SECTION 4

WAVEFORM INTERPRETATION

4-1 WAVEFORM AND PHASE DEVELOPMENT

The primary objective of Section 4 is the interpretation of the common voltage and current waveforms as observed on an oscilloscope or other test equipment. A complete understanding of the appearance and reason for a normal waveform will help to recognize an abnormal waveform, and may help to understand the reason for the abnormality. Terms such as "loss of high frequencies" and "loss of response" will have special meanings when referenced to a particular waveform. The causes of deteriorated voltage and current waveforms will be discussed. The information regarding complex waveforms and phase development given in this section illustrates and explains the normal waveform and abnormal waveform, as well as their causes. This material does not include a mathematical analysis of waveforms; it is presented only as an aid to those individuals having a vital interest in waveforms and who cannot, or do not, desire to interpret mathematical explanations. A voltage or current waveform, as encountered in the electronics field, may be graphically represented in both height and width. The height of a graphically displayed waveform represents the quantity or amplitude of voltage or current, the width of the displayed waveform represents the elapsed time or waveform duration. The voltage or current waveform is normally represented in a two-dimensional (horizontal and vertical) plane without depth. The horizontal ("X") axis on a graph will represent time measured in either whole or parts of a second; the vertical ("Y") axis will represent amplitude, quantity, or intensity of the subject waveform measured, either in whole or in parts of volts or amperes. Any portion of the waveform extending above the horizontal (zero amplitude) reference line is considered positive, while any portion of the waveform extending below the horizontal reference line is considered as being negative.

4-2 SINUSOIDAL WAVEFORMS

The sine wave is the basis of all other waveforms. It represents the simple action of a swinging pendulum, a bouncing or vibrating spring, a free-running self-excited oscillator output, etc. When any outside source changes the shape of the sine wave, the wave is said to be distorted. However, it will be found that the original sine wave is still present in combination with other sine waves introduced by the distorting agency to produce a single resultant waveform. Therefore, any waveform, no matter how complex, may be reduced to its individual sine wave components. The original sine wave components cannot be reduced further, because they are the final remaining single-frequency, basic components. No waveform that is composed of more than one frequency is a true sine wave. Since the primary purpose of this portion of the section is to familiarize you with the basic structure of a sine wave, the sine wave is graphically presented in Figure 4-1. Since a 60 hertz wave is represented by one complete cycle in Figure 4-1, the total time duration for this cycle, is therefore 1/60 of one second. Half of this cycle time is above the horizontal reference (zero amplitude) line and is considered positive, while the other half is below the horizontal reference line and is considered negative. These two halves of the sine wave do not cancel out or nullify each other because each half-cycle occurs during a different time. During the positive half-cycle, the negative half-cycle has not occurred, therefore it does not exist. During the time the negative half-cycle is present the positive half-cycle does not exist. The single cycle illustrated in Figure 4-1 must not have any bumps or kinks on either its increasing or decreasing side. The top (positive peak) and the bottom (negative peak) must be smoothly curved, with no appearance of either a point or a flat spot in this region. The positive and negative half-cycles of the sine wave must be exactly equal in both amplitude and time duration. In Figure 4-1, the maximum positive peak amplitude is represented by positive 1 volt, while the maximum negative peak amplitude is represented by negative 1 volt. The time duration illustrated is exactly 1/120 of a second for each half wave. Considering that a sine wave represents a complete mechanical revolution or circle, it requires 90 degrees to traverse 1/4 of the circle; 180 degrees to traverse 1/2 of the circle, 270 degrees to traverse 3/4 of the circle; and 360 degrees (or back to 0 degrees) to complete the entire circle. There are an infinite number of points represented on a sine wave. For example, there are 360 different points, each representing an advance of 1 degree, on each sine wave. However, only four points are shown to illustrate the shape of a basic sine wave. This was done because it is only necessary to become familiarized with the general features of the sine wave curve,

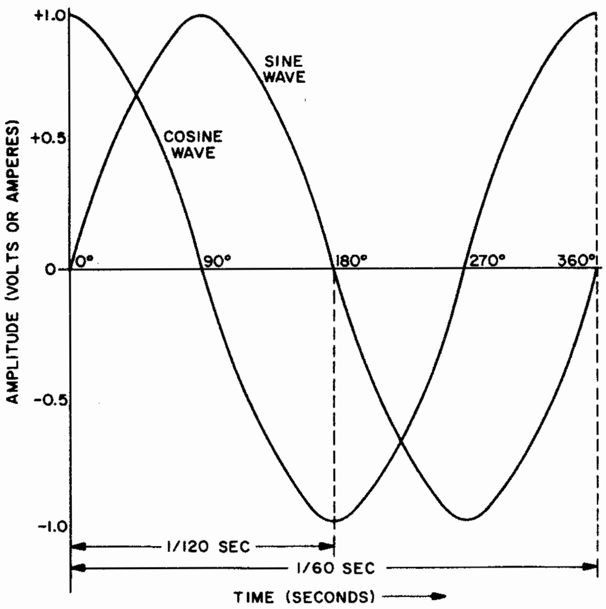


Figure 4-1. Sine and Cosine Waveforms

rather than provide a point-by-point analysis. Amplitude will not affect the general outline of a sine wave, provided that the positive and negative portions of the waveform contain equal amplitudes. If the sine wave is viewed on some form of oscilloscope, the instrument controls may be incorrectly positioned, thus presenting the wave as either too narrow or too wide for its height. This will not affect the actual waveform, but may affect the viewers perspective; the peaks of the normal waveform may appear sharp or peaked for the narrow vertical horizontal version, and flat or broad on the wide or stretched horizontal version. A point-by-point examination of the wave will prove it to be a true sine wave, disregarding its disproportional appearance. A cosine waveform is also shown in Figure 4-1. The cosine wave is the same in all respects as the sine wave except for one difference; it leads, or begins 90 degrees (1/240 of a second in this case) before, the sine wave time begins. The cosine wave is superimposed on the same graph as the sine wave to illustrate the cosine lead. These two waveforms are not used to provide a resultant waveform. The half-sine waveform consists of a series of undirectional pulses, each resembling a half-cycle of a sine waveform. The half-sine wave may exist either above (positive) or below (negative) the horizontal reference line. The half-sine waveform is produced by removing from the complete sine wave any amplitude variations in one direction for a period of one-half the time duration of the complete cycle. As illustrated in Figure 4-2, there are two types of half-sine waves. In part (A) of the figure, the negative portion of the sine waveform has been removed and its 1/2-cycle time interval remains as a zero dc reference level. In this type of half-waveform, the frequency remains the same as the original full sine wave frequency. In part (B) of the figure, the negative half of the original full sine wave has been inverted over the horizontal reference line; consequently, the average dc voltage level is increased. Since each alternation occurs in one-half the time interval of the original full sine waveform, inverting the negative alternations to occupy the empty spaces between the positive alternations causes the frequency to double. Half-sine waveforms are composed of the original fundamental frequency in conjunction with a dc component and an infinite series of even numbered harmonics of progressively decreasing amplitude.

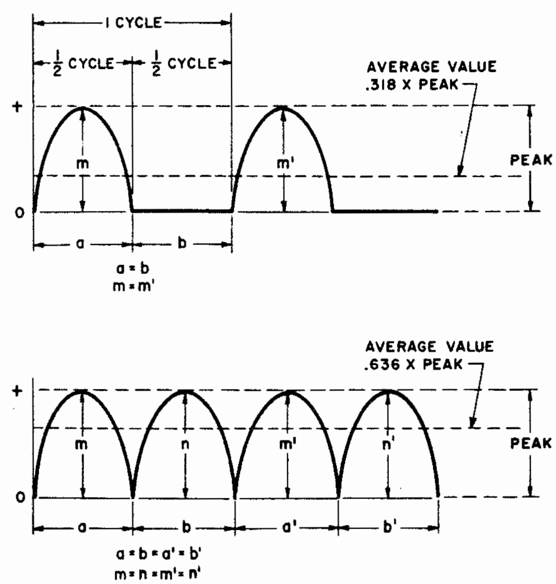


Figure 4-2. Half-Sine Waveforms

4-3. NON-SINUSOIDAL WAVEFORMS

The sine wave is the basic or standard alternating current (or voltage) waveform used in combinations with phase or time differences and amplitudes to algebraically form all other waveforms. The sine wave is the wave most commonly used as an input to circuits under test because it does not introduce distortions commonly associated with nonsinusoidal waveforms. All nonsinusoidal waveforms can be reduced into their individual component sine waves. A nonsinusoidal waveform is composed of more than one sine wave; other frequencies, usually harmonically related, are algebraically added to the fundamental frequencies to produce the resultant nonsinusoidal waveform. In this case, the sine wave of lowest frequency is normally considered to be the fundamental frequency, and higher frequencies which are exact multiples of the fundamental frequency are considered as harmonics of the fundamental. However, in some cases, the nonsinusoidal waveform being considered may be composed of only harmonic frequencies because the fundamental sine wave may have been intentionally removed. The algebraic addition of the fundamental sine wave (f), and the second harmonic of the fundamental (h=2f) will provide a resultant nonsinusoidal waveform (R) as shown in Figures 4-3, 4-4, and 4-5. However, only the resultant would be shown on an oscilloscope or equivalent test instrument. The second harmonic of Figure 4-3 is shown in phase with the fundamental because its amplitude increases in the same direction as the

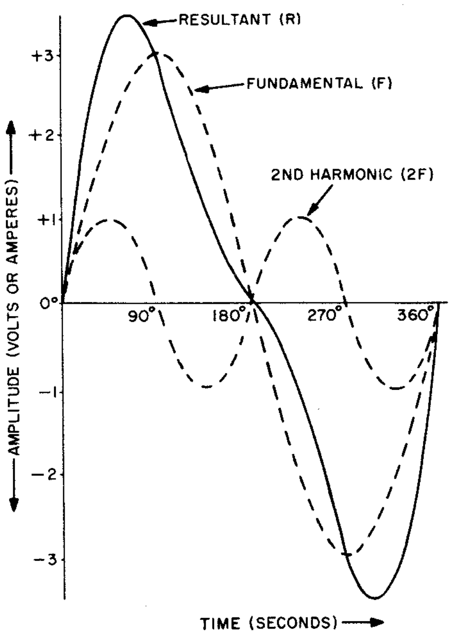


Figure 4-3. Waveform Resulting from Algebraic Addition of a Fundamental Sine Wave with Its Second Harmonic in Phase

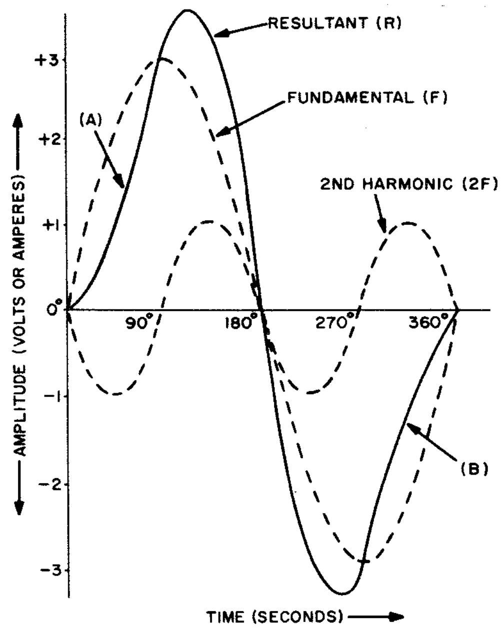


Figure 4-4. Waveform Resulting from Algebraic Addition of a Fundamental Sine Wave with Its Second Harmonic Delayed 180 Degrees

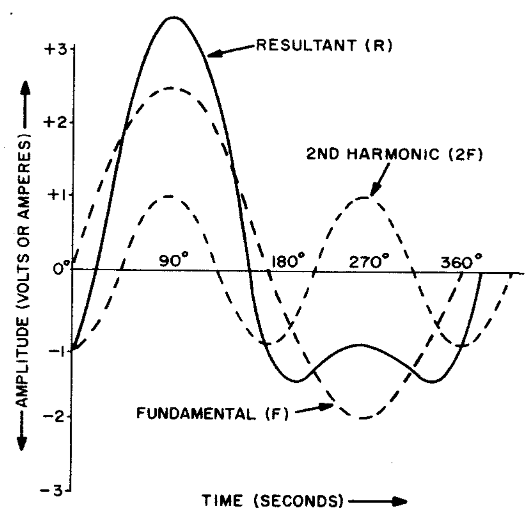


Figure 4-5. Waveform Resulting from Algebraic Addition of a Fundamental Sine Wave with Its Second Harmonic Delayed 90 Degrees

fundamental from the horizontal and vertical zero reference level. The second harmonic of Figure 4-4 is shown 180 degrees out of phase with the fundamental because it proceeds from the horizontal and vertical zero reference level in a direction exactly opposite (negative direction) from the fundamental. The second harmonic shown in Figure 4-5 is shifted 90 degrees behind (lagging) the fundamental. The resultant waveforms contained in Figure 4-3, 4-4, and 4-5 are the only waves you will see on an oscilloscope or other equivalent test instrument. The amplitude of the second harmonic, relative to the fundamental sine wave, will either increase or decrease the amount of dip at points A and B of Figure 4-4, represented by two heavy arrows. However, if the phase of the second harmonic is changed with respect to the fundamental sine wave, the appearance of the resultant waveform will change completely. As shown in Figure 4-5, the positive half of the resultant waveform looks like part of a sine wave, but the negative half does not. Therefore, this resultant is definitely not a true sine wave. At any point along the horizontal zero reference line the vertical amplitude of the fundamental can be added directly to the vertical amplitude of the harmonic to obtain the final amplitude of the resultant waveform at that particular point (that is if $f=3$ units and $h=-1$ unit then $R=2$ units; and if $f=-6$ units and $h=-3$ units, then $R=-9$ units). The resultant waveforms (Figures 4-3, 4-4, and 4-5) show that the algebraic addition of a harmonic waveform to the original fundamental sine wave produces a new waveform that is no longer a sine wave. In many cases these new waveforms are created deliberately to perform functions beyond the capabilities of the original signal. For example, the new waveforms resulting from the addition of a fundamental and its harmonics may be used as timing pulses.

4-3.1 PHASE DISTORTION

Figure 4-5 represents the results of feeding the fundamental and its second harmonic through a network which delays the second harmonic by 90 degrees. This normal resultant waveform obtained by the algebraic addition of the second harmonic (without phase shift) to the fundamental sine wave, is known as "phase distortion". This type of time delay can be recognized only by becoming familiar with the proper waveforms. If the phase of a fundamental sine wave is shifted, its shape will not change. Therefore, special methods must be employed to recognize any change in phase. These methods are described in the following paragraphs.

4-3.2 HARMONIC DISTORTION

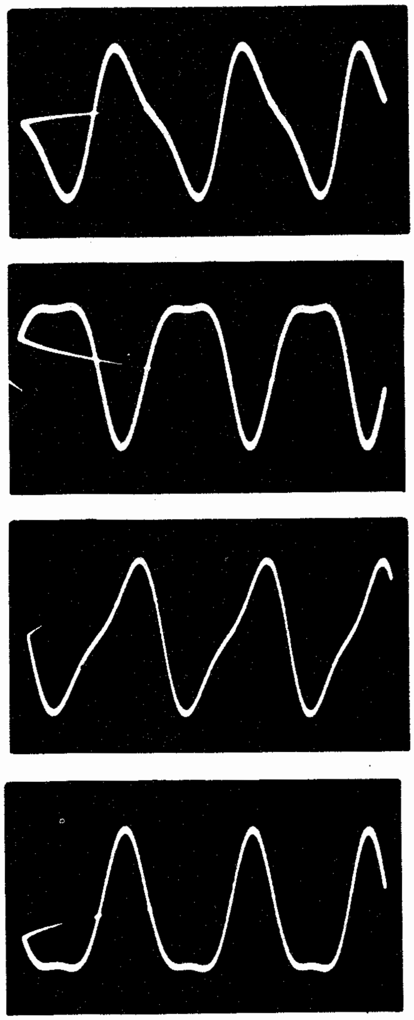
The addition of harmonics to the fundamental wave shape creates a new resultant waveform. The resultant is a distortion of the original waveform and, if undesirable, is termed "harmonic distortion". The resultant waveform created by the addition of only one harmonic to the original waveform can probably be recognized. However, the addition of several harmonics to the fundamental sine wave, in and out of phase, will create a resultant waveform of pure confusion. As was mentioned previously, any waveform can be separated or removed from its resultant with the aid of suitable filters. Therefore, by removal of all except one frequency component, you can extract a pure sine wave of some specific frequency which was not evident in the original wave or its harmonics. This newly extracted sine wave can then become the fundamental sine wave input to a circuit under test.

4-3.3 COMPLEX WAVEFORMS

The resultant waveforms shown in this section are created either by adding harmonics to the fundamental waveform; by changing the phase of the harmonic with respect to the fundamental; or by a combination of harmonic addition and phase change. Therefore, all resultants are termed "complex waveforms", no matter how simply or easily they are recognized. Actually, other effects are realized in the resultant waveform, depending on the addition of even (2nd,4th,6th,8th,etc.) harmonics or odd (3rd,5th,7th, 9th,etc.) harmonics, on the percentage of harmonic waveform amplitude injected, and on the phase of the introduced harmonic with respect to the fundamental sine wave. All so-called "distorted waveforms" are classified as complex waveforms. However, complex waveforms are grouped into types of complex waveforms. To familiarize the technician with some of the primary features of nonsinusoidal waveforms consisting of a single harmonic addition to the fundamental sine wave, a series of illustrations are provided as Figures 4-6 through 4-8.

4-3.4 MIRROR SYMMETRY

The term "mirror symmetry" refers to the fact that if the positive part of the resultant wave is inverted over the horizontal reference line, it will exactly match the negative part of the resultant waveform; or conversely, if the negative part of the resultant wave is inverted over the horizontal reference line, it will have the same shape, outline, and appearance of the positive part of the resultant waveform. By viewing the resultant waveform, it can be definitely determined



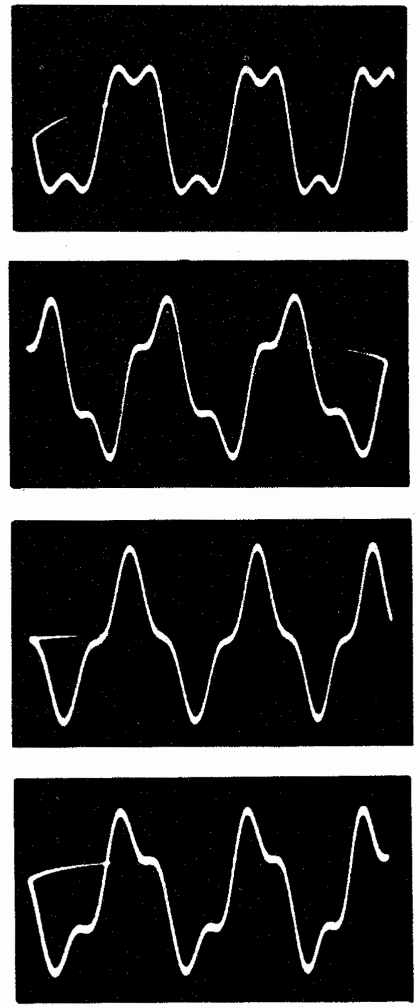
A
IN PHASE

B
PHASE SHIFTED
90 DEGREES

C
PHASE SHIFTED
180 DEGREES

D
PHASE SHIFTED
270 (-90) DEGREES

Figure 4-6. Resultant Waveforms Created by Algebraic Addition of Second Harmonic to Fundamental Sine Wave When Second Harmonic Amplitude is 30 Percent of Fundamental



A
IN PHASE

B
PHASE SHIFTED
90 DEGREES

C
PHASE SHIFTED
180 DEGREES

D
PHASE SHIFTED
270 (-90) DEGREES

Figure 4-7. Resultant Waveforms Created by Algebraic Addition of Third Harmonic to Fundamental Sine Wave When Third Harmonic Amplitude is 30 Percent of Fundamental

whether even harmonics (2nd, 4th, 6th, etc.) were added to the fundamental sine wave to create the resultant, or whether odd harmonics were used for this purpose. If even harmonics were algebraically combined with the fundamental sine wave, there will be a lack or mirror symmetry, as illustrated in part (A) of Figure 4-9. If odd harmonics (3rd, 5th, 7th, 9th, etc.) were algebraically combined with the fundamental sine wave, there may be mirror symmetry, as illustrated in part (B) of Figure 4-9, but only when there is an adequate percent of the odd order harmonics amplitude.

4-4. SQUARE WAVEFORMS

The Square waveform is a resultant waveform type composed of a sine waveform in conjunction with odd harmonics. Unlike the original sine

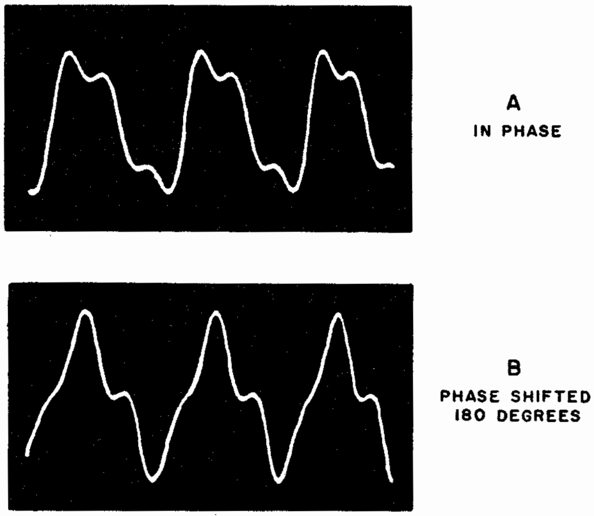


Figure 4-8. Resultant Waveforms Created by Algebraic Addition of Third Harmonic to Fundamental Sine Wave When Third Harmonic Amplitude is 15 Percent of Fundamental

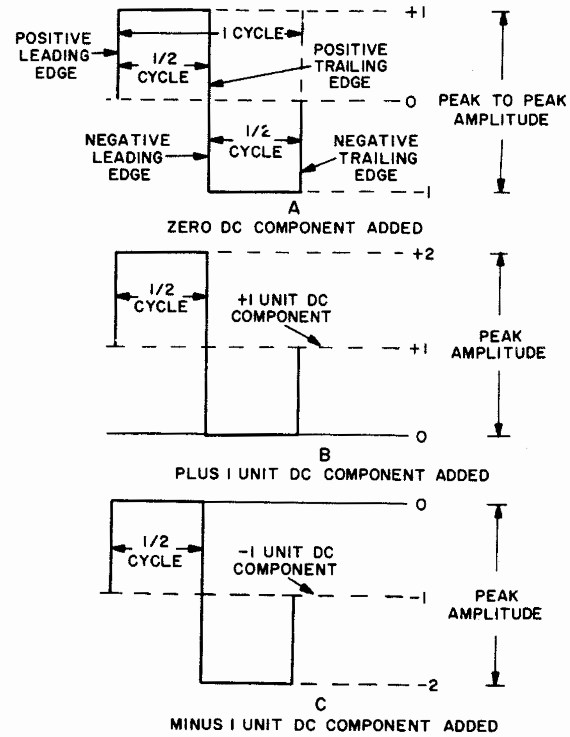


Figure 4-10. Square Waveforms

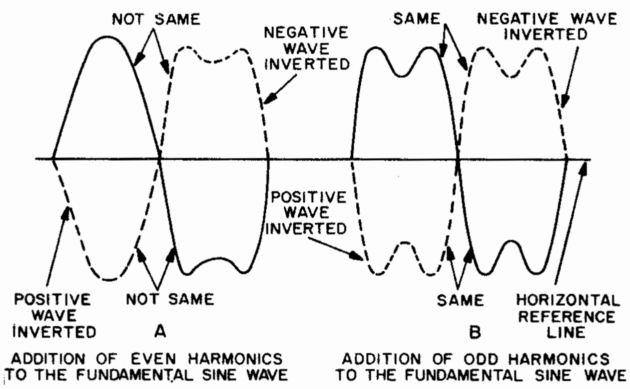


Figure 4-9. Presence or Absence of Mirror Symmetry Due to Harmonic Addition to Fundamental Sine Wave

waveform, the application of a square waveform to either a capacitive or inductive circuit will result in an output of a completely different waveform shape. As illustrated in Figure 4-10, the leading edge of a square wave rises from zero reference value to its maximum value, where it remains as a constant-amplitude wave over a set period of time. It then drops back toward its original zero reference level or beyond, until it reaches a minimum value, where it remains as a constant-amplitude wave over an exact period of time that matches its positive excursion. The rise and fall times are negligible in an ideal square wave. Part (A) of Figure 4-10 illustrates the ac square waveform; it is so called because the waveform extends in a negative direction below the horizontal reference line as well as above. Parts (B) and (C) of Figure 4-10 illustrate the pulsating dc square waveforms; they are so called because they contain a dc component which prevents the waveforms from crossing the horizontal zero dc reference level. However, all three forms are identical except for amplitude. All corners must be square, the sides perpendicular, and the extremities flat. Unfortunately, this idealized square waveform cannot be attained because the waveforming equipment is not perfect. The square wave is formed by the algebraic addition of the fundamental sine wave and an infinite number of odd harmonics of the fundamental sine wave. However, as is basically illustrated in Figure 4-11, as few as three added odd harmonics will produce a reasonable facsimile of a square wave, although a

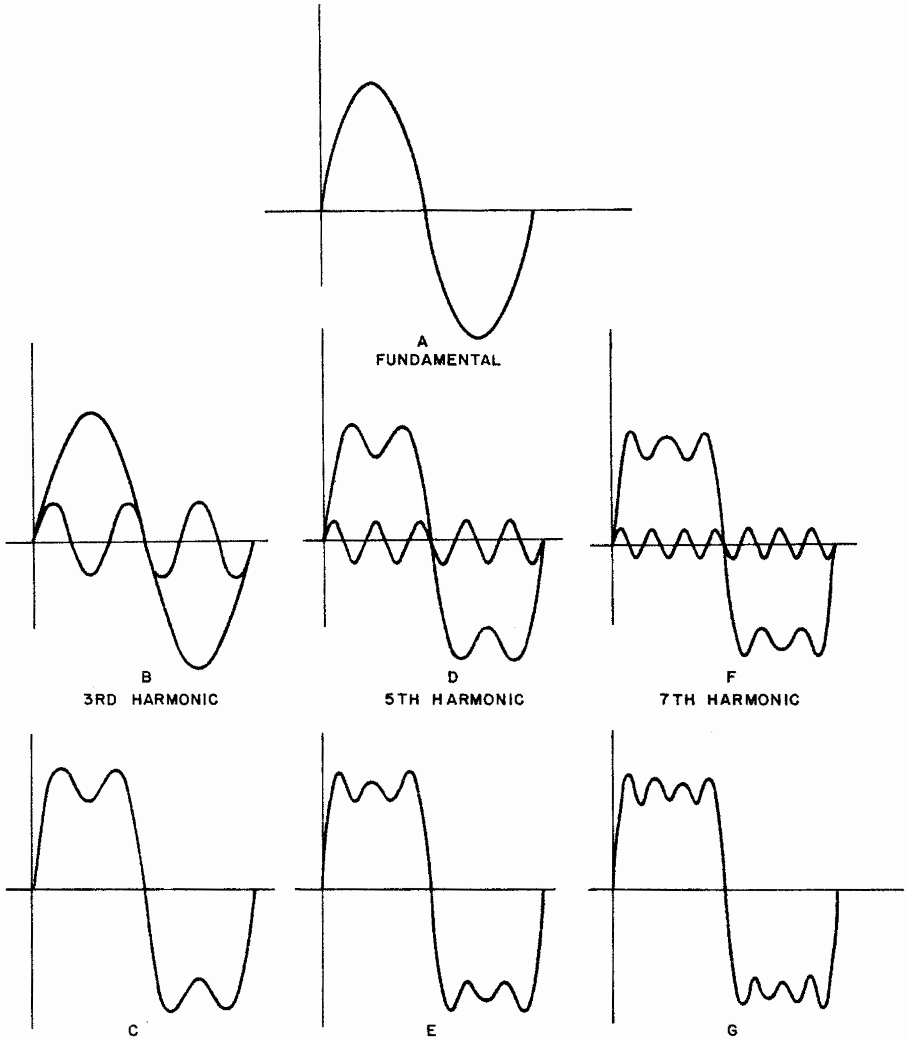


Figure 4-11. Formation of a Square Wave

minimum of 10 added harmonics are required to produce a usable square wave. The fundamental sine wave may start at any phase. For illustrative purposes, Figure 4-11 shows the sine wave as beginning at the horizontal and vertical zero reference level. Figure 4-11 shows the resultant waveforms as additional harmonics are progressively added to the fundamental sine wave. Although this illustration shows the algebraic addition of only three odd harmonics, it conveys the true impression that as each additional harmonic is added, the leading and trailing edges of the resultant waveform become flatter. The frequency (f) of any order odd harmonic can be determined by calculating the value with the aid of the formula:

$$f_N = (2N + 1)f_1$$

where f is the fundamental frequency and N is the order of the harmonic. For example, if the fundamental frequency is 100 Hz, the frequency of the 6th odd harmonic can be determined as follows:

$$f_6 = [(2 \times 6) + 1] \times 100$$

$$f_6 = [(12) + 1] \times 100$$

$$f_6 = (13) \times 100$$

$$f_6 = 1300 \text{ Hz}$$

It also follows that each odd harmonic shown in Figure 4-11 has been added in phase (zero phase difference) with the original fundamental sine wave. In addition, another factor of Figure 4-11 should be inspected. The amplitude of each harmonic is in direct proportion to the harmonic order; that is, the third harmonic contains 1/3 the amplitude of the fundamental, the fifth harmonic contains 1/5 the amplitude of the fundamental, etc. If the harmonic is not in phase, or has incorrect amplitude, etc., the resultant square wave is said to be distorted. However, the type of distortion observed may indicate the kind of trouble, and even the source of trouble, within a circuit.

