COMPARATIVE STUDY OF SEVERAL METHODS FOR DETERMINATION OF TRANSIENT RESPONSE TO A UNIT STEP INPUT FROM FREQUENCY RESPONSE CURVES FOR A STABLE SYSTEM

Thesis for the Degree of M. S.

MICHIGAN STATE COLLEGE

Kurt Elliott Utley

1954

This is to certify that the

thesis entitled

"Comparative Study of Several Methods for Determination of Transient Response to a Unit Step Input Frequency Response Curves for a Stable System"

presented by

Kurt Elliott Utley

has been accepted towards fulfillment of the requirements for

M. S. degree in E.E.

Major professor

Date May 25, 1954

COMPARATIVE STUDY OF SEVERAL METHODS FOR DETERMINATION OF TRANSIENT RESPONSE TO A UNIT STEP INPUT FROM FREQUENCY RESPONSE CURVES FOR A STABLE SYSTEM

By

Kurt Elliott Utley

A THESIS

Submitted to the School of Graduate Studies of Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Electrical Engineering

1954

PREFACE

In the synthesis and design of a servomechanism the frequency response curves are of great importance, being an indication of system stability etc. Such curves are readily obtained, as a rule, either from experimental data or by evaluating the analytical expression of the system. However, from the standpoint of safety and a consideration of the system time requirements, the transient response of the system to discontinuous disturbances should also be known. Excessively high peaks are to be avoided as are semi-sustained oscillations. Unfortunately, the transient response is difficult to determine, either analytically or experimentally, hence relationships are sought between the transient response and the more easily obtained frequency response.

It is the purpose of this thesis to present in an organized manner three different methods of finding the transient response to a unit step input from the frequency response for a stable system. A problem will be solved by each of the methods and comparisons made among the three with respect to the advantages and disadvantages of each.

This paper can be used as a guide for applying any of the three methods to a suitable problem.

ACKNOWLEDGEMENTS

The author wishes to express his sincere thanks to Dr. J. A. Strelzoff for suggesting the topic of this thesis and for his assistance in its preparation, and to the Michigan State College Library for its aid in procuring microfilmed material.

T A B L E   O F   C O N T E N T S

GENERAL

SOLUTION OF A PROBLEM

At the present time techniques for the synthesis of closed-loop servomechanism systems or improvement of existing systems are based primarily upon the frequency response of the system: the output, or some function of the output plotted against frequency in amplitude and phase as the sinusoidal input varies over a wide range of frequencies beginning at zero c.p.s. These responses may indicate maximum gain and stability, as well as other characteristics.

However, it is necessary to have a knowledge of the transient response of the system as well, since all closed-loop systems are likely to be subjected to types of discontinuous inputs which cause transients. The simple closing of a switch, the action of a circuit breaker, component failure in an adjacent system or a sudden change of load may cause transient disturbances. Such disturbances are often negligible, but in a lightly damped system oscillatory transients of damaging magnitude may result, an annoying train of oscillations persisting for some time may be initiated, or both conditions may occur. Hence a knowledge of the transient response of the system is important to the designer.

If the input to a specific system is always to be of a certain specialized nature it is better to study the transient response to this typical input, of course, but often the response of the system to arbitrary inputs is determined by comparison with response to a unit step input: a wave form of zero amplitude for t (time) $\leq$ 0, and of unity amplitude for t $\geq$ 0. Thus the response of a system to a unit step can be considered a standard, and the words "transient response", as used in this paper, refer to the response to a unit step unless otherwise specified.

A basic present-day procedure for finding the response (both transient and steady state) of a system to a certain input is to apply operational calculus e.g. the Laplace transform, to the differential equation of the output, provided such an equation is available, transforming from a time domain to a functional domain, to obtain the characteristic equation and find its roots, and then by means of an inverse transformation return the problem again to the time domain in such a manner that the solution to the original differential equation is obtained.[1] However, due to the fact that the desired roots of the characteristic equation are obtained only after the solution of a polynomial of possibly high order, this analytic method for finding the transient response may be exceedingly slow and laborious.

Another means of determining transient response is by the utilization of a transient analyzer, in which inductances, capacitances, resistors and electronic amplifiers are used as analogues of masses, springs and dampers. Unfortunately, this analyzer cannot be applied until much is known about the individual components of the system, and unless the study is to run over an extended period of time, and the analyzer is flexible and well instrumented, the analytic method, though slow, may take less time.[2]

A third method is to find the transient response of the system from the frequency response curves, which are readily obtainable from the analytic expression of the system or from experimental observations.

---

1. Gordon Brown and Donald Campbell, Principles of Servomechanisms, John Wiley & Sons, Inc., New York, 1948, p. 66.

2. Ibid., p. 19.

It is the purpose of this paper to consider three such methods and to make a comparison among them as to ease of understanding, ease of application, limiting conditions, speed, accuracy and limits of application to involved problems.

The methods are:

I. To determine the roots of s in the characteristic equation of the system by graphic means, from the frequency response curves, and to use either graphic or operational calculus procedures to obtain the transient solution, given the roots.[1]

II. To apply an approximate inverse transform to a succession of trapezoidal waves obtained from the frequency response curve of the output $\Theta_o(j\omega)$, yielding an approximate transient response.[2]

III. To approximate the unit step input by a half cycle of a square wave, resolve the square wave into its sinusoidal components, and to obtain the transient response by summing the system frequency response (known from the output over input frequency response curve $\frac{\Theta_o}{\Theta_i}(j\omega)$) to a sufficient number of these component waves.[3]

All three of these methods presuppose a system linear in the test range, and stable over the frequencies involved.

---

1. Walter Evans, "Graphic Analysis of Control Systems", AIEE Transactions, Vol. 67, pps. 547-551.

2. Brown and Campbell, op. cit., pps. 332-365.

3. C. A. A. Wass and E. G. Hayman, "An Approximate Method of Deriving the Transient Response of a Linear System from the Frequency Response", Royal Aircraft Establishment, C. P. 113 Technical Note no. GW. 148, November, 1951.

Method I

Determine the roots of s in the characteristic equation from the frequency response curves by graphic means and use either graphic or operational calculus procedures to obtain the transient solution, given the roots.[1]

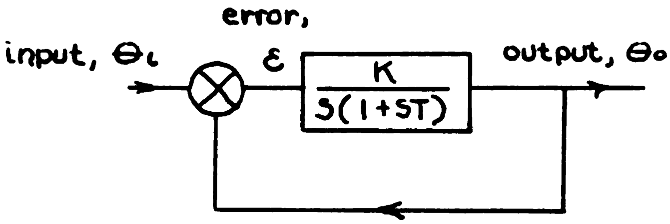
The basic problem in finding the transient response is to determine the roots of the differential equations which correspond to the exponential transient terms dominating the response. The function which describes the system from output to error, $\frac{\theta_o}{\varepsilon}(s)$, or the inverse, $\frac{\varepsilon}{\theta_o}(s)$, is a function of the complex variable s, which has damping (σ) as its real part and frequency (ω) as its imaginary part. The imaginary axis of the s plane corresponds to $\pm j\omega$.[2] The frequency plot of $\frac{\varepsilon}{\theta_o}(j\omega)$ obtained by setting $s = j\omega$ is simply one line of a conformal map, with the roots of s lying at the value of the variable (the point in the s plane, $\sigma_1 + j\omega_1$) which makes the function $\frac{\varepsilon}{\theta_o}(s)$ equal to -1 for a system with unity feed back, or equal to $-K$ for a gain of K in the feedback loop.

That the root lies at a value such that the function $\frac{\varepsilon}{\theta_o}(s)$ is equal to -1 is shown by the following illustration of a simple unity feedback system:

---

1. Walter R. Evans, op. cit., pps. 547-551.

2. Brown and Campbell, op cit., p. 171.



$$\frac{\theta_i}{\theta_o}(s) = \frac{\theta_o + \varepsilon}{\theta_o}(s) = 1 + \frac{\varepsilon}{\theta_o}(s)$$

Fig. 1. Unity Feedback System

$$\frac{\varepsilon}{\theta_o}(s) + 1 = \frac{s(1+sT)}{K} + 1 = \frac{s^2T + s + K}{K} \qquad \frac{\theta_i}{\theta_o}(s)$$

The roots of s in the expression $\frac{s^2T + s + K}{K}$ are obtained by setting the expression equal to zero, or by finding values of s which will make $\frac{\varepsilon}{\theta_o}(s) = -1$ [1], which gives the same effect.

For a general case, Method I goes as follows:

1. Break the feedback loop and apply a sinusoid; make a locus plot of error (error = input in this instance) over output as a function of frequency, either analytically, letting $s = j\omega$ in the $\frac{\varepsilon}{\theta_o}(s)$ expression, or by taking observational data.

2. Considering the vector locus plot to be the base line from which the complex roots of the main damped sinusoidal term can be found, construct an orthogonal lattice of $-\sigma =$ constant and $j\omega =$ constant lines and determine the complex roots of s by noting where on this lattice the point -1 falls.

The justification for this step is easily shown:

---

1. The distinction between $\frac{\varepsilon}{\theta_o}(s)$ and $\frac{\varepsilon}{\theta_o}(j\omega)$ lies in the fact that $\frac{\varepsilon}{\theta_o}(s)$ may lie almost anywhere in the plane, depending on the complex value of $s = -\sigma \pm j\omega$ while $\frac{\varepsilon}{\theta_o}(j\omega)$ is a special case of $\frac{\varepsilon}{\theta_o}(s)$ where $s = j\omega$, and is a single curved line.

Let $s = -\sigma + j\omega$, from which $\frac{\partial s}{\partial \sigma} = -1$, and $\frac{\partial s}{\partial \omega} = j$. $\frac{\varepsilon}{\theta_o}(s)$ is a function of s, which has a particular derivative with respect to s at each point in the region of interest.

$$\frac{\partial}{\partial \omega}\left(\frac{\varepsilon}{\theta_o}(s)\right) = \frac{\partial}{\partial s}\left(\frac{\varepsilon}{\theta_o}(s)\right) \cdot \frac{\partial s}{\partial \omega} = j\frac{\partial}{\partial s}\left(\frac{\varepsilon}{\theta_o}(s)\right) \quad (3)$$

$$\frac{\partial}{\partial \sigma}\left(\frac{\varepsilon}{\theta_o}(s)\right) = \frac{\partial}{\partial s}\left(\frac{\varepsilon}{\theta_o}(s)\right) \cdot \frac{\partial s}{\partial \sigma} = -1\frac{\partial}{\partial s}\left(\frac{\varepsilon}{\theta_o}(s)\right) \quad (4)$$

The change of the function with respect to $-\sigma$ is seen to be displaced 90° ccw. from the change with respect to $j\omega$, which lies along the $\frac{\varepsilon}{\theta_o}(j\omega)$ curve. The change of the function with respect to s is 90° cw. from the $\frac{\varepsilon}{\theta_o}(j\omega)$ curve. For equal changes in $-\sigma$ and $j\omega$, a set of curvilinear squares will be formed with conformal map properties as shown. If the expression of $\frac{\varepsilon}{\theta_o}(s)$ is known, points on this lattice near the -1 point may be plotted by solving the expression using different values of $-\sigma$ and $j\omega$ in $s = -\sigma + j\omega$.

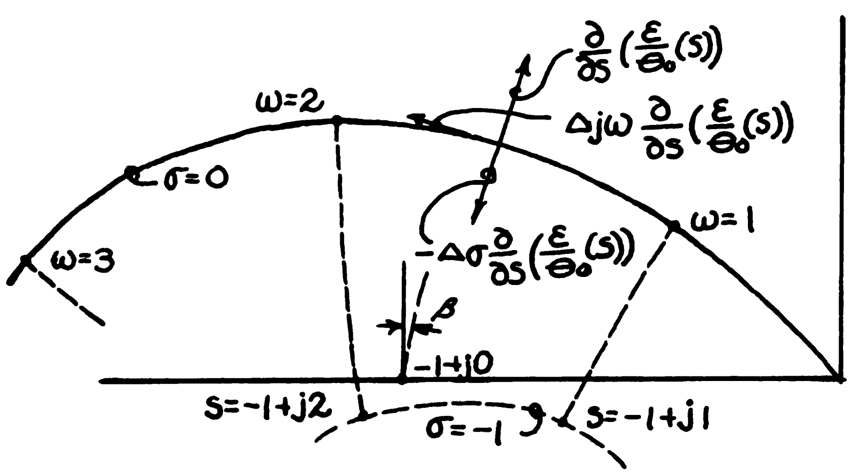


Fig. 2. General $\frac{\varepsilon}{\theta_o}(j\omega)$ locus showing curvilinear squares.

For this illustrative example the root lies at $s = -0.9 + j\,2.8$, and since complex roots occur in conjugate pairs, s is also $= -0.9 - j\,2.8$.

