z_a and Z_b in a number of ways either with or without surplus factors; the method will work for any of them. For example,

$$Z_a = \frac{s + 2.49}{(s+1.06)(s+3.24)}$$

$$Z_b = \frac{s + 1.76}{(s+0.96)(s+2.49)}$$

Since the given voltage transfer function has a simple zero at infinite frequency and a double zero at zero frequency, the following continued-fraction expansion is appropriate:

$$\frac{1}{Z_a} = s + 1.37 + \cfrac{1}{\cfrac{0.056}{s} + 0.023}$$

$$\frac{1}{Z_b} = s + 1.37 + \cfrac{1}{\cfrac{0.056}{s} + 0.031}$$

The unbalanced ladder network can now be formed by inspection, as shown in Fig. 16. The two elements on the extreme left can be omitted if a voltage excitation is applied, leaving a resistor in series with the voltage source.

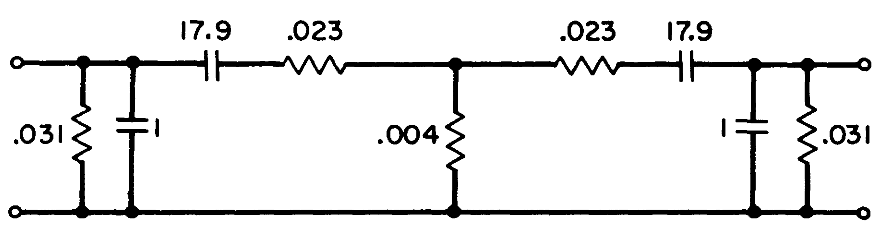

Fig. 16. An unbalanced lattice.

3. NEGATIVE REAL TRANSMISSION ZEROS

If all of the finite, nonzero transmission zeros are on the negative real axis (for RC networks), a procedure that is analogous to zero-shifting is possible. This process, like zero-shifting, involves using surplus factors. Since the function

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN}$$

possesses the double alternation property, a polynomial, P (s), can always be found whose zeros are all on the negative real axis and alternate with the zeros of both D – KN and D + KN, so that

$$Z_a = \frac{D - KN}{P(s)}$$

and

$$Z_b = \frac{D + KN}{P(s)}$$

are both realizable RC impedances. For this particular method, the surplus polynomial is chosen in the following way.

$$P(s) = KN'Q(s)$$

The polynomial KN′ consists of all the factors of the numerator polynomial KN, which can be used and still have the impedances, Z_a and Z_b, RC. If, for instance, KN has two zeros between two of the roots of D – KN or D + KN, then only one of these factors can be used in KN′. The remaining part, Q(s), is the smallest polynomial that it is necessary to add in order to make the impedances RC. Its zeros are arbitrary, subject only to the realizability conditions of the impedances. Two separate cases are recognizable:

1. KN′ contains all the factors of KN.
2. KN does not contain all the factors of KN.

These two cases will be considered separately.

1. KN′ = KN. In this case we can write

$$Z_a = \frac{D - KN}{KN\ Q(s)}$$

$$Z_b = \frac{D + KN}{KN\ Q(s)}$$

If Z_a and Z_b are expanded in partial fractions, the following results will be obtained.

A. The poles corresponding to the zeros of KN have the same residue (as they must, since $Z_a = Z_b$ wherever the transfer function has a zero), and these poles can be completely removed from the lattice as series elements. This step corresponds to total pole removal in zero-shifting.

B. The poles corresponding to the zeros of Q(s) will not have the same residues in Z_a and Z_b but some of each pole (the minimum of Z_a and Z_b) can be removed from the lattice. This step corresponds to partial pole removal in zero-shifting.

C. The lattice that remains has a transfer impedance

$$Z_{12} = \frac{1}{2}(Z_b - Z_a)$$

$$= \frac{1}{2}\left(\frac{D + KN}{KN\ Q(s)} - \frac{D - KN}{KN\ Q(s)}\right) = \frac{1}{Q(s)}$$

Since all of the transmission zeros are at infinite frequency, the methods of the previous section can be used to unbalance the remaining lattice. The general form of the over-all structure appears in Fig. 17, in which the extraneous series elements at the output end have been omitted.

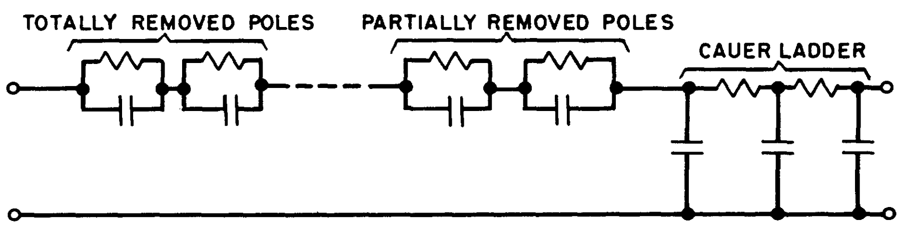

Fig. 17. A zero-shifted ladder.

EXAMPLE. The function

$$A = \frac{(s + 2)(s + 4)}{(s + 1)(s + 3)(s + 5)(s + 7)}$$

has its transmission in the proper places, but it is not an RC function. Then

$$\frac{Z_a}{Z_b} = \frac{D - KN}{D + KN} = \frac{s^4 + 16s^3 + 85s^2 + 170s + 97}{s^4 + 16s^3 + 87s^2 + 182s + 113}$$

Split this into

$$Z_a = \frac{s^4 + 16s^3 + 85s^2 + 170s + 97}{s(s + 2)(s + 4)(s + 6)}$$

$$Z_b = \frac{s^4 + 16s^3 + 87s^2 + 182s + 113}{s(s+2)(s+4)(s+6)}$$

where the poles at s = -2 and -4 will produce the transmission zeros, and the remaining poles are arbitrary. Expanding these functions in partial fractions, we obtain

$$Z_a = 1 + \frac{\frac{15}{16}}{s+2} + \frac{\frac{9}{16}}{s+4} + \frac{\frac{23}{48}}{s+6} + \frac{97}{48s}$$

$$Z_b = 1 + \frac{\frac{15}{16}}{s+2} + \frac{\frac{9}{16}}{s+4} + \frac{\frac{7}{48}}{s+6} + \frac{113}{48s}$$

As was expected, the poles caused by the transmission zeros of A can be completely removed, producing the transmission zeros, while the surplus poles at s = 0 and -6 can only be partially removed. After these common factors have been removed, the remaining lattice has the impedances:

$$Z_a' = \frac{\frac{1}{3}}{s+6} = \frac{1}{3s+18}$$

$$Z_b' = \frac{1}{3s}$$

This lattice has all its transmission zeros at infinite frequency and can be unbalanced by inspection. The resultant unbalanced network is shown in Fig. 18,in which the extraneous elements on the right have been omitted.

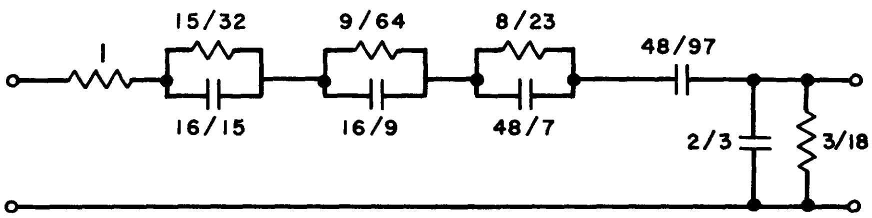

Fig. 18. An unbalanced lattice.

2. KN′ does not contain all the factors of KN. This case is inherently more complicated than the first; it corresponds to zero-shifting when more than one shift is necessary to realize one zero.

In this case,

$$Z_a = \frac{D - KN}{KN' Q(s)}$$

$$Z_b = \frac{D + KN}{KN' Q(s)}$$

The following steps are necessary.

A. Since the residues in the poles of Z_a and Z_b attributable to KN' are equal, these poles can be completely removed from the lattice as series elements, thus realizing these transmission zeros.

B. After removing these poles, the impedances of the remaining lattice are of the form

$$Z_a = \frac{Q_1(s)}{Q(s)}$$

$$Z_b = \frac{Q_2(s)}{Q(s)}$$

where Q_1 (s) and Q_2 (s) must be equal at all of the undeveloped transmission zeros. The object of this step is to reduce the complexity of the lattice by creating zeros of admittance (or impedance) simultaneously in both lattice admittances, and then removing the corresponding poles of impedance (or admittance). The conventional techniques of zero-shifting are used. By partially removing, either one of the poles attributable to Q (s), or the constant value at infinite or zero frequency, a zero is shifted to the correct place simultaneously in both admittances (or impedances). Since the admittances (and impedances) are equal at the frequencies of the transmission zeros, if a zero is created in one of them, at one of these frequencies, because of the subtraction of a suitable element, then subtraction of that element from the other will also produce a zero at the same frequency.

The zeros that are produced in both admittances (or impedances) will not have the same slope unless the original transmission zero was of second order so that the poles in the inverse function will not have the same residue. (The fact that a first-order transmission zero exists, just makes the impedances equal at that frequency; a second-order zero makes their slopes equal, and so forth.) Therefore only part of that pole can be removed from the lattice and it may appear that the transmission zero has not been developed. However, this is not true; part of the poles has been removed at each end of the lattice, and if the entire unbalanced lattice is considered as a single development of a driving-point function, then one of the terms represents the partial removal of a pole and the other term represents the removal of the rest of the pole, thus producing the transmission zero. This point will be clarified in the example at the end of the section.

After the development of one transmission zero, the process is repeated; the lattice impedances must still be equal to all the undeveloped zeros. (Sometimes it is possible to shift more than one set of zeros simultaneously by removing some of each of the poles of Q(s) at the same time.) The only difficulty that may be encountered, in addition to those usually encountered in zero-shifting, is that the shifting must be done at the same time in both Z_a and Z_b; thus the same factors must be removed from both. At the first step, both impedances have the same denominator so that no difficulty arises, but after one zero-shifting step, the numerators and denominators of both impedances will, in general, be different. Then it is necessary to perform the zero-shifting with the zero or infinite frequency values of the impedances or admittances. It cannot be proved that these methods will always work, but it is considered highly probable because of the freedom allowed in the order of the shifting process, as well as in the choice of Q(s). If these methods do not work, the methods described in the next section can always be used to unbalance the lattice.

C. After all the finite transmission zeros have been developed, the remaining lattice has all of its transmission zeros at zero or infinite frequency and can be unbalanced by the methods appropriate to such lattices.

EXAMPLE.

$$A = \frac{1}{8}\,\frac{(s+2)^2\,(s+3)}{(s+1)\,(s+4)\,(s+6)}$$

For this function,

$$\frac{Z_a}{Z_b} = \frac{D-KN}{D+KN} = \frac{7s^3+81s^2+256s+180}{9s^3+95s^2+288s+204}$$

One of the transmission zeros can be produced by the methods of the previous section. Let

$$Z_a = \frac{7s^3+81s^2+256s+180}{s(s+2)(s+5)}$$

$$Z_b = \frac{9s^3+95s^2+288s+204}{s(s+2)\,(s+5)}$$

where the pole at $s = -2$ will produce the transmission zero; the other poles were picked arbitrarily. Removing the pole at $s = -2$ yields

$$Z_a = \frac{\frac{32}{3}}{s+2} + \frac{7s^2+\frac{169}{3}s+90}{s(s+5)}$$

$$Z_b = \frac{\frac{32}{3}}{s + 2} + \frac{9s^2 + \frac{199}{3} s + 102}{s(s + 5)}$$

The zero at s = −2 is now developed. The remaining lattice still contains the other zeros at s = −2 and s = −3; therefore Z_a' must be equal to Z_b' at these frequencies.

$$Z_a' (-2) = Z_b' (-2) = -\frac{8}{9}$$

$$Z_a' (-3) = Z_b' (-3) = \frac{8}{3}$$

This is a partial check on the work.

