Angle modulation: Basic Concepts

Either the phase or frequency of the carrier wave is varied according to the message signal \( m(t) \). In this method of modulating, the amplitude of the carrier wave is maintained constant.

Modulated wave in the general form can be written as

\[
S(t) = A_c \cos \left[ \omega_c t + \phi(t) \right]
\]

where the carrier amplitude \( A_c \) is maintained constant and the angular argument \( \phi(t) \) is varied by a message signal \( m(t) \).

The mathematical form of this variation is determined by the type of angle modulation of interest. In any event, a complete oscillation occurs whenever \( \phi(t) \) changes by \( 2\pi \) radians. If \( \phi(t) \) increases monotonically with time, the average frequency inherent over an interval from \( t \) to \( t+\Delta t \), is given by

\[
\phi(t+\Delta t) - \phi(t) \over 2\pi \Delta t
\]

We define the instantaneous frequency \( f(t) \) of the angle modulated wave \( S(t) \) by

\[
f(t) = \lim_{\Delta t \to 0} \frac{1}{2\pi \Delta t} \left[ \phi(t+\Delta t) - \phi(t) \right]
\]

\[
= \frac{1}{2\pi} \frac{d\phi(t)}{dt}
\]

Thus, according to Eq. (92), we may interpret the angle-modulated wave \( S(t) \) as a rotating phasor of length \( A_c \) and angle \( \phi(t) \). The angular velocity of such a phasor is \( \frac{d\phi(t)}{dt} \), in accordance with Eq. (92). In the simple case of an unmodulated carrier, the angle \( \phi(t) \) is

\[
\phi(t) = 2\pi ft + \phi_c
\]

and the corresponding phasor rotates with a constant angular velocity equal to \( 2\pi ft \). The constant \( \phi_c \) is the value of \( \phi(t) \) at \( t=0 \).
There are an infinite number of ways in which the angle \( \theta(t) \) may be varied in some manner with the message signal. However, we will consider only two commonly used methods, phase modulation and frequency modulation.

1. Phase modulation (PM) is that form of angle modulation in which the angular argument \( \theta(t) \) is varied linearly with the message signal \( m(t) \), as shown by:

\[
\theta(t) = 2\pi ft + k_p m(t) \quad (93)
\]

The term \( 2\pi ft \) represents the angular argument of the unmodulated carrier, and the constant \( k_p \) represents the phase sensitivity of the modulator, expressed in radians per volt. This assumes that \( m(t) \) is a voltage waveform. For convenience, we have assumed in eq. (93) that the angular argument of the unmodulated carrier is zero at \( t=0 \). The phase-modulated wave \( s(t) \) is thus described in the time domain by:

\[
s(t) = A_c \cos [2\pi ft + k_p m(t)] \quad (94)
\]

2. Frequency modulation (FM) is that form of angle modulation in which the instantaneous frequency \( f_i(t) \) is varied linearly with the message signal \( m(t) \), as shown by:

\[
f_i(t) = f + k_f m(t) \quad (95)
\]

The term \( f \) represents the free or unmodulated carrier and the constant \( k_f \) represents the free sensitivity of the modulator, expressed in hertz per volt. This assumes that \( m(t) \) is a voltage waveform. Integrating eq. (95) w.r.t.time and multiplying the result by \( 2\pi \), we get:

\[
\theta(t) = 2\pi ft + 2\pi k_f \int m(t) \, dt \quad (96)
\]

where for convenience, we have assumed that the angular argument of the unmodulated carrier is zero at \( t=0 \). The frequency-modulated wave is therefore described in the time domain by:

\[
s(t) = A_c \cos [2\pi ft + 2\pi k_f \int m(t) \, dt] \quad (97)
\]
A consequence of allowing the angular argument $\theta(t)$ to become dependent on the message signal $m(t)$ as in Eq. (3) or on its integral as in Eq. (4) is that the zero crossings of a PM wave or FM wave no longer have a perfect regularity in their spacing; zero crossings refer to the instant of time at which a waveform changes from a negative to a positive value or vice versa. This is an important feature that distinguishes both PM and FM waves from an AM wave.

Another important difference is that the envelope of a PM or FM wave is constant (equal to the carrier amplitude), whereas the envelope of an AM wave is dependent on the message signal.

Comparing Eqs. (4) with (7) reveals that an FM wave may be regarded as a PM wave in which the modulating wave is $\sin(2\pi f t)$ in place of $m(t)$. This means that an FM wave can be generated by first integrating $m(t)$ and then using the result as the input to a phase modulator as in Fig. 35(a). Conversely, a PM wave can be generated by first differentiating $m(t)$ and then using the result as the input to a freq. modulator, as in Fig. 35(b). We may thus deduce all the properties of PM waves from those of FM waves and vice versa.

![Diagram of modulation process](image)

By 35(a).

Relation betw. FM and PM. (a) Generation of an FM wave by using a phase modulator. (b) Generation of a PM wave by using a frequency modulator.
7.36. (a) Sinusoidal modulating wave \( v(t) \). (b) FM wave.
(c) Derivative of \( v(t) \) w.r.t. time. (d) PM wave.

**Sinusoidal modulation:**

Consider a sinusoidal modulating wave \( v(t) \), two full cycles of which are plotted in fig. 7.36 (a). The FM wave produced by this modulating wave is plotted in fig. (b).

To determine the PM wave for \( v(t) \), we note that it is the same as the FM wave produced by \( \frac{dv(t)}{dt} \), the derivative of \( v(t) \) w.r.t. time. In fig. (c), we plot the derivative \( \frac{dv(t)}{dt} \), which consists of the original sinusoidal modulating wave shifted in phase by \( 90^\circ \). The derived PM wave is plotted in fig. (d). From the plots of fig. 7.36, we see that for sinusoidal waves, a distinction between FM and PM waves can be made only by comparing with the actual modulating wave.
**Square Modulation:**

(a) Square wave modulating wave $m(t)$. 
(b) FM wave.

(c) Derivative of $m(t)$ w.r.t. time.

(d) PM wave.

Fig. (c) is $\frac{dm(t)}{dt}$ which consists of a periodic sequence of alternating delta functions. The desired PM wave is plotted in Fig. (d).

Unlike the case of sinusoidal modulating wave, we see that for square wave modulating wave, the FM and PM waves are distinctly different from each other.

**Frequency Modulation:**

The FM wave $s(t)$ defined by $s(t) = c(t) + kAm(t)$ is a non-linear function of the modulating wave $m(t)$. Hence, FM is a non-linear modulation process. Unlike AM, the spectrum of an FM wave is not related in a simple manner to that of the modulating wave.

**Single-Tone Frequency Modulation:**

Consider a sinusoidal modulating wave defined by

$$m(t) = Am \cos(2\pi ft)$$  \(\text{(76)}\)

The instantaneous frequency of the resulting FM wave equals

$$f(t) = f_c + kAm \cos(2\pi ft)$$  \(\text{(77)}\)

where

$$k = \frac{\Delta f}{Am}$$  \(\text{(100)}\)

The quantity $\Delta f$ is called the frequency deviation, representing the maximum departure of instantaneous frequency.
of the FM wave from the carrier freq. \( f_c \). A fundamental characteristic of an FM wave is that the freq deviation \( \Delta f \) is proportional to the amplitude of the modulating wave and is independent of the modulating frequency.

Using eqn (99), the angular argument \( \Theta(t) \) of the FM wave is obtained as

\[
\Theta(t) = 2\pi \int f(t) dt = 2\pi f_c t + \frac{\Delta f}{f_m} \sin (2\pi f_m t). \tag{101}
\]

The ratio \( \frac{\Delta f}{f_m} \) is called the modulating index or the mod index of the FM wave. We denote it by \( \beta \), so that we may write

\[
\beta = \frac{\Delta f}{f_m} \tag{102}
\]

and

\[
\Theta(t) = 2\pi f_c t + \beta \sin (2\pi f_m t). \tag{103}
\]

From eqn (103), we see that in a physical sense, the parameter \( \beta \) represents the phase deviation of the FM wave; that is, the maximum departure of the angular argument \( \Theta(t) \) from the angle \( 2\pi f_c t \) of the unmodulated carrier.

