HEWLETT hp PACKARD

SQUARE WAVE AND PULSE TESTING OF LINEAR SYSTEMS

application note 17

1 APR 66

SQUARE WAVE & PULSE TESTING OF LINEAR SYSTEMS

B. M. Oliver
Hewlett-Packard Co.
Palo Alto, Calif.

Square waves or pulses are often employed in the testing of a wide variety of non-linear systems. In fact many such systems will operate properly with only certain specific input waveforms. As an example, the receiver for a pulse-position modulation telemetering system can be expected to operate properly only when supplied with pulse pairs having certain widths, amplitudes, repetition rates, and separations between the 1st and 2nd pulse. To test such a receiver with sine wave inputs would in all likelihood be meaningless. In this type of application the square wave or pulse generator serves merely as a simulator of the actual input signal for which the system was designed. While this is an extremely important field of application, the variety of cases is unlimited and any attempt on my part to make general remarks on the subject would only result in 99.44% pure platitudes. Rather we shall confine our attention to linear systems, where square waves or pulses may be used as test signals to give the same information about the system as could be obtained with sine wave inputs, and to do so more quickly.

Square wave or pulse testing has the following advantages:

1. Extreme rapidity. The entire system characteristic transient may be displayed many times per second. If the desired transient is known, the actual transient may be adjusted to the desired one by altering the circuit while watching the response. Square wave and pulse tests are dynamic measurements.

2. Often it is the square wave or pulse response which is of primary interest. In such cases sine wave testing is not only more tedious and time consuming, it is less direct.

Linear Systems

Before going any further let us try to define more precisely what we mean by a "linear system". The following definition is concise and adequate. If in a given system, an excitation f(t) produces a response h(t) and an excitation g(t) produces a response k(t), then the system is linear if and only if the response to

$$a\,f(t) + b\,g(t) \text{ is } a\,h(t) + b\,k(t), \qquad (1)$$

where a and b are arbitrary real numbers.

Now let's see what this says.

Since a and b are arbitrary, let's set b = o. Equation (1) then says the response to a f(t) is a h(t). This is a statement of the proportionality property. Twice the input produces twice the output. π times as much input produces π times as much output. We note that the output need not be proportional to the input on an instantaneous basis: it is not necessary that h(t) be equal, say, to some constant times $f(t - t_o)$. Such a requirement, while admitting delay, rules out linear filters and defines a system which is not only linear but distortionless.

Now let us set a and b = 1. Equation (1) then says that the response to f(t) + g(t) + h(t) + k(t). This is a statement of the superposition property. Each input produces its own output regardless of what other signals may be present. There is no interaction - no intermodulation. If f(t) is the piccolo and g(t) the drum, the response to the piccolo, h(t), is unaffected by the response to the drum k(t). Superposition is the property which enables one speaker cone to reproduce the whole orchestra, and your eardrum to distinguish the separate instruments.

Finally Equation (1) says that proportionality and superposition are simultaneously satisfied.

Actually we shall require one further property of our linear system: invariance with time. This may be stated as follows: If the response to f(t) is h(t),

1501 Page Mill Road, Palo Alto, California, U.S.A., Cable: "HEWPACK" Tel: (415) 326-7000
Europe: 54 Route Des Acacias, Geneva, Switzerland, Cable: "HEWPACKSA" Tel. (022) 42.81.50
Printed in U.S.A.

the response to $f(t - t_0)$ is $h(t - t_0)$. Delaying the input simply delays the output the same amount. This says there is nothing in the system which causes the response to change with time.

The Response of Linear Systems to Sine Waves

An arbitrary input wave f(t) applied to a linear system may produce an output h(t) which is different from f(t) in almost every respect. h(t) may be bigger, smaller, longer, shorter, more wiggly, smoother or what have you. (Of course, h(t) can't begin before f(t). Systems which might otherwise behave this way are unstable. Also there can be no frequencies in h(t) which are not in f(t), but this is getting ahead of the story.) There is, however, one class of functions which any linear system can alter in only two very simple ways. These are the sine waves:

The response of any linear system to a sine wave is another sine wave differing at most in amplitude and phase. An electric power distribution system is enormously complex, yet if the load is constant (system invariance) sine waves of the generated frequency appear at every outlet.

It is this cardinal fact which gives sine waves their unique position in communication theory and which gives physical significance to Fourier analysis and to the whole concept of frequency spectra. For if any input can be represented as the sum of a number of sinusoids, then the output will consist of these same sinusoids modified only in amplitude and phase, each in the same way as if it alone were present (superposition), and the simple sum of the output sinusoids will be the output wave.

The way in which the system modifies sine waves of all frequencies - the amplitude and phase characteristic - thus consititutes a complete description of the system in that it enables us to compute the output in response to any input. The procedure may be represented graphically as shown in Figure 1.

The spectrum of the input wave, if not already known, is found by evaluating the Fourier transform. The system modifies the input spectrum by its transmission (amplitude and phase) characteristic to give the spectrum of the output. The inverse Fourier transform of the output spectrum is the output time function. Now this seems like a long way around to get from f(t) to h(t). Why not take a short cut such as indicated by the dotted line? The answer is that like most short cuts this direct path is often more difficult and treacherous. It involves the evaluation of an integral known as a convolution integral. Thus:

$$h(t) = \int_{-\infty}^{\infty} f(\tau)\ \phi(t - \tau)\ d\tau. \qquad (2)$$

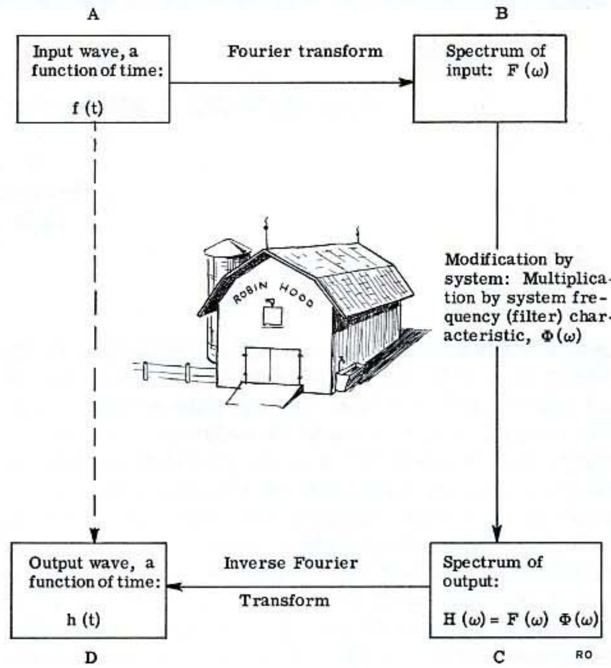


Figure 1

$\phi(\tau)$ is the inpulse response of the system so $\phi(t - \tau)$ is this same function reversed left to right and displaced by an amount t. What (2) says is that to find the output we have to integrate the product of the input and this reversed, displaced inpulse response. The result will be a function of the displacement, t. In other words, we have to SCAN the input wave with the reversed impulse response.

Equation (2) is very easy to derive if one considers the successive ordinates of f(t) to be a succession of impulses and applies superposition. It is sometimes called the superposition (or deHamel's) integral.

While (2) is often easy to evaluate, it is in general rather difficult. By contrast, rather complete tables of transforms exist so that getting from A to B and from C to D often involves merely using a table. The intermediate step from B to C is simple multiplication by the steady state transmission. The situation is rather analogous to the use of logarithms when raising a number to some strange power. One looks up the logarithm (A to B), multiplies by the power (B to C), looks up the antalog and arrives at D. How else would you compute 7^{π} ?

Even more dramatic is the case where one knows f(t) and h(t) and desires the impulse response. The direct solution involves solving (2) as an integral equation. By transforms we have simply

$$\Phi(\omega) = \frac{H(\omega)}{F(\omega)} .$$

$\phi(t)$ is then the inverse transform of $\Phi(\omega)$.

Another reason steady state (sine wave) measurements are important is that often the particular input, f(t), is not known. All that may be known is that all inputs will have spectra containing frequencies over a certain band (e.g. speech or music) and the problem may be to transmit this whole class of inputs without distortion. This requires that

$$h(t) = K f(t - t_o)$$

where K is a constant and t_o a permissible delay. In the frequency domain this requires that*

$$H(\omega) = K F(\omega) e^{-i\omega t_o} \qquad \omega_1 < \omega < \omega_2$$

and hence

$$\Phi(\omega) = K e^{-i\omega t_o} \qquad \omega_1 < \omega < \omega_2$$

Thus the steady state transmission must simply be flat in amplitude and have linear phase over the band of interest. The steady state requirements in other cases are also simply expressed.

The Response of Linear Systems to Inpulses

While steady state measurements are very useful for the reasons pointed out in the last section, they are quite time consuming and for many purposes observing the transient response of the system to certain particular types of input waves may provide all the information necessary. In fact if the input wave is properly chosen such a transient measurement provides exactly the same information as the steady state measurement, but in a different form.

Consider, for example, the case where the input, f(t), is an impulse of negligible duration and, say, unit area. The spectrum of such an impulse contains all frequencies at equal amplitude, and in phase in the sense that they all add at t = o. In other words the spectrum is a constant $F(\omega) = 1$. Now by superposition, it obviously makes no difference whether all frequencies are introduced one after the other as in steady state testing, or simultaneously by applying an impulse. The frequencies will all be modified in the same way by the linear system. You might say that an impulse test is equivalent to an instantaneous steady state test. They both give the same information. The problem is merely one of interpreting the results in either case.