4-5 RECTANGULAR WAVEFORMS

The rectangular waveform contains all but one feature of the square waveform discussed previously and illustrated in Figures 4-10 and 4-11. This one different feature is that the former wave has identical periods of positive and negative pulsations, whereas the latter wave has unlike periods (time duration) of the positive pulse with respect to the negative pulse. The rectangular pulse can be either bidirectional or unidirectional, in that the waveform may be entirely above or entirely below the horizontal zero reference level, as shown in Figure 4-12. The period of the rectangular waveform, like that of the square waveform, is the total time required to complete both half-cycles together as one unit. For example, in part (A) of Figure 4-12, if the positive half-cycle duration is 50 microseconds and the negative half-cycle duration is 150 microseconds, the total period for one cycle of this frequency is 200 microseconds, the totai period for one cycle of this frequency is 200 microseconds. Considering that the frequency of a cycle is the reciprocal of the time required for that cycle, 1/200 μ s, thus the frequency of this example waveform is 5000 Hz. This means that this particular cycle will repeat itself 5000 times every second; it is said to have a frequency repetition of 5000 Hz. The shorter pulse durations require the presence of higher-frequency components, whereas longer pulse durations require the presence of lower frequency components. The rectangular waveform is rarely used as a test voltage. However, it may be used in many special applications to perform a specific function. Figure 4-13 illustrates such a function. The rectangular wave in this case has a rectangular wave riding atop the first wave. This is a practical situation in the transmission and reception of a television signal cycle. The video information is shown riding on the minimum amplitude portion of the first rectangular wave between the blanking pulses, representing the positive excursions of this rectangular pulse. However, rectangular synchronization pulses are

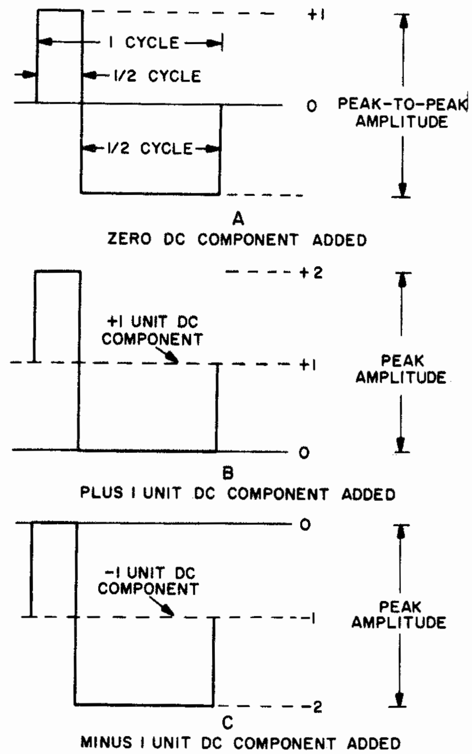


Figure 4-12. Rectangular Waveforms

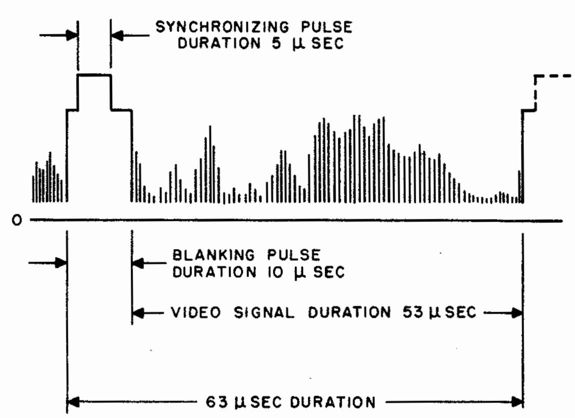


Figure 4-13. Rectangular Waves Used in Television

shown riding atop the blanking pulse. This rectangular pulse represents the so-called "front porch", the synchronizing pulse proper, and the so-called "back porch".

4-6. SAWTOOTH WAVEFORMS

The sawtooth waveform, like all other waveforms except the fundamental sine wave, is composed of sine-wave components. As illustrated in Figure 4-14, the wave consists of a gradual linear change from a maximum negative-going peak to its maximum positive-going peak. It then experiences a rapid drop to its original amplitude. Considering that this waveform is composed of many sine-wave components which may differ in both frequency and phase, you cannot apply the sawtooth waveform to any inductive or capacitive device to cause different lead or lag times between the sine wave components that compose the sawtooth waveform. Therefore, the output from any component device, other than a pure resistance, would not be the same as the original sawtooth input. In the ideal waveform illustrated in part (A) of Figure 4-14, the retrace time is shown as zero seconds. This is not the true practical case because any action or reaction requires a definite time for accomplishment. Parts (B) and (C) of Figure 4-14, therefore reflect a practical case where the retrace time is some finite time, rather than zero. However, the retrace time is normally assigned the smallest practical duration consistent with the design of the equipment with which it is to be used. If the voltage amplitude increases at a constant rate during the forward trace, the waveform is said to be a "linear sawtooth". The fact that half of the waveform shown in Figure 4-14 is above the horizontal zero reference level and the other half is below this reference level will normally not be seen on an oscilloscope because the reference level line (time base line) is absent from the display. Unlike the square waveforms which were produced by the algebraic addition of odd frequency, in-phase harmonic waves with the fundamental sine wave, the sawtooth waveform is the resultant wave produced by the algebraic addition of both even and odd frequency harmonics to the fundamental sine waveform. A positive-going sawtooth waveform is produced by the algebraic addition of all harmonics to the fundamental sine wave, but the fundamental harmonic components must begin in-phase and start in a negative direction, as shown in Figure 4-15. A negative-going sawtooth is the resultant of the same sine-wave components, but the fundamental and the in-phase harmonics must start in a positive direction. Figure 4-15 illustrates the method of progressive

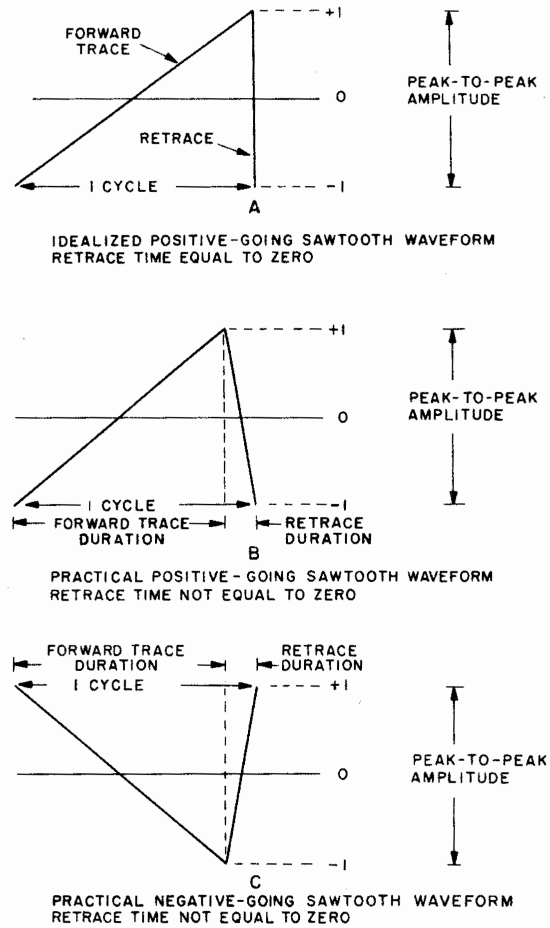


Figure 4-14. Sawtooth Waveforms

algebraic addition of each higher-frequency harmonic to the fundamental sine wave in order to gradually obtain the ultimate sawtooth waveform. However, only the first two harmonics (2nd and 3rd) have been combined with the fundamental sine wave to form the resultant shown in part (D) of Figure 4-15. As more harmonics are progressively added, the resultant wave will approach more and more closely the required sawtooth form.

4-7 TRAPEZOIDAL WAVEFORMS

The trapezoidal waveform is the resultant of the algebraic addition of sine waves, but it is more eaily understood in terms of a sawtooth rectangular waveform, both of which are composed of basic

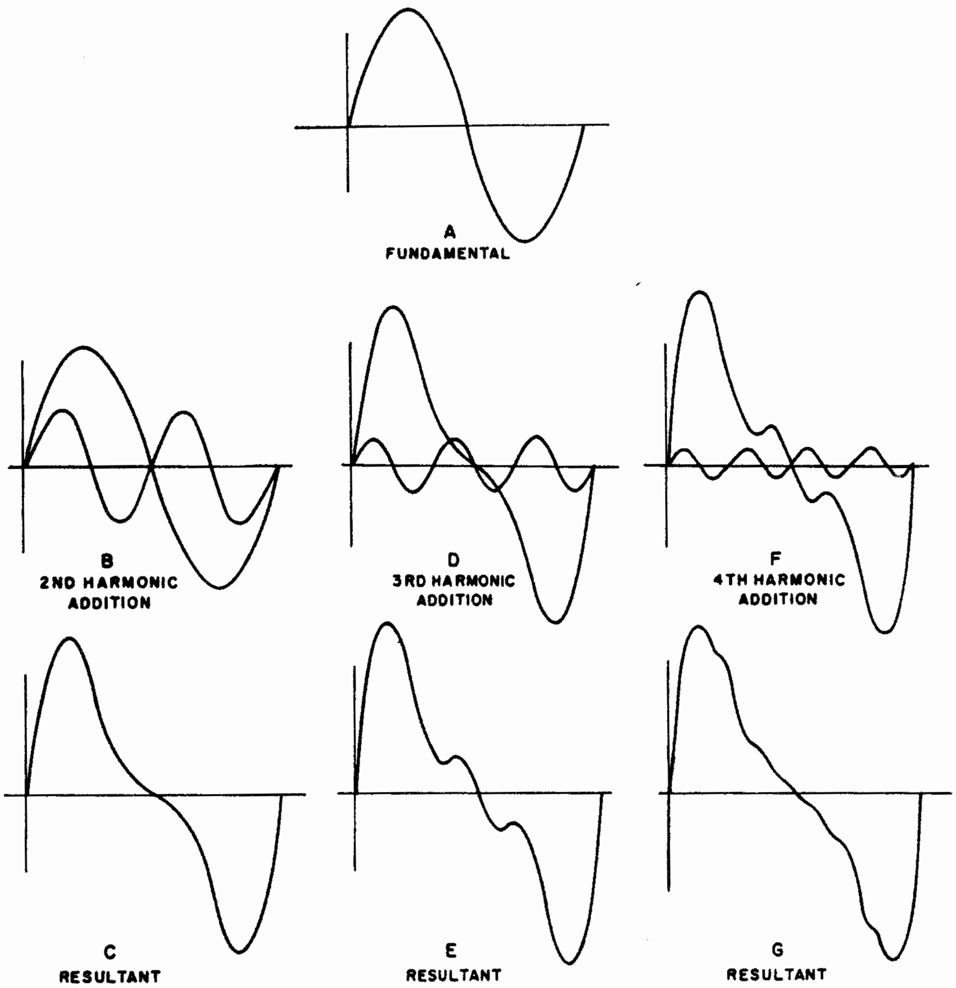


Figure 4-15. Formation of Sawtooth Waveform

sine waves. As previously stated, a sawtooth of voltage applied to the input of either an inductive or capacitive device will not appear at the output of the reactive component or device as a sawtooth waveform. Therefore, those applications which require a sawtooth current waveform must obtain the sawtooth current from a device which has a sawtooth current output produced from a trapezoidal voltage input. The trapezoidal waveform has the necessary characteristics to cause a linear change in the amplitude of the current, with respect to time, as it passes through the resistive and inductive components of a coil. Figure 4-16 illustrates the resultant output of a resistor versus a coil with an applied input sawtooth waveform. The sawtooth wave passing through the pure resistive element produces no change in the output waveform. However, the output from the coil with an applied sawtooth waveform is essentially a rectangular waveform. Figure 4-17 illustrates the algebraic addition of a sawtooth waveform to a rectangular waveform in order to produce a resultant trapezoidal waveform for application to a series resistive-inductive circuit, the output of which is a sawtooth current waveform. The trapezoidal waveform occurs in numerous varieties because of the amplitude differences of the sawtooth voltages and rectangular voltages prior to algebraic addition. A comparison of two varieties of trapezoidal waveforms is illustrated in Figure 4-18. The resultant waveform in part

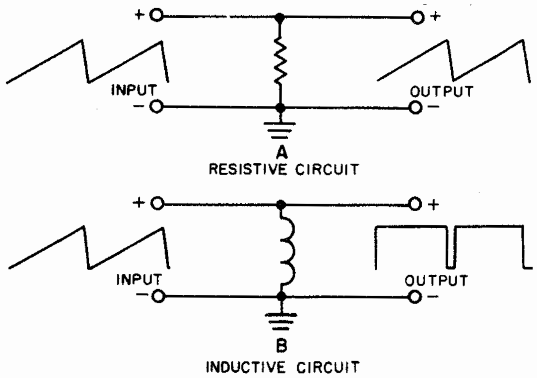


Figure 4-16. Output Current Waveforms for Resistive and Inductive Circuits Resulting from a Sawtooth Voltage Input

(C) is not the same as the resultant waveform in part (F). Part (C) is the resultant trapezoidal waveform most commonly used in electronic technology. Sawtooth current waveforms are generated by deflection circuits for cathode-ray-tube deflection coils. When a deflection coil has a small internal resistance as compared with its inductive reactance, the sawtooth current waveform is produced by a sweep voltage that is a combination of a small sawtooth waveform and a large rectangular waveform.

4-8 DIFFERENTIATED VOLTAGE WAVEFORMS

Various complex waves can be resolved into their component sine-wave frequencies, and any group of frequencies can be extracted from a complex wave by means of a filter. In the case of differentiation, the differentiated waveform extracts the high-frequency sine-wave components, while the integrator

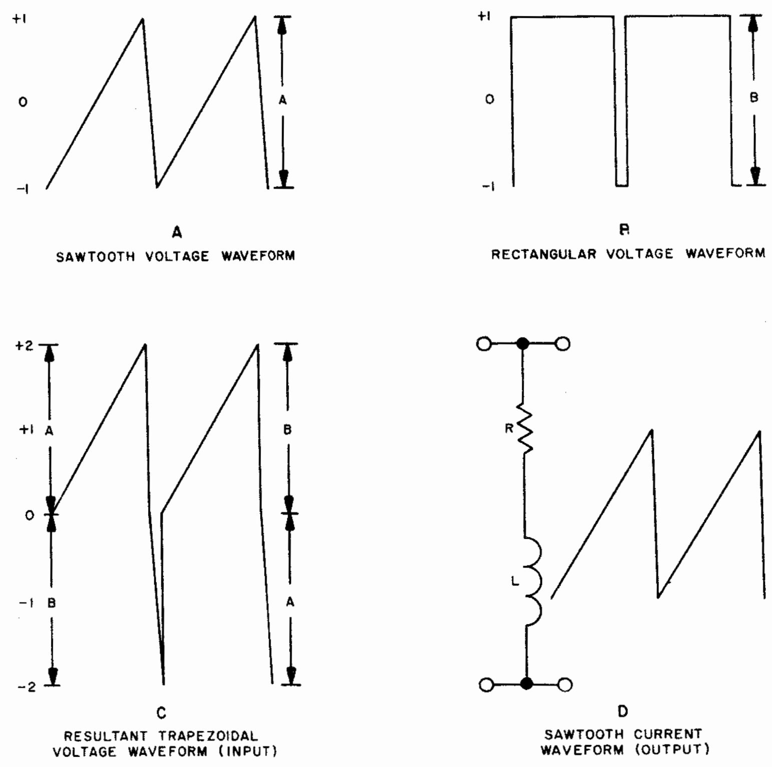


Figure 4-17. Output Current Sawtooth Waveform Resulting from Application of a Trapezoidal Input Voltage Waveform to an Inductor

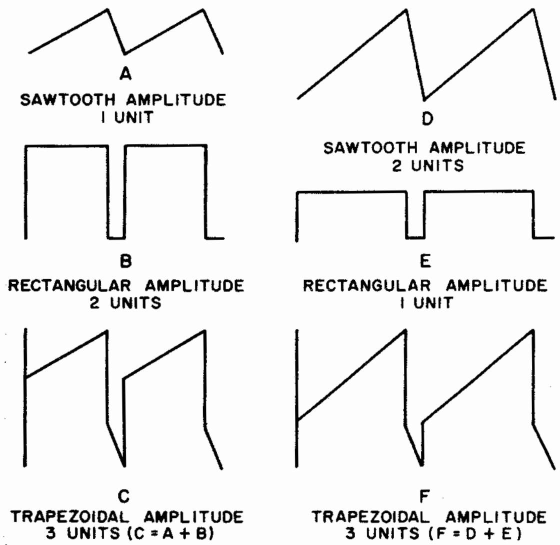


Figure 4-18. Trapezoidal Voltage Waveform Varieties

extracts the low-frequency sine-wave components. A differentiated waveform is obtained by the process of differentiation. This process is simply the procedure whereby a waveform is passed through inductive or capacitive components to provide a voltage output proportional to the rate of change of the input voltage waveform. The most popular method of differentiation employs capacitive-resistance (RC) network. The time constant of the circuit, in microseconds, is the product of the resistance and capacitance in ohms and farads, respectively. A rapid change occurring in the input voltage waveform will produce a narrow sharp-peak (spike) in the output. The peak amplitude of the output pulse is therefore directly proportional to the rate of change in the input waveform. Figure 4-19 illustrates a square-wave input to each of the two most common differentiating circuits, and also the output obtained from each circuit. A square wave was used because of its rapid amplitude change and high harmonic frequency content on both the leading and trailing edges of the applied pulse. The flat horizontal portions of the square wave will produce zero output because they contain zero slope (change), and also because the time constant of the differentiator will not pass the lower frequencies contained in the square wave. A sine wave is not used as an input voltage to a differentiation circuit because these circuits accomplish their function by shifting the phase of the input waveform. In the case of a sine wave, the output will be shifted in phase and will have a smaller amplitude, but will still be a sine wave. This output may be particularly useful as a phase-shifted wave formed from a continuously variable time-constant function.

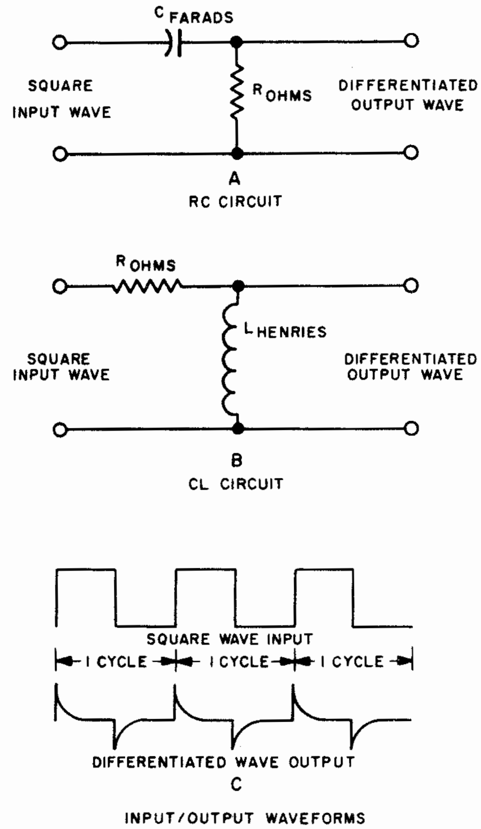


Figure 4-19. Input to and Output from Differentiating Circuits

4-8.1 RECTANGULAR VOLTAGE WAVEFORMS

The rectangular waveform has the same characteristics as the square waveform with respect to differentiation. It is important to note that the output voltage has a sharp peak only when the differentiating circuit contains a short time-constant. Also, it should be realized that the sharp peak is produced only during the rapid rise or rapid fall of the input voltage. Therefore, the differentiator circuit is known as a "peaker" circuit. In a circuit containing a time constant of less than 1/10 the time required for one cycle of input voltage, the time constant is said to be "short"; it is said to be "long" if the circuit components permit a time of

10 times the duration of one cycle of input voltage. A square waveform produces an output differentiated wave with evenly spaced positive and negative excursions, whereas a rectangular waveform produces a positive and negative peak spaced close together (paired), with a distance separation from the next pair of peaks as shown in Figure 4-20.

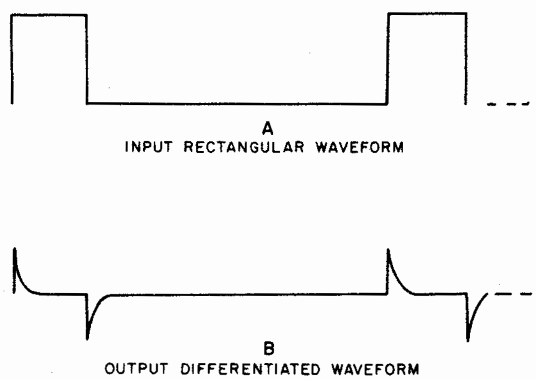


Figure 4-20. Rectangular Input and Resultant Differentiated Output Waveform

4-8.2 SAWTOOTH VOLTAGES

When a sawtooth voltage is applied to a differentiating circuit containing a short time-constant, the output from that circuit will be a rectangular waveform. If the applied sawtooth is positive-going, the negative spike of the output rectangular waveform will increase in amplitude as the retrace duration time is made smaller. This condition is illustrated in Figure 4-21. However, as the time-constant of the differentiated circuit is increased, the output progressively takes on the appearance of the input sawtooth waveform, shown in Figure 4-22.

4-8.3 RESISTOR-INDUCTOR DIFFERENTIATION

The RL differentiator, consisting of a resistor and an inductor in series, serves the same purpose as the RC differentiator. The output of the RC circuit is taken from across the resistor, whereas the output of the RL circuit is taken from across the inductive element. However, using either form of differentiator, the time-constant of the circuit represents the actual time required for the voltage to charge the capacitor in the RC network, or for the current to charge the coil in the RL network. The actual time-constant of a differentiator (in microseconds) can be obtained by use of the applicable formula:

$$T = RC$$

or

$$T = L/R$$

where R is in ohms, C is in microfarads, and L is in microhenries. The shape of the voltage waveform across the capacitor and the waveform of the current through the coil are identical. Therefore, technical data pertaining to the output voltage waveform of the RC network is the same for the output current waveform of the RL network.

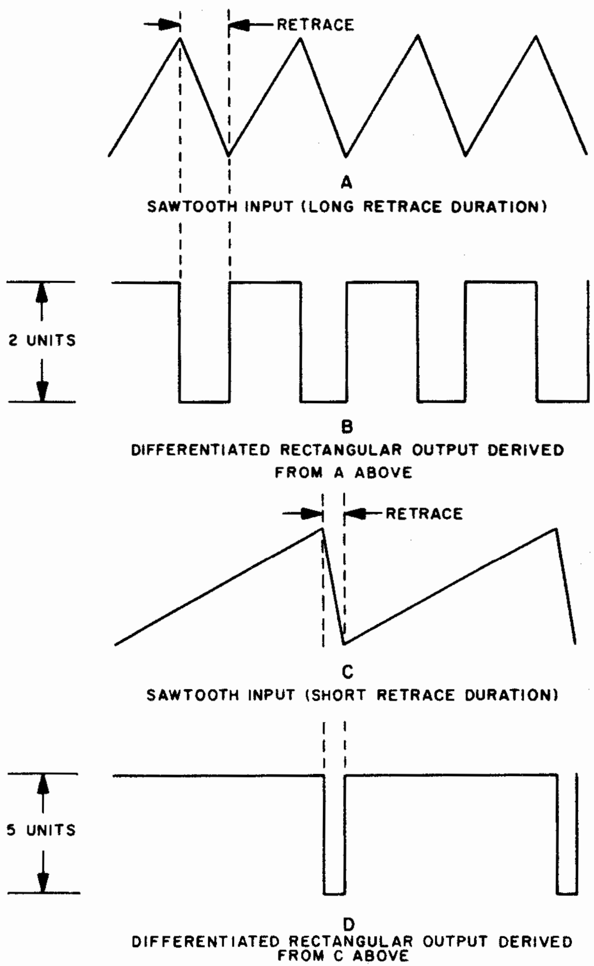


Figure 4-21. Differentiated Wave Amplitude Changes Resulting from Sawtooth Input Rate-of-Change

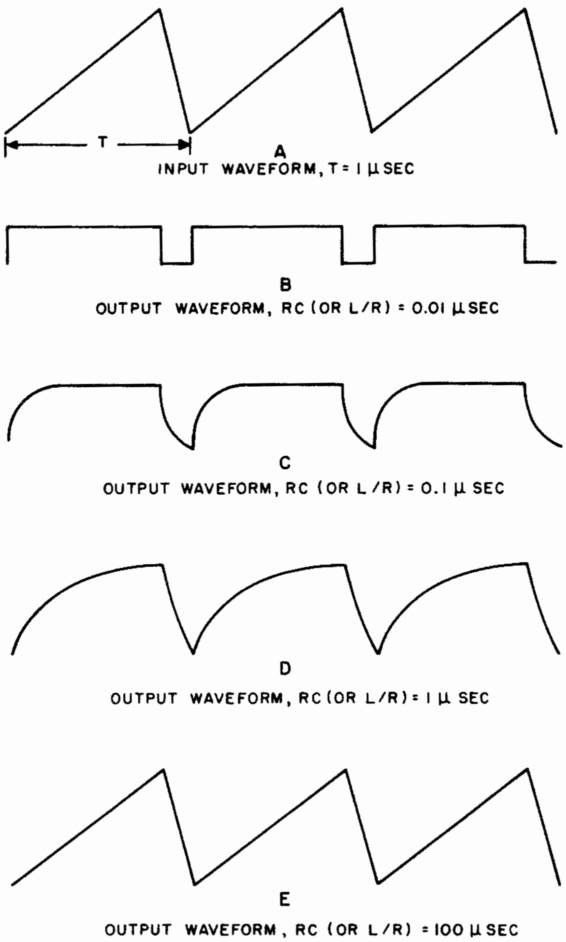


Figure 4-22. Differentiated Output Waveforms for Sawtooth Input Waveform Progressively Illustrating an Increasing RC or RL Circuit Time-Constant

4-9. INTEGRATED VOLTAGE WAVEFORMS

In contrast to the differentiator circuit, which is actually a high-pass filter, the integrator sums up the applied voltages and discriminates against high frequencies; it is thus a low-pass filter. The integrator can use exactly the same components as the differentiator circuit. However, in the case of an RC integrator circuit the output is taken from across the capacitor, whereas in the differentiator circuit the output was taken from across the resistor. The reverse is also true of the RL integrator circuit. The output is taken from across the resistor, whereas in the differentiator circuit the output was taken from across the coil . Figure 4-23

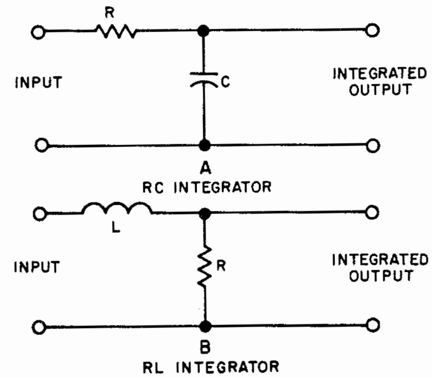


Figure 4-23. Typical Integrator Circuits

illustrates the two common forms of integrator circuits. When a square wave is applied to an integrator circuit and the integrator circuit time-constant is increased, the output waveform will gradually take on the appearance of a sawtooth waveform and will decrease in amplitude. However, the shorter the circuit time-constant, the more closely the shape of the output will resemble the shape of the input. Figure 4-24 shows the various forms and representative amplitudes for an input square or rectangular waveform. When a square waveform is applied to the input of an integrating network having a long time-constant, the output waveform will approximate a sawtooth wave, with the charge or trace portion equal to the discharge or retrace portion. However, if a rectangular waveform is used as the input to the same integrating circuit, the output will not have equal charge and discharge times; therefore, the resultant waveform will build up in either a positive or negative direction. A positive buildup of the output waveform as a result of an input rectangular waveform with longer positive pulse durations than negative pulse durations is illustrated in Figure 4-25. The shorter time duration provided by the negative portion of the input rectangular pulse will provide less time for the output discharge and, as a consequence, the output waveform will charge more rapidly to its maximum value.

4-10. MODULATED WAVEFORMS

The art of superimposing on, combining with, or changing the original carrier frequency by the addition of intelligence in the form of electrical energy, is termed "modulation". The three primary types of modulation are: amplitude modulation; frequency (or phase) modulation; and pulse modulation.