Because (101): Data: \( f_m = 50 \), \( f_c = 1 \) kHz. \( \Delta f = 1 \) kHz.

For: Sensitivity = 40 Hz/V. = \( f_c \).

(a) \( \Delta f > f_c \) \( \beta \)?

\[
\beta = \frac{\Delta f}{f_m} = \frac{200}{1000} = 0.2.
\]

\[
\Theta(t) = 2\pi f_c t + \beta \sin (2\pi f_m t) = 2\pi f_c t + 0.2 \sin (2\pi f_m t).
\]

\[\text{Spectrum analysis of sinusoidal FM wave:}\]

The FM wave for sinusoidal modulating is given by

\[
\Theta(t) = A_t \cos [2\pi f_c t + \beta \sin (2\pi f_m t)]. \tag{104}
\]

Using a trigonometric identity, we may expand this relation as
\[ S(t) = A \cos(2\pi ft) \cos[\beta \sin(2\pi ft)] \\
- A \sin(2\pi ft) \sin[\beta \sin(2\pi ft)] \] 

we see the in-phase and quadrature components of the PM wave \( S(t) \) for the case of sinusoidal modulator as follows:

\[ S_x(t) = A \cos[\beta \sin(2\pi ft)] \]  \hspace{1cm} (10a)

\[ S_q(t) = A \sin[\beta \sin(2\pi ft)] \]  \hspace{1cm} (10b)

Hence, the complex envelope of the PM wave equals

\[ S(t) = S_x(t) + j S_q(t) \]

\[ = A \exp[j \beta \sin(2\pi ft)] \]  \hspace{1cm} (10c)

The complex envelope \( S(t) \) retains complete information about the modulating process. Indeed, we may readily express the PM wave \( S(t) \) in terms of the complex envelope \( S(t) \) by writing

\[ S(t) = \Re \left[ A \exp \left( j \beta \sin(2\pi ft) \right) \right] \]

\[ = \Re \left[ S(t) \exp(j 2\pi ft) \right] \]  \hspace{1cm} (10d)

From eqn (10d) we see that the complex envelope is a periodic function of time, with a fundamental frequency equal to the modulating frequency \( f_m \). We may therefore expand \( S(t) \) in the form of a complete Fourier series as follows.

\[ S(t) = \sum_{n=-\infty}^{\infty} C_n \exp(j 2\pi n ft) \]  \hspace{1cm} (11a)

where the complex Fourier coefficient \( C_n \) equals

\[ C_n = \Re \left[ A \int_{-\frac{1}{2}f_m}^{\frac{1}{2}f_m} S(t) \exp(-j 2\pi n ft) \, dt \right] \]

\[ = \Re \left[ A \int_{-\frac{1}{2}f_m}^{\frac{1}{2}f_m} \exp[j \beta \sin(2\pi ft) - j 2\pi n ft] \, dt \right] \]  \hspace{1cm} (11b)

for convenience, we define the variable

\[ \kappa = 2\pi n f \]  \hspace{1cm} (11c)

in terms of which we may rewrite eqn (11b) as

\[ C_n = \frac{A}{2\pi} \int_{-\frac{\pi}{n}}^{\frac{\pi}{n}} \exp \left[ j (\beta \sin \kappa - n \kappa) \right] \, d\kappa \]  \hspace{1cm} (11d)
The integral on the right side of eq. (13) is recognized as the \textit{n}\textsuperscript{th} order Bessel function of the first kind and argument \( \beta \). This function is commonly denoted by the symbol \( J_n(\beta) \); that is,

\[
J_n(\beta) = \frac{1}{2\pi} \int^\pi_{-\pi} \exp \left[ i(\beta \sin x - \omega t) \right] \, dx - (14).
\]

Hence, we may rewrite eq. (13) as

\[
c_n = A e^{J_n(\beta)} - (15).
\]

Substituting eq. (15) in (10) we get, in terms of the Bessel function \( J_n(\beta) \), the following expansion for the complex envelope of the FM wave:

\[
S(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \exp \left( i 2\pi n \Omega t \right) - (16).
\]

Next, substituting eq. (16) in (109), we get

\[
S(t) = A \Re \left[ \sum_{n=-\infty}^{\infty} J_n(\beta) \exp \left( i 2\pi \left( \Omega t + n \Omega m \right) t \right) \right] - (17).
\]

Order changing the order of summation and evaluating the real part of the right side of eq. (17), we get

\[
S(t) = A \sum_{n=-\infty}^{\infty} J_n(\beta) \cos \left( 2\pi \left( \Omega t + n \Omega m \right) t \right) - (18).
\]

This is the desired form for the Fourier series representation of the single tone FM wave \( S(t) \) for an arbitrary value of \( \beta \). The demodulating spectrum \( S(t) \) is obtained by taking the Fourier transform of both sides of eq. (18):

\[
S(f) = \frac{A}{2} \sum_{n=-\infty}^{\infty} J_n(\beta) \left[ \delta \left( f + \Omega \right) + \delta \left( f - \Omega \right) \right] - (19).
\]

In Fig. (19), we have plotted the Bessel function \( J_n(\beta) \) versus the mode index \( \beta \) for \( n = 0, 1, 2, 3, 4 \). These plots show that for fixed \( n \), \( J_n(\beta) \) alternates between positive and negative values for increasing \( \beta \) and \( |J_n(\beta)| \) approaches zero as \( \beta \) approaches \( \infty \).
Note also that for fixed $\beta$, we have

$$T_n(\beta) = \begin{cases} \sin(\beta), & n \text{ even} \\ \cos(\beta), & n \text{ odd} \end{cases}$$

Accordingly, we need only plot or tabulate $T_n(\beta)$ for positive values of order $n$.

From eq. (97) and (118), we deduce the following properties of FM waves:

**Property 1: Narrow band FM.**

For small values of the modulation index $\beta$ compared to one radian, the FM wave assumes a narrow band form consisting essentially of a carrier, an upper side-frequency component, and a lower side-frequency component.

This property follows from the fact that for small values of $\beta$, we have

$$F_0(\beta) \approx 1$$
$$F_1(\beta) \approx \beta$$
$$F_n(\beta) \approx 0, \quad n > 1$$

The approximations indicated in eq. (12) are strictly justified for values of modulation index defined by $\beta \leq 0.3$ rad.

Thus, substituting eq. (12) in (118), we get

$$s(t) \approx A_c \cos(2\pi f_c t) + \frac{\beta A_c}{2} \cos\left[2\pi \left(f_c + \frac{f_m}{2}\right) t\right]$$
$$- \frac{\beta A_c}{2} \cos\left[2\pi \left(f_c - \frac{f_m}{2}\right) t\right]$$

This expression shows that for small $\beta$, the FM wave $s(t)$ may be closely approximated by the sum of a carrier of amplitude $A_c$, an upper side-frequency component of amplitude $\frac{\beta A_c}{2}$, and a lower side-frequency component of amplitude $\frac{\beta A_c}{2}$ and phase shifted equal to $180^\circ$. An FM wave so characterized is said to be narrow-band.
Property 2: Wideband FM:

For large values of the mode index $\beta$ compared to one radian, the FM wave contains a carrier and an infinite number of side-fre components located symmetrically around the carrier.

This second property is a restatement of Eq. (11) with no approximations made. An FM wave thus defined is said to be wideband. Note that the amplitude of the carrier component contained in a wide-band FM wave varies with the mode index $\beta$ in accordance with $\delta_0(\beta)$.

Property 3: Constant Average Power.

The envelope of an FM wave is constant, and that the average power of such a wave dissipated in a 1-ohm resistor is also constant.

