2

Distortion in Amplifiers

2.1 Introduction

In any discussion of amplifier linearity and of methods of ensuring that linearity is maintained to a high degree, it is necessary to examine the nature of amplifier distortion in all of its various forms and to establish techniques for determining its level accurately and simply.

Audio amplifier distortion has been of concern for very many years and considerable design effort has resulted in its virtual elimination from modern high-fidelity amplifiers. The feedback techniques conventionally used at audio frequencies are, however, not generally applicable to many radio-frequency amplifiers due to problems of stability at high bandwidths and of cost for high gains in RF stages. As a result many RF amplifier designs need to address the compromise of linearity vs. efficiency. This chapter examines the various forms of RF amplifier distortion, together with some of the standard methods of measurement and characterisation.

2.2 Amplitude Distortion

A perfect amplifier would have a linear transfer characteristic, where the output voltage would be a scalar multiple of the input voltage, that is,

$$V_{out}(t) = K_1 V_{in}(t) \tag{2.1}$$

where K_1 is the *voltage gain* of the amplifier. This situation is illustrated in Figure 2.1.

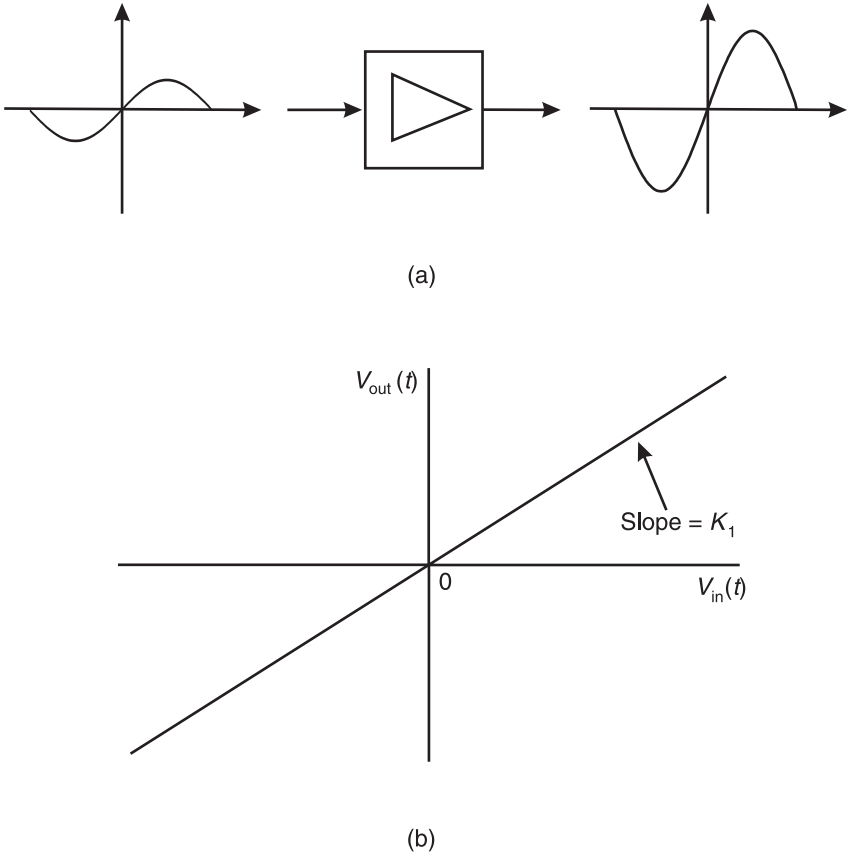


Figure 2.1 Ideal amplifier (a) with a linear transfer characteristic (b).

The output waveshape from such an amplifier will be identical to that of the input and no new (additional) frequency components will be introduced either within or outside of the amplifier bandwidth.

2.2.1 Square-Law Characteristic

The simplest form of amplitude nonlinearity may be illustrated by the addition of a second term to the transfer characteristic (2.1): a term proportional to the square of the input voltage.

$$V_{out}(t) = K_1 V_{in}(t) + K_2 V_{in}^2(t) \tag{2.2}$$

This form of transfer characteristic is referred to as *second-order* due to the power of two which has now been introduced. Figure 2.2 illustrates an

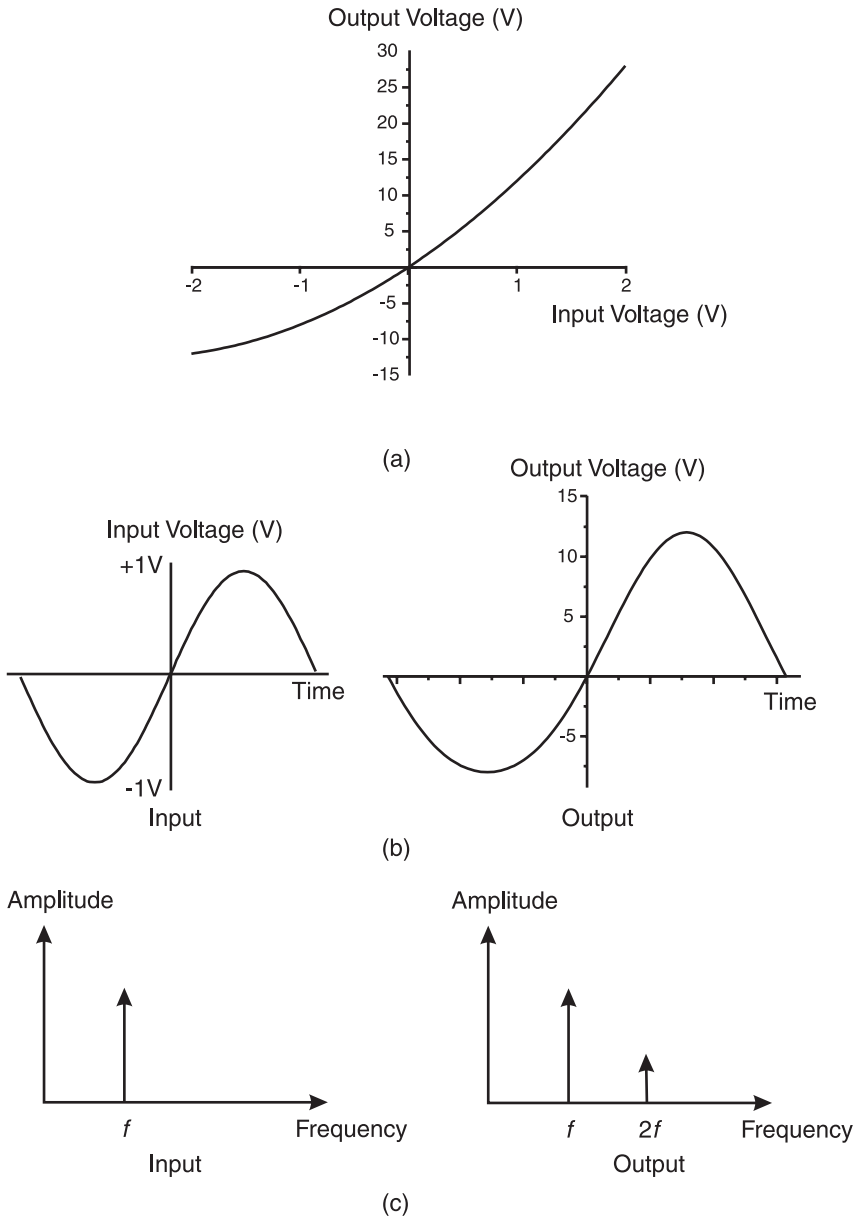


Figure 2.2 Transfer characteristic (a) and effect on a sinusoid in the time domain (b) and frequency domain (c) of an amplifier with transfer characteristic: $V_{out}(t) = 10V_{in}(t) + 2V_{in}^2(t)$.

example characteristic for the case where $K_1 = 10$ and $K_2 = 2$ and demonstrates the effect of such a characteristic on a pure sinusoid in both the time and frequency domains.

The larger the coefficient of the second-order term (K_2), the more curved the transfer characteristic will appear and hence the greater the distortion of the input waveshape. Note that in the frequency domain a second signal component has now appeared at twice the original frequency ($2f_1$) and this gives rise to the term *second harmonic distortion* used to describe the form of nonlinear distortion introduced by the second-order term. Note further that a DC term also results from the second-order term in the transfer characteristic.

Examination of the amplitude of the second harmonic component indicates that it will increase in proportion to the square of the input signal (and also in proportion to the constant, K_2). The amplitude of the fundamental frequency component, however, will only increase in proportion to the fundamental gain, K_1. As a result, it is evident that the amplitude of the second harmonic will increase at a greater rate than that of the fundamental component. A point can thus be envisaged where the fundamental and second harmonic components are of equal level; the signal level at which this would occur is termed the *second-order intercept point*, usually expressed as a power in dBm. This may be quoted as either an input or an output intercept point; the former is most commonly found in receiver front-end specifications and the latter is the usual form for medium- and high-power amplifiers.

The characteristics of the fundamental and second harmonic amplitude levels, with varying input level, are shown in Figure 2.3 for the transfer characteristic illustrated previously ($K_1 = 10$ and $K_2 = 2$). The latter parts of the two characteristics are shown dotted since the input and output levels required to obtain these parts of the characteristics in practice would be impossible without destroying the device. In this example the second-order (output) intercept point may be quoted as approximately 50 volts ($+47$ dBm for a 50Ω load), corresponding to the output signal level where the two characteristics cross.

The advantage of using an intercept point to indicate the linearity performance of an amplifier is that it is a fixed quantity from which the distortion level at a particular operating point may be predicted. The percentage of harmonic distortion which is generally specified in audio amplifiers must be referenced to a particular output power level (usually the maximum for which the amplifier is rated) and gives no indication of the amplifier's performance below that level. A compromise often adopted in RF amplifiers is to de-rate them from their maximum power level, in order to

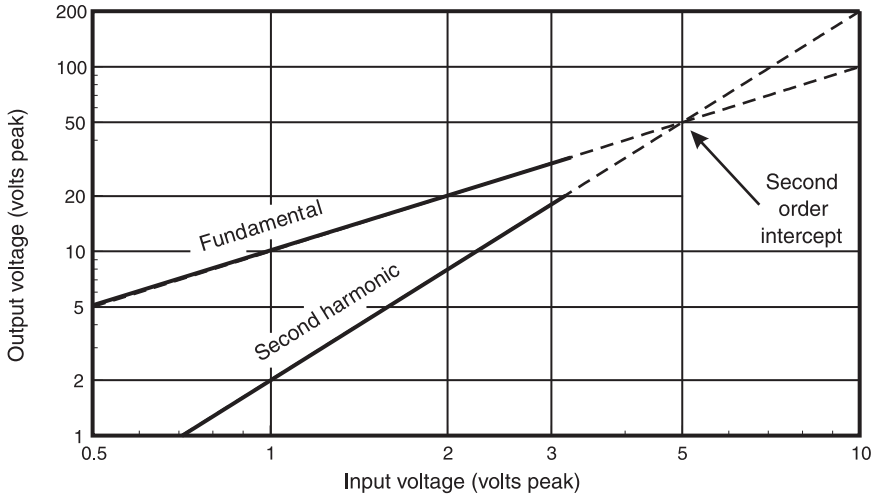


Figure 2.3 Illustration of the second-order intercept point of a nonlinear amplifier.

achieve an improved distortion performance. It would be impossible to predict the level of de-rating required from a percentage distortion measurement unless it was either tabulated or presented graphically.

Finally, note that a second-order characteristic produces harmonic distortion, as outlined above, but does not produce in-band intermodulation distortion (see below). This is an important distinction, in general, between even-order and odd-order nonlinearities: even-order nonlinearities do not generate in-band intermodulation distortion.

2.2.2 Third-Order Characteristic

A very different set of problems occur if an amplifier has a third-order term in its transfer characteristic.

$$V_{out}(t) = K_1 V_{in}(t) + K_3 V_{in}^3(t) \tag{2.3}$$

This characteristic is shown for $K_1 = 10$ and $K_3 = -3$ in Figure 2.4. Note that the output waveshape is now symmetrical above and below the horizontal axis and that a term at three times the original input signal frequency has appeared in the spectrum. This signal is the third harmonic and gives rise to its description as *third harmonic distortion*. Note further that no DC component exists for third-order distortion unlike that present with second-order distortion.

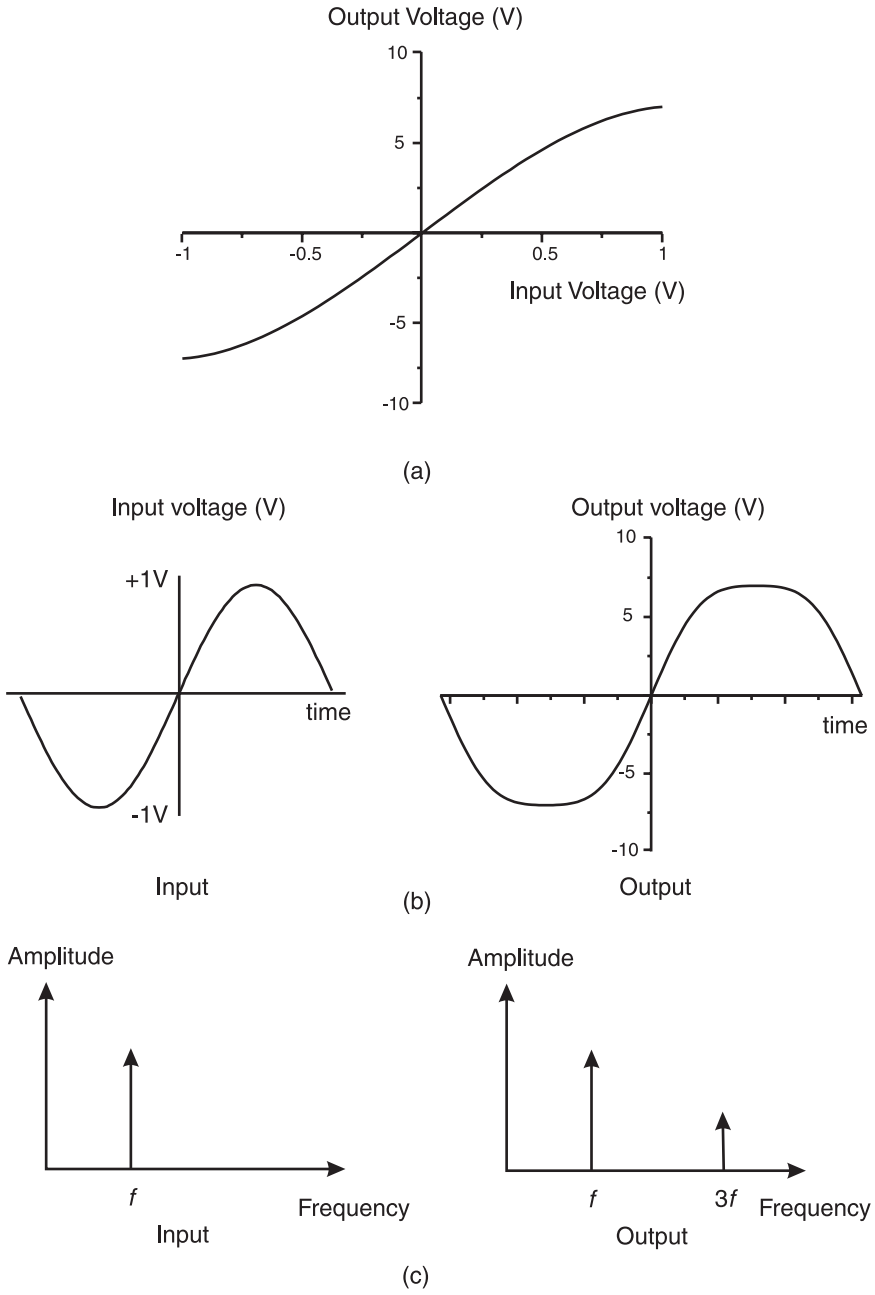


Figure 2.4 Transfer characteristic (a) and effect on a sinusoid in the time domain (b) and frequency domain (c) of an amplifier with transfer characteristic: $V_{out}(t) = 10V_{in}(t) - 3V_{in}^3(t)$.

2.2.2.1 1dB Compression Point

The *1 dB compression point* of an amplifier refers to the output power level at which the amplifier's transfer characteristic deviates from that of an ideal, linear, characteristic by 1 dB. The 1 dB compression point of the amplifier with a third-order nonlinearity is illustrated in Figure 2.5.

A further problem arises when considering two amplitude modulated carriers as the input signals, instead of the simple case of two unmodulated tones examined previously. The amount of compression experienced by a given signal will depend on the instantaneous level of the other signal being amplified. It is thus possible for the amplitude modulation appearing on one carrier to transfer to the other carrier and vice-versa. This problem is known as *cross-modulation* and can be a major problem in AM receivers when they are faced with very strong signals, as well as in transmitters operating close to saturation. This is described in more detail in Section 2.12.

