#3

ON BASIC EXISTENCE THEOREMS
III. THEORETICAL CONSIDERATIONS ON RATIONAL FRACTION EXPANSIONS FOR NETWORK FUNCTIONS

MANUEL V. CERRILLO
ERNEST A. GUILLEMIN

LOAN COPY only

TECHNICAL REPORT NO. 233

JUNE 4, 1952

RESEARCH LABORATORY OF ELECTRONICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

The research reported in this document was made possible through support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research) and the Air Force (Air Materiel Command), under Signal Corps Contract No. DA36-039 sc-100, Project No. 8-102B-0; Department of the Army Project No. 3-99-10-022.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 233 June 4, 1952

ON BASIC EXISTENCE THEOREMS

III. THEORETICAL CONSIDERATIONS ON RATIONAL FRACTION EXPANSIONS FOR NETWORK FUNCTIONS

Manuel V. Cerrillo
Ernest A. Guillemin

Abstract

A basic discussion on certain aspects concerning the rational expansion of "transfer" and "positive-real" functions is presented here. The submitted results are a partial conclusion extracted from an extensive study on the foundation of network theory, particularly in connection with the problem of network synthesis.

The present report is divided into two parts. Part I deals with the establishment of a set of integrals which characterize transfer functions. These integrals express the necessary and sufficient condition for the existence of transfer functions. Part II deals with some basic aspects of the rational expansion of transfer functions. The integral representations above serve as the fundamental mathematical tools for the construction of such rational expansions. Particularized aspects of the expansion problem are dealt with here, but they are general enough to show the main ideas and trends which are followed in a more advanced investigation on the subject. The presentation of such ideas and methods constitutes the principal aim of this report.

Table of Contents

ON BASIC EXISTENCE THEOREMS
III. THEORETICAL CONSIDERATIONS ON RATIONAL FRACTION EXPANSIONS FOR NETWORK FUNCTIONS

Part I. A New Set of Integral Representation for Laplace Transforms

1.0 Introduction.

The theory of construction of rational approximation of transfer functions is primarily based on a set of new integral representations for direct and inverse Laplace transformations. Hence, we will present them first. Lack of space compels us to omit proof and detailed discussion on the subject. Only final results will be provided. The reader will find detailed information in reference 1.

1.1 Direct and Inverse Laplace Transformations.

Let

$$f(t) \begin{cases} = 0 & \text{for } t < 0 \\ \neq 0 & \text{for } t > 0, \text{ and such that } |f(t)| < Me^{ct} \end{cases} \qquad 1(1.1)$$

where M and c are finite positive constants. Let c_o, the abscissa of convergence, be the lowest bound of c which satisfies this inequality. Then, as is well known, f(t) is Laplace transformable. Its transform F(s) is analytic in the right half of the s plane beyond the line parallel to the ω axis and at a distance c_o from it.

The direct and inverse Laplace transformations are given, respectively, by the well-known integrals

$$\left.\begin{aligned} F(s) &= \int_0^\infty f(t)\ e^{-st}\ dt \\ f(t) &= \frac{1}{2\pi i} \int_{\overline{\Gamma}} F(s)\ e^{st}\ ds \end{aligned}\right\} \quad s = \sigma + i\omega \qquad 2(1.1)$$

where the contour $\overline{\Gamma}$ is a line going from $c_o - i\infty$ to $c_o + i\infty$ in such a way that all the singularities of F(s) lie continuously to the left of $\overline{\Gamma}$.

1.2 Contours of Integration.

For the purpose of our representation we shall select $\overline{\Gamma}$ as a symmetric curve having branches Γ and Γ'. Γ' is, by construction, the mirror image of Γ with respect to the real axis. (See fig. 1(1.2).) An arbitrary point of the s plane will be denoted by

$$s = \sigma + i\omega \qquad 1(1.2)$$

For points on the contour $\bar{\Gamma}$ we shall use, for convenience, the notation

$$S = \gamma + i\lambda \qquad 2(1.2)$$

In case the contour Γ, from γ_0 to ∞, is an analytic curve, we will assume that its analytic expression is given by

$$\phi(\gamma,\lambda) = 0 \qquad 3(1.2)$$

Alternately, we may determine a point P on Γ by its arc distance ℓ to the point γ_0 on the real axis. Other needed notation is indicated in Fig. 2(1.2).

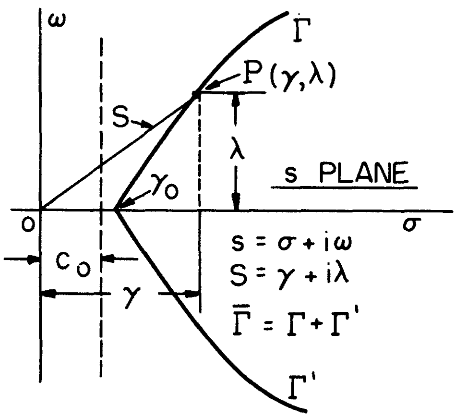

Fig. 1 (1.2)

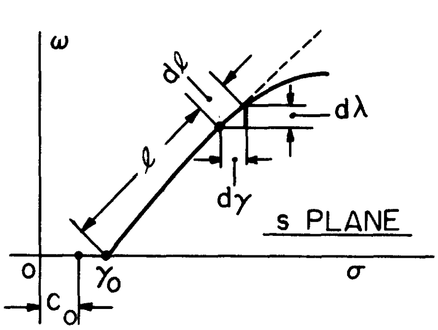

Fig. 2 (1.2)

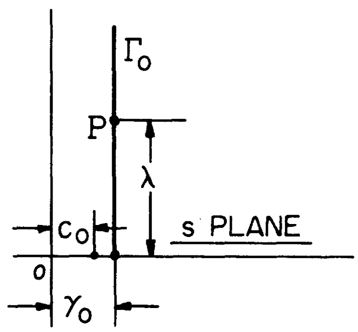

Fig. 3 (1.2)

A particular type of Γ contour, designated by Γ_0 and formed by the semi-infinite straight line parallel to the ω axis, will be frequently used during the discussion. (See fig. 3(1.2).) Finally, the real and imaginary part of F(s) will be denoted by

$$F(s) = U(\sigma,\omega) + iV(\sigma,\omega) \qquad 4(1.2)$$

For points on the contour $\bar{\Gamma}$,

$$F(S) = U(\gamma,\lambda) + iV(\gamma,\lambda) \qquad 5(1.2)$$

1.3 A New Set of Integral Representation for Laplace Transforms.

In this section we will introduce a first set of integral representation for direct and inverse Laplace transformations.

Let f(t) be defined as in 1(1.1). On the assumption that the contour of integration $\bar{\Gamma}$ runs from $c_o - i\infty$ to $c_o + i\infty$ without touching any singularity of F(s) and running in such a way that these singularities continuously remain to the left of the contour $\bar{\Gamma}$, then the following set of integrals is valid:

$$\left.\begin{aligned} F(s) &= \frac{2}{\pi} \iint_{\Gamma} \left\{ \lambda d\gamma + (s-\gamma)\, d\lambda \right\} \frac{U(\gamma,\lambda)}{(s-\gamma)^2 + \lambda^2} \\ f(t) &= \frac{2}{\pi} \int_{\Gamma} e^{\gamma t}\, U(\gamma,\lambda) \left[\cos \lambda t\, d\lambda + \sin \lambda t\, d\gamma \right] \end{aligned}\right\} \quad 1(1.3)$$

$$\left.\begin{aligned} U(\sigma,\omega) &= \int_0^{\infty} f(t)\, e^{-\sigma t} \cos \omega t\, dt \\ U(\sigma,\omega) &= \text{Real } F(s) \end{aligned}\right\} \quad 2(1.3)$$

For the particular contour Γ_o, these expressions (1(1.3)) take a simpler form

$$\left.\begin{aligned} F(s) &= \frac{2(s-\gamma_o)}{\pi} \int_{\Gamma_o} \frac{U(\gamma_o,\lambda)}{(s-\gamma_o)^2 + \lambda^2}\, d\lambda \\ f(t) &= \frac{2e^{\gamma_o t}}{\pi} \int_{\Gamma_o} U(\gamma_o,\lambda) \cos \lambda t\, d\lambda \\ & \qquad \gamma_o > c_o \end{aligned}\right\} \quad 3(1.3)$$

Under the restrictions imposed on f(t) and the contour Γ, this set of integrals exists in the Riemann sense. Details in this connection can be found in reference 1.

1.4 Stieltjes Generalization of the Set of Integral Representation for Laplace Transforms.

For the purpose of producing a mathematical tool which yields to the construction of rational expansions of transfer functions, as well as to discover some basic properties of such functions, we will generalize the set of integral representations given above. For simplicity in the presentation we will first consider the generalization of the set of integrals 3(1.3) which correspond to the particular contour Γ_o, Fig. 3(1.2). For the general contour Γ the generalization will be made later.

We shall now explain the character of the intended generalization.

(a) The integral representation 3(1.3) is valid when γ_o is larger than the abscissa of convergence c_o. The representation 3(1.3) breaks down if we allow the contour Γ_o to go through c_o, since in this position the contour Γ_o touches at least one of the singularities of F(s). Under this circumstance the integrals do not exist in the Riemann sense.

(b) The theory of Laplace transforms shows that the function $U(\sigma, \omega)$ is continuous and bounded at every point of the open half of the s plane defined by $\sigma > c_o$. But this is not necessarily the case for every other point of the remaining s plane. For example, let $f(t) = \cos \Omega t$, Ω being a real constant. We know that $c_o = 0$, and $F(s) = 1/(s^2 + \Omega^2)$. The real part of this function along the imaginary axis is zero almost everywhere on the axis except at the points $s = \pm i\Omega$, where the real part has an infinite discontinuity of the impulse type (the impulse area is constant). This example shows clearly a case in which the Riemann integral 3(1.3) fails to exist if we set $\gamma_o = 0$.

(c) The purpose of our first generalization is to establish an integral representation which may be valid for $\gamma_o = c_o$ and such that it coincides with 3(1.3) when $\gamma_o > c_o$. This is done by the introduction of Stieltjes integrals.

Let us introduce the so-called distribution function defined by

$$\phi(\gamma_o, \lambda) = \int_0^\lambda U(\gamma_o, \mu)\, d\mu \qquad 1(1.4)$$

$$\tau(\gamma_o, t) = \int_0^t f(\mu)\, e^{-\gamma_o \mu}\, d\mu \qquad 2(1.4)$$

The theory of Laplace transforms tells us that the integrals 1(1.4) and 2(1.4) exist if f(t) is Laplace-transformable. Furthermore, they are of "bounded variation"; this implies the existence of the integrals

$$\int_0^\infty |U(\gamma_o, \mu)|\, d\mu < \infty \qquad 3(1.4)$$

$$\int_0^\infty |f(\mu)\, e^{-\gamma_o \mu}|\, d\mu < \infty \qquad 4(1.4)$$

which are called the total or complete variations of ϕ and τ, respectively.

We shall use the distribution functions 1(1.4) and 2(1.4) to define the Stieltjes generalization of the integral representation 3(1.3). The new set of extended integrals reads

$$f(t) = \frac{2\, e^{\gamma_0 t}}{\pi} \int_{\Gamma_0} \cos \lambda t \; d\phi(\gamma_0, \lambda) \qquad 5(1.4)$$

$$F(s) = \frac{2(s-\gamma_0)}{\pi} \int_{\Gamma_0} \frac{d\phi(\gamma_0,\lambda)}{(s-\gamma_0)^2 + \lambda^2} \qquad 6(1.4)$$

$$U(\gamma_0,\lambda) = \int_0^{\infty} \cos \lambda t \; d\tau(\gamma_0, t) \qquad 7(1.4)$$

$$\gamma_0 \geqslant c_0$$

It can be shown with ease, (See reference 1.) that:

(a) the above Stieltjes representation is valid for $\gamma_0 \geqslant c_0$;

(b) it coincides with 3(1.3) for $\gamma_0 > c_0$.

Examples of the application of these integrals will be found later.

1.5 Some Preparatory Steps for Deriving a Set of Fundamental Properties of Transfer Functions.

We wish to derive from our integrals a set of fundamental properties of transfer functions. For this derivation we need some preparatory steps which are given in this section.

Also, we need the following well-known theorem:

Theorem 1(1.5). "Let $\phi(\lambda)$ be a real function of bounded variation in the interval (a, b). Then, there exist two functions $\phi^{(+)}(\lambda)$ and $\phi^{(-)}(\lambda)$ which in (a, b) are: (i) nonnegative, (ii) nondecreasing, (iii) bounded and vanishing at $\lambda = a$, and (iv) discontinuous at the same points as $\phi(\lambda)$, such that

$$\left.\begin{aligned} \phi(\lambda) - \phi(a) &= \phi^{(+)}(\lambda) - \phi^{(-)}(\lambda) \\ V(a,\lambda) &= \phi^{(+)}(\lambda) + \phi^{(-)}(\lambda) \end{aligned}\right\} \qquad 1(1.5)$$

where $V(a,\lambda)$ is the variation of $\phi(\lambda)$ in the interval (a,λ)." The proof is simple and can be found in elementary books on Stieltjes integrals.

Our concern here is to construct the functions $V(a,\lambda)$, $\phi^{(+)}(\lambda)$, and $\phi^{(-)}(\lambda)$ in connection with our integral representations 5(1.4), 6(1.4), and 7(1.4).

Consider first 5(1.4) and 6(1.4).