This procedure is valid only for systems linear in the test range, with results applicable to the range. This restriction is necessary since the justification of the curvilinear lattice depends on the derivative being independent of the amplitude of the function, depending only on the nature of the function.

If the $\frac{\varepsilon}{\theta_o}(s)$ expression is not available, and the $\frac{\varepsilon}{\theta_o}(j\omega)$ curve has been experimentally obtained, a sketch of the curvilinear squares may be made quite accurately without knowing points in the lattice, provided the baseline, or the $\frac{\varepsilon}{\theta_o}(j\omega)$ locus is not so strongly curved as to cause undue distortion and overlapping near the -1 point. In general, this method is much more accurate if points on the curvilinear system are carefully plotted by several evaluations of the $\frac{\varepsilon}{\theta_o}(s)$ expression using complex values of s which will cause $\frac{\varepsilon}{\theta_o}(s)$ to fall near the -1 point.

Once the complex roots are known, assuming a single pair is had, they can be divided out of the polynomial expression for $\frac{\theta_i}{\theta_o}(s)$. The simplified expression may then be either solved by straight forward algebraic methods, or perhaps graphically from the original plot, having found the general location of the real roots by inspection of the simplified function. The author of this paper has found that, once the complex conjugate roots have been found, the real roots are most easily obtained through factoring the simplified expression, since plotting the curvilinear squares in the region of $\omega = 0$ is rather critical.

3. Having found the roots, determine the general form[1] and the amplitude of the transient solution for each root by means of operational calculus. The general form of the solution due to the conjugate complex roots $s = -\sigma + j\omega$ is $f_1(t) = K_1 e^{(-\sigma + j\omega)t} + K_2 e^{(-\sigma - j\omega)t}$ (5)

For a real root, $s = \rho$ $\quad f_2(t) = K_3 e^{-\rho t}$ (6)

The amplitude of a transient is given in terms of its root by:

$$A = \left. \frac{1}{s \frac{\partial}{\partial s}\left(\frac{\theta_i}{\theta_o}(s)\right)} \right|_{s = s_1 = -\sigma_1 + j\omega_1} \quad (7)$$

Since $\frac{\varepsilon}{\theta_o}(s)$ differs from $\frac{\theta_i}{\theta_o}(s)$ by only a constant, $\frac{\partial}{\partial s}\left(\frac{\varepsilon}{\theta_o}(s)\right) = \frac{\partial}{\partial s}\left(\frac{\theta_i}{\theta_o}(s)\right)$
Thus, the angle of the derivative can be read directly from the plot (See angle β , Fig. 2) and the magnitude can be found by taking the change in $\frac{\varepsilon}{\theta_o}(s)$ (distance from origin) divided by the change in s (as measured on curvilinear lattice) and averaging over several measurements near the -1 point.

Letting $s = a e^{\alpha}$ $\quad \left(\alpha = \tan^{-1} \frac{-\sigma}{\omega}\right)$ and $\frac{\partial}{\partial s}\left(\frac{\theta_i}{\theta_o}(s)\right) = b e^{\beta}$

the transient solution for conjugate complex roots $s = -\sigma_1 + j\omega_1$ and $s = -\sigma_1 - j\omega_1$ is:

$$f_1(t) = \left. \frac{e^{(-\sigma_1 + j\omega_1)t}}{a_1 e^{\alpha_1} b_1 e^{\beta_1}} \right|_{s = -\sigma_1 + j\omega_1} + \left. \frac{e^{(-\sigma_1 - j\omega_1)t}}{a_2 e^{\alpha_2} b^2 e^{\beta_2}} \right|_{s = -\sigma_2 - j\omega_1} \quad (8)$$

From relationships between the denominators,

$$a_1 = a_2, \quad b_1 = b_2, \quad \alpha_1 = -\alpha_2, \quad \beta_1 = -\beta_2$$

$$f_1(t) = \frac{e^{-\sigma_1 t}}{a_1 b_1}\left(e^{j\omega t - \alpha_1 - \beta_1} + e^{-j\omega t + \alpha_2 + \beta_2}\right) \quad (9a)$$

1. Stanford Goldman, Transformation Calculus and Electrical Transients, Prentice-Hall, Inc., New York, 1949, p. 419.

$$f_1(t) = \frac{e^{-\sigma t}}{a_1 b_1} \cos(\omega t - \alpha_1 - \beta_1) \tag{9b}$$

For real roots, $s = \rho$, the transient solution is:

$$f_2(t) = \frac{e^{-\rho t}}{a_3 e^{\alpha_3} b_3 e^{\beta_3}} \tag{10}$$

Since there is no j term in s, α_3 and β_3 will be 0 or π , and

$$f_2(t) = \frac{\pm e^{-\rho t}}{a_3 b_3} \tag{11}$$

The entire transient solution is then:

$$f(t) = \frac{e^{-\sigma t}}{a_1 b_1} \cos(\omega t - \alpha_1 - \beta_1) + \frac{\pm e^{-\rho t}}{a_3 b_3} \tag{12}$$

Occasionally systems will have two pairs of quadratic roots in about the same frequency range, and the $\frac{\varepsilon}{\theta_o}(j\omega)$ plot will circle the -1 point with the result that the curvilinear lattice built up from one side will overlap that built up from the other side. A series of successive approximations will simplify. Start with a pair of conjugate roots suggested by one lattice and divide the $\frac{\theta_i}{\theta_o}(s)$ function by these. This may be done graphically[1] or analytically. The resulting plot will suggest a second pair, and $\frac{\theta_i}{\theta_o}(s)$ can be divided by these to give more closely the first pair.

Continued use of this method warrants the use of specialized protractors and pivoted scales, as explained by those who developed this method, to simplify the reading of angles, the plotting of the curvilinear lattice, and any necessary graphic division. The author of this paper has attempted to explain this method without mention of special equipment. An illustrative problem will be worked by this method a little later to more completely explain the procedure.

1. Brown and Campbell, <u>op. cit.</u>, pp. 165-166.

METHOD II

Apply an approximate inverse transform to a succession of trapezoidal waves obtained from the output frequency response curve $\Theta_o(j\omega)$, and obtain an approximate transient response.[1]

Let the overall output function of a system be $H(s) = \Theta_0(s)$. The general problem of finding h(t), the transient response, knowing H(s) is in performing the inverse Laplace transform.

It is assumed throughout the following treatment that:

1. H(s) may be written as a ratio of two rational polynomials in s with real and constant coefficients.

2. $\lim_{s \to \infty} H(s) = 0$

3. H(s) has no poles on the imaginary axis or in the right half-plane. In general the inverse transform is given by the integral

$$h(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} H(s)\, e^{ts}\, ds \tag{13}$$

where the path of integration is parallel to the imaginary axis and c is a value such that the path of integration is to the right of all poles of H(s). Since by assumption there are no poles on or to the right of the imaginary axis, the path of integration may be made this axis.

So
$$h(t) = \frac{1}{2\pi j} \int_{-j\infty}^{+j\infty} H(s)\, e^{ts}\, ds \tag{14}$$

Let $s = j\omega$
$H(s) = \operatorname{Re} H(j\omega) + j \operatorname{Im} H(j\omega)$

---

1. Ibid., pp. 332-365.

2. David Widder, The Laplace Transform, Princeton University Press, Princeton, New Jersey, 1941, p. 241.

Re and Im refer to real and imaginary parts of the function. Make these substitutions into 14.

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\mathrm{Re}\,H(j\omega)\cos\omega t - \mathrm{Im}\,H(j\omega)\sin\omega t\right]d\omega + \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\mathrm{Re}\,H(j\omega)\sin\omega t + \mathrm{Im}\,H(j\omega)\cos\omega t\right]d\omega \qquad (15)$$

Since H(s) is a ratio of polynomials in s with real and constant coefficients, Re $H(j\omega)$ is an even function in ω, and Im $H(j\omega)$ is an odd function in ω. Cos ωt, by series expansion, is an even function in ω, and sin ωt is odd. The bracketed term of the first integral is then an even function in ω, while that of the second integral is odd. Since the limits of integration are from $-\infty$ to $+\infty$, the value of the first integral is twice that from 0 to ∞, while the value of the second integral is zero. Equation 15 then becomes:

$$h(t) = \frac{1}{\pi}\int_{0}^{\infty}\mathrm{Re}\,H(j\omega)\cos\omega t\,d\omega - \frac{1}{\pi}\int_{0}^{\infty}\mathrm{Im}\,H(j\omega)\sin\omega t\,d\omega \qquad (16)$$

Both integrals contain functions of time in the sine and cosine terms, the first integral being an even function of time, the second, odd. For the integration indicated in equation 14 along a closed path which takes in the imaginary axis from $-j\infty$ to $+j\infty$, plus a large semi-circle in the right half-plane, the value of h(t) is zero for all negative values of time, since, by assuming no poles of H(s) in the right half-plane, the sum of the residues at the poles of $H(s)e^{ts}$ within this contour is zero.[1] This means that numerically, the two functions of time in equation 16 are equal for all values of time.

---

1. E. A. Guillemin, The Mathematics of Circuit Analysis, John Wiley & Sons, Inc., New York, 1949, p. 304.

$$\left| \frac{1}{\pi} \int_0^{\infty} \operatorname{Re} H(j\omega) \cos \omega t \, d\omega \right| = \left| \frac{1}{\pi} \int_0^{\infty} \operatorname{Im} H(j\omega) \sin \omega t \, d\omega \right| \qquad (17)$$

And for positive values of time:

$$h(t) = \frac{2}{\pi} \int_0^{\infty} \operatorname{Re} H(j\omega) \cos \omega t \, d\omega \qquad (18)$$

This is the exact inverse transform for an H(s) satisfying the stated conditions.

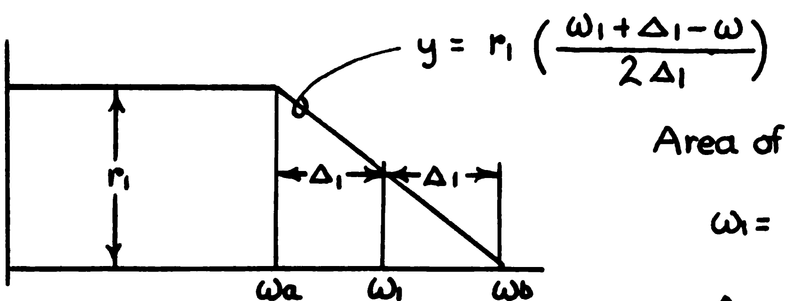
Few practical problems are of such a form the Re $H(j\omega)$ can be integrated as such, so a general approximate procedure is used.

1. Plot Re $H(j\omega)$ against ω and approximate the curve with a series of straight line segments. See Fig. 4.

2. Write the straight line approximation as a sum of trapezoidal functions, and apply equation 18 or a simplification to each one. The sum of the resulting time functions is an approximate h(t).

The general trapezoidal function used to approximate Re $H(j\omega)$ is shown in Fig. 3.