Although gain compression is obviously a problem for amplifiers with a third-order nonlinearity, of greater concern are the intermodulation products appearing at $2f_2 - f_1$ and $2f_1 - f_2$. These distortion products appear 'in-band' and hence will distort the desired waveshape of the original input signal. Furthermore, since these products appear within the band of interest, it is usually impossible to filter them out, unlike the harmonic

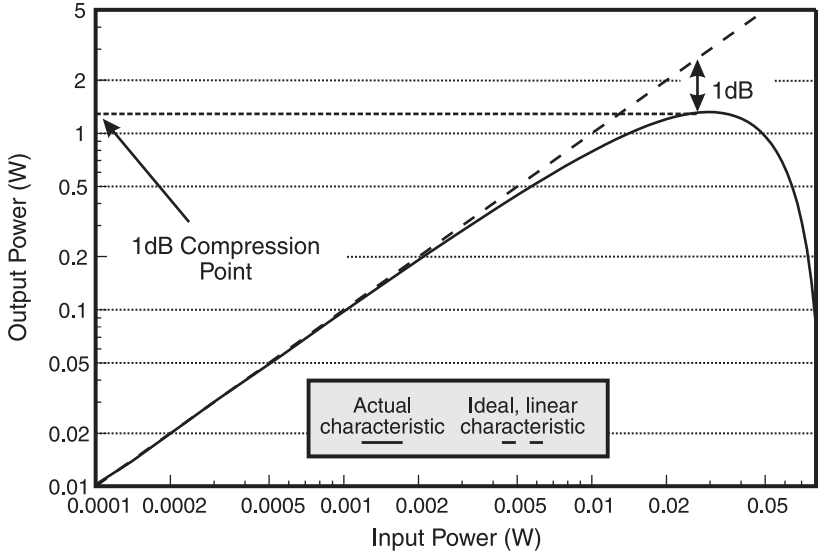


Figure 2.5 1 dB compression point of an amplifier with a characteristic given by: $V_{out}(t) = 10V_{in}(t) - 3V_{in}^3(t)$.

products at $3f_1$ and $3f_2$. For this reason advanced amplifier linearisation techniques, such as those discussed in this book, are required in order to secure their elimination.

2.2.2.2 Third-Order Intercept Point

A *third-order intercept point* may be defined in a similar manner to that of the second-order intercept point examined above; however, the form of the fundamental characteristic requires further explanation. The fundamental and third-order characteristics of an amplifier with a transfer function of the form shown in Figure 2.4 are shown in Figure 2.6. At low signal levels the magnitude of the fundamental component increases almost linearly with input signal level, however, it then begins to deviate from a linear characteristic and eventually decreases again.

This result (Figure 2.6) may be explained as follows. Suppose the input sinusoid is of the form:

$$V_{in}(t) = V_P \sin(\omega t) \tag{2.4}$$

This signal forms the input of an amplifier with a transfer characteristic:

$$V_{out}(t) = 10V_{in}(t) - 3V_{in}^3(t) \tag{2.5}$$

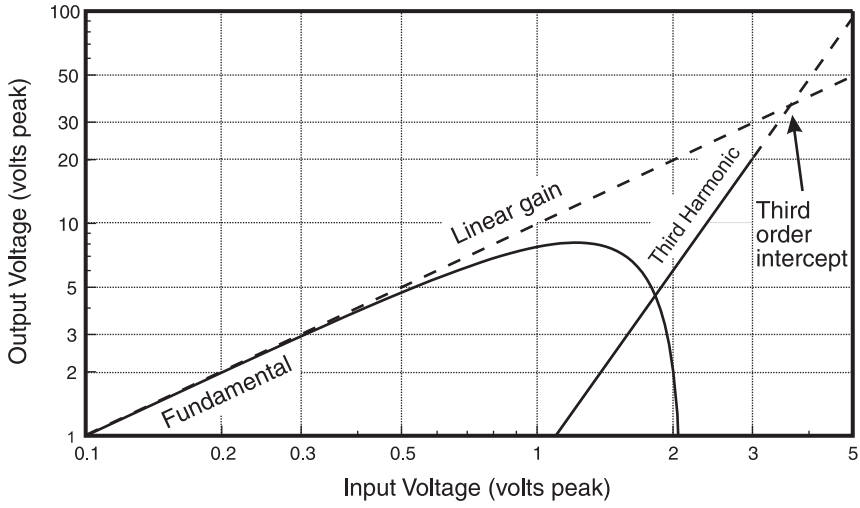


Figure 2.6 Illustration of the third-order intercept point of a nonlinear amplifier.

The resulting output signal is thus:

$$V_{out}(t) = 10V_P \sin(\omega t) - 3\{V_P \sin(\omega t)\}^3 \tag{2.6}$$

which reduces to:

$$V_{out}(t) = 10V_P \sin(\omega t) - \frac{9V_P^3}{4}\sin(\omega t) + \frac{3V_P^3}{4}\sin(3\omega t) \tag{2.7}$$

The first term represents the linear amplification of the fundamental (input signal) and the final term represents the third harmonic distortion. The middle term gives rise to the unusual shape of the fundamental characteristic in Figure 2.6 as it causes partial cancellation of the fundamental due to it appearing at the same frequency. The level of this cancelling signal is proportional to the cube of the input signal amplitude and hence it can quickly have a significant effect on the level of the fundamental in the output signal.

Thus, in an amplifier where third-order distortion predominates, the linear gain characteristic of the fundamental must be extrapolated in order to obtain the third-order intercept point. This is indicated by the dashed line in Figure 2.6.

2.3 Two-Tone Test

The two-tone test is an almost universally accepted method of assessing amplifier linearity and can illustrate both amplitude and phase distortions present in an amplifier. The effect of the two-tone test is to vary the envelope of the input signal throughout its complete range in order to test the amplifier over its whole transfer characteristic. It is thus arguably the most severe test of an amplifier's linearity performance, although it is being challenged as a 'standard' test by alternative techniques more suited to characterising nonlinearities in digital modulation transmitters (e.g., white noise or multicarrier test signals).

When viewed in the frequency domain, the spectrum of a two-tone test signal is shown in Figure 2.7. The individual signals are unmodulated carriers and if the test is carefully constructed, no other products should be in evidence within the band of interest (there will inevitably be some signals present at the generators' harmonics, but these should be small).

If the two-tone signal is viewed in the time domain, the envelope variation can clearly be seen (Figure 2.8). The signal levels should be

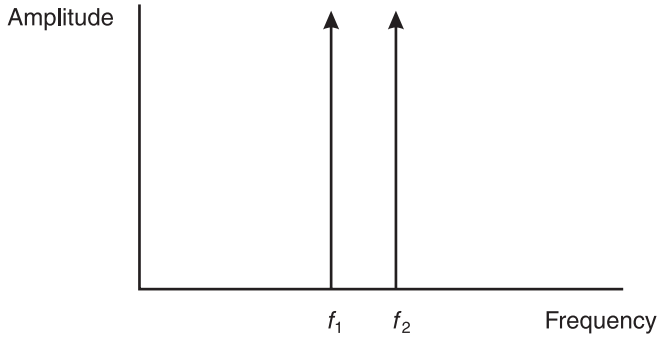


Figure 2.7 Two-tone test in the frequency domain.

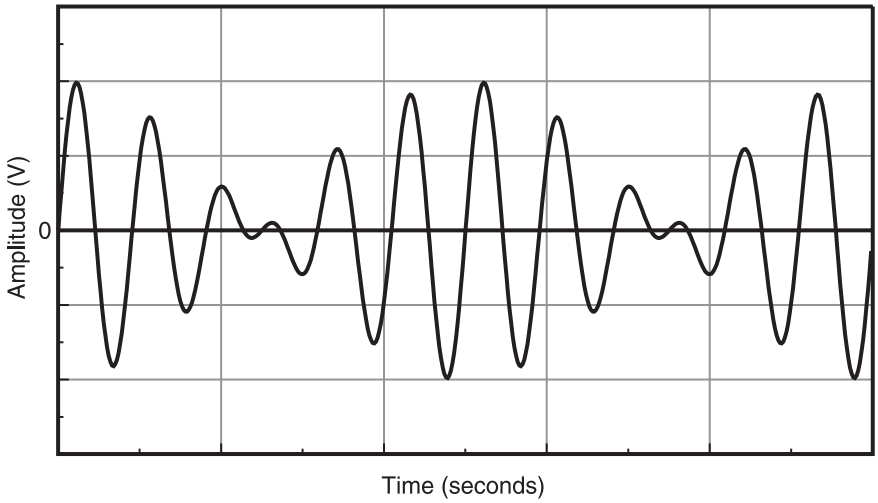


Figure 2.8 Time domain representation of a two-tone test.

arranged such that the peak envelope power (PEP) from the two-tone signal is equal to that of the full power rating at which the amplifier will be used (which may or may not be the same as its maximum CW power rating). For two unmodulated sinusoidal tones of equal level, the peak envelope power of the resulting two-tone signal is 6 dB greater than the CW power in either of the tones (the mean power being 3 dB higher than the CW power in either tone).

2.3.1 Two-Tone Test Applied to an Amplifier With a Second-Order (Square-Law) Nonlinearity

The two-tone test can be applied to an amplifier of the form discussed in Section 2.2.1. Each of the two tones will have a second harmonic and additional sum and difference frequencies will appear. These additional tones will occur at frequencies of: $f_2 - f_1$ and $f_2 + f_1$ and are known as *second-order intermodulation products* since they are created by the $V_{in}^2(t)$ term in the transfer characteristic.

In many RF applications, these distortion products are not significant as they will occur out of the bandwidth of interest, many RF applications requiring bandwidths of less than one octave. Thus the harmonics and intermodulation products falling outside of this bandwidth can be filtered out, the only penalty in this process being the insertion loss of the filter.

A square-law characteristic is often usefully applied in frequency mixers where the two 'tones' are formed by the local oscillator and incoming RF signal. A tuned circuit or monolithic filter is then used to select the required upconverted or downconverted signal.

2.3.2 Two-Tone Test Applied to an Amplifier With a Third-Order Nonlinearity

In the general case of distortion created by any order of nonlinearity when supplied with a two-tone input signal, new frequencies will be generated, as in the case of the second-order nonlinearity discussed above, and these will be of the form:

$$f_{im} = mf_1 \pm nf_2 \tag{2.8}$$

where m and n are positive integers (including zero) and $m + n$ is equal to the order of the distortion.

Thus for a third-order nonlinearity, the additional output frequencies will be:

$$\begin{aligned} f_{im1} &= 3f_1 \\ f_{im2} &= 3f_2 \\ f_{im3} &= 2f_1 + f_2 \\ f_{im4} &= f_1 + 2f_2 \\ f_{im5} &= 2f_1 - f_2 \\ f_{im6} &= f_1 - 2f_2 \end{aligned} \tag{2.9}$$

The original tones will, of course, also appear amplified at the output.

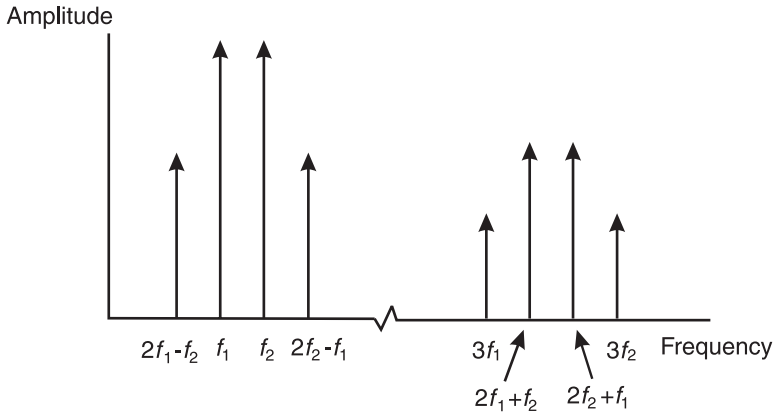


Figure 2.9 Spectrum of an amplifier with linear gain and a third-order nonlinearity.

Figure 2.9 shows these intermodulation products in the frequency domain, along with the amplified tones. The amplified tones will again not appear as large as would be expected from the linear gain portion of the amplifier characteristic (that is, the $10V_{in}(t)$ term in the example given in Figure 2.4) due to the partial cancellation of each tone as described in Section 2.2.2. This effect is termed *compression* and leads to a commonly used specification of maximum output level for an amplifier, that of the 1 dB compression point.

2.3.3 Higher-Order Nonlinearities

When applying a two-tone test to an amplifier exhibiting a number of orders of nonlinearity, a large number of harmonics and intermodulation products are generated and these are all given by the general expression of (2.18). Thus, if the amplifier contains nonlinearities up to and including a seventh-order then m and n will run between 0 and 7 in that equation (provided, of course, that $m+n \leq 7$). If the band of interest is filtered such that out-of-band products are removed, then the remaining intermodulation products will show the classic shape shown in Figure 2.10, for a well-behaved nonlinearity.

The most commonly used measure of intermodulation distortion (IMD) is the ratio of the largest intermodulation product to the amplitude of one of the two tones (assuming that they are equal in level). For a class-A amplifier, this will generally be in the range −30 dB to −35 dB. For most semiconductor amplifiers, the largest IMD product (in-band) will be the third-order product, although this is not always the case.

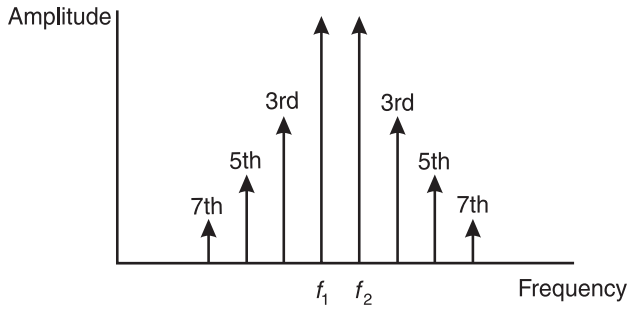


Figure 2.10 In-band intermodulation spectrum of an amplifier with up to a seventh-order nonlinearity.

2.3.4 Combined Effect of Harmonic and Intermodulation Distortion

If an amplifier with all orders of distortion up to a seventh-order is supplied with a two-tone test and the out-of-band products are not assumed to be removed by filtering, the complete spectrum will be that shown in Figure 2.11. The harmonic zones refer to the lowest order of nonlinearity which will produce a signal in that zone. Thus, for example, the second harmonic zone will contain products generated by a second-order nonlinearity (and above).

2.3.5 Effect of IMD on Carrier-to-Noise Ratio

The effect of IMD on a system also corrupted by noise will be to degrade the carrier-to-noise ratio (CNR) from that arising due to noise alone. The IMD products may therefore be considered to add to the received noise level, resulting in a new CNR level.

If the IMD products can be considered noise-like in their properties (this is not true of CW carriers, but is true of most digital modulation formats and of multicarrier signals consisting of a large number of modulated carriers), then it is possible to provide a simple mechanism for analysing the degradation in CNR resulting from the addition of IMD.

The resultant CNR ratio (in dB) for a system corrupted by both IMD and noise is given by:

$$CNR_{N\&IMD} = CNR_I - 10\log\left(1 + 10^{C_{ID}/10}\right) \qquad (2.10)$$

where CNR_I is the intrinsic carrier-to-noise ratio (in dB), C_{ID} is the difference between the carrier-to-interference ratio and the intrinsic CNR

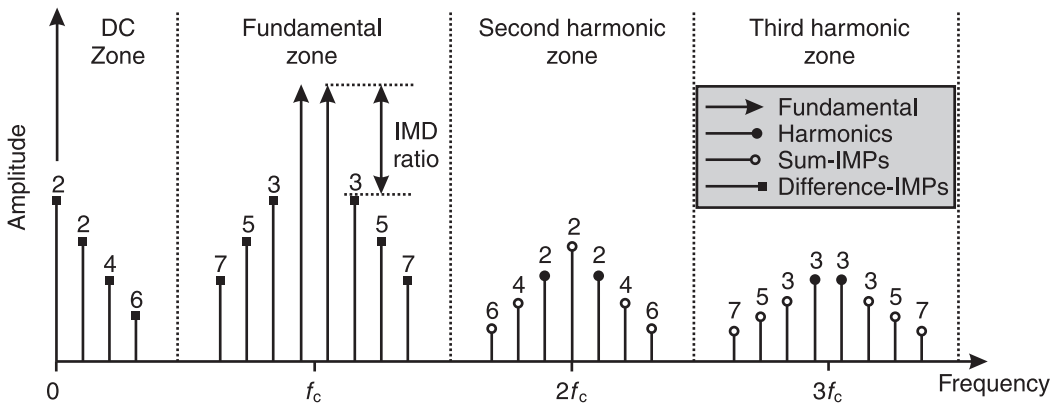


Figure 2.11 Complete frequency-domain response of a nonlinear amplifier supplied with a two-tone test input signal.

(i.e., without any IMD); it is usually negative, implying that the system noise level is higher than the IMD 'noise' level. Thus, for example, if a multicarrier FM signal requires a CNR of 18 dB per carrier (10 dB for the FM capture-effect to operate satisfactorily and 8 dB to allow for a satellite channel fading margin), then a carrier-to-interference ratio of $\geq$28 dB is required in order for the IMD to have a negligible effect (approximately 0.4 dB degradation).

2.4 Calculation of Intermodulation Distortion Ratio

2.4.1 Two-Tone Intermodulation

For an amplifier with a transfer characteristic given by:

$$V_{out}(t) = K_1 v_{in}(t) + K_2[v_{in}(t)]^2 + K_3[v_{in}(t)]^3 \tag{2.11}$$

given an input signal of the form:

$$v_{in}(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t) \tag{2.12}$$

will yield an output signal containing both harmonic and intermodulation distortion (IMD) terms, as outlined in Section 2.3.5.

The intermodulation distortion ratio is defined as the ratio of the amplitude of the highest intermodulation product to the amplitude of one of the tones in the two-tone test. In the ideal case assumed above, with equal tone levels for the two-tone test (i.e., $A_1 = A_2$) and a simple polynomial model for the amplifier characteristic, the highest intermodulation products will generally be the third-order products and both will be equal in level.

In this case, the intermodulation distortion power will vary as the cube of the input signal power, giving:

$$P_{IMD} = (K_{IMD} P_1)^3 \tag{2.13}$$

where P_{IMD} is the power in a single third-order IMD product, K_{IMD} is a constant and $P_1 = A_1^2/2$ is the power in one of the input signal components. Thus for each change in input power of 1 dB, the IMD output power will change by 3 dB.