The functions indicated in the theorem can be constructed as follows.

$$\left.\begin{aligned} V(a,\lambda) &= \int_0^{\lambda} |U(\gamma_o,\mu)|\, d\mu \\ \phi^{(+)}(\gamma_o,\lambda) &= \int_0^{\lambda} U_{(1)}(\gamma_o,\lambda)\, d\lambda \\ \phi^{(-)}(\gamma_o,\lambda) &= \int_0^{\lambda} U_{(2)}(\gamma_o,\lambda)\, d\lambda \end{aligned}\right\} \qquad 2(1.5)$$

where

$$\left.\begin{aligned} U_{(1)}(\gamma_o,\lambda) &= \begin{cases} |U(\gamma_o,\lambda)| & \text{for } U(\gamma_o,\lambda) > 0 \\ \equiv 0 & \text{for } U(\gamma_o,\lambda) < 0 \end{cases} \\ U_{(2)}(\gamma_o,\lambda) &= \begin{cases} |U(\gamma_o,\lambda)| & \text{for } U(\gamma_o,\lambda) < 0 \\ \equiv 0 & \text{for } U(\gamma_o,\lambda) > 0 \end{cases} \end{aligned}\right\} \qquad 3(1.5)$$

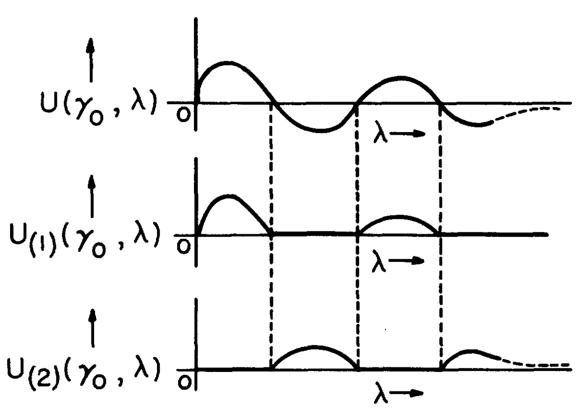

Fig. 1 (1.5)

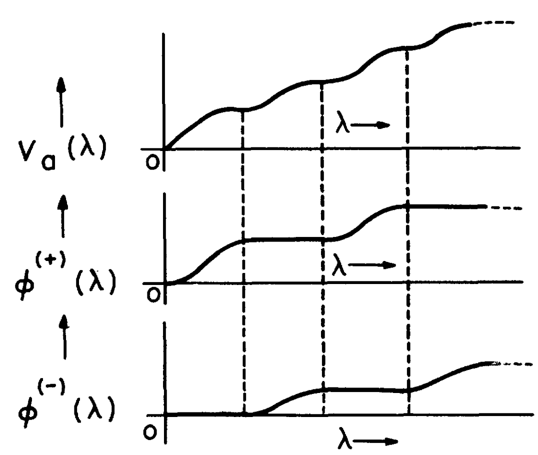

Fig. 2 (1.5)

Figure 1(1.5) provides a clear, graphical illustration of the process of extraction of $U_{(1)}(\gamma_o,\lambda)$ and $U_{(2)}(\gamma_o,\lambda)$ from $U(\gamma_o,\lambda)$. Figure 2(1.5) produces the corresponding graphs of $V_a(\lambda)$, $\phi^{(+)}(\lambda)$, and $\phi^{(-)}(\lambda)$.

1.6 Separation of the Set of Integral Representation.

The boundedness of the ϕ and τ functions and theorem 1(1.5) justify the writing of the integrals 5(1.4), 6(1.4), and 7(1.4) in terms of the nonnegative, nondecreasing, etc., functions $\phi^{(+)}$, $\phi^{(-)}$, $\tau^{(+)}$, $\tau^{(-)}$ as follows:

$$\left.\begin{aligned} f(t) &= f_{(a)}(t) - f_{(b)}(t) \\ F(s) &= F_{(a)}(s) - F_{(b)}(s) \\ U(\gamma_o,\lambda) &= U_{(a)}(\gamma_o,\lambda) - U_{(b)}(\gamma_o,\lambda) \end{aligned}\right\} \qquad 1(1.6)$$

where

$$\left.\begin{aligned} \begin{matrix} f_{(a)}(t) = \\ \\ f_{(b)}(t) = \end{matrix} &\left\{ \frac{2e^{\gamma_o t}}{\pi} \int_{\Gamma_o} \cos \lambda t \begin{Bmatrix} d\phi^{(+)}(\gamma_o,\lambda) \\ \\ d\phi^{(-)}(\gamma_o,\lambda) \end{Bmatrix} \right. \\ \begin{matrix} F_{(a)}(s) = \\ \\ F_{(b)}(s) = \end{matrix} &\left\{ \frac{2(s-\gamma_o)}{\pi} \int_{\Gamma_o} \frac{1}{(s-\gamma_o)^2+\lambda^2} \begin{Bmatrix} d\phi^{(+)}(\gamma_o,\lambda) \\ \\ d\phi^{(-)}(\gamma_o,\lambda) \end{Bmatrix} \right. \\ \begin{matrix} U_{(a)}(\gamma_o,\lambda) = \\ \\ U_{(b)}(\gamma_o,\lambda) = \end{matrix} &\left\{ \int_0^\infty \cos \lambda t \begin{Bmatrix} d\tau^{(+)}(\gamma_o,t) \\ \\ d\tau^{(-)}(\gamma_o,t) \end{Bmatrix} \right. \end{aligned}\right\} \qquad 2(1.6)$$

where the functions $\tau^{(+)}$ and $\tau^{(-)}$ are constructed from 2(1.4) similarly to the construction of $\phi^{(+)}$ and $\phi^{(-)}$.

1.7 The Herglotz Theorem. Conditions for $F_{(a)}(s)$ and $F_{(b)}(s)$ to be (p,r) Functions.

The main result of this section is the establishment of a fundamental property of the functions $F_{(a)}(s)$ and $F_{(b)}(s)$. The first property reads:

Theorem 1(1.7). "In the open half of the s plane defined by $\sigma > c_o$, $(s = \sigma + i\omega)$, the functions $F_{(a)}(s)$ and $F_{(b)}(s)$ are real for s real and have a positive, nonvanishing real part at every point of the half-plane above." (See reference 1.)

We will assume that the reader is acquainted with:

(1) the definition of (p,r) functions (not only when they appear as ratios of polynomials);

(2) the following fundamental theorem of network (and potential) theory:

Theorem 2(1.7). (Herglotz-Cauer). "The necessary and sufficient condition for a function Ƶ(s) to be analytic and have a nonvanishing real part for $\sigma > 0$ and to be real for s real (in other words, to be (p,r)) is that it can be expressed by the Stieltjes integral

$$Ƶ(s) = \frac{2s}{\pi} \int_0^\infty \frac{d\theta(\lambda)}{s^2 + \lambda^2} + sC \qquad 1(1.7)$$

where $\theta(\lambda)$ is a positive, nondecreasing, real function of bounded variation and C is a positive constant equal to $\lim_{s \to \infty} sƵ(s)$." (The notation Ƶ(s) is a contraction of the letters Z and Y, since impedances and admittances are (p,r) functions.)

In the light of theorem 2(1.7) and from 2(1.6) there follows:

Theorem 3(1.7). "Suppose that F(s) is such that $c_o \leq 0$. Then, the corresponding functions $F_{(a)}(s)$ and $F_{(b)}(s)$ are each (p,r) functions." This theorem compels us to investigate the subclass of F(s) for which $c_o \leq 0$. An answer is provided by the following:

Theorem 4(1.7). "Let f(t) be a real, single-valued, bounded-almost-everywhere function, which is zero for $t < 0$; f(t) may possess a denumerable set of isolated points of simple discontinuity with a finite jump and a set of isolated points (the exceptional points), where f(t) suffers an impulse of finite area, such that the sum of these areas is finite; then $c_o \leq 0$, and $F_{(a)}(s)$ and $F_{(b)}(s)$ are (p,r)." (See reference 1.)

1.8 A New Fundamental Theorem in the Network Theory.

We are now prepared to produce a fundamental theorem in the theory of networks. Two corollaries follow it. If, by extension of well-known network theory ideas, we define a "transfer function" as the difference of two (p,r) functions (in Part II we justify this extension), then we can produce the fundamental theorem of existence of transfer functions.

Theorem 1(1.8). "The necessary and sufficient condition for a function F(s) to be a transfer function is that it can be represented by the Stieltjes integral

$$F(s) = \frac{2(s-\gamma_0)}{\pi} \int_{\Gamma_0} \frac{d\phi(\gamma_0,\lambda)}{(s-\gamma_0)^2+\lambda^2} = \frac{2s}{\pi} \int_0^{\infty} \frac{d\phi(0,\lambda)}{s^2+\lambda^2} \qquad 1(1.8)$$

for every contour Γ_0 which shall be made to coincide with the upper part of the imaginary axis, as indicated by the last integral. $\phi(\gamma_0,\lambda)$ and $\phi(0,\lambda)$ are functions of bounded variation."

The result is, of course, independent of Γ_0 down to the limiting position $\gamma_0 = c_0$.

Note. Theorem 1(1.8) is given for integrals along contours of the type Γ_0. For compactness in this report we omit the corresponding theorems for other Γ contours.

The following theorems are corollaries of the previous ones.

Theorem 2(1.8). "Let f(t) be as in theorem 4(1.7). Then its Laplace transform is necessarily a transfer function."

Theorem 3(1.8). "Let F(s) be a general transfer function as defined above. Then its inverse Laplace transform is necessarily a function as defined in theorem 4(1.7) (except for a set of zero measure)."

1.9 Examples.

We will now give some illustrative examples concerning the application of our integral representation, particularly in connection with the basic theorem of existence 1(1.8).

Example 1. Take the function e^{-s}, $c_0 = 0$, which represents the unit delay. Since it is a transfer function, as we already know it, then it must necessarily satisfy the integral representation 1(1.8).

From e^{-s} we get

$$U(\gamma_0,\lambda) = e^{-\gamma_0} \cos \lambda$$

Check of the Direct Theorem (necessary condition). Let us set $\gamma_0 > 0$, so that 1(1.8) can be taken as a Riemann integral.

Then

$$F(s) = \frac{2(s-\gamma_0)}{\pi} e^{-\gamma_0} \int_{\Gamma_0} \frac{\cos \lambda \, d\lambda}{(s-\gamma_0)^2+\lambda^2}$$

To evaluate this integral we will make use of the well-known result

$$\int_0^\infty \frac{\cos x}{a^2 + x^2}\, dx = \frac{\pi}{2a} e^{-a}$$

Hence

$$F(s) = \frac{2(s - \gamma_0)}{\pi} e^{-\gamma_0} \frac{\pi}{2(s - \gamma_0)} e^{-(s - \gamma_0)} = e^{-s}$$

as predicted by the direct part of the theorem.

<u>Example 2 (necessary condition).</u> Let us consider the transfer function

$$F(s) = \frac{s - a}{s + b} \qquad a, b, \text{ being positive-real constants}$$

The abscissa of convergence is evidently $c_o = -b$. Then if one sets $\gamma_o = 0$, the integral above exists in the Riemann sense, since $\gamma_o = 0 > -b$.

By direct computation one gets

$$U(0,\lambda) = \frac{\lambda^2 - ab}{\lambda^2 + b^2}$$

and

$$\frac{U(0,\lambda)}{s^2 + \lambda^2} = \frac{L_1}{\lambda^2 + b^2} + \frac{L_2}{\lambda^2 + s^2}$$

where

$$L_1 = -b\frac{a + b}{s^2 - b^2}; \qquad L_2 = \frac{s^2 + ab}{s^2 - b^2}$$

By direct substitution in 1(1.8) and using the well-known result

$$\int_0^\infty \frac{dx}{a^2 + x^2} = \frac{\pi}{2a}$$

we finally obtain

$$F(s) = \frac{2s}{\pi} \int_0^\infty \frac{\lambda^2 - ab}{\lambda^2 + b^2} \frac{d\lambda}{s^2 + \lambda^2} = \frac{s - a}{s + b}$$

as it should be.

Example 3. This example serves to illustrate the converse part of theorem 1(1.8) (sufficient condition). We shall consider a case in which the integral 1(1.8) exists only in the Stieltjes sense. The example selected here is also intended to illustrate a process which will later play a fundamental role in the rational function approximation of transfer functions. Suppose that $U(\gamma_o, \lambda)$, $\gamma_o = 0$, is chosen as indicated in Fig. 1(1.9)a. For this selection the Riemann integral evidently fails to exist. Figure 1(1.9) b, c, and d, shows the construction of the distribution functions $\phi(0,\lambda)$, $\phi^{(+)}(0,\lambda)$, and $\phi^{(-)}(0,\lambda)$.

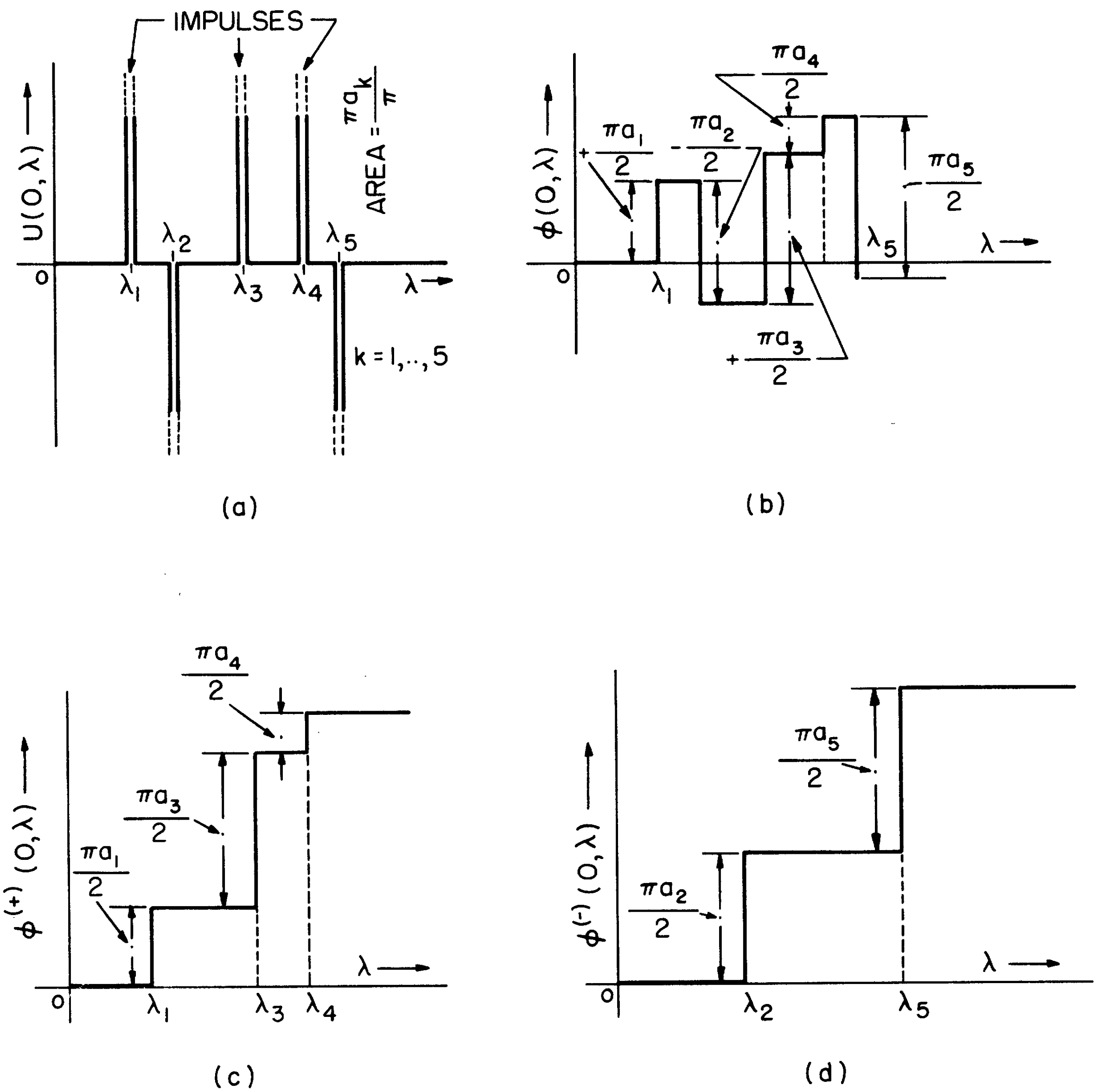

Fig. 1 (1.9)

By using elementary properties of Stieltjes integrals, we reduce the integrals corresponding to $F_{(a)}(s)$ and $F_{(b)}(s)$ to a sum of a finite number of terms

$$F_{(a)}(s) = s\left\{\frac{a_1}{s^2 + \lambda_1^2} + \frac{a_3}{s^2 + \lambda_3^2} + \frac{a_4}{s^2 + \lambda_4^2}\right\}$$

$$F_{(b)}(s) = s\left\{\frac{a_2}{s^2 + \lambda_2^2} + \frac{a_5}{s^2 + \lambda_5^2}\right\}$$

$$F(s) = F_{(a)}(s) - F_{(b)}(s)$$

from which the (p,r) character of $F_{(a)}$, $F_{(b)}$ and the transfer character of F(s) become apparent.