Let Re $H(j\omega)$ aproximation = $R(\omega)$



$$y = r_1 \left(\frac{\omega_1 + \Delta_1 - \omega}{2\Delta_1} \right) \qquad \omega_a \leq \omega \leq \omega_b$$

Area of trap., $A_1 = r_1 \omega_1$

$$\omega_1 = \frac{\omega_b + \omega_a}{2}$$

$$\Delta_1 = \frac{\omega_b - \omega_a}{2}$$

Fig. 3. General trapezoidal function.

$$h_1(t) = \frac{2}{\pi}\int_0^{\infty} R(\omega)\cos\omega t \, d\omega \tag{19}$$

$$h_1(t) = \frac{2}{\pi}\int_0^{\omega_1-\Delta_1} r_1\cos\omega t\, d\omega + \frac{2}{\pi}\int_{\omega_1-\Delta_1}^{\omega_1+\Delta_1} r_1\left(\frac{\omega_1+\Delta_1-\omega}{2\Delta_1}\right)\cos\omega t\, d\omega \tag{20}$$

Expand:

$$h_1(t) = \frac{2r_1}{\pi}\int_0^{\omega_1-\Delta_1}\cos\omega t\, d\omega + \frac{2r_1(\omega_1+\Delta_1)}{2\pi\Delta_1}\int_{\omega_1-\Delta_1}^{\omega_1+\Delta_1}\cos\omega t\, d\omega - \frac{2r_1}{2\pi\Delta_1}\int_{\omega_1-\Delta_1}^{\omega_1+\Delta_1}\omega\cos\omega t\, d\omega. \tag{21}$$

$$h_1(t) = \frac{2r_1}{\pi t}\sin\omega t\Big|_0^{\omega_1-\Delta_1} + \frac{r_1(\omega_1+\Delta_1)}{\pi\Delta_1 t}\sin\omega t\Big|_{\omega_1-\Delta_1}^{\omega_1+\Delta_1} - \frac{\omega r_1}{\pi\Delta_1 t}\sin\omega t\Big|_{\omega_1-\Delta_1}^{\omega_1+\Delta_1} - \frac{r_1}{\pi\Delta_1 t^2}\cos\omega t\Big|_{\omega_1-\Delta_1}^{\omega_1+\Delta_1} \tag{22}$$

Substitute limits; sine terms cancel, leaving:

$$h_1(t) = \frac{-r_1}{\pi\Delta_1 t^2}\left[\cos(\omega_1+\Delta_1)t - \cos(\omega_1-\Delta_1)t\right] \tag{22}$$

$$h_1(t) = \frac{2r_1}{\pi\Delta_1 t^2}\left(\frac{\omega_1}{\omega_1}\right)\sin\omega_1 t \;\sin\Delta_1 t \tag{23}$$

$r_1\omega_1 = A_1$, trapezoidal area

$$h_1(t) = \frac{2A_1}{\pi}\;\frac{\sin\omega_1 t}{\omega_1 t}\;\frac{\sin\Delta_1 t}{\Delta_1 t} \tag{24}$$

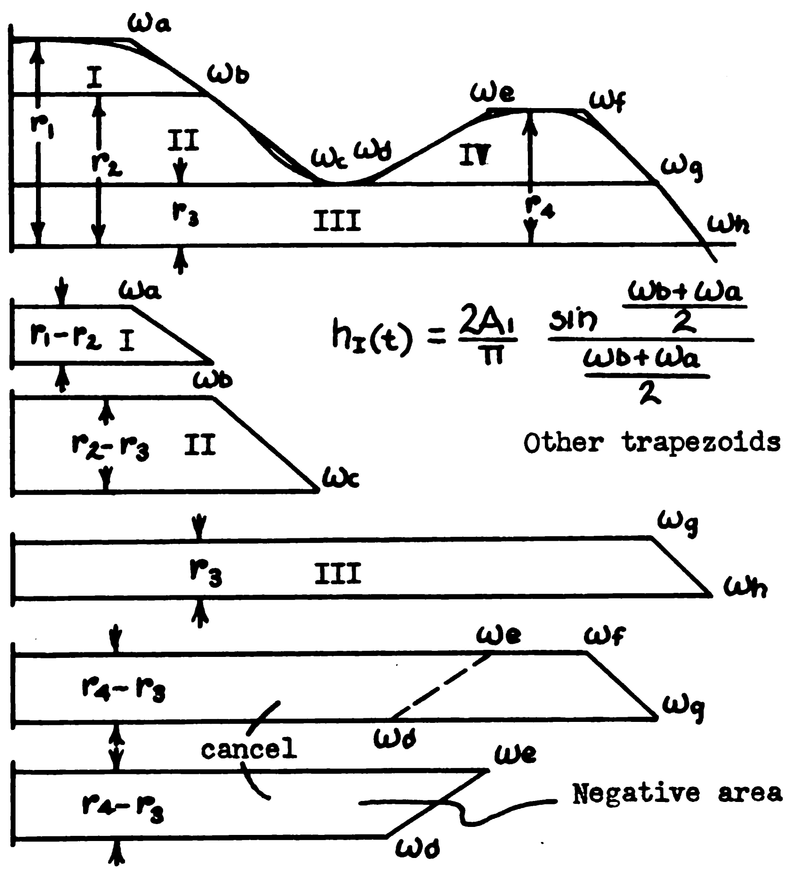
For a general function Re $H(j\omega)$,

$$h(t) = \sum_{K=1}^{n}\frac{2A_K}{\pi}\;\frac{\sin\omega_K t}{\omega_K t}\;\frac{\sin\Delta_K t}{\Delta_K t} \tag{25}$$

This equation can be evaluated more easily by using tables of $\frac{\sin x}{x}$.[1]

Re $H(j\omega)$, or Re $\Theta_o(j\omega)$ can be derived from a locus plot of $\frac{\Theta_o}{\Theta_i}(j\omega)$, its inverse, $\frac{\Theta_i}{\Theta_o}(j\omega)$, or by manipulation of $\frac{\epsilon}{\Theta_o}(j\omega)$. Take $\Theta_o(j\omega)$ and multiply the magnitude of the function at each ω value by the cosine of the angle at which the function lies to find Re $\Theta_o(j\omega)$.

Having obtained Re $\Theta_o(j\omega)$, express it as a series of trapezoidal waves as shown:



$$h_I(t) = \frac{2A_1}{\pi} \frac{\sin \frac{\omega_b+\omega_a}{2}}{\frac{\omega_b+\omega_a}{2}} \frac{\sin \frac{\omega_b-\omega_a}{2}}{\frac{\omega_b-\omega_a}{2}}$$

Other trapezoids treated similarly.

Fig. 4. Expression of Re $\Theta_o(j\omega)$ as series of trapezoidal waves.

---

1. Zeitschrift für Krystallographie, Vol. 85, p. 404, Berlin, 1933, $\frac{\sin x}{x}$ tables prepared by Dr. J. Sherman, reproduced in Brown and Campbell, op. cit., p. 357.

It is advisable, in order to plot the results of the summation of equation 25, to tabulate the evaluation of each term of the summation over a range of t sufficient to plot the transient waveform h(t) until it has reached to within a few percent of its final value.

The following form is suggested:

$h(t) \backslash t$	0	0.1	0.2	0.3	→
$h_1(t)$	$h_1(t)\|_{t=0}$	$h_1(t)\|_{t=0.1}$			
$h_2(t)$	$h_2(t)\|_{t=0}$	$h_2(t)\|_{t=0.1}$			
⋮	. . . .	. . . .	. .		
sum	. . . .	. . .	.		

Table 1.

The plot of the sum versus time yields h(t), the transient response. An illustrative example will be worked by this method a little later to more completely explain the procedure.

Method III

Approximate the unit step input by a half cycle of a square wave, resolve the square wave into its sinusoidal components and sum the system frequency response (known from $\frac{\theta_o}{\theta_i}(j\omega)$ frequency response curve) to a sufficient number of these component waves to obtain the transient response of the system to a unit step input.[1]

---

1. Mass and Hayman, op. cit., p. 1-13.

The frequency response curve $\frac{\theta_o}{\theta_i}(j\omega)$ shows the system gain and phase shift of the output for each single input frequency. If the general input to a system could be expressed as a sum of sinusoidal components, the output could be calculated as a sum of responses to the individual input frequencies.

A step function, which is the standard input for the determination of the transient response of a system, cannot be expressed as a sum of individual frequencies, but a square wave can be, and if the fundamental frequency of the square wave is selected so that the time necessary for the transient response of the system to settle to a few percent of its final value is less than the time for a half cycle of the square wave, then the transient response of the system to the first half cycle of the square will closely represent the transient response to a unit step, for a not too lightly damped system.

An input square wave, of maximum value unity, of minimum value zero and of fundamental frequency ωf is expressed by the formula:

$$F_i(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)\,\omega_f t \qquad (26)$$

The amplitude and phase of the individual frequencies will be altered by the system, and the general output wave will appear as:

$$F_o(t) = \frac{1}{2} + \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{A_{2n-1}}{2n-1} \sin\left[(2n-1)\,\omega_f t + B_{2n-1}\right] \qquad (27)$$

where A_{2n-1} is the gain of the system at a frequency of $(2n-1)\omega_f$ and B_{2n-1} is the phase shift at this frequency. Thus, an approximate transient response curve can be obtained by evaluating equation 27 over a range of t from zero to slightly less than a half period of ω_f, using values of A_{2n-1} and B_{2n-1} gotten from the frequency response curve $\frac{\theta_o}{\theta_i}(j\omega)$ of the system, or $\frac{\theta_i}{\theta_o}(j\omega)$.

The constant '$\frac{1}{2}$' in equation 27 should not be used when the system has an integrating element (Eg. a capacitor) in series. Such an element can be detected by examination of the transfer function curve (it will have a zero value at $\omega=0$) or the transfer function expression (it will have an s factor in the numerator). This '$\frac{1}{2}$' represents the D.C. level upon which the regular alternating square wave is superimposed, and it is seen that whatever the DC level of the input, the output will be clamped to a zero reference level if it is assumed that the capacitor has adequate time to acquire the DC charge of each half to the square wave input.

Evaluating ω_f, the fundamental frequency of the applied square wave requires some care. If ω_f is too low, an excessive amount of work must be done, since the response to a greater number of odd harmonics will have to be considered in order that the maximum amplitude of the response of the system to the highest harmonic chosen will be a low enough value so that higher harmonics can be neglected. If ω_f is too high, the accuracy of the method will be impaired, since the transient will not have time to die down during a half cycle.

It has been determined that for an ω_f which is $\frac{1}{5}$ of the lowest undamped natural frequency of the system, consideration of the first six terms of equation 27, that is, using odd harmonics one through eleven, will be sufficient.

1. Choose ω_f.

Where the gain curve exhibits a sharp peak, indicating light damping, this peak falls near the natural resonant frequency of the circuit. Let the frequency of maximum gain be ω_o. A choice of $\omega_f = \frac{\omega_o}{5}$ is permissible in this case. As damping increases, ω_o decreases, and ω_f must be chosen greater than $\frac{\omega_o}{5}$ in order to give reliable results. With ω_f too low, the system gain will not be low enough at the 11th harmonic, and extra harmonic terms will have to be considered. If the contribution of the 11th harmonic is greater than 2% of the contribution of ω_f, either raise ω_f or use the 13th harmonic if this gives results in the 2% region. If the contribution of the 9th is less than 2%, ω_f is too high and must be lowered.

If the form of the gain curve is such that is has a late resonant peak near the point where the 11th harmonic would fall were the high frequency trend neglected, consideration of an extra harmonic term or two will increase the accuracy of the results.

If the gain curve falls of monotonically, indicating heavier damping, ω_f maybe tentatively chosen so that at $5\,\omega_f$ the phase shift is 90°.

In general, where the gain curve has some positive value at $\omega = 0$ choose as the 11th harmonic of ω_f a frequency such that the maximum gain is 15 db below the gain of the low frequency portion of the curve. Check as before to see if 11th harmonic contribution $\simeq$ 2% of the contribution of ω_f. If the gain curve is zero at $\omega = 0$, as it will be if the system contains a series integrating component, as a capacitor, choose ω_f by resonant peak method and then apply the 11th harmonic check. Experience will soon give a measure of skill in choosing ω_f.

2. Read from the gain and phase shift curves the values needed to make the summation of equation 27:

$$F_o(t) = \left(\frac{1}{2}\right) + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{A_{2n-1}}{2n-1} \sin\left[(2n-1)\omega_f + B_{2n-1}\right] \tag{27}$$

and extend this over a sufficient range of t values.

Certain measures may be taken to simplify the procedure. Namely: construct a $\frac{2}{\pi}\frac{A_{2n-1}}{2n-1}$ table (column 4, Table 2) and alter

$$\measuredangle \theta_n = (2n-1)\,\omega_f \quad \text{into} \quad \measuredangle \theta_n = (2n-1)\,\frac{\omega_f}{\omega_o}\,\frac{t}{T}\,\frac{360^\circ}{2\pi} \tag{28a}$$

By letting $\omega_o = \frac{1}{T}$, with $\frac{\omega_f}{\omega_o} \cong .2$, $\frac{360^\circ}{2\pi}$ being a unity factor for conversion of radians to degrees, and $\frac{t}{T}$ a dimensionless parameter, equation 28 becomes:

$$\measuredangle \theta_n = (2n-1)\,\frac{36}{\pi}\,\frac{t}{T} \tag{28b}$$ [1]

n	$(2n-1)\omega_f$	phase shift B_{2n-1}	gain A_{2n-1}	$\frac{2}{\pi}\frac{A_{2n-1}}{2n-1}$
1	ω_f	B_1	A_1	$\frac{2}{\pi}A_1$
2	$3\omega_f$	B_3		
⋮				
6	$11\omega_f$			

Table 2.

Gain and phase shift of ω_f harmonics.

1. In the solution of the single illustrative example to follow this simplification will not be used. $(2n-1)\,\omega_f\,\frac{360}{2\pi}\,t$ will be evaluated directly, and t will not be contained in the dimensionless parameter $\frac{t}{T}$, thus simplifying the construction of the desired plot of amplitude versus t which will be compared with that of the exact solution.

n \ t/T	0	0.1	0.2	→
1	B_1	$\theta_1\big\|_{t/T=0.1} + B_1$	. . . .	
2	B_3	$\theta_2\big\|_{t/T=0.1} + B_3$		
...				
6	B_{11}			

Table 3.

Tabular form of $\theta_n + B_{2n-1}$ over range of $\frac{t}{T}$ values.