The intermodulation distortion ratio at the output of the amplifier is given by:

$$P_{IMR} = \frac{P_{IMD}}{P_{O,A1}} \tag{2.14}$$

where $P_{O,A1}$ is the output power in one of the wanted two output tones and the two input tone amplitudes are the same.

Combining the above equations, and recalling that the power level of a wanted tone at the output of the amplifier is linearly proportional to its input power, gives:

$$P_{IMR} = K_C^2 P_1^2 \tag{2.15}$$

where K_C is a constant.

The input third-order intercept point, P_{3rd}; illustrated in Figure 2.6, is the input power at which the IMD power (in the output spectrum) is equal to the output power contributed by the linear term $(K_1 A_1)^2/2$. It is therefore possible to relate this quantity to the IMD ratio and the input power, by realising that at the intercept point, the IMD ratio is unity (by definition), and hence:

$$1 = K_C^2 P_1^2 \tag{2.16}$$

At this point, the input power from a single tone is then the input third-order intercept point, P_{3rd}, that is, $P_1 = P_{3rd}$, and (2.16) then becomes:

$$K_C = \frac{1}{P_{3rd}} \tag{2.17}$$

Therefore (2.15) becomes:

$$P_{IMD} = \left(\frac{P_1}{P_{3rd}}\right)^2 \tag{2.18}$$

If the various values are expressed in dB, this expression may be simplified further to become:

$$P_{IMD,dB} = 2\left(P_{1,dBm} - P_{3rd,dBm}\right) \tag{2.19}$$

Thus, for example, an amplifier with an input intercept point of $+30$ dBm and receiving a two-tone input level of 0 dBm per tone (6 dBm PEP), would produce third-order intermodulation products:

$$P_{IMD,dB} = 2(0-30) = -60 \text{ dBc}$$

that is, -60 dB with respect to the level of each tone. The same would be

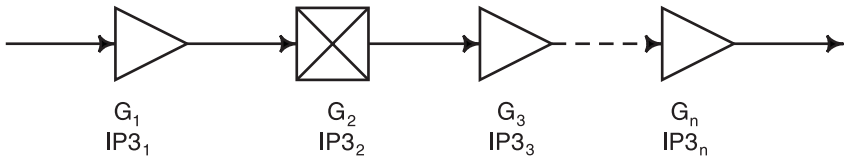


Figure 2.12 Cascaded intercept point calculation.

true of an amplifier with an output intercept point of +30 dBm, when generating an output power level of 0 dBm per tone.

2.4.2 Cascaded Third-Order Intercept Point

It is possible to calculate the effective third-order intercept point of a cascade of elements (such as amplifiers, mixers) from the intercept points of the individual elements. This is useful in determining an approximate value for the intercept point of a complete receiver front-end or of a cascade of amplifier stages forming a power amplifier. The general situation is illustrated in Figure 2.12.

The third-order intercept point of the complete amplifier chain is then given by:

$$\mathrm{IP3}_{tot} = \frac{1}{\dfrac{1}{\mathrm{IP3}_1} + \dfrac{G_1}{\mathrm{IP3}_2} + \dfrac{G_1 G_2}{\mathrm{IP3}_3} + \cdots + \dfrac{G_1 G_2 \cdots G_{n-1}}{\mathrm{IP3}_\mathrm{n}}} \tag{2.20}$$

where each of the IP3 and gain terms is expressed in linear units (not dB), that is,

$$\begin{aligned} G_n &= 10^{G_{n.\mathrm{dB}}/10} \\ \mathrm{IP3}_\mathrm{n} &= 10^{IP3_{n.\mathrm{dB}}/10} \end{aligned} \tag{2.21}$$

It is important not to forget about the intercept point units used initially, as these will usually be specified in dBm, hence producing an intercept point in milliwatts in linear units. Note also that the above intercept points are at the input to each stage and that the overall result is therefore an input intercept point.

2.4.3 Effect of a Driver Stage on Overall Amplifier IMD

It is tempting to assume that the output IMD of an amplifier consisting of a cascade of two or more stages is approximately equal to the IMD from the

final stage, assuming that each of the previous stages has some headroom in driving the following stage. Unfortunately, the IMD produced by a driver stage can have a significant effect upon the overall IMD of the amplifier and hence the amount of 'headroom' allowed must be carefully considered.

Examining (2.20), it is clear that when the output intercept point of the driver stage is equal to the input intercept of the final stage (that is, if the stages were individually tested with low-distortion input signals, each would produce the same relative level of IMD), then the overall output IMD level will degrade by 3 dB. For example, if each of the stages individually (one driver and one output stage) was running at an IMD level of −30 dBc, then the cascaded amplifier would produce an output IMD level of −27 dBc. This can be expressed mathematically as:

$$\text{Degradation in IM3 (dB)} = 10\log\left[1 + 10^{\frac{\text{IM3(Driver)}-\text{IM3(Final)}}{20}}\right] \qquad (2.22)$$

where 'IM3(Driver)' and 'IM3(Final)' refer to the relative third-order intermodulation levels of the driver and final stages in the (two-stage) amplifier and both are expressed in −dBc. This equation is illustrated in Figure 2.13 for a representative range of relative IM3 levels.

As an example, consider a driver amplifier with an output intercept point which is 6 dB greater than the input intercept point of the final stage. It will therefore produce third-order IMD, when tested independently, at

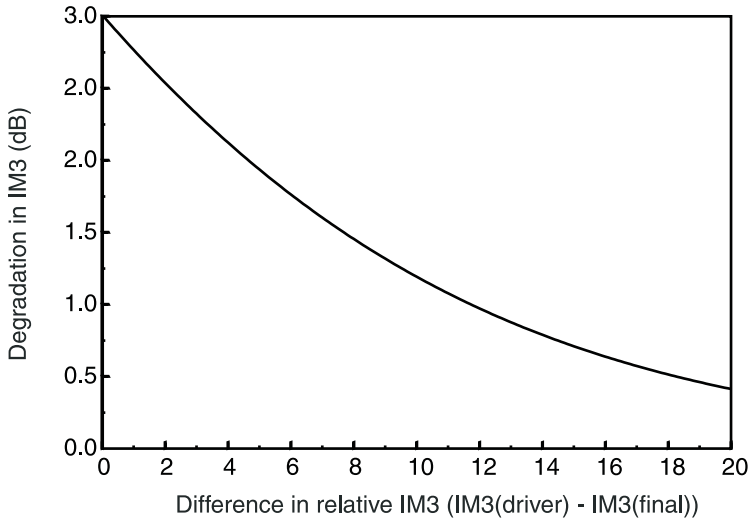


Figure 2.13 Degradation in IMD for a two-stage cascaded amplifier.

12 dBc lower than that of the final stage (i.e., IM3(Driver) −IM3(Final) = −12 in (2.22)). This results in an IM3 degradation of 1 dB for the output IMD of the final stage, when the two stages are cascaded.

2.5 Nonlinearity Measures for Multitone and Modulated Signals

There are a number of standard measurements for determining the degree of unwanted signal energy added by a nonlinear device. The two-tone test has been covered extensively above and provides a good indication of the degree of nonlinearity present over the whole of the amplifier characteristic, under simple signal conditions. Most modern systems are not, however, simple in this respect and hence a number of other measures are required; these will be outlined below.

2.5.1 Adjacent Channel Power Ratio

Adjacent channel power ratio (ACPR) is a measure of the degree of signal spreading into adjacent channels, caused by nonlinearities in the power amplifier. It is defined as the power contained in a defined bandwidth (B_1) at a defined offset (f_o) from the channel center frequency (f_c), divided by the power in a defined bandwidth (B_2) placed around the channel center frequency. The two bandwidths B_1 and B_2 need not be the same (and indeed are not for many current standards). The concept is illustrated in Figure 2.14.

2.5.2 Noise Power Ratio

Noise power ratio (NPR) is a measure of the unwanted in-channel distortion power caused by the nonlinearity of the power amplifier. This can be

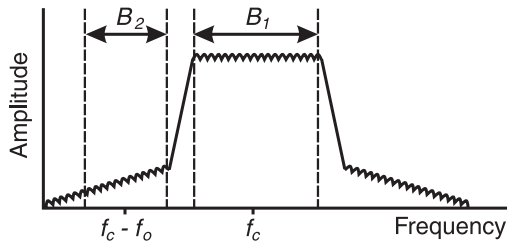


Figure 2.14 Adjacent channel power ratio.

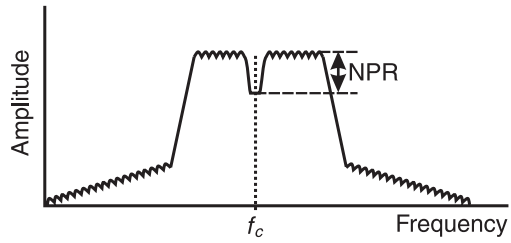


Figure 2.15 Noise power ratio.

measured by extracting a portion of the input signal, using a notch filter, and examining the level of distortion 'filling in' the space within the resulting gap. Further details of this test are provided in Figure 2.15.

NPR is defined as the ratio between the noise power spectral density of a white noise signal passing through the amplifier, measured at the center of the notch, to the noise power spectral density without the notch filter, where the amplifier is driven at the same power level in each case. The concept is illustrated in Figure 2.15.

2.5.3 Multitone Intermodulation Ratio

Multitone intermodulation ratio (M-IMR) is a measure of the effect of nonlinearity on a multicarrier signal. This could be a multicarrier modulation format (e.g., OFDM) or a multicarrier signal from, for example, a base-station transmitter. It is defined as the ratio between the wanted tone power (of one of the multiple tones) and the highest intermodulation tone power just outside of the wanted band. The concept is illustrated in Figure 2.16.

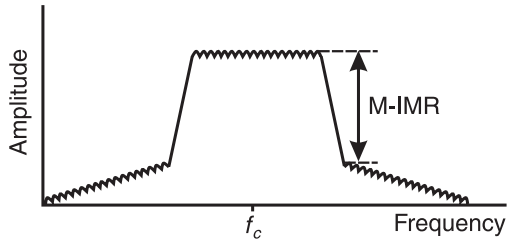


Figure 2.16 Multitone intermodulation ratio.

2.5.4 Relationship Between Two-Tone IMD and Complex Signal IMD

Based on simple third-order modelling of an amplifier characteristic, it is possible to derive empirical relationships between the above complex-signal IMD measures and their two-tone counterparts [1,2]. It is therefore possible to approximate the multitone or complex signal behavior of an amplifier based on only a simple two-tone measurement. This process has many potential areas of inaccuracy, but it does serve to provide an approximate idea of the desired response prior to more detailed measurements.

The process is based upon Volterra–Weiner theories [3,4], which state that any third-order system may be completely characterised by a three-tone test. Increasing the number of tones, therefore, will not yield any further information about the system and any additional effects seen in practice, when performing multitone testing, are due to the presence of higher-order products.

By examination of the statistics of a multitone signal with uncorrelated phases and greater than about 10 carriers, the central limit theorem indicates that the resulting waveform tends toward a narrow-band noise excitation and the response of the system therefore approximates a noise power ratio test. Examination of a multicarrier signal utilising an analytical approach can therefore be used to derive approximate measures of the three complex-signal IMD measures discussed above (ACPR, NPR, and M-IMR).

Based on the definitions below, the following equations may be derived:

$\mathrm{IMR}_{2\text{-tone}}$	Two-tone intermodulation ratio (dBc)
IP_3	Device third-order intercept point (dBm)
P_{ave}	Mean output power from the device (dBm)
n	Number of tones
r	Adjacent channel product number ($r = 1$ for the first—closest—IMD product)
b	Number of tones on one side of the 'gap', where one or more tones is removed for NPR testing. For removal of a single tone at the center of the band, $b = f[(n-1)/2]$, where $f(\bullet)$ extracts the integer part of the expression in brackets (by rounding down).

2.5.4.1 ACPR

$$\mathrm{ACPR}_{\mathrm{dBc}} = \mathrm{IMR}_{2\text{-tone}} - 6 + 10\log\left(\frac{n^3}{4\mathcal{A}+B}\right) \tag{2.23}$$

where

$$A = \frac{2n^3 - 3n^2 - 2n}{24} + \frac{\text{mod}(n/2)}{8} \tag{2.24}$$

and

$$B = \frac{n^2 - \text{mod}(n/2)}{4} \tag{2.25}$$

Note that mod(x/y) is defined as the remainder when x is divided by y.

2.5.4.2 NPR

$$\text{NPR}_{\text{dBc}} = \text{IMR}_{\text{2-tone}} - 6 + 10\log\left(\frac{n^2}{4C + D}\right) \tag{2.26}$$

where

$$C = \left(\frac{n - b - 2}{2}\right)^2 - \frac{\text{mod}[(n + b)/2]}{4} + \left(\frac{b - 1}{2}\right)^2 - \frac{\text{mod}[(b + 1)/2]}{4} + b(n - b - 2) \tag{2.27}$$

and

$$D = \left(\frac{n - b - 2}{2}\right) - \frac{\text{mod}[(n + b)/2]}{2} + \left(\frac{b - 1}{2}\right) + \frac{\text{mod}[(b + 1)/2]}{2} \tag{2.28}$$

2.5.4.3 M-IMR

$$\text{MIMR}_{\text{dBc}} = \text{IMR}_{\text{2-tone}} - 6 + 10\log\left(\frac{n^2}{4E + F}\right) \tag{2.29}$$

where

$$E = \left(\frac{n - r}{2}\right)^2 - \frac{\text{mod}[(n + r)/2]}{4} \tag{2.30}$$

and

$$F = \left(\frac{n-r}{2}\right) + \frac{\text{mod}[(n+r)/2]}{2} \tag{2.31}$$

For all of equations (2.23), (2.26), and (2.29), the relationship between the two-tone intermodulation ratio, $\text{IMR}_{2\text{-tone}}$, and the total *mean* output power, P_{ave}, is given by:

$$P_{ave} = \frac{-\text{IMR}_{2\text{-tone}}}{2} + \text{IP}_3 + 3\text{dB} \tag{2.32}$$

The additional factor of 3 dB (when compared with (2.19)) arises from the fact that (2.32) uses the average power of the output signal and not the power per tone considered previously.

Figure 2.17 plots these characteristics, taking the same equivalent two-tone IMD level in each case (0 dB). This allows the relative level of each measure to be judged and indicates that ACPR is the closest to a two-tone equivalent, with NPR being the furthest from the two-tone case.

2.5.5 Examples

Consider an unlinearised multicarrier power amplifier with a third-order intercept point of + 60 dBm, which is required to amplify 16 carriers, each of

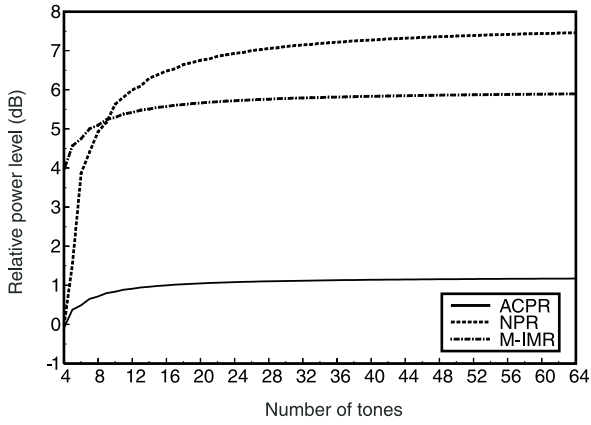


Figure 2.17 Normalised relationship between the various nonlinearity measures for multitone signals.

1W (+30 dBm). Using this as an example, it is possible to predict (and compare) the three multicarrier IMD specifications outlined above (ACPR, NPR, and M-IMR).

2.5.5.1 ACPR

Calculating the two intermediate values gives:

$$\begin{aligned} A_{16} &= 308 \\ B_{16} &= 64 \end{aligned} \tag{2.33}$$

and hence:

$$\mathrm{ACPR_{dBc}} = \mathrm{IMR_{2\text{-}tone}} - 1.023 \tag{2.34}$$

substituting for $\mathrm{IMR_{2\text{-}tone}}$ from (2.32):

$$\begin{aligned} \mathrm{IMR_{2\text{-}tone}} &= 2(IP_3 - P_{ave}) + 6 \\ &= 2[IP_3 - (10\log(n) + P_{tone})] + 6 \\ &= 2[60 - (10\log(16) + 30)] + 6 \\ &= 41.9 \text{ dBc} \end{aligned} \tag{2.35}$$

where n is the number of carriers (16 in this example) and P_{tone} is the per-carrier power of the signal (+30 dBm in this example).

Therefore:

$$\begin{aligned} ACPR &= 41.9 - 1.023 \\ &= 40.9 \text{ dBc} \end{aligned} \tag{2.36}$$

2.5.5.2 NPR

Calculating the two intermediate values gives:

$$\begin{aligned} C_{16} &= 69 \\ D_{16} &= 7 \end{aligned} \tag{2.37}$$

and hence:

$$\mathrm{NPR_{dBc}} = \mathrm{IMR_{2\text{-}tone}} - 6.456 \tag{2.38}$$

substituting for $IMR_{2\text{-tone}}$ from (2.35) above gives:

$$\begin{aligned} NPR_{dBc} &= 41.9 - 6.456 \\ &= 35.5 \text{ dBc} \end{aligned} \tag{2.39}$$

2.5.5.3 M-IMR

Calculating the two intermediate values gives:

$$\begin{aligned} E_{16} &= 56 \\ F_{16} &= 8 \end{aligned} \tag{2.40}$$

and hence:

$$MIMR_{dBc} = IMR_{2\text{-tone}} - 5.593 \tag{2.41}$$

substituting for $IMR_{2\text{-tone}}$ from (2.35) above gives:

$$\begin{aligned} MIMR_{dBc} &= 41.9 - 5.593 \\ &= 36.3 \text{ dBc} \end{aligned} \tag{2.42}$$

This analysis in no way represents a 'worst-case' for n carriers being amplified by a single amplifier—the phases have been assumed to be random and uniformly distributed and the nonlinearity has been assumed to be purely third-order. It can therefore be said to represent an 'average' level of intermodulation if a device were to be tested with a number of carriers a large number of times. It is therefore not a conservative estimate of IMD level (or power rating), although it does serve to illustrate, in some measure, the effect of multiple tones on a nonlinear amplifier. A more conservative 'rule-of-thumb' often used in the absence of more rigorous information or analysis, is to rate a multicarrier amplifier at a mean power level 9 dB or 10 dB below its peak power rating, for a large number of carriers (>16).