Example 4. This example illustrates the use of 5(1.4) in computing the corresponding time response, f(t), associated with the density distribution $U(0,\lambda)$ as defined in example 3.

By setting $\gamma_o = 0$ in 5(1.4), we reduce the corresponding Stieltjes integral to the finite sum of terms.

$$f(t) = a_1 \cos \lambda_1 t - a_2 \cos \lambda_2 t + a_3 \cos \lambda_3 t + a_4 \cos \lambda_4 t - a_5 \cos \lambda_5 t$$

The reader must note that examples 3 and 4 show the synthesis of a transfer function, and of its time response, from an arbitrary density distribution function $U(0,\lambda)$, subject only to the condition of boundedness.

1.10 Summary.

We close Part I by pointing out three basic features of the integral representation previously introduced:

(1) The functions f(t) and F(s) are uniquely determined in terms of the function $U(\gamma,\lambda)$ along a contour that satisfies conditions already stated, but otherwise are arbitrary.

(2) If the function $U(\gamma,\lambda)$ and the contour Γ can be arbitrarily chosen, except for restrictions already given, then for any selection the integrals above allow us to generate a transfer function F(s) and its corresponding inverse Laplace transform.

(3) Conversely, given f(t) or F(s), we can find the corresponding value of $U(\sigma,\omega)$ along a prescribed contour such that it satisfies the respective integral representations.

Part II. Some Basic Aspects of Rational Expansions for Transfer Functions

2.0 Introduction.

In Part I we have shown the importance of the function $U(\gamma,\lambda)$ in the generation and construction of a transfer function $F(s)$ and of its inverse Laplace transform $f(t)$. The function $U(\gamma,\lambda)$ will be called the "density distribution function" of the corresponding Stieltjes integral representation. For a given $F(s)$ this density distribution function is equal to the real part of $F(s)$ along the contour Γ. Conversely, for a given density distribution function of bounded variation, but otherwise arbitrary, there is a corresponding transfer function.

A transfer function generated by such arbitrary density distribution, say $U(0,\lambda)$ along Γ_o with $\gamma_o = 0$, is not necessarily a rational function. The consultation of a table of Laplace transforms tells us that the transforms of bounded time functions may have transforms which may be algebraic, transcendental, multivalued, or other similar functions of s. For example:

$$\mathscr{L}\left\{e^{-bt}\frac{\sin at}{t}\right\} = \arctan\frac{a}{s+b}$$

$$\mathscr{L}\left\{e^{-bt}\, t^{\frac{1}{\nu}-1}\right\} = \frac{\Gamma\left(\frac{1}{\nu}\right)}{\sqrt[\nu]{s+b}} \qquad \nu > 0$$

An analytic characterization of transfer functions is outside the scope of this paper. The reader will find in reference 1, theorem 1(1.4) a more precise characterization of such a class of functions.

The purpose of Part II is first, to lay some fundamental principles, in a particular way, regarding the rational expansion of transfer functions. The main task is to show that we can construct a sequence of rational functions, say $\left\{P_n(s)/Q_n(s)\right\}_0^{\infty}$, in which each member is a transfer function for itself such that the limiting function, when n tends to ∞, tends uniformly to $F(s)$ for every point of the open half of the s plane defined by $\sigma > 0$, $(s = \sigma + i\omega)$. Corresponding to this sequence, there is a set of exponential functions which simultaneously converge toward $f(t)$ for almost every value of t such that $t > 0$. The mathematical tool for constructing the sequences is supplied by the integral representation, particularly the Stieltjes forms introduced in Part I. We will begin our discussion by considering the contour Γ_o. The procedure followed will show certain basic aspects of the problem of rational expansions, a procedure which will be extended with ease to a more elaborate contour Γ.

For simplicity in the presentation we will consider directly the rational expantion of the functions $F_{(a)}(s)$ and $F_{(b)}(s)$. The rational expansions for $F(s)$ follow immediately from them. Since $F_{(a)}(s)$ and $F_{(b)}(s)$ both have integral representations which have the same analytical structure, it is sufficient to consider one of them directly, say $F_{(a)}(s)$, and to show how to construct the rational sequences associated with $F_{(a)}(s)$. The corresponding procedure for $F_{(b)}(s)$ will be the same after the substitution of $\phi^{(-)}(\gamma,\lambda)$ for $\phi^{(+)}(\gamma,\lambda)$.

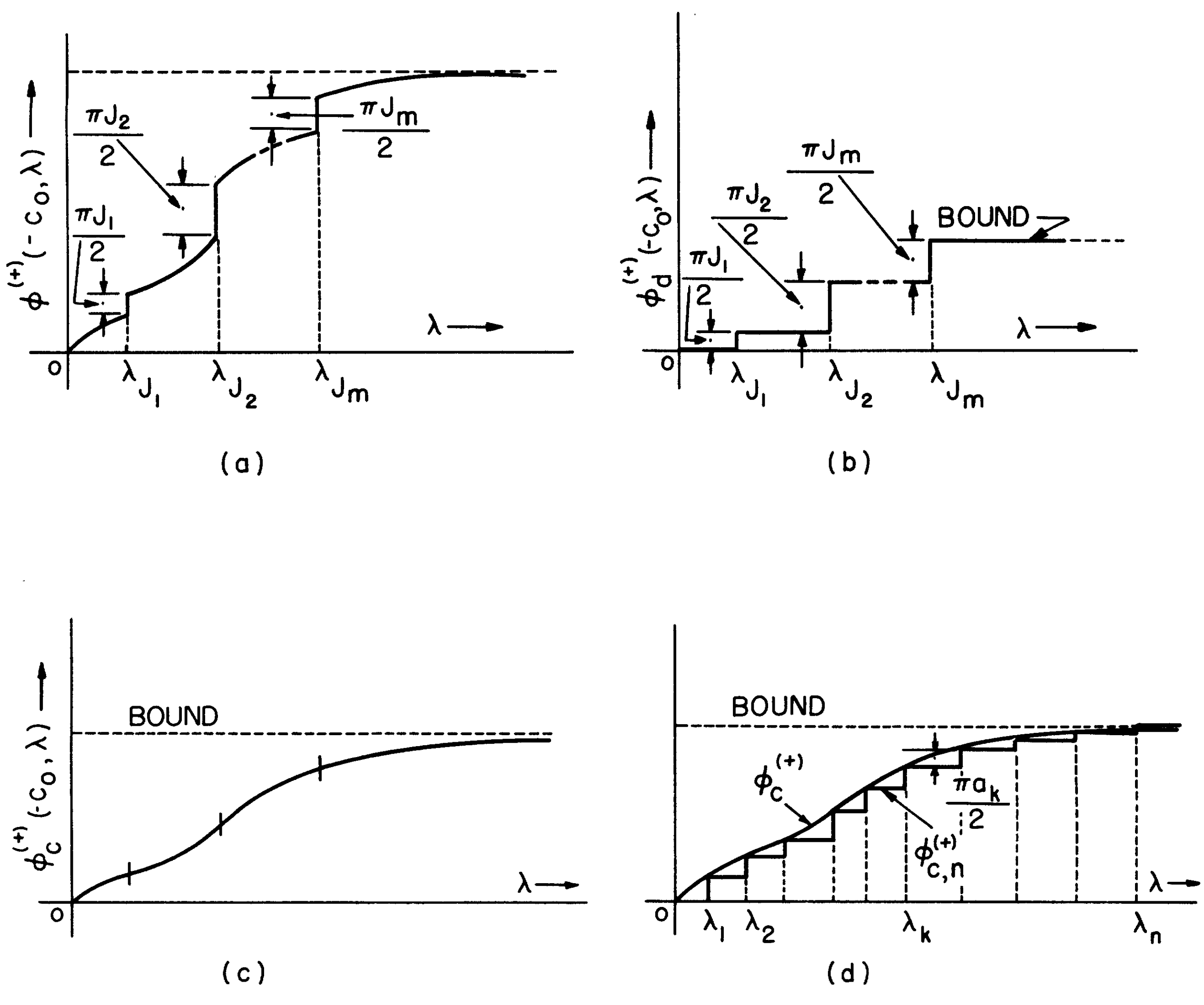

Fig. 1 (2.1)

2.1 Construction of Rational Expansions for Transfer Functions. (First Method.)

Let us assume that F(s) is a given transfer function. We have shown that $c_o \leqslant 0$. Let us set $\gamma_o = -c_o$. From F(s) we can compute $U(-c_o,\lambda)$. By the method indicated in Part I, we can construct $\phi^{(+)}(-c_o,\lambda)$ and $\phi^{(-)}(-c_o,\lambda)$, which are given by

$$\left.\begin{aligned} \phi^{(+)}(-c_o,\lambda) &= \int_0^{\lambda} U_1(-c_o,\mu)\, d\mu \\ \phi^{(-)}(-c_o,\lambda) &= \int_0^{\lambda} U_2(-c_o,\mu)\, d\mu \end{aligned}\right\} \qquad 1(2.1)$$

We will illustrate our method directly with $\phi^{(+)}(-c_o,\lambda)$, which generates $F_{(a)}(s)$. A similar procedure can be followed with $\phi^{(-)}(-c_o,\lambda)$, which generates $F_{(b)}(s)$.

Suppose that $\phi^{(+)}(-c_o,\lambda)$ has a graph as in Fig. 1(2.1)a. We can split $\phi^{(+)}(-c_o,\lambda)$ into its continuous part, $\phi_c^{(+)}(-c_o,\lambda)$, and its discontinuous part, $\phi_d^{(+)}(-c_o,\lambda)$. Let $\lambda_{d,1}, \lambda_{d,2}, \ldots \lambda_{d,n}$ be the points of discontinuity of $\phi_d^{(+)}(-c_o,\lambda)$. See Fig. 1(2.1) b and c. The next step is to approximate the continuous function $\phi_c^{(+)}(-c_o,\lambda)$ by an ascending stair function, say by approaching from below.

The corresponding steps occur at a set of points λ_k, k = 1,2 . . ., n. Let us call $\phi_{c,n}^{(+)}(-c_o,\lambda)$ the resulting stair function. See Fig. 1(2.1)d. The set λ_k may, or may not, be equally spaced. In fact, uneven spacing of the steps, following a strategic array, proves to be convenient in order to produce a better and faster convergence. (For example, the spacing must be closer in the intervals of faster variations of $\phi_c^{(+)}(-c_o,\lambda)$ and far apart in the intervals of slow variation, as becomes apparent from the geometrical picture of $\phi_c^{(+)}(-c_o,\lambda)$.) In Fig. 1(2.1) we did not intentionally follow any specific distribution of the steps because we first want to show the method of expansion in its general aspect.

Since both functions $\phi_d^{(+)}$ and $\phi_{c,n}^{(+)}$ are nonnegative, nondecreasing functions of bounded variation, then the transfer function constructed with them is, according to a previous theorem, necessarily a (p,r) function. This function is given and denoted by

$$
\left.\begin{aligned}
F_{(a),n}(s) &= Z_{(a),n}(s) \\
&= \frac{2(s+c_o)}{\pi}\left\{\int_0^\infty \frac{d\phi_d^{(+)}(-c_o,\lambda)}{(s+c_o)^2+\lambda^2} + \int_0^\infty \frac{d\phi_{c,n}^{(+)}(-c_o,\lambda)}{(s+c_o)^2+\lambda^2}\right\} \\
&= (s+c_o)\sum_{\nu=1}^{\nu=m} \frac{J_\nu}{(s+c_o)^2+\lambda_\nu^2} + (s+c_o)\sum_{k=1}^{k=n} \frac{a_k}{(s+c_o)^2+\lambda_k^2}
\end{aligned}\right\} 2(2.1)
$$

on account of an elementary property of Stieltjes integrals. The function $F_{(a),n}(s)$ is (p,r) for every value of n. This can be verified by the observation of the third member of 2(2.1) together with the fact that J_ν, $\nu = 1, 2, \ldots, m$ and a_k, $k = 1, 2, \ldots, n$ are positive quantities. It is evident that the function $F_{(a),n}$ is a rational function, whose poles lie at $s_\nu = -c_o \pm i\lambda_\nu$ and $s_k = -c_o \pm i\lambda_k$. It must be observed that the poles s_ν, $\nu = 1, 2, \ldots, n$ are really contained in $F_{(a)}(s)$. The poles $s_k = -c_o \pm i\lambda_k$ come from the method of expansion and are spurious with respect to $F_{(a)}(s)$. From Fig. 1(2.1) it is evident that $F_{(a),n}(s) \to F_{(a)}(s)$ uniformly as $n \to \infty$.

A remark is pertinent with regard to the expression 2(2.1). The function $\phi^{(+)}(-c_o,\lambda)$ is computed from the value of $U(\gamma_o,\lambda)$, $\gamma_o = (-c_o)$, along the contour Γ_o. Clearly, for $-c_o < \gamma_o < \infty$, the function $U(\gamma_o,\lambda)$ changes in value. Consequently, the magnitude of the steps J_ν and a_k changes for the same set of numbers λ_ν and λ_k. Then J_ν and a_k are functions of $-c_o$. We want to make this relationship apparent by writing, more explicitly,

$$
F_{(a),n}(s) = (s+c_o)\left\{\sum_{\nu=1}^{\nu=m} \frac{J_\nu(-c_o)}{(s+c_o)^2+\lambda_\nu^2} + \sum_{k=1}^{k=n} \frac{a_k(-c_o)}{(s+c_o)^2+\lambda_k^2}\right\} \qquad 3(2.1)
$$

We can now explain the following situation. Since c_o is the abscissa of convergence, then the function $F_{(a),n}(s)$ must be independent of the position of the contour Γ_o for $-c_o < \gamma_o < \infty$. This means that the residues $J_\nu(-c_o)$ and $a_k(-c_o)$ will change accordingly for other positions of Γ_o in such a way as to keep 3(2.1) invariant.