This property follows directly from the definition given in Eq. (97) for an FM wave. Specifically, the FM wave set defined in Eq. (97) has a constant envelope equal to $A_c$. Accordingly, the average power dissipated by $\delta(t)$ in a 1-ohm resistor is given by

$$P = \frac{1}{2} A_c^2 \tag{123}$$

The result may also be derived from Eq. (11). In particular, we note from the series expansion of Eq. (11) that the average power of a single tone FM wave $\delta(t)$ may be expressed in the form of a corresponding series as:

$$P = \frac{1}{2} A_c^2 \sum_{n=-\infty}^{\infty} I_n^2(\beta) \tag{124}$$

Next, we note that

$$\sum_{n=-\infty}^{\infty} I_n^2(\beta) = 1 \tag{125}$$

Thus, substituting Eq. (125) in (124), we get the result given in Eq. (123).
Example 8: Effect on the spectrum of the FM wave based on variations of the amplitude and frequency of a sinusoidal modulating wave.

\[ \beta = 1.0 \]

\[ f \]

\[ -2 \Delta f \rightarrow \Delta f \rightarrow f \]

\[ \beta = 2.0. \]

\[ f \]

\[ -2 \Delta f \rightarrow \Delta f \rightarrow f \]

\[ \beta = 5.0 \]

\[ f \]

\[ -2 \Delta f \rightarrow \Delta f \rightarrow f \]

**(c)**

From the diagram, we have normalized the spectrum with respect to the unmodulated carrier amplitude.

Keeping the modulator frequency fixed, we find that the amplitude spectrum of the resulting FM wave is as plotted in fig. 39 for \( \beta = 1.2, 5.0 \). In this diagram, we have normalized the spectrum with respect to the unmodulated carrier amplitude.

Consider the next case when the amplitude of the modulating wave is fixed; that is, the free deviation \( \Delta f \) is maintained constant and the modulator frequency is varied. Amplitude spectrum of the resulting FM wave is as plotted in fig. 40 for \( \beta = 1, 2, 5.0 \).

When \( \Delta f \) is fixed and \( \beta \) is increased, we have an increasing number of spectral lines crowding in to a fixed free interval defined by \( f - \Delta f < f < f + \Delta f \).

That is, when \( \beta \) approaches infinity, the B.W. of the FM wave approaches the limiting value of \( 2 \Delta f \).
Transmission Band width of FM wave.

FM wave contains an infinite number of side frequencies so that the b.w. required to transmit such a signal is similarly infinite in extent. In practice, we find that the FM wave is effectively limited to a finite number of significant side frequencies compatible with a specified amount of distortion. We may therefore specify the effective b.w. required for the transmission of an FM wave.

Consider first the case of an FM wave generated by a single-tone modulating wave of frequency \( f_m \). In such an FM wave, the side frequencies that are separated from the carrier frequency \( f_c \) by an amount greater than the frequency deviation \( \Delta f \) decrease rapidly toward zero, so that the b.w. always exceeds the total frequency excursion but nevertheless is limited. Specifically, for large values of the modulation index \( \beta \), the b.w. approaches and is only slightly greater than the total frequency excursion \( 2\Delta f \).
on the other hand, for small values of the mode index \( \beta \), the spectrum of the \( \text{FM} \) wave is effectively limited to the carrier frequency \( f_c \) and one pair of side frequencies at \( f_c \pm f_m \), so that the b.w. approaches \( 2f_m \). We may thus define an approximate rule for the transmission b.w. of an \( \text{FM} \) wave generated by a single tone modulating wave of frequency \( f_m \) as:

\[
B = 2f(t + 2f_m) = 2f(t + \frac{1}{15}) = (2f).
\]

This relation is called the Carson's rule.

For a more accurate assessment of the b.w. requirement of an \( \text{FM} \) wave, we may use a method based on retaining the largest number of significant side frequencies with amplitudes all greater than some selected value. A convenient choice for this value is \( 1.5 \) of the unmodulated carrier amplitude. We may thus define the 99 percent b.w. of an \( \text{FM} \) wave as the separation between the two frequencies beyond which none of the side frequencies is greater than \( 1.5 \) of the carrier amplitude obtained when the mode is removed. That is, we define the transmission b.w. as \( 2f_{\text{max}} \), where \( f_{\text{max}} \) is the maximum frequency, and \( f_{\text{max}} \) is the maximum value of the integer \( n \) that satisfies the requirement \(| \text{ln}(n) | < 0.01 \). The value of \( f_{\text{max}} \) varies with the mode index \( \beta \) and can be determined readily from tabulated values of the function \( \frac{1}{\text{ln}(n)} \).

<table>
<thead>
<tr>
<th>Mode index</th>
<th>No. of significant side frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>0.3</td>
<td>4</td>
</tr>
<tr>
<td>0.5</td>
<td>4</td>
</tr>
<tr>
<td>1.0</td>
<td>6</td>
</tr>
<tr>
<td>2.0</td>
<td>8</td>
</tr>
<tr>
<td>5.0</td>
<td>16</td>
</tr>
<tr>
<td>10.0</td>
<td>26</td>
</tr>
</tbody>
</table>
Fig. 7.41. Universal curve for evaluating the 99.1% b.w. gain frnwave.

Table 7.2 shows the total no. of significant-side frequencies (including both the upper and lower side frequencies) for different values of $\beta$, calculated on the 1% basis just explained. The maintained B.W. B calculated using this procedure can be presented in the form of a universal curve by normalizing it with respect to the free deviation $\Delta f$, and then plotting $B$ versus $\beta$. This curve is shown in Fig. 7.41, which is drawn as a best-fit through the set of points obtained by using Table 7.2. On Fig. 7.41 we note that as the mode index $\beta$ is increased, the B.W. occupied by the significant side frequencies drops toward that over which the carrier freq actually deviates. This means that small values of the mode index $\beta$ are relatively more extra vagrant in a given B.W. than are the larger values of $\beta$.

Consider next an arbitrarily modulating wave with its highest freq component denoted by $w$. The B.W. required to transmit an FM wave generated by this modulating wave is estimated by using a worst-case tone mode analysis. Specifically, we first determine the so-called deviation ratio $\delta$, defined as the ratio of the ratio of the free deviation $\Delta f$ to the highest modulating freq $w$. These conditions represent the extreme cases possible. The deviation ratio $\delta$ plays the same role for nonminimum modulators that the mode index $\beta$ plays for the case of minimum modulators.
Then, replacing $\beta$ by $\Delta$ and replacing $f$ by $w$, we use Carson's rule given by eqn. (12) of the universal wave of fig. (11) and obtain a value for the deviation $\Delta f$ of the FM wave. From a practical viewpoint, Carson's rule somewhat underestimates the deviation requirement of an FM system where $\alpha$ is used as the universal wave of fig. (11). It yields a somewhat conservative result. Thus the choice of a transmission $\Delta f$ that lies half the bounds provided by these two rules of thumb is acceptable for most practical purposes.

In North America, the maximum value of frequency deviation $\Delta f$ is fixed at 75 kHz for commercial FM broadcasting by radio. If we take the modulating frequency $w = 15$ kHz, which is typically the maximum audio frequency in FM transmission, we find that the corresponding value of $\Delta f$ required by the deviation ratio is

$$\Delta f = \frac{75}{15} = 5 \text{ kHz}$$

Using Carson's rule of eqn. (26), replacing $\beta$ by $\Delta$ and replacing $f$ by $w$, the approximate value of the deviation $\Delta f$ of the FM wave is obtained as

$$\Delta f = 2(75 + 15) = 180 \text{ kHz}.$$

On the other hand, using the curve of fig. (11) gives the deviation $\Delta f$ of the FM wave to be

$$\Delta f = 3 \times 2 = 2 \times 75 = 240 \text{ kHz}.$$

Thus Carson's rule underestimates the deviation $\Delta f$ by 25%. Compared with the result of using the curve of fig. (11).

To repeat the calculation of Eq. (9), assuming that the frequency deviation is decreased to 50 kHz.
Generation of FM Waves:

1. Indirect FM and 2. Direct FM.

In the indirect method of producing FM, the modulating wave is first used to produce a narrow-band frequency, and frequency multiplication is next used to increase the frequency deviation to the desired level. On the other hand, in the direct method of producing FM, the carrier frequency is directly varied in accordance with the incoming message signal.

