Spectrum Analysis of Pulse Modulated Waves

By J. C. LOZIER

The problem here is to find the frequency spectrum produced by the simultaneous application of a number of frequencies to various forms of amplitude limiters or switches. The method of solution presented here is to first resolve the output wave into a series of rectangular waves or pulses and then to combine the spectrum of the individual pulses by vectorial means to find the spectrum of the output. The rectangular wave shape was chosen here as the basic unit in order to make the method easy to apply to pulse modulators.

Introduction

The rapidly expanding use of pulse modulation[1] in its various forms is bound to make the frequency spectrum of pulse modulated waves a subject of increasing practical importance. The purpose of this paper is to show how to determine the frequency spectrum of these waves by methods based as far as possible on physical rather than mathematical considerations. The physical approach is used in an attempt to maintain throughout the analysis a picture of the way in which the various factors contribute to a given result. To further this objective the fundamentals involved are reviewed from the same point of view.

The method is used here to analyze two distinct types of pulse modulation, namely, pulse position and pulse width modulation.[2] These two cases are especially important for illustrative purposes because their spectra can be tied back to more familiar methods of modulation. Thus it will be shown that, as the ratio of the pulse rate to the signal frequency becomes large, pulse position modulation becomes a phase modulation of the various carrier frequencies that form the frequency spectrum of the unmodulated pulse wave, and pulse width modulation becomes a form of amplitude modulation of its equivalent carriers. The analysis also shows certain interesting input-output relationships that may be obtained from such modulators, treating them as straight transmission elements at the signal frequency.

These relationships are of more than theoretical interest. The pulse position modulator has already been used as phase or frequency modulator to good advantage.[3] The use of a pulse width modulator as an amplifier is

[1] E. M. Deloraine and E. Labin, "Pulse Time Modulation", *Electrical Communications*, Vol. 22, No. 2, pp. 91-98, Dec. 1944; H. S. Black "AN-TRC-6 A Microwave Relay System", *Bell Labs. Record*, V. 33, pp. 445–463, Dec. 1945.

[2] By *pulse position* modulation is meant that form of pulse modulation in which the length of each pulse is kept fixed but its position in time is shifted by the modulation, and by *pulse width* modulation that form in which the length of each pulse varies with the modulation but the center of each pulse is not shifted in position.

[3] L. R. Wrathall, "Frequency Modulation by Non-linear Coils", *Bell Labs. Record*, Vol. 23, pp. 445–463, Dec. 1945.

another practical application, of which the self oscillating or hunting servomechanism is an example.

The quantitative analysis of such systems depends on the ratio of the pulse repetition rate to the signal frequency. When this ratio is low, the solution can be obtained by a method shown here for resolving the modulated waves into selected groups of effectively unmodulated components. This technique is powerful since it can be done by graphical means whenever the complexity of either the system or the signal warrants it. When the ratio of pulse rate to signal frequency becomes high enough, such methods are no longer practical. However, under these conditions other methods become available, especially in cases like those mentioned above where the spectrum of the modulation approaches one of the more familiar forms. An important example of this occurs in the case of the pulse position modulator where, as the spectrum approaches that of phase modulated waves, the solution can often be found by the conventional Bessel's function technique used in analyzing phase and frequency modulators.

The method proposed here for obtaining the spectrum analysis of pulse modulated waves is based on the use of the magnitude-time characteristic of the single pulse and its frequency spectrum as a pair of interchangeable building blocks, so that the analysis will develop this relationship. Before doing this the elementary theory of spectrum analysis will be reviewed

Review of the Elementary Theory of Spectrum Analysis

A complex wave may be represented in two ways. One way is by its magnitude at each instant of time. The other way is by its frequency spectrum, that is, by the various sinusoidal components that go to make up the wave. The two representations are interchangeable.

The transformation from a given frequency spectrum to the corresponding magnitude vs. time function is straight-forward, for it is apparent that the various components in the frequency spectrum must add up to the desired magnitude-time function. The necessary additions may be difficult to make in some cases but they are not hard to understand.

The reverse process of finding the frequency spectrum when the magnitude-time characteristic is given is more involved, though using Fourier analysis, the problem can generally be formulated readily enough. Furthermore the mathematical procedures involved can be interpreted physically in broad terms by modulation theory. However, these procedures become more difficult to perform, and the physical relationships more obscure, as the wave form under analysis becomes more complex. This is particularly true when general or informative solutions rather than specific answers are required. Pulse modulated waves are sufficiently new and complex to give such difficulties.

The process of finding the frequency spectrum of a complex wave from its magnitude-time function has a simple mathematical basis. It depends on the fact that the square of a sinusoidal wave has a positive average value over any interval of time, whereas the product of two sinusoidal waves of different frequencies will average zero over a properly chosen interval of time.[4]

In theory then, as the magnitude-time function of a complex wave is the sum of all the components of the frequency spectrum, we have only to multiply this magnitude-time function by a sinusoidal wave of the desired frequency and then average the product over the proper time interval to find the component of the spectrum at this frequency.[5]

One physical interpretation of this procedure can be given in terms of modulation theory. The product of the magnitude-time function with a sinusoidal wave will produce the beat or sum and difference frequencies between the frequency of the sinusoid and each component of the frequency spectrum. Thus, if the spectrum contains the same frequency, a zero beat or dc term is produced, and this term may be evaluated by averaging the product over an interval that is of the proper length to make all the ac components vanish.

The application of this principle for spectrum analysis is simple when the magnitude of the wave in question is a periodic function of time. The very fact that the wave is periodic is sufficient proof that the only frequencies that can be present in the wave are those corresponding to the basic repetition rate and its harmonics. Thus the frequency spectrum is confined to these specific frequencies and so it takes the form of a Fourier series. Knowing that the possible frequencies are restricted in this way, the problem of finding the frequency spectrum of a complex periodic wave is reduced to one of performing the above averaging process at each possible frequency. The period of the envelope of the Complex Wave is the proper time interval for averaging, and the integral formulation for obtaining this average is that for determining the coefficients in a Fourier series.

The principle holds equally well when the magnitude-time function is nonperiodic, but the concept is complicated by the fact that the frequency spectrum in such cases is transformed from one having a discrete number of components of harmonically related frequencies to one having a continuous band of frequencies.[6] Such spectra contain infinite numbers of sinusoidal

[4] The proper time interval is generally some integral multiple of the period corresponding to the difference in frequency of the two sinusoid waves.

[5] In practice it is generally necessary to multiply by both sine and cosine functions because of possible phase differences.

[6] One exception to this statement is the fact that any wave made up of two or more incommensurate frequencies is nonperiodic. Yet such waves will have a discrete spectrum if the number of components is finite. This incommensurate case is neglected throughout the discussion.

components, each of infinitesimal amplitude and so close together in frequency as to cover the entire frequency range uniformly.

The continuous band type of frequency spectrum is just as characteristic of non-periodic waves as the discrete spectrum is of periodic waves. This can be shown as a logical extension of the Fourier series representation of periodic waves. The transition from a frequency spectrum consisting of a series of discrete frequencies to one consisting of a continuous band of frequencies can be made by treating the non-periodic function as a periodic function in which the period is allowed to become very large. As the period approaches infinity the fundamental recurrence rate approaches zero, so that the harmonics merge into a continuous band of frequencies.

This does not of course change the basic realtionship between the frequency spectrum of a wave and its magnitude-time function. The magnitude-time function is still the sum of the components of the frequency spectrum. Also the frequency spectrum can still be obtained frequency by frequency, by averaging the product of the magnitude-time function and a unit sinusoid at each frequency. However, the actual transformations in the case of the non-periodic functions require summations over infinite bands of frequencies and over infinite periods of time and so fall into the realm of the Fourier and similar integral transforms.

However, in any case the problem of spectrum analysis reduces to an averaging process. The process can be performed by mathematical integration in all cases where a satisfactory analytical expression for the magnitude-time function is available. Fourier analysis provides a very powerful technique for setting up the necessary integrals in such cases.

This averaging process can also be done graphically. It is apparent from the theory that if the product of the magnitude-time function and the sinusoid is sampled at a sufficient number of points, spaced uniformly over the proper time interval, then the average of the samples gives the desired value. This technique is fully treated elsewhere[7] so that it will not be considered in detail here. However, use will be made of it in a qualitative way to augment the physical picture.

Non-Linear Aspects

The use of the frequency spectrum in transmission studies is generally limited to cases where the system in question is linear; that is, where the transmission is independent of the amplitude of the signal. However, the same techniques can still be used on systems employing successive linear and non-linear components, in cases where the transmission through the non-linear elements is independent of frequency. Under these conditions, the magnitude-time representation of the wave can be used in computing

[7] Whittaker and Robinson, Calculus of Observations.

the transmission over each non-linear section, where the transmission is dependent only on the amplitude, and the frequency spectrum used over each linear section, where the transmission is dependent only on the frequency. This a technique can be used on most pulse modulating systems because such non-linear elements as the modulators and limiters generally encountered are substantially independent of frequency.