Modulation applied to the carriers, specifically that involving a significant degree of phase modulation, may serve to create the 'random' phase behavior, assumed in this analysis, as a continuous feature.

2.6 Crest Factor and Crest Factor Reduction Techniques

2.6.1 Crest Factor Definition

There are a range of terms and often contradictory definitions for '*crest factor*', '*peak factor*', '*peak to average ratio*' and '*peak to mean ratio*'. Essentially they are all methods of defining the statistics of a modulated signal in a manner which an amplifier designer can understand and interpret. The two definitions which have become standard are:

Crest factor: the ratio of the peak to r.m.s. *amplitude* of a signal [5]

$$\mathrm{CF} = \frac{\left|\hat{S}(t)\right|}{\sqrt{|S(t)|^2}} \tag{2.43}$$

Peak-to-mean ratio: the ratio of the peak *power* to r.m.s. *power* of a signal (for example, 6). Some typical values for common modulation formats are provided in Table 2.1.

$$\mathrm{PMR} = \frac{\text{Peak Power}}{\text{Average Power}} \tag{2.44}$$

Note that:

$$\mathrm{PMR} = \mathrm{CF}^2 \tag{2.45}$$

2.6.2 Crest Factor Reduction for Multicarrier Signals

There are a range of techniques for reducing the crest factor (and hence by implication the peak-to-average ratio) of a multicarrier signal and this can have a significant impact upon the power rating required of a linear or linearised power amplifier. This is a substantial (and highly mathematical) subject and a detailed treatment is beyond the scope of this book—however, a brief summary of the primary techniques is appropriate. These techniques involve either careful phasing of RF carriers to minimise the overall peak-to-mean ratio, coding of the modulation information in order to minimise signal transitions which increase the peak-to-mean ratio (e.g., those traversing the origin in the complex plane) or appropriate distribution (mapping) of the required data across the carriers (e.g., in an OFDM system). Clearly not all techniques are appropriate to all systems and much of the work in this area has been directed toward OFDM systems.

The primary methods include *Shapiro–Rudin sequences* [7,8], *Golay codes*

Table 2.1
Peak to mean ratios for some common modulation formats

Modulation scheme	Parameter	Crest factor (dB)
Wideband CDMA (see Note 1)		
	16 occupied channels	10.5
	32 occupied channels	11.1
	64 occupied channels	12.2
	128 occupied channels	13.6
$\pi/4$-DQPSK (see Note 2)		
	$\alpha = 0.20$	4.86
	$\alpha = 0.25$	4.55
	$\alpha = 0.30$	4.23
	$\alpha = 0.35$	3.87
	$\alpha = 0.40$	3.38
	$\alpha = 0.50$	3.21
16-QAM (see Note 3)		
	$\alpha = 0.20$	6.03
	$\alpha = 0.25$	5.92
	$\alpha = 0.30$	5.66
	$\alpha = 0.35$	5.40
	$\alpha = 0.40$	5.18
	$\alpha = 0.50$	4.94
GSM EDGE (8-PSK) (see Note 4)		3.21

Notes:

1. Wideband CDMA parameters: chip-rate: 4.096Mchips/s, filter type: root-raised cosine, $\alpha = 0.22$. All channels occupied with statistically-independent pseudo-random data.
2. $\pi/4$-DQPSK parameters: symbol rate: 24.3ksymbols/s, filter type: root-raised cosine, DAMPS coding applied.
3. 16-QAM parameters: symbol rate: 25ksymbols/s, filter type: root-raised cosine, no coding applied.
4. GSM-EDGE: as per ETSI SMG2 technical document WPB 386/98.

[9,10], *maximal-length sequences* [11], *Barker codes* [12,13], *Newman phases* [14,5], *Schroeder phases* [15], *block coding* [16], *selected mapping* [6] and *partial transmit sequences* [17]. More recently, attention has turned to multicarrier CDMA and the reduction of peak-to-mean power in this type of system [18].

2.7 Phase Distortion

Any deviation from perfect linearity of the amplitude transfer characteristic of an amplifier will lead to distortion of the output waveshape. The same is also true of the phase characteristic for that amplifier, although the reason why is perhaps less obvious. A nonlinear phase characteristic will result if an amplifier does not delay all frequency components within the input signal by the same amount when they reach the output. The effect of such a characteristic on a 1 kHz 'squarewave', comprised of only the first three frequency components of its Fourier analysis, is shown in Figure 2.18.

The fundamental (1 kHz) has been delayed by 100 μs, the third harmonic (3 kHz) by 200 μs and the fifth harmonic (5 kHz) by 300 μs; the resultant waveshape bears little relation to that of the original squarewave. Similar distortion is present in long cable runs, such as those of the public switched telephone network (PSTN), and hence the use of direct baseband data modulation on such systems is inadvisable (FSK or M-ary QAM modems are more usual).

The relationship between time delay, τ and phase shift, ϕ is:

$$\tau = \frac{\phi}{2\pi f} \tag{2.46}$$

where f is the fundamental frequency of the waveform and ϕ is in radians. The time delay will be identical for all frequency components making up a complex waveform, if the phase-shift increases in proportion to the frequency. Thus it is the time delay imposed by the amplifier, and not its phase-shift, which must remain constant over the bandwidth of interest, to eliminate distortion of the waveshape.

If these ideas are extended to an amplifier operating at radio frequency, then it is the modulating signal's waveshape which must be preserved. As an example, consider the three-frequency 'squarewave' used above. If it is used to amplitude modulate a high frequency carrier, then the spectrum shown in Figure 2.19 will result. Since the modulation will be recovered from the carrier further on in the communication system it must be preserved in the transmitter to a high degree of fidelity. Thus the delay through the RF amplifiers for each of the frequencies including the carrier must be constant. The actual magnitude of the delay is, in general, not important.

With modern network analysis techniques it is a relatively simple matter to obtain the delay characteristics of an amplifier over the frequency range of interest. The deviation from a flat characteristic indicates the degree of phase distortion which may be expected for the amplifier.

Alternatively, the delay may be measured using a high-frequency oscilloscope and by passing a sinusoidally modulated carrier through the device under test (DUT). The envelope delay may be measured as illustrated in Figure 2.20 and this should be constant as the modulating frequency is varied. The absolute value of envelope delay is given by:

$$T_e = \frac{\Delta\phi}{2\pi\Delta f} \quad \text{seconds} \tag{2.47}$$

where $\Delta\phi$ is the phase difference between the carrier and the modulation sideband (in radians) and Δf is the frequency difference between carrier and sideband.

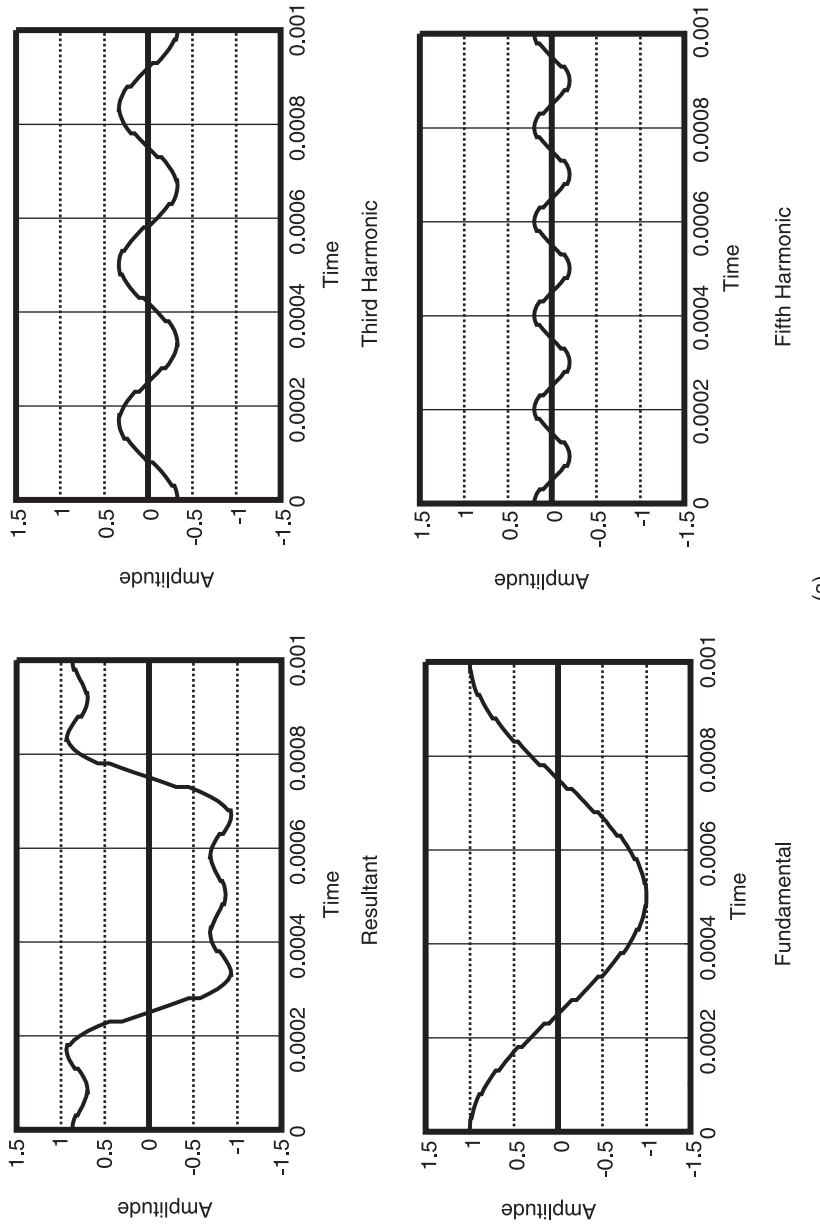
Note that envelope delay and group delay are *not* the same and are measured in a different manner. The terms are frequently used interchangeably and the error resulting from this is small in a narrow-band system, however care should be exercised to ensure that the correct term is used [19].

2.8 Practical Creation of a Multitone Test Signal

The creation of a multitone test signal suitable to fully characterise the in-band distortion characteristics of a highly linear RF amplifier is a non-trivial problem. The distortion present on the test signal must be appreciably below the anticipated distortion level from the amplifier and this may prove difficult if the amplifier is expected to perform to an IMD specification of 50 dBc or better.

The problem may be broken down into that of testing at RF and that of testing using baseband information (such as when a complete transmitter is being tested). RF testing will be covered first.

Most two-tone or multitone testing is performed with the aid of signal generators of some description, at least during the design phase. Modulated carriers from 'real' system sources are usually only applied when satisfactory performance of a prototype with either simulated or CW signals has been achieved. Most laboratory signal generators contain power-levelling diodes electrically close to their RF output port(s) and any signal energy leaking through the power combiner (combining the various signal sources) will cause intermodulation to occur in this diode. This will occur at the output of each signal generator, resulting in a distorted output signal from the power combiner (see Figure 2.21). The level of this distortion will vary with the brand of generator employed, the isolation of the power combiner and the required output level. However as a 'rule of thumb' a distortion level in the vicinity of 50–55 dBc will result from a two-tone test, employing commer-


Amplitude
1.5
1
0.5
0
-0.5
-1
-1.5
0
0.0002
0.0004
0.0006
0.0008
0.001
Time
Resultant
Fundamental
Third Harmonic
Fifth Harmonic

(a)

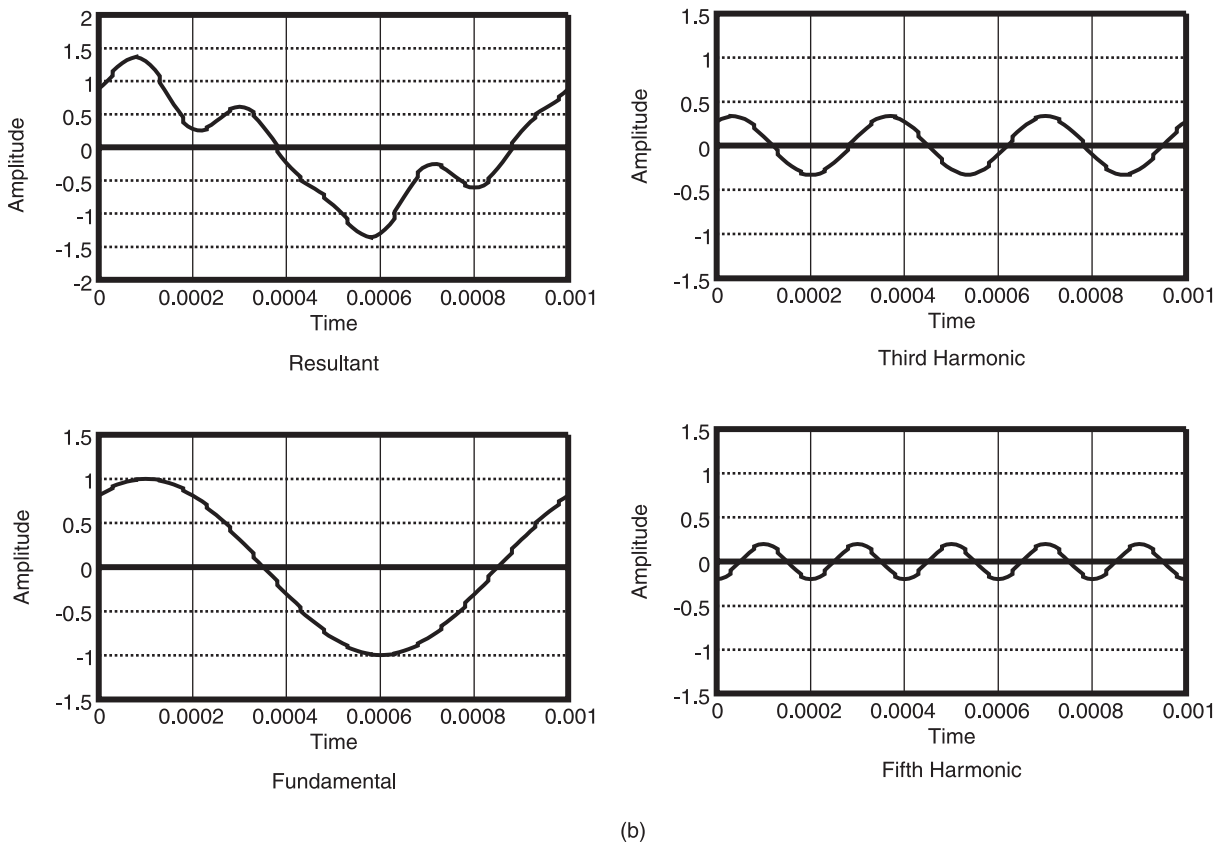


Figure 2.18 Phase distortion of a 1 kHz squarewave. (a) Original signal and (b) Distorted resultant.

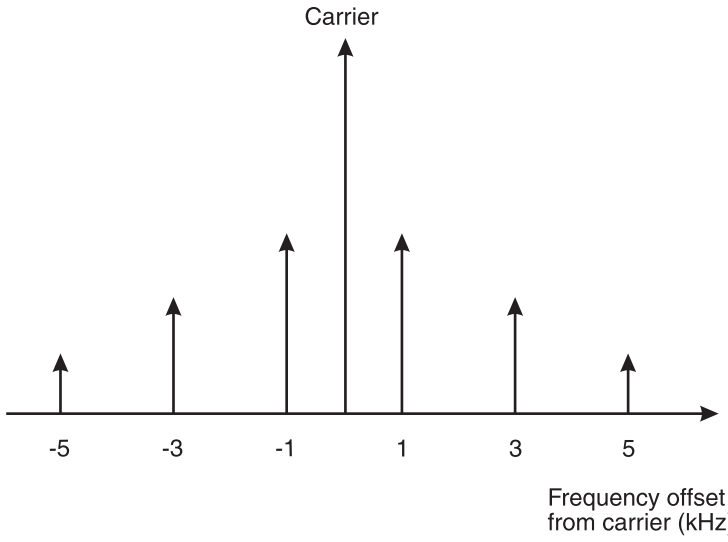


Figure 2.19 Spectrum of a carrier plus 1 kHz squarewave amplitude modulation.