Multiply the sine of each angle in the n = 1 row of Table 3 by the n = 1 row value of $\frac{2}{\pi}\frac{A_{2n-1}}{2n-1}$ from Table 2. Record as in Table 4 and repeat with remaining rows.

n \ t/T	0	0.1	0.2	→
1	$\frac{2}{\pi}A_1 \sin B_1$	$\frac{2}{\pi}A_1 \sin(\theta_1+B_1)\big\|_{0.1}$	. . . .	
2	$\frac{2}{\pi}A_3 \sin B_3$	. . . .		
. .	. . .	. . .		
+0.5	+0.5	+0.5		
sum				

Table 4.

Tabular data necessary to plot **output** vs t (or $\frac{t}{T}$).

Add 0.5 to the sum of each column of Table 4 (this is the DC reference level of equation 27), except when the gain curve shows that the system has an integrating component in series by having a gain of zero and a 90° phase shift at $\omega = 0$, and plot the column sums against time to obtain the approximate transient response to a step input.

In order to compare these three methods of finding transient response from frequency response, an illustrative problem will be worked by each of the three in order.

SOLUTION OF PROBLEM

A single problem which will illustrate all three of the methods and yield in addition an exact solution for purposes of comparison is difficult to frame. Its expression must have a pair of conjugate roots and several real roots, to illustrate Method I; it must have no poles in the right half plane, or on the imaginary axis (or at zero) in order to satisfy the conditions of Method II; it must have an exact analytic form for application of the inverse Laplace transform.

The problem shown in Fig. 5, with a judicious choice of numerical values, satisfies the conditions atated above.

Given: system as shown.

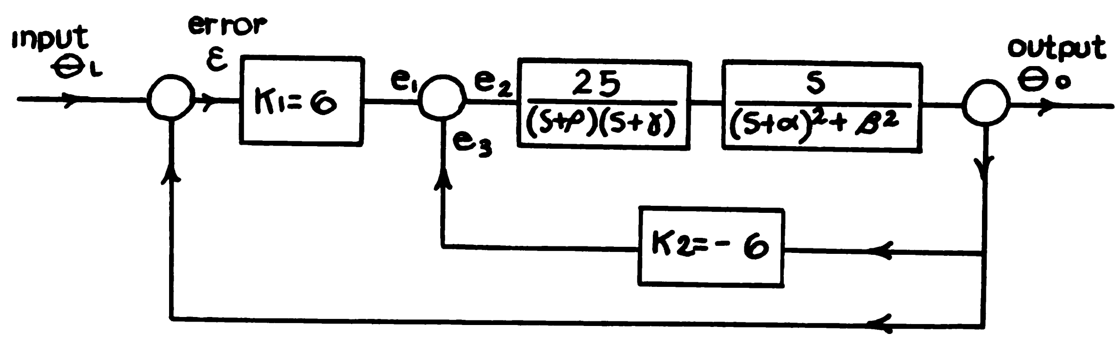


Fig. 5. Block diagram of illustrative problem.

$$6\varepsilon = e_1 = e_2 + e_3$$

$$e_3 = -6\theta_o \quad , \qquad e_2 = \frac{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho)}{25s}\,\theta_o$$

$$6\varepsilon = \theta_o\left\{\frac{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho)}{25s} - 6\right\}$$

$$\therefore \frac{\varepsilon}{\theta_o}(s) = \frac{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho) - 150s}{150s} \qquad (29a)$$

$$\frac{\theta_i}{\theta_o}(s) = \frac{\varepsilon}{\theta_o}(s) + 1 = \frac{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho)}{150s} \qquad (30)$$

Note the selection of $K_2 = -6$ in equation which makes $\frac{\theta_i}{\theta_o}(s)$ factorable. By inversion,

$$\frac{\theta_o}{\theta_i}(s) = \frac{150\,s}{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho)} \tag{31a}$$

Since, for a unit step input, $F(s) = \frac{1}{s}$,

$$\theta_o(s) = \frac{150}{[(s+\alpha)^2+\beta^2](s+\gamma)(s+\rho)} \tag{32a}$$

The s coefficient of 150 in the numerator of the r.h. side of equation is necessary to cancel the s introduced in the denominator when removing $\theta_i(s) = \frac{1}{s}$ form the expression, which would violate the conditions of Method II that there be no poles at zero.

In the system of Fig. 5 assign values:

$\alpha = 1.2$	$\gamma = 6$
$\beta = 2.5$	$\rho = 3.3$

Applying the inverse Laplace transform to

$$\theta_o(s) = \frac{150}{[(s+1.2)^2+2.5^2](s+6)(s+3.3)} \tag{32b}$$

yields the transient solution

$$\theta_o(t) = -3.35\,e^{-1.2t}\sin(2.5t+102.5) - 1.88\,e^{-6t} + 5.16\,e^{-3.3t} \tag{33}$$

which is shown plotted in Figs. 7, 9, and 11.

At this point, the vector locus diagram has been made for $\frac{\theta_i}{\theta_o}(j\omega)$ (Fig. 6, note) and from this, vector locus plots of $\frac{\varepsilon}{\theta_o}(j\omega)$, required for Method I (Fig. 6 and note), $\frac{\theta_o}{\theta_i}(j\omega)$, required by Method III (Fig. 10) and Re $\theta_o(j\omega)$, required by Method II.

Although these curves were obtained analytically for the most part, it is possible to carry through the graphic solutions of Methods II and III without knowing the analytic expression for the system. It is almost necessary, however, to know the analytic expressions in order to obtain an accurate solution by Method I, since, in the event of strongly curved $\frac{\varepsilon}{\theta_o}(j\omega)$ near the -1 point, the curvilinear lattice is plotted by evaluating $\frac{\varepsilon}{\theta_o}(s)$ for several complex values of s.

ω	$\frac{\theta_i}{\theta_o}(j\omega)$ mag.	angle	$\frac{\theta_o}{\theta_i}(j\omega)$ mag.	angle	$Re\ \theta_o(j\omega)$
0	∞	-90	0	90	.99
.1	10.10	-85.5	.10	85.5	.99
.2	5.10	-81	.20	81	.98
.5	2.0	-67.5	.49	67.5	.91
.8	1.25	-53.6	.80	53.6	.81
1.0	.99	-43.6	1.01	43.6	.69
1.5	.60	-14.5	1.66	14.5	.26
2.0	.49	11.8	2.04	-11.8	-.20
3.0	.49	79	2.06	-79	-.67
4.0	.79	125	1.26	-125	-.26
6.0	2.05	169	.49	-169	-.015
8.0	4.31	192	.23	-192	.006
10.0	8.33	207	.12	-207	.005
20.0	55.50	237	.02	-237	$.8 \times 10^{-4}$

Table 5. Tabular data for plotting system frequency response curves.

METHOD I

To solve: Problem given in Fig. 5.

Given: the $\frac{\varepsilon}{\Theta_o}(j\omega)$ locus plot, baseline of the $-\sigma + j\omega$ system.

Find the conjugate roots of s by determining where the point $-1+j0$ lies on the $-\sigma+j\omega$ coordinate system. In Fig. 6 a rough sketch indicates that the point $-1+j0$ will lie between $\sigma = -2$ and $\sigma = -1$, and $\omega = 2$ and $\omega = 3$. In order to plot the curvilinear squares more accurately, substitute for s in the analytic expression of $\frac{\varepsilon}{\Theta_o}(s)$, whose expansion is:

$$\frac{\varepsilon}{\Theta_o}(s) = \frac{s^4 + 11.7s^3 + 49.8s^2 - 31s + 152.3}{150s} \tag{29b}$$

the four values	obtaining respectively
$s = -1 + j2$	$\frac{\varepsilon}{\Theta_o}(s) = .90\, e^{-j175.7}$
$= -1 + j3$	$= 1.10\, e^{-j186}$
$= -2 + j2$	$= .85\, e^{-j174}$
$= -2 + j3$	$= 1.12\, e^{-j174}$

Plot these points on Fig. 6 and draw the orthogonal lines. The complex conjugate roots are read as $s \simeq -1.25 \pm j2.6$.

The two remaining roots may be found by dividing the quadratic portion of the output expression $\Theta_o(s)$ by $(s + 1.25)^2 + 2.6^2$

$$\frac{s^4 + 11.7s^3 + 49.8s^2 + 119s + 152.3}{(s + 1.25)^2 + 2.6^2} \simeq$$

$$s^2 + 9.2s + 18.48 = (s + 6.23)(s + 2.97) \tag{34}$$

$$s_2 = -2.97 \quad , \quad s_3 = -6.23$$

The transient solution resulting from all the roots is

$$h(t) = \frac{e^{-1.25t}\cos(2.6t - \alpha_1 - \beta_1)}{a_1 b_1} + \frac{e^{-2.97t}}{a_2 e^{\alpha_2} b_2 e^{\beta_2}} + \frac{e^{-6.23t}}{a_3 e^{\alpha_3} b_3 e^{\beta_3}} \tag{35}$$

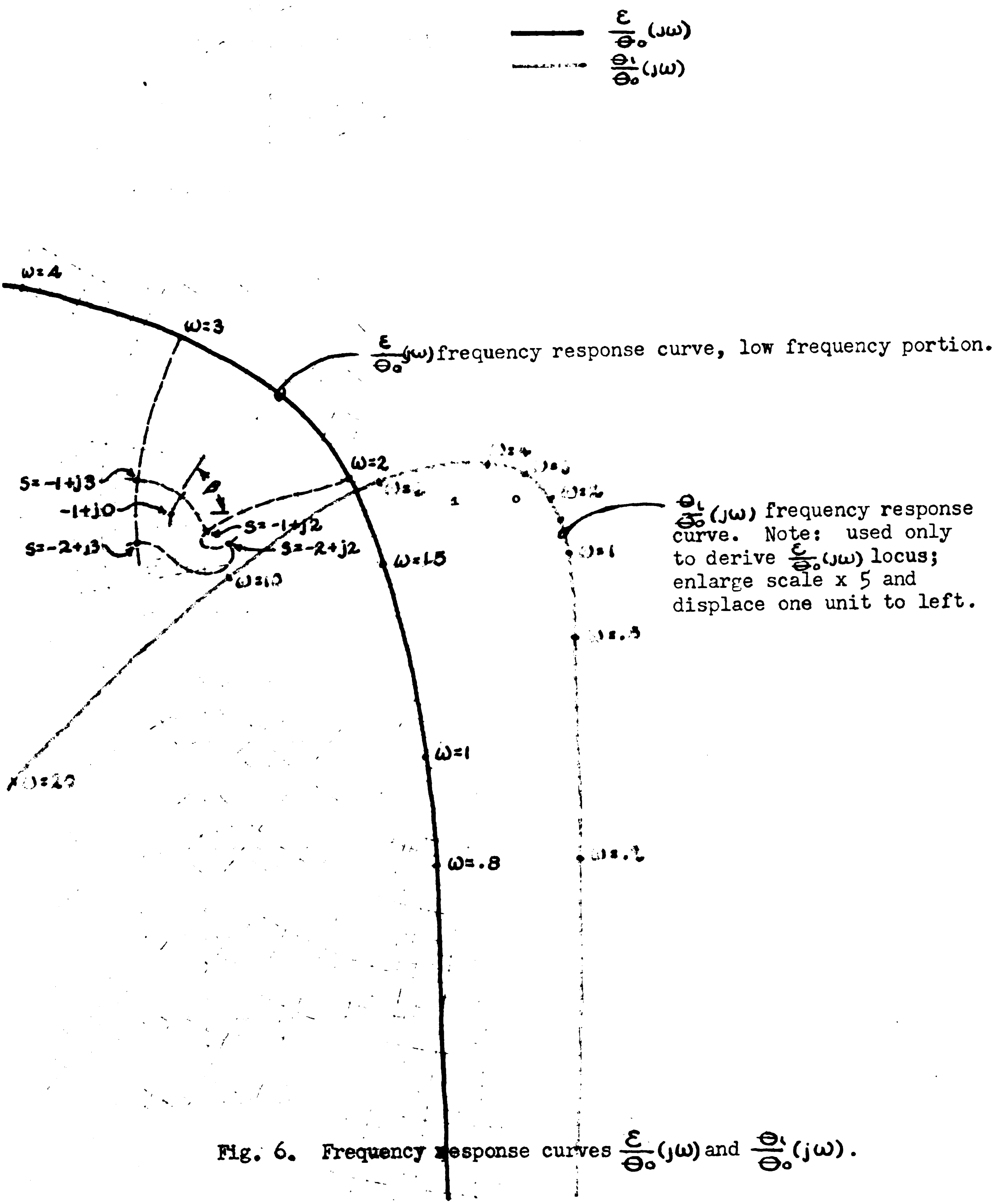


Fig. 6. Frequency response curves $\frac{\mathcal{E}}{\Theta_o}(j\omega)$ and $\frac{\Theta_i}{\Theta_o}(j\omega)$.