When the input spectrum is $F(\omega) = 1$, the output spectrum $H(\omega) = \Phi(\omega)$. That is the output spectrum has same amplitude and phase at all frequencies as would be found by steady state measurement. What you see as the output is the time function which has this spectrum. The response of a linear network to an impulse is a pulse whose spectrum is the amplitude and phase characteristics of the network.

In the Table of Transformers which follows, a number of time functions and their spectra are illustrated. Assuming, as we have, that the impulse is applied at t = o, a physically realizable network can give no output for t < o.

Thus, those functions illustrated for which a response exists for t < o (pairs 3, 6, 7, 8, 9, 10, 12) can only be realized as network impulse responses by having enough delay in the network (enough additional linear phase in the phase characteristic) to shift the entire function to the right of t = 0.

Suppose you applied an impulse of current to a network and the time function of pair 2S appeared as a voltage output. What is the network? Well, its frequency characteristic is of the form constant x $\frac{1}{P}$. If you put in the dimensions you will find the constant must have the dimensions of (capacity) -1. Thus your network can be represented as an impedance of the form $\frac{1}{pc}$. A shunt condenser for example. It integrated the current impulse input to give the step function voltage out.

If instead of the response 2S you found the response 1, you would say that the condenser was being discharged by a parallel resistance. Such a combination would have an impedance

$$Z = \frac{1}{C} \quad \frac{1}{P + \frac{1}{RC}} .$$

Sure enough the frequency function is of this form with $a = \frac{1}{RC}$.

Similarly pair 4 could be obtained by applying a unit voltage impulse to a series L, shunt C half section. Pair 5 would result if a current impulse were applied to a parallel RLC combination and the voltage observed, or if a voltage impulse were applied to a series RLC circuit and the current observed.

Obviously the interpretation of impulse test involves largely a familiarity with the spectra associated with a wide variety of time functions and vice versa. Fortunately, for qualitative work it suffices to know

*See "Table of Important Transforms", Table of Properties Section.

a few key cases. Most responses can then be represented as the sum of one or more of these key cases. By superposition the frequency characteristic is then the sum of the component spectra. For example a response of the form $(1 + a\,t)\,e^{-a\,t}$ is the sum of uses 1 and 2 in the table and has the spectrum

$$\frac{1}{p + a} + \frac{a}{(p + a)^2} = \frac{p + 2a}{(p + a)^2}$$

Inpulse testing, while satisfactory or even preferable in certain instances, has two severe drawbacks.

1. The impulse must be short compared with the duration of the finest detail of the output transient which is to be reproduced accurately. Alternatively one may say that the spectrum of the input must be flat over the entire frequency band of the device under test. To get appreciable response then, often requires the impulse to be so large in amplitude that the device under test is driven out of its range of linear operation.

2. In testing wide band devices, the low end effects are hard to observe. This is because with a flat input spectrum the low end cutoff deletes an insignificant amount of the spectrum. For example pair 11 shows the impulse response of a resistance-coupled amplifier with simple 6 db/octave cutoff at the low end and high end. The low end cutoff produces the negative tail on the output pulse. The peak amplitude of this tail is less $\frac{\beta}{a - \beta}$ where β is the low end 3 db point, a the high end 3 db point. Obviously if $a > 100\,\beta$ as is typically the case, the tail is hard to see without blowing up the pattern to the point where the system overload can occur during the initial spike. It is principally for these two reasons that impulses are rarely used.

The Response of Linear Systems to Step Functions

The drawbacks of impulse tests are avoided by using step functions. The rise time of the step may be made as short as desired without requiring an increase in amplitude. The spectrum of a unit step is $\frac{1}{p}$. Because of the concentration of amplitude at the low end of the spectrum, low end and high end cutoff effects are placed on a more nearly equal footing.

A unit step is the integral of a unit impulse. Thus we might obtain the step function response of a system in any of the three ways indicated in Figure 2.

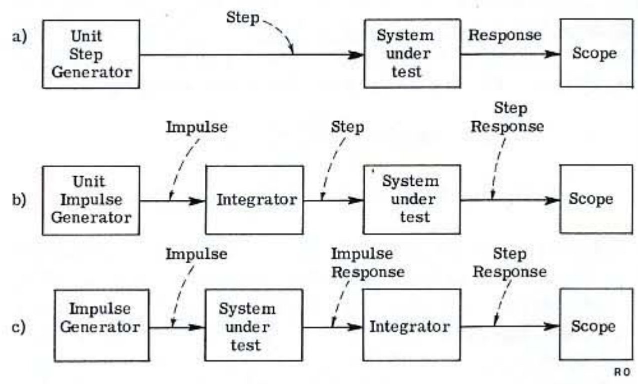


Figure 2

In (a) we have a straightforward test using a step function generator. In (b) the step generator is replaced by a combination of impulse generator and integrator. In (c) the integrator and the system under test have been interchanged. Since both are linear systems the output is unaffected. Case (c) illustrates the fact that: The step function response of a linear system is the integral of the impulse response. Its spectrum is $\frac{1}{p}$ times the steady state amplitude and phase characteristic of the system.

Conversely, of course, the impulse response is the derivative of the step response.

Consider pairs 1 and 4S in the table. If 1 represents the impulse response 4S is the step response. Similarly if 5 is the impulse response of a system 4 will be the step function response.

Just as steady state measurements are especially convenient when the signals to be handled by the system are defined only in terms of their spectra-their frequency band - so unit step testing is especially convenient in testing systems which are to handle signals characterized by having step-like transitions (such as TV signals). In both cases you might say each method is the natural one to use. And as a matter of fact, historically, sine wave testing came into prominence in the testing of audio and carrier telephone systems, step function testing came in along with the advent of TV and other pulse modulation systems.

Since step function testing has become so common we will devote the next section to a study of the effects on the unit step response produced by various typical frequency characteristics.

HEWLETT - PACKARD
COMPANY
PALO ALTO, CALIFORNIA

TABLE OF IMPORTANT TRANSFORMS

(See inside pages)

EXPLANATION OF THE TABLE

The time functions and corresponding frequency functions in this table are related by the following expressions:

$$F(\omega) = \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t}\, dt \qquad \text{(Direct transform)}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\, e^{i\omega t}\, d\omega \qquad \text{(Inverse transform)}$$

The $1/2\pi$ multiplier in the inverse transform arises merely because the integration is written with respect to ω, rather than cyclic frequency. Otherwise the expressions are identical except for the difference of sign in the exponent. As a result, functions and their transforms can be interchanged with only slight modification. Thus, if $F(\omega)$ is the direct transform of f(t), it is also true that $2\pi f(-\omega)$ is the direct transform of F(t). For example, the spectrum of a $\frac{\sin x}{x}$ pulse is rectangular (pair 6) while the spectrum of a rectangular pulse is of the form $\frac{\sin x}{x}$ (pair 7). Likewise pair 1S is the counterpart of the well-known fact that the spectrum of a constant (d–c) is a spike at zero frequency.

The frequency functions in the table are in many cases listed both as functions of ω and also of p. This is done merely for convenience. F(p) in all cases is found by substituting p for $i\omega$ in $F(\omega)$. (Not simply p for ω as the notation would ordinarily indicate. That is, in the usual mathematical convention one would write $F(\omega) = F(\frac{p}{i}) = G(p)$ where the change in letter indicates the resulting change in functional form. The notation used above has grown through usage and causes no confusion, once understood.) Thus, in the p-notation

$$F(p) = \int_{-\infty}^{\infty} f(t)\, e^{-pt}\, dt \qquad f(t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} F(p)\, e^{pt}\, dp$$

The latter integral is conveniently evaluated as a contour integral in the p-plane, letting p assume complex values.

The frequency functions have been plotted on linear amplitude and frequency scales, and where convenient, also on logarithmic scales. The latter scales often bring out characteristics not evident in the linear plot. Thus, many of the spectra are asymptotic to first or second degree hyperbolas on a linear plot. On a log plot these asymptotes become straight lines of slope –1 or –2 (i.e., –6 or –12 db/octave).

The time functions in the table have all been normalized to convenient peak amplitudes, areas or slopes. For any other amplitude, multiply both sides by the appropriate factor. Thus, the spectrum of a rectangular pulse 10 volts in amplitude and 2 seconds long is (from pair 7) $20 \frac{\sin \omega}{\omega}$ volt-seconds.

Again, upon multiplication by a constant having appropriate dimensions, the frequency functions become filter transmissions. Thus, if pair 1 is multiplied by α, the frequency function represents a simple RC cutoff. A one coulomb impulse (pair 1S) applied to this filter would produce an output (impulse response) with the spectrum $\frac{\alpha}{p+\alpha} \times 1$ coulomb, representing the time function $\alpha e^{-\alpha t}$ coulombs (which has the dimensions of amperes). Or a 1 volt step function (pair 2S) would produce the output spectrum $\frac{\alpha}{p+\alpha} \times \frac{1}{p}$ volts, which represents the time function $(1-e^{-\alpha t})$ volts (pair 4S).

The entries 1S through 6S in the table are singular functions for which the transforms as defined above exist only as a limit. For example, 1S may be thought of as the limit of pair 7 (multiplied by $\frac{1}{\tau}$) as $\tau \longrightarrow 0$.