We close this section by remarking that the procedure given above produces a sequence of (p,r) rational functions which approximate $F_{(a)}(s)$, as closely as we wish, by using intrinsic elements of F(s). A similar treatment produces a sequence of rational functions, say $F_{(b),n}(s)$, which approximate $F_{(b)}(s)$. By writing $F_n(s) = F_{(a),n}(s) - F_{(b),n}(s)$, we evidently obtain the corresponding rational expansion for F(s). We can therefore formally conclude that we have found a certain method of rational expansion for transfer functions.

2.2 Some Properties of Monogenic Functions.

The objective of this section is to study some basic properties of the function $F_{(a),n}(s)$, particularly the position of its zeros as related to the position of its poles. We will show that poles and zeros interlace along Γ_o, $\gamma_o = -c_o$.

The study of these properties is facilitated by the separation of the function $F_{(a),n}(s)$ into two parts:

$$\left.\begin{aligned} F_{(a),n}(s) &= \left\{ (s + c_o) \sum_{\nu=1}^{\nu=m} \frac{J_\nu}{(s + c_o)^2 + \lambda_\nu^2} \right\} + M_{(a),n}(s) \\ M_{(a),n}(s) &= (s + c_o) \sum_{k=1}^{k=n} \frac{a_k}{(s + c_o)^2 + \lambda_k^2} \end{aligned}\right\} \quad 1(2.2)$$

We have indicated before that the poles at $s_\nu = -c_o \pm i\lambda_\nu$, $\nu = 1, 2, \ldots, m$ are contained in the original function F(s), while the poles at $s_k = -c_o \pm i\lambda_k$, $k = 1,2, \ldots, n$ of $M_{(a),n}(s)$ are spurious or nonexisting in $F_{(a)}(s)$ or F(s). We will now study the zero-pole configuration of the function $M_{(a),n}(s)$ and show that poles and zeros interlace along Γ_o, $\gamma_o = -c_o$, and become dense everywhere in the intervals of Γ_o where the function $\phi_c^{(+)}(-c_o,\lambda)$ is not constant. The function $M_{(a),n}(s)$ is called the "monogenic" component of $F_{(a),n}(s)$.

We will consider a few basic properties of $M_{(a),n}(s)$.

<u>First Property.</u> "The function $M_{(a),n}(s)$ has, for every value of n, a real part equal to zero almost everywhere along the line Γ_o, $\gamma_o = -c_o$. The exceptional points are the poles at $-c_o \pm i\lambda_k$, $k = 1,2,\ldots, n$, where the real part suffers a positive impulse of area equal to $\pi a_k/2$." On Γ_o we have $s = -c_o \pm i\omega$. Substituting this value in 1(2.2), we get

$$M_{(a),n}(s) = i\omega \sum_{k=1}^{k=n} \frac{a_k}{\lambda_k^2 - \omega^2} \qquad 2(2.2)$$

which is purely imaginary.

At the exceptional points $s = -c_o \pm i\lambda_k$, the function $\phi^{(+)}_{c,n}(-c_o,\lambda)$ (see fig. 1(2.1)d.) has a jump equal to $\pi a_k/2$. This means that the real part of $M_{(a),n}(s)$ suffers an impulse of $\pi a_k/2$ area.

Second Property. "Let $M'_{(a),n}(s)$ be the derivative of $M_{(a),n}(s)$ with respect to s. The function $M'_{(a),n}(s)$ is real, positive, and nonvanishing along the contour Γ_o, $\gamma_o = -c_o$." A simple direct proof can be given. We will use, however, the general integral representation to produce the proof because of future similar uses of the integral. From our general representation we have

$$M_{(a),n}(s) = \frac{2}{\pi} \int_{\Gamma_o} \frac{(s + c_o)}{(s + c_o)^2 + \lambda^2} \, d\phi^{(+)}_{c,n} \qquad 3(2.2)$$

On account of the uniform convergence we can differentiate, for $s > -c_o$, inside the integral sign. Thus

$$M'_{(a),n}(s) = \frac{2}{\pi} \int_{\Gamma_o} \frac{\lambda^2 - (s + c_o)^2}{\left[(s + c_o)^2 + \lambda^2\right]^2} \, d\phi^{(+)}_{c,n} = \sum_{k=1}^{k=n} a_k \frac{\lambda_k^2 - (s + c_o)^2}{\left[(s + c_o)^2 + \lambda_k^2\right]^2} \qquad 4(2.2)$$

Along Γ_o we have $s = -c_o + i\omega$. Hence

$$M'_{(a),n}(s) = \sum_{k=1}^{k=n} a_k \frac{\lambda_k^2 + \omega^2}{\left[\lambda_k^2 - \omega^2\right]^2} \qquad 5(2.2)$$

which is definite positive for all values of ω. (All a_k are positive.)

Third Property. "The derivative $M'_{(a),n}(s)$ cannot vanish at every point of the s plane defined by $\sigma_o \geqslant -c_o$." The proof follows with ease from 4(2.2).

Fourth Property. "Let $-c_o + i\lambda_k$ and $-c_o + i\lambda_{k+1}$ be two successive poles of $M_{(a),n}(s)$. Then, the imaginary part of $M_{(a),n}(s)$ along Γ_o vanishes once at some point between the poles above." For simplicity, we will supply a strategical heuristic proof. In the vicinity of a pole, say at $-c_o + i\lambda_k$, the function $M_{(a),n}(s)$ behaves as

$$\frac{a_k}{2} \frac{1}{(s + c_o) - i\lambda_k} \qquad (a_k > 0) \qquad 6(2.2)$$

By setting $s + c_o - i\lambda_k = Re^{i\vartheta}$ we obtain for the expression for the imaginary part

$$- \frac{a_k}{2R} \sin \phi \qquad 7(2.2)$$

The sign distribution of the imaginary part in the neighborhood of this pole is given by Fig. 1(2.2)a.

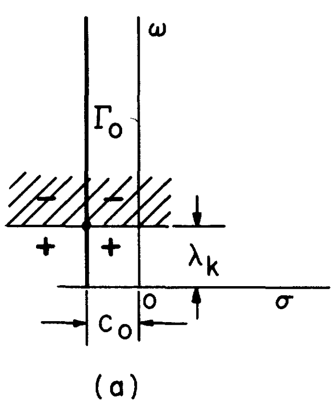

Fig. 1 (2.2)

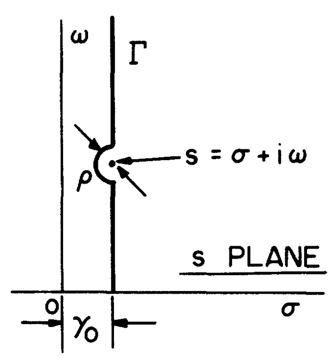

Fig. 2 (2.2)

A similar situation exists for the sign distribution of the imaginary part of $M_{(a),n}(s)$ around the pole at $-c_o + i\lambda_{k+1}$. (See fig. 1(2.2)b.) By combining the results given above and by considering that the imaginary part of $M_{(a),n}(s)$, 2(2.2), is continuous in the interval between the two successive poles, then we can deduce that an odd number of zeros must exist inside such an inverval. Otherwise, the imaginary part would jump from a negative to a positive value in this interval, which cannot take place. Now, since the derivative of $M_{(a),n}(s)$ does not vanish on Γ_o, then there can be one, and only one, zero between these consecutive poles.

Fifth Property. "The position of the intermediate zero located between two consecutive poles, $-c_o + i\lambda_k$ and $-c_o + i\lambda_{k+1}$, can be closely computed for large values of n." Above the lower pole, say $-c_o + i\lambda_k$, the imaginary part is described by $-(a_k/2R_k)$, where R_k is the variable distance from this pole. (See 7(2.2).) Below the upper pole, $-c_o + i\lambda_{k+1}$, the imaginary part, is given by $+(a_{k+1}/2R_{k+1})$, where R_{k+1} is the variable distance from the upper pole. Then, the zero lies at

$$\left.\begin{aligned} \frac{1}{2}\left\{\frac{a_{k+1}}{R_{k+1}} - \frac{a_k}{R_k}\right\} &= 0 \\ R_{k+1} + R_k &= (\lambda_{k+1} - \lambda_k) \end{aligned}\right\} \qquad 8(2.2)$$

from which we obtain

$$a_{k+1}\, R_k = a_k\, R_{k+1} \qquad 9(2.2)$$

or

$$R_k = \frac{a_k}{a_{k+1} + a_k}\left[\lambda_{k+1} - \lambda_k\right] \qquad 10(2.2)$$

above λ_k.

Equation 9(2.2) shows that the zero points can be interpreted as the center of gravity of two concentrated masses a_{k+1} and a_k, separated by the distance $(\lambda_{k+1} - \lambda_k)$.

Sixth Property. "The zeros of function $M_{(a),n}(s)$ are simple; they lie on the line Γ_o, $\gamma_o = -c_o$, and they interlace with the poles at $-c_o \pm i\lambda_k$."

Proof. At a zero of $M_{(a),n}(s)$ the real and imaginary parts vanish together. Along Γ_o, $\gamma_o = -c_o$, the real part vanishes identically at every point which is not a pole. Then, the set of points 10(2.2) are necessarily zeros of $M_{(a),n}(s)$, since both real and imaginary parts of the function vanish at the points of this set.

Hence, there is a sequence of zeros at

$$\overset{o}{S}_k = -c_o \pm i\left(\frac{\lambda_k a_{k+1} + a_k \lambda_{k+1}}{a_{k+1} + a_k}\right) \qquad 11(2.2)$$

which are simple and interlaced with the poles.

This sequence contains the only zeros because there are no zeros at the right of Γ_o, $\gamma_o = -c_o$, since $M_{(a),n}(s)$ is a (p,r) function with respect to Γ_o.

(The real part of a (p,r) function cannot vanish at a finite point to the right of Γ_o.) (Symmetric conditions of $M_{(a),n}(s)$ show that there are also no zeros at the left of Γ_o.)

Seventh Property. "Let n increase without limit. Then distances $\lambda_{k+1} - \lambda_k$ between successive zeros of $M_{(a),n}(s)$ become smaller and smaller, so that the poles become denser and denser at the points where $\phi_c^{(+)}(-c_o,\lambda)$ is not constant. (No steps can come from $\phi_c^{(+)}(-c_o,\lambda)$ = constant.) Simultaneously, the zeros also become denser and denser due to their interlacing with the poles. Then the segments of lines in which $\phi_c^{(+)}(-c_o,\lambda)$ is not constant become 'singular lines' of the function $M_{(a),n}(s)$, but not necessarily singular lines of $F_{(a)}(s)$. Therefore, the function $M_{(a),n}(s)$ does not represent the function $F_{(a)}(s)$ along such singular lines."

Eighth Property. "The function $M_{(a),n}(s)$ is a single value of s." Therefore, when it is multivalued, it can only represent the function $F_{(a)}(s)$ in the corresponding Riemann sheet, where $F_{(a)}(s)$ has a (p,r) branch.

Ninth Property. "The pole-zero chains along Γ_o may produce a confusing situation at first sight. The function $M_{(a),n}(s)$ become successively zero and infinite as we progress along Γ_o. Since this pole-zero configuration does not exist in $M_{(a)}(s)$, then we must investigate the convergence of $M_{(a),n}(s)$ when $n \to \infty$ at points s infinitesimally close to Γ_o."

We will show that "the limiting function $\left\{ \lim_{n\to\infty} M_{(a),n}(s) \right\}$ attains at any point of Γ_o, which is not a singular point of $F_{(a)}(s)$, limiting values, for the real and imaginary parts, which are equal to the corresponding values of $F_{(a)}(s)$, when we approach such a point of Γ_o from the right side of this line."

For a future application of this property, we will not confine the proof to the case of (p,r) functions such as $F_{(a)}(s)$. The property holds in general for F(s), as we shall prove. The corresponding proof for $F_{(a)}(s)$ follows logically.

Proof:

Let us consider the integral representation*

$$F(s) = \frac{2}{\pi} \int_{\Gamma} \frac{\lambda d\gamma + (s-\gamma)\, d\lambda}{(s-\gamma)^2 + \lambda^2} U(\gamma,\lambda) \qquad 12(2.2)$$

*We will also need the equivalent integral representation in terms of the imaginary part which reads (See ref. 1.):

$$F(s) = \frac{2}{\pi} \int_{\Gamma} \frac{(s-\gamma)\, d\gamma - \lambda d\lambda}{(s-\gamma)^2 + \lambda^2} V(\gamma,\lambda) \qquad 12'(2.2)$$

and consider a contour Γ_o. Let us assume: first, that the point s approaches Γ_o from the right of Γ_o; second, that the limiting point on Γ_o is a regular point of F(s); third, that $F'(s) \neq 0$.

When s touches Γ_o, the integrand has a pole at s. In order to avoid this difficulty, we consider that the contour of integrations Γ is composed of two parts: the contour Γ_o and a semicircle surrounding s.
See Fig. 2(2.2). We can write

$$F(s) = \frac{2}{\pi}\int_0^{\omega-\rho} \frac{(s-\gamma_o)\ U(\gamma_o,\lambda)\ d\lambda}{(s-\gamma_o)^2 + \lambda^2} + \frac{2}{\pi}\int_{\rho \curvearrowleft s} \frac{\lambda d\gamma + (s-\gamma)\ d\lambda}{(s-\gamma)^2 + \lambda^2}\ U(\gamma,\lambda)$$

$$+ \frac{2}{\pi}\int_{\omega+\rho}^{\infty} \frac{(s-\gamma_o)\ U(\gamma_o,\lambda)\ d\lambda}{(s-\gamma_o)^2 + \lambda^2}$$

$$= I_1 + I_2 + I_3 \qquad 13(2.2)$$

Let us consider the middle integral. In accordance with Fig. 2(2.2) let

$$s = \sigma + i\omega; \quad \gamma = \sigma + \rho \cos\phi; \quad \lambda = \omega + \rho \sin\phi \qquad 14(2.2)$$

After a few simple algebraic manipulations we get

$$I_2 = -\frac{2}{\pi}\int_{\frac{\pi}{2}+\pi}^{\frac{\pi}{2}} \frac{-\rho + i\omega e^{i\phi}}{\rho^2 - 2i\omega\rho e^{i\phi}}\ U(\sigma + \rho\cos\phi,\ \omega + \rho\sin\phi)\ \rho d\phi \qquad 15(2.2)$$

Let $\rho \to 0$. On the assumption that the point s is a regular point of F(s) and is such that $F'(s) \neq 0$, then

$$U(\gamma,\lambda) \to U(\sigma,\omega) \quad \text{uniformly, as } \rho \to 0$$

Hence

$$I_2 = \frac{2}{\pi} U(\sigma,\omega) \int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\pi} \frac{d\phi}{2} = U(\sigma,\omega) \qquad 16(2.2)$$

Therefore, I_2 is equal to the real part of F(s).