Frequency Spectrum of the Single Pulse

The single pulse is a non-periodic function of time and so has a continuous frequency spectrum. In this case the Fourier transforms are simple. They are derived in Appendix A. Figure 1 gives a graphical representation of the magnitude-time function and the frequency spectrum of the pulse. The expressions are general and hold for pulses of any length or amplitude.

It is instructive to note that the frequency spectrum in this case can be

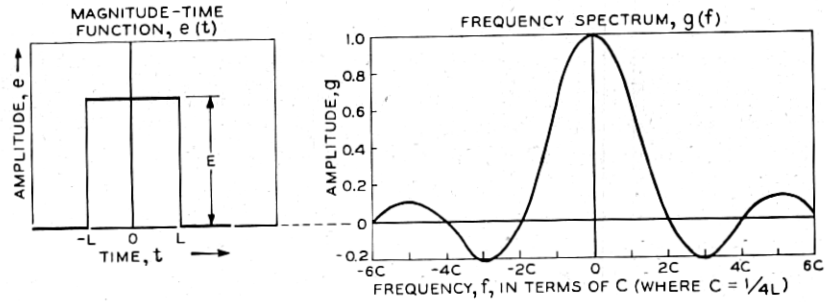

Fig. 1—Magnitude time and frequency spectrum representations of a single pulse.

determined by using the graphical technique mentioned previously. For example, consider the product of the magnitude-time function of the single pulse with a sinusoidal wave of given frequency and unit amplitude, so arranged in phase that its peak coincides with the center of the pulse. Theoretically the average of this product taken over the infinite period will give the relative magnitude of the component in the frequency spectrum of the pulse having the same frequency as the sinusoidal wave. In this case however, the average need only be taken over the length of the pulse, since the product vanishes everywhere else. Thus at very low frequencies, where the period of the sinusoidal wave is very much greater than the length of the pulse, the average is proportional to $2EL$ where E is the amplitude and $2L$ the length of the pulse. Then as the frequency increases, the average of the product, and hence the relative amplitude of the component in the spectrum, will first decrease. For the particular frequency such that the length of the pulse is one half the period, the relative amplitude will have fallen to $2EL \times \frac{2}{\pi} \left(\frac{2}{\pi}\right.$ being the average value of a half wave of unit ampli tude$\left.\right)$. Similarly when the frequency is such that the length of the pulse is a full wavelength, the average will vanish, and when the pulse length is one and a half times the wavelength, the average is negative, having two negative and one positive half waves over the length of the pulse, and the relative magnitude is $2EL \times \frac{2}{3\pi}$. These products are shown graphically on Fig. 2. Since these amplitudes correspond to those given in Fig. 1, for the spectrum components at $f = f_0 = 1/4L$, $2f_0$, and $3f_0$, it is apparent that the spectrum could be determined in this way.

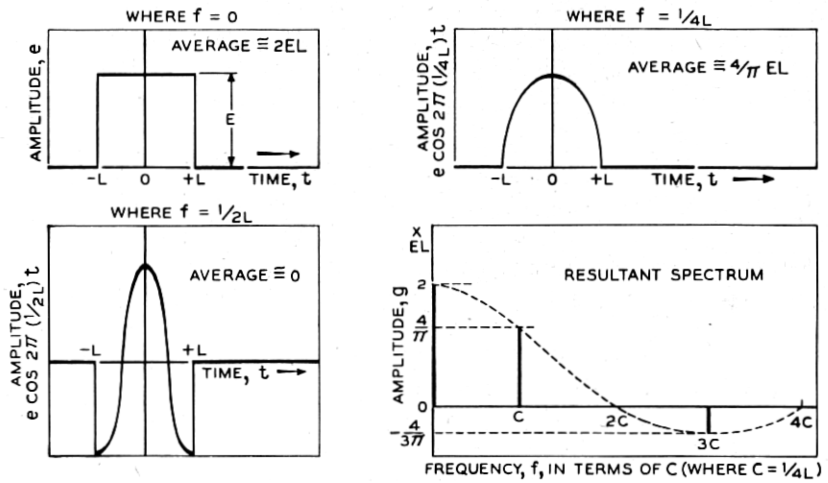

Fig. 2—Graphical derivation of spectrum of single pulse by averaging product of pulse with sinusoidal waves of various frequencies.

Basic Technique

In the analysis presented here, the single pulse and its spectrum will be used in such a way that the need for individual integral transforms for each complex wave form under study is avoided. The theory is simple.

A complex wave form may be approximated to any desired accuracy by a series of pulses, varying with respect to time in length, in amplitude, and in position. Now the spectra of these individual pulses are already known. Therefore, to find the frequency spectrum of the complex wave in question, it is necessary only to combine properly the spectra of the various pulses representing the complex wave.

Thus the process is theoretically complete. The procedure is first to break down the given complex wave into a series of single pulses. Next the spectrum of each pulse is determined separately. Then the spectrum of the complex wave is obtained by combining the spectra of the various single pulses involved. One of the things to be demonstrated here is that it is perfectly feasible in many cases to perform these summations graphically, even though basically it does involve the handling of spectra each containing an infinite number of frequency components.

There are other wave forms that could be used as the fundamental building block instead of the single pulse. The unit step function is one possibility, since it is used in transient analysis for a similar purpose. However, the single pulse has obvious advantages when the complex wave to be analyzed is itself a series of pulses, as in pulse modulation. Again it would be nice to be able to choose as the fundamental unit a wave that has a discrete rather than a continuous band frequency spectrum, but it seems that any wave flexible enough to make a satisfactory building unit is inherently nonperiodic and so has a continuous frequency spectrum. However the fact that the fundamental units have continuous spectra does not of itself complicate the results. If for example, the wave to be analyzed is periodic, the sum of the spectra of the various pulses must reduce to a discrete frequency spectrum. In the cases of interest here, when the pulse train under analysis is repetitive, combinations of identical pulses will be found to occur with the same fundamental period, and generally the first step in the summation of such spectra is to group the series of pulses into periodic waves with discrete spectra.

Manipulations of Single Pulses

In its use, the single pulse may be varied in amplitude, in length, and in position with respect to time. These changes have independent effects on the frequency spectrum. A variation in the amplitude of a pulse does not change its spectrum, except to increase proportionately the magnitudes of all components. A change in position of a pulse with time does not change the amplitude vs. frequency characteristic of the spectrum, but it does shift the phase of each component by an amount proportional to the product of the frequency and the time interval through which the pulse was shifted. A change in the length of a pulse will change the shape of the amplitude vs. frequency characteristic of the spectrum. Figure 3 shows this effect. However, if the center point of the pulse is not shifted in time, the relative phases of the components are not affected by such changes in length.

The single pulse can also be modulated to aid in the resolution of more complicated wave forms. This process is based on the use of the pulse as a function having a value of unity over a chosen time interval and a value of zero at all other times. Thus, to show a part of a sinusoidal wave, we need only multiply this wave by a pulse of the correct length and proper phase with respect to the sinusoid to show only the desired piece of the wave. In this simple case it is not difficult to derive the spectrum because what are produced are the sum and the difference products of the modulating frequency with the spectrum of the pulse. This gives two single pulse spectra shifted up and down in frequency by the frequency of the modulation. An example of this is shown in Fig. 4, where the spectrum of a single half cycle is determined.

Pulse Position Modulation

For the first example, a simple form of pulse position modulation will be analyzed. The pulse train in this case is made up of pulses spaced T seconds

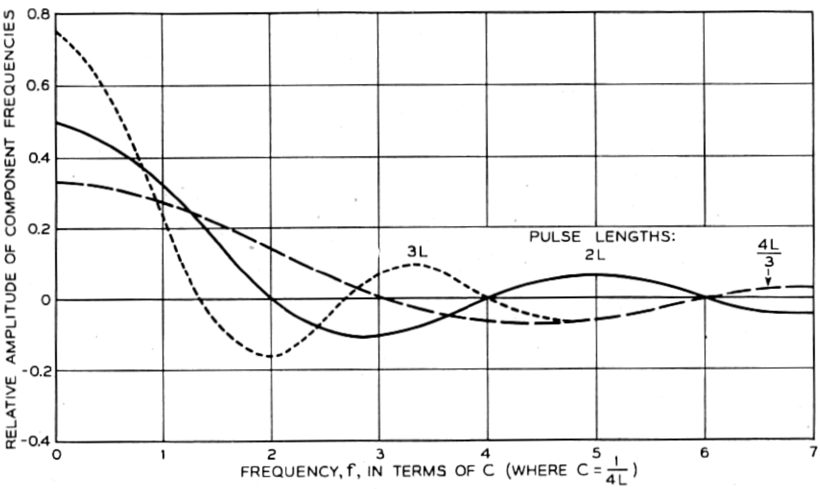

Fig. 3—Change in frequency spectrum with pulse length.

apart and the width of each pulse is a very small part of the spacing T. Such a pulse train is shown on Fig. 5. The pulse train is modulated by advancing or retarding the position (time of occurance) of the pulses by an amount proportional to the instantaneous amplitude of the signal at sampled instants T seconds apart. Figure 5 also shows the signal, in this case a sine wave of frequency $1/10T$, and the resulting modulated pulse train. The peak amplitude of the modulating sine wave is assumed to shift the position of a pulse by $1/4T$. The length and the amplitude of the pulses are the same since neither is affected in this type of modulation.