**Indirect FM (Amplitude Modulation):**

Consider first the generation of a narrow-band frequency. To do this, we begin with the expression for an FM wave $s(t)$ for the general case of a modulating wave $m(t)$, which is written in the form

$$s(t) = A_1 \cos \left[ 2 \pi f_0 t + \phi(t) \right] - (27).$$

where $f_0$ is the carrier frequency and $A_1$ is the carrier amplitude. The angular argument $\phi(t)$ of $s(t)$ is related to $m(t)$ by

$$\phi(t) = 2 \pi f_0 \int_0^t m(\tau) \, d\tau - (2)$$

where $f_0$ is the frequency sensitivity of the modulator. Provided that the angle $\phi(t)$ is small compared to one radian for all $t$, we may use the following approximation:

$$\cos \left[ \phi(t) \right] \approx 1 - (29)$$

$$\sin \left[ \phi(t) \right] \approx \phi(t) - \frac{\phi(t)^3}{3!}$$

Correspondingly, we may approximate eq 27 as follows:

$$s(t) \approx A_1 \cos \left[ 2 \pi f_0 t \right] - A_1 \sin \left[ 2 \pi f_0 t \right] \phi(t)$$

$$= A_1 \cos \left[ 2 \pi f_0 t \right] - 2 \pi f_0 A_1 \sin \left[ 2 \pi f_0 t \right] \int_0^t m(\tau) \, d\tau - (13).$$

Eq (13) defines a narrow-band FM wave. Indeed, we may use this eq to set up the scheme shown in Fig 4(9) for the generation of a narrow-band FM wave: the oscillator feeds 20kHz, followed by the product modulator. Moreover, bearing in mind the relationship that exists between frequency, modulator and phase modulators, we see that the output of the 1st modulator that lies inside the
The modulated wave produced by the narrow band modulator of Fig. 419 differs from an ideal FM wave in two respects:

1. The envelope contains residual amplitude modulation and therefore varies with time.

2. For a sinusoidal modulating wave, the phase of the FM wave contains harmonic distortion in the form of third and higher-order harmonics of the modulating frequency.
However, by restricting the modulation index \( \beta \leq 0.3 \),
the effects of residual AM and harmonic PM are limited to
negligible levels.

The next step in the indirect FM method is that of
frequency multiplication. Basically, a frequency
multiplier consists of a nonlinear device (e.g.,
diode or transistor) followed by a
band-pass filter, as in fig 42(c). The nonlinear device is
assumed to be memoryless, which means that there is no energy
storage. In general, a memoryless nonlinear device is represented
by the \( \Delta f \)-\( \Delta f \) relation
\[
S_2(t) = a_1 S_1(t) + a_2 S_1^2(t) + \ldots + a_1^n S_1^n(t) \quad (12)
\]
where \( a_1, a_2, \ldots, a_n \) are constant coefficients.

Substituting eq. (13) in (12), expanding and then collecting
 terms, we find that the \( \Delta f \)-\( \Delta f \) has \( \Delta f \)-components
and \( \Delta f \)-frequency-modulated waves with carrier frequencies
\( f_1, 2f_1, \ldots, nf_1 \) and \( \Delta f \)-deviation \( \Delta f_1, 2\Delta f_1, \ldots, n\Delta f_1 \),
respectively. The value of \( \Delta f_1 \) is determined by the \( \Delta f \)-sensitivity
of the narrow-band \( \Delta f \)-modulator and the \( \Delta f \)-amplitude
of the \( \Delta f \)-modulating wave itself. We now see the motivation
for using the band-pass filter in fig 42(c). Specifically, the
filter is designed with two aims in mind:

1. To pass the \( \Delta f \)-wave centered at the carrier \( f_1 \) and
   with \( \Delta f \)-deviation \( n\Delta f_1 \).

2. To suppress all other \( \Delta f \)-spells.

Thus, connecting the NB FM and the \( \Delta f \)-multiplier as
depicted in fig(c), we may generate a wideband FM wave \( \Delta f \)
with carrier \( f_c \) of \( f_1 \) and \( \Delta f \)-deviation \( \Delta f = n\Delta f_1 \),
as desired. Specifically, we may write
\[ s(t) = A e^{\cos \left[ \frac{2\pi}{n_1} t + 2\pi f_0 \int_0^t e^{\cos \lambda \theta} \, dt \right]} \] 

where \( \lambda = n_1 \) \[ n_1 - 1.34. \]

In other words, the wide band FM of fig. 42(c) has a freedom sensitivity 4'14. Note that the 118 FM of fig. 42(b), where \( n \) is the freq. multiplication ratio. In fig. 42(c) we show a crystal-controlled oscillator as the source of carrier; this is done for freq. stability.
Hans Neumann on frequency modulated waves:

FM wave contains an infinite number of side frequencies and consequently the bandwidth required for transmitting such signals is also infinite in extent.

Most of the power of the FM wave resides in a finite number of sidebands, decided by the magnitude of \( \ln(\beta) \) and hence \( \beta \).

For \( \beta \) smaller compared to one radian, only \( \ln(\beta) \) and \( \ln(\beta^2) \) are significant. Hence, the significant side frequencies are \( f_m + \beta \) and \( f_m - \beta \). Thus, the bandwidth due to transmission \( B_m \) is \( \ln(\beta) + (f + \beta) - (f - \beta) = 2\beta \).

For large values of \( \beta \), that is, \( \beta \) large compared to one radian, \( \ln(\beta) \approx 0 \) for \( n \beta \) only. This means that only \( n \beta \) frequency components are significant. In other words, the significant components range from \( f - n\beta \) to \( f + n\beta \), giving
\[
B_n = (f + n\beta) - (f - n\beta) = 2n\beta.
\]

Carter developed generalised the \( B \) formula for an FM wave. According to him, the approximate formula for computing the \( B \) of an FM wave generated by a single tone
The modulating signal of b.f.m. is
\[ B_T = 2(1 + \beta)B_m \]

Substituting \( f_m \beta = 4f \), we get:
\[ B_T = 2B_m \beta + 2B_m = 24f + 2B_m. \]
\[ B_T = 24f(1 + \frac{1}{\beta}) \]

A more accurate value of the B.W. of F.M. wave can be obtained by using the universal curve for evaluating F.M. B.W.

The effective B.W. is the separation b/w the two extreme significant side frequencies on either side of the carrier.

A side frequency is considered to be significant if its amplitude is at least one percent of the unmodulated carrier amplitude.

On this basis, the bandwidth B.W. of the carrier wave is defined as the separation b/w two side frequencies beyond which none of the side frequencies has an amplitude greater than one percent of the unmodulated carrier amplitude.
Suppose $n_{\text{max}}$ is the largest value of $n$ integer in such that $|F_n(\beta)| > 0.01$. Then, we define the equivalent B.W. as

$$\beta_T = 2n_{\text{max}} f_m.$$  \(\text{(1)}\)

where $f_m$ is the modulating signal freq.

The value of $n_{\text{max}}$ depends on the value of $\beta$ and $n_{\text{max}} \approx \beta$.

The bandwidth B.W. as computed by eq(1) can be presented in the form of universal curve, a plot of normalized B.W., $\frac{\beta_T}{\beta_f}$ vs. mode index $\beta$.

Universal curve for evaluating the $1\%$ B.W. of an Airwave
At $\beta$ unity, $B_f = 24f$.

Generally, $m \mu (f)$ is an arbitrary signal with the highest significant frequency $W$ Hz. Called $B_w$ or $m \mu (f)$.

If $A_f$ is the peak deviation, then the ratio of $A_f$ to $W$ is called deviation ratio. That is,

$$D = \frac{A_f}{W}$$

$A_f$ depends on the peak amplitude of $m \mu (f)$.