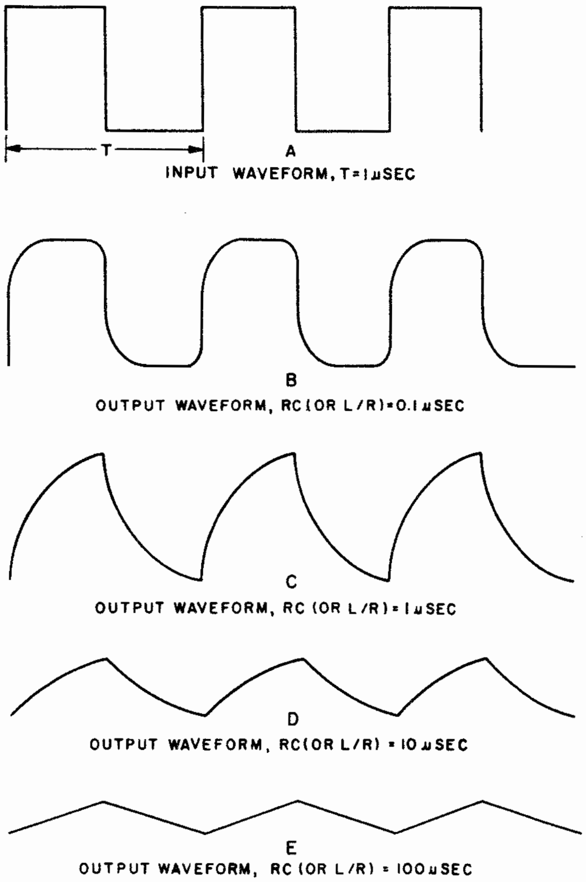


Figure 4-24. Integrated Output Waveforms Progressively Illustrating an RC or RL Circuit Time-Constant

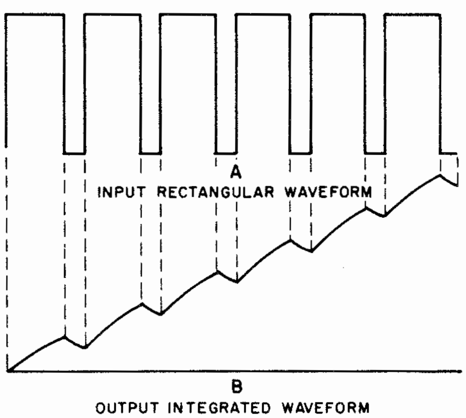


Figure 4-25. Cumulative Wide Integrated Pulse Obtained from Narrow Pulse Rectangular Waveform

4-10.1 AMPLITUDE MODULATION

The radio-frequency (RF) carrier is normally generated with the characteristics of a constant frequency and a constant amplitude. However, the amplitude of this carrier can be varied in direction in direct accordance with the intelligence to be transferred (the spoken word, music, etc.) by simply adding the amplitude of the intelligence algebraically to the amplitude of the RF carrier. This is accomplished by means of some form of mixing circuit. Figure 4-26

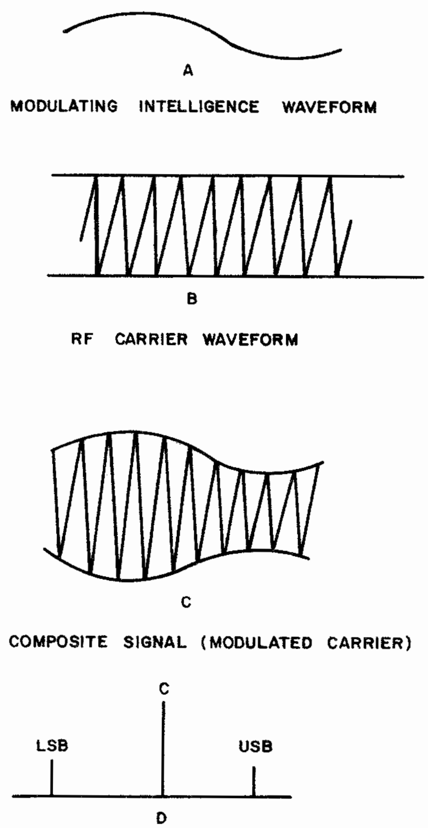
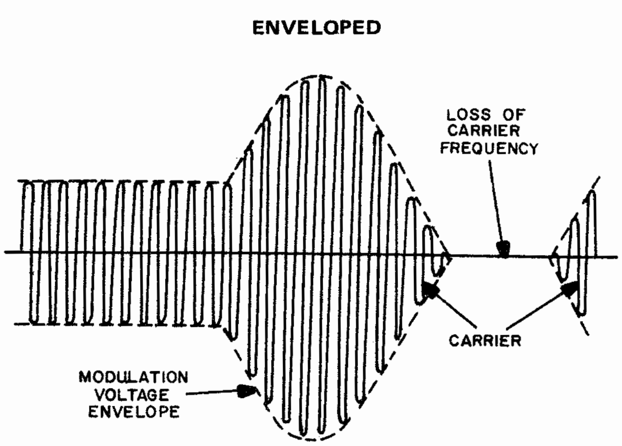


Figure 4-26. RF Carrier Amplitude Modulated by a Sine Wave

illustrates a hypothetical waveform representing the intelligence; a hypothetical RF carrier; the resulting composite signal obtained by algebraic addition of the carrier and intelligence frequencies; and the spectral display obtained through mixing the two signals. This resultant amplitude-modulated waveform is transmitted and received, and the amplitude is then detected (separated from the carrier) and converted back to a facsimile of the orginal intelligence information. The amplitude-modulated carrier illustrated in part (C) of Figure 4-26 is actually a composite of one carrier frequency component, part (B), in addition to two modulating frequency components originally contained in part (A). The two frequency displacements from the carrier, as illustrated in part (D), are the sidebands, upper and lower, generated from linear mixing of the carrier and the modulation voltages. For example, if the carrier frequency shown in part (B) of the figure were 1000 kilohertz and the modulation voltage shown in part (A) were 10 kilohertz, the resultant modulated RF carrier shown in part (C) and (D) would contain the original carrier frequency component of 1000 kilohertz plus a lower sideband component obtained by algebraically subtracting the modulating frequency from the carrier frequency (1000 kHz — 10 kHz = 990 kHz) and an upper side band component obtained by algebraically adding the modulating frequency to the carrier frequency (1000 kHz + 10 kHz = 1010 kHz). If the same carrier was modulated with two modulating frequencies, such as 10 kHz and 20 kHz, the resultant modulated carrier would be composed of five frequency components (that is, 1000 kHz, 990 kHz, 1010 kHz, 980 kHz, and 1020 kHz. Thus the original frequencies of the modulating voltages are not apparent, and the intelligence is now contained within the parameters of the sidebands created. Do not assume, however, that the original carrier now contains intelligence, because the carrier will be completely eliminated after it has served its purpose. The modulating voltage can increase the amplitude of the carrier any amount above and below the horizontal zero reference level without creating any operational difficulty. However, if it decreases the amplitude of the carrier to the zero reference level, it will remove the existing carrier frequency and thus create distortion. This type of distortion is termed "overmodulation", as is illustrated in Figure 4-27. Figure 4-28 illustrates the modulation pattern displayed on an oscilloscope when 100% modulation or less is employed. Superimposed modulation is normally undesirable. It is readily recognized because the negative portion of the modulated carrier is not inverted. Superimposed modulation is normally encountered as a result of hum or noise modulation. Figure 4-29 provides an example of superimposed modulation without the aid of actual modulator equipment. In addition, no sidebands are evident. Thus, only the modulation frequency and the carrier frequency are present. The superimposed distortion signal illustrated in Figure 4-29 has limited application in that this type of signal is sometimes applied to the input of an amplifier. If the output from the amplifier is reasonably close to being identical with the input, the amplifier



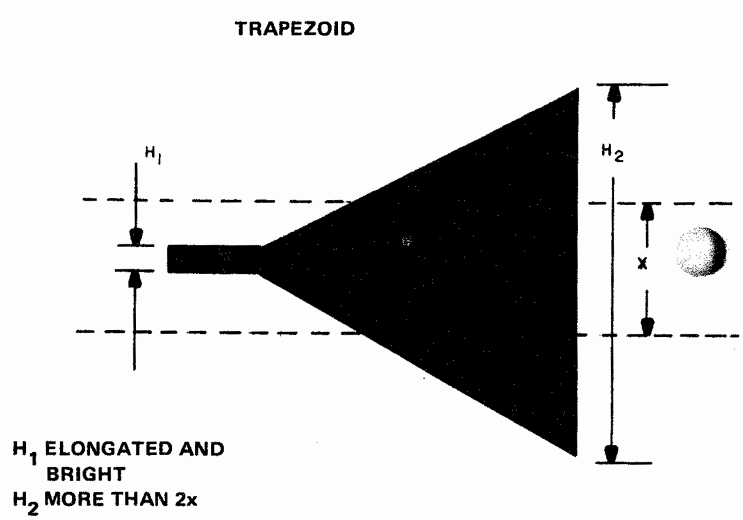


Figure 4-27. Overmodulated RF Carrier

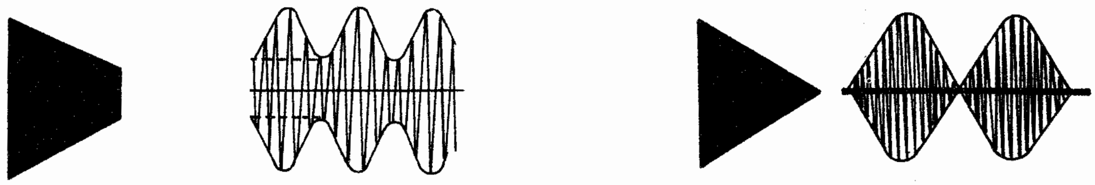


Figure 4-28. Undistorted Modulation

stage is relatively distortionless and therefore linear. However, if the amplifier stage is not distortionless, or linear (amplifies all applied frequencies within its range equally), the output will be amplitude-modulated by the amplifier stage. The output will therefore contain the upper and lower sideband-frequency harmonic components which were not contained within the original waveform. These frequency components will then modulate each other. This effect, termed "intermodulation distortion", is illustrated in Figure 4-30.

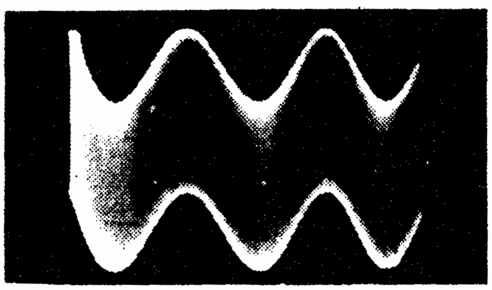


Figure 4-29. Superimposed Modulation

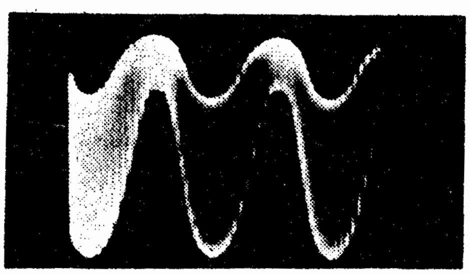


Figure 4-30. Intermodulation Distortion

4-10.2 FREQUENCY MODULATION

This method of modulating the constant-frequency, constant-amplitude carrier will also permit the transference of intelligence. However, considering that the majority of man-made or natural noise interference is amplitude modulation, the frequency modulation method is relatively free from noise or other interference during the process of intelligence transference. In the frequency modulation technique, the modulation signal represents the intelligence by frequency changes, rather than by amplitude changes. Therefore, when the modulation signal is used to modulate the constant-frequency carrier signal, no amplitude change occurs. The resultant RF modulated carrier signal will contain no amplitude variation, but its frequency will vary in accordance with the frequency and amplitude of the modulating voltage, as illustrated in part (B) of Figure 4-31. Increasing the modulation voltage causes the carrier frequency to decrease proportionally. Decreasing the modulation voltage has the reverse effect on the carrier. Increasing the frequency of the modulation voltage would increase the rate at which the carrier changed frequencies, and decreasing the frequency of the modulating voltage would decrease the rate at which the carrier frequency changes. The percentage of modulation is not a consideration in frequency modulation since FM is measured in terms of deviation and modulation index. Deviation is defined as the amount of shift to either side of the carrier frequency, and is directly proportional to the amplitude of the modulating signal. For example if a 1 MHz carrier were to be shifted 10 kHz to either side of its center frequency for each cycle of the modulation frequency, the resulting deviation would be 10 kHz. Modulation index on the other hand is defined as the ratio of deviation frequency to modulation frequency. Thus, if the same

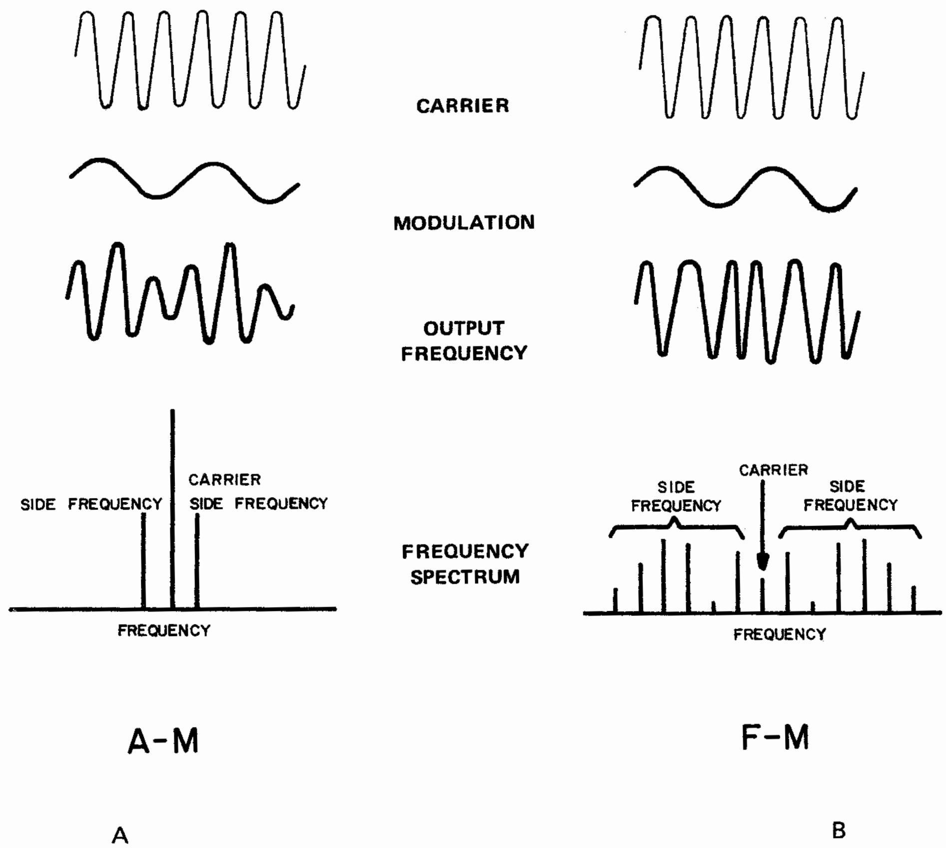


Figure 4-31. FM Patterns Compared to AM Patterns

1 MHz signal were to be modulated by 2 kHz at the same amplitude as in the foregoing example, the modulation index would then be 5 (10 kHz deviation/2 kHz modulation). Figure 4-32 illustrates the variations in carrier and sideband amplitude as the modulation index increases. The spectral display of Figure 4-31 illustrates that the power within the FM spectrum is distributed throughout the sideband in an amount proportional to the amplitude of the modulation voltage. In the AM spectrum, only one-fourth of the rated output power can be attained in the sidebands, and then only with 100% modulation. Furthermore, the sideband frequencies in the FM spectrum are separated from the carrier and from each other by an amount equal to the frequency of the modulation voltage. Theoretically, each modulating frequency, creates an infinite number of sideband frequencies, but in actuality these are limited by the response of the transmitter circuitry.

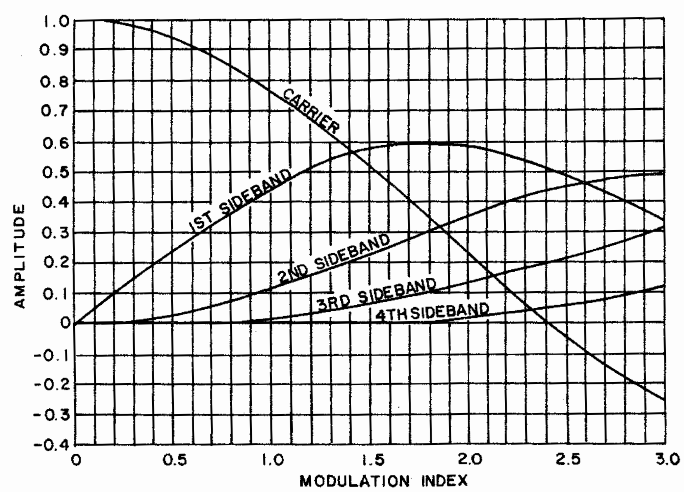


Figure 4-32. Bessel Curve for Frequency Modulation

4-10.3 PHASE MODULATION

Phase modulating of a constant-amplitude, constant-frequency carrier will result in basically the same type of transference characteristics as FM modulation. Phase modulation involves changing the carrier phase in direct accordance with the intelligence. The primary difference between frequency modulation and phase modulation is that in FM the deviation frequency of the carrier is a function of the modulation signal's voltage, whereas in phase-modulation it is a function of the modulation signal's frequency and voltage, as illustrated in Figure 4-33.

4-10.4 PULSE MODULATION

Pulse modulation is accomplished by periodically interrupting the carrier frequency. Either the amplitude or the width of the pulse can be varied as a means of transferring intelligence. In some applications, both width and amplitude are varied. The most common applications of pulse modulation are found in cw (continuous-wave) keying and in radar circuitry.

4-11 RESPONSE AND DISCRIMINATOR WAVEFORMS

4-11.1 RESPONSE CURVES

A response curve is a form of graph showing the relationship between output voltage and frequency. The response curve can indicate the degree of acceptance, amplification, or rejection, by either a component or a circuit, as the signal frequency is varied over a desired range. There are three primary types of response curves: single-peaked; double-peaked; and triple peaked, as shown in Figure 4-34. As shown in the figure, frequency is plotted along the horizontal axis, while amplitude of the output current or voltage is plotted along the vertical axis. The circuit response for a given input frequency is the measured amplitude separation between that point on the response curve representing the frequency and the horizontal zero reference line. The amplitude of the response curve may be shown either above or below the horizontal zero reference line, as illustrated in Figure 4-35. The half-power points shown in Figure 4-35 are 3 dB down from the peak or maximum amplitude point on the curve. The term 0.707 is an amplitude value above or below the horizontal zero reference level, and is obtained from the reciprocal of the square root of two ($1/\sqrt{2}$). A single-peaked response curve indicates the circuit is tuned to a single frequency, and will naturally provide a very narrow frequency pass band. A double-peaked waveform is the result of the deliberate design of transformer-type circuits which, when tuned to a single frequency, will provide a voltage maximum peak above the resonant frequency, and a voltage maximum peak below the resonant frequency. The resonant frequency will be represented by the dip between the two peaks, as illustrated by part (B) of Figure 4-34. The purpose of this type of waveform is to increase the frequency pass band by increasing the amplitude of a greater number of frequencies adjacent to the center frequency. The greater the dip between the two peaks, the greater the coupling between the primary and secondary windings of the transformer. Too great a dip is undesirable. A flat-topped curve is ideal because all frequencies within the pass band will then be the same amplitude. Several response curves may be algebraically added through a mixing circuit to produce a flat-topped, broad resultant response curve, as illustrated in Figure 4-36. The terms "overcoupled"

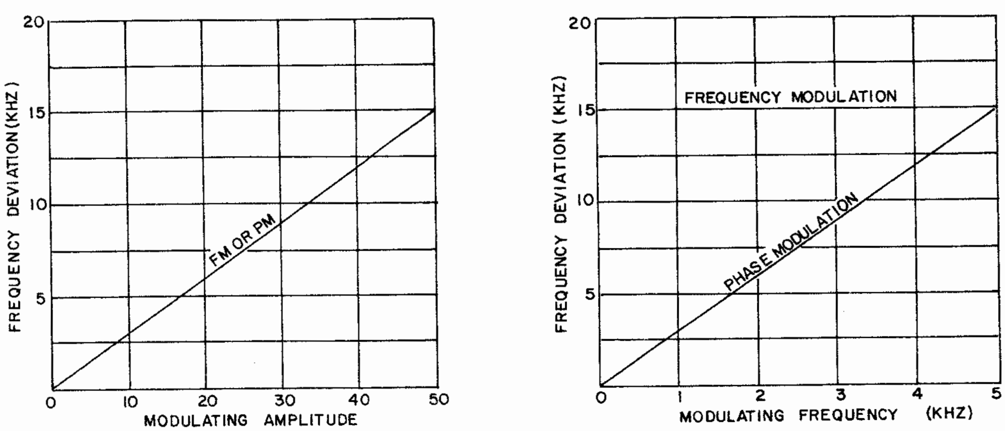


Figure 4-33. Frequency and Phase Modulation Characteristics

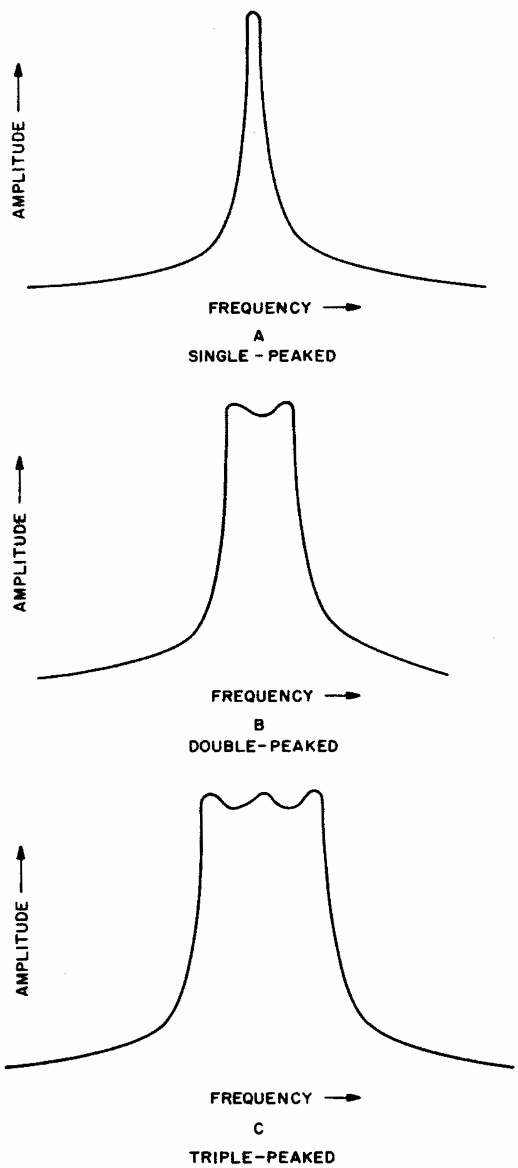


Figure 4-34. Primary Types of Response Curves

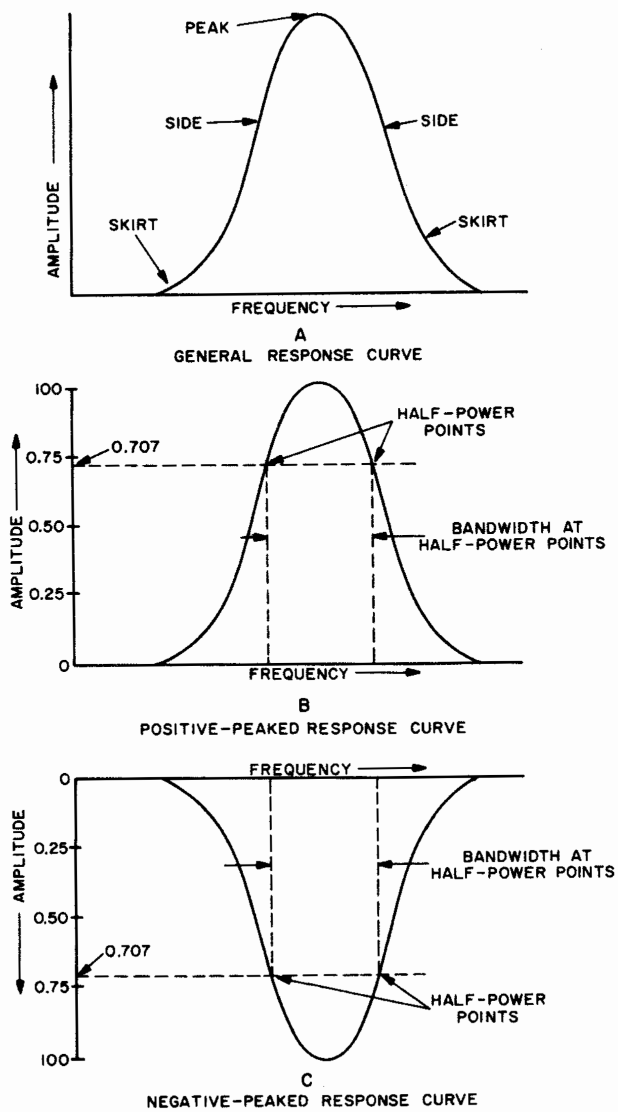


Figure 4-35. Positive and Negative Single-Peaked Response Curves

(close coupling) and "undercoupled" (loose coupling) refer to the spacing between the primary and secondary windings of the transformer. For example, if the primary is brought closer to the secondary (overcoupled), all frequencies within the bandwidth will be transferred from the primary to the secondary with approximately the same amplitude; this provides a wider pass band, less frequency selectivity, and greater over-all amplitude. However, if the primary and secondary windings are moved farther apart, more impedance is effectively placed between the two windings, and only the frequencies containing the greatest amplitude will have sufficient energy to bridge the gap. This will create a sharply-peaked waveform in the output, representing a very narrow bandwidth, high frequency selectivity, and less over-all amplitude, even

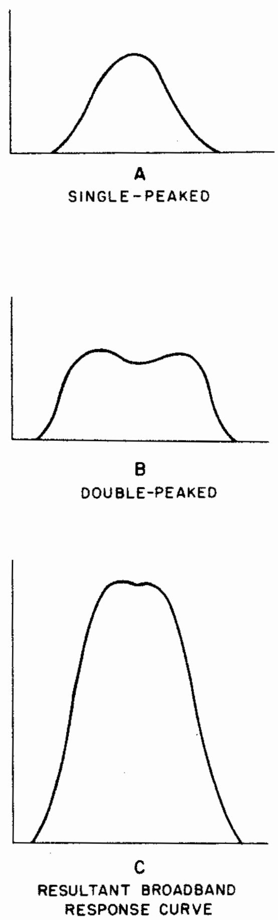


Figure 4-36. Response Curve Combinations To Produce a Required Resultant Wide-Band Response Curve

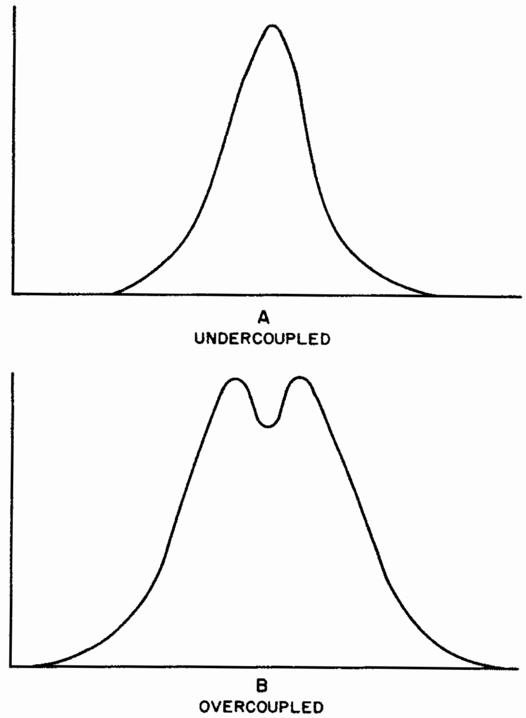


Figure 4-37. Response Curve Coupling

though the waveform peak is more pronounced. These effects are illustrated in Figure 4-37. Broad-banding, or the technique of increasing the bandwidth to permit a greater number of frequencies to pass, is accomplished by two primary methods: overcoupling, as was just discussed; and stagger tuning. The term "stagger tuning" refers to the tuning of a series of circuit stages to slightly different frequencies. For example, three stages could be tuned 1000 cycles apart from one another, as illustrated in Figure 4-38, to combine as a resultant waveform. This resultant waveform is considered as a triple-peaked response curve.

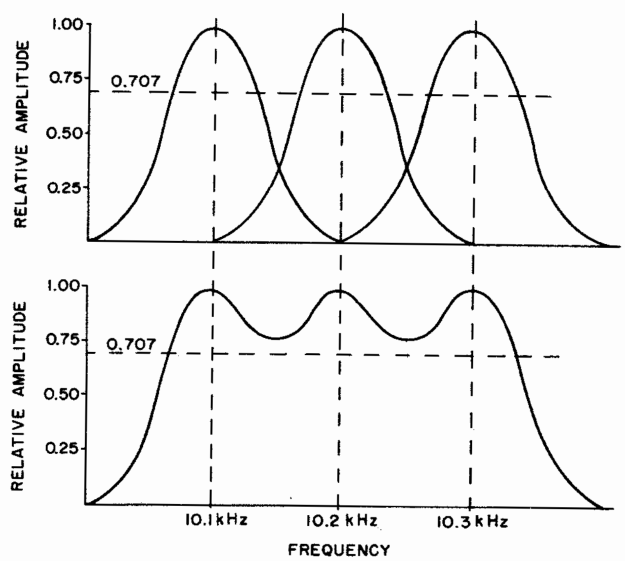


Figure 4-38. Response Curve Resulting from Stagger-Tuned Stages

4-11.2 DISCRIMINATOR CURVES

The output from a discriminator circuit is sometimes referred to as an "S" curve. Figure 4-38 illustrates the ideal form of the "S" curve used as a reference standard. Any deviation from this shape represents incorrect tuning of the primary or secondary transformer windings, or other improper circuit adjustment. The "S" curve is linear and always crosses the horizontal zero reference axis at the point on the curve representing the center frequency. Many times a marker pulse is electronically added so that it appears at some point on the curve. However, this marker will disappear at the center frequency because this point occurs at zero voltage amplitude. The positive amplitude and low-frequency components on one side of the center frequency should equal the negative amplitude and high-frequency components, respectively, on the opposite side of the center frequency. In other words, A = A (amplitude) and B = B (frequency separation) in Figure 4-39. The audio frequency response curve shown in Figure 4-40 is ideal. The constant height of this response curve proves that the circuit under test has a flat response from its lowest to its highest frequency. The horizontal zero reference base line is useful for measuring relative amplitudes. Considering that the portion of the wave below the reference or base line is exactly the same as the wave above the base line, only the top half of any "S" curve may be observed for full information. Any peaks extending above the average amplitude of the waveform represent accentuation of the frequencies within that region of the pass band, and valleys or dips reflect attenuation of the frequencies at those points. Therefore, this waveform as an input not only shows the circuit behavior as a whole, but instantly reflects any unusual frequency characteristic of any recently-added components, filters, or circuits. For example, Figure 4-41 represents an undesirable voltage-frequency characteristic within an LC filter circuit resonant at 4000 Hz. Video and other high-frequency response curves are similar to low-frequency (audio) response curves. However, in high-frequency curves the frequency band pass is wider (broader), with an extremely low-frequency limit (60 Hz) and an extremely high-frequency limit (in the megacycle range). Two different types of markers may be used to designate exact frequencies. The first marker is a disturbance along the response curve at a particular frequency, whereas the second is produced by a tuned circuit which removes or absorbs energy from the response curve at a particular frequency. Both types of markers are illustrated on the typical high-frequency response curve shown in Figure 4-42.

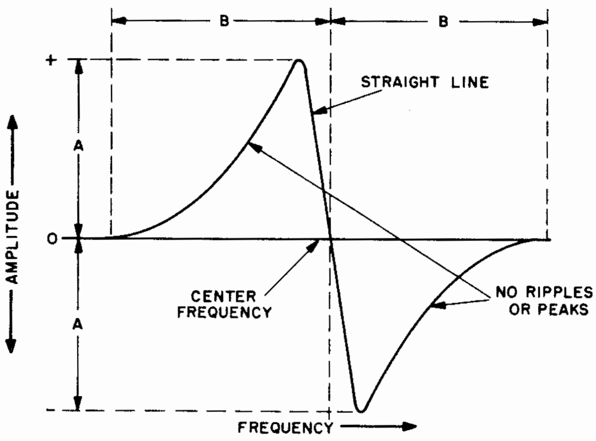


Figure 4-39. Discriminator "S" Curve

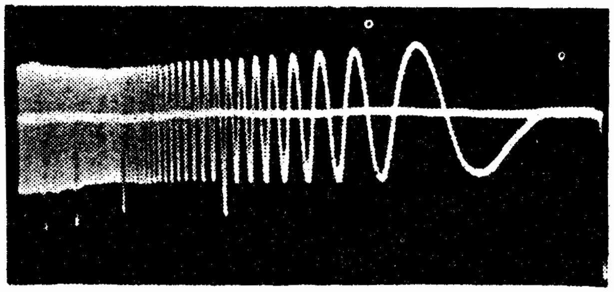


Figure 4-40. Ideal Audio-Frequency Response Curve

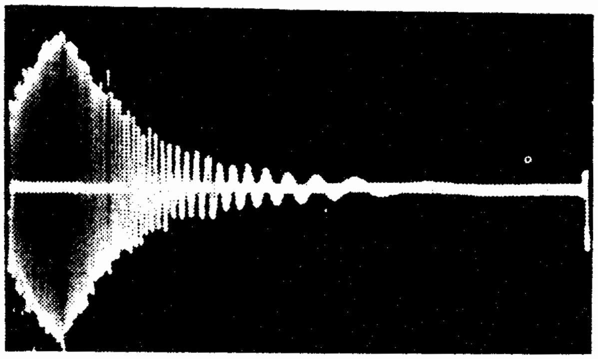


Figure 4-41. Resonant Circuit Audio-Frequency Response Curve

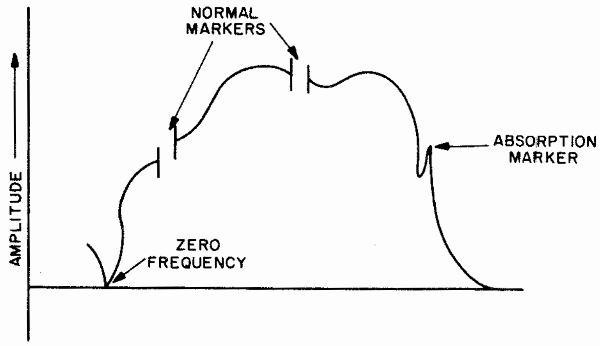


Figure 4-42. High-Frequency Response Curve

Figure 4-43 illustrates still another type of high-frequency response curve. However, the curve shown in Figure 4-43 has not been demodulated and is not popular because frequency markers are very difficult to discern on this type waveform.