Consider the first and third integral. From Fig. 2(2.2) let

$$\gamma_o = \sigma, \qquad s = \sigma + i\omega$$

After a few algebraic operations we get

$$I_1 + I_3 = \frac{2}{\pi} i\omega \left\{ \int_0^{\omega-\rho} \frac{U(\sigma,\lambda)\ d\lambda}{\lambda^2 - \omega^2} + \int_{\omega+\rho}^{\infty} \frac{U(\sigma,\lambda)\ d\lambda}{\lambda^2 - \omega^2} \right\} \qquad 17(2.2)$$

which is a pure imaginary quantity. We will show that $I_1 + I_3 = iV(\sigma,\omega)$ as $\rho \to 0$. This can be recognized immediately because the integral is the "Hilbert transform," (ref. 2, p. 342), of the imaginary part $V(\sigma,\omega)$. Hence

$$I_1 + I_3 = iV(\sigma,\omega) \qquad 18(2.2)$$

For future purposes we will produce an alternative proof of 17(2.2) without the aid of Hilbert transforms. In the new proof, we obtain directly the function $V(\sigma,\omega)$ from an integration in the vicinity of the point s in a similar method to that used for the function $U(\sigma,\omega)$.

We start, here, with the integral representation 12'(2.2).

The integral surrounding the point s is obtained by using 14(2.2). After a few algebraic manipulations we get

$$I_2' = -\frac{2}{\pi} \int_{\pi+\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\omega e^{i\phi}}{2i\omega\rho e^{i\phi} - \rho^2} V(\sigma + \rho\cos\phi,\ \omega + \rho\sin\phi)\ \rho\ d\phi \qquad 19(2.2)$$

By setting $\rho \to 0$, and from reasoning similar to that used for integral I_2, we get

$$I_2' = iV(\sigma,\omega)$$

Note: Here the integrals I_1' and I_2' (corresponding to I_1 and I_2) are real. By applying Hilbert transforms to them, they reduce to $U(\sigma,\omega)$.

The result given above can be translated with ease to prove the Ninth Property of the sequence $M_{(a),n}(s)$. First, we recall that $M_{(a)}(s)$ belongs to the F(s) class. Second, we set $\sigma = -c_o$. At a regular point of Γ_o, for which $F'(s) \neq 0$, the following limits exist:

$$\left.\begin{aligned} \lim\ (\text{real}\ M_{(a),n}(s)) &= \text{real}\ M_{(a)}(s) \\ \lim\ (\text{imag}\ M_{(a),n}(s)) &= \text{imag}\ M_{(a)}(s) \end{aligned}\right\} n \to \infty \qquad 20(2.2)$$

Consequently, the pole-zero distributions are such that in the limit they produce unique limiting finite functions at regular points, and so forth, of Γ_o.

Tenth Property. Suppose that n is sufficiently large in $M_{(a),n}(s)$, so that two consecutive poles at $-c_o + i\lambda_k$ and $-c_o + i\lambda_{k+1}$ (having residues equal respectively to a_k and a_{k+1}) are very close together. Then, there is a simple way to predict what the limiting value of the real part of $M_{(a)}(s)$ must be. Let U_a be the value of $M_{(a)}(s)$ at a regular point of the boundary. By using the integral I_2 in 13(2.2), we can show with ease that between λ_{k+1} and λ_k

$$U_a \approx \frac{\pi}{2} \frac{a_{k+1} - a_k}{\lambda_{k+1} - \lambda_k} \quad \text{on the boundary } \Gamma_o \qquad 21(2.2)$$

Eleventh Property. By setting $\gamma_o = 0$, the expansions 2(2.1) produce the Foster canonical forms.

2.3 Time Functions.

In this section we will momentarily direct our discussion from the frequency to the time domain. Our next objective is to construct, by means of our basic integrals, the time functions which correspond to $F_{(a)}(s)$, $F_{(a),n}(s)$ and, in particular, to $M_{(a),n}(s)$. Similar methods of construction and of reasoning can be followed to produce the respective time functions corresponding to $F_{(b)}(s)$, $F_{(b),n}(s)$, $M_{(b),n}(s)$.

The construction of the time domain functions is accomplished from the set of functions $\phi^{(+)}$, $\phi_d^{(+)}$, $\phi_c^{(+)}$, and $\phi_{c,n}^{(+)}$ (fig. 1(2.1)) and by using the basic integral 5(1.4) or its alternate forms 2(1.6). We will write in full the corresponding integral in order to illustrate the appropriate notation in the time domain.

$$\left.\begin{array}{l} f_{(a)}(t) = \\ f_{(a),d}(t) = \\ f_{(a),c}(t) = \\ f_{(a),c,n}(t) = \end{array}\right\} \underbrace{\frac{2e^{\gamma_o t}}{\pi} \int_{\Gamma_o} \cos \lambda t}_{\gamma_o = -c_o} \left\{\begin{array}{l} d\phi^{(+)}(\gamma_o,\lambda) \\ d\phi_d^{(+)}(\gamma_o,\lambda) \\ d\phi_c^{(+)}(\gamma_o,\lambda) \\ d\phi_{c,n}^{(+)}(\gamma_o,\lambda) \end{array}\right. \qquad 1(2.3)$$

If we use the corresponding graphs of the $\phi^{(+)}$, $\phi_d^{(+)}$, $\phi_c^{(+)}$, and $\phi_{c,n}^{(+)}$ functions as given in Fig. 1(2.1), then we obtain

$$f_{(a),n}(t) = e^{-c_o t}\left\{\left[\sum_{\nu=1}^{\nu=m} J_\nu(-c_o)\cos\lambda_\nu t\right] + \left[\sum_{k=1}^{k=n} a_k(-c_o)\cos\lambda_k t\right]\right\} \quad 2(2.3)$$

where

$$f_{(a),n} = f_{(a),d} + f_{(a),c,n} \quad 3(2.3)$$

The first summation term represents the time functions corresponding to the contribution of the poles (supposed to be simple) of $F_{(a)}(s)$ which lie on Γ_o. The second summation term represents the time function corresponding to the contribution of the poles of the monogenic function $M_{(a),n}(s)$. This last contribution is equivalent to the contribution of the actual singularities of $F_{(a)}(s)$ which lie to the left of Γ_o.

From the construction of the function $\phi^{(+)}_{c,n}$ it becomes clear that

$$\lim f_{(a),n}(t) \to f_{(a)}(t), \qquad \text{as } n \to \infty$$

uniformly almost everywhere for $t > 0$. The exceptional points are the set of points which contain the discontinuities and impulses of f(t), as admitted in the postulation of theorem 4(1.7). For $t < 0$, the function $f_{(a),n}(t)$ vanishes identically, because of a well-known property of Laplace transformations. In a future section we will resume the discussion of the series 2(2.3).

2.4 A Second Method of Construction of Rational Expansions for Transfer Functions.

In section 2.2 we developed a method of construction of rational expansion of F(s), a method which is based on the stair-like approach of the continuous part of the function $\phi^{(+)}_c(-c_o,\lambda)$. The objective of the present section is to develop another method of construction of rational expansion, a method which is based on a somewhat different trend of almost self-suggestive ideas. This new method renders simple results and it finds application in the cases where the function $U(\gamma_o,\lambda)$ vanishes or becomes negligible beyond a certain finite value of λ.

For simplicity in the explanation we will consider a compound contour Γ_o as follows. It is formed by the vertical line above γ_o, $\gamma_o = -c_o$, plus a finite number of semicircles of radius ρ surrounding the singular and critical points of F(s) which lie along Γ_o. (See fig. 1(2.4).)

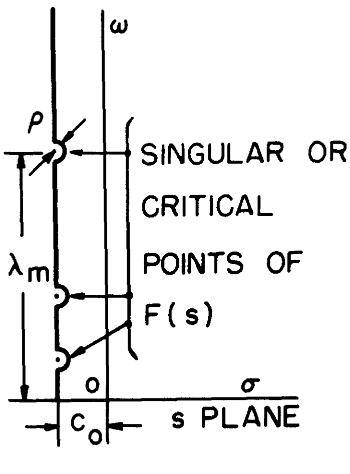

Fig. 1 (2.4)

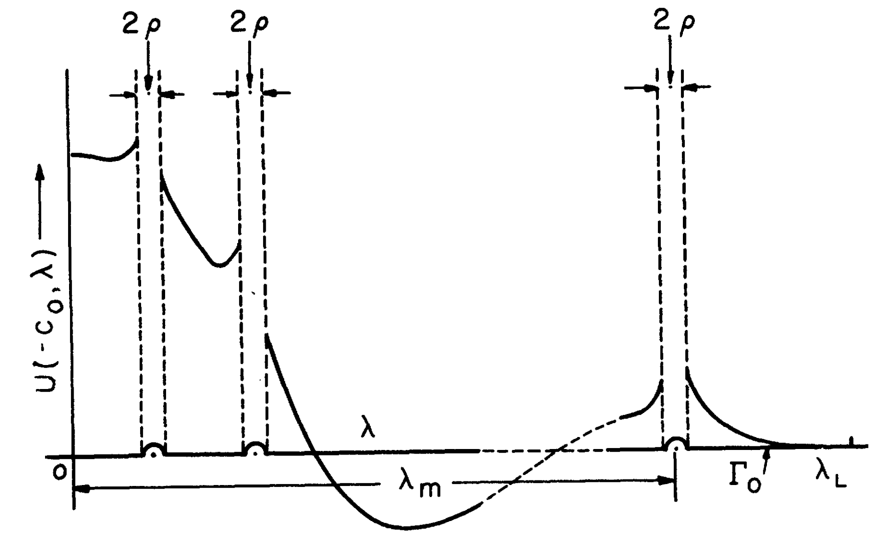

Fig. 2 (2.4)

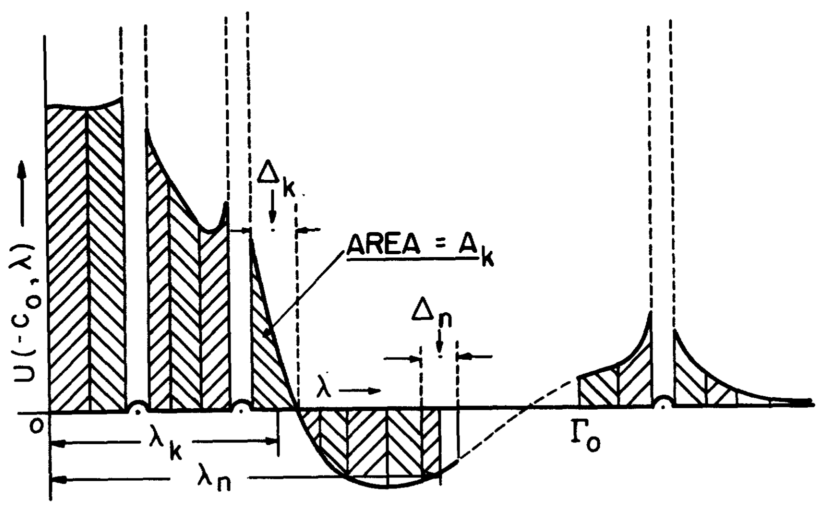

Fig. 3 (2.4)

Along the nonexceptional points of Γ_o, the function $U(\gamma_o,\lambda)$ is uniformly continuous and bounded. Therefore, for the nonexceptional points of Γ_o we can set

$$d\phi(\gamma_o,\lambda) = U(\gamma_o,\lambda)\ d\lambda \qquad 1(2.4)$$

In the semicircles around the exceptional points, the function $U(\gamma,\lambda)$ $(\gamma \neq \gamma_o)$ is still bounded and continuous for $\rho > 0$. But when $\rho \to 0$, we cannot assume the boundedness, the uniform continuity of $U(\gamma,\lambda)$, and the validity of 1(2.4) at the limit $\rho = 0$. For this reason, we shall study this situation in two parts. We shall consider the integral representation and the corresponding expansion: first, along the contour Γ_o itself; and second, around the small semicircles at the exceptional points.

Let us assume that the function $\phi(-c_o,\lambda)$ is of bounded variation on Γ_o. This implies that

$$\phi(-c_o,\infty) = \int_{\Gamma_o} |U(-c_o,\mu)| \, d\mu < \infty \qquad 2(2.4)$$

Figure 2(2.4) illustrates a possible graph of $U(-c_o,\lambda)$. Now we proceed as follows: Let $\left\{\Delta_k\right\}_0^n$ be a sequence of n finite disjointed open intervals. The lengths of the intervals are small and fixed but not necessarily equal. The union of Δ_k covers a segment of Γ_o, from $\lambda = 0$ to $\lambda = \lambda_n$, as indicated in Fig. 3(2.4). Note that λ_n does not necessarily cover the complete Γ_o contour but it will do so if we allow n to approach infinity, keeping each interval Δ_k, $k = 1,2, \ldots, n$ of finite length. (Later we may assume that $\Delta_k \to 0$ when $n \to \infty$, but this assumption is not now postulated.) Associated with the subdivision above we introduce the sequence of functions

$$F_n(s) = \sum_{\nu=1}^{\nu=m} \frac{2}{\pi} \int_{\lambda_k \to \curvearrowright_{\rho \to 0}} \left\{ \lambda d\gamma + (s-\gamma) \, d\lambda \right\} \frac{U(\gamma,\lambda)}{(s-\gamma)^2 + \lambda^2} + M_n(s)$$

where

$$M_n(s) = \frac{2(s + c_o)}{\pi} \sum_{k=1}^{k=n} \int_{\Delta_k} \frac{U(-c_o,\lambda) \, d\lambda}{(s + c_o)^2 + \lambda^2} \qquad 3(2.4)$$

The reader must take due note that the first summation sign includes the contribution of all the m exceptional points, even if they are not covered by λ_n. Our first preoccupation is a study of the sequence of functions $\mathcal{M}_n(s)$, which represent the contribution of the straight line segment of Γ_o from 0 to λ_n.

Let A_k be equal to $2/\pi$ times the area under $U(-c_o,\lambda)$ corresponding to the subinterval Δ_k. Now we make the first basic assumption. Let Δ_k be so small (but not necessarily tending to zero) that the denominator $(s-\gamma)^2 + \lambda^2$ suffers small change as λ varies inside Δ_k, so that we can set

$$(s + c_o)^2 + \lambda^2 \approx (s + c_o)^2 + \lambda_k^2 \qquad \lambda_k < \lambda < \lambda_{k+1} \tag{4(2.4)}$$

where λ_k is a point inside Δ_k. (See fig. 3(2.4).)

The reader may note that 4(2.4) holds good for $|c_o|$ large and mainly for $|s|$ large. Introducing in 3(2.4) the assumption 4(2.4), we get a new sequence

$$\left.\begin{aligned} \mathcal{M}_n(s) &= \frac{2(s + c_o)}{\pi} \sum_{k=1}^{k=n} \frac{A_k}{(s + c_o)^2 + \lambda_k^2} \\ A_k &= \int_{\Delta_k} U(-c_o,\lambda)\, d\lambda \end{aligned}\right\} \tag{5(2.4)}$$

Now let $n \to \infty$. The convergence of this series toward a limiting function, say $\lim \mathcal{M}_n(s) \to \mathcal{M}(s)$ as $n \to \infty$, is guaranteed by the convergence of

$$\sum_0^\infty |A_k| \leqslant \mu \int_{\Delta_k} |U(-c_o,\lambda)|\, d\lambda < \infty$$

as postulated in 2(2.4).