The first step in the analysis is to determine the spectrum of the pulse train before modulation. Each pulse contributes a spectrum of the form shown on Fig 1. Now the phase of each component in such a spectrum is so arranged that the spectrum forms a series of cosine terms all of which have zero phase angle at the center of the pulse. From successive pulses T

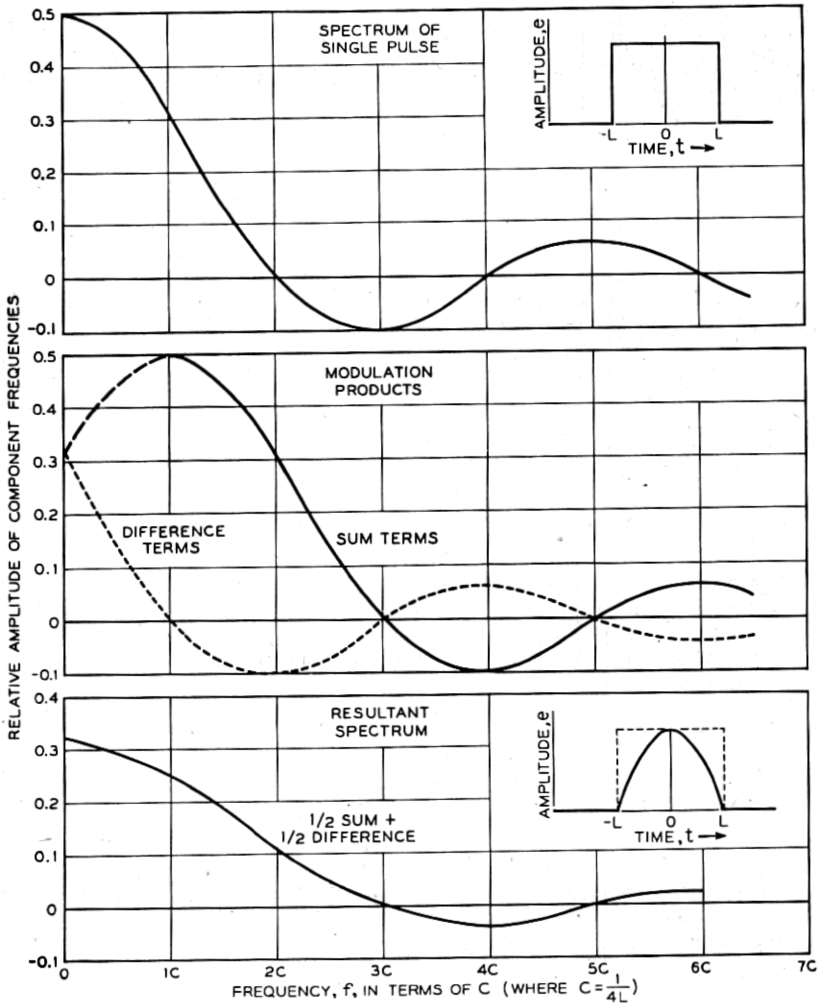

Fig. 4—Determination of spectrum of single half sine wave by modulation of single pulse spectrum with cos $2\pi ct$.

seconds apart, the component at any given frequency will have the same amplitudes, but the relative phases will be $2\pi fT$ radians apart. It is apparent that frequencies for which $2\pi fT$ is 2π or some multiple of 2π radians apart, the contributions from all pulses add in phase. These are the frequencies nc, where $n = 1,2,3$ and $c = \frac{1}{T}$. It is also apparent that at frequencies for which the phase differences between the components are not an exact multiple of 2π radians apart, the contributions from enough pulses must be spread in phase over an effective range of 0 to 2π radians in such a way as to cancel one another. For example, take the particular frequency for which the difference in phase between pulses is 361° instead of 360°.

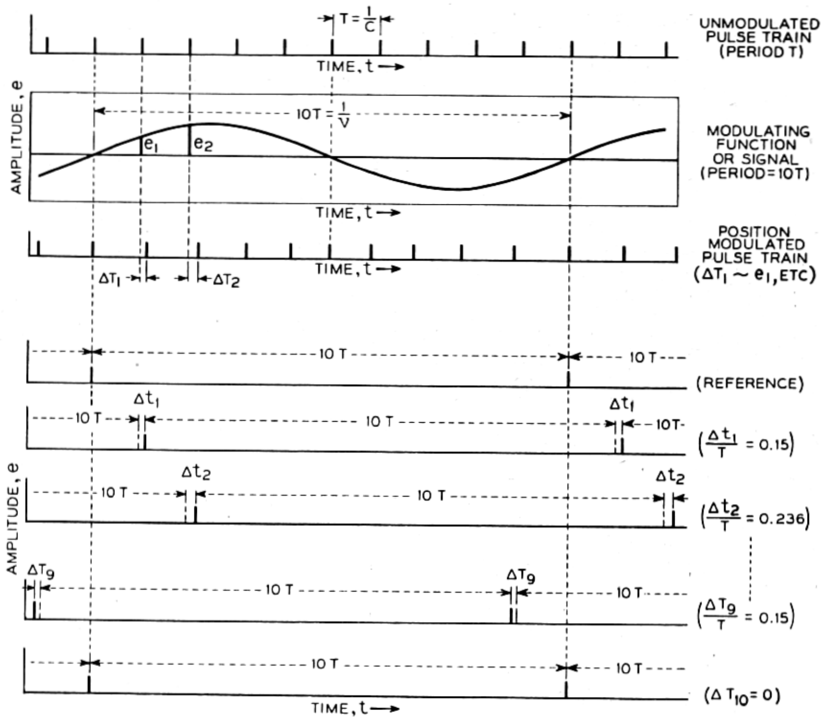

Fig. 5—Formation of pulse position modulated pulse train and its resolution into subsidiary unmodulated pulse trains.

The contribution from each preceding pulse will be effectively advanced in phase 1° with respect to its successor, so that the contributions from pulses 180 periods apart will be exactly 180° out of phase. Therefore over a sufficient number of pulses, the net contribution is zero.

The spectrum of the unmodulated pulse train is thus made up of a dc term plus harmonics of the frequency $C = 1/T$. The dc term is the average, and therefore is equal to $E \times 2L/T$, where E is the magnitude of the pulse. All of the other components have the same relative magnitudes that they have in the single pulse spectrum. This gives a spectrum like that shown on Fig. 6. Figure 6 also shows for comparative purposes the spectrum of the subsidiary pulse wave consisting of every 6th pulse.

Thus in the unmodulated case, the pulses have a uniform recurrence rate and the resultant spectrum, found by adding those of the individual pulses, reduces to a train of discrete frequencies comprised only of the harmonics of the recurrence rate of the pulses. The fundamental frequency, correspond-

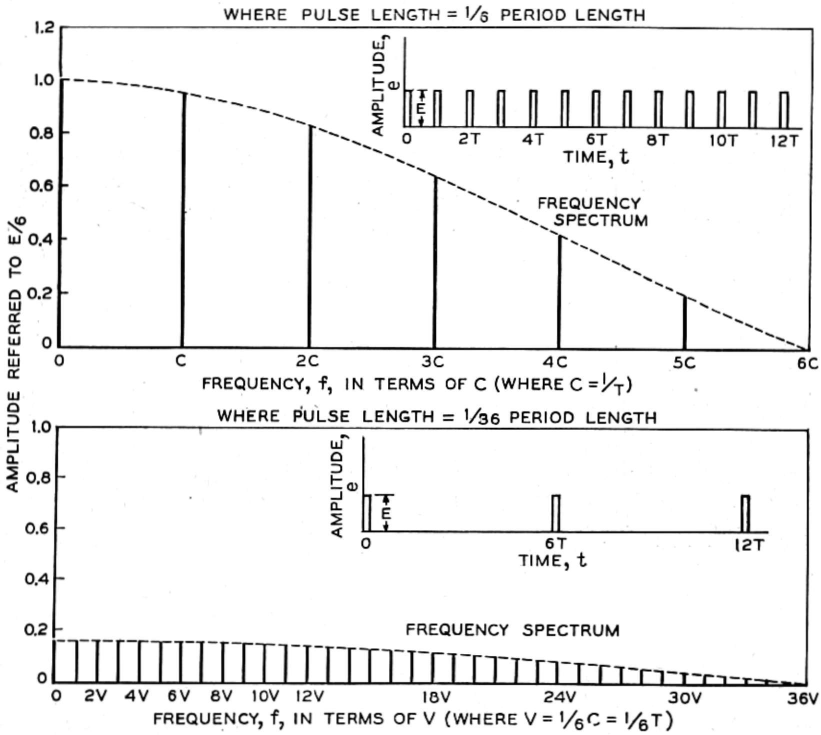

Fig. 6—Frequency spectrum of pulse trains where the spacing between the pulses is 6 and 36 times the pulse length respectively.

ing to the recurrence rate, and its harmonics will be called the carrier frequencies of the pulse train. The effect of modulating the pulse train is to modulate each of these carriers, producing sidebands of the signal about them.