In this problem we cannot simultaneously shift both of these zeros to the proper places because then Z_a' and Z_b' would have zeros at s = −2 and −3 which would not be interlaced with the poles. It is, therefore, necessary to zero-shift one zero at a time. Partially removing the value at infinity will shift one of the zeros to s = −3.

$$Z_a = \frac{\frac{16}{3}}{s + 2} + \frac{8}{3} + \frac{(s + 3)(\frac{13}{3} s + 30)}{s(s + 5)}$$

$$Z_b = \frac{\frac{16}{3}}{s + 2} + \frac{8}{3} + \frac{(s + 3)(\frac{17}{3} s + 34)}{s(s + 5)}$$

Inverting the remaining function and expanding in partial fractions, we obtain

$$Z_a = \frac{\frac{16}{3}}{s + 2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{17} s}{s + 3} + \cfrac{0.113\,s}{s + 6.9}}$$

$$Z_b = \frac{\frac{16}{3}}{s + 2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{15} s}{s + 3} + \cfrac{0.0246\,s}{s + 5.36}}$$

The residues in the pole at s = −3 are not equal. This is to be expected; we have already used the fact that $Z_a' = Z_b'$ at a transmission zero, when we shifted a zero to that frequency in both impedances at the same time. We cannot expect that the slope at that zero will be the same in both impedances, unless the original transmission zero was of double order. Therefore, only part of the pole at s = −3 can be removed at this time, and it appears that we have not created a transmission zero. However, we are really removing part of the pole at each end of the lattice; we shall see later that this actually will produce the transmission zero. Removing part of this pole at s = −3, we have

$$Z_a = \frac{\frac{16}{3}}{s+2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{17}s}{s+3} + \cfrac{0.113\,s}{s+6.9}}$$

$$Z_b = \frac{\frac{16}{3}}{s+2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{17}s}{s+3} + \cfrac{0.04035\,s(s+3.91)}{(s+3)(s+5.36)}}$$

The process can be continued in a straightforward fashion, but the remaining lattice is so simple that it can be unbalanced by inspection.

$$Z_a = \frac{\frac{16}{3}}{s+2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{17}s}{s+3} + \cfrac{1}{8.85 + \frac{61}{s}}}$$

$$Z_b = \frac{\frac{16}{3}}{s+2} + \frac{8}{3} + \cfrac{1}{\cfrac{\frac{2}{17}s}{s+3} + \cfrac{1}{8.85 + \frac{61}{s} + \cfrac{1}{\frac{0.04s}{s+2} + \frac{0.023\,s}{s+5}}}}$$

The unbalanced network is shown in Fig. 19, in which the extraneous series elements at the right end have been omitted. At this point, we might consider how the different zeros have been developed. There is actually no question as to whether or not the ladder has the proper zeros, since the original lattice had all the zeros and any development of the lattice must also have those zeros, but it is interesting to see how the zeros are produced.

1. Branch A produces one transmission zero at s = −2.
2. Branch C produces the other transmission zero at s = −2.
3. Branches B and F produce the transmission zero at s = −3. Branch B represents a partial

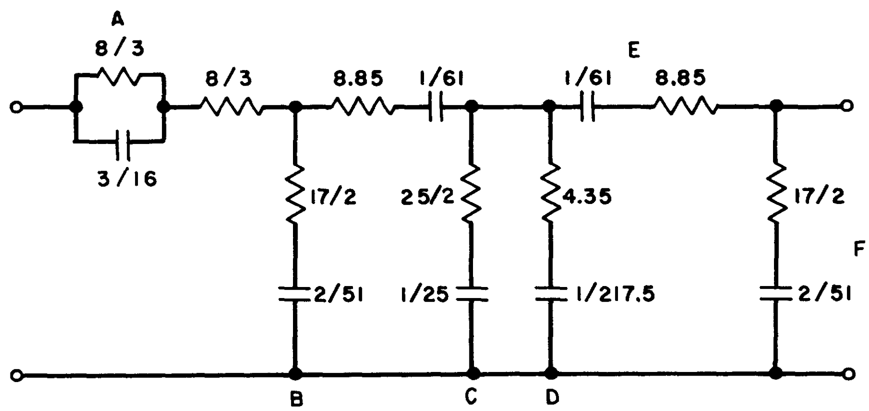

Fig. 19. A zero-shifted ladder.

removal of a pole of admittance and branch F represents the removal of the rest of the pole, thus producing the transmission zero.

4. Branch D does not produce a transmission zero at its resonant frequency ($s = -4$) because the network to its right (E and F) also contains a zero of impedance at this frequency.

4. PARALLEL LATTICES

When some of the zeros of the transfer function are off the negative real axis (for RC functions), the lattice cannot be converted into a single ladder. Some portion of the resultant unbalanced network must contain parallel paths in order to bring about the cancellation that produces these zeros. In order to accomplish this cancellation in an unbalanced equivalent of a lattice, the given lattice is first split into two or more parallel lattices and then each of these lattices is converted into a ladder. The resultant parallel ladder's configuration produces the required cancellation. This splitting into parallel lattices need not be done at the beginning of the procedure; it can be done after zeros have been developed at zero frequency, at infinite frequency, and on the negative real axis.

A simple example of the parallel lattice method has been given by Guillemin (13). He shows that a lattice of the type shown in Fig. 20a can be split into two equivalent lattices, as in Fig. 20b, and then each lattice can be converted to unbalanced form, as in Fig. 20c, provided that the following condition is satisfied by the original impedances.

$$\frac{R_b}{R_a} + \frac{C_a}{C_b} \geqslant 1$$

It can be easily shown that these are the most general conditions under which this lattice can be unbalanced by simply writing the voltage transfer function for this lattice:

$$A = \frac{Z_b - Z_a}{Z_b + Z_a} = \frac{s^2 + \frac{1}{R_bC_a}\left(\frac{R_b}{R_a} + \frac{C_a}{C_b} - 1\right) s + \frac{1}{R_aR_bC_aC_b}}{s^2 + \frac{1}{R_bC_a}\left(\frac{R_b}{R_a} + \frac{C_a}{C_b} + 1\right) s + \frac{1}{R_aR_bC_aC_b}}$$

The condition stated by Guillemin is seen to be the condition in which A has positive conditions, and, therefore, it is the most general condition for unbalancing the lattice.

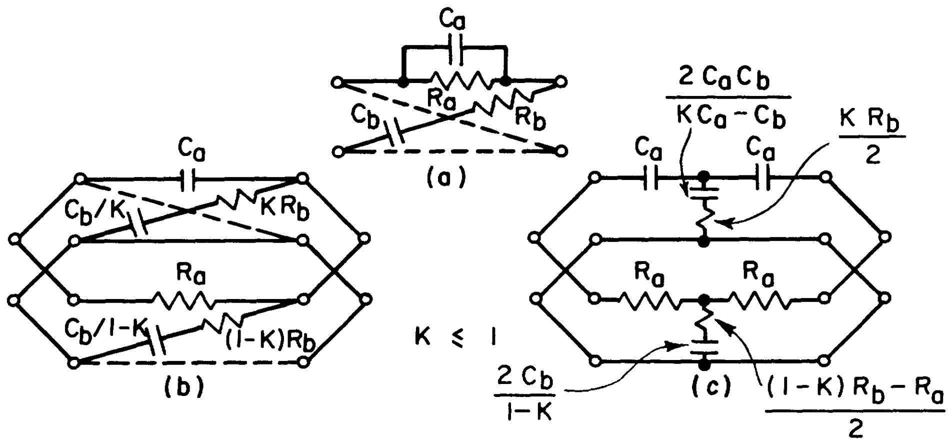

Fig. 20. Guillemin's parallel lattices.

a. General Procedures

In order to split a lattice into parallel lattices and insure that each of the parallel lattices is convertible to unbalanced form, the following rules must be followed:

1. $Y_a = Y_{a1} + Y_{a2} + --$
   $Y_b = Y_{b1} + Y_{b2} + --$
2. $Y_{a1} - Y_{b1}$ must have all positive coefficients
   $Y_{a2} - Y_{b2}$ must have all positive coefficients

In the following discussion, impedances or admittances without numerical subscripts (Y_a, Z_b) refer to the original lattice, while elements with subscripts (Y_{a1}, Z_{b2}) refer to the individual parallel lattices. The first rule states that the parallel combination of all the lattices must equal the original lattice, and the second rule states that each parallel lattice must be convertible to unbalanced form. If each parallel lattice is to be convertible to one of the ladders described in the previous sections, rule 2 must read,

2'. $Y_{a1} - Y_{b1}$ must have all negative real zeros or it must have all its zeros at zero or infinite frequency, depending on the ladder desired.

It is relatively easy to satisfy each of these conditions individually but it is more difficult to satisfy both conditions simultaneously. For example,

1. To satisfy condition 1, expand Y_a and Y_b in partial fractions and consider part of each term as belonging to each lattice. This splitting can be done in an infinite number of ways; however, for any particular way, there is no guarantee that the transfer function of each individual lattice will have positive coefficients. This method has not yet yielded a satisfactory realization method.

2. To satisfy the positive coefficient condition, consider the transfer admittance

$$y_{12} = \frac{1}{2}(Y_a - Y_b) = \frac{KN}{P(s)}$$

where KN is assumed to have all positive coefficients. Split KN into two or more polynomials, each of which has all positive coefficients.

$$\frac{KN}{P(s)} = \frac{N_1}{P(s)} + \frac{N_2}{P(s)} = y_{12,1} + y_{12,2}$$

Consider each of these functions to be the transfer admittance of one parallel lattice. To obtain the lattice impedances of each of these lattices, expand $y_{12,1}$ and $y_{12,2}$ in partial fractions, calling the terms with positive residues $\frac{1}{2}Y_{a1}$ and $\frac{1}{2}Y_{a2}$ and those with negative residues $\frac{1}{2}Y_{b1}$ and $\frac{1}{2}Y_{b2}$. However, when this is done there is no guarantee that condition 1 has been satisfied, namely, that

$$Y_a = Y_{a1} + Y_{a2}$$

$$Y_b = Y_{b1} + Y_{b2}$$

To clarify this last point, consider a lattice in which

$$Y_a = 2s + \frac{1}{3} + \frac{7s}{s+2}$$

$$Y_b = \frac{\frac{52}{3}}{s+3}$$

and

$$y_{12} = \frac{1}{2}(Y_a - Y_b) = \frac{s^3 + 1}{(s+2)(s+3)} = \frac{KN}{P(s)}$$

If y_{12} were broken into

$$y_{12,1} = \frac{s^3}{(s+2)(s+3)} = s + \frac{4s}{s+2} - \frac{9s}{s+3}$$

$$y_{12,2} = \frac{1}{(s+2)(s+3)} = \frac{1}{6} + \frac{\frac{1}{3}s}{s+3} - \frac{\frac{1}{2}s}{s+2}$$

it is apparent that the parallel combination of these two lattices would not be equivalent to the original lattice, since one of the terms from branch A of the original lattice appears in branch B of one of the parallel lattices, and one of the terms of branch B from the original lattice appears in branch A of one of the parallel lattices. The two lattices in parallel have the same transfer admittance as the original (i.e., by construction), but they have different input admittances and hence different voltage transfer functions.

In the general case, in order to insure that the parallel combination of the two lattices is equivalent to the original lattice, it is sufficient (but not necessary, as shown later) to require that, at each of the poles y_{12}, the three polynomials KN, N_1, and N_2 have the same sign. (The signs may be positive at one pole and negative at another in any order but, at any particular pole, the three polynomials have the same sign.) Then, in the partial-fraction expansion of y_{12}, $y_{12,1}$, $y_{12,2}$, all three of the terms corresponding to a particular pole will have the same sign so that, if that sign is positive, the term will belong to branch A of all the lattices and if negative, to branch B.

This requirement is automatically satisfied in one very important practical case—when the numerator is an odd or an even polynomial. (This includes the case with all the zeros on the j-axis.)

5. UNBALANCING THE LATTICE WHEN THE NUMERATOR IS ALL EVEN OR ALL ODD

The details of this method will be presented for the numerator that is even; the extension to the other case is straightforward.

The lattice admittances Y_a and Y_b are formed from the function

$$\frac{Y_b}{Y_a} = \frac{D - KN}{D + KN}$$

in any convenient manner. Then the transfer admittance is computed.

$$y_{12} = \frac{1}{2}(Y_a - Y_b) = \frac{KN}{P(s)}$$

where KN is of the form

$$KN = a_0 + a_2 s^2 + a_4 s^4 + - - a_{2n} s^{2n}$$

This polynomial is split as follows:

$$N_1 = a_0$$

$$N_2 = a_2 s^2$$

$$N_3 = a_4 s^4$$

$$N_{n+1} = a_{2n} s^{2n}$$

The transfer admittances of the individual lattices are then

$$Y_{12,1} = \frac{a_0}{P(s)}$$

$$Y_{12,2} = \frac{a_2 s^2}{P(s)}$$

$$Y_{12,n+1} = \frac{a_{2n} s^{2n}}{P(s)}$$

The sum of these transfer admittances is clearly the original transfer admittance. Now the individual lattice admittances of each parallel lattice can be formed from a partial-fraction expansion of its transfer admittance, and, since the numerators of all the transfer admittance are positive on the whole negative real axis, as is the polynomial KN (provided that KN has all positive coefficients), the conditions on the residues are satisfied, and the sum of all these lattice admittances, which represents the driving-point admittance of all the lattices connected in parallel, will indeed be equal to the driving-point admittance of the given lattice.

Since each of the individual parallel lattices has all its transmission zeros at zero or infinite frequency, they can all be easily unbalanced by using an appropriate Cauer expansion.