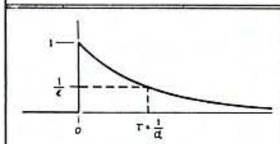
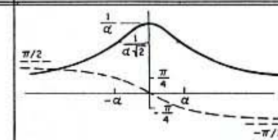
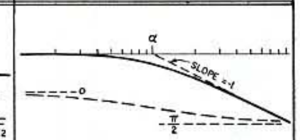
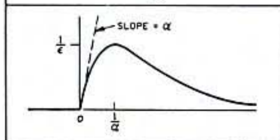
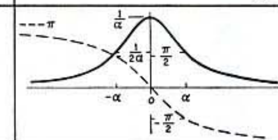
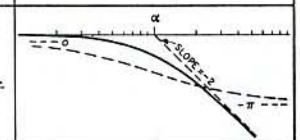
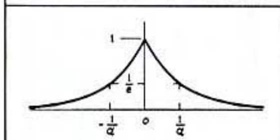
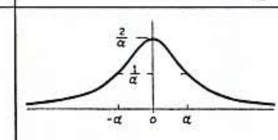
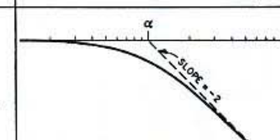
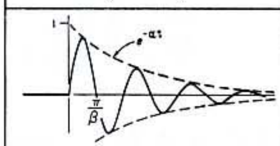
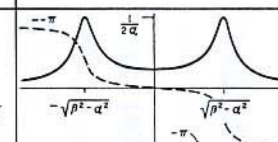
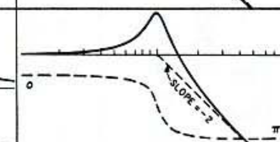
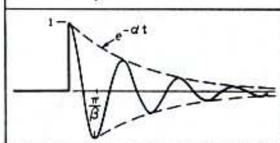
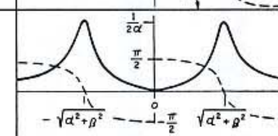
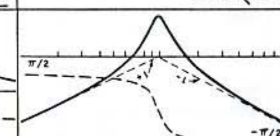
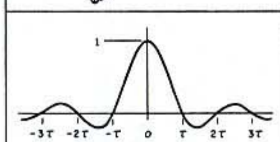
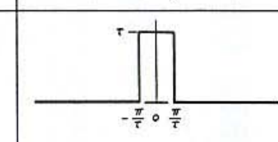
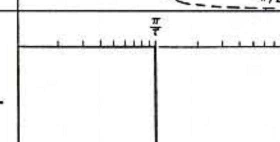
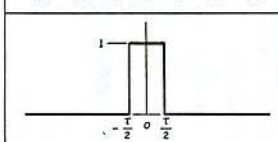
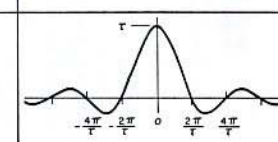
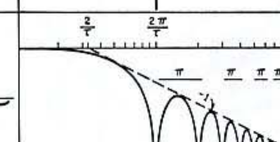
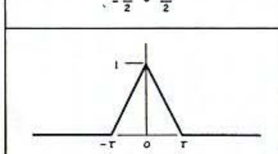
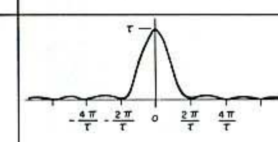
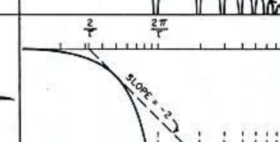
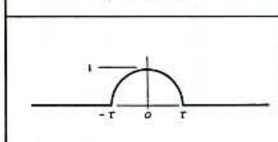
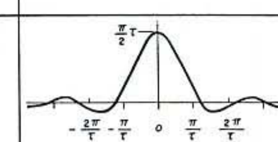
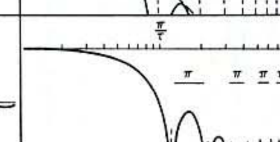
PROPERTIES OF TRANSFORMS

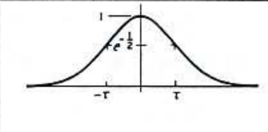
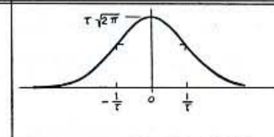
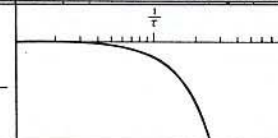
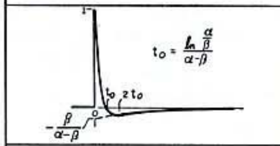
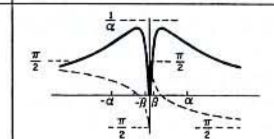
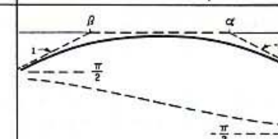
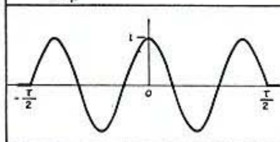
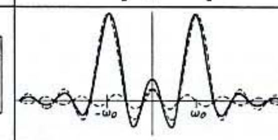
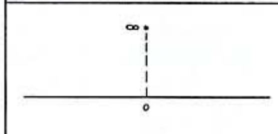
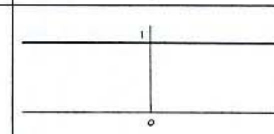
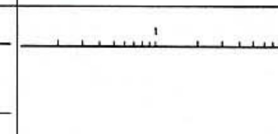
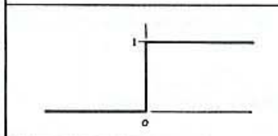
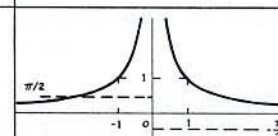
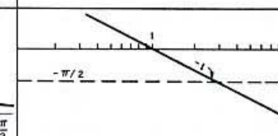
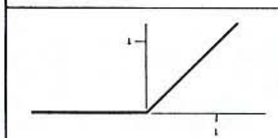
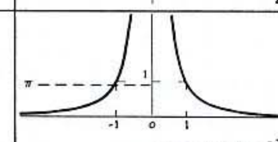
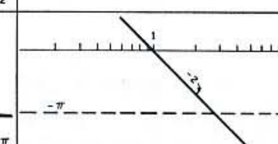
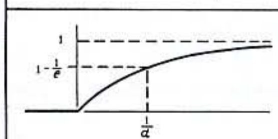
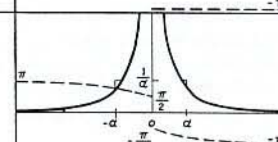
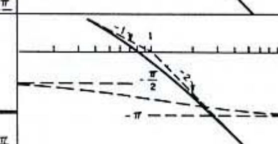
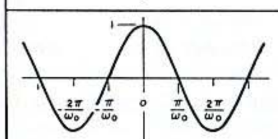
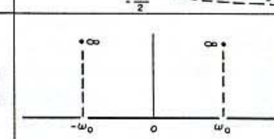
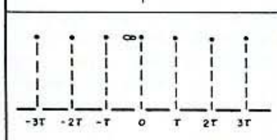
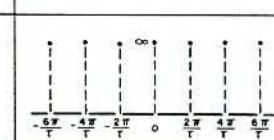
There are a number of important relations which describe what happens to the transforms of functions when the functions themselves are added, multiplied, convolved, etc. These relations state mathematically many of the operations encountered in communications systems: operations such as linear amplification, mixing, modulation, filtering, sampling, etc. These relations are all readily deducible from the defining equations above; but for ready reference some of the more important ones are listed in the Table of Properties (back page).

Again, because of the similarity of the direct and inverse transforms, a symmetry exists in these properties. Thus, delaying a function multiplies its spectrum by a complex exponential; while multiplying the function by a complex exponential delays its spectrum. Multiplying any two functions is

(Continued on back page)

TABLE OF IMPORTANT TRANSFORMS

TIME FUNCTIONS		NO.	FREQUENCY FUNCTIONS	(LINEAR SCALES)	(LOG AMPL. - LOG FREQ.)
	$f(t)=\begin{cases}0, & t<0\\ e^{-\alpha t}, & t>0\end{cases}$	1	$F(p)=\frac{1}{p+\alpha}$ $F(\omega)=\frac{1}{\alpha+i\omega}$		
	$f(t)=\begin{cases}0, & t<0\\ \alpha t e^{-\alpha t}, & t>0\end{cases}$	2	$F(p)=\frac{\alpha}{(p+\alpha)^2}$ $F(\omega)=\frac{\alpha}{(\alpha+i\omega)^2}$		
	$f(t)=e^{-\alpha\vert t\vert}$	3	$F(p)=\frac{2\alpha}{\alpha^2-p^2}$ $F(\omega)=\frac{2\alpha}{\alpha^2+\omega^2}$		
	$f(t)=\begin{cases}0, & t<0\\ e^{-\alpha t}\sin\beta t, & t>0\end{cases}$	4	$F(p)=\frac{\beta}{(p+\alpha)^2+\beta^2}$ $F(\omega)=\frac{\beta}{(\alpha^2+\beta^2)-\omega^2+i2\alpha\omega}$		
	$f(t)=\begin{cases}0, & t<0\\ e^{-\alpha t}(\cos\beta t-\frac{\alpha}{\beta}\sin\beta t), & t>0\end{cases}$	5	$F(p)=\frac{p}{(p+\alpha)^2+\beta^2}$ $F(\omega)=\frac{i\omega}{(\alpha^2+\beta^2)-\omega^2+i2\alpha\omega}$		
	$f(t)=\frac{\sin(\pi\frac{t}{\tau})}{(\pi\frac{t}{\tau})}$	6	$F(\omega)=\begin{cases}\tau, & \vert\omega\vert<\frac{\pi}{\tau}\\ 0, & \vert\omega\vert>\frac{\pi}{\tau}\end{cases}$		
	$f(t)=\begin{cases}1, & \vert t\vert<\frac{\tau}{2}\\ 0, & \vert t\vert>\frac{\tau}{2}\end{cases}$	7	$F(\omega)=\tau\,\frac{\sin(\frac{\omega\tau}{2})}{(\frac{\omega\tau}{2})}$		
	$f(t)=\begin{cases}1-\frac{\vert t\vert}{\tau}, & \vert t\vert<\tau\\ 0, & \vert t\vert>\tau\end{cases}$	8	$F(\omega)=\tau\,\frac{\sin^2(\frac{\omega\tau}{2})}{(\frac{\omega\tau}{2})^2}$		
	$f(t)=\begin{cases}\sqrt{1-(\frac{t}{\tau})^2}, & \vert t\vert<\tau\\ 0, & \vert t\vert>\tau\end{cases}$	9	$F(\omega)=\frac{\pi}{2}\,\tau\,\frac{2J_1(\omega\tau)}{(\omega\tau)}$		