When the pulse train is position modulated, the pulses are shifted in position by an amount ΔT, corresponding to the instantaneous amplitudes of the modulating function. The spectrum of each pulse is unchanged, since the pulse length remains constant. However, components of successive pulses at the carrier frequency c and its harmonics will no longer add directly, because of the phase shifts that accompany the change in position. This phase shift is equal to ΔT, the shift in position, times the radian frequency of the component in question.

However, when the signal function is periodic, each pulse will have the same shift in position as any other pulse that occurs at the same relative instant in a later modulating cycle. Furthermore, when the carrier frequency is an exact multiple of the signal frequency i.e., $c = nv$, there will be a pulse recurring at the same relative instant in each cycle of v. Under these conditions, the pulse position modulated wave can be broken down into a group of unmodulated waves, each being made up of that series of pulses that recur at a given part of each modulating cycle, as shown in Fig. 5. These subsidiary waves are effectively unmodulated because, as each pulse recurs at the same instant in the modulating cycle, they are shifted to the same extent and hence will be uniformly spaced. This uniform spacing between pulses in a given wave is equal by definition to the period of the modulating function, and there will be as many of these unmodulated pulse trains as there are pulses in a single cycle. Thus, if $c = nv$, there will be n such pulse trains.

The reason for grouping the pulses into these unmodulated pulse tains is that unmodulated periodic trains have spectra of discrete frequencies. Since the pulse widths are all equal, and since the spacing between pulses is the same for each wave, the spectra of these unmodulated waves will all be identical. Furthermore, these spectra will be the same as that of the original carrier wave of pulses before modulation, except for two factors. First, the fundamental frequency is now v, corresponding to the modulating period, so that there are n times as many components as before. Secondly the amplitudes are reduced by the factor $\frac{1}{n}$ because there is only one pulse in these new waves to every n pulses in the original wave. Thus, instead of having a spectrum made up of the carrier frequency and its harmonics, we now have one made up of harmonics of v. Since $c = nv$, such frequencies as c, $c, \pm v$, $c \pm 2v$, etc., are included. An example of the spectra of both the subsidiary and original pulse waves is shown on Fig. 6, for the case where $n = 6$.

Thus the problem of finding the spectrum of such a pulse position modulated wave is reduced by this procedure to adding up the n equal components at each of the frequencies of interest, such as c and $c \pm v$, allowing for the phase difference between components corresponding to the position of one pulse with respect to that of the other n–l pulses in one modulating cycle. As an example, suppose $n = 10$ and the frequency to be computed is $c + v$. Now $c + v$ is 10% higher in frequency than c. Thus in the unmodulated case, when the n pulses are equally spaced, they are 360° apart at c and consequently 360° + 36 or 396° at $c + v$. Therefore in the unmodulated case, each component would be advanced in phase 36° with respect to the previous one, so that the diagram of the 10 components would form the

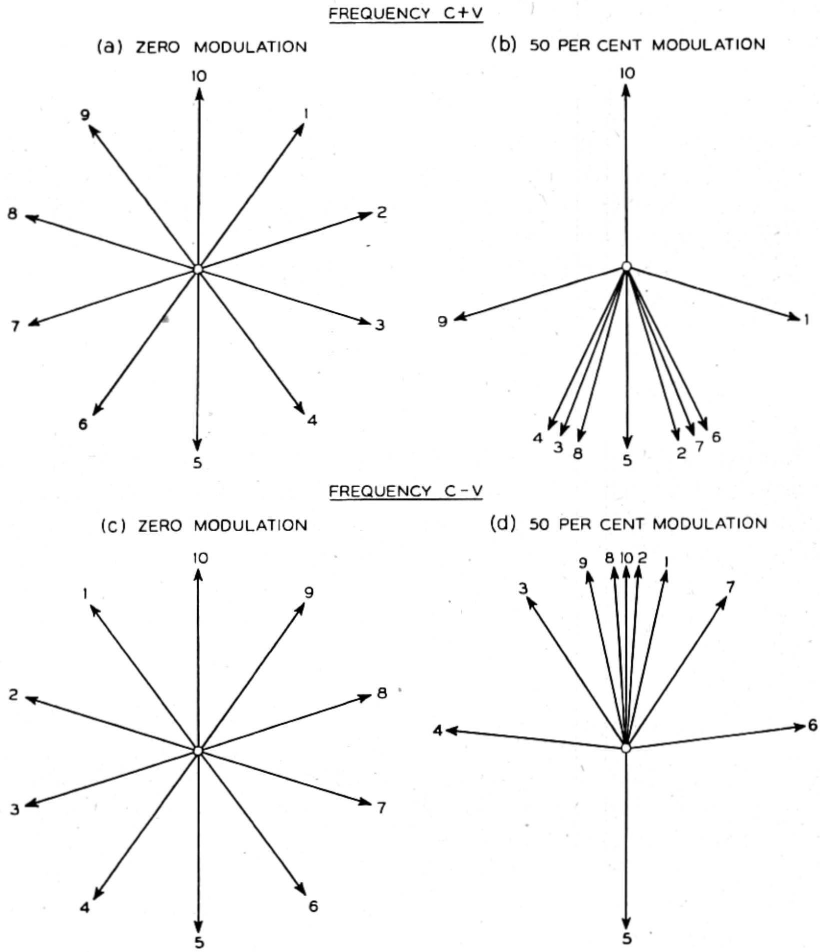

Fig. 7—Vector pattern of subsidiary pulse components.

vector pattern shown on Fig. 7A. The successive components are numbered 1 to 10. The sum in this unmodulated case is of course zero.

Now the effect of modulation is to shift the relative phases of these components by an amount determined by the shift in position of the corresponding pulses. When these relative phase shifts are such as to spoil the cancellation of the 10 components, a net component of this frequency is produced in the frequency spectrum of the pulse wave. Taking the example shown in Fig. 5, the 10 components in Fig. 7A would be shifted to the positions shown in Fig. 7B. These shifts in relative phase are determined in the following way. Figure 5 shows that the number 1 pulse is retarded an amount ΔT_1 equal to 15% of T, the normal spacing between pulses. Thus at the carrier frequency c, the phase shift between the component from this retarded pulse and the reference pulse is 15% more than 360° or 414°. Thus the component at the carrier frequency c from the first subsidiary pulse train is shifted 54° from its unmodulated position.

At $c + v$, since the frequency is 10% higher, the net shift is 10% more than at c or 59.5°. Thus the number 1 component on the vector diagram of Fig. 7B is rotated 59.5° clockwise from its unmodulated position shown on Fig. 7A.

Similarly pulses 2 and 3 are each shifted in position by equal amounts, ΔT_2 and ΔT_3. These shifts in position give 85° phase shift at the carrier frequency. Hence components 2 and 3 at $c + v$ are each rotated 10% more or 93.5° from their respective unmodulated reference positions shown on Fig. 12A. Component number 4 is shifted 59.5° clockwise just as number 1. Component 6 and 9 are also shifted 59.5° each, but in this case the modulating function has the reverse polarity so that the components are rotated counterclockwise. Similarly components 7 and 8 are rotated 93.5° counterclockwise.

The sum of these components in the vector diagram of Fig. 7B gives a resultant that is negative with respect to the reference direction and the magnitude that is 58% of the reference magnitude, where the reference magnitude and direction are those for the carrier c with no modulation.

This gives the relative magnitude and phase of the $c+v$ term produced by pulse position modulation for the case where the modulating function is a sine wave of frequency $v = c/10$ with a peak amplitude just large enough to shift a pulse by 1/4 of T, where T is the spacing between unmodulated pulses. A shift of this magnitude will be defined here as 50% modulation on the basis that 100% modulation should be $1/2T$, the maximum displacement that can be used without possible interference between pulses.

In the same way the other component frequencies in the spectrum such as $c, c - v, c \pm 2v$, etc., have been computed for the above case of 50% modulation, and for other peak amplitudes of the modulating sine wave giving 25%, 70% and 100% modulation. In all cases the frequency of the modulating function was held at $v = c/10$. This information is plotted on Fig. 8, showing v, c and the various components of the frequency spectrum that represent the sidebands about the carrier frequency c, as a function of the peak % modulation.