EXAMPLE. The set of short-circuit admittances

$$y_{12} = \frac{s^4 + 1}{(s+1)(s+2)(s+3)} = \left(s + \frac{1}{6} + \frac{\frac{17}{2}s}{s+2}\right) - \left(\frac{s}{s+1} + \frac{\frac{41}{3}s}{s+3}\right)$$

$$y_{11} = y_{22} = \frac{s^4 + \frac{88}{3}s^3 + 92s^2 + \frac{200}{3}s + 1}{(s+1)(s+2)(s+3)} = s + \frac{1}{6} + \frac{\frac{17}{2}s}{s+2} + \frac{s}{s+1} + \frac{\frac{41}{3}s}{s+3}$$

correspond to a voltage transfer function

$$A = \frac{s^4 + 1}{s^4 + \frac{88}{3}s^3 + 92s^2 + \frac{206}{3}s + 1}$$

To split this into two parallel lattices,

$$y_{12,1} = \frac{s^4}{(s+1)(s+2)(s+3)} = s + \frac{8s}{s+2} - \frac{\frac{1}{2}s}{s+1} + \frac{\frac{27}{2}s}{s+3}$$

$$y_{12,2} = \frac{1}{(s+1)(s+2)(s+3)} = \frac{1}{6} + \frac{\frac{1}{2}s}{s+2} - \frac{\frac{1}{2}s}{s+1} + \frac{\frac{1}{6}s}{s+3}$$

where it is readily apparent that

$$Y_a = Y_{a1} + Y_{a2}$$

$$Y_b = Y_{b1} + Y_{b2}$$

Since each of the individual lattices has all its zeros at zero or infinite frequency, an appropriate Cauer form is used.

Lattice 1.

$$Y_{a1} = \frac{2(s^2 + 10s)}{s+2} = \cfrac{1}{\cfrac{1}{10s} + \cfrac{1}{\cfrac{1}{25} + \cfrac{1}{\cfrac{2}{5s}}}}$$

$$Y_{b1} = \frac{2(14s^2 + 15s)}{(s+1)(s+3)} = \cfrac{1}{\cfrac{1}{10s} + \cfrac{1}{\cfrac{1}{25} + \cfrac{1}{\cfrac{2}{5s} + \cfrac{1}{3}}}}$$

Lattice 2.

$$Y_{a2} = \frac{2(\frac{2}{3}s + \frac{1}{3})}{s+2} = \cfrac{1}{\cfrac{3}{4} + \cfrac{1}{\cfrac{8}{9}s + \cfrac{1}{\cfrac{9}{4}}}}$$

$$Y_{b2} = \frac{2\left(\frac{2}{3}s^2 + \frac{10}{9}s\right)}{(s+1)(s+3)} = \cfrac{1}{\cfrac{3}{4} + \cfrac{1}{\cfrac{8}{9}s + \cfrac{1}{\cfrac{9}{4} + \cfrac{1}{\cfrac{2}{9}s}}}}$$

The complete unbalanced network is shown in Fig. 21.

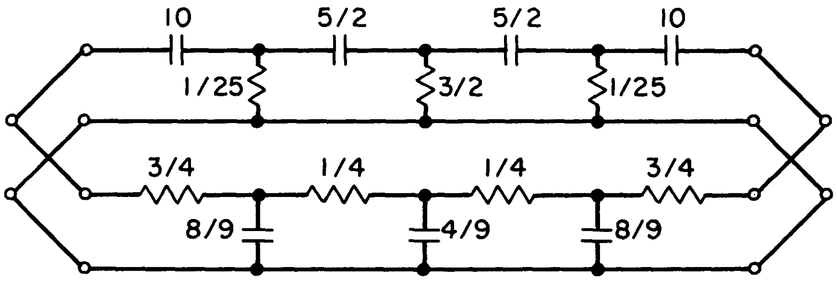

Fig. 21. Parallel ladder development.

6. THE GENERAL CASE

In the general case, when the numerator is neither odd nor even, it is not readily apparent how the numerator can be split in the required manner. We present here two methods that will always unbalance the lattice provided KN has all positive coefficients, but which require a reduction from the maximum gain and will also use surplus factors. It is strongly suspected that a method can be developed which does not involve either of these inconveniences.

Both methods consist of, first, splitting the numerator into its odd and even parts, $KN = N_o + N_e$, that is, splitting the given lattice into two lattices, one with an even numerator and one with an odd numerator, each of which can be unbalanced by using the methods of the previous section. It is immediately apparent that this split does not obey the sufficient condition given previously; N_o and N_e do not have the same sign at each of the poles of y_{12}; in fact, they have opposite signs everywhere in the negative real axis. Nevertheless, it will be shown that the split can still be made, but at the expense of surplus factors and a decrease in gain.

Method 1. The given voltage transfer function is of the form

$$A = \frac{KN}{D} = \frac{N_o + N_e}{D} = \frac{y_{12}}{y_{22}}$$

We are interested in splitting this into parallel lattices, using the expression

$$A = \frac{y_{12}}{y_{22}} = \frac{y_{12,1} + y_{12,2}}{y_{22,1} + y_{22,2}}$$

Introducing a surplus polynomial P(s), which will be determined later, we can write the voltage transfer function in the form:

$$A = \frac{\dfrac{KN_o}{P(s)} + \dfrac{KN_e}{P(s)}}{\dfrac{\frac{1}{2}D}{P(s)} + \dfrac{\frac{1}{2}D}{P(s)}}$$

The lattice has now been split into two parallel lattices.

For lattice 1, $\quad y_{12,1} = \dfrac{KN_o}{P(s)} \qquad y_{22,1} = \dfrac{D}{2P(s)}$

$$A_1 = \frac{2\ KN_o}{D}$$

For lattice 2, $\quad y_{12,2} = \dfrac{KN_e}{P(s)} \qquad y_{22,2} = \dfrac{D}{sP(s)}$

$$A_2 = \frac{2\ KN_e}{D}$$

Forgetting about the surplus polynomial P(s) for a moment, we see that the maximum gain obtainable is the value of K for which both lattices are realizable; that is, the value of K for which

$$D - 2\ KN_e \qquad D - 2\ KN_o$$

$$D + 2\ KN_e \qquad D + 2\ KN_o$$

all have their zeros on the negative real axis. This value of K is of the order of one-half the maximum K obtainable from the original lattice and may be considerably less than this value.

There is no choice in the construction of the lattice admittances

$$Y_{b1} = \frac{D - 2\ KN_e}{P(s)}$$

$$Y_{a1} = \frac{D + 2\ KN_e}{P(s)}$$

$$Y_{b2} = \frac{D - 2\,KN_o}{P(s)}$$

$$Y_{a2} = \frac{D + 2\,KN_o}{P(s)}$$

Notice that the ones and minus ones of the lattice are produced only by the zeros of Y_a and Y_b, while the poles are surplus factors. The surplus polynomial P(s) is determined by the requirement that its zeros alternate with the zeros of each of the polynomials

$$D - 2\,KN_o \qquad D - 2\,KN_e$$

$$D + 2\,KN_o \qquad D + 2\,KN_e$$

The methods of root locus show that for small enough value of K, it is always possible to find such a polynomial P(s). In fact, if we consider the polynomials two at a time, a polynomial P(s) can always be found whose zeros alternate with the zeros of both $D - 2\,KN_o$ and $D + 2\,KN_o$, for example, for any K less than that required for the polynomials to have all their zeros on the negative axis. If we consider the four polynomials at once, a slight reduction in gain may be necessary.

This completes the synthesis; the previously derived methods can be used to unbalance the lattice.

EXAMPLE. Consider the voltage transfer function

$$A = \frac{s^3 + 1}{s^3 + \frac{52}{3} s^2 + \frac{104}{3} s + 1}$$

for which K = 1 is the highest obtainable gain. For this method, a constant of one-half, at most, can be associated with the function. Using this constant, we can split the function in the appropriate manner

$$A = \frac{\dfrac{\frac{1}{2} s^3}{P(s)} + \dfrac{\frac{1}{2}}{P(s)}}{\dfrac{\frac{1}{2}(s^3 + \frac{52}{3} s^2 + \frac{104}{3} s + 1)}{P(s)} + \dfrac{\frac{1}{2}(s^3 + \frac{52}{3} s^2 + \frac{104}{3} s + 1)}{P(s)}}$$

P(s) will be determined later. The lattice is now split into two parallel lattices.

For lattice 1,

$$\frac{Y_{b1}}{Y_{A1}} = \frac{\dfrac{D - 2\,KN_o}{P(s)}}{\dfrac{D + 2\,KN_o}{P(s)}} = \frac{\dfrac{17.3\,(s + 0.029)\,(s + 1.971)}{P(s)}}{\dfrac{s\,(s + 0.0295)\,(s + 3.06)\,(s + 5.57)}{P(s)}}$$

For lattice 2,

$$\frac{Y_{b2}}{Y_{A2}} = \frac{\dfrac{D - 2\,KN_e}{P(s)}}{\dfrac{D + 2\,KN_e}{P(s)}} = \frac{\dfrac{s\,(s + 2.29)\,(s + 15.0)}{P(s)}}{\dfrac{(s + 0.060)\,(s + 2.24)\,(s + 14.96)}{P(s)}}$$

Now P(s) can be chosen so that its zeros alternate with the zeros of all four polynomials. For example,

$$P(s) = (s + 0.5)\,(s + 4)$$

Then, for lattice 1,

$$Y_{b1} = \frac{17.3\,(s + 0.029)\,(s + 1.971)}{(s + 0.5)\,(s + 4)} = \frac{1}{2} + \frac{6.85s}{s + 0.5} + \frac{9.97s}{s + 4}$$

$$Y_{a1} = \frac{2\,(s + 0.0295)\,(s + 3.06)\,(s + 5.57)}{(s + 0.5)\,(s + 4)} = \frac{1}{2} + 2s + \frac{6.98s}{s + 0.5} + \frac{0.817s}{s + 4}$$

and for lattice 2,

$$Y_{b2} = \frac{s\,(s + 2.29)\,(s + 15.0)}{(s + 0.5)\,(s + 4)} = s + \frac{7.49s}{s + 0.5} + \frac{5.36}{s + 4}$$

$$Y_{A2} = \frac{(s + 0.060)\,(s + 2.24)\,(s + 14.96)}{(s + 0.5)\,(s + 4)} \quad s + 1 + \frac{6.35s}{s + 0.5} + \frac{5.79s}{s + 4}$$

The common factors can be removed as parallel admittances at the input and output.

$$Y_{par} = \frac{1}{2} + s + \frac{13.2s}{s + 0.5} + \frac{6.61s}{s + 4}$$

The remaining elements are

Lattice 1:

$$Y'_{A1} = \frac{9.15s}{s+4} = \frac{1}{\frac{0.44}{s} + \frac{1}{9.5}}$$

$$Y'_{b1} = 2s + \frac{0.14s}{s+0.5} = \frac{1}{\frac{0.44}{s} + \frac{1}{9.5}} + \frac{1}{\frac{0.07}{s}}$$

Lattice 2:

$$Y'_{a2} = \frac{0.14s}{s+0.5} = \frac{1}{0.875 + \frac{1}{\frac{s}{0.440}}}$$

$$Y'_{b2} = 1 + \frac{0.143}{s+4} = \frac{1}{0.875 + \frac{1}{\frac{s}{0.440} + \frac{1}{1.25}}}$$

The unbalanced network is shown in Fig. 22 in which the extraneous shunt elements at the left have been omitted.

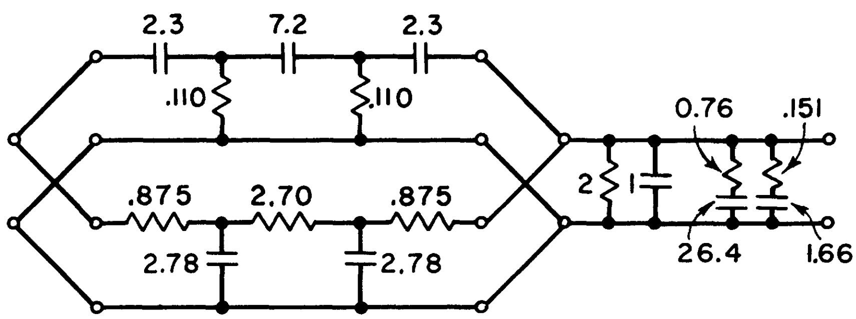

Fig. 22. General Development.