In $\frac{e^{-1.25t}}{a_1 b_1} \cos(2.6t - \alpha_1 - \beta_1)$,

$a_1 e^{\alpha_1} = s_1 = -1.25 + j2.6 = 2.86\, e^{j115.8}$

$$b_1 = \left| \frac{\Delta \frac{\varepsilon}{\Theta_o}(s)}{\Delta s} \right|_{s = s_1 = -1.25 + j2.6}$$

b_1 is the average of several $\frac{\Delta \frac{\varepsilon}{\Theta_o}(s)}{\Delta s}$ ratios in the region of the root.

		Δs	$\Delta \frac{\varepsilon}{\Theta_o}(s)$	ratio
Between s =	$-1+j3$, $-2+j3$	1	.02	.02
	$-1+j2$, $-2+j2$	1	.05	.05
	$-1+j2$, $-2+j3$	$\sqrt{2}$	.241	.17
	$-1+j2$, $-1+j3$	1	.20	.20
			4	.44
				.11

β_1 can be read directly from the plot as the angle between the real axis and the $\omega = 2.6$ line at the point $-1 + j0$ $\qquad \beta_1 \simeq 60°$

Thus the first term becomes:

$$\frac{e^{-1.25t} \cos(2.6t - 115.8 - 60)}{2.86 \times .11} = -3.18\, e^{-1.25t} \sin(2.6t + 94.2) \tag{36}$$

since $\cos(2.6t - 175.8) = -\sin(2.6t + 94.2)$.

In $\frac{e^{-2.97t}}{a_2 e^{\alpha_2} b_2 e^{\beta_2}}$, $\quad a_2 e^{\alpha_2} = s_2 = -2.97 + j0 = -2.97\, e^{j0}$

$$b_2 = \left| \frac{\partial}{\partial s}\left(\frac{\Theta_i}{\Theta_o}(s)\right) \right|_{s = s_2 = -2.97}$$ Find absolute value of derivative algebraically.

$$b_2 = \left| \frac{\partial}{\partial s} \left(\frac{s^4 + 11.7s^3 + 49.8s^2 + 119s + 152.3}{150s} \right) \right|_{s = -2.97}$$

$b_2 = .065$

The line of constant frequency associated with the root $s = -2.97$ passing through the point $-1+j0$ is the $\omega = 0$ line, hence the angle of intersection of this line with the real axis is 0° or 180°.

At $s = -2.97$, $\frac{\varepsilon}{\theta_0}(s) = -1$, while at $s = -2.97 + \Delta$, $\frac{\varepsilon}{\theta_0}(s) > -1$. Hence, β_2 180°. (As σ decreases, the ω const. line runs left to right).

Then $$\frac{e^{-2.97t}}{a_2 e^{\alpha_2} b_2 e^{\beta_2}} = \frac{e^{-2.97t}}{-2.97 \times .065\, e^{j180}} = 5.18\, e^{-2.97t}. \qquad (37)$$

For the third term, $\frac{e^{-6.23t}}{a_3 e^{\alpha_3} b_3 e^{\beta_3}}$,

$$a_3 e^{\alpha_3} = s_3 = -6.23 + j0 = -6.23$$

$$b_3 = 0.11$$

$$\beta_3 = 0^\circ$$

The third term is $$\frac{e^{-6.23t}}{-6.23 \times 0.11} = -1.46\, e^{-6.23t} \qquad (38)$$

The entire transient solution is the sum of equations 36, 37, and 38.

$$h(t) = -3.18 e^{-1.25t} \sin(2.6t + 94.2^\circ) + 5.18 e^{-2.97t} - 1.46\, e^{-6.23t} \qquad (39)$$

Compare the plot of this output waveform with the exact solution in Fig. 7. The merits and disadvantages of this method will be discussed at a later point and compared with the two methods remaining to be illustrated.

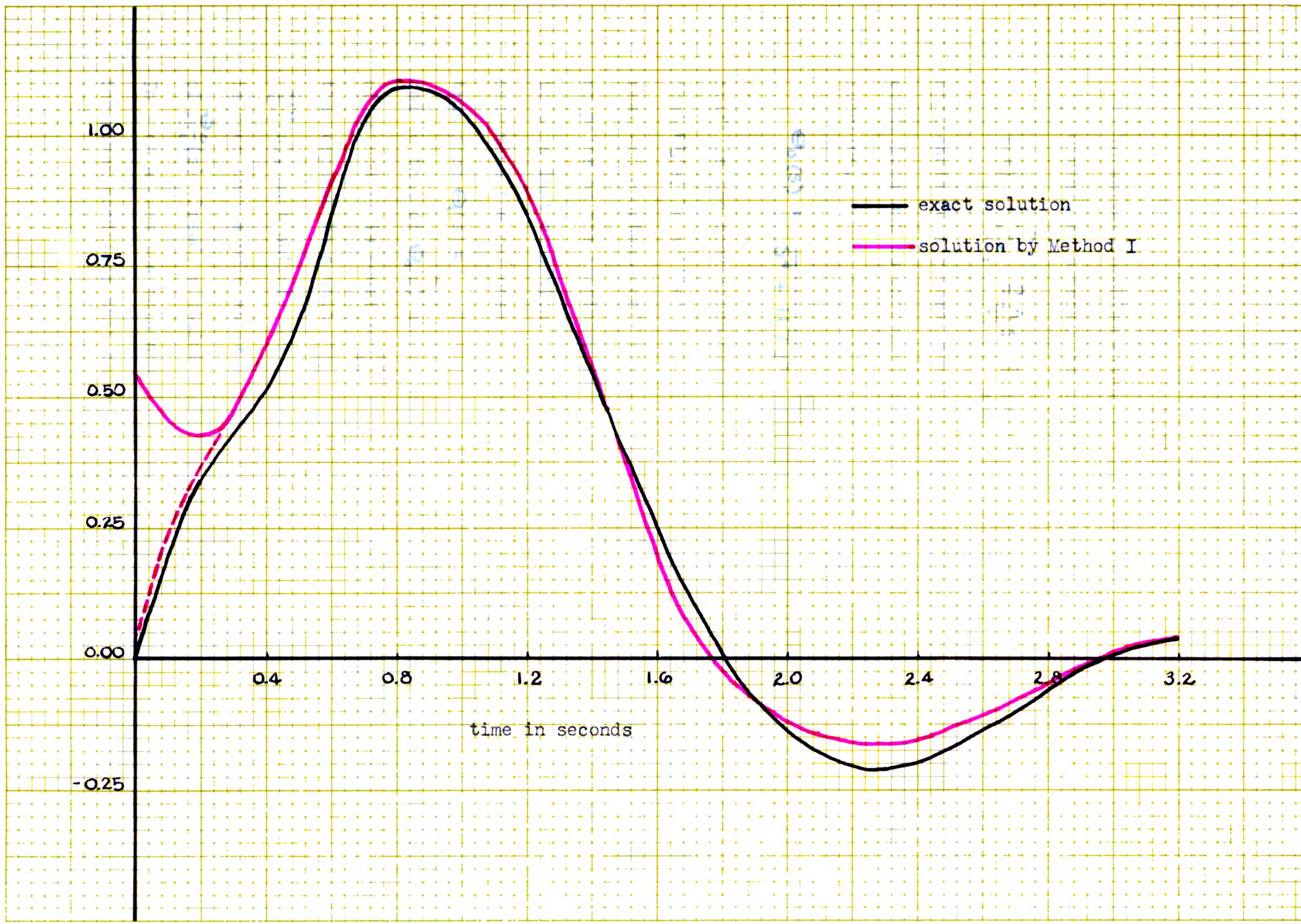


Fig. 7. Transient Response Curves.

METHOD II

To solve: problem given in Fig. 5.

Given: an approximate inverse transformation equation

$$h(t) = \sum_{K=1}^{n} \frac{2A_K}{\pi} \frac{\sin \omega_K t}{\omega_K t} \frac{\sin \Delta_K t}{\Delta_K t} \qquad (25)$$

applicable to a general function $Re\,\Theta_o(j\omega)$ which satisfies certain conditions and is expressed as a sum of trapezoidal waves. The output function,

$$\Theta_o(s) = \frac{150}{S^4 + 11.7s^3 + 49.8s^2 + 119s + 152.3} \qquad (32c)$$

fulfills the conditions necessary for application of this method, namely: that $\Theta_o(s)$ be written as a ratio of two rational polynomials in s, with real and constant coefficients; that lim. $\Theta_o(s)_{s \to \infty} = 0$; and that $\Theta_o(s)$ have no poles on the imaginary axis or in the right half-plane. With regard to this last condition: This method was specifically developed to determine the response of a system to a unit impulse, whose Laplace transform $\Theta_i(s) = 1$. Because of this, the expression of $\frac{\Theta_o}{\Theta_i}(s)$ is identical to that of $\Theta_o(s)$ alone, and the generalized H(s) in the explanation of Method II is equal to $\frac{\Theta_o}{\Theta_i}(s) = \Theta_o(s)$. In using this method to find the response to a unit step, (whose transform $\Theta_i(s) = \frac{1}{s}$), $\Theta_o(s)$, not $\frac{\Theta_o}{\Theta_i}(s)$, must satisfy the stated conditions. Since $\Theta_o(s) = \frac{1}{s}\frac{\Theta_o}{\Theta_i}(s)$ it is seen that a pole at s = 0 will exist in $\Theta_o(s)$ unless the expression of $\frac{\Theta_o}{\Theta_i}(s)$ already contains a zero at the origin (an s in the numerator). In such a case the pole and zero will cancel, and $\Theta_o(s)$ will satisfy the given condition that there be no poles on the imaginary axis or in the right half-plane. The method cannot be used to find the unit <u>step</u> response of any system save one whose transfer function $\frac{\Theta_o}{\Theta_i}(s)$ has one or more zeros at the origin, whereas it can be used to find the unit <u>impulse</u> response to any system whose transfer function has no poles at the origin.

Since the transfer function of the problem has a zero at the origin,

$$\frac{\Theta_o}{\Theta_i}(s) = \frac{150s}{[(s+1.2)^2+2.5^2](s+3.3)(s+6)} \tag{31b}$$

Then

$$\Theta_o(s) = \frac{1}{s}\frac{\Theta_o}{\Theta_i}(s) = \frac{150}{[(s+1.2)^2+2.5^2](s+3.3)(s+6)} \tag{32b}$$

and fulfills the three conditions.

The data necessary to plot $Re\,\Theta_o(j\omega)$ has been entered in column 3, Table 5. Plot the function and approximate the curve with straight line segments in Fig. 8.

Write the straight line approximation as a sum of trapezoidal function and apply

$$h(t) = \sum_{k=1}^{n} \frac{2A_k}{\pi} \frac{\sin \omega_k t}{\omega_k t} \frac{\sin \Delta_k t}{\Delta_k t} \tag{25}$$

where A_k is the area of each trapezoid,

ω_k is the mean base,

Δ_k is half the difference between the bases.

From Fig. 8, tabulate trapezoidal data.

Evaluate h(t) over a sufficient range of t values (see Table 7) and plot, in Fig. 9, along with the exact solution for purposes of comparison.