$D$ plays the same role for arbitrary $\beta$ modulation as $\beta$ plays for sinusoidal waves. Replacing $\beta$ by $D$ and $f$ by $W$, we get in eq. $B_f$.

$$B_f = 2W (1 + D)$$

The above relation is called Carson's formula.

A comparison of the values of $B_w$ obtained from Carson's rule and universal curve reveals that the former scheme slightly underestimate the $B_w$ while the latter technique yields somewhat conservative results.
1. Unlike AM where there are only two sidebands, FM wave for continuous waves contains an infinite set of sidebands located symmetrically on either side of the carrier at freq. separations of \( \pm f_m, \pm 2f_m, \ldots \) etc.

2. The diffr. sidebands are proportional to Bessel coefficients \( J_0(\beta) \) for small values of \( \beta \), only \( J_0(\beta) \) and \( J_1(\beta) \) have significant values, with the result that FM wave is effectively composed of carrier and a single pair of side frequencies. This situation corresponds to the special case of narrow band F.M. \( \beta \ll 1 \) from the table \( J_0(\beta) = 1, J_1(\beta) = \sqrt{\frac{2}{\pi}} \beta \) so that the carrier component amplitude \( = A_c \) and the first set of side frequencies \( \pm f_m \) has an amplitude \( \beta A_c \).

3. In F.M. unlike A.M., the amplitude of the carrier component does not remain constant, but varies with \( \beta \) according to \( J_0(\beta) \).

4. In A.M., increasing the depth of modulation increases the sideband power, thereby increasing the total transmitted power. In F.M., the total transmitted power remains constant as the envelope is of constant amplitude, so that the mean power is \( P = \frac{1}{2} A_c^2 \). When the carrier is modulated to generate F.M., the power in the side frequencies appears at the expense of power originally in the carrier, thereby making the amplitude of the carrier dependent on \( \beta \).
Power distribution in F.M.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>carrier-frequency ( f_c )</td>
<td>( A_c F_0(\beta) )</td>
</tr>
<tr>
<td>first sideband ( f_c + \Delta f )</td>
<td>( A_c F_1(\beta) )</td>
</tr>
<tr>
<td>second ( f_c + 2\Delta f )</td>
<td>( A_c F_2(\beta) )</td>
</tr>
<tr>
<td>( \cdots )</td>
<td>( A_c F_n(\beta) )</td>
</tr>
</tbody>
</table>

For a fixed load resistance \( R \), the total power \( P_T \) is given by:

\[
P_T = \frac{\left( A_c F_0(\beta) \right)^2}{R} + \frac{2\left[ A_c F_1(\beta) \right]^2}{R} + \frac{2\left[ A_c F_2(\beta) \right]^2}{R} + \cdots
\]

\[
P_T = \frac{A_c^2}{2R} F_0^2(\beta) + 2 \frac{A_c^2}{2R} F_1^2(\beta) + 2 \frac{A_c^2}{2R} F_2^2(\beta) + \cdots
\]

\[
\therefore P_T = \frac{A_c^2}{2R}
\]

\[
P_T = P_c F_0^2(\beta) + 2P_c F_1^2(\beta) + 2P_c F_2^2(\beta) + \cdots
\]

\[
= P_c \left[ F_0^2(\beta) + 2F_1^2(\beta) + 2F_2^2(\beta) + \cdots + F_n^2(\beta) + \cdots \right]
\]

One of the properties of the Bessel function is that the sum

\[
\sum F_0^2(\beta) + \sum F_1^2(\beta) + \sum F_2^2(\beta) + \cdots = 1
\]

This shows that the total avg. power is equal to the unmodulated carrier power.
Demodulation of FM signals

Frequency demodulation is the process that enables us to recover the original modulating signal from a frequency-modulated signal. The objective is to produce a transfer characteristic that is the inverse of that of the FM, which can be realized directly or indirectly.

Direct Method:

Frequency discriminators, whose instantaneous amplitude is directly proportional to the instantaneous frequency of the input FM signal, freq. discriminators consisting of a resistor followed by an envelope detector.

(a) Freq. response ideal slope cat.

(b) Slope cat's response

(c) Freq. response of complex I/OF equivalent to ideal slope cat complementary to that of part (a)
An ideal slope cat is characterized by a transfer function that is purely imaginary, varying linearly with frequency inside a prescribed frequency interval.

Consider the transfer function plotted in fig. (a) which is defined by

$$H(f) = \begin{cases} \frac{2\pi a (f - a + \frac{B_r}{2})}{B_r}, & a - \frac{B_r}{2} \leq f \leq a + \frac{B_r}{2} \\ \frac{2\pi a (f + a - \frac{B_r}{2})}{B_r}, & -a - \frac{B_r}{2} \leq f \leq -a + \frac{B_r}{2} \\ 0, & \text{elsewhere} \end{cases}$$

where $a$ is a constant parameter.

We wish to evaluate the response of the slope cat denoted by $S_{1}(t)$, which is produced by an FM signal $s(t)$ of carrier frequency $a$ and bandwidth $B_r$. It is assumed that spectrum of the signal is essentially zero outside the base interval $[a - B_{1/2}, a + B_{1/2}]$.

Replace the slope cat with an equivalent LPF and driving with $s(t)$ the complex envelope of the FM signal $s(t)$. Let $h_{1}(t)$ denote the complex transfer function of the slope cat defined by $H_{1}(a)$. This complex transfer function is related to $H_{1}(f)$ by

$$H_{1}(f-k) = 2\pi a (f-k), \quad f > 0$$

---
Hence why cases 1 and 2 we get:

\[ \tilde{f}_2(t) = \begin{cases} \frac{1}{2} \pi m(t) (1 + B_2), & -B_2 \leq f < B_2 \\ 0, & \text{elsewhere} \end{cases} \]  

- (3).

which is plotted in fig. (b).

The incoming FM signal \( S(t) \) is defined by

\[ S(t) = A_c \cos \left( 2\pi f_c t + 2\pi k \int_0^t m(\tau) d\tau \right) - (4). \]

Given that the carrier frequency is high compared to the deviation \( B_m \) of the FM signal \( S(t) \), the complex envelope \( \tilde{S}(t) \) is

\[ \tilde{S}(t) = A_c \exp \left[ i 2\pi f_c t \int_0^t m(\tau) d\tau \right] - (5). \]

Let \( \tilde{S}(t) \) denote the complex envelope of the response of the filter in definition by fig. (c) due to \( S(t) \).

P.

If \( \tilde{S}(t) \) is

\[ \tilde{S}(t) = \begin{cases} \frac{1}{2} \pi m(t) \tilde{S}(t) \\ \right. \\ \left. \begin{cases} \frac{1}{2} \pi m(t) (1 + B_2), & -B_2 \leq f < B_2 \\ 0, & \text{elsewhere} \end{cases} \end{cases} \]  

- (6).

where \( \tilde{S}(t) \) is the P.T. of \( \tilde{S}(t) \).