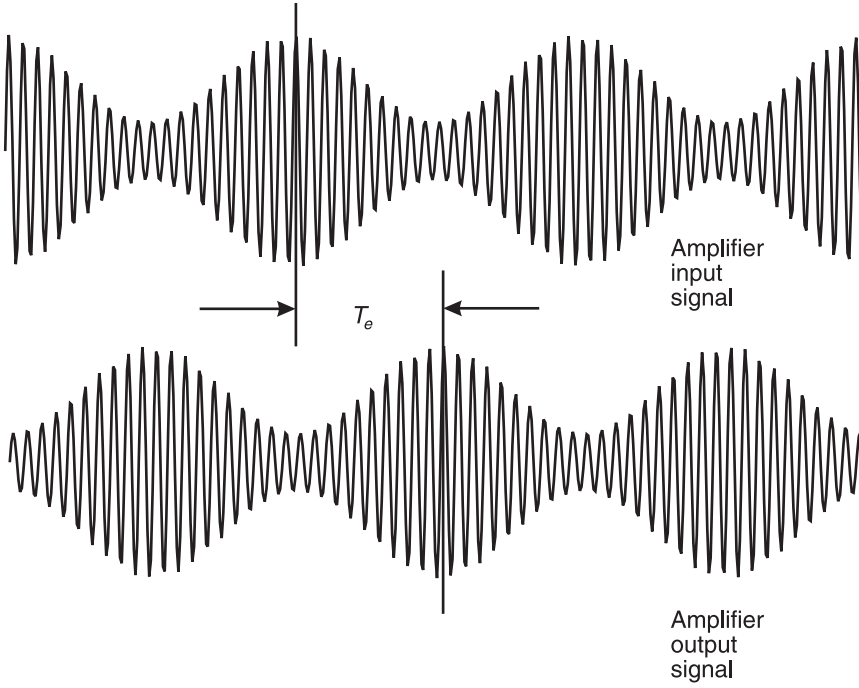


Figure 2.20 Measurement of envelope delay using an oscilloscope.

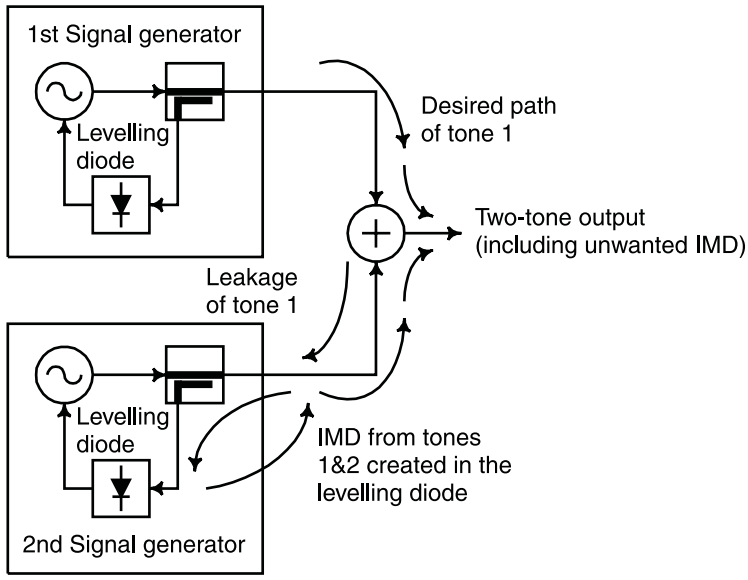


Figure 2.21 Unwanted IMD generation when using commercial signal generators to create a two-tone test (note that a similar process, not shown, occurs in the first signal generator).

cial bench signal generators at a power output of +10 dBm per tone, using a transformer-based 3 dB hybrid as the combiner. It must be stressed that this figure is only a rough guide and a particular test arrangement may yield better or worse performance than this.

To enable the satisfactory testing of higher performance amplifiers, such as feedforward systems for base station applications, it is necessary to achieve an IMD specification for the test signal approaching 80 dBc. To achieve this level of performance requires a high degree of isolation to be introduced between the signal generators and the power combiner.

A typical test arrangement is illustrated in Figure 2.22. In this arrangement, each signal source is provided with either an attenuator or an isolator in its output path. An attenuation value of at least 10 dB is recommended for 60–70 dBc IMD performance and must be provided externally, as the levelling diode is often placed after the internal attenuators in the generator.

The use of an attenuator will obviously reduce the drive level of the test signal and this is usually undesirable. A better solution is to use one or more isolators in each generator output. In this manner, multitone test

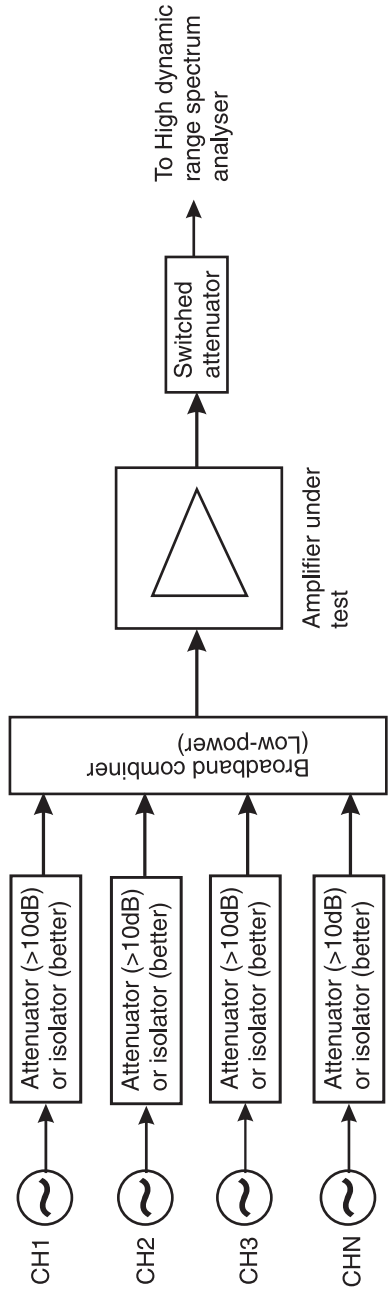


Figure 2.22 Practical realisation of a multitone test at RF.

signals with an IMD performance of 80 dBc or greater may be created, with relatively little loss of drive level (1–2 dB is typical).

High-quality baseband test signals are generally easier to create, but require careful thought if an in-phase and quadrature test signal is required (such as for Cartesian loop and CALLUM transmitters). It is possible to arrange the test tones to appear above, below or either side of the carrier, with the latter usually being the preferred option. In the case of a complete transmitter these test tones can also be used to judge the gain and phase balance achieved between the I and Q channels in the upconversion process, since they may be slightly offset (i.e., non-symmetrical about the carrier) and hence allow the image signals to be seen. The gain and phase balance of the input signals can then be adjusted until the desired image cancellation is achieved and the gain and phase 'error' required to do this, noted. Further details on diagnostic testing of this nature are provided in Chapter 4.

Table 2.2 illustrates the I and Q signals required to provide a two-tone test surrounding the carrier in an SSB-type system. Such a signal is ideal for testing both Cartesian loop and CALLUM transmitters and may also be used in adaptive baseband predistortion systems. The tones should be straightforward sinewaves, with an amplitude appropriate for the transmitter under test. Modern four-channel audio-frequency synthesisers are ideal for this task.

After I/Q upconversion, the two tones will have a frequency spacing of $f_1 + f_2$ and should be of equal level. If the carrier is at a frequency, f_C, the tones will be located at: $(f_C - f_1)$ and $(f_C + f_2)$; the image sidebands will occur at: $(f_C + f_1)$ and $(f_C - f_2)$. Thus, for example, to achieve a tone spacing of 20 kHz, set $f_1 = 9$ kHz and $f_2 = 11$ kHz. The terms δa and $\delta\phi$ indicate the gain and phase error respectively in the upconversion process (note that in the case of transmitters involving baseband feedback, for example, Cartesian loop, the gain and phase error is set in the error path, that

Table 2.2
I and Q baseband signals to create a two-tone test around a carrier in a quadrature-based transmitter

Channel	Frequency	Amplitude	Phase
I	f_1	a	0°
I	f_2	a	180°
Q	f_1	$a + \delta a$	$90^\circ + \delta\phi$
Q	f_2	$a + \delta a$	$90^\circ - \delta\phi$

is, by the downconverter). Both δa and $\delta\phi$ must be varied by equal amounts in the directions indicated by their respective signs. Thus, for example, to correct for a 3° phase error in the upconversion process, the phase of f_1 in the Q channel should be increased by 3° and the phase of f_2 in the Q channel should be reduced by 3°.

2.9 Two-Tone Test With Unequal Tone Powers

The above discussions have all assumed that the two tones used are identical in level, hence providing an infinite (theoretically) envelope variation, although the peak to mean ratio is only 3 dB. If the tone amplitudes are made unequal, then it is possible to test particular parts of the amplifier's transfer characteristic. In particular it is often useful to be able to test the top few dB of the characteristic (up to the 1 dB compression point), as this allows the amplifier's effect on quasi-linear modulation formats to be assessed (e.g., $\pi/4$-DQPSK). These formats have envelope variations which do not fall to zero, despite (in many cases) having a typical peak-to-mean ratio of around 3 dB.

Figure 2.23 shows the peak-to-minimum ripple obtained from a two-tone test for a range of values of amplitude difference between the tones.

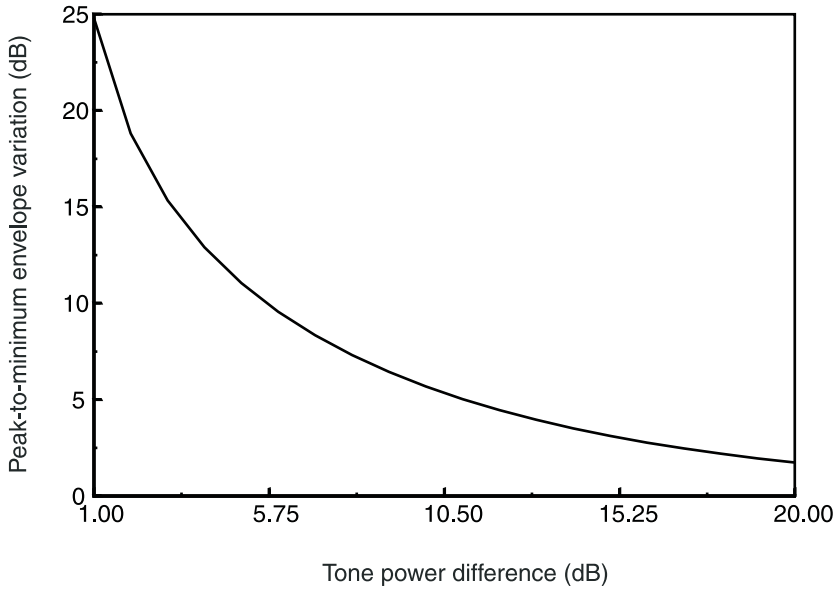


Figure 2.23 Peak-to-minimum envelope variation of a two-tone test with unequal tone powers.

Table 2.3
Peak-to-minimum envelope variation obtained from a two-tone test for a range of values of amplitude difference between the tones

Tone Amplitude Difference (dB)	0	1	2	3	4	5	6	7	8	9
Peak-to-minimum Envelope Variation (dB)	∞	24.8	18.8	15.3	12.9	11.1	9.6	8.3	7.3	6.4
Tone Amplitude Difference (dB)	10	11	12	13	14	15	16	17	18	19
Peak-to-minimum Envelope Variation (dB)	5.7	5.0	4.5	4.0	3.5	3.1	2.8	2.5	2.2	2.0

This data is also reproduced in Table 2.3. Using these it is possible to select an appropriate tone inequality in order to simulate the effect of amplifier nonlinearity for a given modulation format, without the need for advanced test equipment. Note that the results will only be approximately correct, but that they will generally give a more realistic idea of performance than would a conventional, equal-power two-tone test.

A comparison between the (measured) equal power two-tone performance of an amplifier and the equivalent TETRA $\pi/4$DQPSK performance is provided in Chapter 4. This shows clearly that a conventional two-tone test is very pessimistic for this application.

2.10 White Noise Testing of Amplifier Linearity

Although the two-tone test is widely used to measure the linearity of amplifier stages it does not accurately simulate the effects of speech signals or of linear data schemes (such as 16-QAM). It is also a poor indicator of linearity for TWTA systems involving a large number of carriers (>10). A more realistic test for these systems is to use a band-limited white noise signal which has a flat power spectrum across a single channel within the operational bandwidth of the amplifier (or an appropriate band representative of a number of carriers in a multicarrier amplifier). At all frequencies outside of the channel (or band, in the case of a multicarrier system), the signal should be zero.

If this signal forms the input to an amplifier, then any signals in the output spectrum which appear outside of the channel are due to amplifier nonlinearity. In the case of a multicarrier system, the white noise is simulating a large number of carriers of random amplitude and phase and

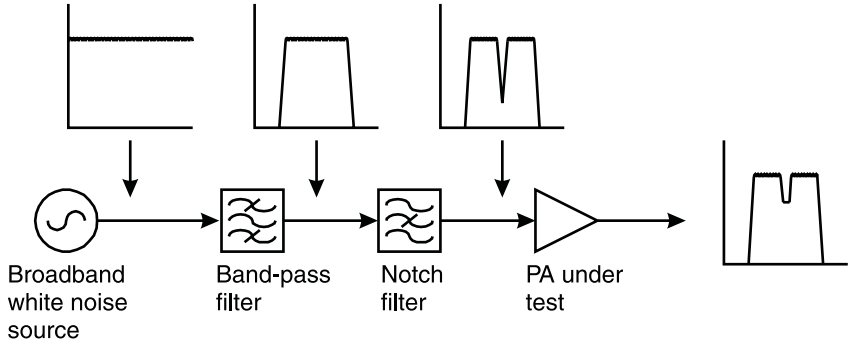


Figure 2.24 White noise testing of amplifier linearity.

a narrow notch filter may be used to leave a 'gap' in the center of the band of interest. This gap will then be 'filled' with any IMD resulting from nonlinearities in the power amplifier, and the ratio of the noise power in this notch to that in the bands either side is the noise power ratio (NPR) for the amplifier.

A block-diagram of this approach is shown in Figure 2.24. The width of the notch should be about 1% or less of the width of the bandpass-filtered noise source and this may most easily be produced at IF, with subsequent (high-linearity) upconversion to the required RF band of interest. The notch depth of the resulting test signal should be at least 10 dB greater than the maximum NPR to be measured [20].

In practice, it may be difficult to produce an input signal of the form described above, particularly for narrow-band systems, and so a carrier, amplitude modulated by a pseudo-random binary sequence (PRBS) of ± 1 in level, may be used instead. The resulting signal has zeros at the edges of the channel and at the edges of the adjacent channels throughout the band, assuming that the bit-rate is chosen appropriately. Thus for a channel of bandwidth $2f_b$, the bit rate of the PRBS must be f_b and zeros will occur at $\pm f_b$, $\pm 3f_b$, $\pm 5f_b$, and sequentially. Any energy present in the output of the amplifier at these frequencies is then a measure of its nonlinearity [21].

2.11 Spurious Signals

Some signals appearing at the output of an amplifier bear no obvious relationship to the input signals being amplified. They may appear and disappear at random and may change frequency and level at will. Such

signals are referred to as *spurious products* and consist of parasitic and subharmonic oscillations together with unwanted external interference which is inadequately shielded by the amplifier casing.

Most spurious signal problems are the result of poor RF constructional techniques and inadequate case design or build quality. The problems are usually made more acute by the constant desire to make RF circuitry as small as possible, particularly in the hand-portable radio market. Amplifiers for use in products in this area must be extremely stable and robust.

2.12 Cross-Modulation

In a nonlinear system in which two input signals are being processed, one desired and one undesired, amplitude modulation present on the undesired signal can be imposed upon the desired signal and this can cause potentially severe interference to the wanted signal (see Figure 2.25). This is a particular problem for receiver systems, as amplitude modulation present on a strong unwanted signal can transfer onto a weaker, wanted signal and interfere with, or even mask, the wanted amplitude modulation present on that carrier. Figure 2.25 illustrates this problem for a two-tone modulated 'unwanted' signal and a CW 'wanted' (weaker) signal. Note that third-, fifth-, and seventh-order distortion only is assumed in this example.

This affect can also, however, have a bearing upon multicarrier power-amplifier systems in a power-controlled environment. In such a scenario, it is possible to have a large signal, required to serve a user close to the cell

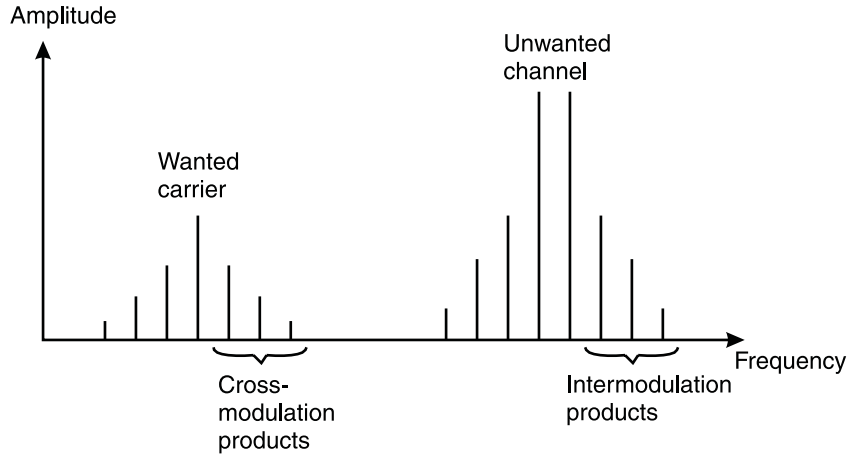


Figure 2.25 Cross-modulation affecting a weak carrier.

boundary, in the presence of a much smaller signal, required to serve a user close to the base station transmitter. Amplitude modulations, such as those present on filtered digital modulation schemes, can transfer from the higher powered signal to the lower powered signal, and thereby degrade the signal vector error of the lower powered signal.

The likely effect of this in a practical system scenario is that the mobile will report a poor signal quality and the base station will therefore increase its output power in order to provide the mobile with a better signal. This mechanism will succeed, at some power level; however, this power level will be higher than that strictly necessary to provide the required service quality (in the absence of nonlinearity in the power amplifier). The result of this will be the potential for additional interference elsewhere in the network, as unnecessary power is being transmitted.