The following remarks are pertinent in regard to the expansion 5(2.4):

(a) Because of the approximated assumption 4(2.4) for finite Δ_k, it becomes clear that the limits

$$\left.\begin{aligned} \lim M_n(s) &= M(s) \\ \text{and} \qquad \lim \mathcal{M}_n(s) &= \mathcal{M}(s) \end{aligned}\right\} \text{ as } n \to \infty$$

are not necessarily equal. However, for $|c_o|$ relatively large, and particularly for $|s|$ large, the expression 4(2.4) is almost an equality, and then we can get

$$\mathcal{M}(s) \approx M(s) \quad \text{asymptotically} \tag{6(2.4)}$$

(b) Let us assume that the set $\left\{\Delta_k\right\}_0^\infty$ of disjointed open intervals covers the complete Γ_o contour. Now let us allow that each $\Delta_k \to 0$. Then

$$\lim \mathcal{M}(s) = M(s), \ k = 1,2,\ldots\infty \text{ as } \Delta_k \to 0$$

Under the assumption $\Delta_k \to 0$, the two types of expansion 5(2.4) and 3(2.1) coincide.

(c) The expansion 5(2.4) is particularly useful when the function $U(-c_o,\lambda)$ vanishes rapidly beyond some finite value of λ, as λ_L in Fig. 2(2.4), because a large number of intervals is not needed for obtaining a good degree of approximation.

(d) The exact position of the end of λ_k, inside Δ_k, is almost immaterial when Δ_k is small. Its exact position is chosen in order to minimize the error committed by the approximation. Sometimes it is convenient, but not necessary, to end λ_k at the center of gravity of the respective element of area inclosed in Δ_k. This facilitates the finding of the zeros of $\mathcal{M}_n(s)$, as mentioned before.

The separation of $\mathcal{M}_n(s)$ in its (p,r) components is done with ease. Let $A_k^{(+)}$ and $A_k^{(-)}$ be, respectively, the coefficients coming from the positive and negative parts of $U(-c_o,\lambda)$. We get

$$\left.\begin{aligned} \mathcal{M}_{(a),n}(s) &= \frac{2(s+c_o)}{\pi} \sum_k \frac{A_k^{(+)}}{(s+c_o)^2+\lambda_k^2} \\ \mathcal{M}_{(b),n}(s) &= \frac{2(s+c_o)}{\pi} \sum_k \frac{A_k^{(-)}}{(s+c_o)^2+\lambda_k^2} \end{aligned}\right\} \qquad 7(2.4)$$

The time response functions associated with the expansion 5(2.4) can be computed directly from the integral representation 3(1.3) by following a similar procedure to that given above. We get

$$\mathcal{L}^{-1}\mathcal{M}_n(s) = e^{-c_o t} \sum_{k=1}^{k=n} A_k \cos \lambda_k t \qquad 8(2.4)$$

the convergence of which is guaranteed by the condition of 2(2.4).

We will close this section by saying a few words regarding the integration of the first summation term in 3(2.4), which represents the contribution of the singularities and critical points of F(s). We have shown before that for simple poles of positive residue, the contribution of the integral around the poles is given simply by

$$\sum_{\nu} \frac{J_\nu(s + c_o)}{(s + c_o)^2 + \lambda_k^2}$$

The computation of the integral for other types of singularities will be discussed later. Our main interest in this section is to develop a method of rational expansion corresponding to the continuous part of the function $U(-c_o,\lambda)$. This is not a limitation of our explanations, because in practice we can remove from F(s) the function containing the singularities closest to the imaginary axis, say at c_o. The remaining function, say $F_r(s)$, has a longer (negative) abscissa of convergence, say c_{or}. The function $U_r(-c_o,\lambda)$ is consequently continuous on Γ_o, $\gamma_o = -c_o$, and we can proceed with the integration as indicated before.

2.5 Example.

We wish to illustrate with a particular example what may happen if we proceed with the construction of a rational expansion of the type 5(2.4) in violation of the assumption 2(2.4). We will show that we may end up with a limiting function $\mathcal{M}(s)$ which is completely different from the original one. Consider the function e^{-s}. Here $c_o = 0$ ($s = \infty$ is an essential singularity). Along Γ_o, $\gamma_o = 0$ we have

$$U(0,\lambda) = \cos \lambda \qquad 1(2.5)$$

Evidently

$$\int_0^\infty |U(0,\lambda)| \, d\lambda = \infty; \text{ (violation)}$$

Let us proceed with the expansion as follows: Let us take, for simplicity in the results, $\Delta_o = \pi/2$; $\Delta_k = k\pi$, $k = 1,2, \ldots n$. By direct computation we get $A_o = 1$, $A_k = 2(-1)^k$, $k = 1,2, \ldots n$. Let us set $\lambda_o = 0$, $\lambda_k = \pi k$ (center of each interval). We get

$$\mathcal{M}_n(s) = \frac{2}{\pi}\left\{\frac{1}{s} + 2s \sum_{k=1}^{k=n} \frac{\cos k\pi}{s^2 + (k\pi)^2}\right\} \qquad 2(2.5)$$

As a limiting function, we obtain

$$\mathcal{M}(s) = \frac{2}{\pi}\left\{\frac{1}{s} + 2s \sum_{k=1}^{\infty} \frac{\cos k\pi}{s^2 + (k\pi)^2}\right\} = \frac{4}{\pi \sinh s} \qquad 3(2.5)$$

instead of e^{-s}.

The time response is

$$\theta(t) = \mathcal{L}^{-1}\mathcal{M}(s) = \left\{\frac{1}{2} + \sum_{k=1}^{\infty} (-1)^k \cos k\pi t\right\}\frac{4}{\pi} \qquad 4(2.5)$$

which is a divergent series. The series represents a singular function which is zero almost everywhere except at the points t = (1 + 2k), where there is an impulse of area $\pi/2$. A quick glance at 3(2.5) tells us immediately that for large $|s|$ we have

$$\frac{1}{2\sinh(s)} \approx e^{-s}; \quad \text{hence } e^{-s} \approx \frac{\pi}{8}\mathcal{M}(s)$$

as we could expect. This asymptotic behavior tells us that $\theta(t)$ must behave at the beginning of the transient as it does when $\mathcal{L}^{-1} e^{-s}$ = unit impulse at t = 1. Evidently 4(2.5) has an impulse at t = 1, but it also contains impulses at t = 1 + 2k, k = 1,2, . . . n, which are not present in $\mathcal{L}^{-1} e^{-s}$.

It is important to point out that this unpleasant phenomenon of convergence is not a characteristic of this particular type of expansion 5(2.4) when 2(2.4) is violated. It may appear when arbitrary rational methods of expansion are used, even when the certain conditions of convergence are completely satisfied in the frequency domain.

2.6 The Network Representation.

Before we proceed to evaluate the measures of approximation committed in both the frequency and time domain when the expansions above are used, we must produce a simple network interpretation of the previous result. This network interpretation will help us to understand the electrical meaning of some mathematical steps which are followed in the evaluation of error. We will use a lattice structure in a formal way, because this structure produces a simple interpretation of our results, but the actual network construction can be followed with other pertinent structures.

The function F(s) is represented by the transfer function of the lattic structure given in Fig. 1(2.6). The function f(t) represents the output response at the terminal 2, 2' of this network when it is excited at the terminals 1, 1' by a unit impulse. An analogous situation occurs for the network composed of the functions $F_{(a),n}(s)$ and $F_{(b),n}(s)$.

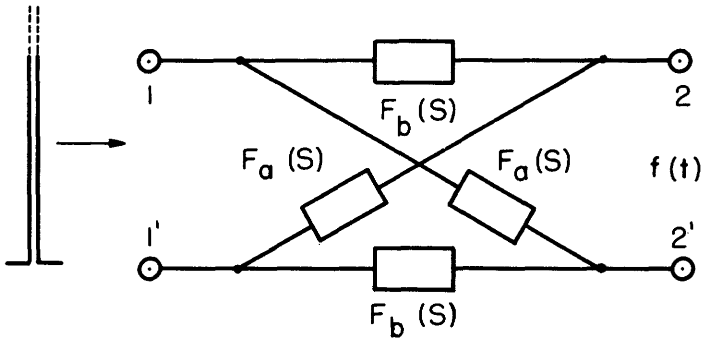

Fig. 1 (2.6)

2.7 The Error in the Frequency Domain.

We will now study the errors which are committed both in frequency and time domains when we substitute the structure corresponding to F(s), by the structure corresponding to $F_n(s)$. This general study will be directed afterwards toward setting the questions of tolerances, modes of convergence, and the like.

In the present section we will limit ourselves particularly to a detailed study of the evaluation and meaning of the errors which are committed when we substitute the function $F_{(a)}(s)$ by $F_{(a),n}(s)$. A similar treatment can be applied to the errors corresponding to the substitution of $F_{(b)}(s)$ by $F_{(b),n}(s)$; therefore, the last case is not explicitly developed here.

Suppose that a transfer function F(s) is given. The corresponding abscissa of convergence c_o, therefore, satisfies the condition $c_o \leqslant 0$ on account of the transfer character of F(s). Let us set a contour Γ_o so that $c_o < (-\gamma_o) < 0$. If $\gamma_o = -c_o$, then the contour Γ_o will contain at least one singularity of F(s); consequently, the function $U(-\gamma_o,\lambda)$ shows a singular behavior at one point, at least, of Γ_o. The study of errors associated with the segments of Γ_o, where the function U is continuous, becomes a remarkably simple and illuminating process. The evaluation of the errors which come from the singular points of F(s) on Γ_o may produce certain difficulties in the study. We will start our study by considering first that U is continuous along Γ_o. This is not a serious limitation because:

(a) If $c_o \neq 0$ then one can select Γ_o by making γ_o as follows: $c_o < (-\gamma_o) < 0$ in which case U is always continuous on Γ_o.

(b) If $c_o = -\gamma_o$, then we may assume that $F_{(a),n}(s)$ or $F_{(b),n}(s)$ already contains the singularities of F(s) along Γ_o, as in the expansion 3(2.1), where F(s) may have a set of m simple poles on Γ_o.

(c) We can extract from F(s), when possible, the singularities on Γ_o and construct a new function, H(s), whose singularities all remain on or to the left of c_o. This case is therefore equivalent to (a).

When these possibilities cannot be exploited or when it is not convenient to use them, then we can use some tricks, suggested from the cases above, which facilitate the study of errors.

Let us first consider the function $\phi_c^{(+)}(-c_o,\lambda)$ which is constructed from the function F(s) as indicated in Part I. The function $\phi_c^{(+)}(-c_o,\lambda)$ defines the (p,r) function $M_{(a)}(s)$ by means of the integral

$$M_{(a)}(s) = \frac{2(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\phi_c^{(+)}(-c_o,\lambda)}{(s + c_o)^2 + \lambda^2} \qquad 1(2.7)$$

Now we construct the stair-like function $\phi_{c,n}^{(+)}(-c_o,\lambda)$ which approximates from below the function $\phi_c^{(+)}(-c_o,\lambda)$. (See, for example, fig. 1(2.7).) The stair function $\phi_{c,n}^{(+)}(-c_o,\lambda)$ defines the (p,r) and, at the same time, a sequence of (p,r) rational functions

$$M_{(a),n}(s) = \frac{2(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\phi_{c,n}^{(+)}(-c_o,\lambda)}{(s + c_o)^2 + \lambda^2} = \frac{2(s + c_o)}{\pi} \sum_{k=1}^{k=n} \frac{a_k}{(s + c_o)^2 + \lambda_k^2} \qquad 2(2.7)$$

We may immediately commit ourselves to use the difference $\left[M_{(a)}(s) - M_{(a),n}(s)\right]$ as a measure of the error which is obtained when we use the function $M_{(a),n}(s)$ instead of $M_{(a)}(s)$. The following consideration tells us that such an evaluation may be meaningless because:

(a) At the points $s = -c_o \pm i\lambda_k$, the function $M_{(a)}(s)$ is regular and has finite and continuous real and imaginary parts, while the function $M_{(a),n}(s)$ becomes infinite. Hence the error there is infinitely large. Moreover, at the zeros of $M_{(a),n}(s)$ (See 11(2.2).) the error is equal to the function $M_{(a)}(s)$ itself. The large oscillation of this "difference" function in the vicinity of Γ_o would create an undesirable situation for making any formulation of tolerances if the needed information is in the vicinity of Γ_o.

For values of s far away from the poles of $M_{(a),n}(s)$, the difference $\left[M_{(a)}(s) - M_{(a),n}(s)\right]$ is small and shows slow oscillation; consequently at great distances from Γ_o the difference above becomes quite suitable for settling the question of tolerances.

(b) The convenient goal in the estimation of errors is that we may express the error function in a compact and simple mathematical expression. We will see that the direct use of the integrals 1(2.7) and 2(2.7) may lead to undesirable results in connection with the evaluation of errors. Take

$$\Delta M = \left[M_{(a)}(s) - M_{(a),n}(s)\right] = \frac{2(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\left[\phi_c^{(+)} - \phi_{c,n}^{(+)}\right]}{(s + c_o)^2 + \lambda^2} \qquad 3(2.7)$$

The function $(\phi_c^{(+)} - \phi_{c,n}^{(-)})$, as extracted from Fig. 1(2.7), is plotted in Fig. 2(2.7).