Since multiplication of the P.T. of a signal by the factor \( \frac{1}{2} \pi f \) is equivalent to differentiating the signal in the time domain, we deduce from eq. (6) that

\[ \tilde{S}_1(t) = \pi \left[ \frac{d \tilde{S}(t)}{dt} + \right. \\ \left. \frac{i}{2} \pi B_2 \tilde{S}(t) \right] - (7). \]

Substituting \( \tilde{S}(t) \) in (7), we get

\[ \tilde{S}_1(t) = \frac{1}{2} \pi B_2 A_c \left[ 1 + \frac{2B_2}{\pi f} \right] \exp \left[ i 2\pi f \int_0^t m(\tau) d\tau \right] - (8). \]
The desired response of the slope circuit is therefore

\[ S(t) = \text{Re} \left[ S_i(t) \exp (j2\pi f t) \right] \]

\[ = \hat{V}_i B r a A c \left[ 1 + \frac{2k^2 m(t)}{B^2} \right] \cos \left[ 2\pi f_0 t + 2\pi f y \int_{t_0}^{t} S_i(t) dt + \phi \right] \]

(3)

The signal \( S(t) \) is a hybrid-modulated signal, in which both amplitude and frequency of the carrier wave vary with the message signal \( m(t) \). However, provided that we choose \( \left| \frac{2k^2 m(t)}{B^2} \right| \ll 1 \), for all \( t \)

then we may use an envelope detector to recover the amplitude variations and thus, except for a bias term, obtain the original message signal. The resulting envelope detector output is therefore

\[ \tilde{S}(t) = \hat{V}_i B r a A c \left[ 1 + \frac{2k^2 m(t)}{B^2} \right] \]

(10)

The bias term \( \hat{V}_i B r a A c \) in the L.H.S. of (10) is proportional to the slope \( \frac{d}{dt} \) of the transfer function of the slope circuit. This suggests that the bias may be removed by subtracting from the envelope detector output \( \tilde{S}(t) \) the output of a second envelope detector

preceded by the complementary slope circuit with the transfer function \( H_2(t) \) plotted in fig. (1C)

that is, the respective complex transfer functions of the two slope circuits are related by

\[ H_2(t) = \overline{H_1(-t)} \]

(11)
Let $S_{2}(t)$ denote the response of the complementary slope circuit produced by the incoming FM signal $S(t)$. Then, the envelope of $S_{2}(t)$ is

$$
S_{2}^{0}(t) = \pi B T A_{c} \left[ 1 - \frac{2 R_{T}}{B T} \right] - (12)
$$

where $S_{2}^{0}(t)$ is the complex envelope of the signal $S_{2}(t)$.

The difference between the two envelopes in eqs (13) & (12) is

$$
S_{0}(t) = (S_{1}^{0}(t) - S_{2}^{0}(t))
$$

$$
= 4\pi I T A_{c} m(t)
$$

which is free from bias.

We may thus model the ideal frequency discriminator as a pair of slope circuits with their complex transfer functions related by eq. (11), followed by envelope detectors and finally a summer, as in fig 2(a). This scheme is called a balanced freq. discriminator.

![Diagram of frequency discriminator](image)

B2. CRT diagram
(C) Frequency response.

The linearity of the useful portion of the total response in Fig. 2(c), centered at 0, is determined by the separation of the two relevant frequencies. The frequency separation of 2B gives satisfactory results, where 2B is the 3-dB
B.W. of either filter. However, there will be distortion in the
clip of this frequency discriminator due to the following
factors:

1. The amplitude of the dip for signal S1D is not exactly
zero for frequencies outside the range \(10^{-6} \leq f \leq 10 \text{kHz}\).

2. The linear filter steps are not strictly band limited, and
so some distortion is introduced by the low-pass RC 
filter following the diodes in the envelope detector.

3. The linear filter characteristics are not linear over the
whole frequency band of the input for signal S1D.

By proper design, it is possible to maintain the
maximum distortion produced by these factors within tolerable
limits.
FM stereo multiplexing:

Stereo multiplexing is a form of FM designed to transmit two separate signals via the same carrier. It is widely used in FM radio broadcasting to send two different elements of a program (a vocalist and an accompanist) so as to give a spatial dimension to its perception by a listener at the receiving end.

The specification of standards for FM stereo transmission is influenced by two factors:

1. The transmitted signal has to operate within the allocated FM broadcast channels.
2. It has to be compatible with monophonic radio receivers.

The first requirement sets the permissible frequency parameters, including frequency deviation. The second requirement controls the way in which the transmitted signal is configured.

Fig. 8 shows the block diagram of the multiplexing system used in an FM stereo transmitter. Let \( m_L(t) \) and \( m_R(t) \) denote the signal picked up by left-hand and right-hand microphones at the transmitting end of the system. They are applied to a single-multiplexer that generates the sum signal and difference signal.
The sum signal is left unprocessed with baseband form. It is available for monophonic reception.

The difference signal and a 38-kHz subcarrier (derived from a 19-kHz crystal oscillator by free doubling) are applied to a product modulator, thereby producing a DSB-SC modulated wave.

In addition to the sum signal and this DSB-SC modulated wave, the multiplexed signal m(t) also included a 19-kHz pilot to provide a reference for in-coherent detection of the difference signal at the stereo receiver; thus the multiplexed signal is described by

\[ m(t) = [m_1(t) + m_2(t)] + [m_1(t) - m_2(t)] \cos(2\pi f_c t) + k \cos(2\pi f_c t) \]

\[ f_c = 19 \text{ kHz}, \quad k = \text{Amplitude of the pilot tone.} \]

The multiplexed signal m(t) then frequency-modulates the main carrier to produce the transmitted signal.

![Diagram](image.png)

(a) Multiplexer in Braunmiller for stereo.
(b) **Demodulate the receiver of FM stereo.**

At a stereo receiver, the multiplexed signal \( m(t) \) is recovered by frequency demodulating the incoming waveform.

Then \( m(t) \) is applied to the demultiplexing system (shown in fig. (b)).

The individual components of the multiplexed signal \( m(t) \) are separated by the use of three appropriate filters.

The recovered pilot is frequency doubled to produce the desired 28-KHz subcarrier. The availability of this subcarrier enables the coherent detection of the DSB-SC modulated wave, thereby recovering the difference signal, \( m_y(t) \).

The basebands low-pass filtered in the top path of fig. (b) is designed to pass the sum signal, \( m_y(t) + m(t) \).

Finally, the simple matrix reconstitutes the left-hand signal \( m_y(t) \) and right-hand signal \( m(t) \) and applies them to their respective speakers.
Phase-Locked Loop.

The PLL is a negative F. B. System, the operation of which is closely linked to frequency modulation. It can be used for synchronization, frequency division/multiplication, frequency modulation and indirect frequency demodulation.

![ PLL Diagram ]

PLL consists of three major components: a multiplier, a loop filter, and a VCO combined together in the form of an F. B. loop as in fig. The VCO is a sinusoidal generator whose frequency is determined by a voltage applied to it from an external source. In effect, any FM may serve as VCO.

Initially we assume that we have adjusted the VCO so that when the control voltage is zero, two conditions are satisfied:

1. The frequency of the VCO is precisely set at the unmodulated carrier frequency f_c.
2. The VCO output has a 90-degree phase shift with the unmodulated carrier wave.

Input to the PLL is an FM Signal defined by

\[ s(t) = A_c \sin \left( 2\pi f_c t + \phi(t) \right) \]

\[ A_c \] - carrier amplitude.
with a modulating signal \( m(t) \), the angle \( \phi_1(t) \) is related to \( m(t) \) by the integral

\[
\phi_1(t) = 2\pi f_0 \int m(t) \, dt - 1.
\]

\( f_1 \) is the sensitivity of the FM.

let VCO output in the PLL be defined by

\[
v(t) = A_0 \cos \left( 2\pi f_0 t + \phi_2(t) \right) - 2.
\]

\( A_0 \) is the amplitude.

with a control voltage \( v(t) \) applied to the VCO input, the angle \( \phi_2(t) \) is related to \( v(t) \) by the integral

\[
\phi_2(t) = 2\pi f_0 \int v(t) \, dt - 3.
\]

\( A_0 \) is the sensitivity of the VCO, \( v(t) \) is the input.

To develop an understanding of the PLL, it is desirable to have a model of the loop. We first develop a nonlinear model, which is subsequently linearized to simplify the analysis.