2.13 Modelling of Amplifier Nonlinearities

2.13.1 Ideal Transfer Characteristic

The model of an ideal linear amplifier, with an optimal transfer characteristic, may be represented by an ideal linear limiter, with a transfer function given by:

$$f\{A(t)\} = \begin{cases} A(t) & \forall\, |A(t)| \leq A_0 \\ K & \forall |A(t)| > A_0 \end{cases} \tag{2.48}$$

The form of this (amplitude) characteristic is illustrated in Figure 2.26 for both positive and negative input signal swings. The phase characteristic is assumed to be linear (constant delay) at all input amplitude levels.

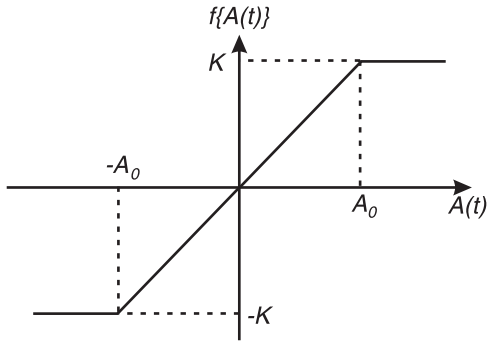


Figure 2.26 Ideal linear limiter as a model of a linear amplifier characteristic.

2.13.2 Memoryless (Instantaneous) Nonlinear Model

As its name suggests this form of model assumes that the power amplifier has no memory effects, that is, that it has no knowledge of past events and hence the present output signal is only a function of the present input signal(s). It is possible to further simplify this type of model by restricting the analysis to signals and distortion contained within the first harmonic zone; such models are then referred to as *bandpass memoryless nonlinear models* and are further simplified since they do not need to model even-order nonlinearities (none result in distortion in the first-harmonic zone).

When applying a memoryless nonlinear model to an RF amplifier, the characteristics of primary interest are the nonlinear gain response (AM/AM characteristic) and amplitude to phase conversion (AM/PM characteristic). These may be modelled in polar (amplitude and phase) (Section 2.13.5) or Cartesian (I and Q) form (Section 2.13.7).

2.13.3 AM–AM and AM–PM Conversion in a Nonlinear Amplifier

Most of the discussions so far in this chapter have concentrated on amplitude nonlinearity, that is the nonlinear relationship between input power and output power present in all practical amplifiers. This is often termed *AM–AM conversion* since it is a conversion between the amplitude modulation present on the input signal(s) and the modified amplitude modulation present on the output signal (the 'modification' being due to the amplitude nonlinearity of the amplifier).

Another effect is, however, also present and that is a conversion from amplitude modulation on the input signal to phase modulation on the output signal; this is known as *AM–PM conversion*. Consider a sinusoidally-modulated input carrier, $P_{in}(t)$, with a modulating signal defined by:

$$M(t) = A_M \cos(\omega_M t) \tag{2.49}$$

For an ideal amplifier, the output phase is given by:

$$\Phi(P_{in}(t)) = K_\phi \tag{2.50}$$

where K_ϕ is a constant. In other words, the output phase remains constant, irrespective of the amplitude (or envelope level) of the input signal.

In the case of a real amplifier, however, amplitude modulation present on the input signal will result in phase modulation of the output signal (i.e., AM–PM conversion), hence (2.50) becomes:

$$\Phi(P_{in}(t)) = K_\phi \cos[\omega_C t + A_M \cos(\omega_M t)]$$
$$= K_\phi \sum_{n=-\infty}^{\infty} J_n(A_M) \cos[(\omega_C + n\omega_M)t] \tag{2.51}$$

where J_n is a Bessel function of order n and ω_C is the angular frequency of the carrier signal. The resulting spectrum is effectively that of a phase-modulated carrier with a sinusoidal modulating signal, hence forming IMD products similar to those resulting from an amplitude nonlinearity.

A real amplifier will, of course, suffer from both forms of nonlinearity to some degree and the output spectrum will consist of a superposition of both effects. This superposition can result in an asymmetry of the IMD products, since the upper and lower IMD products resulting from AM–AM

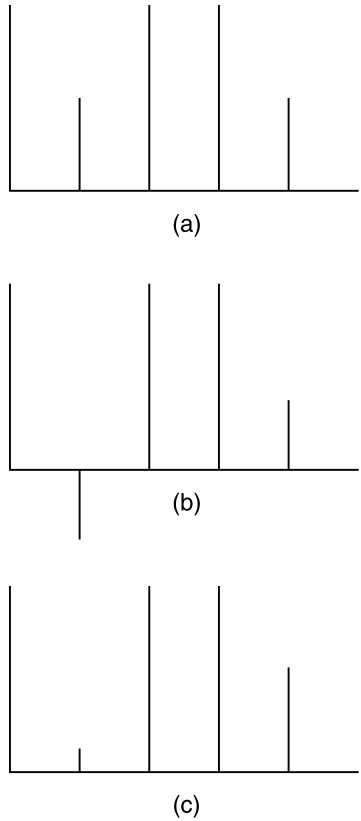


Figure 2.27 IMD resulting from a combination of AM–AM and AM–PM conversion. (a) IMD from AM–AM conversion. (b) IMD from AM–PM conversion, (c) Resultant IMD from combined AM–AM and AM–PM conversion.

conversion are in-phase, whereas a component of those resulting from AM–PM conversion can be 180° out of phase. The phase difference in the IMD products created by the AM–PM conversion process can arise from the mechanisms described in Section 3.13. The effect of the superposition of the AM–AM and AM–PM characteristics is illustrated in Figure 2.27.

2.13.4 Measurement of AM/AM and AM/PM Characteristics

The most convenient method of determining these characteristics from a practical power amplifier is by means of a vector network analyser, although the AM/AM characteristic alone may be measured by a scalar network analyser.

Most modern network analysers contain a 'power sweep' function which permits the source power to be linearly varied by the instrument over a range of 20 dB or more, at a single operating frequency. The amplitude and phase characteristics which result from this measurement are then the approximate AM/AM and AM/PM characteristics of the amplifier under test.

Care must be exercised when making this measurement (as with any other power amplifier measurement) to ensure that the instrument is not damaged by the application of an excessive RF input level.

The measurement is most useful if the (say) 20 dB power variation produced by the network analyser is arranged to just drive the amplifier into compression (by perhaps 1 dB or 2 dB). This will then illustrate what is likely to be the most nonlinear portion of the characteristic and is usually all that is necessary to characterise most amplifiers (the lower portion of the characteristic is usually relatively linear). If, however, it is suspected that significant nonlinearity is contained in the lower portion of the characteristic, then the measurement can be repeated with 20 dB of attenuation inserted before the amplifier under test and the two characteristics combined. Any effects caused by the attenuator itself (e.g., increased delay) will need to be calibrated out. The required data from the two measurements is usually available via the GPIB interface on the analyser, and hence the two characteristics may easily be combined to form the complete model.

Pulsed versions of these measurements provide a more accurate measure of the 1 dB compression point for signals containing a significant amount of AM.

2.13.5 Polar Form of a Memoryless Nonlinear Model

It is possible to treat the AM/AM and AM/PM components of a nonlinearity as separate items and hence produce a model which is effectively a cascade of the two processes. This may be considered at two levels: first, by considering

the effect of the model on a single RF sinewave (CW carrier) and, secondly, by extension of this to a modulated signal. The latter is of considerably greater value, since it will model the distortion of the signal envelope and hence the important in-band effects which are usually of greatest interest.

Consider a CW signal, $C(t)$, at a frequency f_C and with an amplitude A:

$$C(t) = A\cos(2\pi f_C t + \theta) \tag{2.52}$$

The distorted output signal, $D(t)$, of a bandpass memoryless nonlinearity is then:

$$D(t) = f(A)\cos[2\pi f_C t + \theta + g(A)] \tag{2.53}$$

where $f(A)$ describes the AM/AM characteristic of the nonlinearity and $g(A)$ describes its AM/PM conversion.

Work by Minkoff [22,23] has shown, by means of measurement, computer simulation and analytical methods, that (2.52) and (2.53) can be applied to modulated signals as well as to a single CW carrier. Thus:

$$C(t) = A(t)\cos[2\pi f_C t + \theta(t)] \tag{2.54}$$

where $A(t)$ describes the envelope modulations present on the carrier.

The resulting distorted output signal, $D(t)$, is now:

$$D(t) = f\{A(t)\}\cos[2\pi f_C t + \theta(t) + g\{A(t)\}] \tag{2.55}$$

This model can therefore characterise *envelope nonlinearities* and can be used to describe the in-band distortion of a signal containing some form of envelope variation. Such variations are present on a wide variety of signals, including full-carrier AM, SSB, filtered $\pi/4$-DQPSK, 16-QAM (and above), all multicarrier signals and many other forms of modulation. A block diagram of this form of envelope nonlinearity is shown in Figure 2.28.

The AM/AM and AM/PM characteristics of a typical class-A amplifier

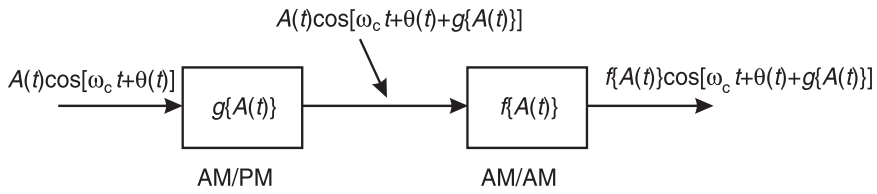


Figure 2.28 Polar form of an envelope nonlinearity for both AM/AM and AM/PM distortion.

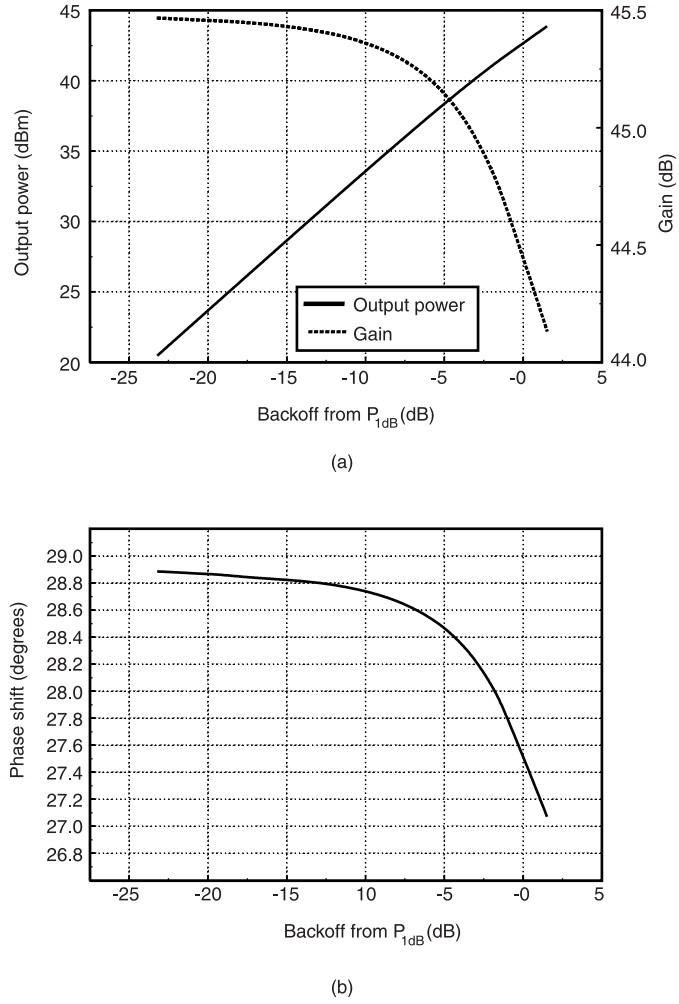


Figure 2.29 AM/AM (a) and AM/PM (b) conversion of a 900 MHz 40W class-A amplifier.

are shown in Figure 2.29 (results obtained from measurements on an actual class-A amplifier at 900 MHz). By eye, this amplifier looks to have an extremely linear characteristic, although its actual third-order IMD performance is only around 27 dBc at its 1 dB compression point.

The AM/AM and AM/PM characteristics of a typical class-C amplifier are shown in Figure 2.30 (results obtained from measurements on an actual class-C amplifier at 400 MHz). This amplifier has a very small amount of bias applied which has the effect of improving amplifier stability at the point

where the devices turn on; it therefore does not exhibit the abrupt 'turn-on' characteristic usually associated with amplifiers operated in class-C.

The overall DC to RF conversion efficiency of the amplifier, even with the addition of the bias, is almost 60%. It is evident therefore that the added bias has very little effect on the power efficiency of the amplifier. The amplifier does, however, exhibit a much steeper transfer characteristic than the class-A stage (roughly a 1.5 dB increase in output signal for every 1 dB increase in input signal over much of its characteristic). Feedback linearisation techniques, for example, the Cartesian loop, must have sufficient

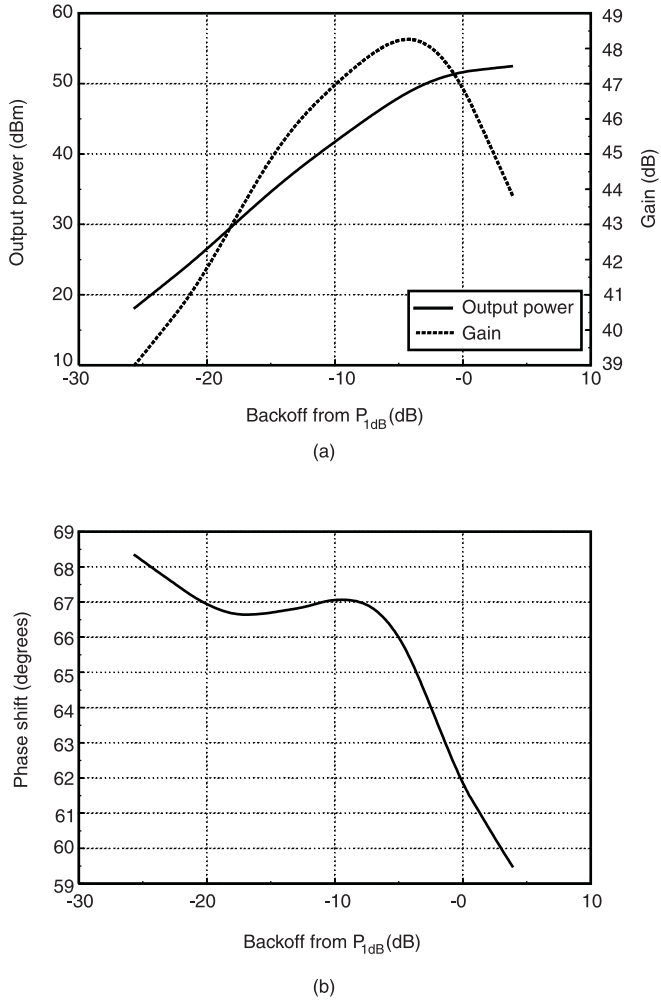


Figure 2.30 AM/AM (a) and AM/PM (b) conversion of a 400 MHz 150W class-C amplifier.

bandwidth to cope with these rapid transitions and it is this factor which often limits the loop gain available from such techniques when applied to a class-C stage. The AM–PM conversion characteristic also has a point of inflection, which is not present on that of the class-A stage.

2.13.6 Effect of AM–AM and AM–PM Conversion on Digital Modulation Formats

Increasing use is being made of digital modulation formats in both voice and data communications. Many of these formats are severely affected by the AM–AM and AM–PM conversion processes which take place in nonlinear amplifiers, irrespective of adjacent-channel considerations. The formats which are affected are those with envelope variations present on the RF signal (i.e., after any relevant filtering) and include two of the most popular schemes currently in use, namely QAM and filtered QPSK.

It is possible to highlight the effects of amplifier nonlinearity on the modulation itself, by examining the constellation and eye diagrams for these two schemes both before and after passing through a highly nonlinear (class-C) power amplifier. The resulting characteristics are shown in Figures 2.31 and 2.32 and the relevant modulation scheme parameters in Table 2.4. The class-C power amplifier had an output power rated at 10W peak and the modulation peaks were arranged to be at this level (i.e., the amplifier was not clipping the signal peaks). The channel center frequency was 435 MHz.

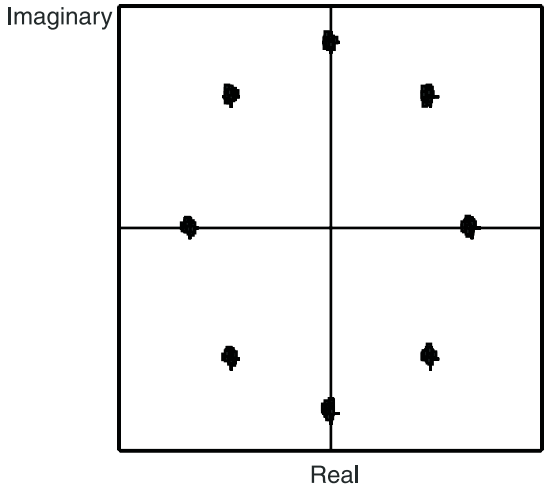
Note that in both cases, the constellation points have become ill-defined; the QAM constellation also betrays the severe AM–PM conversion present in the class-C stage through the obvious rotation of the entire constellation. The eye diagrams show clearly the effect on bit-error performance, as the sampling points (where the eyes should be 'open') are almost completely closed, thus making error-free detection impossible, even in the absence of other noise or interference.