TIME FUNCTIONS		NO.	FREQUENCY FUNCTIONS	(LINEAR SCALES)	(LOG AMPL. - LOG FREQ.)
	$f(t) = e^{-\frac{1}{2}\left(\frac{t}{\tau}\right)^2}$	10	$F(\omega) = \tau\sqrt{2\pi}\; e^{-\frac{1}{2}(\tau\omega)^2}$		
	$f(t) = \begin{cases} 0 & , \lvert t\rvert < 0 \\ \dfrac{\alpha e^{-\alpha t} - \beta e^{-\beta t}}{\alpha - \beta} & , \lvert t\rvert > 0 \end{cases}$	11	$F(p) = \dfrac{p}{(p+\alpha)(p+\beta)}$		
	$f(t) = \begin{cases} \cos \omega_0 t, & \lvert t\rvert < \frac{\tau}{2} \\ 0 & , \lvert t\rvert > \frac{\tau}{2} \end{cases}$	12	$F(\omega) = \frac{\tau}{2}\left[\dfrac{\sin\left(\frac{\omega-\omega_0}{2}\right)\tau}{\left(\frac{\omega-\omega_0}{2}\right)\tau} + \dfrac{\sin\left(\frac{\omega+\omega_0}{2}\right)\tau}{\left(\frac{\omega+\omega_0}{2}\right)\tau}\right]$		
	$f(t) = \lim_{\tau \to 0} \begin{cases} \frac{1}{\tau}, & \lvert t\rvert < \frac{\tau}{2} \\ 0, & \lvert t\rvert > \frac{\tau}{2} \end{cases}$ $= \delta(t)$ (DELTA FUNCTION)	1S	$F(p) = F(\omega) = 1$		
	$f(t) = \int_{-\infty}^{t} \delta(\lambda)d\lambda = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases}$ $= u(t)$ (UNIT STEP)	2S	$F(p) = \dfrac{1}{p}$		
	$f(t) = \int_{-\infty}^{t} u(\lambda)d\lambda = \begin{cases} 0, & t < 0 \\ t, & t > 0 \end{cases}$ $= s(t)$ (UNIT SLOPE)	3S	$F(p) = \dfrac{1}{p^2}$		
	$f(t) = \begin{cases} 0 & , t < 0 \\ 1 - e^{-\alpha t}, & t > 0 \end{cases}$	4S	$F(p) = \dfrac{\alpha}{p(p+\alpha)}$		
	$f(t) = \cos \omega_0 t$	5S	$F(\omega) = \pi\left[\delta(\omega+\omega_0) + \delta(\omega-\omega_0)\right]$		
	$f(t) = \sum_{-\infty}^{\infty} \delta(t - n\tau)$	6S	$F(\omega) = \frac{2\pi}{\tau} \sum_{-\infty}^{\infty} \delta\left(\omega - n\frac{2\pi}{\tau}\right)$		

PROPERTIES OF TRANSFORMS

(Continued from first page)

equivalent to convolving their spectra; multiplying their spectra is equivalent to convolving the functions; etc.

Many of the pairs listed in the Table of Transforms can be obtained from others by using one or more of the rules of manipulation listed in the Table of Properties. For example, the time function in pair 8 is $\frac{1}{\tau}$ times the *convolution* of that in pair 7 with itself. The spectrum should therefore be $\frac{1}{\tau}$ times the *product* of that in pair 7 with itself, as it indeed is. Further, by using these properties, many pairs *not* in the table can be obtained from those given. For example, the spectrum of $f(t) = (1 - \alpha t)\, e^{-\alpha t}$ is (by the addition property) $F(p) = \frac{1}{p+\alpha} - \frac{\alpha}{(p+\alpha)^2} = \frac{p}{(p+\alpha)^2}$.

TIME OPERATION	FREQ. OPERATION	SIGNIFICANCE
LINEAR ADDITION $af(t) + bg(t)$	**LINEAR ADDITION** $aF(\omega) + bG(\omega)$	Linearity and superposition apply in both domains. The spectrum of a linear sum of functions is the same linear sum of their spectra (if spectra are complex, usual rules of addition of complex quantities apply). Further, any function may be regarded as a sum of component parts and the spectrum is the sum of the component spectra.
SCALE CHANGE $f(kt)$	**INVERSE SCALE CHANGE** $\frac{1}{\lvert k \rvert} F\left(\frac{\omega}{k}\right)$	Time—Bandwidth invariance. Compressing a time function expands its spectrum in frequency and reduces it in amplitude by the same factor. The amplitude reduces because less energy is spread over a greater bandwidth. For same energy pulse as for $k=1$, multiply both functions by $\sqrt{\lvert k \rvert}$. The case where $k=-1$ reverses the function in time. This merely interchanges positive and negative frequencies; so for real time functions, reverses the phase.
EVEN AND ODD PARTITION $\frac{1}{2}\left[f(t) \pm f(-t)\right]$	**EVEN AND ODD PARTITION** $\frac{1}{2}\left[F(\omega) \pm F(-\omega)\right]$	Any real function f(t) may be separated into an even part $\frac{1}{2}\left[f(t) + f(-t)\right]$ and an odd part $\frac{1}{2}\left[f(t) - f(-t)\right]$. The transform of the even part is $\frac{1}{2}\left[F(\omega) + F(-\omega)\right]$ which is purely real and involves only even powers of ω. The transform of the odd part is $\frac{1}{2}\left[F(\omega) - F(-\omega)\right]$ which is purely imaginary and involves only odd powers of ω. Note: for f(t) real, $F(-\omega) = \overline{F(\omega)}$.
DELAY $f(t-t_o)$	**LINEAR ADDED PHASE** $e^{-i\omega t_o} F(\omega)$	Delaying a function by a time t_o multiplies its spectrum by $e^{-i\omega t_o}$, thus adding a linear phase $\theta = -\omega t_o$ to the original phase. Conversely a linear phase filter produces a delay of $-\frac{d\theta}{d\omega} = t_o$.
COMPLEX MODULATION $e^{i\omega_o t} f(t)$	**SHIFT OF SPECTRUM** $F(\omega - \omega_o)$	Multiplying a time function by $e^{i\omega_o t}$ "delays" its spectrum, i.e., shifts it to center about ω_o rather than zero frequency. Ordinary real modulation—by cos $\omega_o t$ say—produces the time function $\frac{1}{2}(e^{i\omega_o t} + e^{-i\omega_o t})f(t)$ with the spectrum $\frac{1}{2}\left[F(\omega - \omega_o) + F(\omega + \omega_o)\right]$.
CONVOLUTION $\int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$	**MULTIPLICATION (FILTERING)** $F(\omega)G(\omega)$	The spectrum of the convolution of two time functions is the product of their spectra. In convolution one of the two functions to be convolved is reversed left-to-right and displaced. The integral of the product is then evaluated and is a new function of the displacement. Convolution occurs whenever a signal is obtained which is proportional to the integral of the product of two functions as they slide past each other—in other words, in any *scanning* operation such as in optical or magnetic recording or picture scanning in television. Transform theory states that such scanning is equivalent to filtering the signal with a filter whose transmission is the transform of the scanning function (reversed in time). Conversely, the effect of an electrical filter is equivalent to a convolution of the input with a time function which is the transform of filter characteristic. This function, the so-called "memory curve" of the filter, is identical with the filter impulse response, aside from dimensions. (Note: the convolution of a time function with a unit impulse gives the same function times the dimensions of the impulse.)
MULTIPLICATION $f(t)g(t)$	**CONVOLUTION** $\frac{1}{2\pi}\int_{-\infty}^{\infty} F(s)G(\omega - s)ds$	The spectrum of the product of two time functions is the convolution of their spectra. This is the more general statement of the modulation property. For example, sampling a signal is equivalent to multiplying it by a regular train of unit area impulses. The spectrum of the sampled signal consists of the original signal spectrum repeated about each component of the (line) spectrum of the train of impulses (see pair 6S). For no overlap, highest frequency in signal to be sampled must be less than half sampling frequency. If this is true original signal spectrum (hence signal) can be recovered by low pass filter (Sampling theorem).
DIFFERENTIATION $\frac{d^n f(t)}{dt^n}$	**MULTIPLICATION BY p** $p^n F(p)$	The spectrum of the nth derivative of a function is $(i\omega)^n$ times the spectrum of the function. A "differentiating network" has (over the appropriate frequency range) a transmission $K\frac{p}{\omega_o}$ where K is dimensionless or has the dimensions of impedance or admittance. Thus the output wave is *proportional* to the derivative of the input.
INTEGRATION $\underbrace{\int_{-\infty}^{t} \cdots \int_{-\infty}^{t}}_{n} f(\tau)(d\tau)^n$	**MULTIPLICATION BY $\frac{1}{p}$** $\frac{1}{p^n} F(p)$	The spectrum of the nth integral of a function is $(i\omega)^{-n}$ times the spectrum of the function. Thus, the response of any filter to a step function is the integral of its impulse response. An "integrating network" has (over the appropriate frequency range) a transmission $K\frac{\omega_o}{p}$, where K is dimensionless or has the dimensions of impedance or admittance. Thus, the output is *proportional* to the integral of the past of the input.