Method 2. Method 1, although it allows any lattice to be unbalanced, involves a considerable reduction in gain. Frequently, it is possible to use an alternate method which uses some of the philosophy of the first method but does not involve such a large decrease in gain. Starting with the original lattice which has a voltage transfer function $A = KN/D$, for which

$$\frac{Y_b}{Y_a} = \frac{D - KN}{D + KN}$$

we write

$$Y_b = \frac{D - KN}{P(s)}$$

$$Y_a = \frac{D + KN}{P(s)}$$

where $P(s)$ is a surplus polynomial whose zeros alternate with the zeros of both $D - KN$ and $D + KN$. Now form

$$Y_{12} = \frac{1}{2}(Y_a - Y_b) = \frac{KN}{P(s)}$$

$$Y_{22} = \frac{1}{2}(Y_a + Y_b) = \frac{D}{P(s)}$$

and split KN into its odd and even parts.

$$Y_{12} = \frac{N_e + N_o}{P(s)} = \frac{N_e}{P(s)} + \frac{N_e}{P(s)} = y_{12,1} + y_{12,2}$$

We can call $y_{12,1}$ and $y_{12,2}$ the transfer admittances of two lattices which, if connected in parallel, will yield the desired transfer admittance. However, we are not sure that the input admittances that correspond to these transfer admittances will add up to the desired input admittances. We shall, therefore, compute the input admittance that corresponds to this splitting and compare this with the original input admittance. To find this input admittance, first, we find the lattice admittances of the parallel lattices by expanding the transfer admittances in partial functions and grouping together the positive and negative terms (noting that at every pole of y_{12} the residues in $y_{12,1}$ and $y_{12,2}$ have opposite signs)

$$y_{12,1} = \frac{N_e}{P(s)} = \frac{N_a}{P_1(s)} - \frac{N_b}{P_2(s)} = Y_{a1} - Y_{b1}$$

$$y_{12,2} = \frac{N_o}{P(s)} = \frac{N_c}{P_2(s)} - \frac{N_d}{P_1(s)} = Y_{a2} - Y_{b2}$$

(Note that the polynomials N_a, N_b, N_c, and N_d are not of importance here.) Now, if these two networks are connected in parallel, the equivalent lattice admittances are

$$Y_a' = Y_{a1} + Y_{a2} = \frac{N_a}{P_1(s)} + \frac{N_c}{P_2(s)} = \frac{Q_1(s)}{P(s)}$$

$$Y_b' = Y_{b1} + Y_{b2} = \frac{N_b}{P_2(s)} + \frac{N_d}{P_1(s)} = \frac{Q_2(s)}{P(s)}$$

(Note that $Q_1(s)$ and $Q_2(s)$ are of no importance here.) Notice that each lattice admittance contains all of the poles of y_{12}, as do the admittances of the original lattice. Then the input admittance of the parallel lattice is

$$Y_{in}' = Y_A' + Y_b' = \frac{Q_1(s)}{P(s)} + \frac{Q_2(s)}{P(s)} = \frac{Q(s)}{P(s)}$$

Let us compare this with the input admittance of the original lattice:

$$Y_{in} = \frac{D(s)}{P(s)}$$

If we think of both input admittances as being expanded in partial fractions, the following cases can be distinguished.

1. If, at a particular pole, the residue in Y_{in}' is just equal to the residue in Y_{in}, then no further work is required.

2. If, at a particular pole, the residue in Y_{in}' is less than the residue in Y_{in}, then an admittance corresponding to the difference in the residues can be added in parallel with the input and output of the parallel lattice, thus making the parallel lattices equivalent to the original lattice.

3. If, at a particular pole, the residue in Y_{in}' is greater than the residue in Y_{in}, the method breaks down.

The question naturally arises, as to when the method will work, that is, when will condition 3 not occur at any of the poles of Y_{in}? There is no readily apparent way to determine this before actually carrying through the synthesis. Clearly, for a small enough value of K, the method will always work, because it can be reduced to the previously described method, but, as will be shown in the example, frequently a much higher value of gain can be realized with this method. In addition, the best choice for the surplus polynomial $P(s)$ is not at all apparent, although it is felt that a good first choice would be the polynomial that has been found suitable for the method of the last section.

EXAMPLE. Consider, again, the voltage transfer function

$$A = \frac{s^3 + 1}{s^3 + \frac{52}{3} s^2 + \frac{104}{3} s + 1}$$

for which K = 1 is the highest obtainable gain. Let us try to obtain this value of gain by using the method described in this section. For this value of gain, a lattice can be constructed for which

$$\frac{Y_b}{Y_a} = \frac{D - KN}{D + KN} = \frac{\frac{52}{3} s(s + 2)}{s(s + 3)(s + 5.61)(s + 0.060)}$$

The surplus polynomial P(s) is arbitrarily chosen to be the same polynomial that was used in the previous example,

$$P(s) = (s + 0.5)(s + 4)$$

Then, for this lattice,

$$Y_a = \frac{2(s + 0.06)(s + 3)(s + 5.61)}{(s + 0.5)(s + 4)} = 2s + 1 + \frac{\frac{45}{7} s}{s + 0.5} + \frac{\frac{19}{21} s}{s + 4}$$

$$Y_b = \frac{52}{3} \frac{s(s + 2)}{(s + 0.5)(s + 4)} = \frac{\frac{52}{7} s}{s + 0.5} + \frac{\frac{208}{21}}{s + 4}$$

$$Y_{11} = \frac{Y_a + Y_b}{2} = s + \frac{1}{2} + \frac{\frac{97}{14} s}{s + 0.5} + \frac{\frac{227}{42} s}{s + 4}$$

$$Y_{12} = \frac{Y_a - Y_b}{2} = \frac{s^3 + 1}{(s + 0.5)(s + 4)} = s + \frac{1}{2} - \frac{\frac{1}{2} s}{s + 0.5} - \frac{\frac{9}{2} s}{s + 4}$$

Then, the transfer admittance is split in the appropriate fashion:

$$Y_{12,1} = \frac{s^3}{(s + 0.5)(s + 4)} = s + \frac{\frac{1}{14} s}{s + 0.5} - \frac{\frac{64}{14} s}{s + 4}$$

$$Y_{12,2} = \frac{1}{(s + 0.5)(s + 4)} = \frac{1}{2} + \frac{\frac{1}{14} s}{s + 4} - \frac{\frac{8}{14} s}{s + 0.5}$$

Then, the lattice admittance for the individual lattices is

$$\frac{Y_{a1}}{2} = s + \frac{\frac{1}{14}s}{s + 0.5} \qquad \frac{Y_{b1}}{2} = \frac{\frac{64}{14}s}{s + 4}$$

$$\frac{Y_{a2}}{2} = \frac{1}{2} + \frac{\frac{1}{14}s}{s + 4} \qquad \frac{Y_{b2}}{2} = \frac{\frac{8}{14}s}{s + 0.5}$$

The input admittance for these two lattices in parallel is

$$Y'_{11} = \frac{Y_{a1} + Y_{a2} + Y_{b1} + Y_{b2}}{2} = s + \frac{1}{2} + \frac{\frac{9}{14}s}{s + 0.5} + \frac{\frac{65}{14}s}{s + 4}$$

The residue in each of these poles is equal to or less than the residue in the corresponding pole of the desired transfer admittances. Therefore, the two lattices in parallel can be made equivalent to the desired lattice with the addition, at each end of the lattice, of a parallel admittance equal to the difference $Y_{11} - Y'_{11}$. The appropriate admittance is

$$Y_{11} - Y'_{11} = Y_{par} = \frac{\frac{86}{7}s}{s + 0.5} + \frac{\frac{32}{7}s}{s + 4}$$

The fact that such a realizable admittance exists assures the success of this method.

The individual lattices can then be unbalanced by an appropriate Cauer development:

$$Z_{a1} = \frac{7}{16s} + \cfrac{1}{\cfrac{64}{7} + \cfrac{1}{\cfrac{1}{16}s}}$$

$$Z_{b1} = \frac{7}{64} + \frac{7}{16s}$$

$$Z_{a2} = \frac{7}{8} + \cfrac{1}{\cfrac{16s}{7} + \cfrac{1}{8}}$$

$$Z_{b2} = \frac{7}{8} + \frac{7}{16s}$$

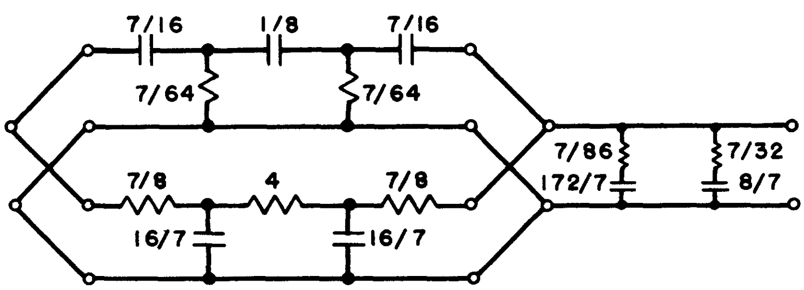

Fig. 23. Another development.

The completely unbalanced ladder appears in Fig. 23, in which the extraneous elements in parallel with the voltage source have been omitted.

VI. GENERAL METHODS

1. INTRODUCTION

When the required gain is greater than that obtainable from a symmetrical network, but less than the maximum obtainable from the given pole-zero plot, realization techniques which are more general than those previously considered must be used. Because the ones and minus ones can no longer be produced by circuit elements that have zeros or poles, but must be produced by cancellation, the concept of the one loses a good deal of its significance if it is applied to the general problem. On account of this, the reasoning for the realization techniques tends to be more numerical than physical and the resultant network configurations tend to be more complicated than those previously considered.

In the general case, the only restriction on the gain is that D – KN and D + KN have positive coefficients. For a given value of K, less than the maximum, some augmentation may be required to meet the positive-coefficient condition, but we assume that any necessary augmentation has been done and the voltage transfer function is in the form:

$$A = \frac{a_n s^n + a_{n-1} s^{n-1} + - - a_1 s + a_o}{b_n s^n + b_{n-1} s^{n-1} + - - b_1 s + b_o}$$

where the magnitude of each numerator coefficient is equal to or less than the corresponding denominator coefficient, $a_j \leqslant b_j$. If the realization is to yield a grounded network, then the numerator coefficients must be all positive (which may require further augmentation); if an ungrounded network is satisfactory, then the numerator coefficients can be positive or negative, provided only that the magnitude of each numerator coefficient is less than or equal to the magnitude of the corresponding denominator coefficient.

The general realization problem was first solved by Fialkow and Gerst (14, 15) in papers that appeared in 1952 and 1954. Similar realization methods will be described here, which, although they use a good deal of the philosophy of the procedures of Fialkow and Gerst, introduce some simplifications and generalizations. Thus the present methods frequently require much fewer elements than do the original methods of Fialkow and Gerst. At every point in the discussion, it will be made clear exactly where the similarities and differences between the two methods lie.

In building up the solution to a general problem like this, it is frequently helpful to consider certain particular cases in which the solution can be easily accomplished, and then to show that the general case can be treated as a combination of these particular cases. This method was applied to the problem of lossless driving-point impedance realization with great success by Foster, who showed that the general LC driving-point impedance can be constructed as the series or parallel combination of certain simple canonic forms, the realization of which is trivial. A similar process will be used here, consisting of the following steps.

1. Simple canonic form networks will be discovered for the realization of particular kinds of

two-element-kind grounded voltage transfer function.

2. It will be shown that the general two-element-kind grounded voltage transfer function can be realized as a combination of these canonic forms.

3. It will be shown that the general three-element-kind grounded voltage transfer function can be realized as a combination of general two-element-kind grounded networks.

4. It will be shown that any ungrounded voltage transfer function can be realized as a combination of grounded voltage transfer functions.

2. CANONIC FORMS FOR TWO-ELEMENT-KIND GROUNDED NETWORKS

The details of this discussion will be given for RC functions, but the results apply to the RL and LC cases as well.

By a canonic form, we mean a simple voltage transfer function which can be realized easily with any obtainable gain, and which can be used to construct more general voltage transfer functions. Almost the simplest nontrivial RC voltage transfer function that can be studied is of the form

$$A = \frac{a_1 s + a_o}{b_1 s + b_o}$$

where

$$0 \leqslant a_1 \leqslant b_1$$

$$0 \leqslant a_o \leqslant b_o$$

This is the most general voltage transfer function of degree 1 (that is, the denominator is of the first degree). Two simple RC realizations in terms of all networks are possible for this function, corresponding to writing two functions in the following form

$$1.\quad A = \frac{a_1 s + a_o}{b_1 s + b_o} = \frac{y_2}{y_1 + y_2} = \frac{C_2 s + G_2}{(C_1 + C_2)s + G_1 + G_2} \qquad \text{(Fig. 24b)}$$

$$2.\quad A = \frac{a_1 + \frac{a_o}{3}}{b_1 + \frac{b_o}{s}} = \frac{z_1}{z_1 + z_2} = \frac{R_1 + \frac{1}{C_1 s}}{R_1 + R_2 + \frac{1}{C_1 s} + \frac{1}{C_2 s}} \qquad \text{(Fig. 24a)}$$

These two ELL-networks are shown in Fig. 24, in which all the element values are guaranteed to be positive by the positive-coefficient condition. These are the canonic forms used by Fialkow and Gerst in their realization procedure.

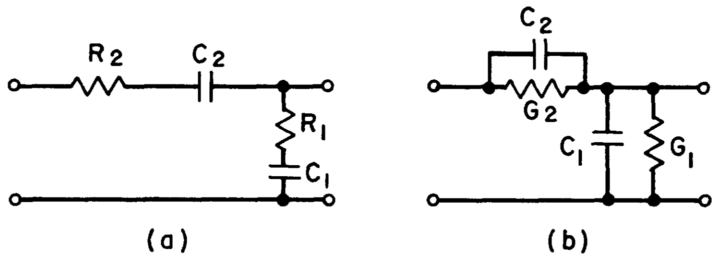

Fig. 24. ELL canonic forms.