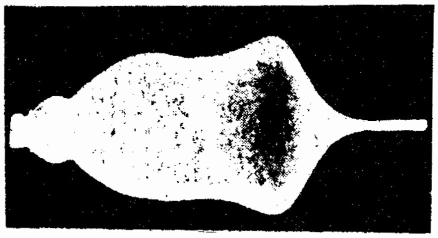


Figure 4-43. Non-demodulated High-Frequency Response Curve

4-12. INTENSITY MODULATED PRESENTATIONS

The most common usage of intensity (relative brightness) modulation occurs in television. However, intensity modulation is also used to display signals on a raster (continuous display), or to brighten the mark to space transition when taking end-distortion measurements of teletype signals with a digital data distortion test set. Intensity modulation is also employed in comparing frequencies in excess of 10:1, provided the frequency ratio involved is an integer such as 10:1, 20:1 and 30:1. In all instances, the information is obtained from the display by noting the degree of intensity of the display. This intensity can vary from zero magnitude to a very bright illumination.

Intensity modulation is accomplished by modulating the Z axis -- the electron beam -- of the display scope. In television, the video signal modulates the beam. In a raster display, intensity modulation could result from a received signal at a specific frequency and in a given spectrum. On oscilloscopes, intensity variation is accomplished by the signal input to the Z axis. Shaping circuits are incorporated in the Z axis circuitry of an oscilloscope to ensure a definitive presentation, regardless of the type of signal being applied.

4-12.1 COMPARING TWO FREQUENCIES

As the ratio of two frequencies being compared increases, the Lissajous pattern becomes more difficult to retain in a stationary position, and the counting of multiple loops becomes a more difficult task. For these reasons, the intensity modulation method of obtaining frequency ratios can be used to advantage. A circular pattern, obtained from the low-frequency signals, is passed through an appropriate phase-shifting network and applied to the vertical and horizontal inputs of the oscilloscope, as illustrated in Figure 4-44. The high-frequency signal is connected to the intensity modulation terminal of the oscilloscope and the low frequency signal then serves as the reference signal. In part (A) of Figure 4-45, the frequency ratio is 10:1. There are therefore 10 blanked-out segments of the original circular display. In part (B) of Figure 4-45, the frequency ratio is 20:1, thus there are 20 blanked-out segments in the pattern. The number of blanks in the pattern is therefore equivalent to the ratio of frequencies. Because of this appearance, such displays are often called "spot-wheel" patterns.

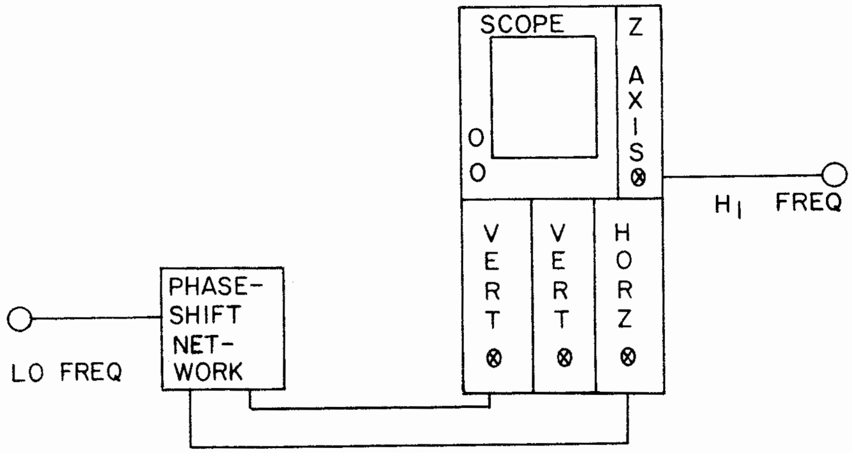


Figure 4-44. Phase-Shifting Circuit for Circular Sweep Displays

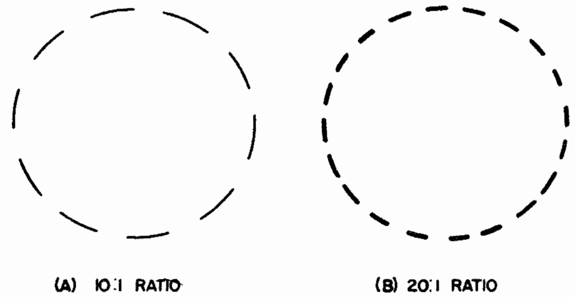


Figure 4-45. Spot-Wheel Patterns

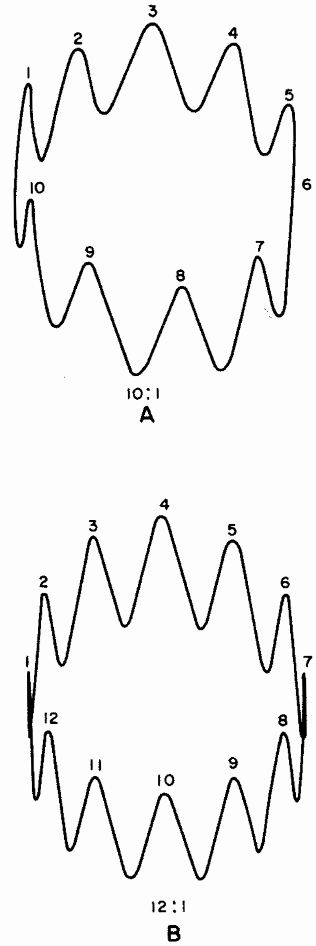


Figure 4-47. Integral Frequency Ratios In Circular Sweeps

4-13 CIRCULAR SWEEP PRESENTATIONS

Figure 4-46 shows the circuit connections for employing the circular sweep method of obtaining frequency ratios greater than 10:1. The circular sweep is developed by the low-frequency signal. The high-frequency signal is then applied to either input terminal of the oscilloscope. Figure 4-46 shows the circuit for only one situation: the case where the high-frequency signal variations are superimposed on the circular sweep through the action of the vertical deflecting plates in the circuit. The resultant patterns of Figure 4-46 are illustrated in Figures 4-47 and 4-48.

4-14 WAVEFORM DISTORTION

Distortion is normally considered as a deviation from the desired waveform. However, the undesirable waveform in one application may be the desired waveform in some other applications. Therefore, the term "distortion" refers to a particular waveform application, and is meaningless if no application is being considered. A normal high-frequency current is characterized by its amplitude, frequency, and phase relationships, and can be altered by changing any one of these characteristics. Actually, any two (or possibly all three) characteristics may be altered by a circuit change. If the circuit change produces the desired signal, this new signal is termed an "undistorted" or pure waveform; if the circuit change produces an undesired signal, it is termed a "distorted" waveform. The factors contributing to waveform distortion in one application may be the same factors required to produce a desired waveform in some other application. This paragraph considers and explains only those

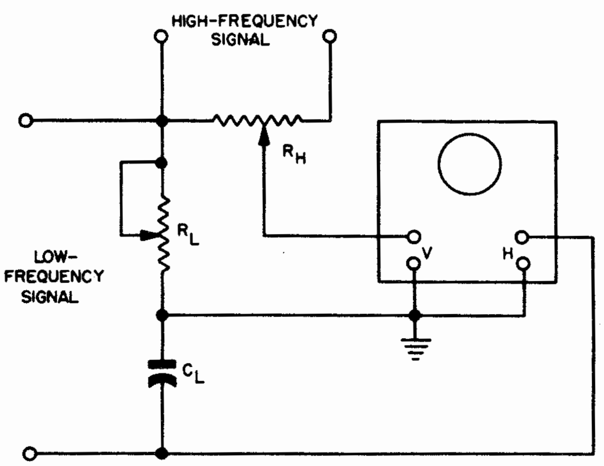


Figure 4-46. Circular Sweep Comparison Circuit, Using Deflection Systems Only

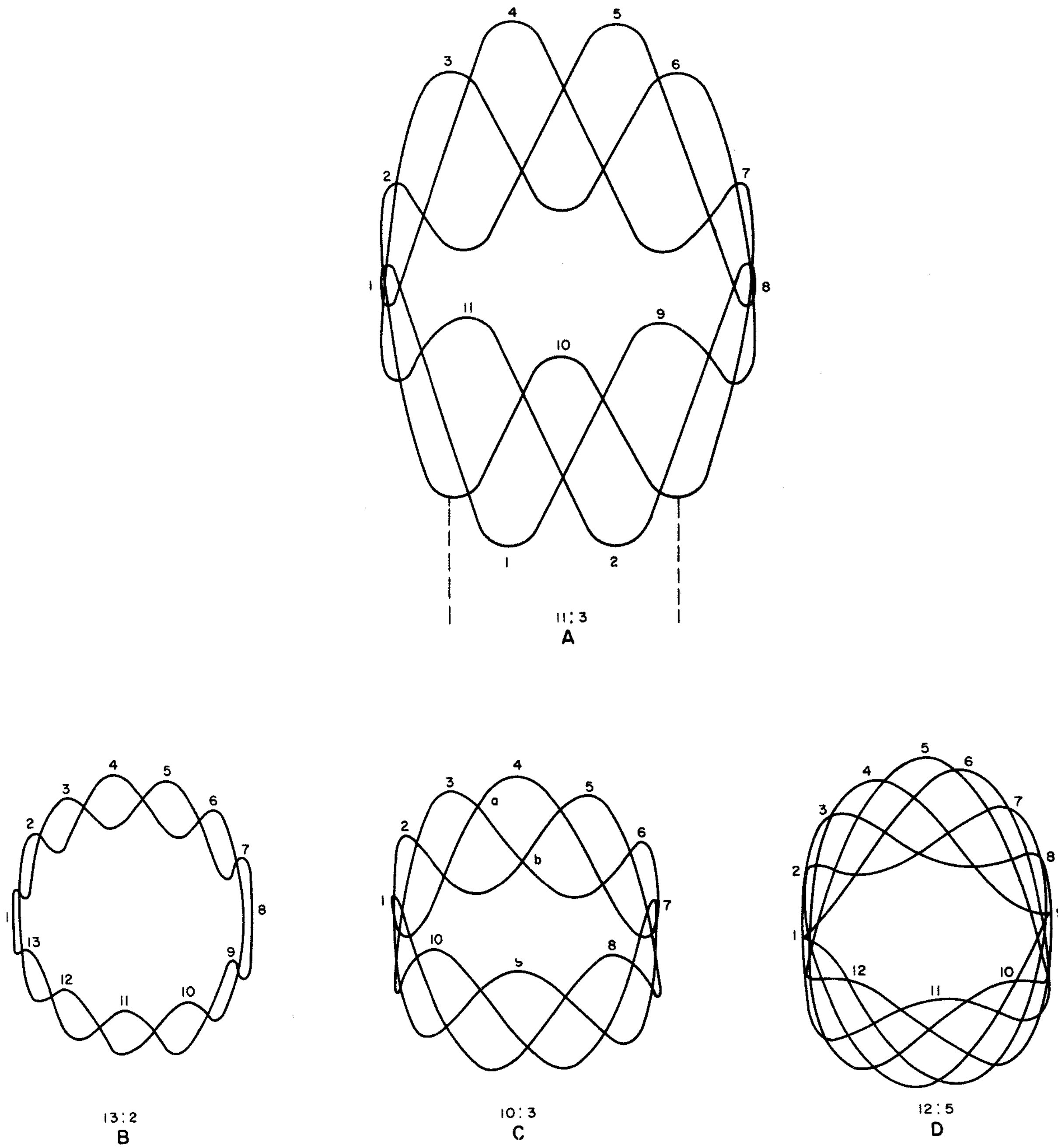


Figure 4-48. Fractional Frequency Ratios in Circular Sweeps

undesirable factors which contribute to distortion. The primary cause of distortion created within an actual circuit can be traced to overloading of the active component in that particular circuit. For example, overloading an active component will cause it to operate on a nonlinear portion of its characteristic curve, and the output waveform will not retain the same shape as that of the input. The same overloading effect can occur if the active component is defective or if one or more of the applied operating inputs is incorrect. For other than active components or circuit defects and incorrect operating inputs, distortion can be eliminated by simply decreasing the input amplitude (volume control) or intensity. The placement of components, wires, or leads may create undesirable feedback voltages in a phase relationship which results in distortion. Distortion-removing circuits, designed to eliminate feedback, may be defective. Neutralization circuits may be used to remove or balance out distortion resulting from undesirable feedback. The elimination or overemphasis of the amplitude of particular frequencies, within a desired band or range of input signal frequencies, will create distortion. The primary and secondary windings of frequency-sensitive transformers may be incorrectly tuned or be spaced an incorrect distance from each other. Therefore, the sideband frequencies, which form an important part of the resultant desirable signal, may be missing from the output signal. Finally, defective input or output components may blank out certain pass-band frequencies or permit undesirable voltages to pass, thereby causing distortion.

4-14.1 AMPLITUDE DISTORTION

Amplitude distortion may be caused by a limitation of bandwidth or by irregularities within the bandwidth. In either event, amplitude distortion is normally expressed in terms of attenuation because it is a logarithmic quantity which is algebraically added for cumulative stages. Amplitude distortion, free of phase distortion, cannot change the symmetry of a symmetrical input pulse. The response of a circuit should be the same for all frequencies present in the input signal voltage. However, if the circuit response is not the same for all input frequencies, suppression or exaggeration of the amplitude of some frequencies will create distortion. The fundamental plus harmonics will be seen or heard in the output waveform when amplitude distortion exists. In the case of an amplifier stage, it can be determined whether amplitude distortion is present by applying a signal voltage of known characteristics to the amplifier input and then viewing or measuring the output signal. The output waveform should be a replica of the input waveform. Amplitude distortion caused by an amplifier is the result of the generation, by the amplifier, of frequencies which were not contained in the input. The result of generating additional frequency components is seen by the change in waveform amplitude.

4-14.2 FREQUENCY DISTORTION.

Frequency distortion occurs when different frequency inputs are not all amplified equally. The distortion may be audible or inaudible, depending on the circuit frequency response limits. In addition, if the circuit output load is composed of reactive components, the low-frequency resonance and the increase in inductive resistance at high frequencies will increase the nonlinear (amplitude) distortion and modify the response. If a feedback network contains reactive elements, then the over-all gain of an associated stage is a function of frequency, and frequency distortion due to feedback will be obtained. However, negative feedback, even when reactive elements are present, will decrease the total circuit distortion at the expense of maximum gain. The distortion in linear amplifiers as a result of the relationship between the input voltage and output voltage is a type of frequency distortion as well as of amplitude distortion. With a square waveform applied to the input of a linear circuit, the output should also be a square waveform. However, if the circuit response is not the same for all frequencies, the output waveform will not be the same shape as the input waveform. Figure 4-49 shows output waveforms for nonlinear circuit response. Frequency modulation distortion is often termed "flutter" distortion. This type of frequency distortion is generally the result of speed fluctuations as a recording is driven by the recorder or reproducer motor. The flutter effect may also be caused by a loudspeaker when it is reproducing two frequencies simultaneously. This is true because the sound pitch is a function of the relative velocities and sources with respect to the listener. Both linear and nonlinear loudspeakers produce this type of distortion.

4-14.3 INTERFERENCE DISTORTION

Figure 4-50 illustrates two signals, separated slightly in frequency and differing in amplitude. The third waveform is the resultant obtained when the desired and undesired signals are combined algebraically at every point. The amplitude of the resultant varies at a rate equal to the difference in frequency between the two original signals. If both signals differ in frequency by 100 Hertz, the resultant waveform amplitude will change 100 times per second. In an AM receiver, this amplitude would be separated by the detector and be heard as a whistle from the speaker. The resultant waveform may at times lead, lag, or be in phase with the desired signal. The resultant

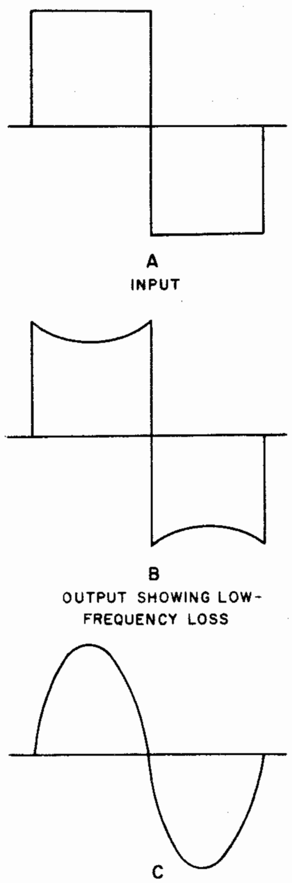


Figure 4-49. Output Waveforms Resulting from Poor Circuit Frequency Response

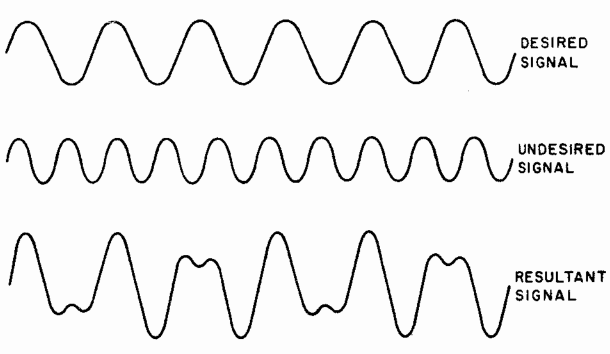


Figure 4-50. Combination of Two Signals Forming an Amplitude-and Phase-Modulated Resultant

is therefore phase-modulated. This phase modulation (and, indirectly, frequency modulation) is directly proportional to the amplitude difference between the two signal carriers. When the amplitude ratio between the two signals is 2 to 1, the phase angle shift is slightly under 30 degrees. The rate of phase shift change is in direct proportion to the frequency difference between the two original signals. Static is primarily a form of amplitude distortion caused by uncontrolled electrical waves associated with thunderstorms and other natural phenomena. The strength of these waves is sometimes great enough to drown out the desired station or prevent clarity or reception. Limiter stages will limit incoming bursts of static amplitude, and, by selecting a narrow bandwidth, much of the continuous crackling variety of static can be removed through frequency selectivity. For FM reception, transmission allocations are in the higher-frequency bands, where static amplitude changes are not very effective; most of the outburst energy is limited to lower frequencies. Even if no external natural disturbances or other station interference is present, internal active component and circuit noises exist which will limit the weakest received signal to some minimum amplitude. Any signal lower than this minimum amplitude will not be amplified with clarity. "Thermal agitation" is the term applied to the noise created by the random motion of electrons in any conductor. The thermal noise produced is proportional to the amplifier bandwidth. "Tube hiss" is the term applied to the noise created by "shot effect", which refers to the fact that electrons moving through a vacuum tube are a congregation of separate particles which do not impinge on the anode as a continuous fluid movement, but rather as sporadic fluctuations. This tube hiss noise is normally distributed evenly throughout the frequency spectrum. Transistors generate noise by the shot effect in the bias current, and their thermal noise is caused by inherent resistance in the base region. It is thought that surface recombination of electrons may also be a source of semi-conductor noise that only becomes significant at very low frequencies. Impulse noise, as distinguished from random noise, consists of external sharp bursts of energy. Normally, this noise is associated with automobile ignition systems and sparking gaps in electrical machines. A limiter stage is required to decrease the effects of this type of interference. Hum interference is normally a result of insufficient filtering in the power supply, of heater-cathode leakage of unshielded transformers and chokes in the DC line, or of defective coupling between circuits. This type of interference, as is true of other types of circuit noise, will combine with the desired signal and produce distortion.

4-15 USE OF LISSAJOUS FIGURES

4-15.1 PHASE RELATIONSHIPS

Lissajous figures can be used to measure the phase relationship existing between two voltages of the same frequency. The patterns involved appear as ellipses with different degrees of eccentricity. As shown in Figure 4-51, the pattern if formed when two sine waves of the same frequency are applied to the vertical and horizontal input terminals of the oscilloscope. Point-to-point plotting of like-numbered projections will verify the formation of the resultant pattern. In order to measure the angle of phase displacement, it is necessary to use an oscilloscope with a cross-section screen, called a "graticule", to provide a graph of the X- and Y-axis coordinates. If two sine waves of unequal amplitude are used, the resultant pattern will always be elliptical in form and could not be used intelligently. In actual phase measurement, unequal amplitudes of the input to the scope are compensated for by adjusting the horizontal and vertical gain controls. The vertical gain is first reduced until a straight horizontal line is obtained. The horizontal gain control is then adjusted for some convenient length: for example, 2 inches. The next step is to place the horizontal function or selector switch in the horizontal input position. A small spot near the center of the oscilloscope will now be obtained, depending upon the relative adjustments of the vertical and horizontal positioning controls. Apply the signal to be measured to the vertical input terminals, and increase the vertical gain control to the same length of line as previously obtained; in this case, 2 inches. Since there is no horizontal deflection, the 2-inch trace will be only a thin vertical line. At this point the gain of both amplifiers is equalized, and the technician may apply the comparison signal to the horizontal input terminals and proceed with the phase measurement technique. However, it is to the technician's advantage to make one further check for equalization. Connect a jumper wire from the vertical input terminal of the scope to the horizontal input terminal so that the same signal is applied simultaneously to both amplifiers. The pattern should tilt over to a 45-degree line intersecting the corners of the 2-inch square, as shown for 0 degrees in Figure 4-52. The procedure just given is not the only method of equalizing the oscilloscope amplifiers, but it is applicable to any oscilloscope, and is not subject to any special switching positions provided on a specific oscilloscope. If the phase angle displacement between two input frequencies remains at a fixed angle, the phase angle may be calculated. Count the number of divisions along the vertical Y axis to the point where the ellipse intersects the Y axis. This number if known as the "Y-axis intercept", or Y_1. Next count the number of divisions along the vertical Y axis to the point on that axis which indicates the maximum vertical amplitude of the ellipse. This number is known as the "Y-axis maximum", or Y_2. The angle of phase difference, θ, is found by the following calculation.

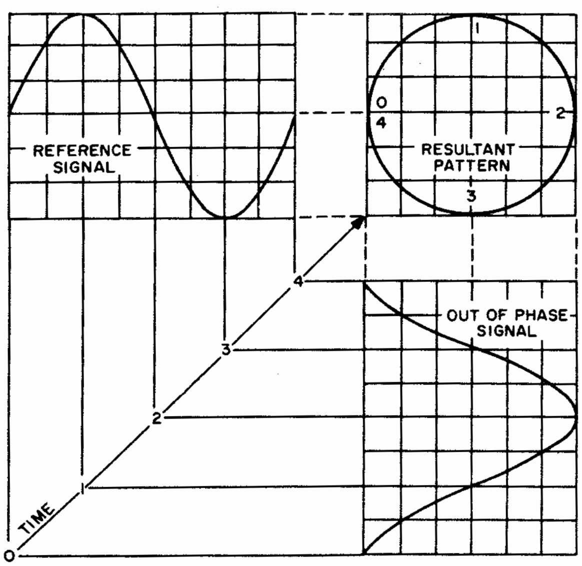


Figure 4-51. Formation of a Lissajous Figure, Illustrating 90 Degrees Phase Difference

$$\text{Sine } \theta = \frac{\text{- Y axis intercept}}{\text{y-axis maximum}} \text{ or } \frac{Y_1}{Y_2}$$

Reference to a trigonometric function table, or use of an electronic calculator or conventional slide rule will enable the technician to convert this ratio into a phase angle expressed in degrees and minutes. A similar procedure, using the horizontal X axis in the same manner, will produce the same results. The direction in which these values are measured can be either positive or negative. The ratios obtained are independent of the direction taken when counting. Assume the calculated ratio is 0.5, which, when converted into angles from the trigonometric tables, gives the phase angle difference of 30 degrees. Reference to Figure 4-53 shows that the major axis of one of the

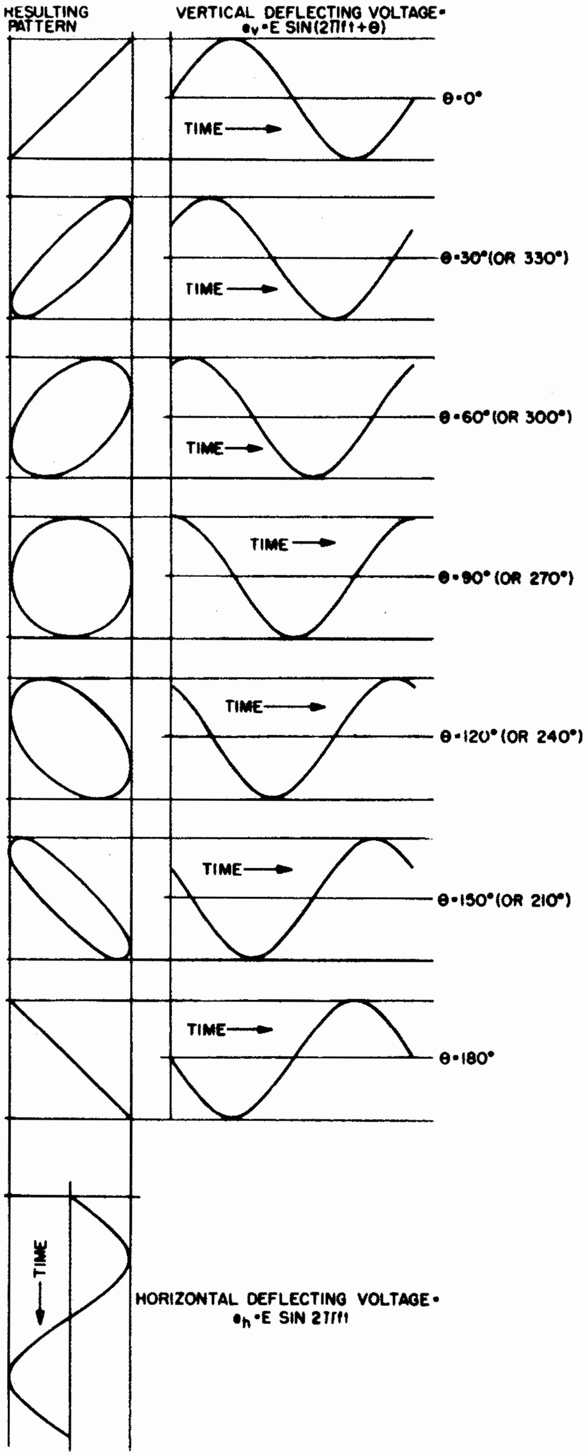


Figure 4-52. 1:1 Lissajous Patterns, Showing Effect of Phase Relationships

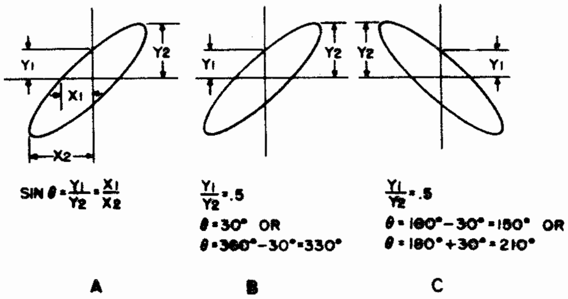


Figure 4-53. Computation of Phase Difference

ellipses lies in the first and third quadrants, and that the phase angle could therefore be 30 degrees or 330 degrees. This ambiguity of two possible phase angles is not surprising. Whether a signal is 30 or 330 degrees ahead of a second signal, the difference in phase is still 30 degrees. It is this difference that a Lissajous figure has the ability to indicate, and not which signal leads or lags the other. Fortunately, it is not difficult to learn which is the leading signal when the information is not known from other sources. Assume that a phase difference of 30 degrees is computed. If it happens that the signal applied to the vertical deflecting plates leads the horizontal signal by 30 degrees, an additional phase advancement of the vertical signal will reduce the eccentricity of the ellipse; that is, it will be made to resemble a circle. On the other hand, if the vertical signal lags by 30 degrees (equivalent to leading by 330 degrees), an advancement in phase will bring the two signals more nearly into phase. Consequently, the ellipse will continue to contract until eventually it becomes a straight line. There are a variety of circuits for shifting the phase of a signal, such as the one shown in Figure 4-54. One of the two signals under investigation, such as the signal applied to the vertical deflecting plates, can be impressed across a series circuit consisting of a potentiometer and a capacitor. At the frequency concerned, the resistance of the potentiometer should be about 10 times the reactance of the capacitor. An output from the network can be taken from either the resistance as shown in Figure 4-54, or from the capacitor. If the signal developed across the capacitor is desired, the ground connection should be made to the input side of the capacitor. If the output signal is derived from the resistance, its phase will be advanced relative to the original signal; if taken from the capacitor, the phase will be retarded. It shall be assumed in this case that the signal across the resistance is applied to the vertical

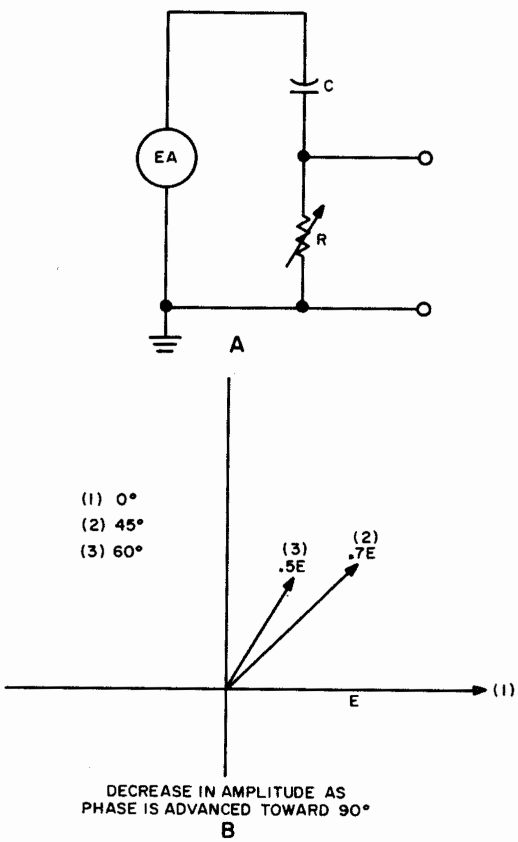


Figure 4-54. Phase-Shift Network

terminals of the oscilloscope. If the vertical signal leads the horizontal signal, the ellipse will become broader as the resistance of the potentiometer is decreased. Most likely a circle will not be obtained, since the amplitude of the signal also decreases as the resistance becomes less. In those cases where one of the two applied frequencies is constantly changing phase with respect to the other, the resultant ellipse changes form, and the plane of the ellipse appears to rotate around either of two imaginary diagonal axes. As the phase difference increases from zero to 90 degrees, the plane of the ellipse appears to rotate around one of the imaginary diagonals; and as the phase difference increases from 90 to 180 degrees, the plane of the ellipse rotates around the opposite diagonal axis. When an oscilloscope is used to determine phase relationships, several precautions must be observed. It is imperative, first of all, to know whether the circuits in the oscilloscope ahead of the deflecting plates have unequal phase-shift characterisitics. If there is an inequality, the indicated phase relationship of the two signals undergoing investigation will be in error by the amount of the inequality. To determine the amount of phase error introduced by the oscilloscope circuits, apply a sine wave simultaneously to both the horizontal and vertical input terminals of the oscilloscope. If a straight line is displayed in the first and third quadrants, no phase shift is introduced by the oscilloscope amplifier. If the straight line appears in the second and fourth quadrants, a 180-degree phase shift is introduced by the amplifying stages of the oscilloscope, probably because the number of stages in the two sections are unequal. It is important to check this possibility, as the design requirements of the sections are not generally the same. The appearance of an ellipse, however, discloses an inherent disparity in the phase characteristics of the two amplifiers, rather than a mere difference in design. This phase difference (in degrees) must be determined, and then added to or subtracted from the result of a phase measurement of the two signals according to which amplifier has the leading characteristic. This check of amplifier characteristics, which should be made over the entire frequency range of interest, is especially important in the low- and high-frequency portions of the band passed by the amplifiers. Astigmatism in an oscilloscope may be so pronounced that accuracy in measuring Y-maximum and Y-intercept is difficult. In this case the trace may be in focus over one region of the tube face, but out of focus in other regions. Wherever the trace is poorly defined, there will be uncertainty in a measurement of distance. For an accurate determination of the sine of the phase angle, it is necessary that Y_1 and Y_2 be measured accurately. This means that the intersection of the X and Y axes must be placed in the exact center of the ellipse.