2.13.7 Cartesian Form of a Memoryless Nonlinear Model

In a similar manner to the above polar nonlinear model, an equivalent Cartesian model can be generated. This model has the advantage that it can be constructed from two nonlinear *amplitude* models, $I\{A(t)\}$ and $Q\{A(t)\}$, thereby avoiding the potential complexity of an AM/PM model.

The distorted output from a bandpass memoryless nonlinearity (from (2.55)) may be expanded to give:

$$\begin{aligned} D(t) = {} & f\{A(t)\}\cos[g\{A(t)\}]\cos[2\pi f_C t + \theta(t)] \\ & - f\{A(t)\}\sin[g\{A(t)\}]\sin[2\pi f_C t + \theta(t)] \end{aligned} \tag{2.56}$$



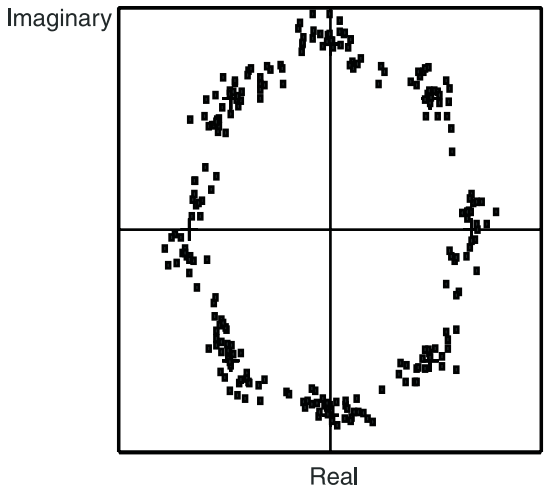
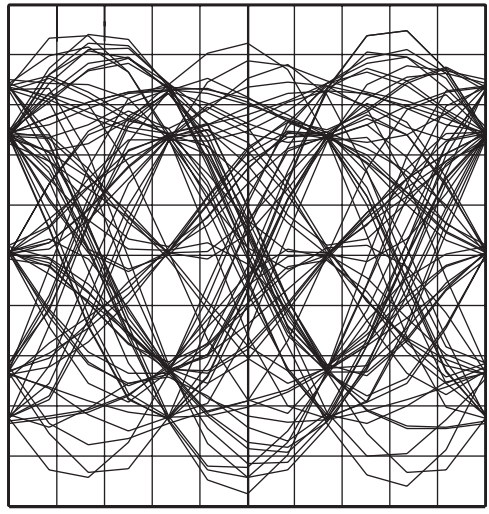
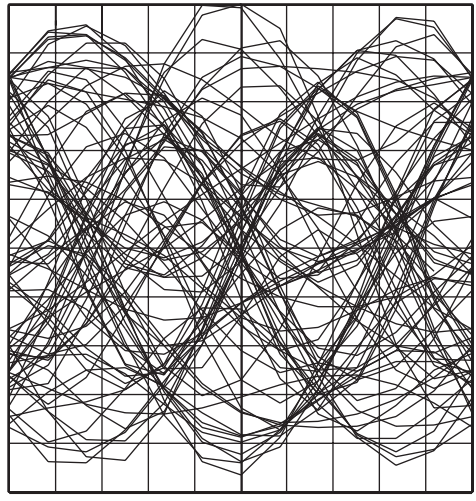


Figure 2.31 The effect of a class-C power amplifier upon a $\pi/4$-DQPSK (TETRA) signal. (a) Input signal constellation. (b) Output signal constellation. (c) Input signal eye diagram. (d) Output signal eye diagram



Symbols

(c)



Symbols

(d)

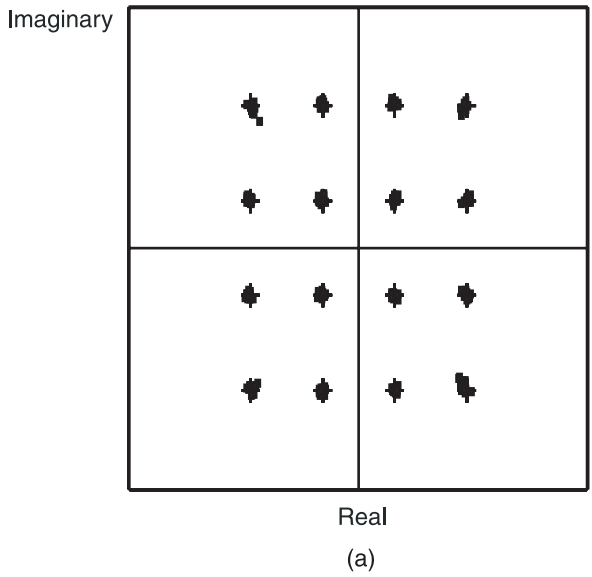
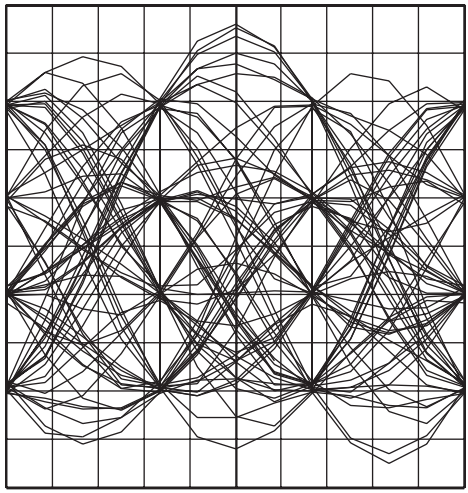
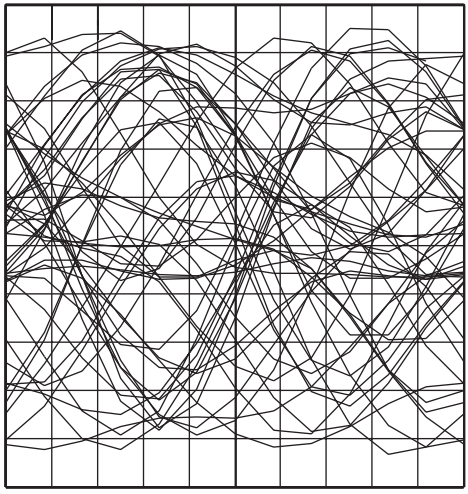


Figure 2.32 The effect of a class-C power amplifier upon a 16-QAM signal. (a) Input signal constellation. (b) Output signal constellation. (c) Input signal eye diagram. (d) Output signal eye diagram.



Symbols

(c)



Symbols

(d)

Table 2.4
Digital modulation scheme parameters used to obtain Figures 2.31 and 2.32

Modulation format	Measurement filter and transmitter filter	Reference filter	Symbol rate
π/4-DQPSK (TETRA)	Root-raised cosine $\alpha = 0.35$	Raised cosine $\alpha = 0.35$	18 kbps
16-QAM 4 bits/symbol	Root-raised cosine $\alpha = 0.35$	Raised cosine $\alpha = 0.35$	15 ks/s (60 kbps)

This may be expressed in quadrature components as:

$$D(t) = I\{A(t)\}\cos[2\pi f_C t + \theta(t)] - Q\{A(t)\}\sin[2\pi f_C t + \theta(t)] \quad (2.57)$$

where:

$$I\{A(t)\} = f\{A(t)\}\cos[g\{A(t)\}] \quad (2.58)$$

and

$$Q\{A(t)\} = f\{A(t)\}\sin[g\{A(t)\}] \quad (2.59)$$

and is shown diagrammatically in Figure 2.33.

The Cartesian form of the AM/AM and AM/PM characteristics of a typical class-C stage are shown in Figure 2.34 (for the same amplifier as

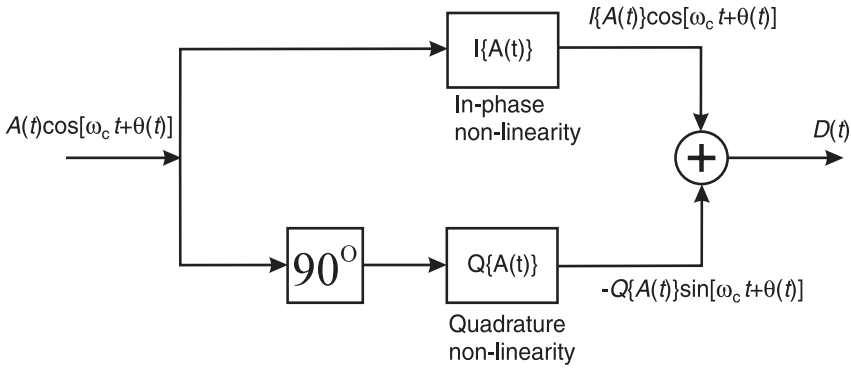


Figure 2.33 Cartesian form of an envelope nonlinearity for both AM/AM and AM/PM distortion.

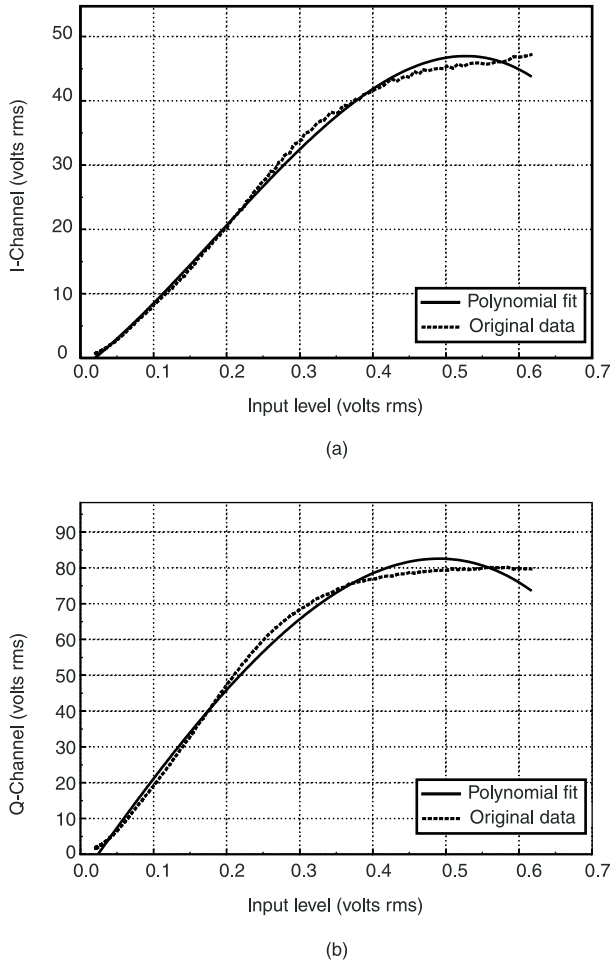


Figure 2.34 Cartesian form of the AM/AM and AM/PM characteristics of a typical class-C amplifier. (a) I-channel, with cubic polynomial fit: $V_{OUT,I} = 86.79V_{in} + 197.45V_{in}^2 - 353.92V_{in}^3$. (b) Q-channel, with cubic polynomial fit: $V_{OUT,Q} = 290.025V_{in} - 67.39V_{in}^2 - 309.70V_{in}^3$.

depicted in Figure 2.30). At first glance, the AM–AM and I-channel characteristics appear very similar and this may be explained as follows.

From (2.56) and (2.57), it can be seen that for low levels of AM/PM conversion, where:

$$\sin[g\{A(t)\}] \approx g\{A(t)\} \tag{2.60}$$

and

$$\cos[g\{A(t)\}] \approx 1 \tag{2.61}$$

then (2.58) and (2.59) may be simplified to become:

$$I\{A(t)\} = f\{A(t)\} \tag{2.62}$$

and

$$Q\{A(t)\} = f\{A(t)\}g\{A(t)\} \tag{2.63}$$

Thus, over most of the AM/PM range depicted in Figure 2.30, these approximations hold true and hence explain the overall shape of the characteristics in Figure 2.34.

2.13.8 Approximate Forms of Memoryless Nonlinear Model

The polar and Cartesian modelling techniques described above require full knowledge (and hence measurement) of the AM/AM and AM/PM characteristics of the nonlinear stage and also often require interpolation between data points, which may prove time consuming. A number of simpler models, involving more analytical functions may be used to improve speed at the expense of model accuracy.

2.13.8.1 Taylor Series

The use of this approach has already been illustrated, for a simple case, in Section 2.2. The concept involves describing the distorted output of the system, $D(t)$, in the form of a Taylor series [24,25] of the input, $v_{in}(t)$:

$$D(t) = \sum_{n=1}^{\infty} a_n v_{in}^n(t) \tag{2.64}$$

where a_n are constants.

This form of model has the advantage that the relative level of each order of distortion is clear from its coefficient a_i and the level of each IMD product may therefore be calculated. It is most useful for systems which contain relatively few orders of distortion, such as TWT amplifiers which have a predominantly third-order characteristic, and hence the series may be truncated after a manageable number of terms without loss of accuracy. Most semiconductor amplifiers, however, generate a very large number of orders of distortion and hence yield a long Taylor series expansion.

A standard Taylor series, as described by (2.64), will only characterise amplitude (AM/AM) distortion. If significant AM/PM distortion exists within the nonlinear device, then a complex power series must be used:

$$D(t) = \sum_{n=1}^{\infty} (x_n + jy_n) v_{in}^n(t) \tag{2.65}$$

where x_n and y_n are constants.

This approach can be shown to be valid, since intermodulation distortion produced by AM/PM conversion is orthogonal to that produced by AM/AM conversion [26,27].

An alternative is the use of two series, one representing the I-characteristic and the other the Q-characteristic. This is effectively a splitting of the terms in (2.65) and results in a model of the form shown in Figure 2.35.

If this approach is used on the class-A and class-C amplifier characteristics presented previously (Figures 2.29 and 2.30), then the characteristics shown in Figures 2.36 and 2.37 result, when terms of up to a seventh-order are used. The resulting models are sufficiently accurate for most purposes and yet are very straightforward to understand and use. Note that u_I and u_Q should be zero (in Figure 2.35), since a properly-functioning amplifier should have an output of zero when no input signal is present.

Note also that in both cases (class-A and class-C), the characteristics and the models extend beyond the 1 dB compression point (see Figures 2.29 and 2.30). In general, this will lead to a poorer model fit than if it were restricted to this range, however, most of the nonlinearity, in the case of the class-C stage, occurs well before the compression point and hence the effect on model accuracy will be relatively small. The models presented do, however, allow 'overload' effects to be modelled and this is useful in assessing the performance of a system under these circumstances.

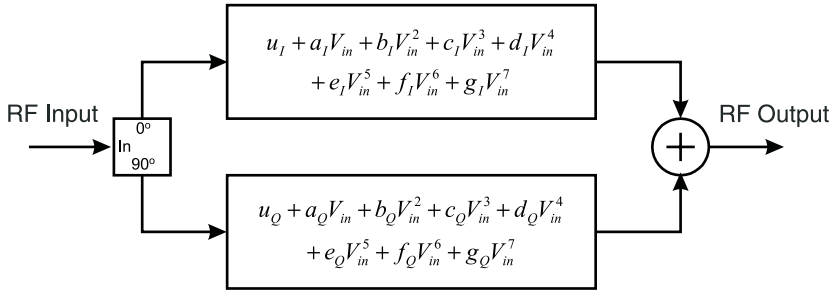
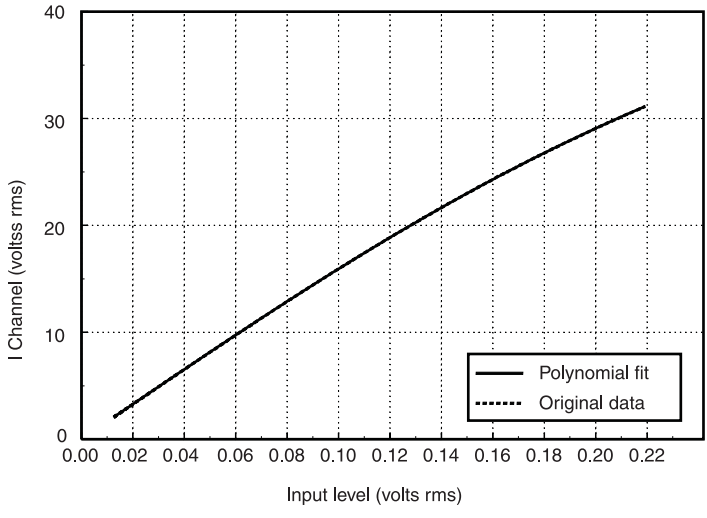
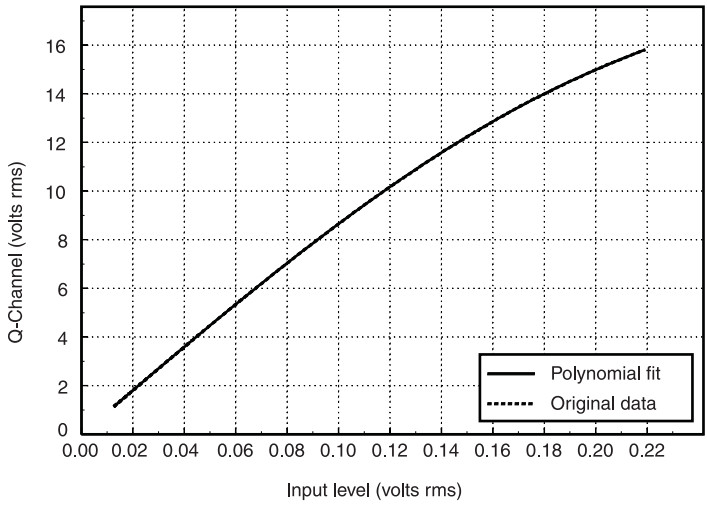


Figure 2.35 Quadrature Taylor series amplifier model.