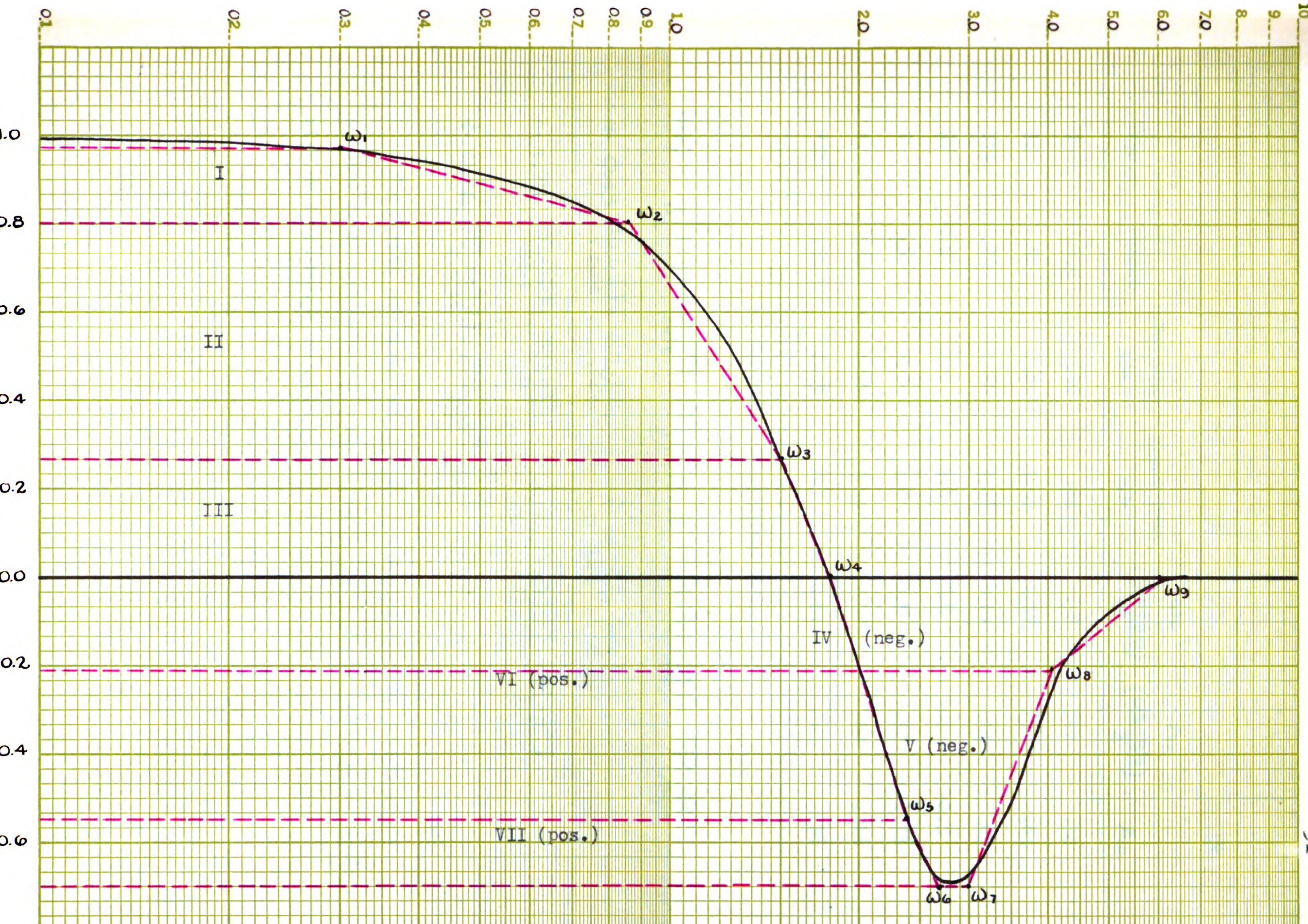


Fig. 8. **Re $\Theta_0(j\omega)$** with trapezoidal approximation.

Trap.	Height	Top Freq. ω_a	Lower Freq. ω_b	ω_K $\left(\frac{\omega_b+\omega_a}{2}\right)$	Δ_K $\left(\frac{\omega_b-\omega_a}{2}\right)$	A_K
I	.175	.30	.86	.58	.27	.10
II	.54	.86	1.50	1.17	.34	.63
III	.26	1.50	1.78	1.64	.14	.43
IV	-2.0	4.0	6.0	5.0	1.0	-1.0
V	-.50	3.0	4.0	3.5	.50	-1.75
VI	.56	1.78	2.35	2.07	.29	1.16
VII	.14	2.35	2.7	2.53	.18	.36

Table 6. Trapezoidal Data.

$$h(t) = \frac{2}{\pi}\left[0.10\,\frac{\sin .58t}{.58t}\,\frac{\sin .27t}{.27t} + 0.63\,\frac{\sin 1.16t}{1.16t}\,\frac{\sin .36t}{.36t}\right.$$
$$+ 0.43\,\frac{\sin 1.64t}{1.64t}\,\frac{\sin .14t}{.14t} - 1\,\frac{\sin 5t}{5t}\,\frac{\sin t}{t}$$
$$- 1.75\,\frac{\sin 3.5t}{3.5t}\,\frac{\sin .5t}{.5t} + 1.16\,\frac{\sin 2.07t}{2.07t}\,\frac{\sin .29t}{.29t}$$
$$\left. + 0.36\,\frac{\sin 2.53t}{2.53t}\,\frac{\sin .18t}{.18t}\right]. \qquad (40)$$

The merits and disadvantages of this method will be discussed at a later point and compared with the other two methods, the last of which will be illustrated immediately.

TIME IN SECONDS

trap. \ t	0	.2	.4	.8	1.2	1.0	2.0	2.4	2.8	3.2
I	.10	.10	.10	.10	.09	.08	.08	.07	.06	.07
II	.63	.63	.61	.54	.43	.31	.18	.07	-.05	-.07
III	.43	.42	.40	.32	.20	.08	-.02	-.08	-.09	-.07
IV	-1.00	-.84	-.44	.17	-.04	.08	.03	.01	-.01	.00
V	-1.75	-1.61	-1.22	-.20	.35	.18	-.14	-.14	-.05	.10
VI	1.17	1.14	1.04	.70	.29	-.06	-.23	-.21	-.09	.05
VII	.36	.35	.30	.16	.01	-.07	-.07	-.01	.04	.04
sum	-.06	.18	.78	1.79	1.40	.44	-.16	-.29	-.09	.10
$\times \frac{2}{\pi}$	-.04	.12	.49	1.13	.89	.28	-.10	-.19	-.06	.06

Table 7. Evaluation of h(t) expression of 7 trapezoids over t range 0 - 3.2 sec.

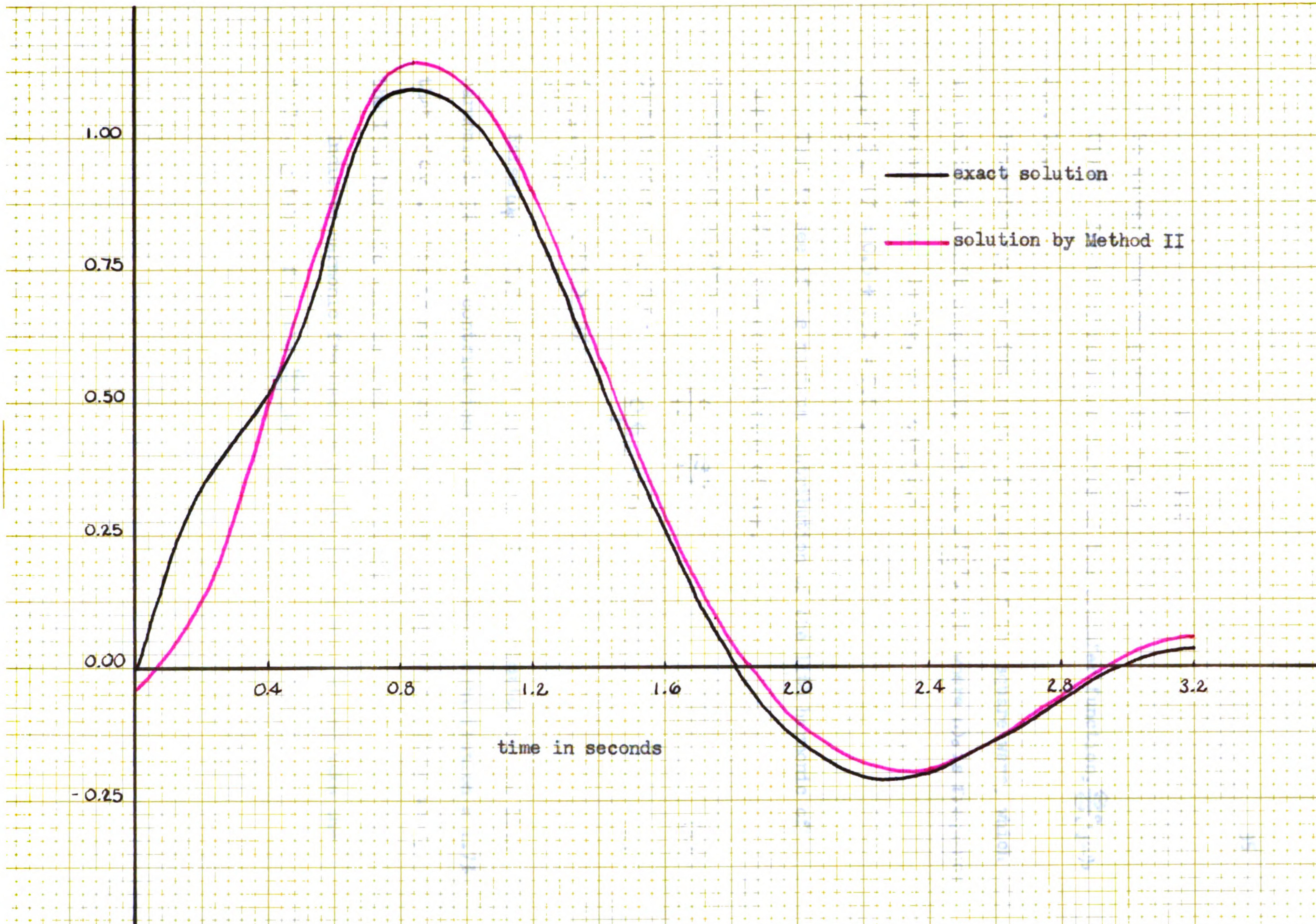
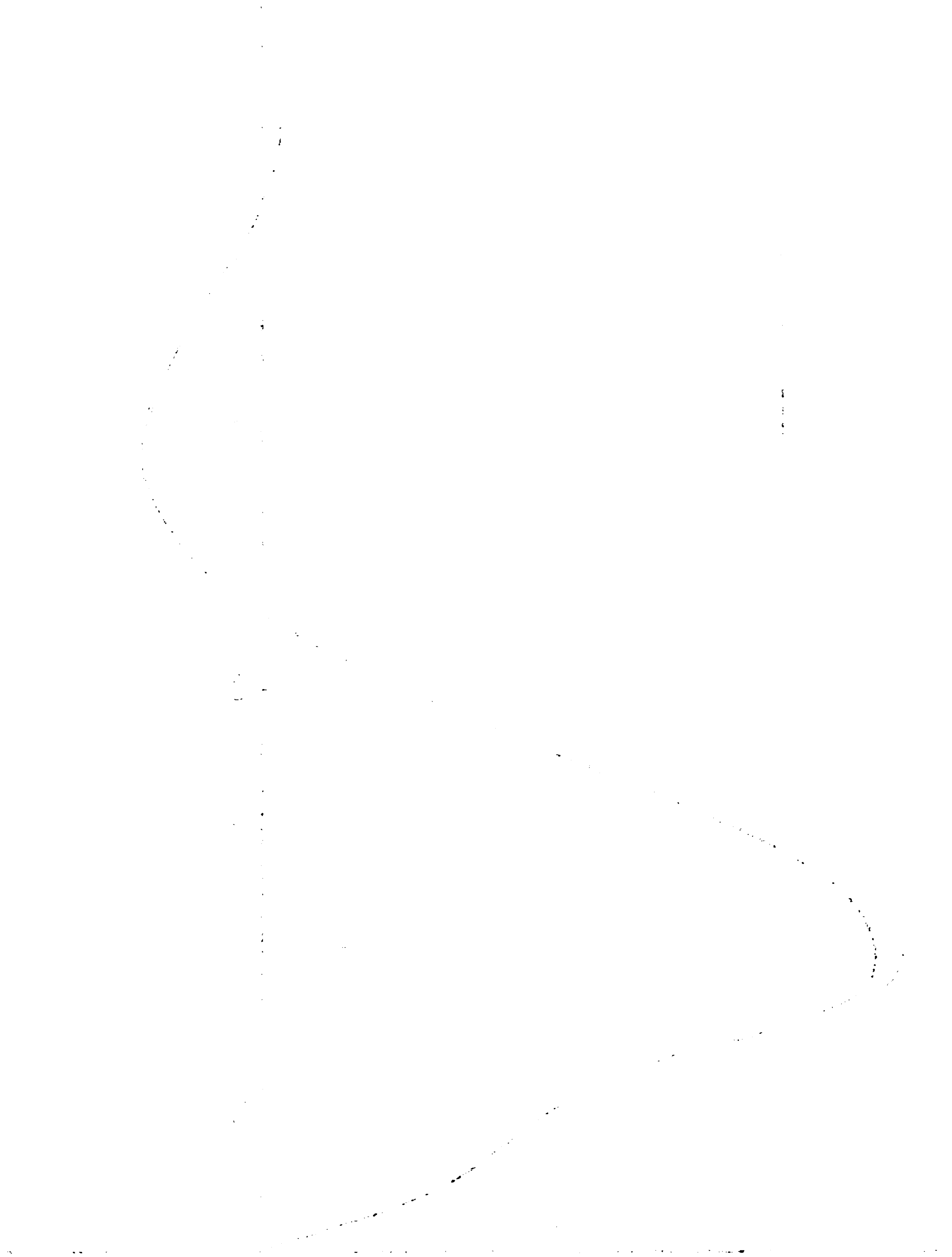


Fig. 9. Transient Response Curves.