4-15.2 2:1 LISSAJOUS PATTERNS

The concept of Lissajous patterns was developed on a limited basis in Section 2 of this manual with respect to multiple patterns and time basis, and with respect to the phase-angle measurement in the preceding paragraphs. A more detailed discussion of Lissajous figures will now be presented, with special attention given to frequency comparisons. There are many possible configurations for any ratio of applied frequencies. One consideration is whether the higher or lower frequency is applied to the horizontal deflecting plates. The most significant consideration, however, is the "phase" of the high-frequency signal with respect to that of the low-frequency signal when the latter is beginning a cycle. Strictly speaking, "phase" in this sense is a misnomer, as the definition is normally in terms of a single frequency. Nevertheless, a cycle of the high-frequency signal is often well advanced at a time when a cycle of the low-frequency signal has just commenced; for convenience, this condition is usually referred to as a "difference in phase". Figure 4-55 shows the situation that prevails when both applied signals start at the same time. The resulting pattern can be likened to a figure "8" resting on a side. In parts (A) and (B) of Figure 4-56 a line drawn against the top edge of the pattern, called a "tangent", would make contact with the pattern at two places. Similarly, a line drawn against a vertical side would be tangent at only one place. Notice that the horizontal tangents correspond to the vertical deflecting voltage, and that the vertical tangents correspond to the horizontal deflecting voltage. Hence, the ratio of the vertical deflecting frequency to the horizontal deflecting frequency is 2:1.

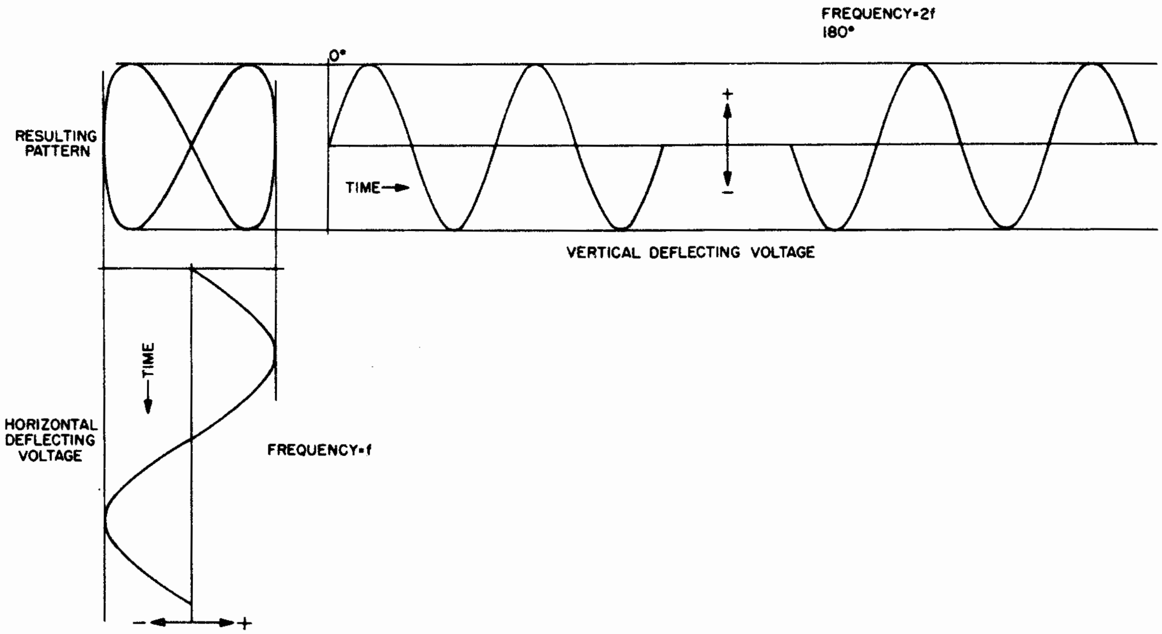


Figure 4-55. 2:1 Lissajous Pattern

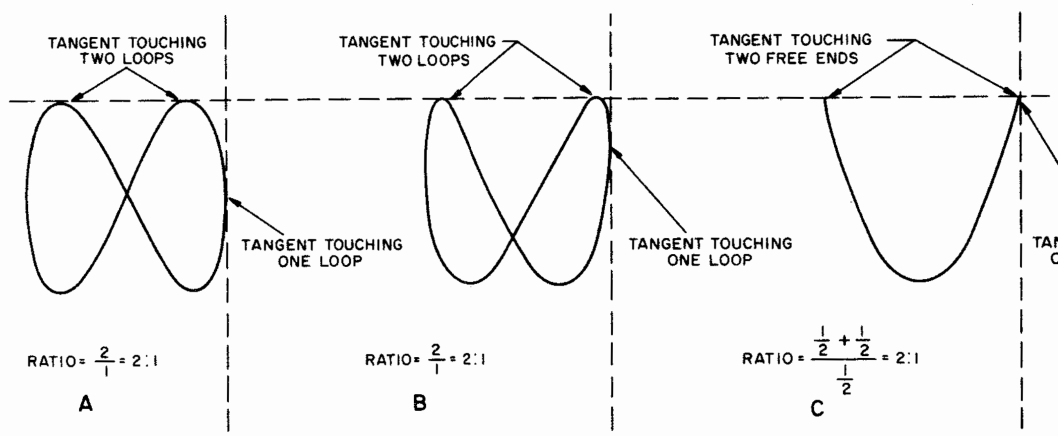


Figure 4-56. Calculation of 2:1 Frequency Ratio

If the two signals were applied to the opposite sets of deflecting plates, the resulting pattern would be rotated 90 degrees. An interesting situation exists when the high-frequency signal is shifted ahead 90 degrees in phase. As shown in part (B) of Figure 4-57, the high-frequency signal may be at its maximum value when the low-frequency signal is just beginning a cycle. When this condition occurs, the two loops are closed into the

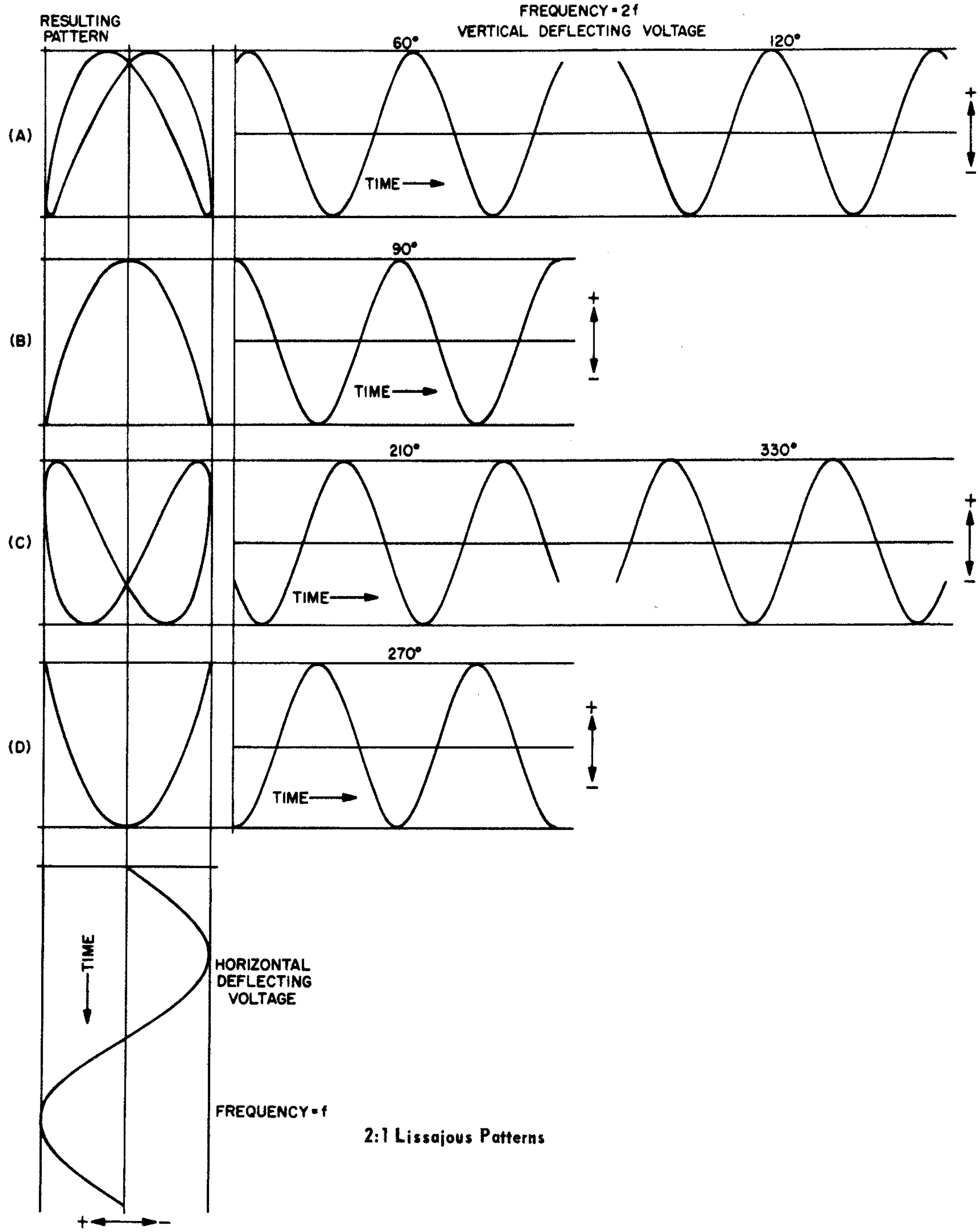


Figure 4-57. 2:1 Lissajous Patterns for Various Phase Relationships

form of a parabola, with its cup pointing downward. Similarly, if the high-frequency signal is at its most negative value when the low-frequency signal is commencing a cycle, the pattern is a parabola with its cup pointing upward, as shown in part (D) of Figure 4-57. This type of pattern is commonly referred to as a "double image" because the electron beam, after reversing its direction, traces out the same path. A double-image pattern is also called an "uncompleted loop" or "closed pattern". Each type of 2:1 Lissajous pattern, except the parabola, is developed for two phase relationships. For example, the pattern of Figure 4-55 is also generated when the high-frequency signal is 180 degrees out of phase with the low-frequency signal. These alternative phases are illustrated in Figures 4-55 and 4-57 by the high-frequency signals which produce the vertical deflection. When a double image such as the parabola is developed, a somewhat different method of evaluating the frequency ratio must be employed. If the contact of an open line against the side of a pattern is counted as one-half, the correct ratio can be determined as shown in part (C) of Figure 4-56. There is only one contact, against the vertical line, giving a figure of 1/2. There are two contacts against the top horizontal line, giving a total of 1. The ratio of the vertical deflecting-plate signal frequency to the horizontal deflection-plate signal frequency is therefore 1:1/2, or still 2:1. Now assume the horizontal line had been drawn against the bottom edge of the pattern. Here the rounded end, or closed loop, of the parabola clearly has a single point of tangency with the line, giving a total of 1. The ratio is now, therefore, 1:1/2, which is still 2:1.

4-15.3 3:1 LISSAJOUS PATTERNS

Analogous conditions hold when the frequency ratio is 3:1. Representations of the various patterns that may be obtained are shown in part (A) of Figure 4-58. If the signals applied to the deflecting plates are interchanged, the resulting patterns are

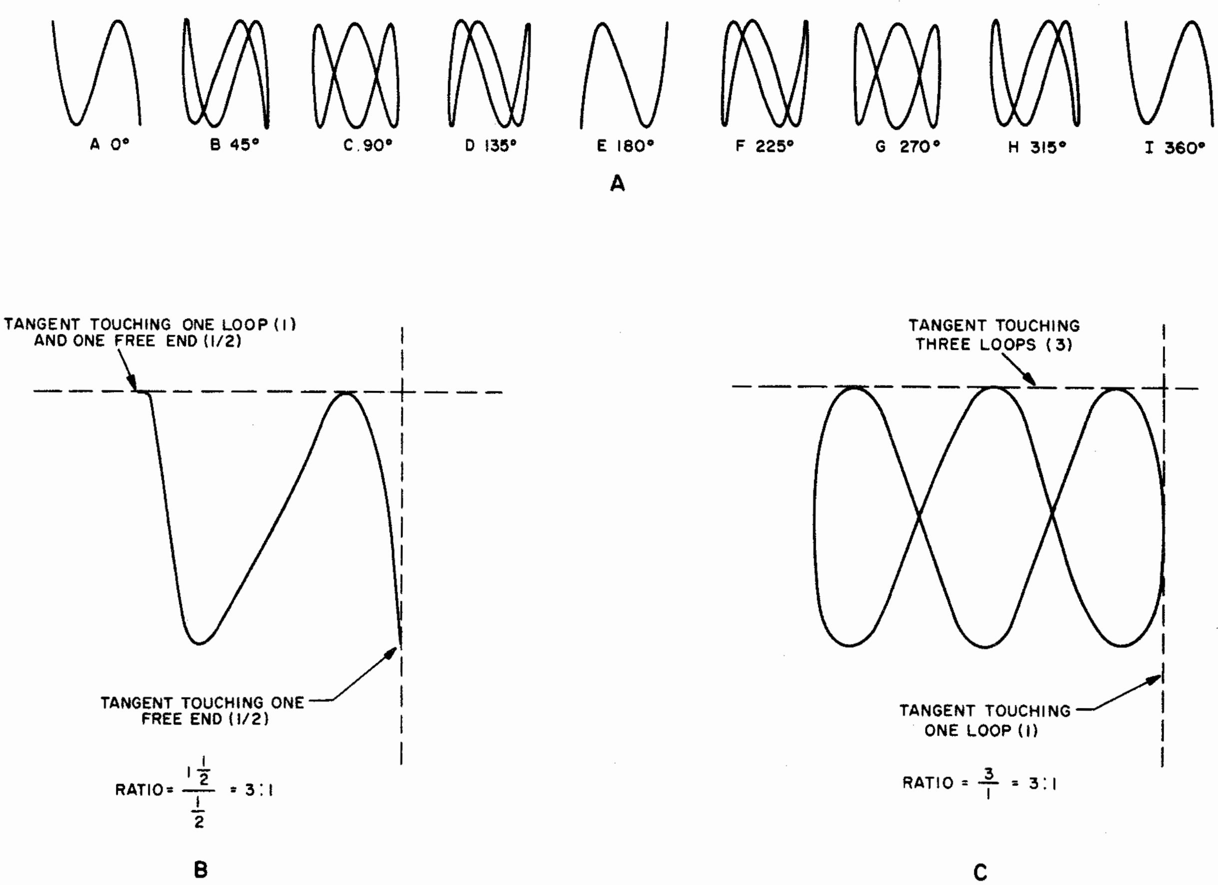


Figure 4-58. 3:1 Lissajous Patterns and Calculation of Frequency Ratio

exactly the same except for an axis rotation of 90 degrees. In the case of the S-shaped curve, the frequency ratio is computed by the same procedure as described for closed patterns in the previous paragraph. To illustrate, there would be a tangency and a contact with respect to a horizontal line drawn against the pattern shown in part (B) of Figure 4-58. This gives a ratio of 1-1/2, or 3/2. If a vertical line were drawn, there would be a single contact, giving a figure of 1/2. The ratio of these two numbers is 3:1, which is consistent with the ratio of frequencies. Figures 4-59 and 4-60 show how the phase relationship of the signals affects the resultant Lissajous pattern. In Figure 4-59, the high-frequency signal starts from its maximum amplitude just as a cycle of the low-frequency signal is ready to begin. This results in a symmetrical pattern comprising three loops. (This figure is also shown in part (C) of Figure 4-58 for ratio computation purposes.) As shown in Figure 4-58(A), the same pattern is formed when the phase difference is 270 degrees instead of 90 degrees. Figure 4-60 shows how an S-shaped pattern is formed when the two signals are in phase. If the high-frequency signal began to swing negative as the low-frequency signal began, the pattern shown in part (C) of Figure 4-58 would result.

4-15.4 OTHER LISSAJOUS PATTERNS

There are two restrictions on the frequencies of the signals applied to the deflecting plates. One was mentioned previously: namely, the frequency

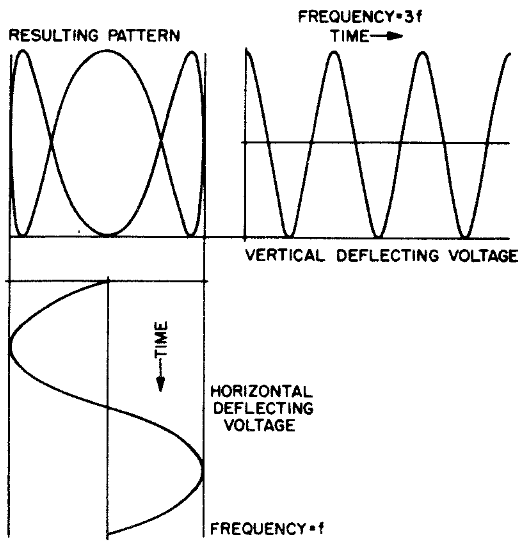


Figure 4-59. OPEN 3:1 Lissajous Pattern

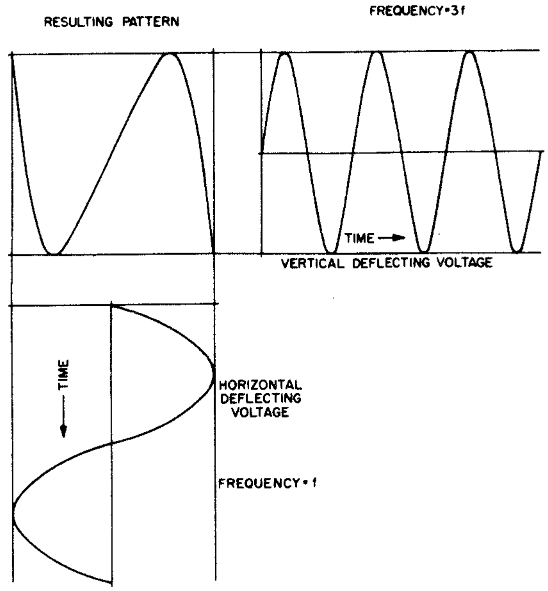


Figure 4-60. CLOSED 3:1 Lissajous Pattern

must lie within the useful pass band of the oscilloscope. Second, the relationship between the applied frequencies must not result in a pattern too involved for an accurate evaluation of the frequency ratio. As a rule, ratios as high as 10:1 and as low as 10:9 can be readily determined. Patterns in which the ratios are 6:1, 8:1, and 10:1, including the corresponding double images, are shown in Figures 4-61, 4-62, and 4-63. The discussion thus far has been limited to integral ratios, such as 1:1, 2:1, 8:1, and 10:1. In addition to these patterns, there are many patterns for which, even though the numerator and denominator of the ratio are whole number (or integers) the ratio is not an integer. For example, there are the 3:2 patterns of Figure 4-64, the 5:4 patterns of Figure 4-65, the 5:3 patterns of Figure 4-66, the 7:2 patterns of Figure 4-67, etc. In every case however, the methods for determining the ratio of the applied frequencies are the same as those previously described.

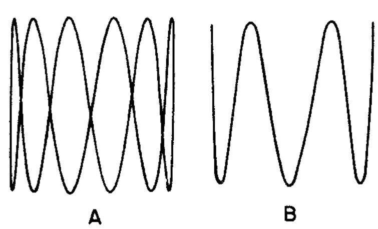


Figure 4-61. 6:1 Lissajous Patterns

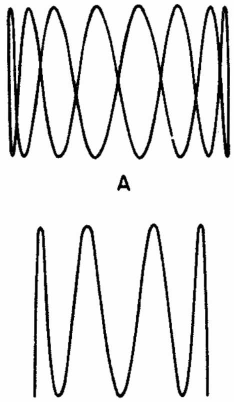


Figure 4-62. 8:1 Lissajous Patterns

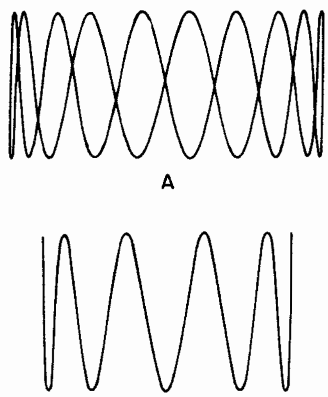


Figure 4-63. 10:1 Lissajous Patterns

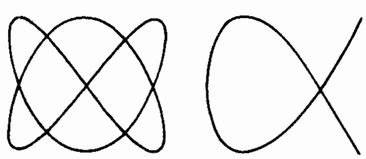


Figure 4-64. 3:2 Lissajous Patterns

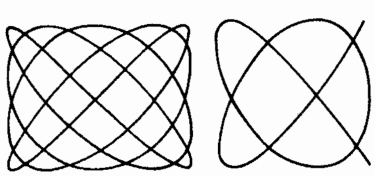


Figure 4-65. 5:4 Lissajous Patterns

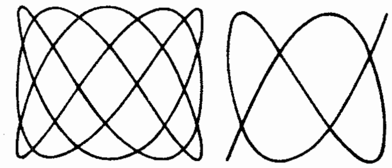


Figure 4-66. 5:3 Lissajous Patterns

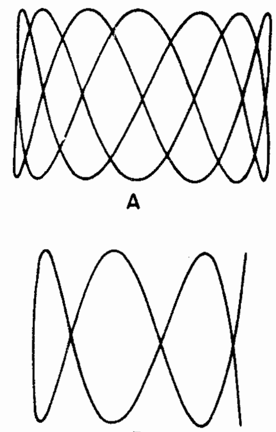


Figure 4-67. 7:2 Lissajous Patterns

4-16 TRANSIENT RESPONSE MEASUREMENT

4-16.1 BACKGROUND

The ability of a linear device to cope with intermittent pulses is determined by the degree of transient response of the device. The quiescent state must be considered and calculated as being zero while the technician determines the transient response of the device. (The calculation of transient response by this method will apply to linear devices only.) When a linear constant parameter device is excited by an input pulse, such as that shown in Figure 4-68, the quiescent component of the device remains constant with respect to time. (Static-state current). Therefore, when making calculations, the total response shape can be considered as dependent only as the input pulse. The response of a linear constant-parameter device can be displayed and measured directly on an oscilloscope. The technician must be aware that if this method is applied to a device that does not have constant parameters, it will be extremely difficult to separate the quiescent response from the transient response. This is because the quiescent component cannot be considered as being "zero" in a non-constant parameters device. The

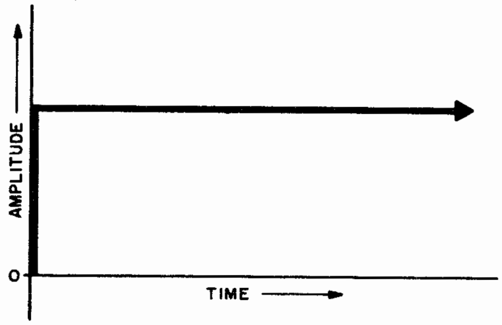


Figure 4-68. Input Step Function

response to a single pulse input, as shown in part A of Figure 4-69, is actually the time duration of the device's response to a single pulse input, as shown in part B of Figure 4-69.

4-16.2 MEASUREMENT TECHNIQUE

The input signal to the amplifier or other electronic device should be a rectangular pulse with the stage output being applied to an oscilloscope to provide a visual display. The half-period of the input rectangular pulse must require a longer time than the transient response duration. The oscilloscope should be adjusted to show the width of a single complete pulse or less. The result of applying a proper pulse versus a too-narrow pulse is shown in Figure 4-70. The transient response of a stage, with an input square wave, is measured by viewing two separate response characteristics. The first of these is related to the leading edge of the output response curve, and is composed of the rise time, time delay, and overshoot. The second of the response curve characteristics is related to the flat-top portion of the curve, called "sag". Sag is possible in a circuit only if the circuit is not capable of passing dc currents. In order to examine the leading edge of the wave, as shown in part B of Figure 4-71, a fast sweep

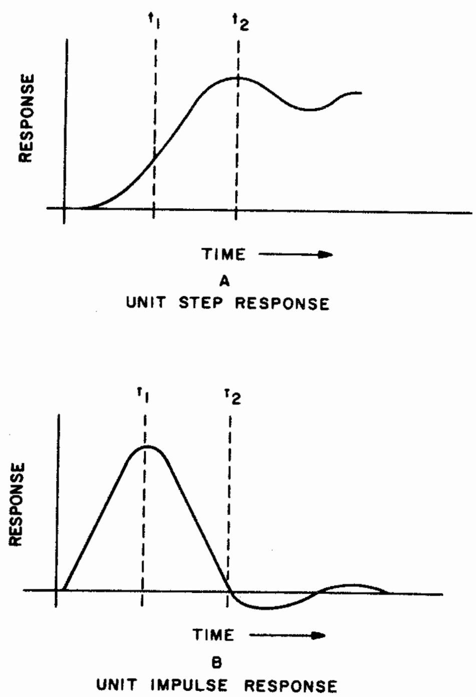


Figure 4-69. Linear Constant Parameter Amplifier Responses

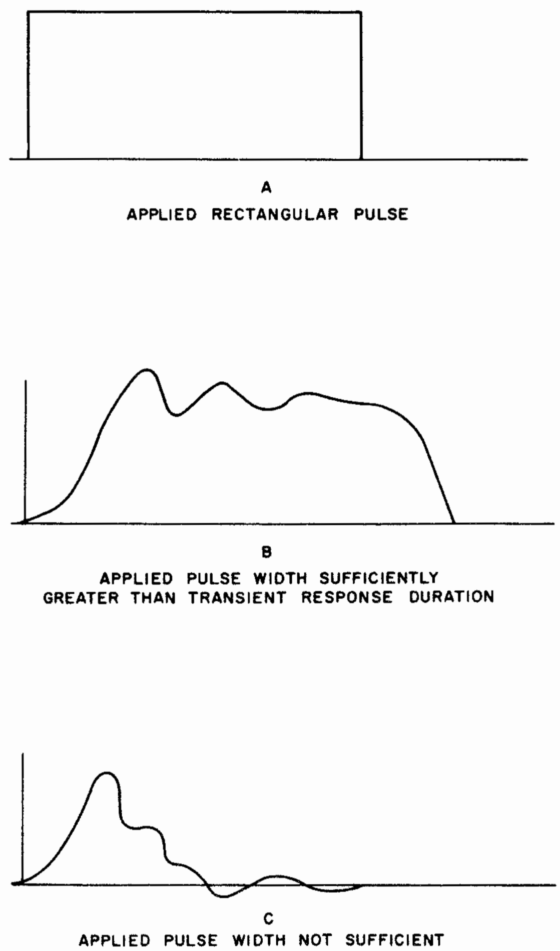


Figure 4-70. Comparison of Applied Pulse Width and Transient Response Times

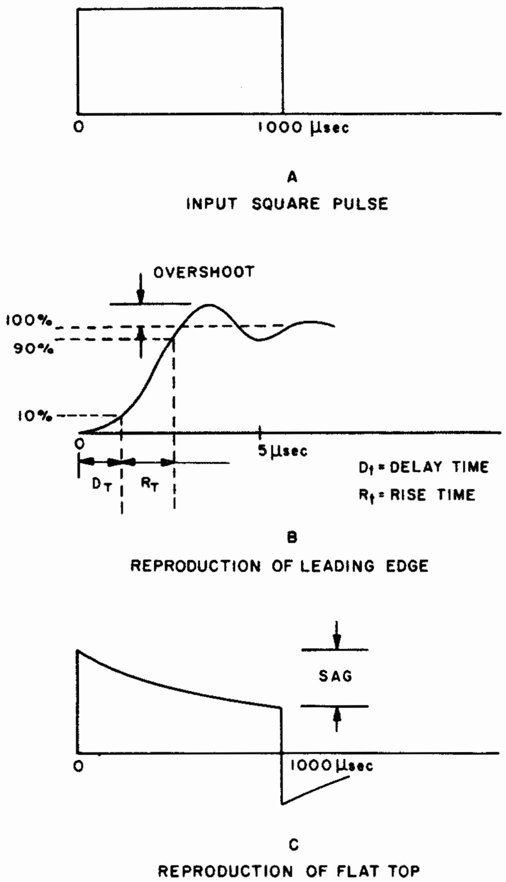


Figure 4-71. Transient Response Characteristics

rate on the oscilloscope should be used, whereas a slow sweep rate should be used to illustrate the flat-top, as shown in part C of Figure 4-71. Wide-input pulses normally have long rise-times. Therefore, narrow-input pulses with short rise-times are used to obtain response pulse leading-edge measurements. Wide pulses should be applied to the input for response pulse sag measurements. Figure 4-71 illustrates a pulse wide of 5 microseconds as being adequate for leading-edge measurements, but 1000-microsecond input pulses are required for the flat-top sag portions of the applied pulse. When a radio frequency signal, rather than a rectangular pulse, is used as the input to a tuned stage or band-pass device undergoing transient response measurements, the parameters of the response are related only to the leading edge because there will be no flat top characteristics. The response of a band-pass device or tuned stage is illustrated in Figure 4-72.