The above solution assumed a special case where c was an exact multiple of v. The purpose of this assumption was to simplify the problem to the extent that the periodicity of the modulated wave would be the same as that of the modulating function. There are two other possible cases. For one, the ratio of c to v could be such that a pulse would occur at the same instant of the modulating period only once every so many periods. The actual periodicity of the modulated pulse wave would be reduced accordingly because it would make the same number of periods of the modulating function before the modulated pulse train is repeated. This is a result of the fact that pulse modulation provides for a discrete sampling rather than a continuous measure of the modulating wave. The technique of spectrum analysis demonstrated above is just as applicable to this case as it was to the simpler one. However, there will be comparatively more terms to be handled. The other possible case is the one where c and v are incommensurate.[8] In this case, the resulting modulated wave is non-periodic. However, on the basis that the spectrum is practically always a continuous function of the signal frequency, this case has received no special attention here.

At frequencies for which c is very much greater than v, so that the number of component pulse trains becomes too numerous to handle conveniently in the above fashion, the sidebands about each carrier or harmonic of the switching frequency can be computed by the standard methods for phase modulation, as the next section will demonstrate. This result follows directly from the theorem that as the carrier frequency c becomes large with respect to v, pulse position modulation merges into a linear phase modulation of each of the carriers.

Pulse Position Modulation vs Phase Modulation

When a pulse, in a pulse position modulated wave, is shifted by 1/2 the spacing between pulses (100% modulation) it is apparent from the previous discussion that the component of the carrier in the frequency spectrum of the pulse is shifted by 180°. Therefore to compare the spectrum of a pulse position modulated wave like that on Fig. 8 with the equivalent spectrum of a phase modulated wave, what is needed is Fig. 9, showing the frequency spectrum of a phase modulated wave of the form $Cos(ct - k \sin vt)$ as a function of k for values of k up to π radians or 180°. The computation of the frequency spectrum of such a phase modulated wave has been adequately covered elsewhere and all that is done here is to give the brief development shown in appendix B.

[8] Mr. W. R. Bennett has pointed out that this incommensurate case is the general one. It requires a double Fourier series, which reduces to a single series when the signal and carrier frequencies are commensurate. This analysis is based on the single Fourier series.

A comparison of the spectra on Figs. 8 and 9 shows that the sidebands have the same general pattern. However comparative sidebands are not

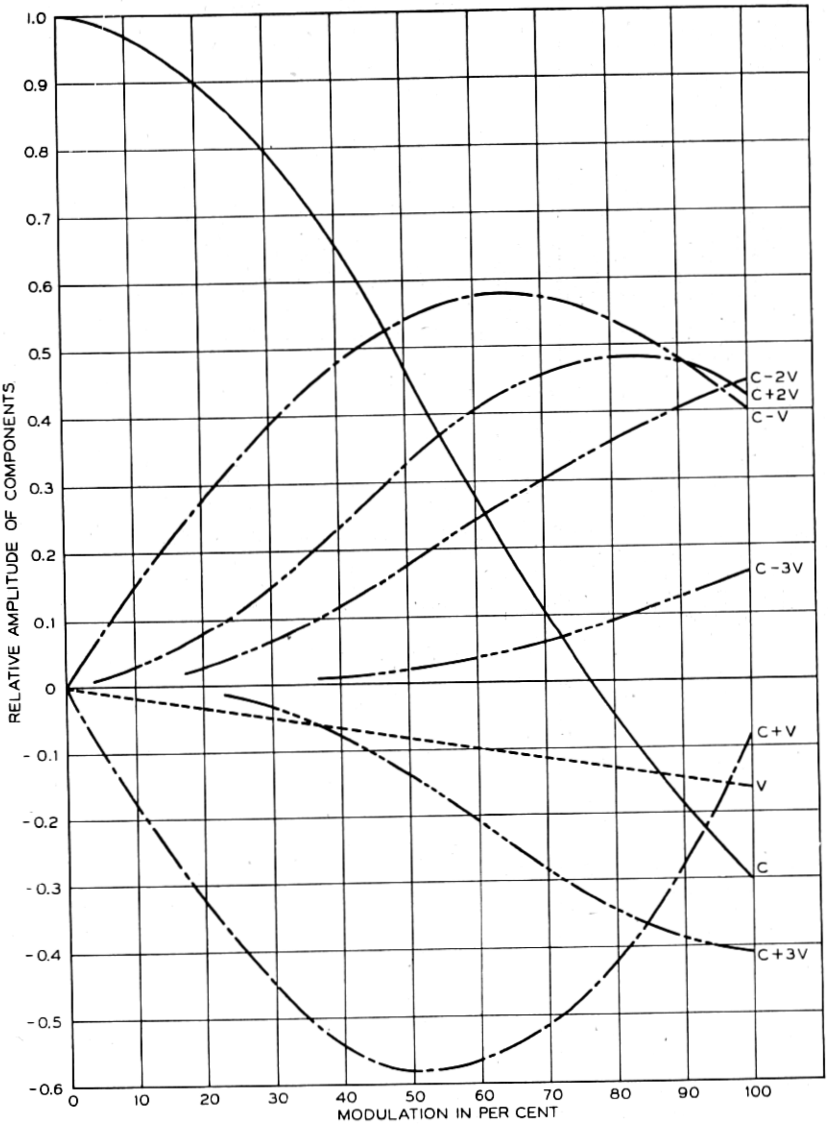

Fig. 8—Spectrum of pulse position modulated wave for case where the carrier frequency C is 10 times the signal frequency v.

quite equal in the two cases. In fact comparable upper and lower sidebands in the case of the pulse modulated wave shown on Fig. 8 are not equal in absolute magnitude to each other. This lack of symmetry is due to the fact that c is only 10 times v.

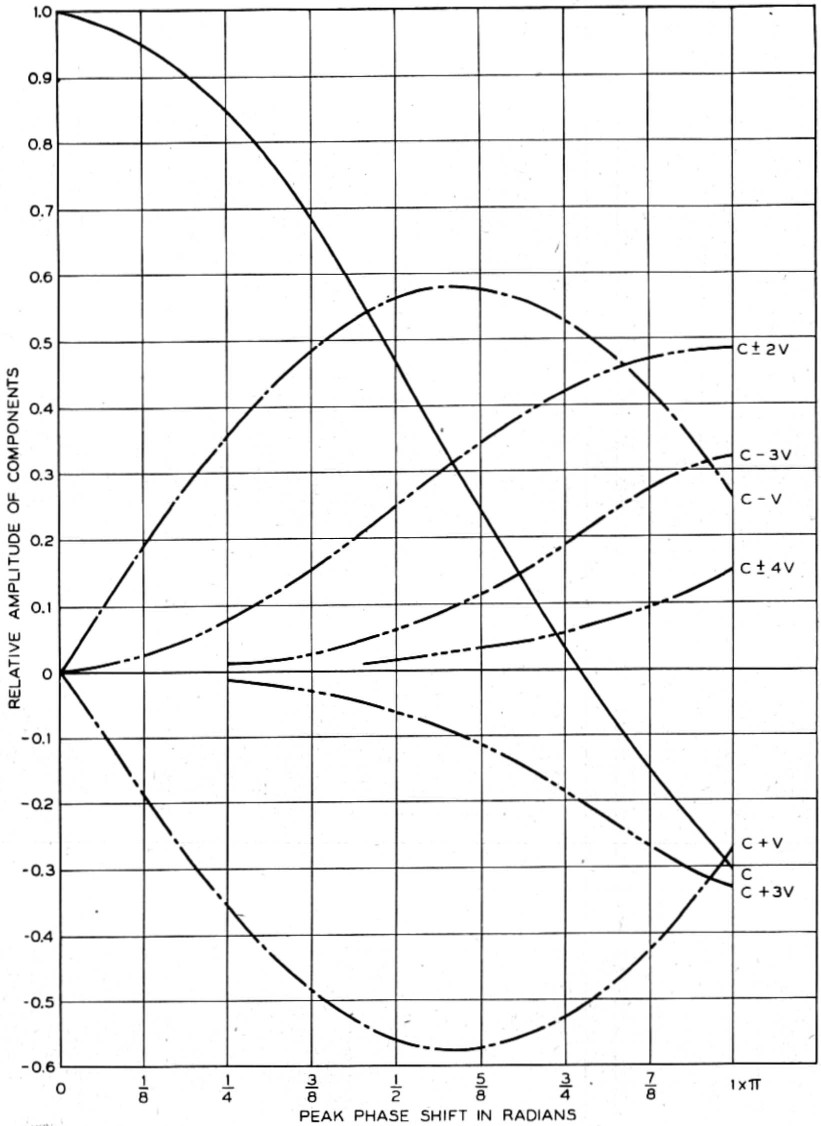

Fig. 9—Spectrum of phase modulated wave cos $(ct + k \sin vt)$ as function of peak phase shift k for values of k up to π radians.