**Non-linear model of PLL:**

According to fig, \( s(t) \), \( \delta(t) \) and \( v(t) \) or \( \phi(t) \) are applied to the multiplier, producing two components:

1. A high-\( f_0 \) component, represented by the double-\( f_0 \) term

\[
K_m A_v s(t) \sin \left( 4\pi f_0 t + \phi_1(t) + \phi_2(t) \right)
\]

2. A low-\( f_0 \) component, represented by the \( 3f_0 \)-term

\[
K_m A_v s(t) \sin \left( \phi_1(t) - \phi_2(t) \right)
\]

\( K_m \) is the multiplier gain, measured in volt^{-1}.
Therefore, discarding the high-frequency component, the SP to the loop filter is reduced to
\[ e_{LT}(t) = k_m A_d u \sin \left[ \phi_e(t) \right] \quad -C \]
\( \phi_e(t) \) is the phase error defined by:
\[ \phi_e(t) = \phi_1(t) - \phi_2(t) \]
\[ = \phi_1(t) - 2\pi k_e \int_0^t e_{LT}(\tau) d\tau \quad -D \]

The loop filter operates on the input \( e_{LT}(t) \) to produce an output \( v_{LT}(t) \) defined by the convolution integral:
\[ v_{LT}(t) = \int_{-\infty}^{0} e_{LT}(\tau) h(t-\tau) d\tau \quad -E \]
where \( h(t) \) is the impulse response of the loop filter.

Using eqs. (4.62) to (4.64) to relate \( \phi_e(t) \) and \( \phi_1(t) \), we obtain the following non-linear integral-differential equation as the discretized dynamic behavior of the PLL:
\[ \frac{d\phi_e(t)}{dt} = \frac{d\phi_1(t)}{dt} - 2\pi k_V \int_0^t \sin \left[ \phi_e(\tau) \right] h(t-\tau) d\tau \quad -F \]

where \( k_V \) is the loop gain parameter defined by:
\[ k_V = k_m k_u A_d u \quad -G \]
\( k_u \), and \( k_m \rightarrow \text{volts}, \quad k_u \rightarrow \text{volt}^{-1}, \quad k_m \rightarrow \text{Hz}\cdot\text{volt}. \]
Hence it follows from eqn. 9 that to has the

dimensions of frequency.

Fig. 10 suggests the model shown in fig. (2) for PLL.

\[ \phi(t) \rightarrow \phi(t) \rightarrow \text{Sim.} \rightarrow k(t) \rightarrow v(t) \]

\[ \text{Fig. 2: Non-linear model of PLL} \]

In this model we have also made the relationship between \( v(t) \) and \( \phi(t) \) as represented by eqns (5) and (7).

Model resembles the B.D.S. type. The multiplier at the input of the PLL is replaced by a subtractor and a sinusoidal nonlinearity and the VCO by an integrator.

The sinusoidal nonlinearity in the model of fig. 2 greatly increases the difficulty of analyzing the behavior of the PLL.

It would be helpful to linearize this model to simplify the analysis, yet give a good approximate description of the loop's behavior in certain modes of operation.
Linear model of PLL

When the phase error \( \phi_{\text{el}} \) is zero, the PLL is said to be in phase-lock. Then \( \phi_{\text{el}} \) is at all times small compared with one radian, we may use the approximation

\[
\sin [\phi_{\text{el}}(t)] \approx \phi_{\text{el}}(t) - \frac{\phi_{\text{el}}(t)^3}{3!}.
\]

In this case, the loop is said to be near phase-lock, and the nonlinearities nonlinearity of fig 2 may be disregarded. Thus, we may represent the PLL by the linearized model shown in Fig 4(a).

![Fig. 4(a) Linearized model.](image)

\[\phi_{\text{el}}(t) \approx \frac{\partial}{\partial t} \phi_{\text{el}}(t) \rightarrow V_{\text{in}}(t)\]

(b) Simplified model (when the loop gain is very large compared to unity).

According to this model (1.49), the phase error \( \phi_{\text{el}}(t) \) is related to the input phase \( \phi(t) \) by the linear integro-differential equation

\[
\frac{d\phi_{\text{el}}(t)}{dt} + \frac{2\pi f_0}{V_{\text{in}}} \int_{-\infty}^{t} \phi(t) \Delta(t-s) dt = \frac{d\phi(t)}{dt}.
\]
Transforming Eq. (2) in the freq. domain and solving for \( \phi(t) \), the F.T. \( \Phi_c(t) \), in Eq. (1), we get:

\[
\Phi_c(t) = \frac{1}{1 + L(t)} \Phi_1(t) \quad - (8).
\]

The function \( L(t) \) in Eq. (3) is defined by

\[
L(t) = \frac{k_0}{k_0} H(t) \quad - (4).
\]

H(t) is transfer function of the loop-filter.

The quantity \( \frac{dH}{df} \) is called open-loop transfer function of the PLL.

If \( L(t) \) is very large compared with unity, then from Eq. (8) we find that \( \Phi_c(t) \) approaches zero.

That is, the phase of the VCO becomes asymptotically equal to the phase of the incoming signal.

Under this condition, phase-lock is established and the objective of the PLL is thereby satisfied.

From Fig. 3(a), we see that \( V(t) \), the F.T. of \( PLL \) of \( V(t) \) is related to \( \Phi_c(t) \) by

\[
V(t) = \frac{k_0}{k_0} H(t) \Phi_c(t) \quad - (5).
\]

\[
V(t) = \frac{k_0}{k_0} \frac{dH}{df} \Phi_c(t) \quad - (6)
\]

Substituting Eq. (5) in (6):

\[
V(t) = \frac{(dH/k_0) L(t)}{1 + L(t)} \Phi_1(t) \quad - (7).
\]
When we make \(|L(f)| > 1\) for the frequency band of interest, we may approximate \(e^{j\Phi(t)}\) as
\[
V(t) \approx \frac{d}{du} \Phi(u)
\]

Corresponding time domain equation is
\[
V(t) = \frac{1}{2\pi f_0} \frac{d}{dt} \Phi(t)
\]

Thus, provided that the magnitude of the open-loop transfer function \(L(f)\) is very large for all frequencies of interest, \(V(t)\) may be modeled as a differentiator with its output scaled by the factor \(1/2\pi f_0\), as in Fig. 3(b).

The simplified model of Fig. 3(b) provides an indirect method of using PLL as a frequency demodulator. When an input is an FSK signal, the angle \(\Phi(t)\) is related to the message signal \(m(t)\) as in Eq. (2)
\[
\Phi(t) = 2\pi f^* \int m(t) dt
\]
Substituting Eq. (2) in Eq. (4), we find that the resulting output signal \(v(t)\) of the PLL is approximately
\[
V(t) \approx \frac{1}{2\pi} \frac{d}{dt} m(t)
\]

Eq. (6) states that when the loop operates in its phase locked mode, the output \(v(t)\) of the PLL is approximately the same, except for the scale factor \(1/2\pi f_0\), as the original message signal \(m(t)\), frequency demodulation of the incoming FSK signal \(s(t)\) is thereby accomplished.
System Noise and Calculations

The inevitable and presence of noise in a communication system causes the reliable transmission of electrical signals through the system to be endangered. It is therefore important to know how noise arises in the system.

There are many potential sources of noise. The sources of noise may be external to the system (e.g., atmospheric noise, galactic noise, man-made noise) or internal to the system.

1. Owing to the spontaneous fluctuations of noise current or voltage in electrical circuits.

This type of noise, in one way or another, is present in every common system and represents a fundamental limitation on the transmission detection of signals.

Electrical noise

In an electrical circuit, noise is generated owing to various physical phenomena.

1. Thermal noise — produced by the random motion of electrons in conducting media and shot noise produced by random fluctuations of current flow in electronic devices. These are the two most common examples of spontaneous fluctuations noise encountered in
Besides thermal noise and shot noise, transistors exhibit a low-frequency noise phenomenon known as flicker noise.

The mean-square value of flicker noise is inversely proportional to frequency; hence, it is referred to as "one-over-f" noise.

Another type of noise encountered in semiconductor devices is burst noise, whose mean-square value falls off as $1/f^2$.

Flicker noise and burst noise are both non-white, with their degrading effects being observed at low frequencies. They can be ignored above a few kHz.

Thermal noise and shot noise are both white for all practical purposes; hence their degrading influence on the operation of a system extends right across the complete frequency band of interest.

**Thermal Noise:**

Thermal noise is an ubiquitous source of noise that arises from thermally induced motion of electrons in conducting media. On a conductor there is a large number of
"Free electrons" and an equally large number of ions bound by strong molecular forces. The ions vibrate randomly about their normal positions. The free electrons move along randomly oriented paths, owing to continual collisions with the vibrating ions. The net effect of this random motion is an electric current that is likewise random.