Typical Step Function Responses

The table which follows shows some typical step function responses together with the amplitude or frequency characteristics of the systems which have these responses. A few explanatory remarks about each case shown in the table are in order.

Case 1. This is the typical simple low-frequency cutoff such as might be produced by a series condenser-shunt resistor combination. The step response shows an abrupt rise to unity followed by an exponential decay. Usually encountered in amplifier interstages and so-called "differentiating networks". In interstages f_0 is typically a few cycles, in differentiating networks f_0 may be as high as several megacycles in which case the step response is very nearly an impulse.

Case 2. Rising simple step in the frequency characteristic. Step response rises initially to amplitude determined by high frequency transmission, falls exponentially to level determined by low frequency (or dc) transmission. This is commonly encountered in improperly compensated resistance-capacity dividers, such as scope probes, dc amplifier interstages.

Case 3. The counterpart of case 2. Here it is the high frequency transmission which is deficient.

Case 4. Typical simple high-frequency cutoff such as is produced by a parallel RC combination. The step response rises exponentially to the final value determined by the low frequency (or dc) transmission. Commonly encountered in simple (not "peaked") interstages, and wherever shunt capacity (as from connection cables) loads down a resistive source.

Case 5. Two simple RC high frequency cutoffs in tandem. Typical rise characteristic of two-stage resistance coupled amplifier without "peaking". Prinicpal differences compared with case 4:

1. Greater rise time for same ω_0.

2. Zero slope at $t = o$.

For each additional high frequency cutoff one more derivative of step response vanishes at $t = o$. Thus if high frequency transmission falls (ultimately) at 6 n db/octave, all derivatives of step response up to the nth are zero at $t = o$.

Case 6. Phase-compensated low end cutoff. Step function response falls to zero eventually, but initial slope is zero. As a result square wave response shows little or no tilt. May be produced in a single network, or by two networks (cases 1 and 3) in tandem. Often found in video amplifiers.

Case 7. Two simple low frequency cutoffs (case 1) in tandem. Typical low frequency transient response of single stage resistance coupled amplifier with input blocking condenser or two stage amplifier with no input blocking condenser. Principal differences compared with case 1:

1. Faster initial rate of fall for same ω_0.

2. Response goes negative crossing axis at $t = \frac{1}{\omega_0}$

With each additional low-end cutoff one additional axis crossing is produced. Thus if the low end response falls off (ultimately) at 6 n db/octave, there will be n - 1 axis crossings. They do not occur at regular intervals - each successive half cycle takes longer. All step function responses produced by n similar simple low frequency cutoff are members of the family of LaGuerre functions.

Case 8. Simple high and low frequency cutoff. The step response rises exponentially at a rate determined by high frequency cutoff, then falls exponentially at a rate determined by low frequency cutoff. Typical complete resistance coupled interstage response. If

$$\frac{\omega_2}{\omega_1} >> 1,$$

then on a slow time scale response looks like case 1, on a fast time scale response looks like case 4. If $\omega_2 = \omega_1 = \omega_0$, we have the case of a critical damped RLC circuit, the response then becomes

$$\omega_0 \, t \, e^{-\omega_0 t}.$$

Case 9. Typical damped oscillation. Exact response shown is current in series RLC circuit in response to series voltage step, or voltage across parallel RLC circuit in response to applied current step. The dotted lines in the frequency characteristic are the asymptotes which the actual characteristic approaches for $\frac{\omega}{\omega_0} << 1$ and $\frac{\omega}{\omega_0} >> 1$.

The peak of the resonance curve is Q times as high as the intersection of these asymptotes. For reasonable Q's, such that $\beta \approx \omega_0$, the Q of circuit may be readily found from the fact that the envelope of the oscillation decays to $\frac{1}{\epsilon}$ in $\frac{Q}{\pi}$ cycles. Thus $Q = \pi n$ where n is the number of cycles to the $\frac{1}{\epsilon}$ point.

Case 10. Small resonance in an otherwise flat characteristic. Response consists of unit step due to flat transmission plus damped oscillation due to resonance - simple superposition. Initial amplitude of oscillation is related to amplitude of hump in frequency characteristic as indicated in figure. For the same amplitude of hump, increasing the Q decreases amplitude of oscillation but oscillation persists longer. If hump is near top of band, time scale will be such that initial rise of response will not appear so abrupt, but will blend with oscillation to give response like that of over-peaked interstage. Mid-band resonances such as shown in case 10 often occur as a result of stray feedback paths such as heater leads, or from attempting to bypass electrolytics with small mica condensers. (Electrolytics become inductive at high frequencies.)

Case 11. Similar to case 10 but here we have a resonant dip. Note that the effect of a complete null ($\blacktriangle = 1$) is no worse than that of a 6 db hump. The pilot separation filters used in the coaxial television system produce this type of dip - a complete null. Because then Q is so high (several thousand), the disturbance they produce, while it persists for a long time, is of such low amplitude as to be invisible in the picture.

Case 12. Positive echo. Associated frequency characteristic has nearly sinusoidal ripple in amplitude and phase. Frequency interval between successive maxima or minima is reciprocal of echo delay. The longer the delay, the closer the ripples. Commonly encountered in systems having faulty or misterminated delay lines. Also in measurements where multipath transmissions can exist - such as acoustic measurements. The reason most speaker characteristics look so ragged (fine structure) is that multipath reflections with long delay were present.

Case 13. Negative echo. Same frequency ripples as in case 11 but reversed 180°. Dc transmission is now $1 - \epsilon$ rather than $1 + \epsilon$.

Case 14. Rectangular pulse response. Can be considered to be a 100% negative echo. Minima of frequency ripples have now become nulls. Shape of amplitude characteristic is that of rectified sine wave. Phase characteristic is sawtooth decreasing from $\frac{\pi}{2}$ linearly to $-\frac{\pi}{2}$ and jumping back to $\frac{\pi}{2}$ at each null. Such a characteristic can be obtained by using a delay line as an interstage with the near end terminated and the far end shorted.

Case 15. "Differentiated Echo". This is the sort of disturbance produced when a delay line is terminated in such a way that the reflection coefficient increases with frequency. Typical causes are:

1. Series inductance or shunt capacitance in the termination of a smooth line.
2. Termination of a constant-k filter in simple resistances.