At least two other canonic forms exist; these comprise functions of the forms

1. $$A = \frac{a_n s^n}{b_n s^n + b_{n-1} s^{n-1} + -- b_1 s + b_o} \qquad 0 \leqslant a_n \leqslant b_o$$

2. $$A = \frac{a_o}{b_n s^n + b_{n-1} s^{n-1} + -- b_1 s + b_o} \qquad 0 \leqslant a_o \leqslant b_o$$

Both of these functions can be realized as ladder networks by using essentially nothing more than simple Cauer continued-fraction expansions. In either case, write

$$A = \frac{KN}{D} = \frac{y_{12}}{y_{22}}$$

associate a surplus polynomial P(s) with the given function

$$A = \frac{\frac{KN}{P(s)}}{\frac{D}{P(s)}} = \frac{y_{12}}{y_{22}}$$

where P(s) is any polynomial that has all its zeros on the negative real axis alternating with the zeros of D in such a way that D/P(s) is an RC driving-point admittance. Then expand this admittance in a Cauer ladder:

1. in the form shown in Fig. 25a, if all the zeros of A are at zero frequency, as in canonic form 1.

2. in the form shown in Fig. 25b, if all the zeros of A are at infinite frequency, as in canonic form 2.

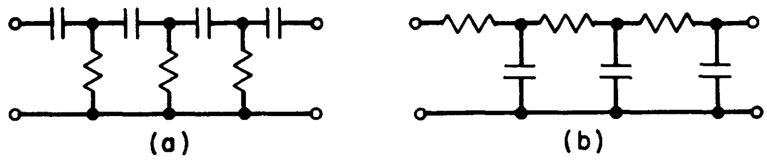

Fig. 25. Cauer canonic forms.

The resultant ladders will have a gain of exactly one at infinite frequency and zero frequency, respectively (corresponding to $a_n = b_n$ or $a_o = b_o$). This is the maximum gain obtainable from such voltage transfer functions. If less than the maximum gain is required, that is, if $a_n < b_n$ or $a_o < b_o$, then these networks can be modified to yield the reduced gain by an application of Thévenin's theorem, as shown in Fig. 26. If a voltage transfer function with a certain gain has been realized with an input series element as in Fig. 26a, any gain smaller than this original value can be achieved with the circuit of Fig. 26b, in which

$$\frac{E_2'}{E_1} = k \frac{E_2}{E_1} \qquad k \leqslant 1$$

as can be readily verified from Thévenin's theorem. With this method, voltage transfer functions that have the canonic forms 1 or 2, can be realized with any obtainable gain.

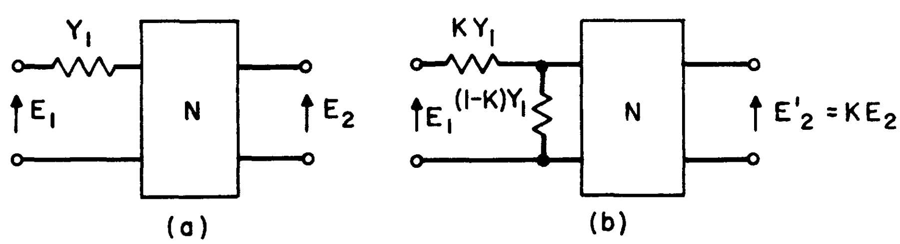

Fig. 26. An application of Thévenin's theorem.

EXAMPLE. The voltage transfer function

$$A = \frac{K}{(s+1)(s+3)(s+5)}$$

can be realized with a gain as high as K = 1. In order to demonstrate the method just described, the function will be synthesized with K = 2/3.

$$A = \frac{\frac{2}{3}}{(s+1)(s+3)(s+5)}$$

An appropriate surplus polynomial is associated with this function.

$$A = \frac{y_{12}}{y_{22}} = \frac{\dfrac{\frac{2}{3}}{(s+2)(s+4)}}{\dfrac{(s+1)(s+3)(s+5)}{(s+2)(s+4)}}$$

Then y_{22} is expanded in the proper Cauer ladder with all its transmission zeros at infinite frequency, as in Fig. 27a. This ladder has the desired voltage transfer function, but with K = 1, since the voltage transmission is obviously equal to one at zero frequency. In order to obtain a K of 2/3, Thévenin's theorem is used on the conductance in series with the voltage source to obtain Fig. 27b, which is the desired network.

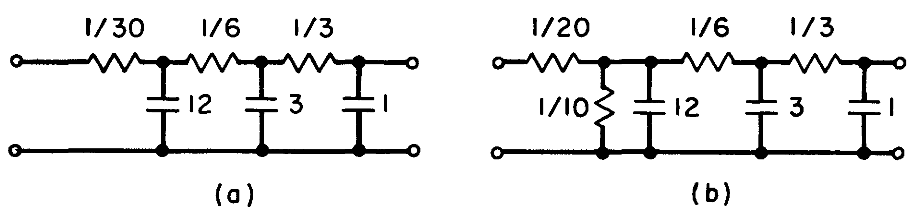

Fig. 27. A ladder network.

3. GENERAL SYNTHESIS OF TWO-ELEMENT-KIND GROUNDED VOLTAGE TRANSFER FUNCTIONS—(WITH RC AS THE EXAMPLE)

The general philosophy of this synthesis method can be summarized as follows.

1. If the given function is exactly one of the canonic forms, then realize it in the appropriate manner.

2. If the given function is not one of the canonic forms, then break the function into smaller functions that are canonic forms, realize these canonic forms and then interconnect them in the appropriate manner.

The only new technique that must be studied is the method of splitting a voltage transfer function into simpler voltage transfer functions in a manner that corresponds to splitting the network which realizes that voltage transfer function into simpler networks. This splitting is done by using the formula

$$A = \frac{y_{12}}{y_{22}} = \frac{y_{12(1)} + y_{12(2)}}{y_{22(1)} + y_{22(2)}}$$

for the voltage transfer function of two networks in parallel. Before considering the general method, it is helpful to consider a simple example that will clarify the problems involved. The simplest voltage transfer function that can be studied is that of second degree (the first-degree function is a canonic form).

EXAMPLE.

$$A = \frac{s^2 + 1}{(s + 1)(s + 3)}$$

At first thought, one might consider dividing this function into

$$A_1 = \frac{s^2}{(s + 1)(s + 3)}, \qquad A_2 = \frac{1}{(s + 1)(s + 3)}$$

associating an appropriate polynomial with each function so as to put it in the form $A = y_{12}/y_{22}$

$$A_1 = \frac{\dfrac{s^2}{s + 2}}{\dfrac{(s + 1)(s + 3)}{s + 2}}, \qquad A_2 = \frac{\dfrac{1}{s + 1}}{\dfrac{(s + 1)(s + 3)}{s + 2}}$$

realizing each function by the methods considered previously for canonic forms, and then connecting the two networks in parallel. However, if this is done, the resultant voltage transfer function is

$$A = \frac{y_{12(1)} + y_{12(2)}}{y_{22(1)} + y_{22(2)}} = \frac{1}{2} \frac{(s^2 + 1)}{(s + 1)(s + 3)}$$

and the maximum gain has not been realized.

One way to split up this voltage transfer function so as to obtain the maximum gain is as follows. (A general method for doing this will be given later.) Starting with

$$A = \frac{s^2 + 1}{(s + 1)(s + 3)} = \frac{s^2 + 1}{s^2 + 4s + 3}$$

associate a surplus polynomial with this function in order to put it in the form y_{12}/y_{22}:

$$A = \frac{\frac{s^2 + 1}{s + 2}}{\frac{(s + 1)(s + 3)}{s + 2}} = \frac{\frac{s^2 + 1}{s + 2}}{\frac{s^2 + 4s + 3}{s + 2}}$$

then split the numerator and denominator as follows:

$$A = \frac{\frac{s^2}{s + 2} + \frac{1}{s + 2}}{\frac{s^2 + 2.1s}{s + 2} + \frac{1.9s + 3}{s + 2}} = \frac{y_{12(1)} + y_{12(2)}}{y_{22(1)} + y_{22(2)}}$$

It is of importance in this particular splitting that the entire first coefficient of the denominator (and numerator) belongs to the first network, while the entire last coefficient belongs to the second network. It is this property that ensures the success of this procedure. Each of these two networks can be realized by either of the canonic forms. The realization will now be carried out for both canonic forms.

a. Cauer Forms

For network 1,

$$y_{22(1)} = \frac{s^2 + 2.1s}{s + 2} \qquad y_{12(1)} = \frac{s^2}{s + 2}$$

The appropriate Cauer continued-fraction expansion is

$$y_{22(1)} = \cfrac{1}{\cfrac{1}{2.1s} + \cfrac{1}{64.1 + \cfrac{1}{\cfrac{1}{21s}}}}$$

The corresponding network is the top ladder in Fig. 28a.

For network 2,

$$y_{22(2)} = \frac{1.9s + 3}{s + 2} \qquad y_{12(2)} = \frac{1}{s + 2}$$

This time the continued-fraction expansion must be followed by an application of Thévenin's theorem. The continued fraction expansion is

$$y_{22(2)} = \cfrac{1}{\cfrac{1}{1.9} + \cfrac{1}{\cfrac{36.1}{8}s + \cfrac{1}{\cfrac{8}{5.7}}}}$$

After the application of Thévenin's theorem to reduce the gain by 1/3, the network appears as the bottom ladder in Fig. 28a. The over-all network is the parallel connection of these two ladders, as in Fig. 28a.

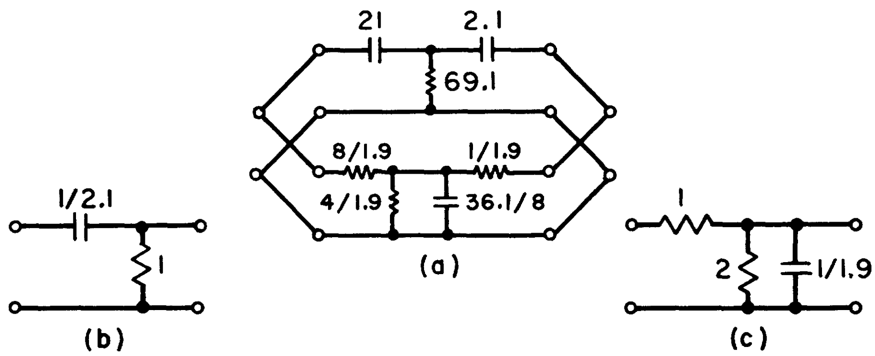

Fig. 28. An example.

b. Fialkow and Gerst Canonic Forms

The Fialkow and Gerst canonic forms require two steps for their realization.

For network 1, since

$$y_{22(1)} = \frac{s^2 + 2.1s}{s + 2}$$

and

$$y_{12(1)} = \frac{s^2}{s + 2}$$

the voltage transfer function of this network is

$$A_1 = \frac{s}{s + 2.1} = \frac{1}{1 + \frac{2.1}{s}} = \frac{z_b}{z_a + z_b}$$

This voltage transfer function can be realized as an ELL-network, as in Fig. 28b. The ELL-network has the correct voltage transfer function, but the incorrect y_{22}. However, if an impedance is added in series with the output terminal, and the impedances in the ELL-network are scaled in magnitude, the output admittance y_{22} can be adjusted without affecting the voltage transfer function. When y_{22} has been adjusted to the correct value, then y_{12} must automatically be correct, since the voltage transfer function remains unchanged. Fialkow and Gerst showed that such a series impedance can always be found (in fact, an infinite number of appropriate impedances exists). In this case no work is involved, because if we multiply R and 1/C by 1/64 and add a 2.1-farad capacitor in series with the output, we obtain the same circuit as the top ladder in Fig. 27a, which is evidentally a correct network.

For network 2, a similar procedure can be used. Since

$$y_{22(2)} = \frac{1.9s + 3}{s + 2} \qquad\qquad y_{12} = \frac{1}{s + 2}$$

the voltage transfer function

$$A = \frac{1}{1.9s + 1}$$

can be realized as a simple canonic ELL-network; this time an admittance ELL is used and the resultant network is shown in Fig. 28c. Again, this network has the incorrect output admittance y_{22}, but if R and 1/C are multiplied by 8/19, and a 1/1.9-ohm resistor is used as a series impedance, the same network as the bottom ladder in Fig. 28a results. These two networks are then connected in parallel, yielding the circuit of Fig. 28a.

In this simple problem, the two canonic forms can be made to yield the same over-all network. In the general case this is not true, and it will be shown that the Cauer forms frequently require fewer circuit elements.

c. The General Synthesis

Given an RC voltage transfer function,

$$A = \frac{KN}{D} = \frac{a_n s^n + a_{n-1} s^{n-1} + \cdots a_1 s + a_o}{b_n s^n + b_{n-1} s^{n-1} + \cdots b_1 s + b_o}$$

$$= \frac{a_n s^n + a_{n-1} s^{n-1} + \cdots a_1 s + a_o}{(s + \sigma_1)\,(s + \sigma_2) \cdots (s + \sigma_n)}$$

where $0 \leqslant a_j \leqslant b_j$, the first step is to associate a surplus polynomial, P(s), with the function, in order to put it in the form

$$A = \frac{\frac{KN}{P(s)}}{\frac{D}{P(s)}} = \frac{y_{12}}{y_{22}}$$

where P(s) is any polynomial, of one degree less than D, that has all simple zeros on the negative real axis alternating with the zeros of D in such a way as to make D/P(s) an RC admittance. Then, if the resultant function is one of the canonic forms, this form is realized and the synthesis is over. If the resultant function is not a canonic form, the second step is to split it up into the parallel combination of two networks, each of which is simpler; this splitting is continued until one of the canonic forms is encountered.