METHOD III

To solve: problem given in Fig. 5.

Given: frequency response curves of the transfer function, $\frac{\theta_o}{\theta_i}(j\omega)$, Fig. 10.

Determine ω_f, the fundamental frequency of the square wave, which approximates a unit step input. Set $\omega_f = \frac{\omega_o}{5} + 20\%$, where ω_o is frequency of maximum gain. In Fig. 10, $\omega_o = 2.5$.

$\omega_f = \frac{2.5}{5} + 20\% = 0.5 + 0.1 = 0.6$.

Test to see if the maximum contribution of the 11th harmonic of $\omega_f = 0.6 \simeq 2\%$ of the contribution of ω_f.

Contribution of $\omega = 6.6$: $\frac{A_1}{11} = \frac{.37}{11}$ (a)

Contribution of $\omega_f = 0.6$: $\frac{0.6}{1} = 0.6$ (b)

$\frac{(a)}{(b)} = \frac{.0336}{0.6} = .056 > 2\%$

Raise ω_f to 0.7 and test as before. The contribution of the 11th harmonic $\simeq 3\%$ of ω_f contribution. The 13th harmonic contribution = 1.7% that of ω_f . This is satisfactorily close, and $\omega_f = 0.7$ will be used; the 13th harmonic will be used in the solution.

Indicate harmonic frequencies of $\omega_f = 0.7$ on Fig. 10 and make the following tables from the $\frac{\theta_o}{\theta_i}(j\omega)$ curves:

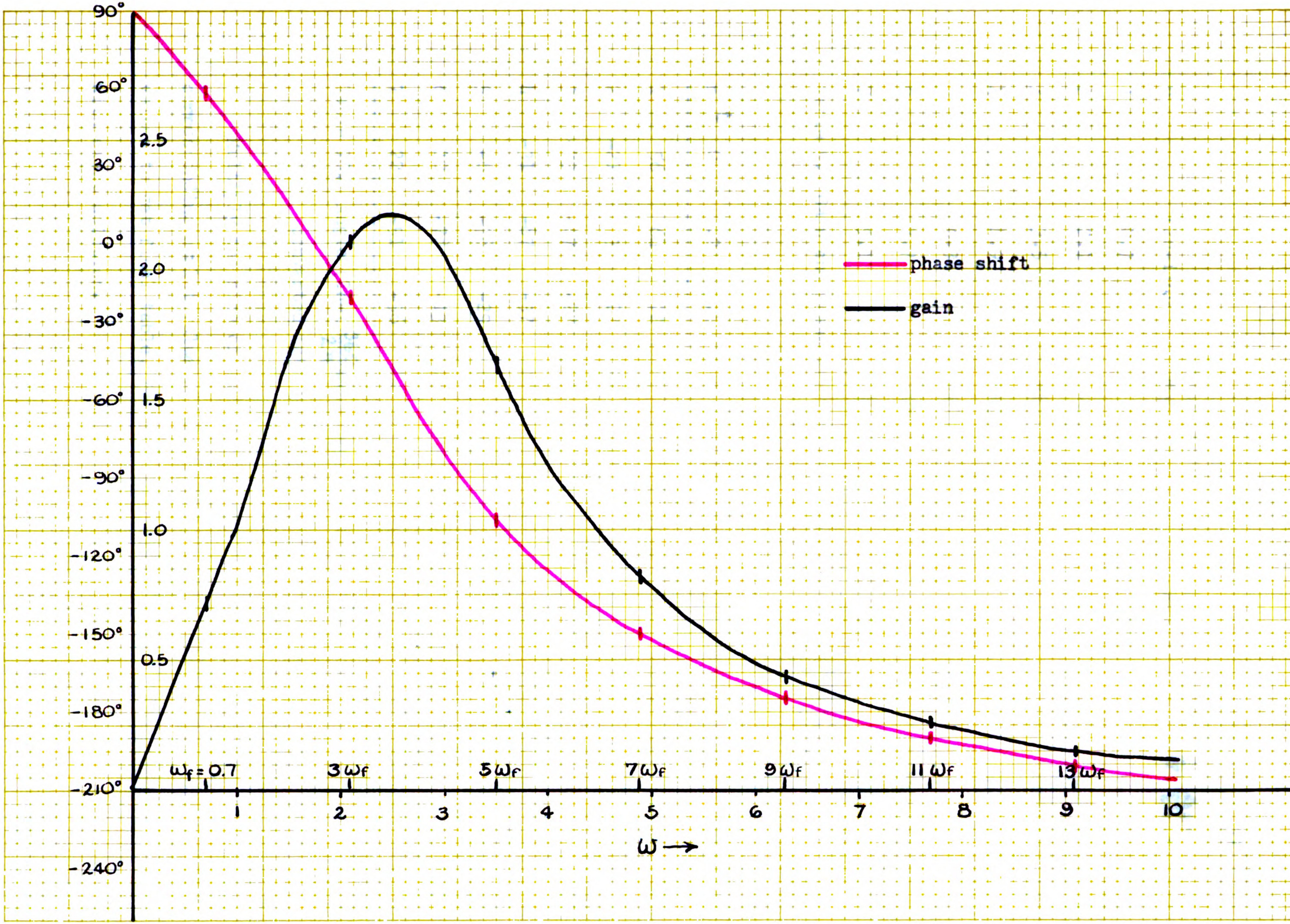
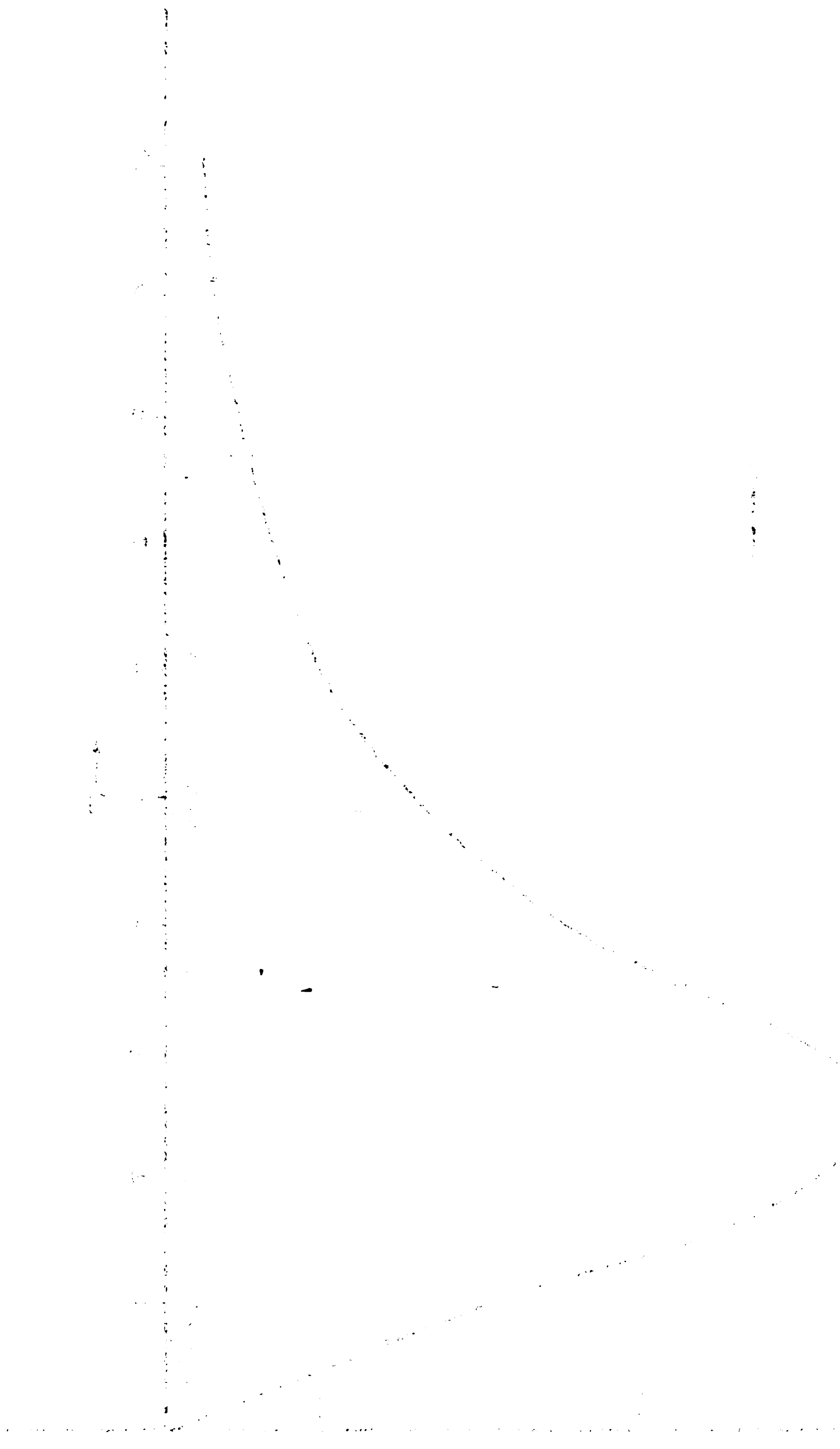


Fig. 10. Curves of System Transfer Function.



n	$(2n-1)\omega_f$	Phase Shift B_{2n-1}	Gain A_{2n-1}	$\frac{2}{\pi}\frac{A_{2n-1}}{2n-1}$
1	0.7	56	0.7	.45
2	2.1	-21	2.12	.45
3	3.5	-107	1.58	.20
4	4.9	-149	.85	.08
5	6.3	-174	.42	.03
6	7.7	-190	.24	.01
7	9.1	-202	.15	.008

Table 9. Gain and phase shift data of $\frac{\theta_o}{\theta_i}(j\omega)$ at odd harmonics 1-7 of $\omega_f = 0.7$.

TIME IN SECONDS

n \ t	0	.2	.4	.8	1.2	1.6	2.0	2.4	2.8	3.2
1	56	64	72	88	76	60	44	27	12	-5
2	-21	3	27	75	57	9	-39	-87	-45	3
3	-73	-67	-27	54	46	-35	-65	15	85	3
4	-31	-87	-37	75	-7	-61	61	17	-85	27
5	-6	-78	-29	64	-81	46	-10	-24	60	-84
6	10	-78	-13	16	-20	23	-26	30	-33	37
7	22	-51	-56	90	-57	24	10	-42	76	-70

Table 10. Angle $(2n-1)\frac{360}{2\pi}\omega_f t + B_{2n-1}$ table with angles in degrees converted into quadrant I & IV equivalents.

Multiply the sine of each term in the n = 1st row of Table 19 by the corresponding n = 1st value of $\frac{2}{\pi}\frac{A_{2n-1}}{2n-1}$, Table 9; repeat with other n values to get final table, Table 11 yielding solution. Plot in Fig. 11 with exact solution.

TIME IN SECONDS

n \ t	00	.2	.4	.8	1.2	1.6	2.0	2.4	2.8	3.2
1	.37	.40	.43	.45	.44	.39	.31	.21	.09	-.04
2	-.16	-.02	.20	.43	.40	.07	-.28	-.45	-.32	.02
3	-.19	-.19	-.09	.16	.15	-.12	-.18	.05	.20	.01
4	-.04	-.08	-.05	.08	.01	-.07	.07	.02	-.08	.04
5	-.003	-.03	-.014	.027	-.03	.02	-.005	-.012	.026	-.03
6	.002	-.01	-.002	.003	-.003	.004	-.004	.005	-.005	.005
7	.003	-.006	-.007	.008	-.007	.003		-.005	.008	-.008
SUM	-.02	.00	.47	1.16	.96	.28	-.09	-.18	-.08	-.003

Table 11. Evaluation of $h(t)$ expression for 7 square wave frequency components over t range 0 - 3.2 sec.