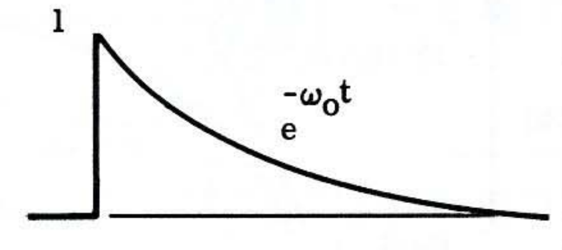
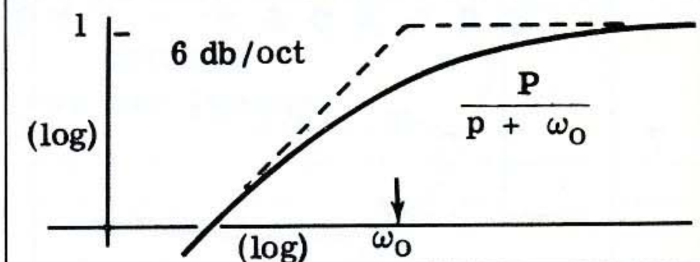
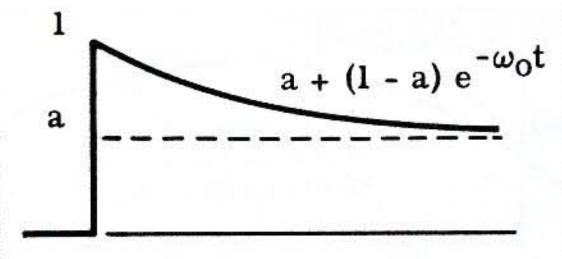
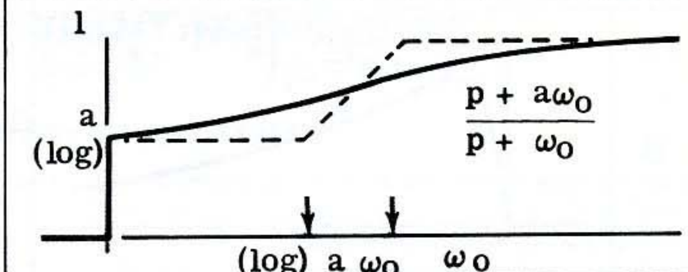
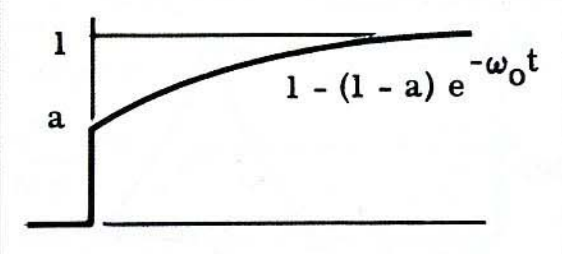
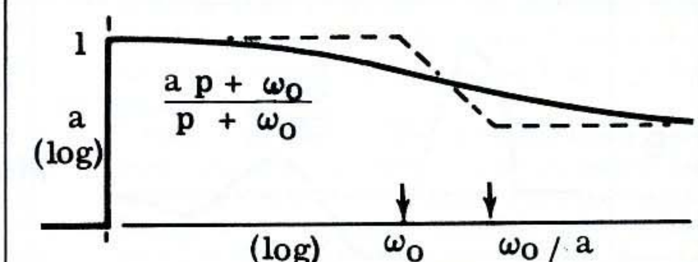
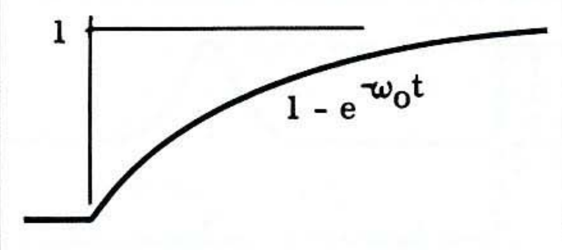
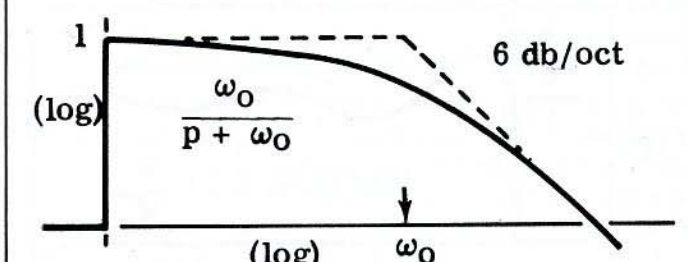
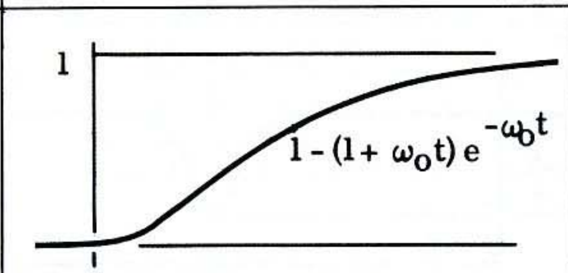
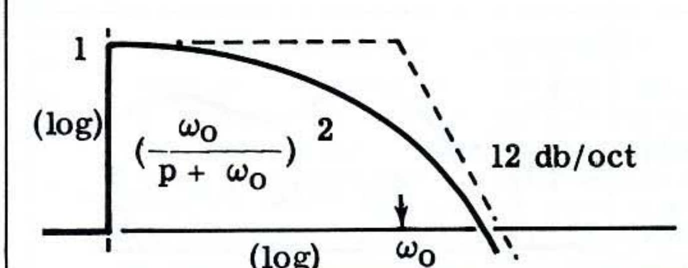
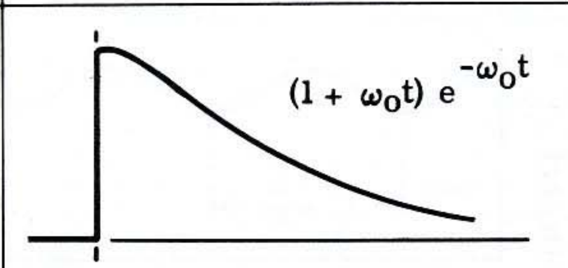
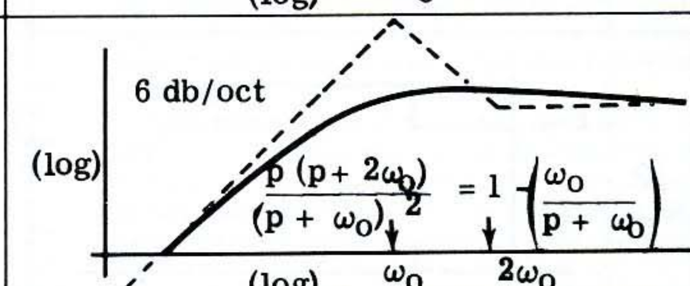
With both ends matched at low frequencies the transmitted echo involves two reflections both of which increase with frequency and so tends to be "double differentiated" and smaller.

Case 16. Rise characteristic (qualitative only) of a low pass filter without phase correction. The initial part of the rise ($t = 0 + \epsilon$) is a high power of time, the exponent depending on the number of sections. Following the rise there is a ripple whose period is not constant but approaches $2\tau = \frac{1}{f_0}$. That is, the apparent frequency approaches the cutoff frequency after several cycles. With an increasing number of sections this ripple increases in amplitude and duration. The "ringing sound" so often attributed to sharp cutoff filters is not due to exaggeration of frequencies near cutoff (there is no) nor to the sharp cutoff per se, but rather to the delay distortion which exists near cutoff causing those upper frequencies which are passed to arrive too late and thus be separately audible. The effect is noticeable only in extreme cases as for example long telephone circuits with many channel filters in tandem. With proper delay equalization the effect disappears.

Case 17. The "ideal" low pass filter passes all frequencies below f_0 with the same amplitude and delay while attenuating completely those above f_0. Its step response is the sine integral, i.e. the $\int_0^t \frac{\sin x}{x} dx$. This function differs from zero (except at regular points) for all $t > -\infty$. Hence the ideal filter cannot be realized without infinite delay. A practical approximation will have a finite delay and its step response therefore will execute only a finite number of wiggles before the main rise. The approximation can be quite good, however. Here again, the ripples in the step response do not indicate high frequency enhancement, but are the "Gibb's effect" encountered in Fourier series, and are properly called band elimination ripples. The rise time from the last zero crossing to the first crossing of the final amplitude level is $\frac{1}{2f_0}$: one half cycle of the cutoff frequency.

Case 18. The ideal high pass filter. By superposition the response of this filter is obtained by subtracting the response of the ideal low pass filter from an equally delayed unit step.

Table of Step Function Responses

No.	Step Response	System Frequency Characteristic
1	 $e^{-\omega_o t}$	 $\frac{P}{p + \omega_o}$; 6 db/oct ; (log) ; ω_o
2	 $a + (1 - a)\, e^{-\omega_o t}$	 $\frac{p + a\omega_o}{p + \omega_o}$; (log) ; $a\,\omega_o$; ω_o
3	 $1 - (1 - a)\, e^{-\omega_o t}$	 $\frac{a\,p + \omega_o}{p + \omega_o}$; (log) ; ω_o ; ω_o / a
4	 $1 - e^{-\omega_o t}$	 $\frac{\omega_o}{p + \omega_o}$; 6 db/oct ; (log) ; ω_o
5	 $1 - (1 + \omega_o t)\, e^{-\omega_o t}$	 $\left(\frac{\omega_o}{p + \omega_o}\right)^2$; 12 db/oct ; (log) ; ω_o
6	 $(1 + \omega_o t)\, e^{-\omega_o t}$	 $\frac{p\,(p + 2\omega_o)}{(p + \omega_o)^2} = 1 - \left(\frac{\omega_o}{p + \omega_o}\right)^2$; 6 db/oct ; (log) ; ω_o ; $2\omega_o$

G-S-62

Table of Step Function Responses (Cont'd)