To split y_{22} into simpler driving-point admittances, first expand

$$\frac{y_{22}}{s} = \frac{D}{sP(s)}$$

in partial fractions, then multiply by s in order to obtain the usual Foster-like expansion of RC driving-point admittances:

$$y_{22} = As + B + \sum_{v} \frac{C_v s}{s + \sigma_v}$$

where A and B are not zero because of the form of P(s). Now form the two simpler driving-point admittances as follows.

For $y_{22(1)}$, use the following terms: (a) all of the pole at ∞ (As); (b) none of the constant value (B); and (c) some arbitrary, nonzero part of each of the finite poles, but not the entire pole.

For $y_{22(2)}$, use the rest of y_{22}, consisting of: (a) none of the pole at ∞ (As); (b) all of the constant value (B); and (c) the rest of each finite pole.

The splitting given here accomplishes the same purpose as that given by Fialkow and Gerst, but our procedure is somewhat simpler. When this procedure has been carried out, we obtain:

$$y_{22} = y_{22(1)} + y_{22(2)}$$

$$\frac{b_n s^n + b_{n-1} s^{n-1} + \text{--} b_1 s + b_o}{P(s)} = \frac{b_n s^n + c_{n-1} s^{n-1} + \text{--} c_2 s^2 + c_1 s}{P(s)}$$

$$+ \frac{d_{n-1} s^{n-1} + \text{--}\ d_1 s + b_o}{P(s)}$$

Of importance in this splitting are the following:

1. Both $y_{22(2)}$ and $y_{22(2)}$ have the same denominator.
2. The numerator of y_{22} has been split between $y_{22(1)}$ and $y_{22(2)}$, so that $b_j = c_j + d_j$ for $j \neq n$ or 0.
3. The entire coefficient of the highest power of the numerator of y_{22}, b_n, is in $y_{22(1)}$ and the entire coefficient of the lowest power b_o is in $y_{22(2)}$, so that the numerators of $y_{22(1)}$ and $y_{22(2)}$ are each one degree simpler than that of y_{22}.

The next step is to split the transfer admittance

$$y_{12} = \frac{N}{P(s)} = \frac{a_n s^n + a_{n-1} s^{n-1} + -- a_1 s + a_o}{P(s)}$$

into $y_{12(1)}$ and $y_{12(2)}$, so that, if each transfer admittance is associated with its driving-point admittance, the ratios

$$\frac{y_{12(1)}}{y_{22(1)}}, \qquad \frac{y_{12(2)}}{y_{22(2)}}$$

are both realizable voltage transfer functions. This is readily done by splitting the numerator of y_{12} as follows: Since in the original voltage transfer function each term in the numerator (a_j) is equal to or less than the corresponding term in the denominator (b_j), that is, $a_j \leqslant b_j$, and, since the denominator has been split so that $b_j = c_j + d_j$, it is always possible to split up $a_j = e_j + f_j$ so that

$$e_j \leqslant c_j$$

$$f_j \leqslant d_j$$

This splitting of the numerator can be done in a great number of ways. The resultant voltage transfer function is of the form:

$$A = \frac{\dfrac{a_n s^n + e_{n-1} s^{n-1} + e_{n-2} s^{n-2} + -- - e_2 s^2 + e_1 s}{P(s)} + \dfrac{f_{n-1} s^{n-1} + -- f_1 s + a_o}{P(s)}}{\dfrac{b_n s^n + C_{n-1} s^{n-1} + C_{n-2} s^{n-2} + -- C_2 s^2 + C_1 s}{P(s)} + \dfrac{d_{n-1} s^{n-1} + -- - d_1 s + b_o}{P(s)}}$$

$$= \frac{y_{12(1)} + y_{12(2)}}{y_{22(2)} + y_{22(2)}}$$

The original network has been split into two parallel networks, each of which has a realizable voltage transfer function. If that function is one of the canonic forms, then that form is realized; if not, the process must be repeated. But before that can be done, each of the network functions must be put in the same form as the original function. This last statement will be clarified as we proceed. The method of reasoning is different for the two networks.

Network 2. For this network,

$$y_{12(2)} = \frac{f_{n-1}s^{n-1} + - - a_o}{P(s)} = \frac{N_2}{P(s)}$$

$$y_{22(2)} = \frac{d_{n-1}s^{n-1} + - - b_o}{P(s)} = \frac{D_2}{P(s)}$$

$$A_2 = \frac{f_{n-1}s^{n-1} + - - a_o}{d_{n-1}s^{n-1} + - - b_o} = \frac{N_2}{D_2}$$

The voltage transfer function is one degree less than the original voltage transfer function. However, the denominator associated with $y_{22(2)}$, $P(s)$, is of the same degree as its numerator (D_2), while in the original function, the denominator of y_{22}, $P(s)$, was of one degree less than its numerator (D). Therefore, in order to put this function in the same form as the original function, we must associate with it a different polynomial $P_2(s)$. This is done as follows. Consider

$$Z_2 = \frac{1}{y_{22}} = \frac{P(s)}{D_2}$$

If a resistor equal to the value of this impedance at infinite frequency is removed from Z_2,

$$Z_2 = R + \frac{P_2(s)}{D} = R + Z_2'$$

and the remaining impedance will have a numerator $P_2(s)$, of one degree less than $P(s)$. This process is shown pictorially in Fig. 29. The reduced network N_2' has the same voltage transfer function as N_2, but with a different y_{12} and y_{22}. For N_2',

$$A_1' = \frac{\dfrac{N_2}{P_2(s)}}{\dfrac{D_2}{P_2(s)}}$$

where $P_2(s)$ is of one degree less than D. Thus, network 2′ is of the same form as the original function (but one degree simpler) and the process can be repeated. The resistor R_∞ in series with network 2′ changes its output admittance to the correct value for its insertion in the over-all network.

Network 1. For this network,

$$y_{12(1)} = \frac{a_n s^n + e_{n-1} s^{n-1} + -- \; e_2 s^2 + e_1 s}{P(s)} = \frac{sN_1}{P(s)}$$

$$y_{22(2)} = b_n s^n + C_{n-1} s^{n-1} + -- \; C_2 s^2 + C_1 s = s\frac{D_1}{P(s)}$$

$$A_1 = \frac{a_n s^{n-1} + e_{n-1} s^{n-2} + -- \; C_2 s + e_1}{b_n s^{n-1} + e_{n-1} s^{n-2} + -- \; C_2 s + C_1} = \frac{N_1}{D_1}$$

Notice that, in the voltage transfer function, the s-factors cancel, thus making it one degree smaller than the original voltage transfer function. Again, we have the problem of associating with this network a new polynomial $P_1(s)$, of one degree less than $P(s)$; we must also remove the s-factor from the numerator of $y_{22(2)}$, since that factor cancels out in A. Again, consider

$$Z_1 = \frac{1}{y_{22(1)}} = \frac{P(s)}{sD}$$

and remove the pole at zero frequency

$$Z_1 = \frac{P_1(s)}{D} + \frac{\frac{1}{C_1}}{s} = Z_1' + \frac{\frac{1}{C_1}}{s}$$

Now $P_1(s)$ is of the proper degree, and, in addition, the s-factor has been removed. This process is shown in Fig. 30. Again, the reduced network N_1' has the same voltage transfer function as N_1, and the capacitor makes the output admittances of the two networks equal, as well as supplying the extra factor that cancels. Physically, this extra factor is another transmission zero at zero frequency, which is supplied by the capacitor. For network N_1'

$$A_1' = \frac{\frac{N_1}{P_1(s)}}{\frac{D_1}{P_1(s)}}$$

This is a function of the form originally considered and the process can be repeated.

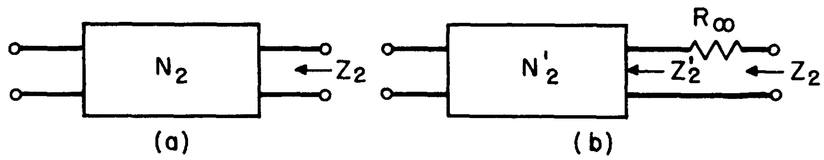

Fig. 29. A reduction step.

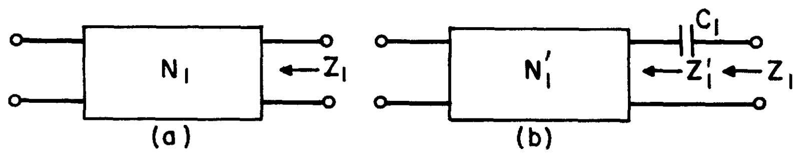

Fig. 30. Another reduction.

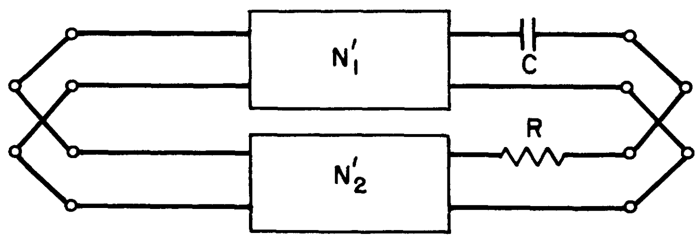

Fig. 31. After one step.

After this one step the over-all network is in the form of Fig. 31. The process is then repeated on N_1' and N_2' until all the networks have been reduced to canonic form; the result is a tree-like arrangement.

The method used by Fialkow and Gerst for removing the series elements is somewhat different from that given here. They pick the polynomials $P_1(s)$ and $P_2(s)$ arbitrarily, subject only to certain realizability conditions, and then find the series impedance that is necessary to obtain the correct output admittance (after suitable impedance scaling). The series impedance is, in general, not a single resistor or capacitor.

In comparing our method with the original method of Fialkow and Gerst, the following statements can be made:

1. Our method uses much of the general philosophy of their method.

2. Certain steps, which are part of both procedures are given here in slightly simplified form. In particular, the removal of only one resistor or one capacitor is needed in the final step in each cycle.

3. The introduction of a new canonic form allows a considerable reduction in the number of elements in certain particular cases.

EXAMPLE.

$$A = \frac{s^4 + s + 1}{(s + 1)(s + 3)(s + 5)(s + 7)}$$

$$= \frac{s^4 + s + 1}{s^4 + 16s^3 + 86s^2 + 176s + 105}$$

First, a polynomial, P(s), must be associated with this function in order to put it in the form y_{12}/y_{22}. Let P(s) = (s + 2) (s + 4) (s + 6). Then

$$A = \frac{y_{12}}{y_{22}} = \frac{\dfrac{s^4 + s + 1}{(s+2)(s+4)(s+6)}}{\dfrac{(s+1)(s+3)(s+5)(s+7)}{(s+2)(s+4)(s+6)}}$$

This is not one of the canonic forms, so it must be split into simpler networks. Expanding y_{22} in partial fractions, we have

$$y_{22} = s + \frac{35}{16} + \frac{\frac{15}{6}s}{s+2} + \frac{\frac{9}{16}s}{s+4} + \frac{\frac{5}{16}s}{s+6}$$

For one of the simpler admittances $y_{12(1)}$, we take the whole pole at ∞, and, arbitrarily, a residue of ¼ at each of the poles

$$y_{22(1)} = s + \frac{\frac{1}{4}s}{s+2} + \frac{\frac{1}{4}s}{s+4} + \frac{\frac{1}{4}s}{s+6}$$

$$= \frac{s^4 + 12.75s^3 + 50s^2 + 59s}{(s+2)(s+4)(s+6)}$$

For $y_{22(2)}$, we take the rest of y_{22}

$$y_{22(2)} = \frac{35}{16} + \frac{\frac{11}{16}s}{s+2} + \frac{\frac{5}{16}s}{s+4} + \frac{\frac{1}{16}s}{s+6} = \frac{3.25s^3 + 36s^2 + 117s + 105}{(s+2)(s+4)(s+6)}$$

Now y_{22} has been split into two parallel driving-point admittances; the next step is to split y_{12} into two transfer admittances, and associate each with one of the driving-point admittances. This can be done in a number of ways. One convenient way is

$$A = \frac{y_{12(1)} + y_{12(2)}}{y_{22(1)} + y_{22(2)}} = \frac{\dfrac{s^4}{(s+2)(s+4)(s+6)} + \dfrac{s+1}{(s+2)(s+4)(s+6)}}{\dfrac{s^4 + 12.75s^3 + 50s^2 + 59s}{(s+2)(s+4)(s+6)} + \dfrac{3.25s^3 + 36s^2 + 117s + 105}{(s+2)(s+4)(s+6)}}$$

This particular way was chosen because network 1 is a Cauer canonic form and can be immediately realized. In a Fialkow and Gerst realization, this function would have to be split and the process repeated.