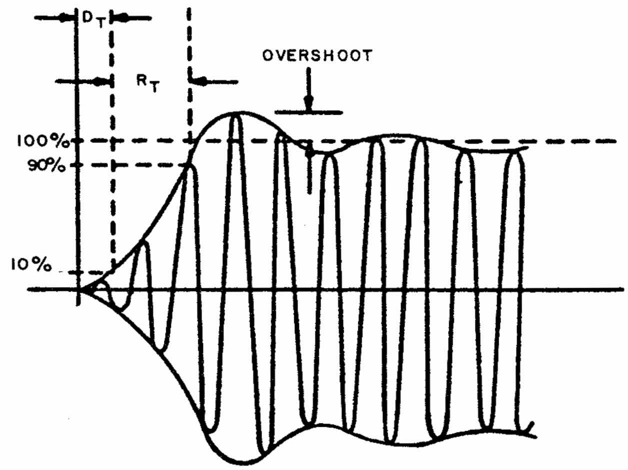


Figure 4-72. Typical Transient Response of a Tuned Stage

4-16.3 TRANSIENTS

The total response in a linear circuit includes all of the individual transients due to the store of energy in each inductor, capacitor, and external energy source connected to the circuit, plus the steady state (forced response) of each external applied energy source. The response can be computed by starting at any arbitrary time (T = O), where all of the initial energy conditions of the proposed circuit are known. As was previously stated, at high frequencies the transit time contributes a conductance element to the grid input admittance. This is because as an electron passes the grid it will introduce a grid current, even if it does not strike the grid. In this case, the grid conductance increases in direct proportion to the applied frequency. The transit time effect on an average tube can be visualized if it is realized that the input resistance to a typical electron tube may be several megohms at a frequency of 5 megahertz, while at 100 megahertz the resistance drops to about 1500 ohms. In triode tubes, the transconductance of the tube decreases and lags behind an increasing amount as the transit time increases. The amplification factor also decreases and the phase angle increases. Much of the tube transit-time difficulty has been solved, however, by placing the electrodes closer together.

4-16.4 REACTIVE ELEMENTS

The discharge of a capacitive element through a resistor requires the time T_2 minus the original starting time (T_1) for the voltage or current to decay down to 37 percent of the original value. The same analogy applies if the capacitive or inductive component charges up to 63 percent. This type of analysis may be used for any periodic applied voltage. The steady-state current and voltage for an applied voltage are determined, the periodic voltage is resolved into its individual harmonic components, and then the transient is determined. The transient waveform does not bear a relationship to the applied voltage, as the transient waveform depends only on the circuit constants and the initial current and voltage conditions.

4-16.5 RESISTIVE ELEMENTS

Time is considered to be zero when no reactive elements are present. Pure resistance elements do not charge or discharge with time.

4-16.5.1 High Frequency Elements

At ultra-high frequencies, the transit time of an electron traveling between the cathode and the plate of a diode constitutes an appreciable part of the input cycle. This causes the dynamic plate resistance of the diode to decrease. Therefore, the cathode-to-plate resistance must be represented by a series resistor capacitor. Figure 4-73 illustrates the curves of both the resistance and the capacitance, each with respect to transit time, as the frequency increases. The resistance R/RP eventually oscillates about a zero reference level, and is sometimes negative.

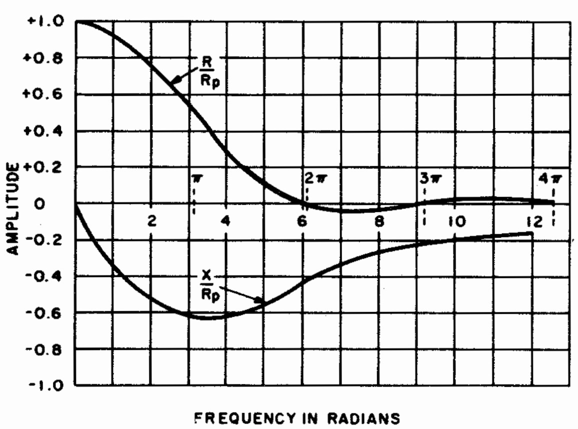


Figure 4-73. Series Resistive and Reactive Diode Components Represented as a Function of Frequency and Transit Times

4-16.6 MEASURING EQUIPMENT

The device being tested contains a definite transit or rise-time to be measured. However, the equipment used to test the device may affect the rise-time. In fact, if the rise-times of the pulse generator and oscilloscope are each less than 10 percent of the rise-time of the device being measured, an accuracy within less than 1 percent can be obtained. To partially compensate for the sag on the rectangular pulse, introduced by the testing equipment, the individual sags may be subtracted from the final measurement to obtain the correct value, that is, assuming that the sag introduced by the equipment is small. Figures 4-74 and 4-75 illustrate the required test setup for the measurement of transient response within linear equipment. The delay time measured with the aid of test setups shown in Figures 4-74 and 4-75 will result in a larger delay time than is actually contained in the equipment itself because of the additional time introduced by the test equipment. However, the test equipment time can be directly subtracted from the final measured result, and the measuring equipment will also tend to reduce the leading edge overshoot of the waveform in the device being tested. The requirements of test equipment used to measure transient response are extremely rigid, as illustrated by several special features of the oscilloscope. A triggered oscilloscope, referred to as a "synchroscope", requires a variety of sweep speeds. The sweep circuit may be triggered by the trigger pulse which starts the pulse generator; the response being measured may be used to trigger the sweep so that later transient responses may be measured; or the applied signal may be used to trigger the sweep circuits of the synchroscope and then be passed through a delay line to the circuit under test. The pulse or square wave generator must contain a wide range of available pulse widths and frequencies; the leading and trailing edges of the output pulse must be short as compared with the pulse width; and the sag should be flat and contain no ringing or oscillation. Finally, the carrier frequency pulses must be relatively free from frequency modulation during the active pulse.

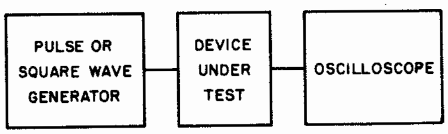


Figure 4-74. Typical Test Setup for Measurement of Transient Response in Low-Pass Equipment

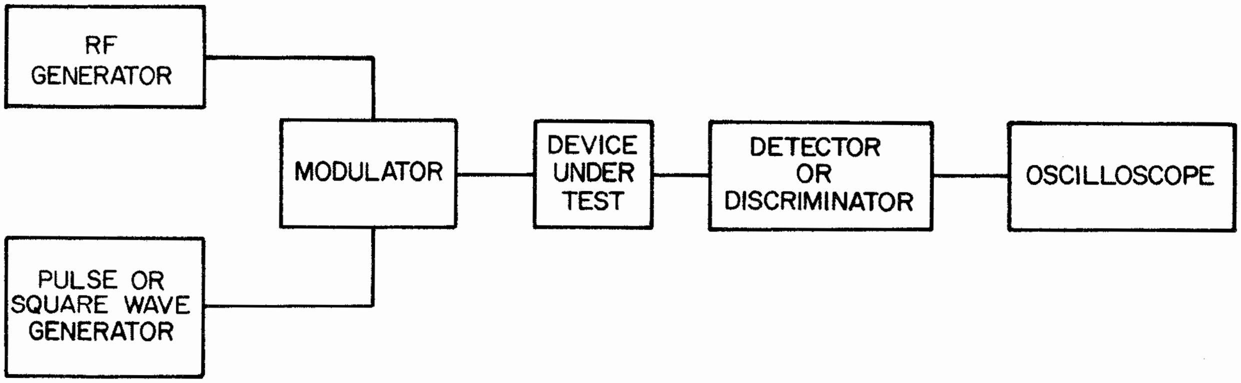


Figure 4-75. Typical Test Setup for Measurement of Transient Response in Bandpass of FM Equipment

4-16.7 TEST EQUIPMENT CONNECTION

The connection of test equipment to the device being tested for transient response is extremely critical because the internal impedance of the test equipment can load down the equipment under test and thus cause response distortion. A cathode-follower stage or other isolating device should be employed between the test equipment and the device under test to minimize the loading effect. All connecting leads should be maintained as short as possible, and the connecting lines must be matched to prevent feedback reflection along the line at high frequencies, which would cause spurious ripples or ringing.

4-16.8 TRANSISTOR CONSIDERATIONS

The transient response of a transistor used as a switch is important because of the time required to turn the transistor switch from the "off" to the "on" position, and vice versa. Normally, either a step or a pulse of current applied to the input is required to turn the transistor from "off" to "on". Referring to Figure 4-76, which is the high frequency equivalent circuit, some assumptions and simple calculations can be made which will provide an approximate rise-time within a transistor circuit operating in an active region. First, assume that load resistor R_L is small enough to represent a short circuit; therefore R_3 (the collector resistance) and C_c (the barrier capacitance) are effectively in parallel with R_2 (the barrier resistance), as shown in Figure 4-76. Now, since R_2 is very much smaller than R_3 in this parallel circuit, the value of R_3 can be disregarded. Furthermore, as C_c (the barrier capacitance) is much larger in reactance than R_2 at the assigned frequency, $2\pi F$, C_c also can be neglected. Thus, for a grounded-base configuration, the equivalent circuit of Figure 4-76 can be reduced to the more practical equivalent circuit shown in Figure 4-77. The value of C_d (the diffusion capacitance) is equal to the reciprocal of the assigned frequency, $2\pi F$, and the emitter resistance, R_1;

$$\left(\frac{1}{2\pi F_1 R_1}\right)$$

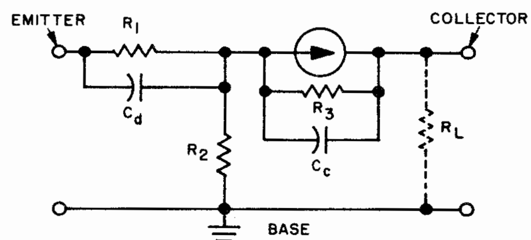


Figure 4-76. Transistor Equivalent Switching Circuit

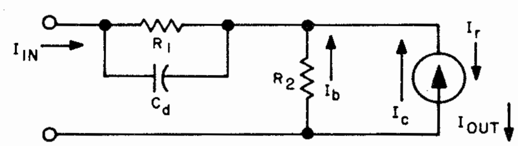


Figure 4-77. Simplified Transistor Equivalent Switching Circuit with Small Output Load

Next, include a subsidiary circuit (R and C), as shown in Figure 4-78. The time of RC equals the reciprocal of $2 \pi F_1$. When the time T equals zero, output I_R equals zero, and when T equals infinity, output I_R equals the input (I_{in}). The time constant is the reciprocal of the

assigned frequency $\left(\frac{1}{2 \pi F_1 R_1}\right)$ and the rise

time is $\left(\frac{1}{2 \pi F_1}\right)$

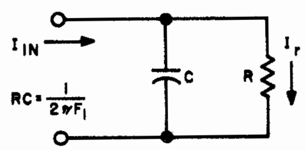


Figure 4-78. RC Circuit, Simulating Frequency Dependence, Used to Calculate I_R

This is the time required for the output current to rise from 0.1 to 0.9 percent of its final value. The effect on the transient response of C_c (the collector capacitor) was not taken into account in the previous calculations because load resistor R_2 was considered small enough to be an effective short circuit. Figure 4-79 illustrates a simplified equivalent circuit which

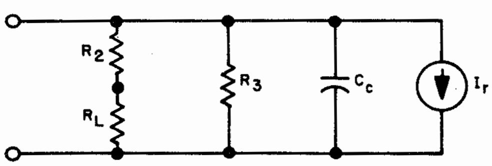


Figure 4-79. Simplified Equivalent Transistor Switch Circuit with Large Output Load

recognizes that R_L does not affect the response when the input is from an infinite series-impedance current and $R_L C_c$ is very much larger than the time-constant

$$\frac{1}{2 \pi F_1}$$

In fact, if $\frac{R2}{R3}$ is very much smaller than 1 and $2 \pi F_1 R_3 C_c$ is much larger than 1, then the time-constant, calculated for the response with a very small or shorted value of R_L, would be increased by the amount, $2 \pi F_1 R_L C_c + 1$. Therefore, the rise times and turn-on times would be increased by this same factor.

4-16.9 TRANSISTOR DELAY TIME

In the previous calculations of transistor response, the finite time required for the impulse signal to diffuse accross the base region was neglected. Under this condition, if a current pulse is applied to the transistor emitter, a response will appear at the collector only after some delay in time. The value of I_R may be obtained for use in the equivalent circuits illustrated in Figure 4-77 and 4-79 by representing this delay time with an equivalent circuit, as indicated in Figure 4-80. The line attenuation will increase in direct proportion to the frequency. The resistive-capacitive delay line indicates a time interval before a response is indicated at the output.

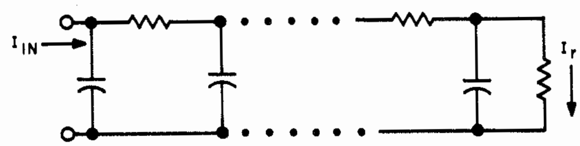


Figure 4-80. Transistor Switch Equivalent Delay Time Circuit

4-16.10 TRANSISTOR STORAGE TIME

Up to this point the discussion has concerned the turn-on transistor process. If the transistor is in the active region, the turn-off process consists simply of applying a pulse of reverse polarity, and the required time-constant is calculated in the same manner, using the equation for the turn-on process. Unfortunately, if the transistor is in the "on" condition and is operating in the saturation region, an abnormally large time-delay will occur before the transistor responds to the turn-off signal. This peculiar delay, termed "storage-time delay", is illustrated by Figure 4-81, which shows the minority carrier density in the base region for three situations. The first situation illustrated is the cutoff condition, with both the emitter and collector junctions back-biased. The minority carrier density is therefore zero at both junctions, and very small throughout the base region. The second situation is illustrated by the active curve in Figure 4-81, where the minority carrier density is high at the emitter junction, and zero at the collector junction. The change in density between the two junctions is the result of the diffusion process, which accounts for current flow across the base of the transistor. However, if an input signal drives the emitter junction to a

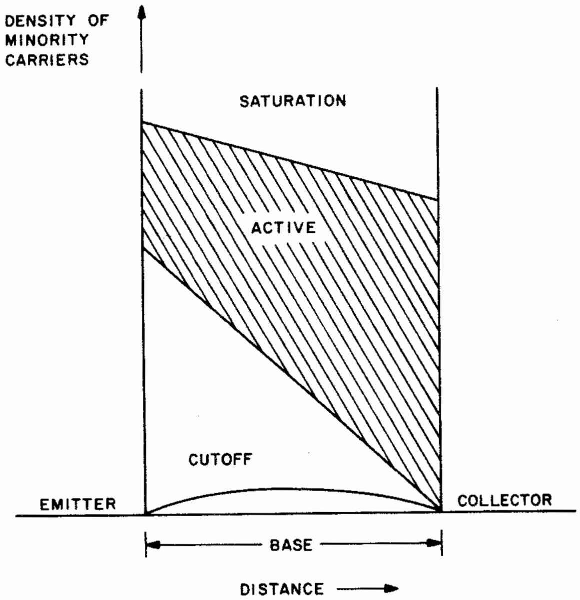


Figure 4-81. Comparative Minority-Carrier Densities in Transistor Base for Cutoff, Active, and Saturation Regions of Operation

back-bias condition, the diffusion process will not continue until the minority carriers in the base region have been removed. The third situation is illustrated by the saturation curve in Figure 4-81. Unlike the cutoff and active conditions, the previously-discussed equivalent circuits used to determine time-delays, response, etc., do not apply in this case. This is because both the emitter and the collector are emitting carriers into the base region. In addition, as both junctions are forward-biased, the junction voltages will be small and the collection process at the junctions will be slow. This, in turn, causes the density of minority carriers in the base region to build up to a relatively large value. This high-density level in the base region must be permitted to decrease before the turn-off process begins to take effect. This long-storage time-delay may represent two or three times the normal rise or fall time in the active region. It is therefore evident that when a transistor switch is used in an application requiring high switching speeds, it must be restrained from entering the saturation region.

4-16.11 TRANSISTOR RESPONSE

Figure 4-82 illustrates the approximate waveforms required to represent the response from a transistor driven from cutoff to saturation and back again. A grounded-emitter switch is used because this type of configuration is the most useful and its response can represent other configurations. In Figure 4-82, the delay-time is represented by the symbol T_D, the rise-time by T_0, the storage-time by T_S, and the decay or fall-time by T_1. The input current reverses at the end of the pulse, rather than falling only to zero. As a result, the output current response falls toward a negative value rather than toward zero. The fall time of the pulse is thereby reduced. The voltage input waveform of Figure 4-82 was terminated in a minus E_2 voltage because the storage of minority carriers in the base region does not permit the transistor input impedance to immediately attain a large

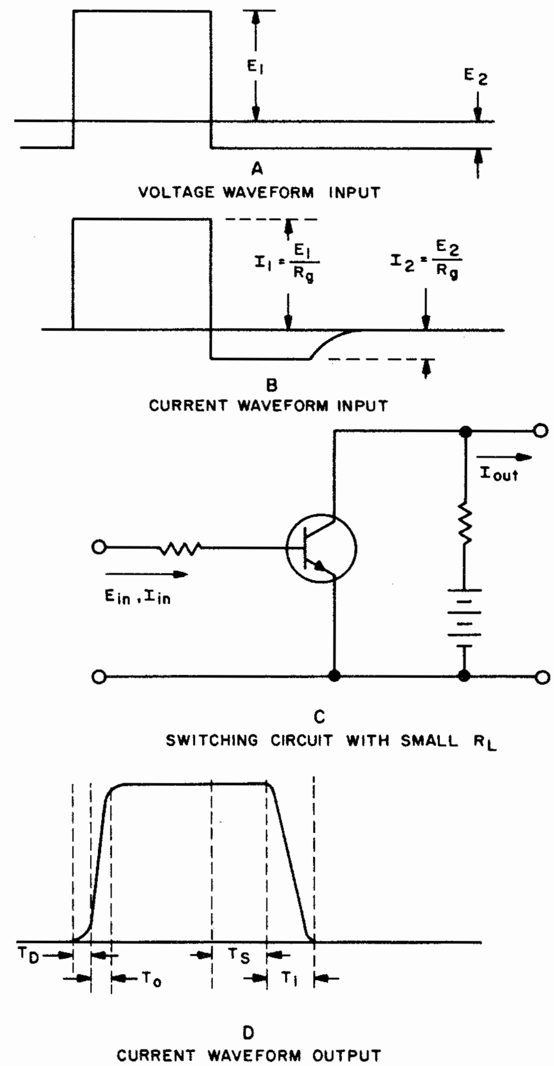


Figure 4-82. Grounded-Emitter Switch Circuit with Input Voltage and Current Waves and Output Current Response Waveform for Small R_L

value. In fact, the input impedance remains small until the minority carriers at the transistor junctions are swept away. At this point the input impedance increases and causes the input current to decrease in direct proportion to the speed with which the minority carriers throughout the base region drift to the junctions.

4-17 SPECTRUM WAVEFORM ANALYSIS AND MEASUREMENTS

The analysis of a complex waveform, prepared in terms of a graphical plot of the amplitude versus frequency, is known as spectrum analysis. Spectrum analysis recognizes the fact that any waveform is composed of the summation of a group of sinusoidal waves, each of an exact frequency and all existing together simultaneously.

4-17.1 ELECTROMAGNETIC FREQUENCY SPECTRUM

A chart showing the electromagnetic frequency spectrum is given in Figure 4-83. This chart indicates the frequency and wavelength of the various frequency bands.

4-17.2 WAVELENGTH - FREQUENCY CONVERSION

The chart in Figure 4-84 and Table 4-1, when used in conjunction with each other, provide a ready means of converting frequency to wavelength and vice-versa. For example, a frequency of 25 MHz has a wavelength of 12.5 meters, and a frequency of 250 MHz has a wavelength of 1.25 meters. Both frequencies enter the chart at the 25 level of f, but the λ multiple in Table 4-1 for 25 MHz is 1.0 while it is .10 for 250 MHz.

4-17.3 ACOUSTIC SPECTRUM

The acoustic spectrum chart illustrated in Figure 4-85 is provided to show the limits of the sound spectrum as set by human ear sensitivity. This chart is based on a musical pitch of 440, a physical pitch of 426.667, and an international pitch of 435.

4-17.4 SPECTRUM ANALYSIS

In the realm of varying frequency, three axes of degree exist: amplitude; time; and frequency. The time domain (amplitude vs. time) plot is used to recover phase relationships and basic timing of the signal, and is normally observed with the aid of an oscilloscope. The frequency domain (amplitude vs. frequency) plot is used to observe frequency response, employing the spectrum analyzer for this purpose. Figure 4-86 illustrates the difference between frequency and time domain plots. View A illustrates a three dimensional coordinate of a fundamental frequency and its second harmonic with respect to time frequency and amplitude. View B depicts the time domain display as it would be seen on an oscilloscope. The solid line, $f_1 + 2f_1$, is the actual display. The dashed lines, f, and $2f_1$, are drawn to illustrate the fundamental and second harmonic frequency relationship used to formulate the composite signal $f_1 + 2f_1$. View C is the frequency domain display of view B as it would be seen on a spectrum analyzer. Note in view C that the components of the composite signal are clearly seen.

4-17.5 FREQUENCY DOMAIN DISPLAY CAPABILITIES

The frequency domain contains information not found in the time domain. The spectrum analyzer can display signals composed of more than one frequency (complex signals); and it can discriminate between its components, while measuring the power level at each one. It is more sensitive to low level distortion than an oscilloscope and its sensitivity and wide dynamic range are also useful for measuring low level modulation, as illustrated in (a) and (b) of Figure 4-87. The spectrum analyzer is useful in the measurement of long and short-term stability such as noise sidebands on an oscillator, residual FM of a signal generator or frequency drift of a device during warm up. This is illustrated in Figure 4-88. The swept frequency response of a filter or amplifier and the swept distortion measurement of a tuned oscillator are also measurable with the aid of a spectrum analyzer. However, in the course of these measurements, a variable persistence display on an X-Y recorder should be employed for simplification of readability. Examples of tuned oscillator harmonics and filter response are illustrated in Figure 4-89. Frequency conversion devices such as mixers, harmonic generators, etc. are easily characterized by such parameters as conversion loss, isolation, and distortion. These parameters can be displayed, as shown in Figure 4-90, with the aid of a spectrum analyzer. Present-day spectrum analyzers can measure segments of the frequency spectra from Ø Hz to as high as 40 GHz.

4-17.6 SPECTRUM ANALYZER USAGES

Although the previously mentioned measurement capabilities are attainable with a

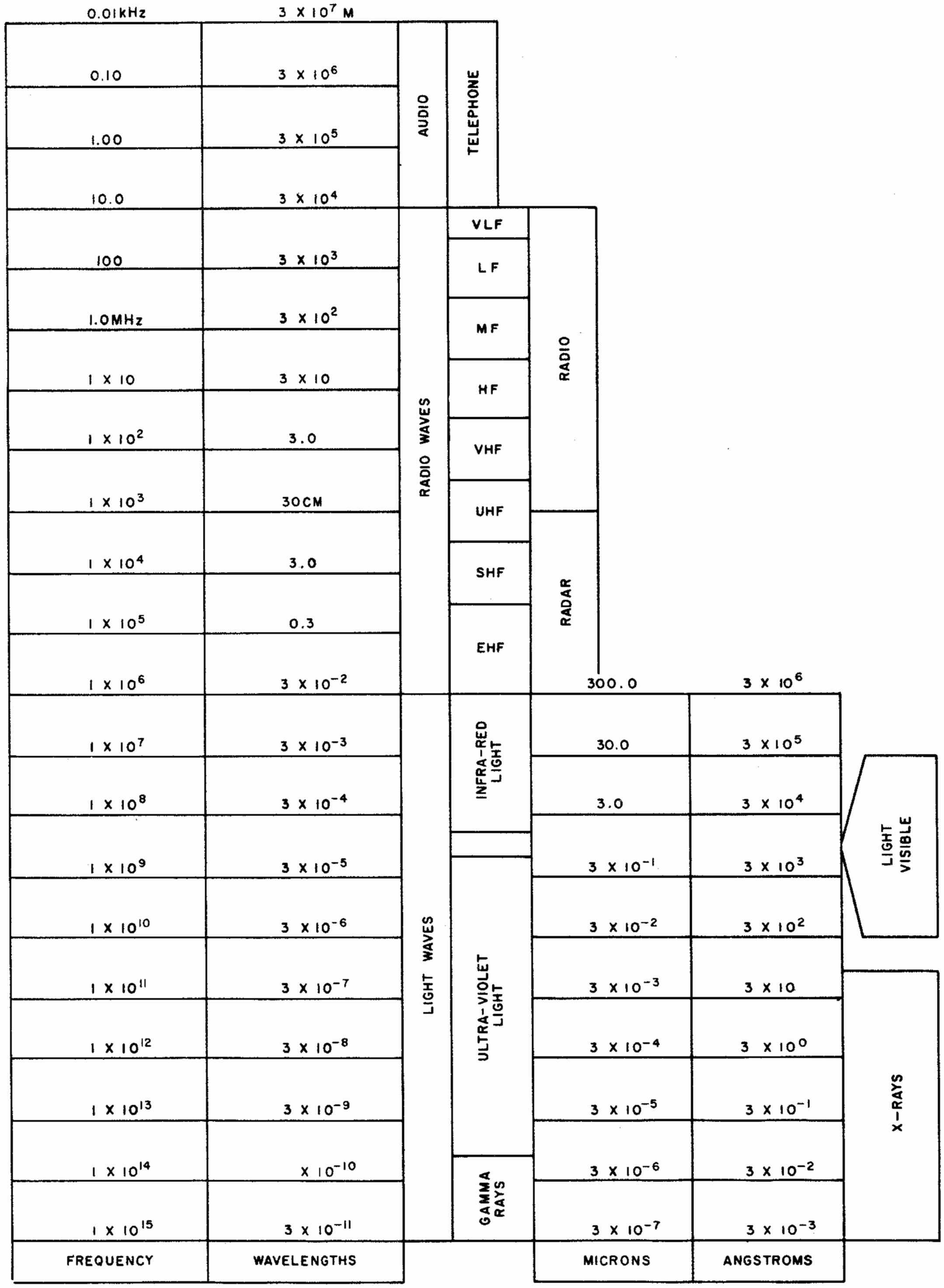


Figure 4-83. Electromagnetic Frequency Spectrum

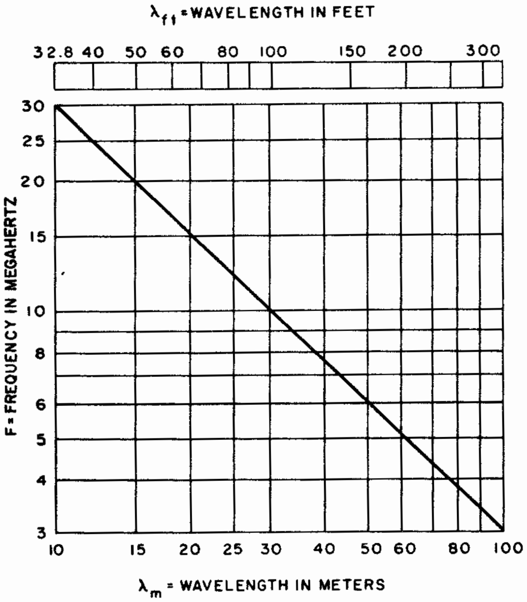


Figure 4-84. Wavelength-Frequency Conversion Chart

Table 4-1. Auxiliary Wavelength-Frequency Conversion Table

FOR FREQUENCIES IN MEGAHERTZ FROM	MULTIPLY f BY	MULTIPLY λ BY
0.03 - 0.3	0.01	100.000
0.30 - 3.0	0.10	10.000
3.00 - 30.0	1.00	1.000
30.00 - 300.0	10.00	0.100
300.0 - 3000.0	100.00	0.010
3000.00 - 30000.0	1000.00	0.001

The formula conversions are:

λ m = 300/f MHz
λ cm = 30/f GHz
λ FT = 984/f MHz
λ IN = 11.8/f GHz

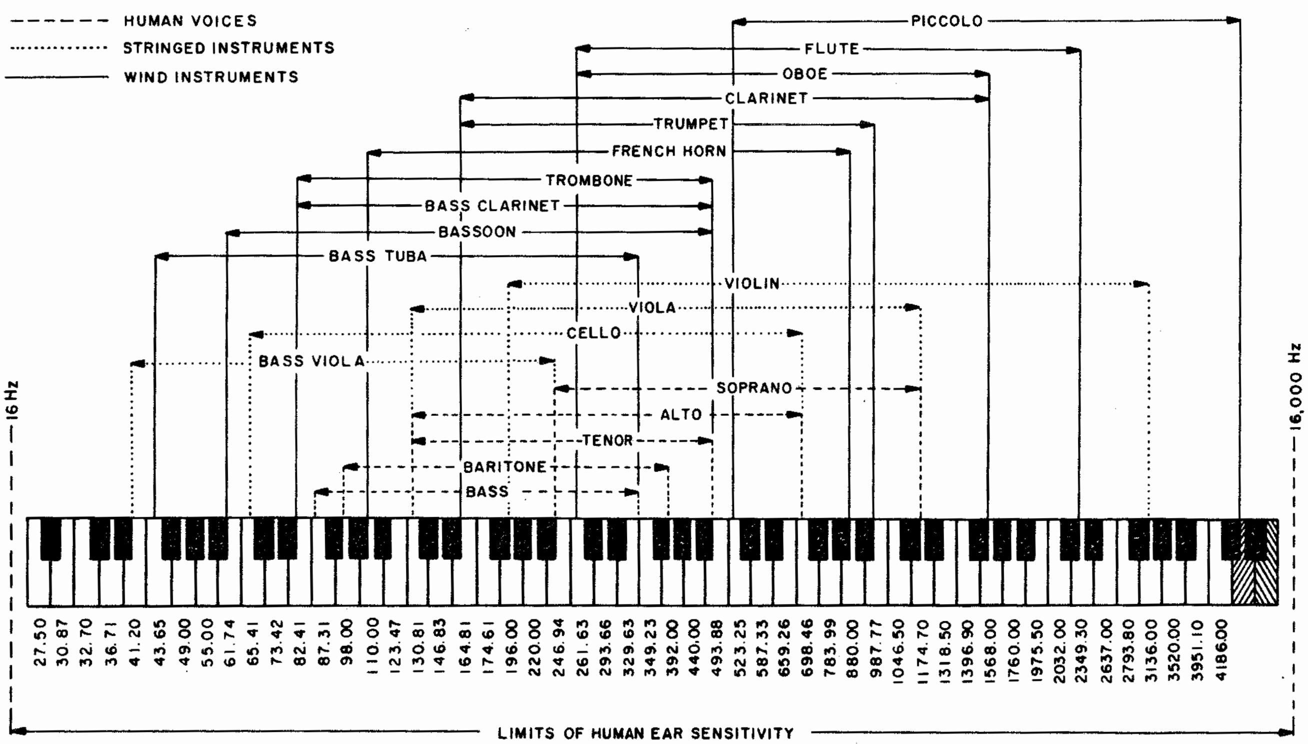


Figure 4-85. Acoustic Spectrum

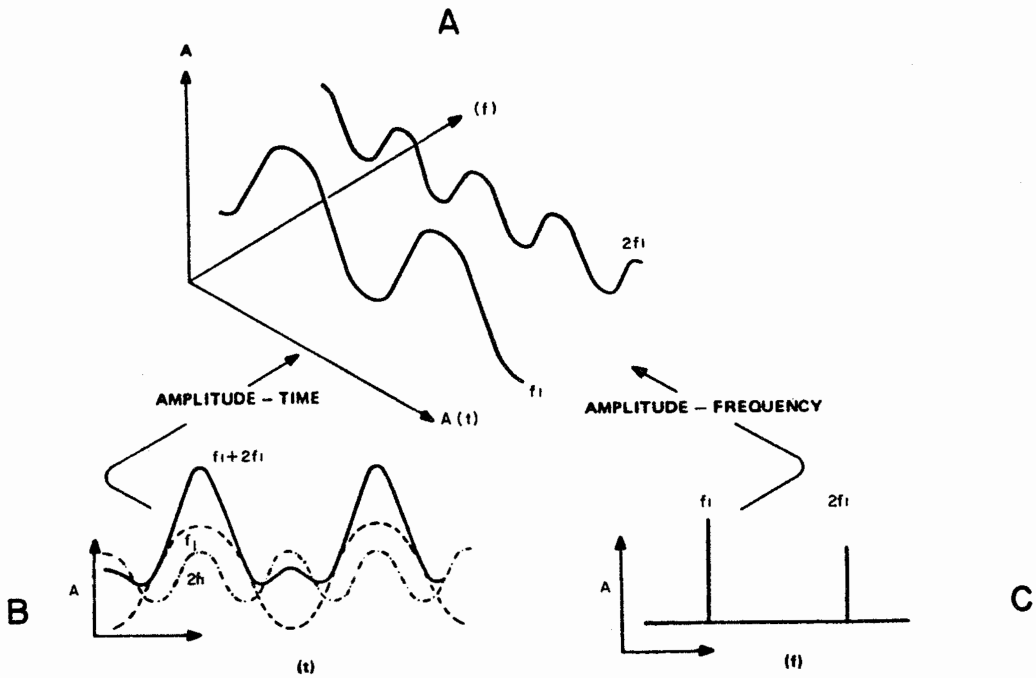


Figure 4-86. Time vs Frequencies

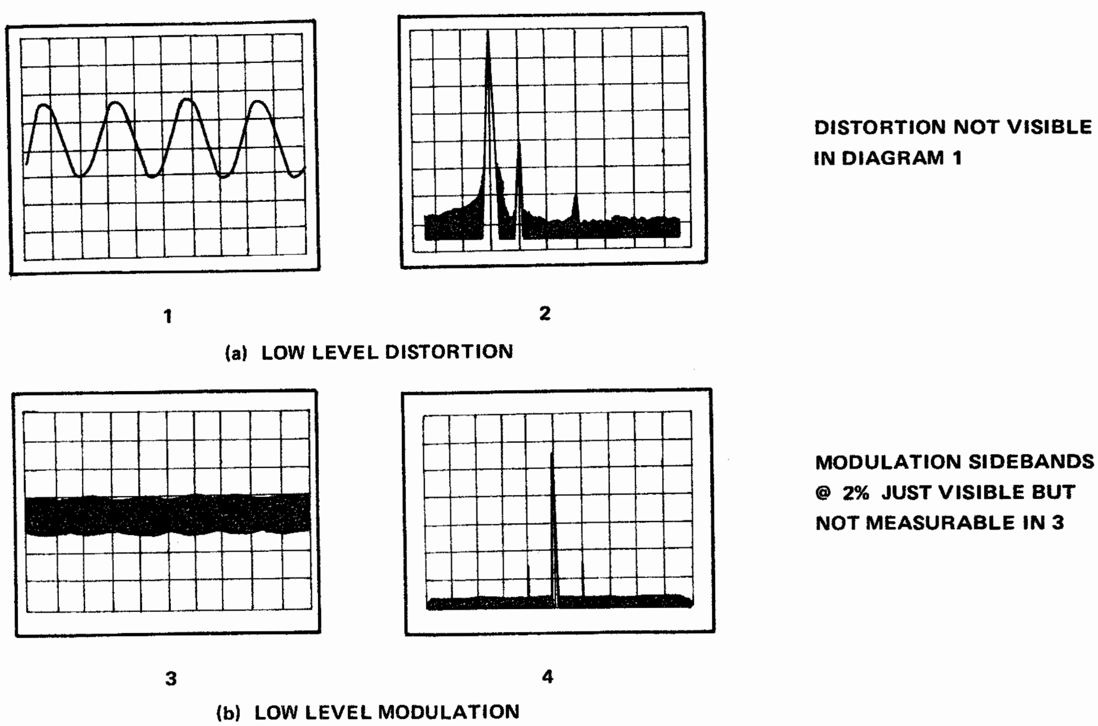


Figure 4-87. Examples of Time Domain (left) and Frequency Domain (right) Low Level Signals

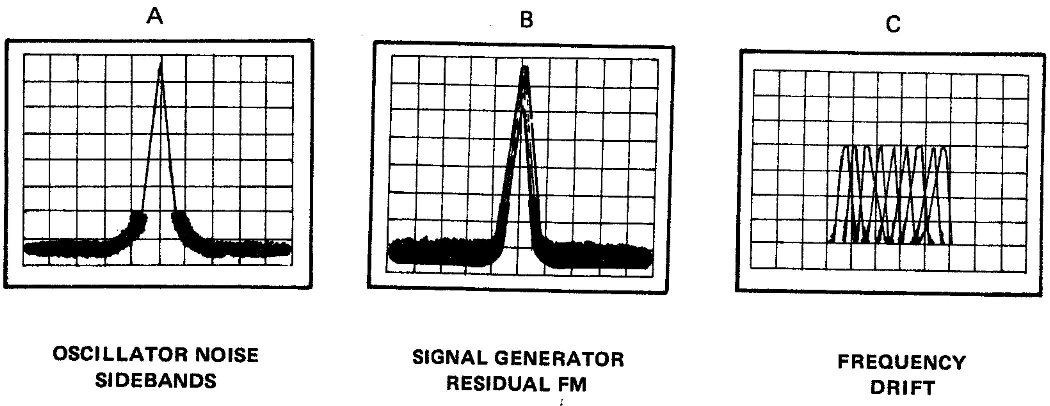


Figure 4-88. Spectrum Analyzer Stability Measurements

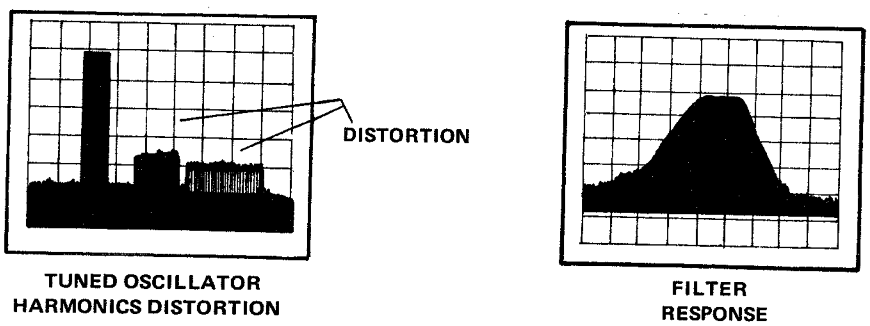


Figure 4-89. Swept Distortion and Response Characteristics

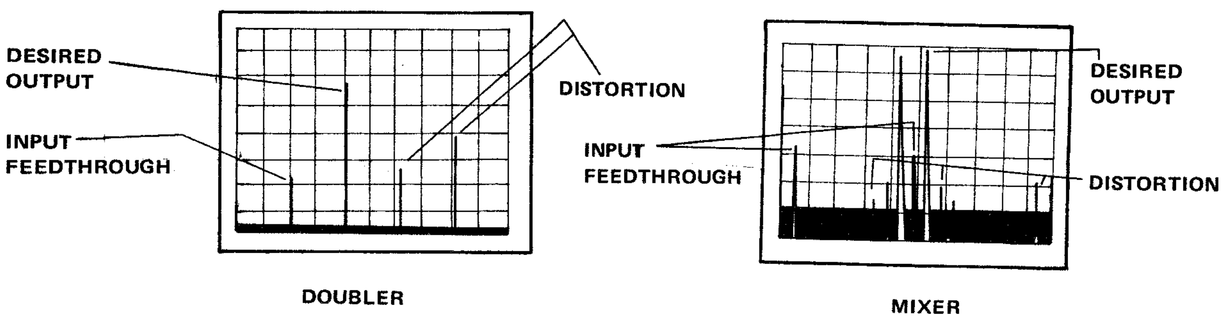


Figure 4-90. Frequency Conversion Characteristics

spectrum analyzer, the technician will find that, in general, the spectrum analyzer is used to measure spectral purity of multiplex signals, percent of modulation of AM signals, modulation characteristics of FM and PM signals, and in the interpretation of the displayed spectra of pulsed RF emitted from a radar transmitter. Occasionally, field strength measurements are required to determine RFI (i.e., radio frequency interference).

4-17.7 COMPLEX WAVEFORMS

Complex waveforms are divided into two groups-periodic waves, and non-periodic waves. Periodic waves contain the fundamental frequency and its related harmonics. Nonperiodic waves contain a continuous band of frequencies, resulting from the repetition period of the fundamental frequency approaching infinity and thereby creating a continuous frequency spectrum.

4-17.8 MODULATION MEASUREMENTS

In all types of modulation, the carrier is varied in proportion to the instantaneous variations of the modulating waveform. The two basic properties of the carrier available for modulation are the amplitude characteristics and the angular (frequency or phase) characteristic.

4-17.8.1 Amplitude Modulation

The modulation energy in an amplitude-modulated wave is contained entirely within the sidebands. Amplitude modulation of a sinusoidal carrier by another sinusoid would be displayed as illustrated in Figure 4-91. For 100% modulation,

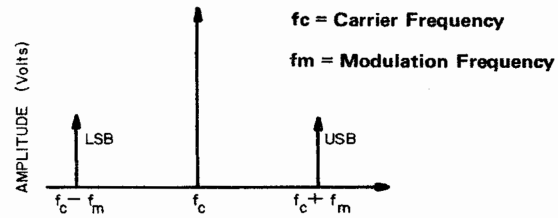
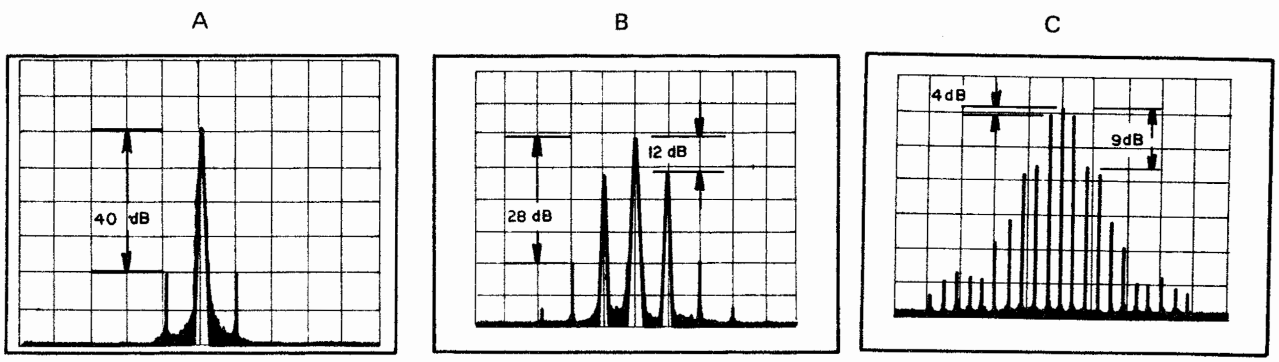


Figure 4-91. Spectrum Analyzer Display of an AM Signal

total sideband power would be one-half of the carrier power; therefore, each sideband will be 6 dB less than the carrier, or one-fourth of the power of the carrier. Since the carrier component is not changed with AM transmission the total power in the 100% modulated wave is 50% higher than in the unmodulated carrier. The primary advantage of the log display of the spectrum analyzer over the linear display provided by the oscilloscopes for percent of modulation measurements, is that the high dynamic range of the spectrum analyzer (up to 70 dB) allows accurate measurements of values as low as 0.06%. It also allows the measurement of low-level distortion of AM signals. Both capabilities are illustrated in Figure 4-92. The chart in Figure 4-93



2% MODULATION
SIDEBANDS 40 dB DOWN

50% MODULATION
SIDEBANDS 12 dB DOWN
DISTORTION 28 dB DOWN

> 100% MODULATION
SIDEBAND 4 dB DOWN
DISTORTION 9 dB DOWN

Figure 4-92. AM Display of Spectrum Analyzer

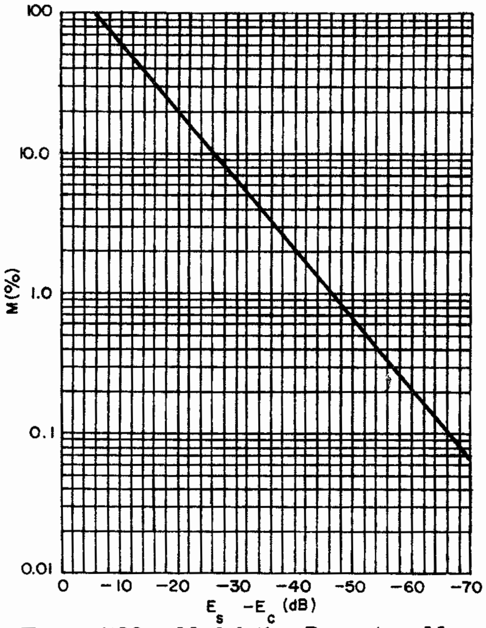
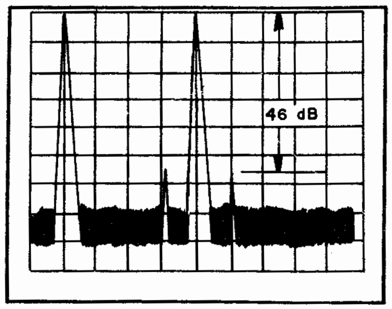


Figure 4-93. Modulation Percentage M vs Sideband Level (Log Display)

provides an easy conversion of dB down from carrier into % of modulation. Note: anything greater than -6 dB is over 100% modulation, therefore producing distortion as illustrated in C of Figure 4-92. In modern long-range HF communication, the most important form of amplitude modulation is SSB (single sideband). Either the upper or lower sideband is transmitted and the carrier is suppressed. Single sideband requires only one sixth of the output power required by AM to transmit an equal amount of intelligence power and less than half the bandwidth. Figure 4-94 illustrates the effects of balancing-out the carrier of an AM signal. The most common distortion experience in SSB is intermodulation distortion, which is caused by non-linear mixing of intelligence signals. The two-tone test is used to determine if any intermodulation distortion exists. Figure 4-95 illustrates the spectrum analyzer display of the two-tone test.

4-17.8.2 Frequency Modulation

Amplitude modulation contains the intelligence in the sideband current pairs spaced symmetrically about the carrier by an amount equal to each modulation frequency. Theoretically, frequency modulation can contain an infinite number of sideband current pairs per modulating frequency with the intelligence spread throughout them as well as along the carrier. The amplitude of a particular pair



CARRIER BALANCED 46 dB

Figure 4-94. Double Sideband Carrier Suppressed

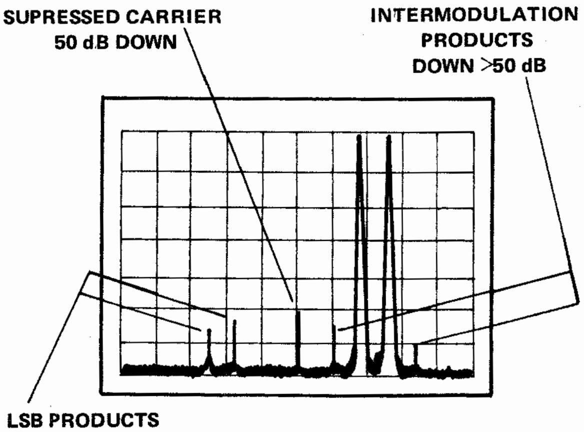


Figure 4-95. Two-Tone Test

of side currents may be larger than the center-frequency component. This fact also holds true for phase modulation, in that, with the same modulation index (β), the same spectrum distribution is obtained. The number of important side currents is larger for low frequencies in the signal band than it is for high frequencies. Table 4-2 provides selected amplitude factors used to multiply the maximum unmodulated carrier current level (I_m) to find the amplitude value of an individual side current pair in the frequency spectrum. For example, use a maximum center frequency swing (Δ F) = $\pm$60 kHz and a 30 kHz signal frequency (f) to find the value of β .

$$\beta = \frac{\Delta F}{f} = \frac{60}{30} = 2$$

Table 4-2. Abbreviated Bessel Factor Table

β	$J_0(\beta)$	$J_1(\beta)$	$J_2(\beta)$	$J_3(\beta)$	$J_4(\beta)$	$J_{14}(\beta)$
0.0	1.0000	0.0000	0.0000	0.0000	0.0000	
0.5	0.9385	0.2423	0.0306			
1.0	0.7652	0.4401	0.1449	0.0196	0.0025	
2.0	0.2239	0.5767	0.3528	0.1289	0.0340	
5.0	- 0.1776	- 0.3276	0.0466	0.3648	0.3912	
10.0	- 1.2459	0.0435	0.2546	0.0584	- 0.2196	0.0120

Table 4-2 illustrates that $J_0(\beta) = J_0(2) = 0.2239$, and this is the value multipled times the maximum unmodulated current (I_m) to obtain the magnitude of the center frequency component (F). The value of $J_1(\beta) = J_1(2) = 0.5767$ multiplied times I_m gives the amplitude of both the first upper side current at the frequency of (F + 30 kHz) and the first lower side current at the frequency of (F — 30 kHz). The spectrum distribution for a modulation index of $\beta = 2$ is illustrated in Figure 4-96. This graph was prepared from data obtained from Table 4-2. As all amplitudes are obtained by multiplying the Bessel factors obtained from Table 4-2 times I_m, the magnitude of the Bessel factors directly determines the intensity of the sideband current pairs in the useful frequency spectrum. The sideband current pairs which are too far down in amplitude from the center frequency (F) are not significant; i.e., they are less than 1% of the unmodulated carrier (I_m). Thus, the bandwidth is determined by the number of significant sideband current pairs. The bandwidth may also be calculated from Table 4-2 for a specific frequency swing. For example, let $\beta = 10$ and read the Bessel function values across the chart, from left to right. As you can see, the 14th

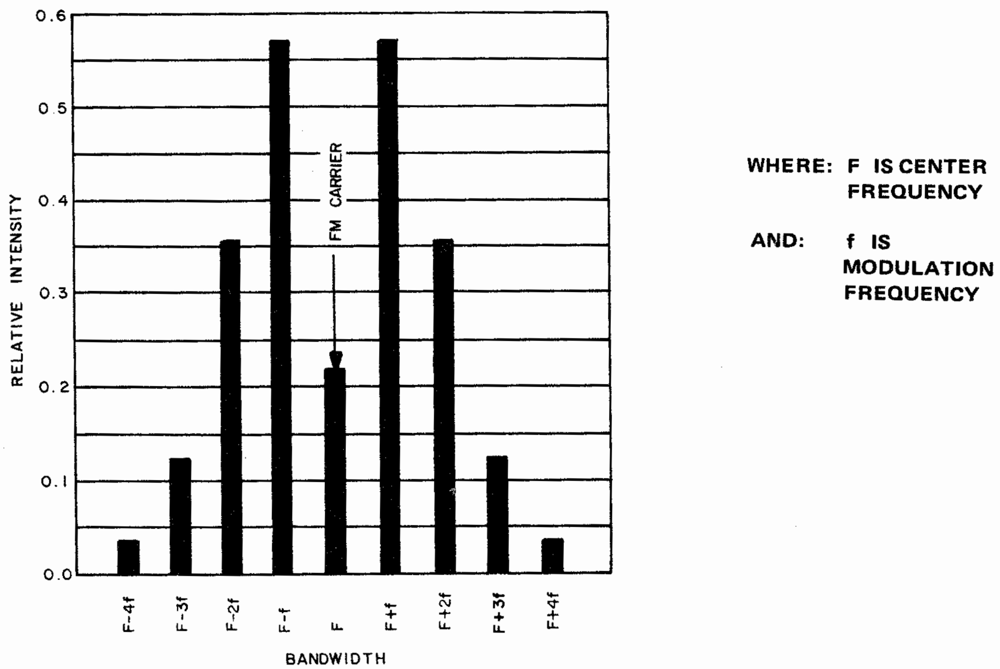


Figure 4-96. Spectrum Distribution for a Modulation Index of $\beta = 2$

sideband pair is 0.012 I_m, which is significant. Therefore, the maximum bandwidth for β = 10 is 2 (both side currents) times 14 sidebands times the signal frequency deviation. For a 30-kHz signal deviation, the bandwidth would be 2 x 14 (30) = 840-kHz, and for 2-kHz it would be 56 kHz. This example shows why a higher-modulation signal requires more frequency space (that is, greater bandwidth) than does a lower modulation signal. Figure 4-97 shows the effect of changing the amplitude or frequency of the modulating signal while holding the other constant. The Bessel factors given in Table 4-2 and Figure 4-98 also show that the greater the modulation index (β), the greater the number of significant sidebands. As the modulation energy spreads out from the carrier frequency (F), it can be determined at what point to establish the bandwidth. For example, in Figure 4-98 which was prepared for an index of β = 24, the largest side currents occurred at the edges of the pass band; thereafter continue to read off the Bessel factors until the side current pairs are less than 1% of the unmodulated carrier current level, I_m. Sideband currents beyond this point do not have significant amplitudes for practical consideration. Frequency modulation, for modulation index values smaller than 0.2, is similar to amplitude modulation in that both types of modulation contain only one significant side-current pair. Therefore, for a value of

$$\frac{\Delta F}{f} = \beta = 0.2$$

or less, FM behaves exactly like AM with respect to spectrum distribution. However, unlike AM, no primary oscillator can be used in the FM transmitter because the carrier frequency must be varied during the modulation cycle to produce FM. The desired intelligence in FM or phase modulation (PM) will create more energy distribution, and thus a larger response in a receiver demodulator, than will noise energy. This is the outstanding desirable feature of FM over AM spectrum distribution.

4-17.8.3 Phase Modulation

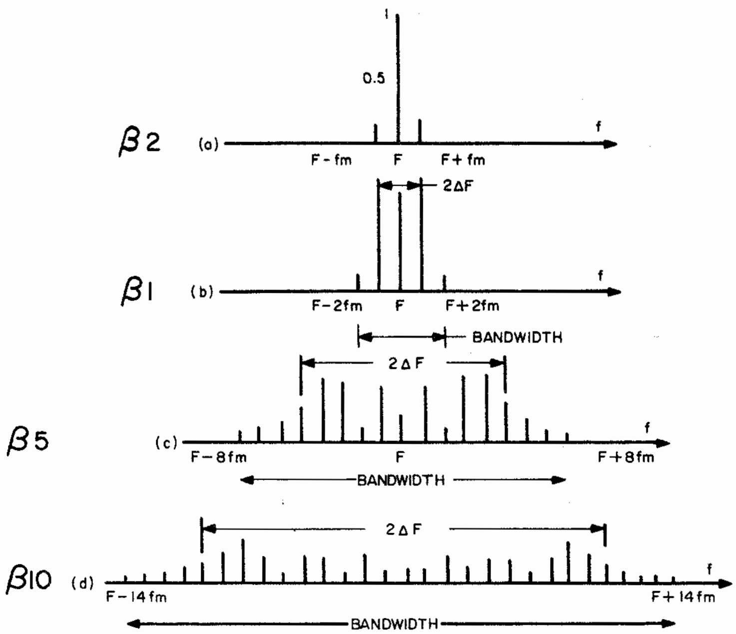
In PM unlike FM or AM, the carrier current level (I_m), as well as the center carrier frequency (F), remains constant. Only the relative phase (θ) changes. The actual value of θ is not important. The deviation from this value is important, and produces the desired PM. For example, if θ equals 30 degrees and the phase deviation is ±40 degrees, this will produce a certain modulation of the carrier. However, if θ had originally been 80 degrees rather than 30 degrees and was subjected to a ±40-degree deviation, the same output PM waveform, containing the same intelligence, would have been produced. The original value of θ indicates only the original amplitude at the start of the phase swing. Actually, the effect of a ± 40-degree deviation would be to give the appearance of wobbling about the carrier frequency (F) as F goes through an angular distance of ± 360 degrees, regardless of its beginning phase, θ. However, the equivalent instantaneous frequency of the modulated carrier would remain the same. Thus PM entails the important fact that it is created not only by the maximum phase deviation ($\Delta\beta$), but also by the applied signal frequency (f). Therefore, the frequency shift is greater for higher frequencies than it is for lower frequencies.

4-17.8.3.1 Sidebands

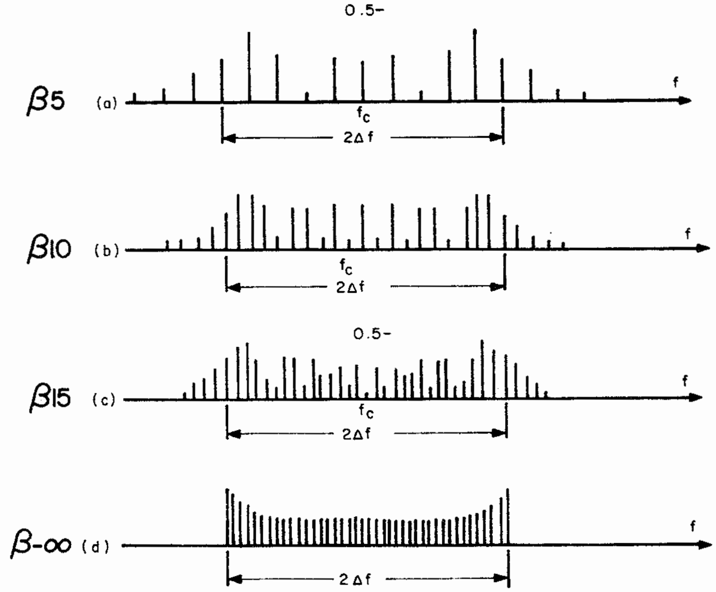
The sidebands for PM are similar to the sidebands for FM and the same general formula for the modulation index holds true. That is:

$$\beta = \left(\frac{\Delta\theta}{f}\right)$$

As with FM, the side currents contain a symmetrical frequency distribution around the carrier. In other words, the first upper sideband and the first lower sideband have the same numerical value, amplitude, and frequency difference from the carrier; the value of the modulation index (β) is proportioned to the phase deviation ($\Delta\theta$); and for a fixed maximum phase swing it does not matter whether you use a 15 Hertz or 15 kHz signal to modulate the phase of the carrier. In either case, you will obtain the same number of important side-current pairs. However, for the 10th upper and lower sideband in the 15 Hertz modulation case, the side-current pair is 10 x 15 Hertz or 150 Hertz above and below the carrier; for the 15 kHz case the 10th side-current pair is 10 x 15 kHz or 150 kHz above and below the carrier, and thus requires a much broader bandwidth. Therefore, PM and FM, unlike AM, have a spectrum distribution of the modulation energy which is proportional to the square of the spectrum amplitudes, and does affect the carrier frequency amplitude. However, it does not matter whether the technician is dealing with a modulation index of 10 for FM or maximum phase deviation of 10 radians (573 degrees) for PM. This is true for PM as long as the maximum phase swing ($\Delta\theta$), which causes the PM, is fixed. The same spectrum distribution is equally true for FM if the maximum deviation (ΔF), which causes the FM, changes directly with the signal



A
AMPLITUDE VARIED FREQUENCY CONSTANT



B
FREQUENCY VARIED AMPLITUDE CONSTANT

Figure 4-97. Modulation Frequency and Amplitude Effects on an FM Carrier

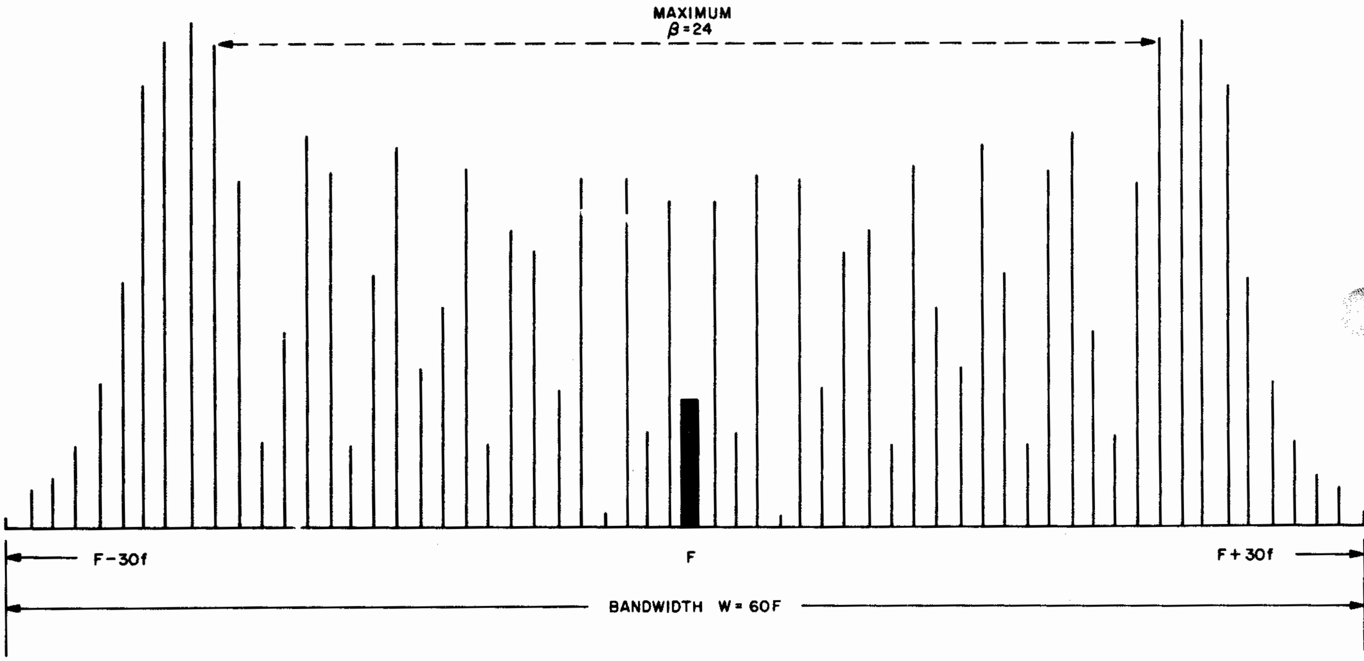


Figure 4-98. Spectrum Distribution for an Index of 24

frequency (f). As the bandwidth of PM remains the same regardless of signal frequency changes, the PM bandwidth can be calculated directly from the modulation index (β). For small values of modulation index (0.3 or less), PM will contain only one pair of significant sidebands, the same as for FM. Actually, this permits the technician to amplitude-modulate a carrier, suppress the carrier, and shift the modulation product by 90 degrees to provide narrow-band FM or PM. The FM or PM effect can be obtained by applying a signal voltage having a magnitude which is inversely proportional to the signal frequency (f). The phase-shifted product is then combined with the unmodulated carrier.