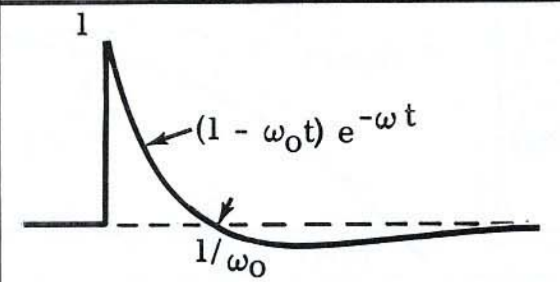
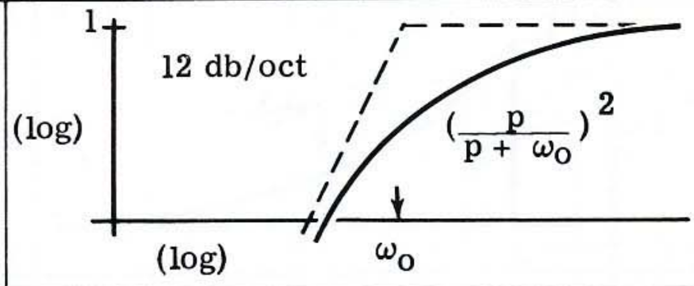
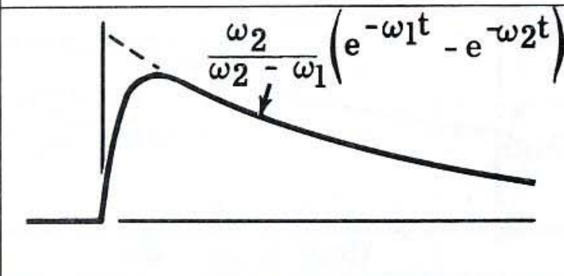
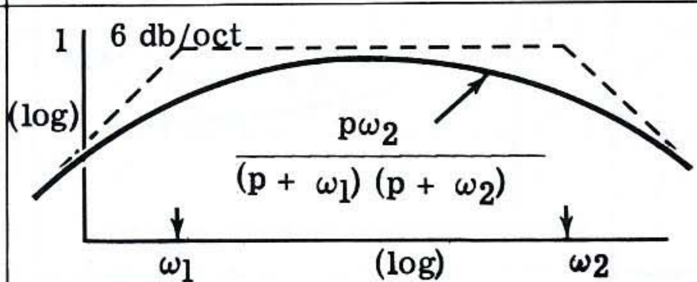
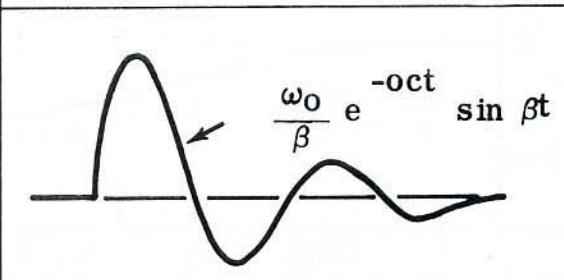
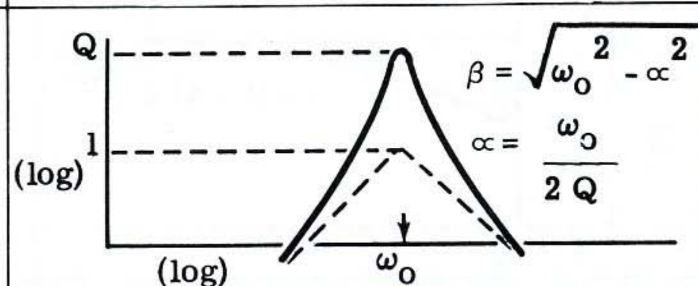
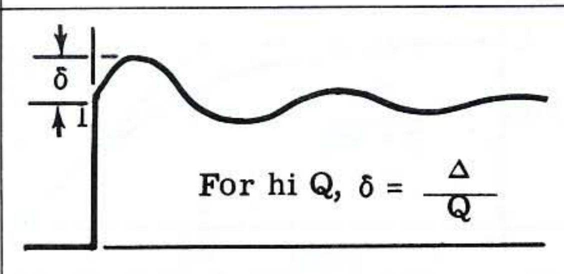
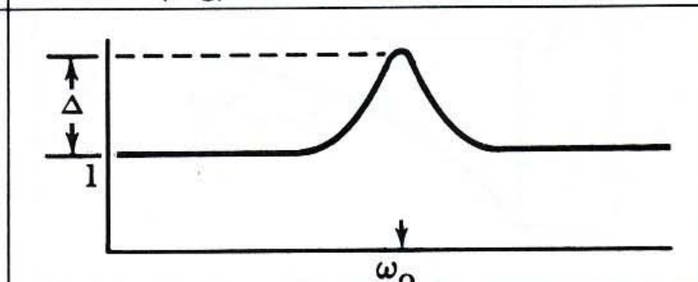
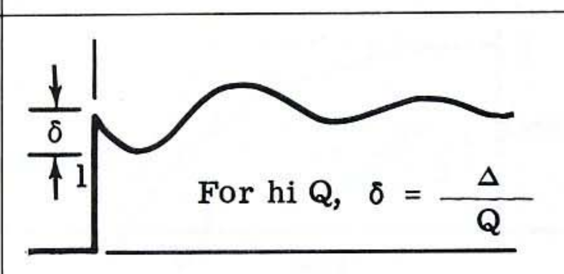
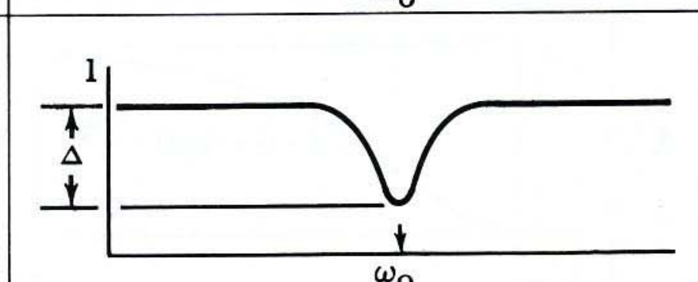
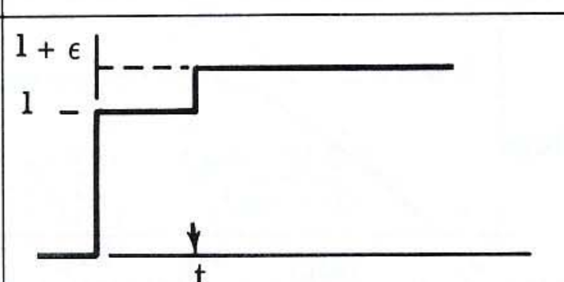
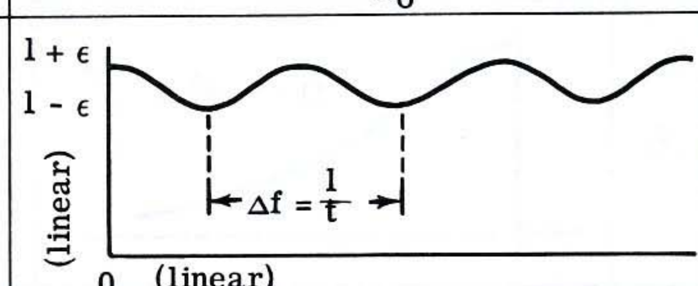
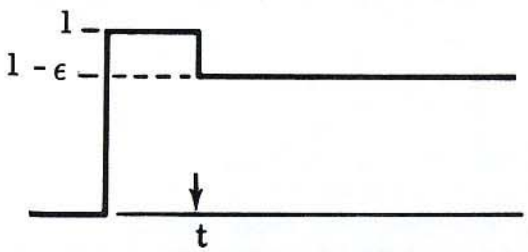
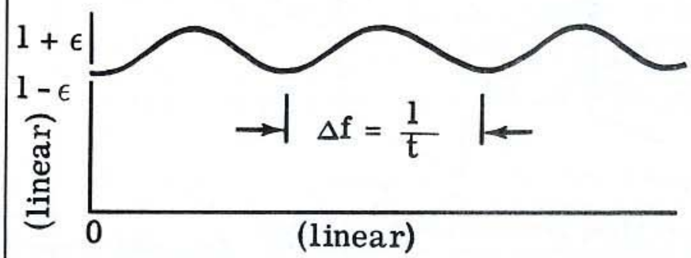
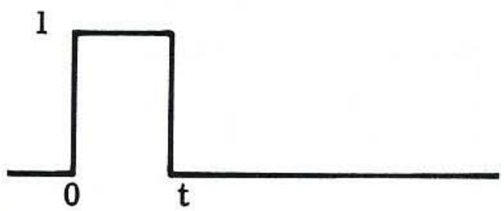
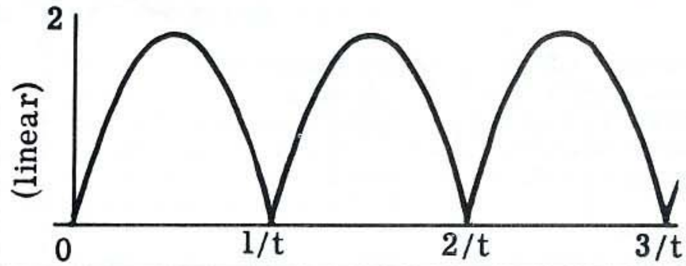
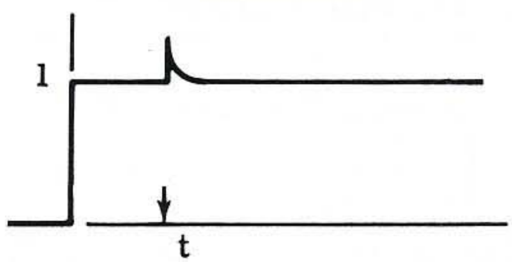
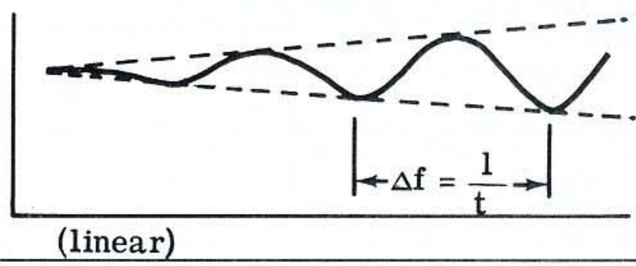
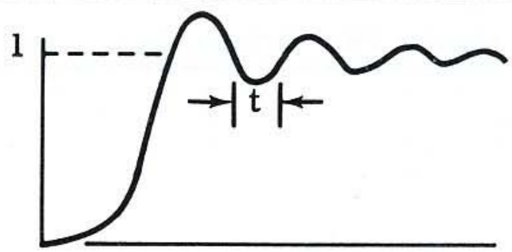
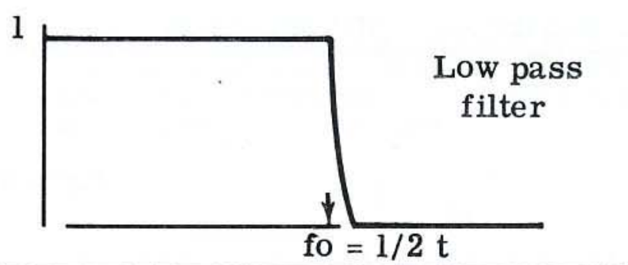
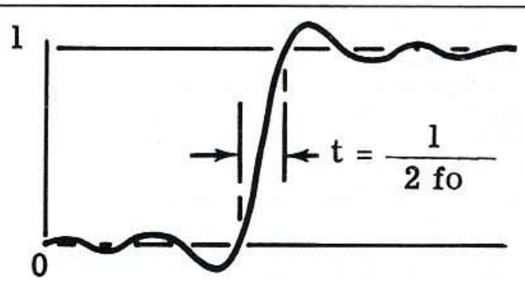
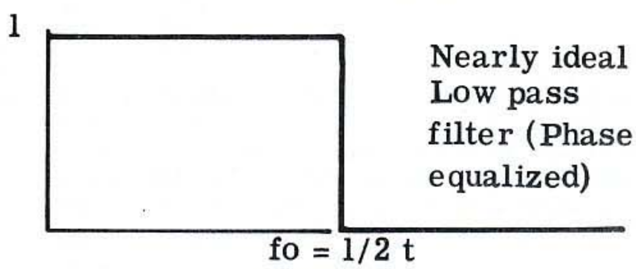
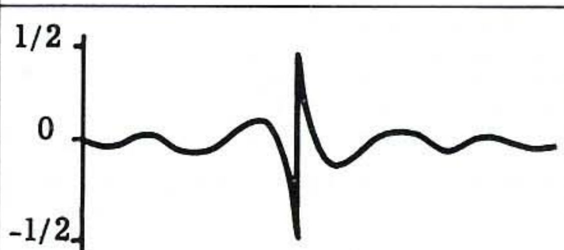
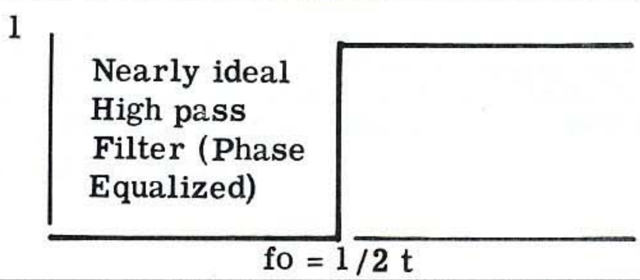
No.	Step Response	System Frequency Characteristic
7		
8		
9		
10		
11		
12		