Network 1. Since $y_{12(1)}$ has all its zeros at zero frequency, an appropriate Cauer expansion of $y_{22(1)}$ will yield the required network.

$$y_{22(1)} = \cfrac{1}{\cfrac{0.815}{s} + \cfrac{1}{17.86 + \cfrac{1}{\cfrac{0.157}{s} + \cfrac{1}{150 + \cfrac{1}{\cfrac{0.0268}{s} + \cfrac{1}{4350 + \cfrac{1}{\cfrac{0.0012}{s}}}}}}}}$$

The corresponding network is shown in Fig. 32.

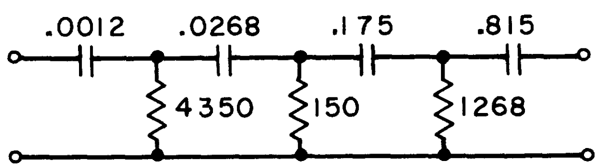

Fig. 32. Canonic form.

Network 2. This network is not a canonic form and must be split again, but before doing this, a series resistor must be removed from $1/y_{22(2)}$

$$Z_2 = \frac{1}{y_{22(2)}} = \frac{s^3 + 12s^2 + 44s + 48}{3.25s^3 + 36s^2 + 117s + 105}$$

$$= 0.308 + \frac{0.92s^2 + 8s + 15.7}{3.25s^3 + 36s^2 + 117s + 105} = 0.308 + \frac{1}{y'_{22(2)}}$$

At this point, network 2 appears as in Fig. 33.

Network N_2'. The rest of the procedure deals with the realization of network N_2', which has the following parameters

$$y'_{22(2)} = \frac{3.25s^3 + 36s^2 + 117s + 105}{0.92s^2 + 8s + 15.7}$$

$$y'_{12(2)} = \frac{s + 1}{0.92s^2 + 8s + 15.7}$$

$$A'_2 = \frac{s + 1}{3.25s^3 + 36s^2 + 117s + 105}$$

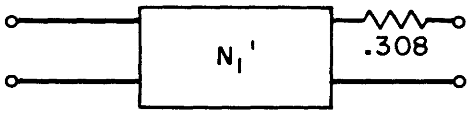

Fig. 33. Partially reduced network.

To split N'_2, into two parallel networks, $y'_{22(2)}$ must be expanded in partial fractions

$$y'_{22(2)} = \frac{3.25s^3 + 36s^2 + 117s + 105}{0.92s^2 + 8s + 15.7} = \frac{3.25s^3 + 36s^2 + 117s + 105}{0.92(s + 5.71)(s + 2.99)}$$

$$= 3.53s + 6.69 + \frac{0.705s}{s + 5.71} + \frac{0.1675s}{s + 2.99}$$

Then write

$$y'_{22(2)A} = y_{22A} = 3.53s + \frac{0.5s}{s + 5.71} + \frac{0.5s}{s + 2.99}$$

$$= \frac{3.25s^3 + 2.86s^2 + 58.0s}{0.92(s + 5.71)(s + 2.99)}$$

$$y'_{22(2)B} = y_{22B} = 6.69 + \frac{0.205s}{s + 5.71} = \frac{1.175s}{s + 2.99}$$

$$= \frac{7.40s + 59s + 105}{0.92(s + 5.71)(s + 2.99)}$$

The voltage transfer function corresponding to N' can be written

$$A'_2 = \frac{\frac{s}{0.92(s + 5.71)(s + 2.99)} + \frac{1}{0.92(s + 5.71)(s + 2.99)}}{\frac{3.25s^3 + 28.6s^2 + 58s}{0.92(s + 5.71)(s + 2.99)} + \frac{7.40s^2 + 59.15s + 105}{0.92(s + 5.71)(s + 2.99)}}$$

$$= \frac{y_{12A} + y_{12B}}{y_{22A} + y_{22B}}$$

Network N′ has now been split into an A and a B network, each of which must be realized separately.

For network A,

$$y_{12A} = \frac{s}{0.92(s + 5.71)(s + 2.99)}$$

$$y_{22A} = \frac{3.25s^3 + 28.6s^2 + 58.0s}{0.92(s + 5.71)(s + 2.99)}$$

$$A_A = \frac{1}{3.25s^2 + 28.6s + 58.0}$$

This voltage transfer function is a Cauer canonic form and can be realized by a continued-fraction expansion of y_{22A} (in which a series capacitor at the output end supplies the cancelled s-factor), followed by an application of Thévenin's theorem at the input terminals. The continued fraction expansion is

$$y_{22A} = \cfrac{1}{\cfrac{0.27}{s} + \cfrac{1}{65s + \cfrac{1}{0.00428 + \cfrac{1}{9.75s + \cfrac{1}{0.0206}}}}}$$

Since the required gain is 1/50 of the maximum obtainable with this canonic form, Thévenin's theorem must be used on the input series resistance, thus yielding the network of Fig. 34.

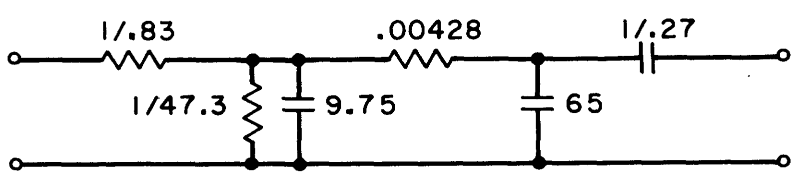

Fig. 34. Network A.

For network B,

$$y_{22B} = \frac{7.40s^2 + 59.1s + 0.05}{0.92(s + 5.71)(s + 2.99)}$$

$$y_{12B} = \frac{1}{0.92(s + 5.71)(s + 2.99)}$$

$$A_B = \frac{1}{7.40s^2 + 59.1s + 105}$$

This voltage transfer function is also a canonic form and can be realized by a Cauer expansion of y_{22B} followed by an application of Thévenin's theorem. The continued-fraction expansion is

$$y_{22B} = \cfrac{1}{0.125 + \cfrac{1}{11.4s + \cfrac{1}{0.022 + \cfrac{1}{102s + \cfrac{1}{0.00276}}}}}$$

Since the required gain is 1/105 of the maximum obtainable with this configuration, an application of Thévenin's theorem is necessary at the input, thus yielding the network of Fig. 35.

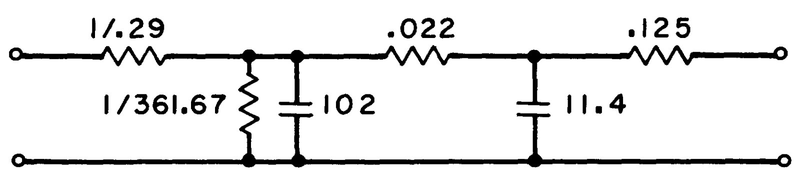

Fig. 35. Network B.

The entire voltage transfer function has now been realized. It consists of the appropriate interconnection of the above network; it is shown in Fig. 36.

4. RLC Grounded Networks

Before considering this general RLC grounded network problem, it is appropriate to note that the second-order RLC voltage transfer function

$$A = \frac{a_2s^2 + a_1s + a_o}{b_2s^2 + b_1s + b_o}$$

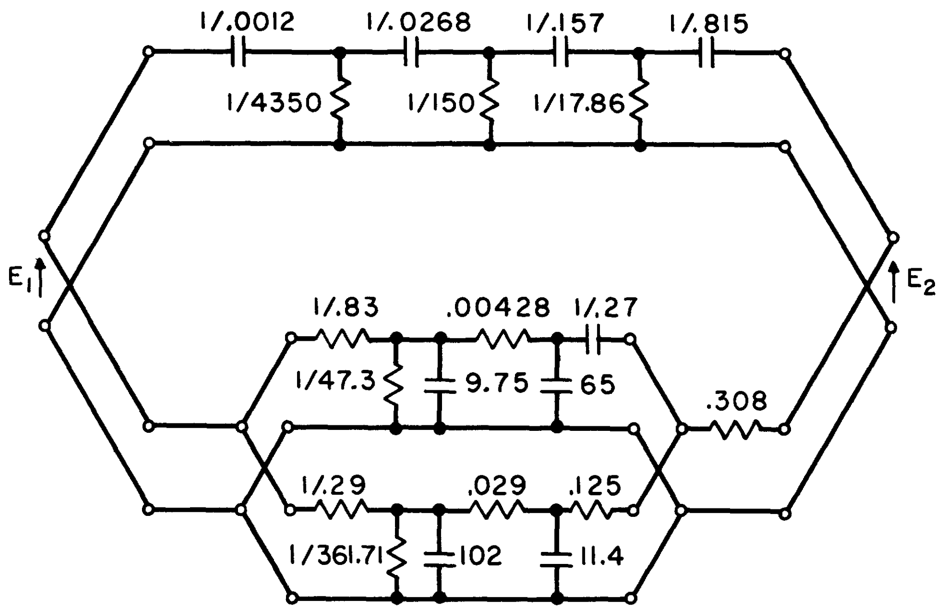

Fig. 36. Complete network.

can always be realized without using these general methods. A can be written either in the form

$$A = \frac{a_2 s + a_1 + \frac{a_o}{s}}{b_2 s + b_1 + \frac{b_o}{s}} = \frac{\dfrac{L_1 s + R + \frac{1}{C_1}}{s}}{(L_1 + L_2)\, s + R_1 + R_2 + \dfrac{\frac{1}{C_1} + \frac{1}{C_2}}{s}} = \frac{Z_1}{Z_1 + Z_2}$$

or in the form

$$A = \frac{C_A s + G_A + \frac{1}{L_A s}}{(C_A + C_B)s + G_A + G_B + \dfrac{\frac{1}{L_C} + \frac{1}{L_B}}{s}} = \frac{y_a}{y_b + y_a}$$

and then can be realized as one of the ELL networks of Fig. 37.

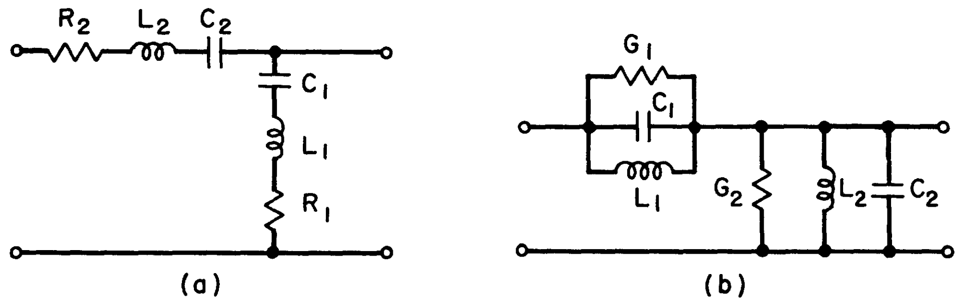

Fig. 37. RLC canonic forms.

One might be tempted to extend the reasoning used on two-element impedances to the RLC case, that is, to break up complicated functions into simpler ones until the canonic form is finally reached. However, the RLC case is sufficiently more complex than the two-element case that it is difficult to visualize the splitting process in terms of pole-zero plots or partial fractions, and, in addition, the Bott and Duffin or a similar procedure must be used to realize the RLC driving-point impedances without transformers.

For these reasons, a different approach will be used in the RLC problem. The method is almost exactly similar to that given by Fialkow and Gerst. It will be shown that the general RLC grounded voltage transfer function can be realized as the parallel combination of two lossless networks, each of which is in series with a one-ohm resistor. The methods of the previous section can then be used to realize the lossless networks.

Consider the general RLC voltage transfer function

$$A = \frac{a_n s^n + a_{n-1} s^{n-1} + -- a_1 s + a_o}{b_n s^n + b_{n-1} s^{n-1} + -- b_1 s + b_o} = \frac{KN}{D}$$

where D is a Hurwitz polynomial and $a_j \leqslant b_j$. Denoting by m and n the even and odd parts of a polynomial, we can rewrite this function as

$$A = \frac{m_a + n_a}{m_b + n_b}$$

To split this into two lossless networks, write it as

$$A = \frac{\dfrac{m_a}{m_b + n_b} + \dfrac{n_a}{m_b + n_b}}{\dfrac{m_b}{m_b + n_b} + \dfrac{n_b}{m_b + n_b}}$$

$$= \frac{y_{12(1)} + y_{12(2)}}{y_{22(2)} + y_{22(2)}}$$

Now consider network 1:

$$y_{12(1)} = \frac{m_a}{m_b + n_b}$$

$$y_{22(1)} = \frac{m_b}{m_b + n_b}$$

$$A_1 = \frac{m_a}{n_b}$$

A_1 is the voltage transfer function of an LC network, but the output admittance $y_{22(1)}$ does not correspond to a lossless network. However, if we write

$$Z_1 = \frac{1}{y_{22(1)}} = \frac{m_b + n_b}{m_b} = 1 + \frac{n_b}{m_b}$$

it becomes apparent that network 1 corresponds to a lossless network in series with a one-ohm resistor (as in Fig. 38a). For this lossless network,

$$y'_{12(1)} = \frac{m_a}{n_b}$$

$$y'_{22(1)} = \frac{m_b}{n_b}$$

$$A'_1 = \frac{m_a}{m_b}$$

The methods of the previous section can now be used to realize the lossless network. In particular, it may be necessary, first, to remove a series lossless element if the degree of n_b is greater than that of m_b.