One way of proving this is to go through the process of computing the $c - v$ term in this pulse modulated wave just as the $c + v$ term was computed earlier. Since the frequency $c - v$ is 10% less than c, the unmodulated pattern of the 10 subsidiary components, as shown on Fig. 7C, is the mirror image of that for $c + v$ in 7A, for the first component is now 360° less 10% or 324°, and subsequent components are each retarded 36° with respect to the previous one. When the pulse train is modulated the effect is similar to the case for $c + v$ and, for the same per cent modulation, the Vector pattern of Fig. 7D is formed. The resultant in this case differs from that of 7B in sign as well as in magnitude. The difference in sign comes from the fact that, since component 1 in 7A corresponds to component 9 in 7C and component 2 in 7A to component 8 etc., the modulation in the case of $c - v$ rotates these corresponding components in opposite directions. The difference in magnitude is due to the fact that since $c - v$ is an appreciabley lower frequency than $c + v$ in this case (approx. 20%), the phase shift corresponding to a given shift in pulse position is proportionately less. Thus the corresponding Vector components are not shifted the same number of degrees. Thus the absolute magnitudes of $c + v$ and $c - v$ are not equal in this case.

It is apparent that this difference in magnitudes of $c + v$ and $c - v$ becomes smaller as the carrier frequency c becomes larger with respect to v. In the limiting case of c very much greater than v, $c + v$ and $c - v$ would each be shifted the same number of degrees as c itself. If this more or less compromise shift of c is used to compute the $c \pm v$, $c \pm 2v$, and $c \pm 3v$ terms, then the resulting frequency spectrum is that of the phase modulated carrier on Fig. 9.

The higher harmonics of c in the pulse position wave are similarly phase modulated and the interesting point is that $2c$ is modulated through twice as many degrees phase shift and $3c$ 3 times as many degrees, etc. Thus a single pulse position modulator could be designed to produce a harmonic of c with almost any desired degree of phase modulation. This is a useful method for obtaining a phase modulated wave, or with a 6 db per octave predistortion of the signal, a frequency modulated wave.

Figure 8 also shows a term in v itself, which has been neglected so far in the discussion. It is apparent that the components at v contributed by the 10 subsidiary unmodulated waves must form the same kind of vector pattern as those of $c + v$ in Fig. 7. However, in this case $c + v$ is eleven times v in frequency, so that the components of v are rotated only one eleventh as much for a given pulse diplacement. Thus the magnitude of v at 100% modulation is equal to that of $c + v$ at approximately 9% modulation. For different frequency ratios of c to v the relationship of the v term to $c + v$ will vary, and it is apparent that for c very much greater than v, the v term will vanish. The relationship is such that the amplitude of the v component out of the modulator at a given per cent modulation is directly proportional to its own frequency v for all frequencies less than approximately one quarter of c, and the phase is 90° with respect to the input. Thus the modulator puts out a signal component that is the derivative of the input signal.

To summarize the case of pulse position modulation, the frequency spectrum may be determined by the methods based on subdividing the modulated pulse train into a series of unmodulated ones when the ratio of c to v is small, and by treating each harmonic of the carrier as a phase modulated wave of the form $\cos n(ct + \theta)$, where θ is the modulating function, when the ratio of c to v is large. In the case treated here, the modulating function was a simple sinusoidal wave. Of course the analysis holds for more complicated wave shapes having frequency spectra of their own. In this event however the restriction on the relative magnitudes of the frequencies v and c should be taken as one on c and the highest frequency in the modulating spectrum. The complexity of the modulating function does not affect the analysis when it is done by this technique of subdividing the pulse train, since all that need be known is how much each pulse is shifted, and this can be done graphically. The analysis given here has neglected the length of the individual pulses. This was done when it was assumed that the individual contributions from the various pulse trains had the same amplitude at all frequencies. For any finite pulse width, the relative magnitudes of the various components must be modified by the $\frac{\sin x}{x}$ factor of the single pulse, as shown on Fig. 6.

As mentioned in the introduction, a complex wave could be analyzed by multiplying its magnitude-time characteristic by unit sinusoids at each frequency in question, sampling the product at a sufficient number of points uniformly spaced over a cycle of the envelope of the complex wave, and then averaging the values of the product thus obtained. This technique is particularly applicable to the analysis of pulse position modulated waves since, by taking the centers of the pulses of the modulated wave as the sampling instants, it is possible, with a finite number of samples (same as the number of pulses) to get the same results as though a very much greater number of uniformly spaced samples were taken. The interesting thing to note here is that the actual computations that would be involved in applying this sampling method of analysis to a pulse position modulated wave are almost identically the same calculations as required by the technique of resolving the pulse train into unmodulated subsidiary pulse trains used here.

Pulse Width Modulation

Pulse Width Modulation as defined here could also be termed "pure" pulse length modulation. The pulse train in the reference or unmodulated condition is a recurrent square wave, and the lengths of the pulses will be varied by the modulation without changing the position of the centers of the pulses. The term "pure" pulse length modulation is applicable to this special case where the phase relationship between spectra of adjacent pulses does not change with modulation because the centers of the pulses are not shifted by the modulation. The conventional form of pulse length modulation, where one end of the pulse is fixed in position, combines both this pulse width modulation and the pulse position modulation previously analyzed. The interest in this case of pulse width modulation arose in connection with the analysis of "hunting" servomechanisms, and the analysis provides a basis for a general solution of the response of a two-position switch or ideal limiter to various forms of applied voltages.

Since the unmodulated wave is a square wave with pulses of length $2L$ recurring at intervals of $T = 4L$, it has the familiar square wave spectrum including a d-c term, a fundamental term or carrier of frequency $c = 1/T$, a 3rd harmonic with a negative amplitude 1/3 that of the fundamental, etc. Figure 10 shows clearly that this spectrum is the sum of single pulses of width $2L$ spaced $T = 4L$ seconds apart. In the summation, all frequencies cancel except harmonics of c and, since they all add directly in phase, the component frequencies in the resultant spectrum have the same relative amplitudes as they have in one single pulse.

When this pulse train is modulated, the width of each pulse becomes $2(L + \Delta L)$, where the magnitude of ΔL depends in some specified way on the magnitude of thhe modulating function at the instant corresponding to the center of the pulse. For simplicity, the case will be taken where ΔL is proportional to the magnitude of the modulating function. For 100% modulation, ΔL will be assumed to vary from $-L$ to $+L$. Figure 3 shows how the relative amplitude of the components of the frequency spectrum of a pulse vary for 3 different values of ΔL , along with the equation that governs these amplitudes.

If the modulating function has a periodicity v such that $c = 10v$, then every 10th pulse, recurring at the same instant in each modulating cycle, will be widened to the same extent and so can be formed into a subsidiary unmodulated pulse train, as was done on Fig. 5 for the pulse position modulated wave.

Again vector diagrams like those in Fig. 7 may be formed showing the contribution of each of these subsidiary pulse trains at various frequencies such as c, $c + v$ and $c - v$. When the waves are unmodulated, the vector diagrams for the same frequencies will be the same as those for the pulse position modulated case, except for the absolute amplitudes of the components, as long as $c = 10v$ in each case. When the pulse width system is modulated, however, the modulation does not rotate the individual vector components as in the pulse position case since the spacing between pulses is not changed. What the pulse width modulation does is to change the length of the individual component vectors exactly as it does in the case of the single pulses shown on Fig. 3. This change of magnitude, of course, can spoil the cancellation of the ten unmodulated components at some frequency like $c + 2v$ just as effectively as rotating them did in the case of the pulse position modulated wave, thus producing a spectrum component at that frequency.