Table of Step Function Responses (Cont'd)

<table>
<tr><th>No.</th><th>Step Response</th><th>System Frequency Characteristic</th></tr>
<tr><td>13</td><td>
</td><td>
</td></tr>
<tr><td>14</td><td>
</td><td>
</td></tr>
<tr><td>15</td><td>
</td><td>
</td></tr>
<tr><td>16</td><td>
</td><td>
</td></tr>
<tr><td>17</td><td>
</td><td>
</td></tr>
<tr><td>18</td><td>
</td><td>
</td></tr>
</table>

G-S-64

LOW END DISTORTION OF SQUARE WAVE

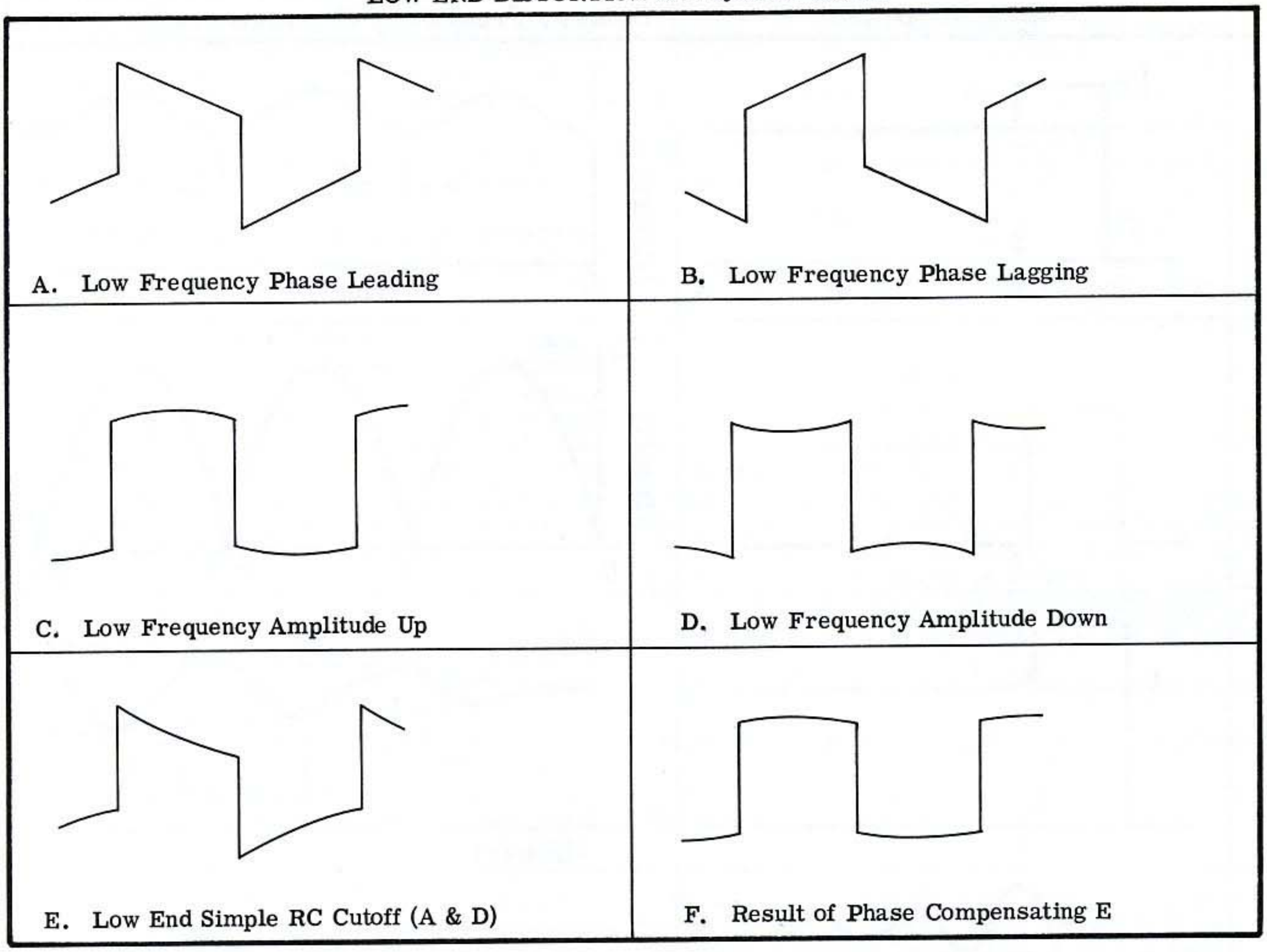


Figure 3

Square Wave and Pulse Testing

If anyone is still listening, we will now discuss what we started out to discuss. Actually, we have already done so, though we didn't say so at the time. Because, for high end effects square wave testing is equivalent to testing with unit steps, while pulse testing is either equivalent to impulse testing (if the pulse is much shorter than the rise time of the response) or to unit step testing (if the pulses are rectangular and of longer duration than the rise time). Obviously, if the entire transient is over before the next step occurs, either pulse or square waves can be considered as merely successions of unit steps alternating in direction. So far as high end effects in wide band circuits are concerned, this will be true for reasonably short period square waves and reasonably short pulses.

With regard to mid-band effects the same remarks apply to both pulses and square waves, while for low end effects short duty cycle pulses are of little use. So let's just talk about square waves from here on in. *

Whereas the spectrum of a unit step contains all frequencies, the spectrum of a square wave contains only certain discrete frequencies: the fundamental and all odd harmonics. The law of fall-off with frequency is the same, but the energy is concentrated into spectral lines leaving completely empty gaps below the fundamental and between harmonics. As a result, if only one frequency of square wave is used, the entire steady state transmission of the system is not measured; only the transmission at a discrete set of frequencies. A sharp mid-band irregularity such as depicted in cases 10 and 11 above could lie entirely between two harmonics and thus escape notice. To avoid this the frequency of the square wave must be varied, thus causing the harmonics to sweep through all parts of the spectrum, or else so low a frequency must be used that the spectrum is covered with sufficient density.

At the low end, there is no component (save perhaps dc) below the fundamental. So it must be possible to reduce the fundamental frequency far enough to place it below or at least in the region of the low end cutoff. Many ac amplifiers have cutoffs so low that it is impractical or undesirable to drop the square wave frequency to the point where each low end transient has died out before the next transient occurs. Failure of persistence of vision, for example, slows the observation. In such cases, tests are commonly made with a fundamental frequency such that considerable transient overlap occurs. Actually the necessary information can usually be gained from such a test particularly if the system under test is not required to handle frequencies below the fundamental square wave frequency used.

Typically what is desired is that the system under test exhibit flat amplitude and linear phase characteristics down to some frequency. In cases where the system phase is of concern, the usual reason is, in fact, that the system will in use be required to handle square waves or pulses. Thus if we choose this lowest frequency for the square wave, we will be making in effect an actual system performance check. All we need to know to wrap this whole thing up is what low end phase shift and amplitude departure do to the square wave. The following pictures tell the story. These pictures are easily constructed by simply shifting the phase of the fundamental of the square wave (and to a lesser extent the first few harmonics) or altering its amplitude.

In Conclusion

Square wave testing offers a quick and informative method of testing the transmission of linear systems. To do this job adequately the square wave generator must:

1. Have a rise time less than half that of the system to be tested.

2. Have a fundamental frequency, variable from a few cycles to a frequency whose period is not more than 100 times the rise time (so as to get good brightness when viewing the rise on fast sweeps).

3. Have negligible (under 1%) departure from flat tops either through overshoot or droop.

4. Have outputs to match cable and line impedances so as to be able to deliver to the system input an undegraded wave.

5. Have adequate output to drive an oscilloscope in spite of considerable loss in the device under test.

* There are of course cases where pulses of 10% duty cycle or less must be used to avoid operating point shifts in tubes, etc.