However, the mean value of the current is zero since, on the average, there are as many electrons moving in one direction as there are in another.

The mean-square value of the thermal noise voltage \( V_{th} \) appearing across the terminals of a rectifier, measured in a bandwidth of \( \Delta f \) hertz, is given by all practical purposes to

\[
E[V_{th}^2] = 4kT R \Delta f \text{ Volts}^2
\]

\( k \to \) Boltzmann's constant equals to \( 1.38 \times 10^{-23} \) Joule per degree Kelvin,

\( T \to \) Absolute temperature in K,

\( R \to \) Resistance in ohms.

\[ R \equiv \begin{array}{c}
-1 \\
E[V_{th}^2] \\
E[V_{th}^2]
\end{array} \]

Models of noisy rectifiers

(a) Thevenin eq. 6.15 
(b) Norton eq. 6.15-

The mean-square value of the noise current generated by

\[
E[I_{th}^2] = \frac{1}{R^2} E[V_{th}^2]
\]

\( = 4kT_0 g_0 \text{ amp}^2 \)

\[ g = 4kT_0 R \]
Number of electrons in a resistor is very large and their random motions inside the resistor are statistically independent of each other, the central limit theorem indicates that thermal noise is Gaussian distributed with zero mean.

Noise calculations involve the transfer of power and so we find that the use of maximum power transfer theorem is applicable to such calculations.

This theorem states that the maximum possible power is transferred from a source of internal resistance \( R_s \) to a load of resistance \( R_L \) when \( R_L = R_s \).

Under this matched condition, the power produced by the source is divided equally between the internal resistance of the source and the load resistance, and the power delivered to the load is referred to as the available power.

Applying maximum power transfer theorem to the Thevenin equivalent circuit of (a) and Norton equivalent circuit of (b), a noisy resistor produces an available noise power equal to \( kT \Delta f \) of Watts.
White Noise:

In the noise analysis of communication systems, it is customary to work with an idealized form of noise called white noise, the PSD of which is independent of the operating frequency.

White light contains equal amounts of all frequencies within the visible band of electromagnetic radiation.

\[
S(f) = \frac{N_0}{2}
\]  \hspace{1cm} (5.121)

\[
N_0 \Rightarrow \text{Watt/Hz}
\]

\[
R_s(\tau) = \frac{N_0}{2} \delta(\tau)
\]

\[
R_s(\tau)
\]

Ff519(a) PSD of white noise.

The parameter \(N_0\) is usually referred to the input stage of the receiver of a communication system, expressed as

\[
N_0 = k_Te
\]  \hspace{1cm} (5.122)

\(k_T\) = Boltzmann constant. \(e\) = equivalent noise current of the Rx.

\(N_0\) depends only on the parameters of the system.

Since the autocorrelation function \(R_s(\tau)\) is the inverse Fourier transform of the PSD,

\[
R_s(\tau) = \frac{N_0}{2} \delta(\tau)
\]  \hspace{1cm} (5.122)
The ACF of white noise is the inverse Fourier transform of the PSD shown in Fig. 2.

\[ R_w(t) = \int_{-B}^{B} \frac{N_0}{2} \exp \left( \frac{-j2\pi ft}{B} \right) df = \frac{N_0B}{2} \operatorname{sinc} \left( \frac{2Bt}{2} \right) \]

\[ R_w(t) \text{ has its maximum value of } \frac{N_0B}{2} \text{ at the origin and its value passes through zero at } t = \pm k\frac{B}{2}, \text{ where } k = 1, 2, 3, \ldots \]
**RC Low-pass Filtered White Noise**

Consider a white Gaussian noise \( W(t) \) of zero mean and PSD \( \text{No}/2 \) applied to a low-pass RC filter, as in Fig 5.21(a).

The transfer function of the filter is

\[
H(j\omega) = \frac{1}{1 + j\omega \tau RC}
\]

The PSD of the noise \( W(t) \) appearing at the low-pass RC filter output is therefore by 5.21(b)

\[
\text{PsD}(f) = \frac{\text{No}/2}{1 + (2\pi f \tau RC)^2}
\]

For \( \tau = \frac{1}{RC} \),

\[
\exp(-a \tau) = \frac{2a}{a^2 + (2\pi f \tau)^2}
\]  -(5.126)

"\( a \) is constant ."

Setting \( a = 1/RC \),

ACF of filtered white noise \( W(t) \) is

\[
\text{RAF}(f) = \frac{\text{No}/2}{4\pi RC} \exp \left(-\frac{Rf}{RC}\right) \]  -(5.127)

which is plotted in Fig 5.21(c).
A convenient measure of the noise performance of a linear two-port device is furnished by the noise figure. Consider a linear two-port device connected to an input signal source of an impedance \( Z_{in} = R_{in} + jX_{in} \) at the input, at the frequency \( f \) in Hz.

The noise voltage \( V_{1n} \) represents the thermal noise associated with the internal noise \( R_{in} \) of the source. The output noise of the device is made up of two contributions, one due to the source and the other due to the device itself.

We define the available output noise power in a band of width \( B \) centered at frequency \( f \) as the maximum average noise power in that band obtainable at the output of the device.

The maximum noise power that the two-port device can deliver to an external load is obtained when the load impedance is the complex conjugate of the output impedance of the device, that is, when its resistance is
Noise figure of the two-port device is defined as the ratio of the total available output noise power (due to the device and the source) per unit bandwidth to the portion of it due solely to the source.

Let the spectral density of the total available noise power at the device output be $S_{N0}(f)$, and the spectral density of the available noise power due to the source at the device output be $S_{N0}(f)$. Let $g(f)$ denote the available power gain of the two-port device, defined as the ratio of the available signal power at the output of the device to the available signal power of the source when the signal is a sinusoidal wave of frequency $f$. Noise figure of the device is

$$F(f) = \frac{S_{N0}(f)}{g(f) S_{N0}(f)}$$

If the device were noiseless, $S_{N0}(f) = g(f) S_{N0}(f)$, and the noise figure would then be unity.

In a physical device, however, $S_{N0}(f)$ is larger than $g(f) S_{N0}(f)$, so that the noise figure is always larger than unity. The noise figure is commonly expressed in decibels, that is, $10 \log F(f)$. 

The noise figure may also be expressed in an alternative form. Let $P_s(t)$ denote the available signal power from the source, which is the maximum average signal power that can be obtained.

For the case of a source voltage $V_s(t) = V_0 \cos(2\pi ft)$, the available signal power is obtained when the load is connected to the source $Z_{0s}(t) = R(t) - jX(t)$, yielding the value

$$P_s(t) = \left(\frac{V_0}{2R(t)}\right)^2 R(t) = \frac{V_0^2}{4R(t)} \quad - (2).$$

The available signal power at the output of the device is therefore,

$$P_0(t) = g(t) P_s(t) \quad - (3).$$

Multiplying both the $V_s(t)$ and $P_s(t)$ of the right side of eqn 1 by $P_s(t) BN$, we have,

$$P(t) = \frac{P_s(t) \cdot S_{Rn}(t) BN}{g(t) R_H(t) \cdot S_{Rn}(t) BN}.$$  

$$= \frac{P_s(t) \cdot S_{Rn}(t) \cdot BN}{P_0(t) \cdot S_{Rs}(t) \cdot BN}.$$  

$$= \frac{P_s(t)}{P_0(t)} \quad - (4).$$

Where $P_s(t) = \frac{P_s(t)}{S_{Rs}(t) \cdot BN} \quad - (5)$ and $P_0(t) = \frac{P_0(t)}{S_{Rn}(t) \cdot BN} \quad - (6).$

$P_s(t)$ is available signal-to-noise ratio of the source.

$P_0(t)$ is available signal-to-noise ratio at the device output.

Both measured in a narrow band of width BN centered at $f$.  