4-17.9 CARRIER FREQUENCY

Both FM and PM contain a true instantaneous and an equivalent instantaneous carrier frequency. Therefore, the number of Hertz must remain constant to prevent the center or mean frequency (F) from drifting to some other nearby frequency not originally assigned to the carrier. If this occurred, the entire spectrum centered around the carrier would drift and infringe on the other nearby FM channels.

4-17.10 PULSED WAVES

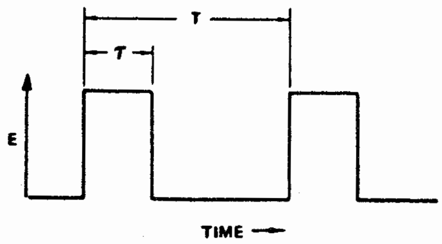
An ideal pulsed radar signal is comprised of a train of RF pulses with a constant repetition rate, constant pulse width and shape, and constant amplitude. To receive the energy reflected from a target, the radar receiver requires close to ideal pulse radar emission characteristics. By observing the spectra of a pulsed radar signal, such characteristics as pulse width, duty cycle, peak, and average power can be measured easily and accurately.

4-17.11 RECTANGULAR PULSE

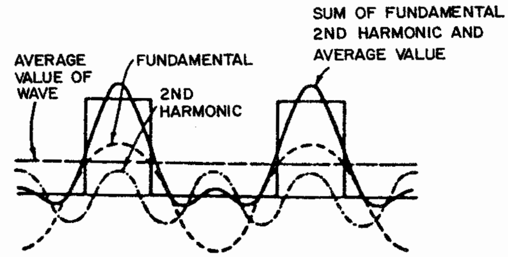
A rectangular wave is used to pulse-modulate the constant frequency RF carrier to produce the pulse radar output. The rectangular wave is comprised of a fundamental frequency and its combined odd and even harmonics. Both in and out of phase harmonics relationships are utilized, depending on the desired pulse width or pulse interval. Figure 4-99 illustrates the development of a rectangular wave and its spectral content.

4-17.12 PULSED WAVE ANALYSIS

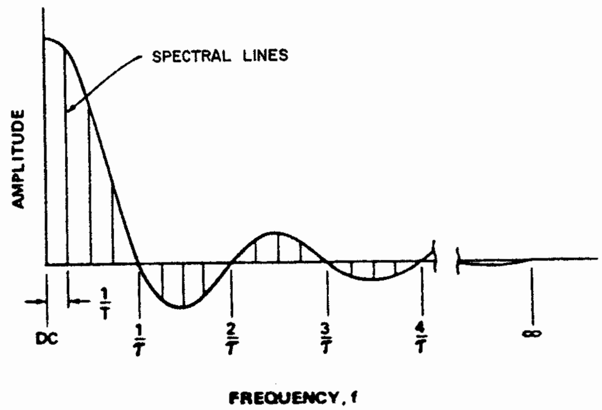
In AM, the sidebands are produced above and below the carrier frequency. The principle



Rectangular Pulse
Time Domain



Rectangular Pulse
Harmonic Development



Rectangular Pulse
Spectra (Frequency Domain)

Figure 4-99. Rectangular Pulse

also applies for a pulse except that the pulse is comprised of many tones. These tones produce multiple sidebands which are commonly referred to as spectral lines or "rails" on the analyzer display. There will be twice as many rails in the pulse radar's modulated output as there are harmonics contained in the modulating pulse (upper and lower sidebands), as illustrated in Figure 4-100. In Figure 4-100, the PRF is equal to the pulse interval of 1/T. The actual spectrum analyzer display would show the lower lobes (shown below the reference line) on top because the spectrum analyzer does not retain any phase information. Changing the pulse interval or pulse width of the modulation signal will change the amount of rails (PRF) or number of lobe minima, as illustrated in Figure 4-101.

4-17.13 ANALYZING THE SPECTRUM PATTERN

The leading and trailing edges of the radiated pulse modulated signal must be extremely steep, with a constant amplitude between them. Incorrect pulse shape will cause frequency spread and pulling, which results in less available energy at the frequency to which the receiver is tuned. The primary reason for analyzing the spectrum is to determine the exact amount of amplitude and frequency modulation present. The amount of amplitude modulation determines the increase in the number of sidebands within the applied pulse spectrum, whereas an increase in frequency modulation increases the amplitude of the side lobe frequencies. In either case, the energy available to the main spectrum lobe is decreased.

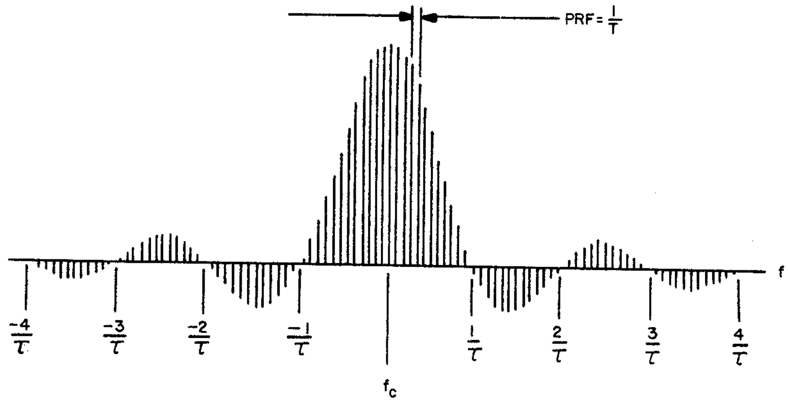


Figure 4-100. Pulsed Radar Output

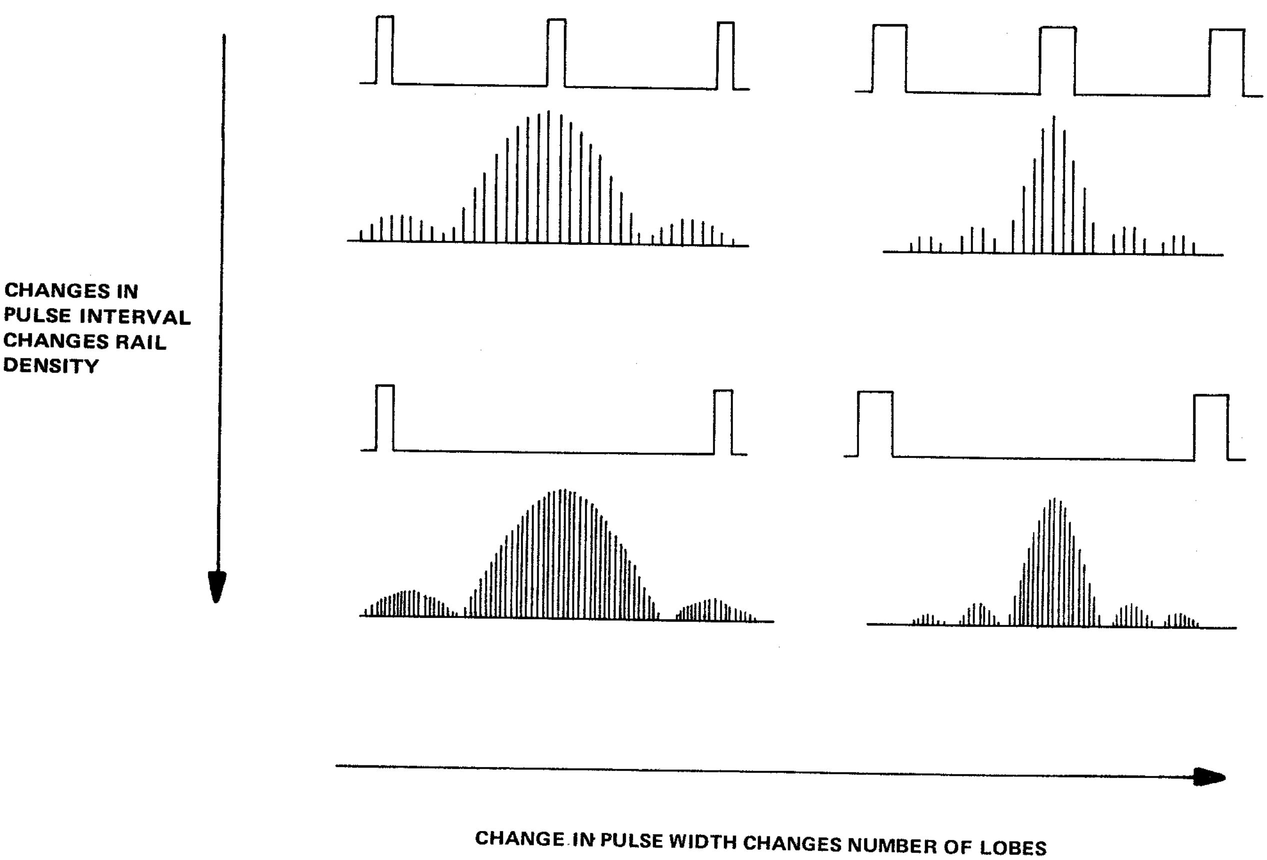


Figure 4-101. Pulsed Radar Chages Affected by Modulating Signal Changes

4-17.13.1 Typical Spectrum Patterns

Figure 4-102 contains several illustrations of commonly obtained patterns. These spectra are the result of pulse-modulated waves which are special types of RF carrier amplitude modulation. As can be seen, amplitude modulation alone does not seriously affect the frequency spectrum on an RF pulse. The type of modulation present can be easily determined because amplitude modulation primarily affects the amplitude of the side lobes and does not affect the shape of the main lobe. Frequency modulation affects the main lobe bandwidth. Spectrum asymmetry, as shown in part F of Figure 4-102 occurs only when both amplitude and frequency modulation occur simultaneously.

4-17.13.2 Spectrum Analyzer Interpretations

A pulsed RF signal has unique properties; therefore, the technician must be careful to correctly interpret the display on a spectrum analyzer. Spectrum analyzer response to a pulsed radar signal can be of two kinds, resulting in displays which seem similar but are of completely different significance. One response is called a "line spectrum" and the other is called a "pulse spectrum". Both are responses to the same pulsed radar signal, and the "line" and "pulse" spectrum terms refer solely to the response of the display on the spectrum analyzer.

4-17.13.2.1 Line Spectrum

A line spectrum occurs when the spectrum analyzer's 3 dB bandwidth (β) is narrow compared to the frequency spacing of the input signal components, as illustrated in Figure 4-103.

4-17.13.2.2 Pulse Spectrum

A pulse spectrum occurs when the spectrum analyzer's bandwidth (β) is equal to or greater than PRF. The spectrum analyzer, in this case, cannot resolve the actual individual frequency components because several rails are within its bandwidth. However, if the spectrum analyzer's bandwidth (β) is narrow as compared to the spectrum envelope, the envelope can then be resolved, as illustrated in Figure 4-104. Figure 4-105 illustrates the effect of varying the scan width bandwidth of the spectrum analyzer in the line analysis interpretation. Figure 4-106 illustrates the effect of the same variations in the pulse analysis interpretation.

4-17.14 SPECTRUM ANALYZER OPERATION

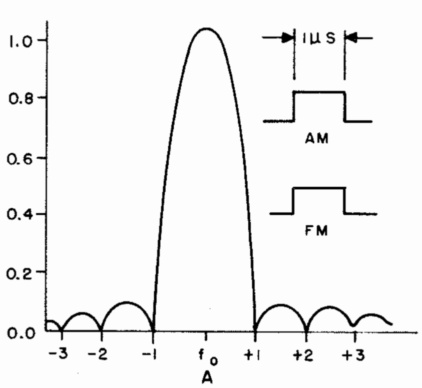
The information desired from the spectra to be analyzed determines the spectrum analyzer requirements. Real time analysis is used if a particular point in the frequency spectrum is to be analyzed, such as line spectra displays. Continuous or swept frequency analysis, which is the most common mode of observation, is used to display a wider portion of the frequency spectrum or (in some cases) the entire range of the spectrum analyzer in use. Changing the spectrum analyzer setting from one mode to another is accomplished by varying the scan time and/or the spectrum analyzer's bandwidth. Most real-time spectrum analyzers, however, are preceded by mechanical filters which limit the input bandwidth of the spectrum analyzer to the desired spectra to be analyzed. Tunable or swept spectrum analyzers function basically the same as heterodyne receivers, the difference being that the local oscillator is not used but is replaced by a voltage control oscillator (VCO). The VCO is swept electronically by a ramp input from a sawtooth generator. The output of the receiver is applied to a CRT which has its horizontal sweep in synchronization with the VCO. The lower frequency appears at the left of the display. As the trace sweeps to the right, the oscillator increases in frequency. Figure 4-107 is a block diagram of a heterodyne spectrum analyzer.

4-17.14.1 Resolution

Before the frequency of a signal can be measured on a spectrum analyzer it must be resolved. Resolving a signal means distinguishing it from other signals near it. Resolution is limited by the narrowest bandwidth of the spectrum analyzer because the analyzer traces out its own I-F bandwidth shape as it sweeps through a signal. Thus, if the narrowest bandwidth is 1 kHz, then the nearest any two signals can be, yet still be resolved, is 1 kHz. Reducing the I-F bandwidth indefinitely would obtain infinite resolution were it not that the usable I-F bandwidth is limited by the stability (residual FM) of the spectrum analyzer. The smaller the shape factor of the I-F bandwidth, the greater the analyzer's capability to resolve closely-spaced signals of unequal amplitude. Signals of equal amplitude can be resolved only when they are separated by the 3 dB bandwidth. Unequal signals can be resolved if they are separated by greater than half the bandwidth at the amplitude difference between them.

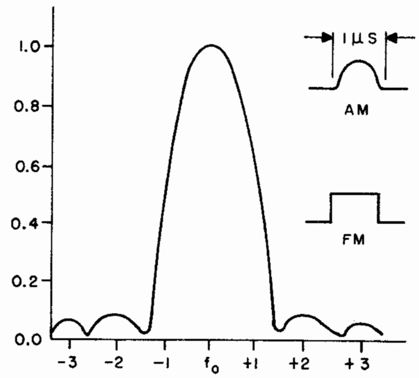
4-17.14.2 Other Spectrum Analyzer Considerations

It is important that the spectrum analyzer be more stable in frequency than the signals being measured. The stability of the analyzer depends on the frequency stability of its VCO. Scan time of



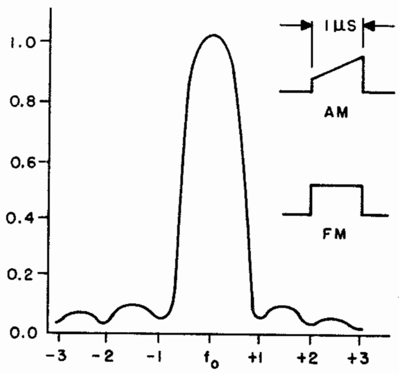
A

IDEAL SPECTRUM ENVELOPE, NOT MODULATED. SYMMETRICAL BECAUSE PULSE DURATION ± CARRIER IS EXACTLY 1 µSEC.



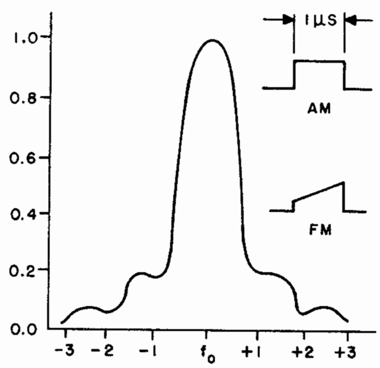
B

WIDE MAIN LOBE CAUSED BY SLOPE-SIDED PULSE OF LESS THAN 1 µSEC.



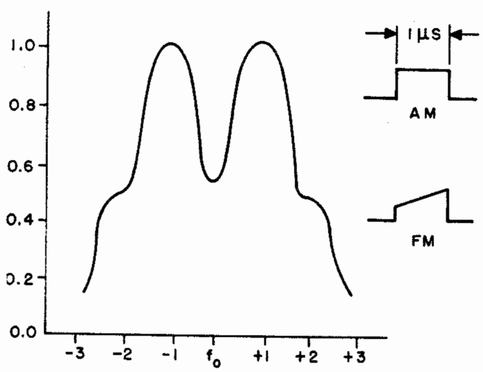
C

MINIMA DO NOT ATTAIN ZERO AMPLITUDE BECAUSE PULSE HAS LINEAR-SLOPED TOP.



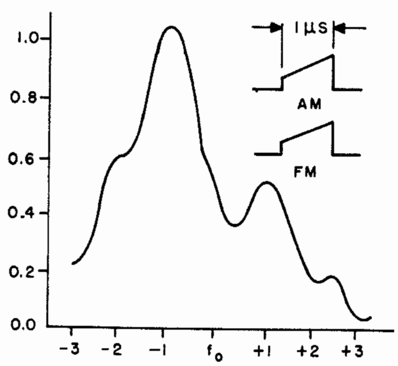
D

DISTORTED AMPLITUDE RISE OF FIRST MINIMA IS DUE TO 2 MHz/SEC LINEAR FREQUENCY MODULATION.



E

DISTORTED AMPLITUDE RISE OF FIRST MINIMA IS DUE TO 6 MHz/SEC LINEAR FREQUENCY MODULATION.



F

DISTORTED TOTAL SPECTRUM IS DUE TO BOTH AMPLITUDE AND FREQUENCY LINEAR PULSE SLOPE. *

* CHANGING THE SLOPE OF ONE TYPE OF MODULATION WITH RESPECT TO THE OTHER WILL PRODUCE A MIRROR IMAGE.

Figure 4-102. Spectrum Patterns

B = SPECTRUM ANALYZER BANDWIDTH

PRF = 1/PULSE INTERVAL

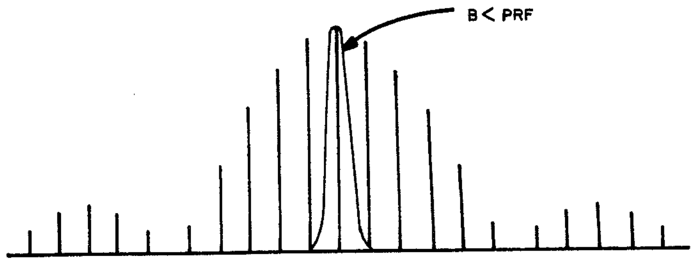


Figure 4-103. Line Spectrum B < PRF

B = SPECTRUM ANALYZER BANDWIDTH

PRF = 1/PULSE INTERVAL

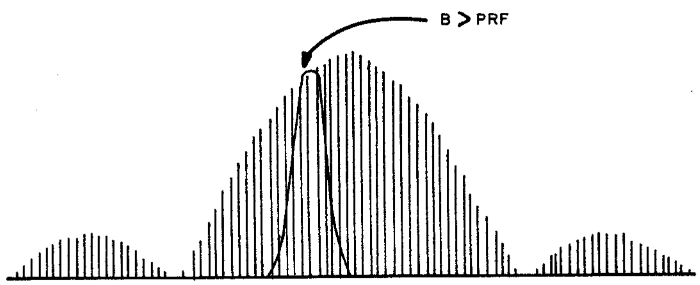
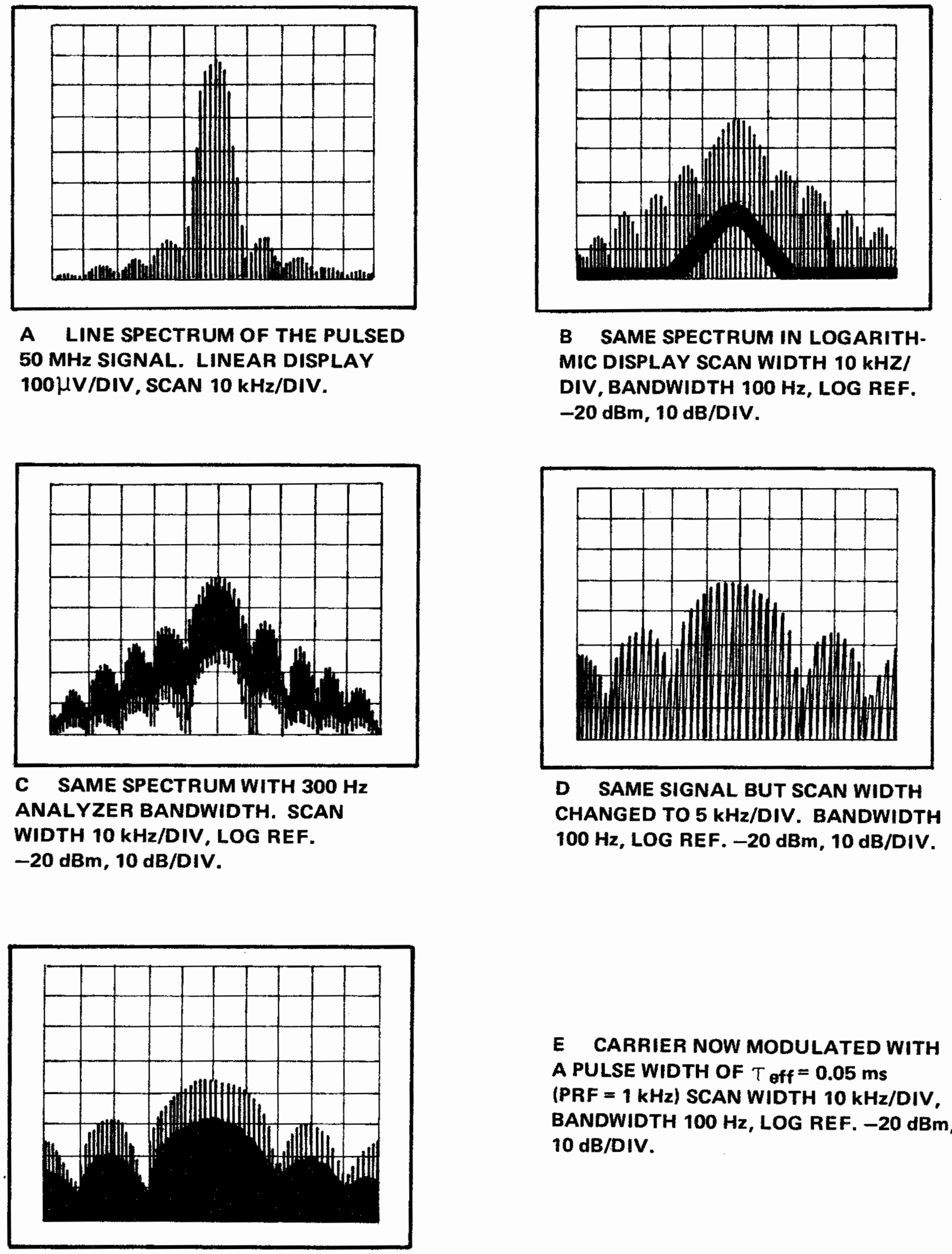


Figure 4-104. Pulse Spectrum B > PRF



A LINE SPECTRUM OF THE PULSED 50 MHz SIGNAL. LINEAR DISPLAY 100µV/DIV, SCAN 10 kHz/DIV.

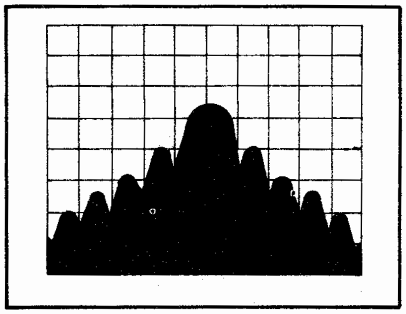
B SAME SPECTRUM IN LOGARITHMIC DISPLAY SCAN WIDTH 10 kHZ/DIV, BANDWIDTH 100 Hz, LOG REF. –20 dBm, 10 dB/DIV.

C SAME SPECTRUM WITH 300 Hz ANALYZER BANDWIDTH. SCAN WIDTH 10 kHz/DIV, LOG REF. –20 dBm, 10 dB/DIV.

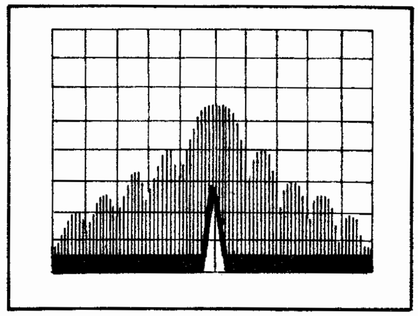
D SAME SIGNAL BUT SCAN WIDTH CHANGED TO 5 kHz/DIV. BANDWIDTH 100 Hz, LOG REF. –20 dBm, 10 dB/DIV.

E CARRIER NOW MODULATED WITH A PULSE WIDTH OF τ_{eff} = 0.05 ms (PRF = 1 kHz) SCAN WIDTH 10 kHz/DIV, BANDWIDTH 100 Hz, LOG REF. –20 dBm, 10 dB/DIV.

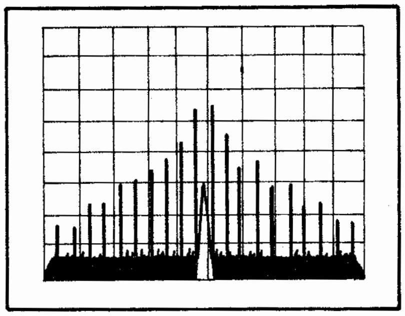
Figure 4-105. Line Spectra of a Pulsed Modulated 50 MHz Carrier



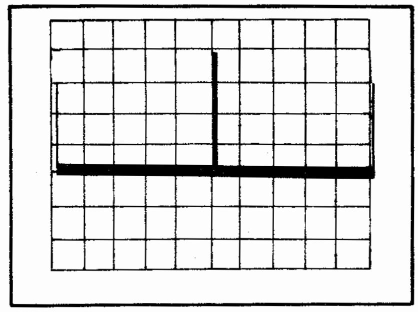
A SIGNAL (PEAK AMPLITUDE –30 dBm) PULSED WITH PRF = 100 Hz, T_{eff} = 1/10 kHz = 100 µs. SCAN WIDTH 10 kHz/DIV, B = 1 kHz SCAN TIME 0.5 s/DIV.



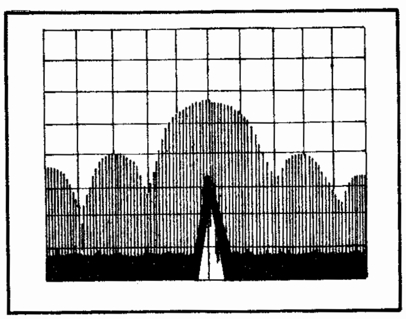
B SAME SIGNAL, BUT SCAN TIME CHANGED TO 0.1 s/DIV.



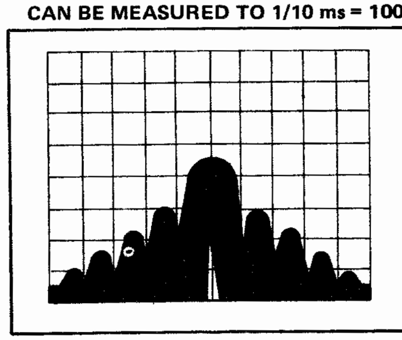
C SAME SIGNAL WITH A SCAN TIME OF 20 ms/DIV.



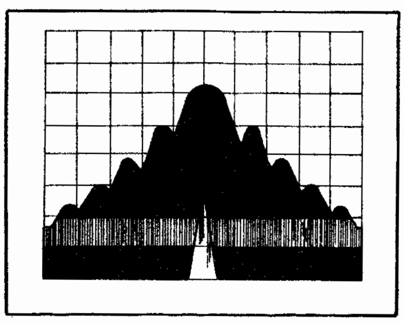
D SAME SIGNAL, BUT B = 300 kHz AND SCAN TIME 2 ms/DIV. THE PRF CAN BE MEASURED TO 1/10 ms = 100 Hz.



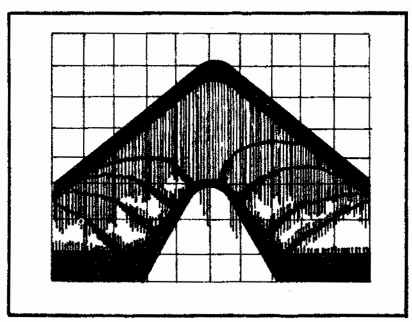
E SAME SIGNAL WITH SCAN WIDTH 5 kHz/DIV, B = 1 kHz, SCAN TIME 0.1 s/DIV.



F SAME SIGNAL WITH B= 300 Hz SCAN WIDTH 10 kHz/DIV, SCAN TIME 0.2 s/DIV.



G SAME SIGNAL WITH B = 3 kHz.



H SAME SIGNAL WITH B = 10 kHz.

Figure 4-106. Pulsed RF Signal in "Pulsed" Spectrum Display

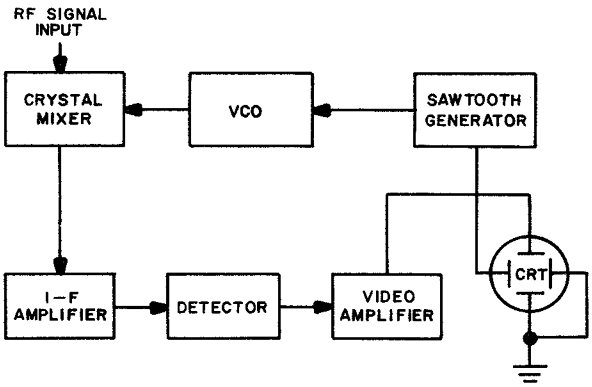


Figure 4-107. Block Diagram of a Heterodyne Spectrum Analyzer

the spectrum analyzer must be long enough, with respect to the amplitude of the signal to be measured, to allow the spectrum analyzer's I-F circuitry to charge and recover. This will prevent amplitude and frequency distortion as illustrated in Figure 4-108.

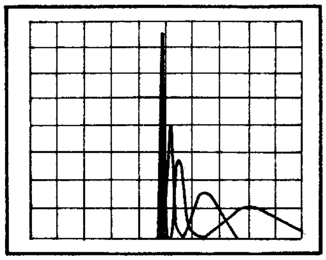


Figure 4-108. Effects of Decreased Scan Time