For network 2,

$$y_{12(2)} = \frac{n_a}{m_b + n_b}$$

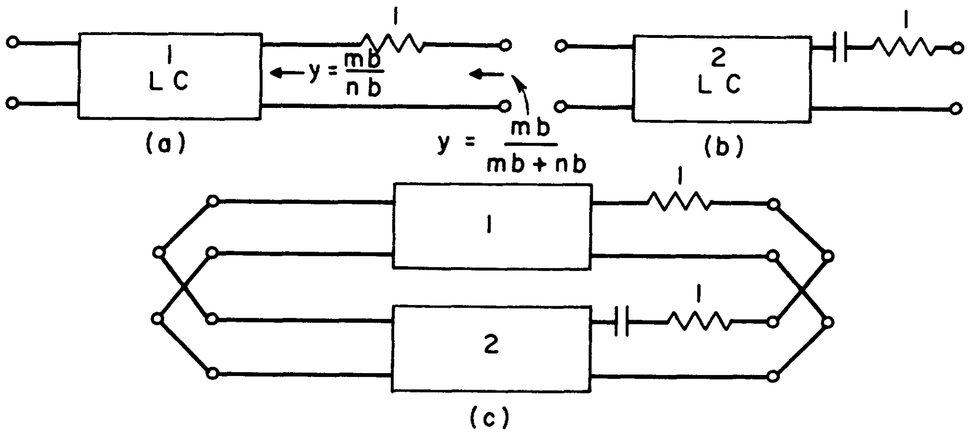

Fig. 38. RLC grounded networks.

$$y_{22(2)} = \frac{n_b}{m_b + n_b}$$

$$A_2 = \frac{n_a}{n_b}$$

A_2 is the voltage transfer function of an LC network; an exactly similar process can be carried out by removing a one-ohm resistor from $1/y_{22(2)}$,

$$Z_2 = \frac{1}{y_{22(2)}} = \frac{m_b + n_b}{n_b} = 1 + \frac{m_b}{n_b}$$

then

$$y'_{12(2)} = \frac{n_a}{m_b}$$

$$y'_{22(2)} = \frac{n_b}{m_b}$$

$$A'_2 = \frac{n_a}{n_b}$$

The only complication is that an s-factor cancels out of the numerator and denominator of A_2, as in one of the previously considered two-element-kind realizations, and a series capacitor must be removed from Z_2' in order to produce the required network. Network 2 then appears as in Fig. 38b. The over-all network is the parallel combination of these two networks and appears as in Fig. 38c. This type of configuration is a logical extension of Cauer's work on single-loaded lossless networks.

EXAMPLE.

$$A = \frac{(s^2 - s + 1)(s + 1)}{(s^2 + s + 1)(s + 1)} = \frac{s^3 + 1}{s^3 + 2s^2 + 2s + 1}$$

To split this into two lossless networks in series with one-ohm resistors, write

$$A = \frac{\dfrac{s^3}{s^3 + 2s^2 + 2s + 1} + \dfrac{1}{s^3 + 2s^2 + 2s + 1}}{\dfrac{s^3 + 2s}{s^3 + 2s^2 + 2s + 1} + \dfrac{2s^2 + 1}{s^3 + 2s^2 + 2s + 1}}$$

$$= \frac{y_{12(1)} + y_{12(2)}}{y_{22(1)} + y_{22(2)}}$$

For network 2,

$$y_{22(2)} = \frac{2s^2 + 1}{s^3 + 2s^2 + 2s + 1}$$

$$y_{12(2)} = \frac{1}{s^3 + 2s^2 + 2s + 1}$$

$$A_2 = \frac{1}{2s^2 + 1}$$

The voltage transfer function is a canonic LC form, and can be realized by a Cauer expansion after the one-ohm resistor has been removed. The whole process can be accomplished by one continued-fraction development,

$$y_{22(2)} = \cfrac{1}{1 + \cfrac{1}{\frac{1}{2}s + \cfrac{1}{\frac{4}{3}s + \cfrac{1}{\frac{3}{2}s}}}}$$

and the resultant network is the top ladder of Fig. 39.

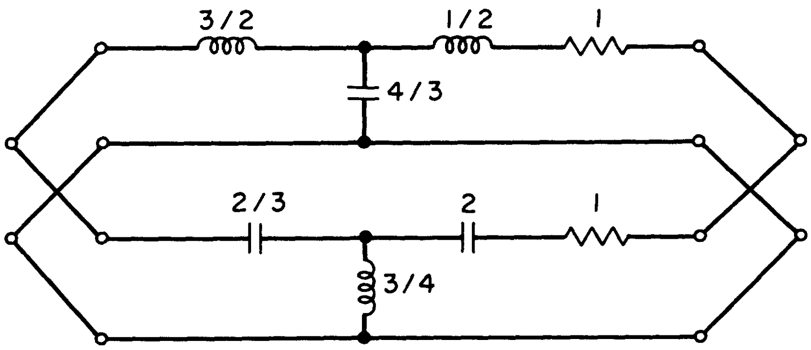

Fig. 39. RLC network.

For network 1,

$$y_{22(1)} = \frac{s^3 + 2s}{s^3 + 2s^2 + 2s + 1}$$

$$y_{12(1)} = \frac{s^3}{s^3 + 2s^2 + 2s + 1}$$

$$A_1 = \frac{s^2}{s^2 + 2}$$

This voltage transfer function is also a canonical form and can be realized as a Cauer form after the one-ohm resistor is removed. As before, the whole process can be accomplished by one continued-fraction expansion

$$y_{22(1)} = \cfrac{1}{1 + \cfrac{1}{2s} + \cfrac{1}{\cfrac{4}{3s} + \cfrac{1}{\cfrac{3}{2s}}}}$$

and the resultant network is the bottom ladder of Fig. 39.

The over-all network is the parallel combination of these two ladders as shown in Fig. 39. In this simple example, the series elements at the output ends of the ladders were developed in the normal course of the Cauer development.

5. TWO TERMINAL-PAIR (UNGROUNDED) NETWORKS; TWO- AND THREE-ELEMENT-KIND FUNCTIONS

The method given here is exactly similar to that given by Fialkow and Gerst. If an ungrounded network is desired, the numerator of the voltage transfer function is not required to have positive coefficients, and the most general form of the function is

$$A = \frac{a_n s^n + a_{n-1} s^{n-1} + - - a_1 s + a_o}{b_n s^n + b_{n-1} s^{n-1} + - - b_1 s + b_o} = \frac{KN}{D}$$

where D is a Hurwitz polynomial for RLC networks, a negative real root polynomial for RC or RL networks or a j-axis polynomial for LC networks, and $a_j \leqslant b_j$. To realize this function, we split KN:

$$KN = N_1 - N_2$$

where N_1 and N_2 consist of all the terms of KN with positive and negative coefficients, respectively. Then we can write

$$A = \frac{N_1}{D} - \frac{N_2}{D} = A_1 - A_2$$

Now A has been split into the difference between two other voltage transfer functions, each of which can be realized as a grounded network by using the previously discussed methods. Then these two grounded networks are connected as in Fig. 40, in order to realize the over-all voltage transfer function.

EXAMPLE. To illustrate this method, a very simple example will be used.

$$A = \frac{s - 1}{s + 1}$$

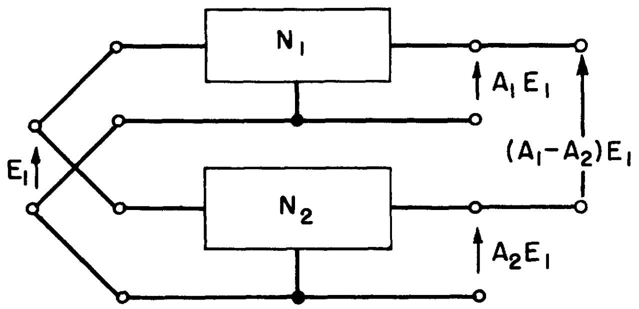

Fig. 40. General ungrounded configuration.

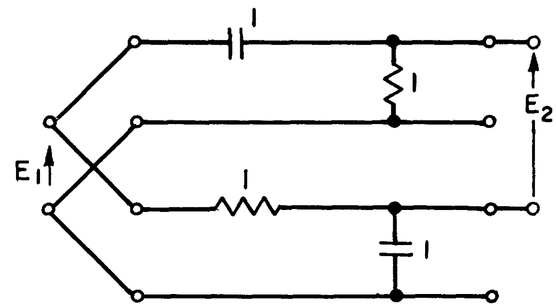

Fig. 41. An Example.

Splitting this into the difference between two grounded voltage transfer functions, we have

$$A = A_1 - A_2 = \frac{s}{s+1} - \frac{1}{s+1}$$

$$= \frac{1}{1+\frac{1}{s}} - \frac{\frac{1}{s}}{1+\frac{1}{s}}$$

Each of these grounded functions can be realized by inspection, and the over-all network is the appropriate interconnection of these networks (as in Fig. 41). Notice that, for this particular function, the resultant network is identical in form with a lattice realization of the same voltage transfer function.

Acknowledgment

The author wishes to express his appreciation to Professor E.A. Guillemin whose keen insight into the basic philosophy of network synthesis has been the inspiration for this work. He also wishes to thank Professor S.J. Mason and Professor F.M. Reza who helped derive some of the theorems.

References

1. R.M. Foster, A reactance theorem, Bell System Tech. J. 3, 259-267 (1924).
2. W. Cauer, Die Verwirklichung von Wechselstromwiderständen vorgeschriebener Frequenzabhängigheit, Arch. Elektrotech. 17, 355-388 (1926).
3. O. Brune, Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency, J. Math. Phys. 10, 191-236 (1931).
4. S. Darlington, Synthesis of reactance four-poles, J. Math. Phys. 18, 257-353 (1939).
5. R. Bott and R. Duffin, Correspondence, J. Appl. Phys. 20, 816 (1949).
6. G.M. Gewertz, Synthesis of a finite four-terminal network from its prescribed driving-point functions and transfer function, J. Math. Phys. 19, 1-257 (1932-33).
7. W. Cauer, Ausgangsseitig Leerlaufende Filter, ENT 16 (1939).
8. E.A. Guillemin, R-C coupling networks, Report 43-10/11/44, Radiation Laboratory, M.I.T., Oct. 11, 1944.
9. J.L. Bower and P.F. Ordung, The synthesis of resistor-capacitor networks, Proc. IRE 38, 263-269 (March 1950).
10. E.A. Guillemin, Synthesis of R-C networks, J. Math. Phys. 28, 22-42 (April 1949).
11. L. Weinberg, New synthesis procedures for realizing transfer functions of RLC and RC networks, Technical Report 201, Research Laboratory of Electronics, M.I.T., Sept. 14, 1951.
12. B.J. Dasher, Synthesis of RC transfer functions as unbalanced two terminal-pair networks, Technical Report 215, Research Laboratory of Electronics, M.I.T., Nov. 20, 1951.
13. E.A. Guillemin, Network Analysis and Synthesis (to be published).
14. A. Fialkow and I. Gerst, The transfer function of general two terminal-pair RC networks, J. Math. Phys. 30, 49-72 (July 1952).
15. A. Fialkow and I. Gerst, The transfer function of networks without mutual reactance, Quart. Appl. Math. 12, 129-134 (July 1954).
16. A. Fialkow and I.Gerst, RLC lattice transfer functions, Proc. IRE 43, 462-469 (April 1935).
17. Theorems 1 and 2 were derived in conjunction with Professor F.M. Reza; P.M. Lewis II and F.M. Reza, A note on the transfer ratio of passive RLC networks, Proc. IRE 42, 1452 (Sept. 1954).
18. P.M. Lewis II and F.M. Reza, Voltage transfer functions, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., April 15, 1954.
19. P.M. Lewis II, Voltage transfer functions—RLC lattice, Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., Oct. 15, 1954, p. 97; Trans. IRE, PGCT, vol. CT-2 (Sept. 1955).
20. P.M. Lewis II, The concept of the one in voltage transfer synthesis, Trans. IRE, PGCT, vol.CT-2, 316-319 (Dec. 1955).
21. R. Kahal, Synthesis of the transfer function of the two terminal-pair networks, Trans. AIEE 71, 129-134 (1952).
22. R.J. Schwarz, A note on the transfer voltage ratio of resistance networks with positive elements, Proc. IRE 43, 1670 (Nov. 1955).
23. H. Poincaré, Analyse Mathématique sur les Equations Algébriques, Compt. rend. 97, 1418 (1883).