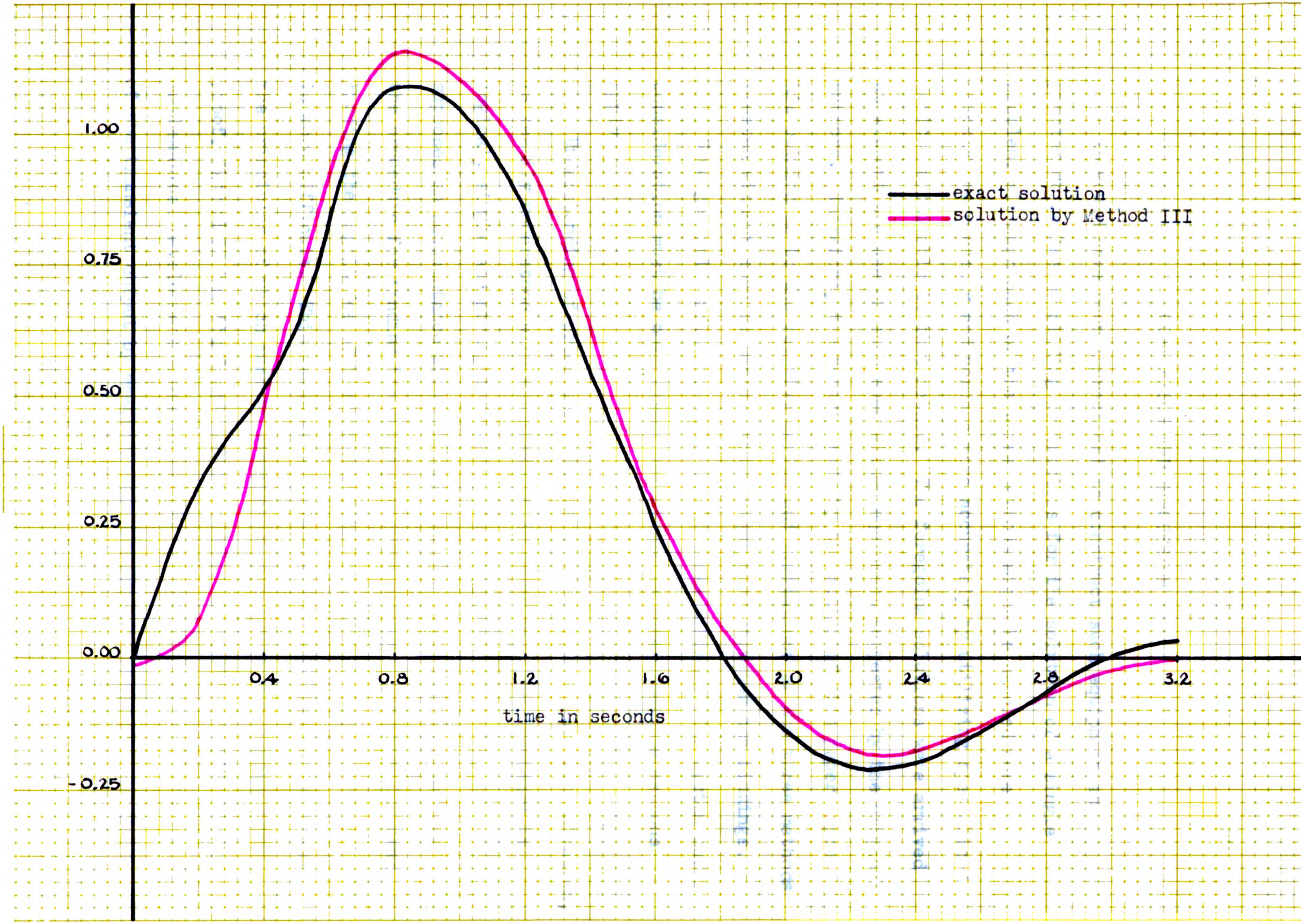
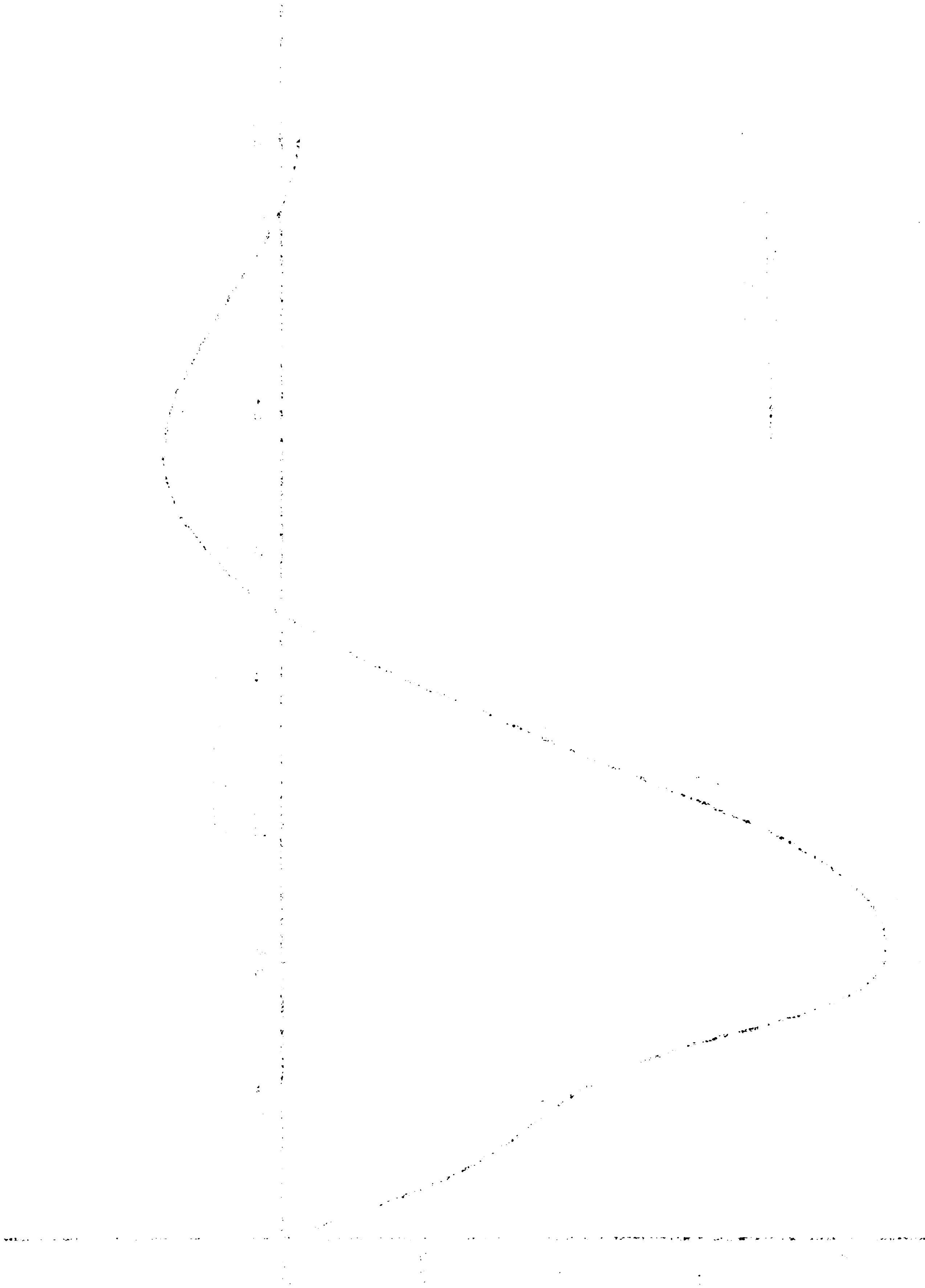


Fig. 11. Transient Response Curves.



COMPARISON AND DISCUSSION

It would have been impractical to construct such equipment as pivoted db scales or sliding protractors for the solution of a single problem, although each of the three methods might be handled more quickly, more easily and perhaps more accurately by utilizing specialized devices applicable to that method. The foregoing solutions were arrived at without the use of such specialized equipment suggested by the original authors. Therefore, the remarks made by the author of this paper express his opinion as to the relative merits and limitations of the three methods treated without special devices which might have affected the findings.

Some general conditions apply to all three methods. The system under consideration must be both linear and stable in order that the methods apply.

Understanding of procedure: Method III, the square wave approximation, is readily understood with a minimum of explanation. Its procedure is logical, simple and straight-forward. Method I, next in order, presupposes an acquaintance with conformal mapping and complex variables; the procedure for finding the conjugate complex roots is easily followed. Finding the magnitudes of the transient terms corresponding to the roots is more difficult. This requires a knowledge of transformation calculus, the principles of which must be accepted on the basis of a mathematical proof. Method II has the least easily understood approach. Here a mathematical development alone is used to develop the final expression of the transient output by means of the inverse Laplace transform, and the end expression is not obvious from an examination of the original inverse transform.

Applicability: Method III applies to any stable system, whatever the form of the curve of the system transfer function, and does not require that the analytic expression of the curve be known. Method II will not find the response to a unit step for a system whose output expression or curve indicates a pole at $s = 0$, thus violating the conditions necessary for the mathematical steps leading to the simplified expression for time response. Hence, the method will determine a unit step response only for a system whose transfer function has one or more zeroes at $s = 0$ as well as having no poles at $s = 0$. The method is applicable, however, to many more systems if it is used to determine the response to a unit impulse, since only the condition rejecting systems with poles at $s = 0$ applies. This method also will work without knowing the analytic expression of the transfer function. Method I is not readily applicable to a system having two or more pairs of conjugate complex roots, especially if the roots lie at about the same point on the complex plane. This is due to the fact that the error-to-output frequency response curve will be strongly bent for such a system, making difficult the plotting of the complex curvilinear square system. In such a case, and indeed, in the general case, where a high degree of accuracy is required, this method should not be applied to a system whose curves cannot be expressed analytically, because such expressions are necessary for the accurate plotting of the curvilinear squares, and for the algebraic determination of the real roots.

Accuracy: Method II, by increasing the number of trapezoids used in approximating the curve, may be made as accurate as desired, at the expense of time and labor. However, even for a small number of trapezoids used to obtain a rough approximation, the result closely follows the exact solution. Method III gives a solution that may agree to within a few percent of the exact response, depending upon the careful and tested selection of ω_f, the fundamental frequency of the square wave. Method I offers the least accurate solution of the three, especially at the beginning of the time range. The accuracy may be improved by making the necessary curvilinear plot to a larger scale and by plotting more points, again at the cost of time and labor. For all but the most critical requirements, however, the solution by Method I should suffice.

Speed: Method III is most quickly carried out, with the data necessary to evaluate the summation expression being read directly from the transfer function curves without delay. Method I allows the rough determination of the complex roots of s almost immediately if the frequency response curve is not too strongly bent. The approximate real roots also follow quickly from dividing the transfer function by the complex roots. However, the magnitudes of the transient terms corresponding to all of the roots require quite some time to find. Method II is regarded as the most time consuming, especially if a large number of trapezoids is considered, since obtaining the expressions for the trapezoids and then evaluating the large summation requires slow, painstaking work.

Advantageous features and general evaluation: Method I yields the equation of the transient response curve. This is desirable, since a change in the transient response due to an altered parameter may be found without again going through the entire procedure for finding the response, merely by considering the effect of the parameter change on the correct term or terms in the equation. This feature is not found in the other two methods. This method also has the advantage of yielding very quickly the approximate roots of the system from which the time response is found. The accuracy of Method II is its chief advantage. In many cases this will offset the additional time used in arriving at the solution. Method III, after due consideration, appears to be the best general method for obtaining a solution quickly and accurately, for, to repeat: the principles and procedures are most clear, the limitations are much fewer in number than those for the other methods, the findings agree favorably with the exact solution and are quickly obtained. Although Methods I and II may have the desirable features of speed and accuracy respectively, these are combined in Method III along with others to make this one of the most favorable graphic means shown for determining transient response from frequency response curves.

	Method I	Method II	Method III
Understanding of procedure	2nd	3rd	1st
Applicability	3rd	2nd	1st
Accuracy	3rd	potentially 1st	2nd
Speed	potentially 1st	3rd	2nd
General Advantages	Gives analytic formula of output waveform; change of system parameters readily handled; quickly gives approximate response and natural frequency of oscillation.	High accuracy possible.	Most reasonable approach; best general method.

Table 12. Tabular summary of Discussion.

BIBLIOGRAPHY

Brown, Gordon S. and Donald P. Campbell. Principles of Servomechanisms, John Wiley & Sons, Inc., New York, 1948.

Chestnut, Harold, and Robert W. Mayer. Servomechanisms and Regulating System Design, John Wiley and Sons, Inc., New York, 1951.

Evans, Walter R. "Graphic Analysis of Control Systems", AIEE Transactions, vol. 67, pp. 547-551.

Gardner, Murray F. and John L. Barnes. Transients in Linear Systems, John Wiley & Sons, Inc., New York, 1952.

Goldman, Stanford. Transformation Calculus and Electircal Transients, Prentice-Hall Inc., New York, 1949.

Guillemin, E. A. The Mathematics of Circuit Analysis, John Wiley & Sons, Inc., New York, 1949.

Sokolnikoff, Ivan S. Advanced Calculus, McGraw-Hill Book Company, Inc., New York, 1939.

Wass, C. A. A. and E. G. Hayman. "An Approximate Method of Deriving the Transient Response of a Linear System From the Frequency Response", Royal Aircraft Establishment, c.p. 113 Technical Note no. GW. 148, November, 1951.

Widder, David. The Laplace Transform, Princeton University Press, Princeton, 1941.