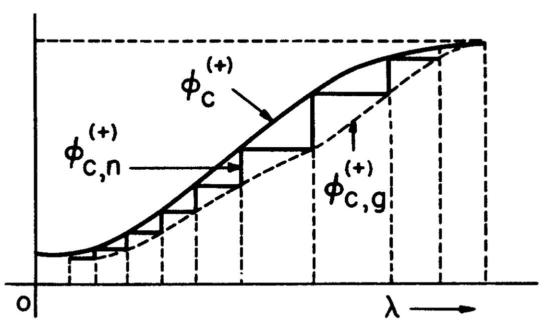

Fig. 1 (2.7)

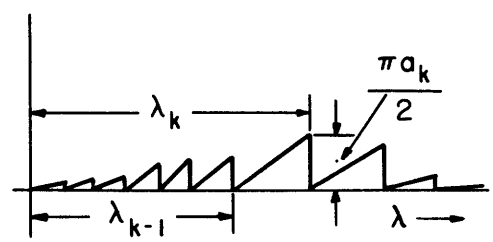

Fig. 2 (2.7)

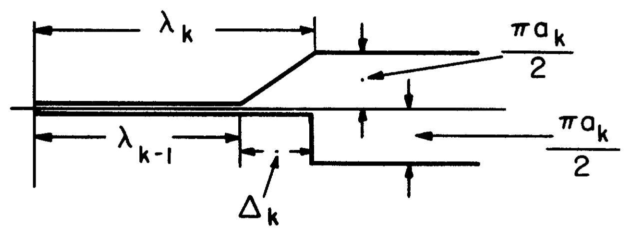

Fig. 3 (2.7)

The integral 3(2.7) can be written

$$\Delta M = \frac{2(s + c_o)}{\pi} \sum_{k=1}^{k=n} \int_{\Delta_k} \frac{d(\phi_c^{(+)} - \phi_{c,n}^{(+)})}{(s + c_o)^2 + \lambda^2} \qquad 4(2.7)$$

where each integral can be taken in its sawtooth-like interval. Each sawtooth, as for example the kth one, can be expressed as the difference of two nonnegative, nondecreasing functions of λ, as shown in Fig. 3(2.7). Since the integrals 4(2.7) must be taken in the Stieltjes sense, then the integration of 4(2.7) leads again, when the summation sign is taken inside the integral, to the original result

$$\Delta M = M_a(s) - (s + c_o) \sum_{k=1}^{k=n} \frac{a_k}{(s + c_o)^2 + \lambda_k^2} \qquad 5(2.7)$$

which is not a simple and compact expression of the wanted error.

A simple mathematical artifice will straighten out this situation.

Suppose that instead of approaching the function $\phi_c^{(+)}$ by a stair-like function, we use a continuous smooth function, say $\phi_{c,g}^{(+)}$. Let $M_{(a),g}(s)$ be defined by

$$M_{(a),g}(s) = \frac{2(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\phi_{c,g}^{(+)}}{(s + c_o)^2 + \lambda^2} \qquad 6(2.7)$$

The function $\phi_{c,g}^{(+)}$ can be constructed, for example, by the smooth curve which joins all the lower corners of the function $\phi_{c,n}^{(+)}$. (See dotted curve in fig. 1(2.7).)

Now, if one increases n without limit, the curve $\phi_{c,g}^{(+)}(-c_o, \lambda)$ approaches the curve $\phi_c^{(+)}$ uniformly. Hence, the function $M_{(a),g}(s)$ tends uniformly toward $M_{(a)}(s)$.

Figure 1(2.7) shows that the function $\phi_{c,n}^{(+)}$ approaches simultaneously both functions $\phi_c^{(+)}$ and $\phi_{c,g}^{(+)}$. Hence, it can be shown that both functions $M_{(a)}(s)$ and $M_{(a),g}(s)$ are simultaneously approximated by the same rational function

$$M_{(a),n}(s) = (s + c_o) \sum_{k=1}^{k=n} \frac{a_k}{(s + c_o)^2 + \lambda_k^2}$$

Consequently

$$M_{(a)}(s) - M_{(a),g}(s) = \left\{ M_{(a)}(s) - M_{(a),n}(s) \right\} + \left\{ M_{(a),n}(s) - M_{(a),g}(s) \right\}$$

$$= \frac{2(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\left[\delta\phi_c^{(+)}\right]}{(s + c_o)^2 + \lambda^2} \qquad 7(2.7)$$

where $\delta\phi_c^{(+)}$ = "variation" of $\phi_c^{(+)} = \phi_c^{(+)} - \phi_{c,g}^{(+)}$.

The proximity of the graphs of $\phi_c^{(+)}$ and $\phi_{c,g}^{(+)}$ allows us to write down an expression to estimate the wanted error. We can assume, due to this proximity, that the function $M_{(a),n}(s)$ approaches $M_{(a)}(s)$ and $M_{(a),g}(s)$ equally with the same relative error. Hence, the following expression is a convenient measure of the error committed.

$$\left[M_{(a)}(s) - M_{(a),n}(s)\right] \approx \frac{(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\left[\delta\phi_c^{(+)}\right]}{(s + c_o)^2 + \lambda^2} \qquad 8(2.7)$$

Formula 8(2.7) leads to results which are very close to the actual ones for values of s away from the immediate vicinity of Γ_o. In the vicinity of Γ_o, the left-hand side of the expression 8(2.7) has a magnitude which suffers very rapid oscillations. The integral does not represent such oscillation but still has a very definite meaning, for it can be shown that the integral still measures the error between the function $M_{(a)}(s)$ and a "sort of average" value of $M_{(a),n}(s)$ in the vicinity of Γ_o. The reader must recall that in the vicinity of Γ_o the average value of the real part, for example, is given by 21(2.2) as

$$U_a \approx \frac{\pi}{2} \frac{a_{k+1} - a_k}{\lambda_{k+1} - \lambda_k} \qquad 9(2.7)$$

Consequently, formula 8(2.7) is quite suitable for synthesis purposes.

2.8 Examples.

The question of errors is continued in this section. In particular, our attention will be directed to the interpretation and manipulation of 8(2.7). The corresponding discussion is primarily intended to illustrate, on an elementary basis, how we can establish a simple (but not necessarily final) criterion relating the accepted errors and the interval decomposition and required number of steps of the function $\phi_{c,n}^{(+)}$. Particular examples will be used, from which a generalization can be visualized with ease.

Example I. Suppose that $\phi_c^{(+)}$ has a graph shape as shown in Fig. 1(2.8). For simplicity in the explanation we have taken a simple curve, but the procedure can be equally applied to more elaborate curves. Let us assume that a curve $\phi_{c,g}^{(+)}$ has already been found in the vicinity of $\phi_c^{(+)}$, such that the committed errors are acceptable. Two simple consequences follow with ease.

(a) If one selects a point, say at $-c_o + i\lambda_k$, as the position of a pole in the expansion 2(2.7), then the position and the number n of poles and residues are well determined, as is graphically illustrated in Fig. 1(2.8).

(b) Let us assume that the curve $\phi_{c,g}^{(+)}$ is slightly modified around $\phi_{c,g}^{(+)}$ itself. The integral 8(2.7) clearly shows that the error committed by using this new slightly deviated curve remains practically the same.

(c) Let us construct a deviated curve $\phi_{c,g}^{(+)}$ such that its vertical distance from the curve $\phi_c^{(+)}$ remains constant in certain nonhorizontal parts of $\phi_c^{(+)}$. (See fig. 2(2.8).)

Figure 3(2.8) shows the graph of $\delta\phi_c^{(+)}$ vs λ for the varied curve $\phi_{c,g}^{(+)}$. The contribution from the interval Δ_{hj}, where $\delta\phi_c^{(+)}$ is constant, is evidently equal to zero. The contribution corresponding to intervals $(0,\lambda_k)$ and (λ_j,λ_L) can be computed in an approximate way by supposing an almost linear change of $\delta\phi_c^{(+)}$.

Hence

$$\left[M_{(a)}(s) - M_{(a),n}(s)\right] = \frac{b}{\pi}\left\{\frac{1}{\lambda_h}\arctan\frac{\lambda_h}{s+c_o} + \frac{1}{\lambda_L-\lambda_j}\left[\arctan\frac{\lambda_L}{(s+c_o)} - \arctan\frac{\lambda_j}{(s+c_o)}\right]\right\} \quad 1(2.8)$$

A detailed discussion of 1(2.8) is omitted, due to lack of space. We confine our attention to the value of this expression for relatively large values of $|s + c_o|$.

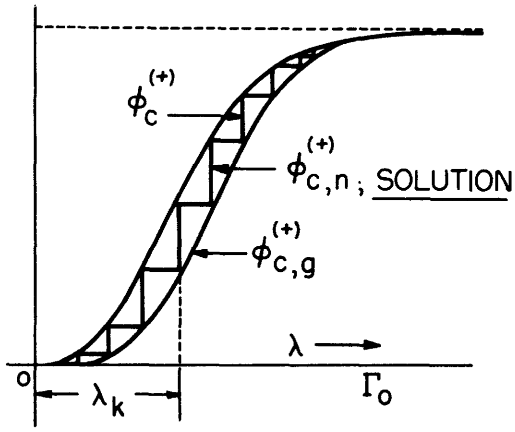

Fig. 1 (2.8)

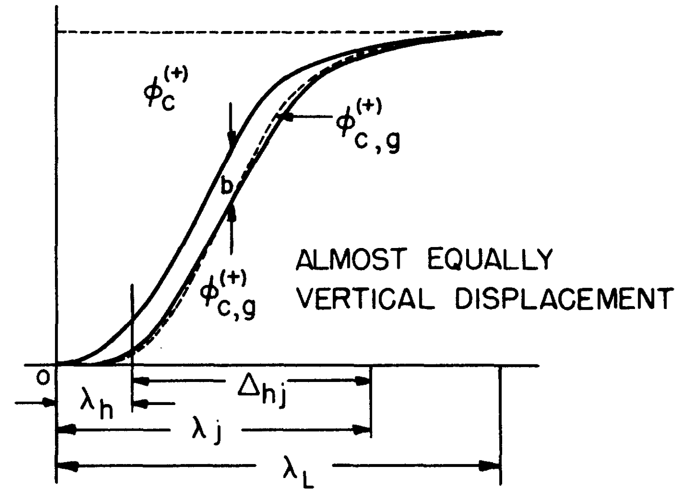

Fig. 2 (2.8)

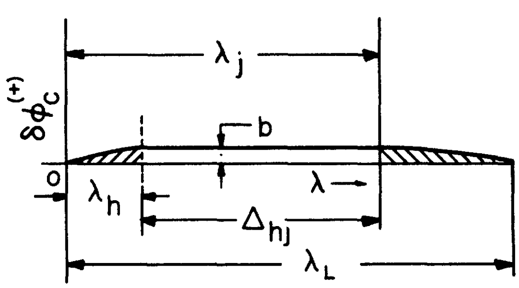

Fig. 3 (2.8)

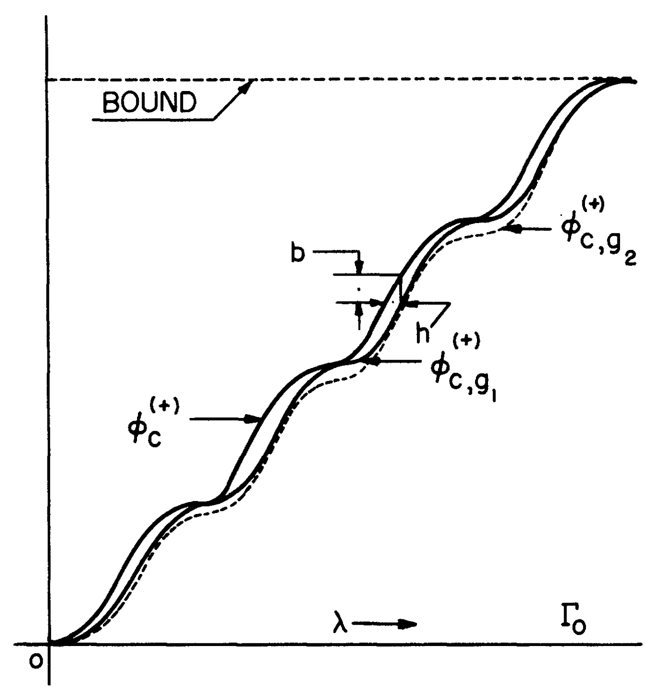

Fig. 4 (2.8)

Then, we obtain the simple expression

$$\left[M_{(a)}(s) - M_{(a),n}(s)\right] \approx \frac{2b}{\pi} \frac{1}{(s + c_o)} \qquad 2(2.8)$$

(d) It is not hard to generalize this particular example to the case in which the function $\phi_c^{(+)}$ has a shape such as is indicated in Fig. 4(2.8). Let N be the number of points where the function $\phi_c^{(+)}$ has a horizontal tangent. It can be shown, as graphically indicated in Fig. 4(2.8), that magnitude of the error committed by approximating the curve $\phi_c^{(+)}$ by a curve $\phi_{c,g}^{(+)}$, whose maximum vertical distance from $\phi_c^{(+)}$ is equal to b, is smaller than or, at most, equal to

$$|M_{(a)}(s) - M_{(a),n}(s)| \leqslant \frac{2b}{\pi} \frac{1}{(s + c_o)}; \qquad |s + c_o| \text{ large} \qquad 3(2.8)$$

(e) The property of section (d) establishes a very simple tolerance equation. For suppose that we want to find the rational expansion $M_{(a),n}(s)$ corresponding to the given function $M_{(a)}(s)$ in such a way that the error committed is less than ϵ_o, (ϵ_o is a small positive quantity) at the points of the s plane whose distance to Γ_o is equal to or larger than R; then the expression 3(2.8) yields the upper bound of b as

$$b = \frac{\epsilon_o R \pi}{2} \qquad 4(2.8)$$

With this value of b, we construct a curve $\phi_{c,g}^{(+)}$, like $\phi_{c,g_2}^{(+)}$ in Fig. 4(2.8), whose vertical distance from $\phi_c^{(+)}$ is constantly equal to b except at $\lambda = 0$ and at the upper boundary of $\phi_c^{(+)}$.

Now, if we fix the position of one, and only one, pole, as indicated in section 2.8(a), then we can determine uniquely a rational function $M_{(a),n}(s)$ which satisfies the required tolerance condition.

<u>Example II.</u> The curve $\phi_{c,g_1}^{(+)}$ in Fig. 4(2.8), clearly suggests that we could take this curve as the horizontal displacement of the curve $\phi_c^{(+)}$ itself. Let h be a small number which represents the horizontal displacement. Then, we can set $\phi_{c,g}^{(+)} = \phi_c^{(+)}\left(-c_o, (\lambda - h)\right)$. This selection leads to a simple interpretation of the meaning of the error term.

If the quantity h is small, we can write

$$\delta\phi_c^{(+)} = \phi_c^{(+)} - \phi_{c,g}^{(+)} = h \frac{d}{d\lambda} \phi_c^{(+)} + \ldots$$

By keeping the first term of the expansion and by using the integral expression for $\phi_c^{(+)}$

$$\phi_c^{(+)} = \int_0^{\lambda} U_c^{(+)}(-c_o,\lambda)\ d\lambda \tag{5(2.8)}$$

where $U_c^{(+)}$ stands for the continuous part of $U^{(+)}$, then one gets

$$\delta\phi_c^{(+)} = h\ U_c^{(+)}(-c_o,\lambda) \tag{6(2.8)}$$

and the error is then

$$\left[M_{(a)}(s) - M_{(a),n}(s)\right] = h\ \frac{(s + c_o)}{\pi} \int_{\Gamma_o} \frac{d\left[U_c^{(+)}\right]}{(s + c_o)^2 + \lambda^2}$$

$$= h\ \frac{(s + c_o)}{\pi} \int_{\Gamma_o} \frac{(U_c^{(+)})'\ d\lambda}{(s + c_o)^2 + \lambda^2} \tag{7(2.8)}$$

where $(U_c^{(+)})'$ stands for the derivative with respect to λ. Equation 7(2.8) indicates:

(a) the error function is proportional to the displacement;

(b) the error function is expressed intrinsically in terms of the real part of the same function $M_{(a)}$ itself;

(c) due to the boundedness of $U_c^{(+)}$ the error term is necessarily a transfer function.