As an example, the case will be taken where the modulating function is a

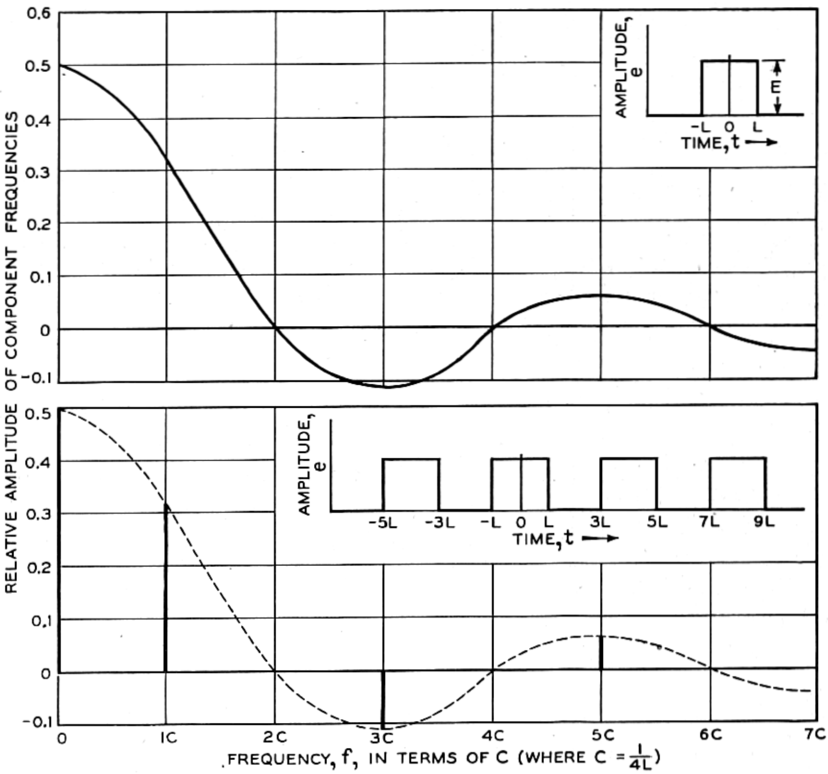

Fig. 10—Comparative spectra of square wave and single pulse.

sinusoid of frequency v. Then the change in width with modulation is given by the formula

$$\frac{\Delta L}{L} = k \sin vt.$$

Since $c = 10v$, the successive subsidiary pulse trains will be modulated an amount $\left(\frac{\Delta L}{L}\right)_m = k \sin\left(2\pi \frac{m}{10}\right)$ as m takes on the values from 1 to 10. Thus the spectra of these subsidiary pulse trains with pulses of length $2(L +$

ΔL_m) recurring every $1/v$ seconds will be a Fourier series of harmonics of v. The amplitude of the nth term of this series will be

$$B_n = \frac{2E}{10\pi n} \sin\left[\frac{\pi n}{2}\left[1 + k \sin\left(\frac{2\pi m}{10}\right)\right]\right].$$

This expression may be found from appendix C, equation (5a). Combining

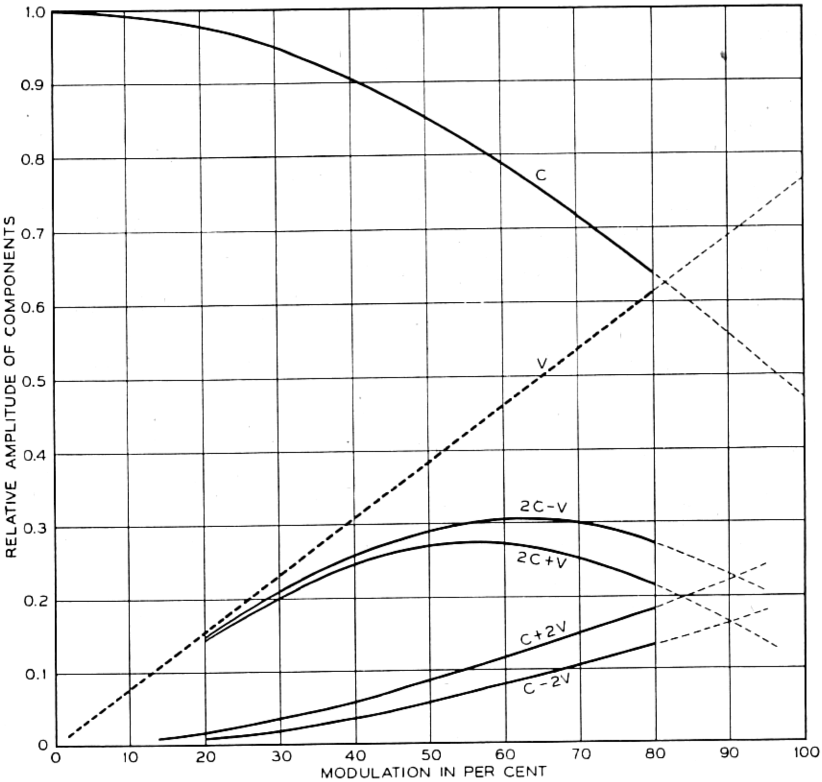

Fig. 11—Spectrum of pulse width modulated wave for case where carrier frequency C is 10 times the signal frequency v.

the 10 such components at each frequency, as shown on Fig. 7 for the case of the pulse position modulated wave, the spectrum for this case of Pulse Width Modulation on Fig. 11 is produced. This spectrum is comparable to that on Fig. 8 for the pulse position modulated case.

Pulse Width vs Amplitude Modulation

That pulse width modulation is a form of amplitude modulation of the carriers of the unmodulated pulse train is shown mathematically by Equa-

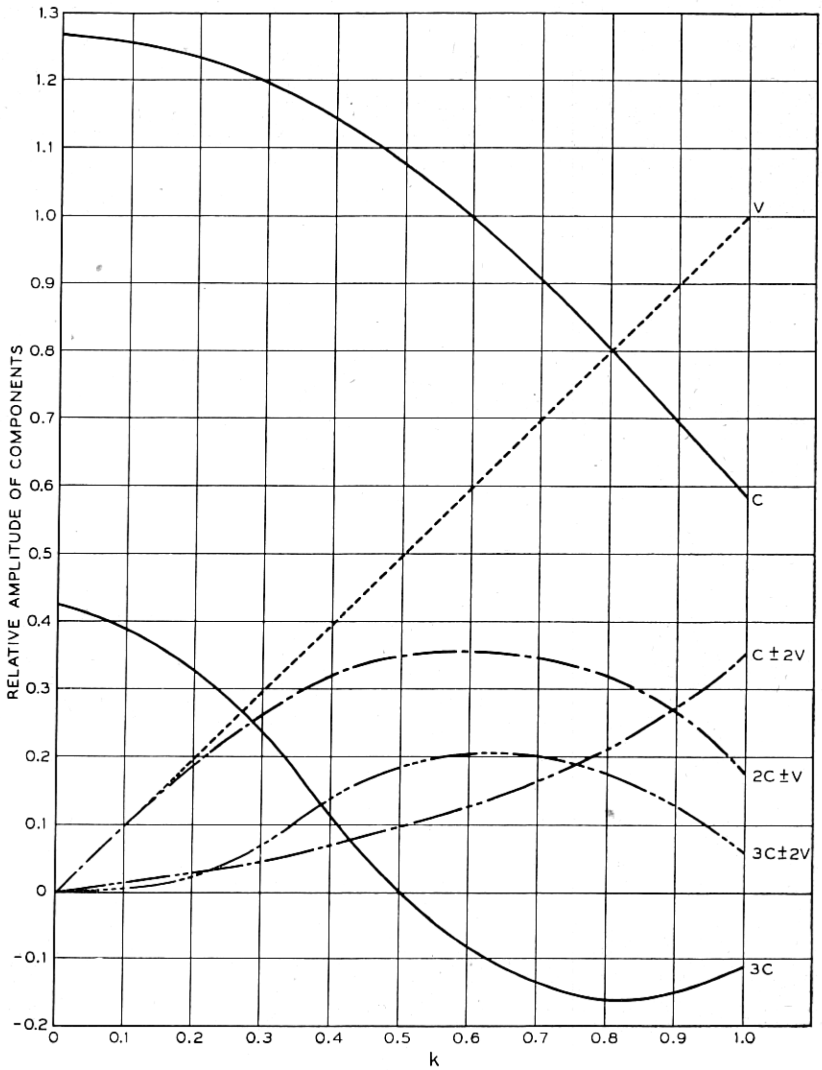

Fig. 12—Response of ideal limiter to simultaneously applied isosceles triangle wave and sine wave inputs. *k* is the ratio of the peak amplitudes of sinusoidal and triangular waves at the input.

tion (8) in Appendix C, where the spectrum is developed as a Fourier series in harmonics of the pulse rate *c* with the modulation affecting only the amplitude of the coefficients.

This mathematical analysis is continued in Appendix D where the frequency spectrum is determined for $\frac{\Delta L}{L} = k \sin vt$. The spectrum thus computed is shown in Fig. 12.

An example of this type of pulse modulator is given by a two position switch or ideal limiter when the signal to be modulated is applied simultaneously to the limiter with an isosceles triangle wave as carrier. The carrier should have a higher peak amplitude than the signal and a recurrence rate based on the desired carrier frequency. Figure 12 is arranged to show the output spectrum for such a limiter in terms of k, when k is the ratio of the peak amplitudes of the sinusoidal signal and triangular carrier wave inputs.

A comparison of this spectrum with that on Fig. 11 shows that the two spectra have almost the same form. c and v have the same amplitude characteristics in each case. The $c \pm 2v$ and $2c \pm v$ terms have differences that are like those found before in comparing the pulse position modulated wave on Fig. 8 and the phase modulated carrier on Fig. 9. As in that case, when c becomes very much greater than v the differences vanish.