A comparison of the curves $\phi_{c,g_2}^{(+)}$ (which correspond to a vertical displacement b from $\phi_c^{(+)}$) and $\phi_{c,g}^{(+)}$ (which corresponds to a horizontal displacement of h from $\phi_c^{(+)}$) immediately reveals that the error committed by a horizontal displacement is smaller than the one corresponding to the vertical displacement. Let us substitute in 7(2.8) for $(U_c^{(+)})'$ its corresponding maximum value. This last value is bounded. Then the integral 7(2.8) leads to a bound similar to 1(2.8) or 2(2.8), when one sets $h(U_c^{(+)})'_{max} = b$, where b is the corresponding maximum vertical distance between curves $\phi_c^{(+)}$ and $\phi_{c,g}^{(+)}$.

2.9 The Network Representation of the Errors.

In this section we will interpret the errors of the last section in terms of its representative network. This network will be used to increase the degree of ap-

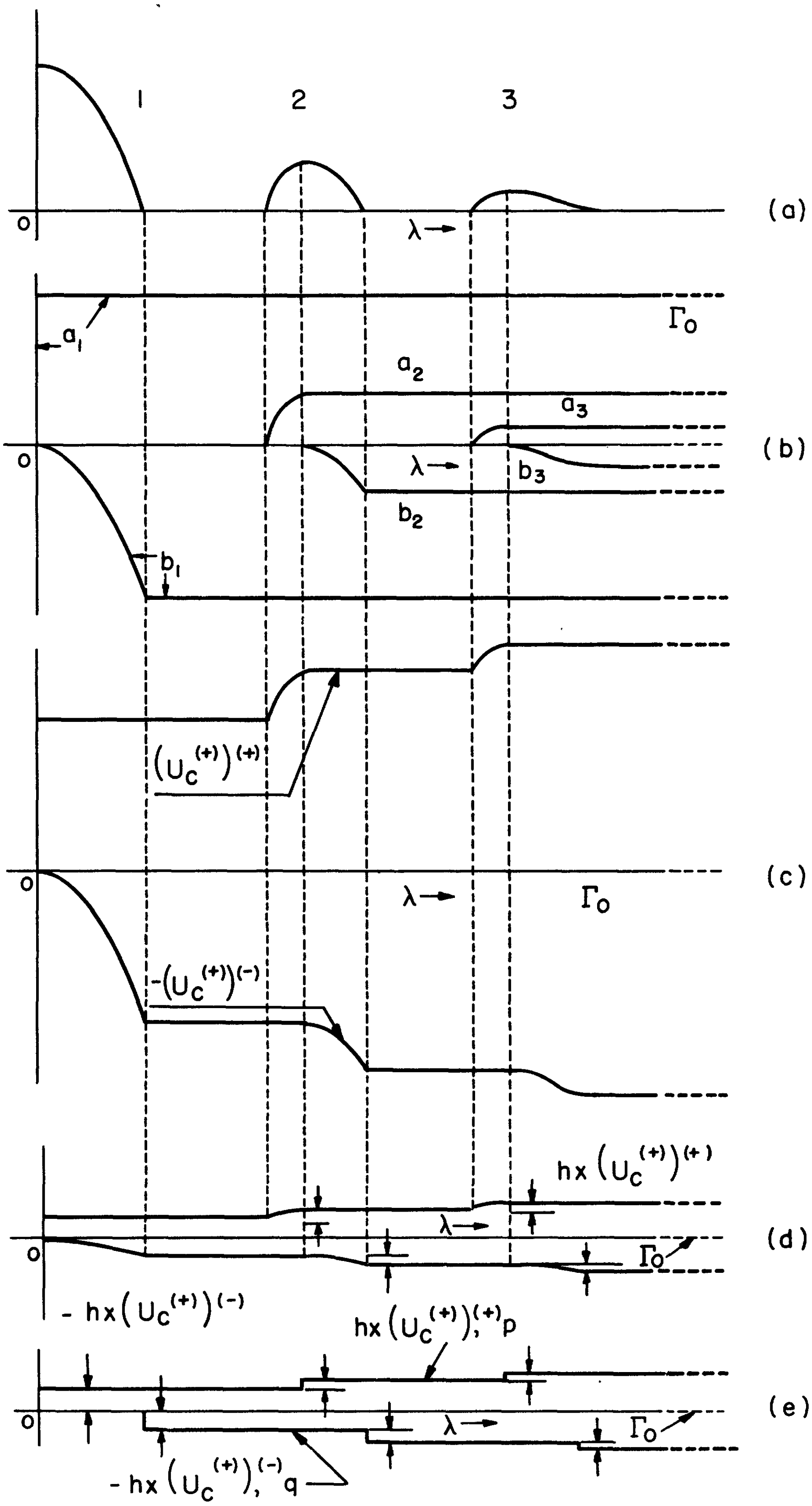
1
2
3
(a)
0
λ→
a1
Γ0
a2
a3
(b)
0
λ→
b3
b2
b1
(Uc(+))(+)
(c)
0
λ→
Γ0
-(Uc(+))(-)
hx (Uc(+))(+)
λ→
(d)
0
Γ0
- hx(Uc(+))(-)
hx (Uc(+)),(+)p
(e)
0
λ→
Γ0
- hx (Uc(+)),(-)q

Fig. 1 (2.9)

proximation of the network which was first obtained from the original rational function approximation. (See section 2.6, and fig. 1(2.6).)

Consider the function $U_c^{(+)}$ in 7(2.8). This quantity is supposed to be known, since from the original transfer function F(s) we can find $U^{(+)}(-c_o,\lambda)$ and extract from the last its continuous part $U_c^{(+)}$.

The error 7(2.8) represents a transfer function because of the boundedness of $U_c^{(+)}$. In order to construct the equivalent network we must find the (p,r) components of 7(2.8). A graphical procedure illustrates quite well the meaning of the required analytical steps.

Figure 1(2.9)a shows a sort of typical graph of $U_c^{(+)}$.

The procedure is as follows:

(a) Take one lobe of $U_c^{(+)}$. Separate the intervals in which the lobe shows, respectively, a monotonic increase and monotonic decrease.

(b) Construct from each lobe the function a_k and b_k, as is clearly shown in Fig. 1(2.9)b.

(c) Let us call $(U_c^{(+)})^{(+)}$ the nonnegative, nondecreasing function of bounded variation which results from the summation of the a_k functions. And let us call $(U_c^{(+)})^{(-)}$ the nonnegative, nondecreasing function of bounded variation which results from the summation, taken with negative sign, of the b_k function. See Fig. 1(2.9)c.

(d) Multiply both $(U_c^{(+)})^{(+)}$ and $(U_c^{(+)})^{(-)}$ by h. The results are, respectively, the distribution functions which generate, by means of the integral 2(2.8), the (p,r) components of the transfer function associated with the error given by this integral 2(2.8). (See fig. 1(2.9)d.)

Let us denote by $H_{(a)}(s)$ and $H_{(b)}(s)$, respectively, the corresponding (p,r) functions constructed with $h(U_c^{(+)})^{(+)}$ and $h(U_c^{(+)})^{(-)}$. These functions represent, in an approximate way, (p,r) components of the error committed in the approximation of the function $M_{(a)}(s)$.

By following almost identical lines, we can find a similar integral expression, like 2(2.8), which represents the error committed in the approximation of $M_{(b)}(s)$ by $M_{(b),n}(s)$. From this error integral representation we can construct the (p,r) components corresponding to the error $\left[M_{(b)}(s) - M_{(b),n}(s) \right]$ by just following the procedure indicated above.

Let us denote by $G_{(a)}(s)$ and $G_{(b)}(s)$, respectively, the corresponding (p,r) functions constructed with $h(U_c^{(-)})^{(+)}$ and $h(U_c^{(-)})^{(-)}$. The set of (p,r) functions

$H_{(a)}(s)$, $H_{(b)}(s)$, $G_{(a)}(s)$ and $G_{(b)}(s)$ can be used for correcting the original lattice structure corresponding to the function $F_n(s)$ so that the new corrected structure represents F(s) with a larger degree of approximation. The functions $H_{(a)}(s)$ and $H_{(b)}(s)$ can be approximated by a rational expansion by following the methods already explained. For example, in Fig. 1(2.9)e there is the suggestion of approaching the functions $h(U_c^{(+)})^{(+)}$ and $h(U_c^{(+)})^{(-)}$ by a simple stair-like function whose steps are equal to the relative maxima of $U_c^{(+)}$ times h.

Let us call $H_{(a),p}(s)$, $H_{(b),p}(s)$, $G_{(a),q}(s)$, $G_{(b),q}(s)$ the rational approximants, respectively, of the H and G functions above. Due to the small variation of the functions in Fig. 1(2.9)d, one pole is suggested for each region, where $h(U_c^{(+)})^{(+)}$, etc., is nonconstant. The position of each pole can be set, for example, in the vicinity of the center of gravity of the regions of increase. When it happens that a pole of $F_{(a),n}(s)$ or of $F_{(b),n}(s)$ lies in the same vicinity, then we may set the corresponding pole associated with the H and G functions in coincidence with the nearest poles of $F_{(a),n}(s)$, etc., so that it is absorbed in the original poles of the expansion. Figure 2(2.9) shows the corrected lattice structure.

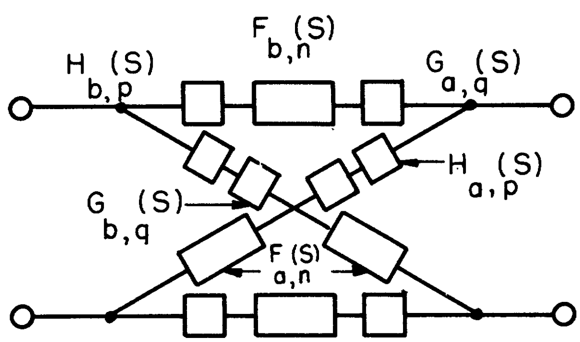

Fig. 2 (2.9)

2.10 Remarks.

The expressions developed above, for the evaluation of the error committed in the rational expansion, all refer to the (p,r) components, $F_{(a)}(s)$ and $F_{(b)}(s)$, of a given transfer function F(s). The methods and procedures given above can be applied to F(s) either by using the previous results corresponding to its (p,r) components or by a direct evaluation by means of an integral expression similar to 8(2.7), when the function U must replace the expression $U_c^{(+)}$, or $U_c^{(-)}$. We do not go into a detailed discussion of the evaluation of the error committed with respect to the function F(s), because the purpose of the present discussion is the presentation of the general ideas followed in the rational expansions presented here.

2.11 The Error in the Time Domain.

The evaluation of the error committed in the time domain by the rational approximation of the corresponding transfer functions can be formally carried out by a similar integral representation. From 5(1.4) of Part I it is evident that the error committed is given by

$$\left[f(t) - f_n(t)\right] = \frac{2e^{-c_o t}}{\pi} \int_{\Gamma_o} \cos \lambda t \; d \left[\delta\phi(-c_o, \lambda)\right] \qquad 1(2.11)$$

where $\delta\phi = \phi - \phi_g$ and represents the variation of the real part of F(s) against the function ϕ_g, which is constructed by means of the corresponding rational expansion of F(s). The construction of ϕ_g is similar to the construction of $\phi^{(+)}_{c,g}$. (See fig. 1(2.7).)

In general, the function $\delta\phi$ attains positive and negative values, since here ϕ is not necessarily a nonnegative, nondecreasing function. For the purpose of illustration let us assume that the graph of $\delta\phi$ is given in Fig. 1(2.11). For simplicity in the presentation we assume that $\delta\phi$ is a continuous function. (The extension of the following consideration is not difficult to perform in the case in which the function ϕ presents jumps.)

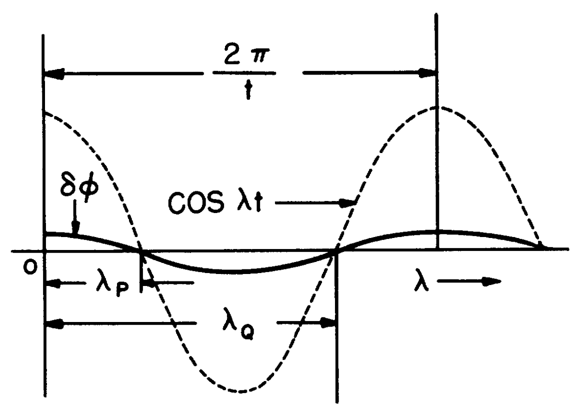

Fig. 1 (2.11)

In order to interpret the meaning of integral 1(2.11) we have also plotted the graph of cos λt. Since the variable of integration is λ, then the period of the cosine function is equal to $2\pi/t$.

A quick glance at the integral 1(2.11) and at Fig. 1(2.11) immediately reveals:

(a) At t = 0 the error is equal to the area under the curve $\delta\phi$.

(b) As the time increases, the error committed changes. From the graph of $\delta\phi$ shown in the figure, it is clear that the error starts decreasing. It reaches approximately a minimum for $t = 2\pi/\lambda_Q$. Then, it increases and attains approximately a maximum, except for the factor $e^{-c_o t}$, at $t = 2\pi/\lambda_P$.

(c) After this last value of time, the error starts decreasing very fast, due to the chopping effect of $\cos \lambda t$ and the attenuation introduced by $e^{-c_o t}$.

The previous explanation immediately suggests that we can select the variation of ϕ in such a way as to minimize the error in the time domain at a certain interval of time, keeping the error below a certain bound, or tolerance, for other values of t.

In spite of the simplicity of the explanation above, the problem of rational expansion of a transfer function F(s) for the purpose of time domain synthesis is a very delicate question which requires careful consideration. Contours of the type Γ_o have, in general, little use in time domain synthesis. A complete study of this situation is outside the scope of this report.

2.12 Summary.

The general discussion of the present notes has been systematically supported by the integral representation along the particular contour Γ_o. This contour selection considerably restricts the ideas and results from the point of view of the general theory of rational expansion of transfer functions. However, this particular selection of contour, in spite of its limitations, illustrates with ease, and renders elementary proof of, some basic ideas involved in the general theory. Similar proofs and situations for a more general type of contour become difficult and involved, even if the final mechanisms of construction are quite similar to those obtained for the particular contour Γ_o. For a more advanced study of this problem the reader is referred to reference 3.

References

1. M. V. Cerrillo: Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology, Jan. 15, 1952

2. E. A. Guillemin: The Mathematics of Circuit Analysis, John Wiley, New York, 1950

3. M. V. Cerrillo: Quarterly Progress Report, Research Laboratory of Electronics, Massachusetts Institute of Technology, April 15, 1952