Application of Pulse Width Modulator

Practical interest in this case lies in the fact that the signal is present in the output spectrum with a linear characteristic that makes such a modulator a linear amplifier. The "on-off" or "hunting" servomechanism is based on a modified form of such an amplifier in which the carrier is supplied by the self oscillation of the system. The term modified form is used because the self oscillations in general are more nearly sinusoidal than triangular in form and so do not give a linear change in pulse length over as wide a range of input amplitudes as does a triangular carrier. No attempt will be made to analyze such a system here since it has been handled elsewhere.[9] However the above method is applicable to such problems regardless of the shape of the carrier or the signal.

Other Forms of Pulse Modulation

Another form of pulse modulation of interest is that of pulse length modulation in which either the start or the end of each pulse is fixed, so that the centers of the pulses vary in position with the length. This is a combination of both the pulse position and the pulse width modulations described above and can be analyzed by a combination of the methods developed.

These same methods are also applicable to the analysis of frequency and phase modulated waves after they have been put through a limiter, as they generally are before detection.

[9] See L. A. Macall, "The Fundamental Theory of Servomechanisms" D. Van Nostrand Company, 1945.

APPENDIX A

Fourier Transforms For Single Pulse

The amplitude $g(f)$ of the component of frequency f in the spectrum of the Complex Magnitude-time function $e(t)$ is given by the d-c component of the Modulation products of $e(t)$ and *cos* $2\pi ft$, found by averaging the product over the period of the complex wave.

Thus, for non-periodic waves, where the period is from $-\infty$ to $+\infty$, the amplitude of the spectrum at f is

$$g(f) \cong \int_{-\infty}^{\infty} e(t) \cos 2\pi ft \, dt. \tag{1}$$

For the single pulse, where $e(t) = E$ for $-L < t < L$ and $e(t) = 0$ for all other values of t, equation (1) reduces to

$$g(f) \cong \int_{-L}^{L} E \cos 2\pi ft \, dt. \tag{2}$$

Integrating,

$$g(f) \cong \frac{F}{2\pi f} \sin 2\pi ft \Big|_{-L}^{L}$$

or

$$g(f) \cong \frac{E}{\pi f} \sin 2\pi fL. \tag{3}$$

Equation (3) is the expression for $g(f)$ plotted on Fig. 1.

Similarly, in the case of the single pulse, each increment in frequency df contributes a factor proportional to $g(f) \cos 2\pi ft \, df$ to the composition of $e(t)$, so that

$$e(t) = \int_{-\infty}^{\infty} g(f) \cos 2\pi ft \, df. \tag{4}$$

Substituting in (4) the expression for $g(f)$ given by equation (3), this becomes

$$e(t) \cong \frac{E}{\pi} \int_{-\infty}^{\infty} \frac{\sin 2\pi fL}{f} \cos 2\pi ft \, df. \tag{5}$$

APPENDIX B

Frequency Spectrum Of Phase Modulated Wave

The Phase Modulated Wave in this case is given by

$$\cos(ct - k \sin vt) = \cos(ct) \cos(k \sin vt) + \sin(ct) \sin(k \sin vt)$$

Now $\cos(ct) \cos(k \sin ct) = J_0(k) \cos(ct)$

$$+ J_2(k) \cos(c - 2v) t$$

$$
\begin{aligned}
& \qquad + J_2(k)\cos(c+2v)\,t + \cdots \\
\text{and } \sin(ct)\sin(k\sin ct) &= J_1(k)\cos(c-v)\,t \\
& \quad - J_1(k)\cos(c+v)\,t \\
& \quad + J_3(k)\cos(c-3v)\,t \\
& \quad - J_3(k)\cos(c+3v)\,t + \cdots \\
\therefore \cos(ct - k\sin vt) &= J_0(k)\cos(ct) \\
& \quad + J_1(k)\cos(c-v)\,t \\
& \quad - J_1(k)\cos(c+v)\,t \\
& \quad + J_2(k)\cos(c-2v)\,t \\
& \quad + J_2(k)\cos(c+2v)\,t \\
& \quad + J_3(k)\cos(c-3v)\,t \\
& \quad - J_3(k)\cos(c+3v)\,t + \cdots
\end{aligned}
$$

APPENDIX C

In this Appendix the spectrum of a train of rectangular pulses of length $2(L + \Delta L)$ recurring every T seconds, will be found from the spectrum of a single pulse of this train.

For the single pulse at any frequency f,

$$g(f) \cong \frac{E}{\pi f} \sin 2\pi f(L + \Delta L). \tag{1}$$

For a series of such pulses recurring with a spacing $T = 1/c$, then the sum of spectra of the individual pulses form a Fourier series of harmonics of c. Thus

$$e(t) = A_0 + \sum_{n=1}^{\infty} A_n \cos 2\pi nct, \tag{2}$$

where A_n is the sum of an infinite number (one from each pulse) of infinitesimal terms $g(nc)$ and $g(-nc)$, shown in (1). Thus

$$A_n \cong 2\Sigma \frac{E}{\pi nc} \sin 2\pi nc(L + \Delta L) \tag{3}$$

Now to put an absolute value to the amplitudes $g(f)$ shown in equation (1), it is necessary to average them over the recurrence period of the single pulse, making them infinitesimals. However, in the train of pulses recurring every $T = 1/c$ seconds, the amplitude of A_n can be determined by averaging the terms in (1) over an interval T. Then

$$A_n = \frac{2E}{\pi ncT} \sin 2\pi nc(L + \Delta L). \tag{4}$$

When $T = 4L = 1/c$, (4) reduce to

$$A_n = \frac{2E}{\pi n} \sin \frac{n\pi}{2}\left(1 + \frac{\Delta L}{L}\right) \tag{5}$$

For the example taken in the text, when the pulse train was subdivided into 10 subsiding pulse trains, the period $T = 1/v = 10/c = 40L$. Thus in this case, the Fourier coefficients of the harmonics of v are

$$B_n = \frac{2E}{10\pi n} \sin \frac{\pi n}{2}\left(1 + \frac{\Delta L}{L}\right). \tag{5a}$$

The expression for A_n in equation (5) can be put in simpler form by using the formula for the sin of the sum of two angles. In this way, we get

$$A_n = \frac{2E}{\pi n}\left[\sin\left(\frac{\pi n}{2}\right)\cos\left(\frac{\pi n}{2}\frac{\Delta L}{L}\right) + \cos\left(\frac{\pi n}{2}\right)\sin\left(\frac{\pi n}{2}\frac{\Delta L}{L}\right)\right]. \tag{6}$$

Now, for n odd, $\sin\frac{\pi n}{2}$ alternately assumes the value ± 1 and $\cos\frac{\pi n}{2}$ vanishes, and for n even, $\cos\left(\frac{\pi n}{2}\right)$ alternately assumes the value ± 1 and $\sin\frac{\pi n}{2}$ vanishes. The A_0 term, being the *d-c* average of the pulse train, is given by

$$\frac{E/2(L + \Delta L)}{T} = \frac{E}{2}\left(1 + \frac{\Delta L}{L}\right). \tag{7}$$

If the pulse train is transformed by shifting the zero so that it alternates between $\pm E/2$ instead of O and E, the first term in equation (7) vanishes and (2) becomes, from (6) & (7),

$$\left.\begin{aligned} e(t) &= A_0 + A_1 \cos 2\pi ct \\ &\quad + A_2 \cos 2\pi\, 2ct + \cdots \\ \text{Where} \\ A_0 &= \frac{E}{2}\left(\frac{\Delta L}{L}\right) \\ A_1 &= \frac{2E}{\pi}\cos\left(\frac{\pi}{2}\frac{\Delta L}{L}\right) \\ A_2 &= \frac{2E}{2\pi}\sin\pi\left(\frac{\Delta L}{L}\right) \\ A_3 &= \frac{2E}{3\pi}\cos\frac{3\pi}{2}\left(\frac{\Delta L}{L}\right) \end{aligned}\right\} \tag{8}$$

etc.

APPENDIX D

The purpose of this section is to compute the spectrum of the carrier given by equation (8) in Appendix C as their amplitudes vary with $\frac{\Delta L}{L} = k \sin vt$.

For the *d-c* term,

$$A_0 = \frac{E}{2}\frac{\Delta L}{L} = \frac{E}{2} k \sin vt.$$

For the fundamental or c term,

$$A_1 \cos 2\pi ct = \frac{2E}{\pi} \cos\left(\frac{\pi}{2} k \sin vt\right) \cos 2\pi ct$$

Using the Bessel's expansion of cos (2 sin θ), we get,

$$A_1 \cos 2\pi ct = \frac{E}{2\pi} \begin{cases} J_0(k) \cos 2\pi c \\ +J_2(k) \cos 2\pi(c - 2v)t \\ +J_2(k) \cos 2\pi(c + 2v)t \\ + \cdots \text{ etc.} \end{cases}$$

In a similar fashion, the other terms can also be computed, giving the spectrum shown on Fig. 12, where $J_0(k)$ becomes the amplitude of c, $J_2(k)$ the amplitude of either $c + 2v$ or $c - 2v$, etc.