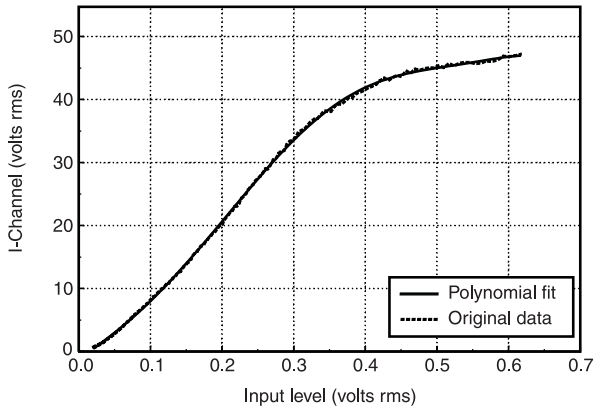


(a)

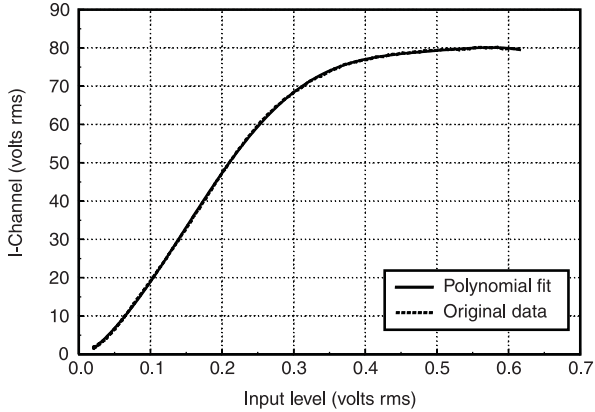


(b)

Figure 2.36 I- and Q-channel characteristics ((a) and (b), respectively) and seventh-order polynomial fit for a 900 MHz 40W class-A amplifier. Fit equations: $V_{OUT,I} = 166.50V_{in} - 109.02V_{in}^2 + 1856V_{in}^3 - 25900V_{in}^4 + 1.53 \times 10^5 V_{in}^5 - 4.63 \times 10^5 V_{in}^6 + 5.75 \times 10^5 V_{in}^7$ $V_{OUT,Q} = 91.80V_{in} - 67.35V_{in}^2 + 1124V_{in}^3 - 17600V_{in}^4 + 1.087 \times 10^5 V_{in}^5 - 3.46 \times 10^5 V_{in}^6 + 4.52 \times 10^5 V_{in}^7$.



(a)



(b)

Figure 2.37 I- and Q-channel characteristics ((a) and (b), respectively) and seventh-order polynomial fit for a 400 MHz 150W class-C amplifier. Fit equations: $V_{OUT,I} = 73.02V_{in} + 159.59V_{in}^2 + 106.8V_{in}^3 + 2270V_{in}^4 - 1750V_{in}^5 + 32700V_{in}^6 - 19100V_{in}^7$ $V_{OUT,Q} = 129.78V_{in} + 654.62V_{in}^2 + 3370V_{in}^3 - 30800V_{in}^4 + 73300V_{in}^5 - 72600V_{in}^6 + 25900V_{in}^7$.

A summary of the quadrature PA model characteristics for the class-A and class-C amplifiers considered above is given in Table 2.5, for a model of the form given in (2.66).

$$V_{OUT,I} = a_I V_{in} + b_I V_{in}^2 + c_I V_{in}^3 + d_I V_{in}^4 + e_I V_{in}^5 + f_I V_{in}^6 + g_I V_{in}^7$$
$$V_{OUT,Q} = a_Q V_{in} + b_Q V_{in}^2 + c_Q V_{in}^3 + d_Q V_{in}^4 + e_Q V_{in}^5 + f_Q V_{in}^6 + g_Q V_{in}^7 \quad (2.66)$$

Table 2.5
Summary of polynomial coefficients for the class-A and class-C quadrature power amplifier models

Amplifier class	Model order	Channel	*a*	*b*	*c*	*d*	*e*	*f*	*g*
Class-A	Cubic	I	165.49	−20.72	−395.87	—	—	—	—
		Q	91.32	−13.34	−340.39	—	—	—	—
	Quintic	I	164.17	−4.64	−356.06	−1182	3641	—	—
		Q	90.58	−8.23	−231.77	−1270	3420	—	—
	Seventh-order	I	166.50	−109.02	1856	-2.59×10^4	1.53×10^5	-4.63×10^5	5.75×10^5
		Q	91.80	−67.35	1124	-1.76×10^4	1.087×10^5	-3.46×10^5	4.52×10^5
Class-C	Cubic	I	86.79	197.45	−353.92	—	—	—	—
		Q	290.025	−67.39	−309.70	—	—	—	—
	Quintic	I	59.99	284.42	127.41	-2.26×10^3	2.25×10^3	—	—
		Q	36.32	2220	−8040	1.06×10^4	-4.90×10^3	—	—
	Seventh-order	I	73.02	159.59	106.80	2.27×10^3	-1.75×10^4	3.27×10^4	-1.91×10^4
		Q	129.78	654.62	3370	-3.08×10^4	7.33×10^4	-7.26×10^4	2.59×10^4

2.13.8.2 Saleh Functions

These functions have been suggested as a simple method of approximating a nonlinear characteristic in either polar or Cartesian form [28]. They have subsequently been applied to the problem of predistortion linearisation of TWT amplifiers and the prediction of carrier-to-intermodulation ratios for multicarrier and Gaussian input signals [29].

The basic concept of Saleh functions is to approximate the AM/AM and AM/PM characteristics of a nonlinearity using two formulas. The constants within these formulas are determined by means of a least-squares curve fit to experimentally-determined AM/AM and AM/PM data. For a polar representation, these functions are:

$$f(r) = \frac{\alpha_a r}{[1 + \beta_a r^2]} \tag{2.67}$$

and

$$g(r) = \frac{\alpha_\theta r^2}{[1 + \beta_\theta r^2]} \tag{2.68}$$

where $\alpha_a,\ \alpha_\theta,\ \beta_a,\ \beta_\theta$ are constants and r is the envelope of the input signal.

The corresponding Cartesian representation is:

$$I(r) = \frac{\alpha_I r}{[1 + \beta_I r^2]} \tag{2.69}$$

and

$$Q(r) = \frac{\alpha_Q r^3}{\left[1 + \beta_Q r^2\right]^2} \tag{2.70}$$

where $\alpha_I,\ \alpha_Q,\ \beta_I,\ \beta_Q$ are constants.

These functions have been shown [29] to provide an accurate, yet simple, model for a bandpass memoryless nonlinearity.

An example of the use of these functions (in polar form) is provided in [30] where the normalised amplitude and phase nonlinearities of a travelling wave tube amplifier are given as:

$$f(r) = \frac{2r}{[1 + r^2]} \tag{2.71}$$

and

$$g(r) = 60^{\circ}\frac{r^2}{1+r^2} \tag{2.72}$$

However, as the nonlinearity of the device increases, the modeling accuracy of Saleh functions decreases. Hence, they are appropriate for quasi-linear amplifiers, such as class-A and class-AB, but are not appropriate for highly nonlinear amplifier classes, such as class-C.

2.13.9 Bandpass Nonlinear Models Incorporating Memory Effects

The instantaneous nonlinear models described in Section 2.13.8 do not account for memory effects within the amplifier (other than AM/PM conversion) and hence are unable to describe the frequency-dependent behavior of the system. Frequency dependent behavior can often be neglected, since the operating bandwidth of most systems is relatively small with respect to the bandwidth capability of the devices used within them. However, in wideband amplifier designs, this assumption is no longer valid and hence frequency dependent models are required.

The frequency-dependence of the amplifier characteristic causes the AM/AM and AM/PM conversion behavior to alter with frequency. It is possible to approximately model this frequency dependent behavior of a nonlinearity with memory and some of the relevant techniques are described below.

2.13.10 Saleh Model

The approximate equations proposed by Saleh and described in Section 2.13.8.2 can be extended to incorporate memory effects [28,29]. This extension does, however, introduce some significant loss of generality [32] since it restricts the shape of the in-phase and quadrature nonlinearities to being independent of frequency; a scaling factor alone provides the frequency dependence. The general shape of the AM/AM and AM/PM characteristics does not alter radically with frequency (the same also being true of the in-phase and quadrature characteristics), however, this loss of generality will still reduce the accuracy which can be obtained using this model.

The model is extended to incorporate frequency-dependent behavior by altering the coefficients α and β in (2.69) and (2.70) to make them a function of frequency. The resulting equations are:

$$I(r) = \frac{\alpha_I(f)r}{[1 + \beta_I(f)r^2]} \tag{2.73}$$

and

$$Q(r) = \frac{\alpha_Q(f)r^3}{\left[1 + \beta_Q(f)r^2\right]^2} \tag{2.74}$$

where $\alpha_I(f)$, $\alpha_Q(f)$, $\beta_I(f)$, $\beta_Q(f)$ are determined by a number of best-fit approximations at a number of frequencies within the band of interest, resulting in an overall best-fit function in each case. These functions can be decomposed into a cascade of three elements: a filter, a memoryless nonlinearity and a further filter. Finally, an additional block, labeled $\phi_0(f)$, is required to model the small-signal phase (the phase which results as $r \rightarrow 0$). The complete model is shown in Figure 2.38.

2.13.11 Blum and Jeruchim Model

A disadvantage with the models described so far is that the interaction effects of adjacent signals on a signal of interest, have not been assessed. These arise due to the measurement method employed in determining the models, namely that of utilising a single frequency tone and stepping it in both power and frequency in order to generate a series of curves. These curves are then the basis for a curve fitting exercise which yields the parameters for the model.

An improved model has therefore been suggested [32] in which the nonlinear device is 'channel-sounded' by a pseudo-random (PN) sequence, binary phase-shift key (BPSK) modulated onto a carrier at the center of the band of interest (thus forming a direct-sequence spread-spectrum signal). A property of this type of modulated carrier is that a number of spectral lines are generated, with a separation determined by the 'chipping-rate' (clock-rate of the PN generator). These spectral lines act like a multicarrier signal and therefore provide the required adjacent signals, assuming that one spectral line is designated as the tone of interest.

The PN modulated carrier signal is applied to the nonlinear device under test and the resulting output signal passed through a narrow bandpass filter. The output of this filter is a measure of the amplitude of a particular frequency component, with the center frequency being swept across the band of interest to generate a complete frequency characteristic. The input power level is also varied, thereby creating a complete series of curves in a similar manner to that required for the Saleh model above. The PN modulated carrier itself is constant envelope (assuming that it is not filtered

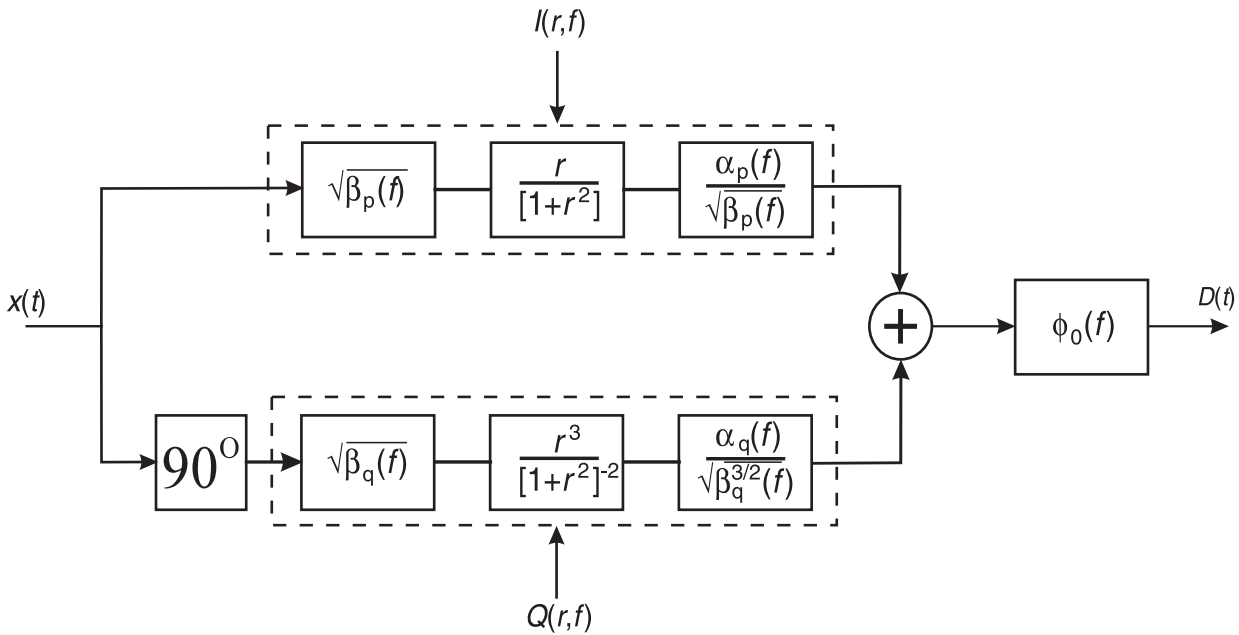


Figure 2.38 Frequency-dependent nonlinear model as proposed by Saleh (after [28]).

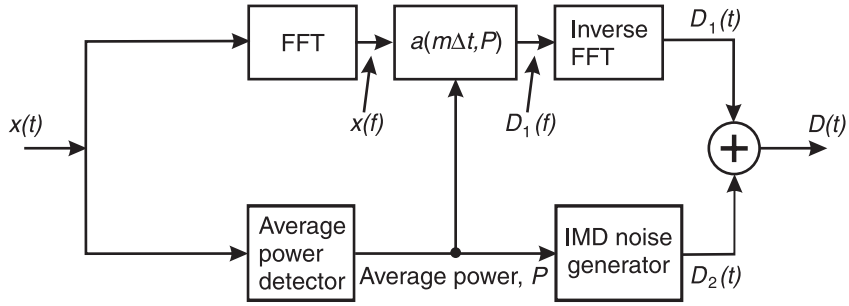


Figure 2.39 Frequency-dependent nonlinear model as proposed by Blum and Jeruchim (after [32]).

before being fed to the input of the device under test) and hence attenuation of its level at the input provides a valid measure of the level dependence of the amplifier's characteristics.

The form of the resulting model is shown in Figure 2.39. The input signal is split, with the upper path feeding a fast-Fourier transform (FFT) to convert the time-domain signal into the frequency domain. The resulting signal is passed through the power-dependent transfer function, with the required average power information supplied from a power measurement algorithm operating on the lower-path split of the input signal. The output of this transfer function is then converted back to the time domain using an inverse FFT to provide the bulk of the modeling process.

It has been acknowledged, however, that the treatment by this model of intermodulation products may be less than ideal and hence an additional block has been added in the lower path. This block takes account of the shortcoming in intermodulation generation by providing a power-dependent intermodulation noise generation function which can then be summed with the output of the inverse FFT. The resulting signal forms the complete model of the nonlinear device.

2.13.12 Volterra Series

A Volterra series [33,34] can be described as a 'Taylor series with memory' [35] and can be used to accurately model an arbitrarily nonlinear system. It does this by describing the distorted output signal, $D(t)$, as an infinite series:

$$D(t) = \sum_{n=0}^{\infty} D_n(t) \tag{2.75}$$

where $D_n(t)$ is the nth order response of the system, formed from an n-fold convolution of the input signal $x(t)$ with the nth-order nonlinear impulse response of the system, $h_n(\tau_1, \ldots, \tau_n)$:

$$D_n(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h_n(\tau_1, \ldots, \tau_n) x(t - \tau_1) \ldots x(t - \tau_n) d\tau_1 \ldots d\tau_n \tag{2.76}$$

The functions $h_n(\tau_1, \ldots, \tau_n)$ are known as the nth-order Volterra kernels of the system and the system can thus be completely characterised by a knowledge of these kernels. Thus, for example, h_0 describes the system response to a DC input signal, h_1 the linear response of the system and the higher order kernels, $h_{2\ldots n}$, the higher-order nonlinearities of the system.

The primary advantage of the Volterra series is in its ability to individually describe all orders of distortion present in the system and to allow comparisons to be made between the different components. It does, however, share a number of disadvantages with the Taylor series, for example that of slow convergence. It also has the disadvantage that computation of the higher-order kernels from measured data is difficult.

Equations (2.75) and (2.76) describe the baseband Volterra series nonlinear model; this can be modified to yield a bandpass model by eliminating the DC term and the even-order terms, since these will yield harmonics of the bandpass signal. The resulting equations are complex and the Volterra kernels difficult to compute, making this form of model unattractive in most circumstances.

2.13.13 Generalised Power Series

It is possible to create a frequency-domain power series representation of the output of a nonlinear system (the Taylor and Volterra series being time-domain representations) and the result is a much more efficient series [36,37] which can deal with severe nonlinearities; it can be shown to be related to the Volterra series [38].

The input signal to a nonlinear system may be described as:

$$x(t) = \sum_{n=1}^{N} |x_n| \cos(\omega_n t + \phi_n) \tag{2.77}$$

where $|x_n|$ are the magnitudes of the individual frequency components, ω_n. The distorted output signal from the nonlinear system, $D(t)$, can then be

described as a generalised power series:

$$D(t) = A\sum_{i=0}^{\infty} a_i \left\{ \sum_{n=1}^{N} b_n x_n(t - \tau_{n,i}) \right\}^i \tag{2.78}$$

where i is the order of the power series, coefficients a_i and b_n are complex and real respectively, and $\tau_{n,i}$ is a time delay term which depends upon frequency and the order of the power series.

The improved modelling efficiency of this technique over the Volterra series representation stems from its operation in the frequency-domain, although it is unclear how to generate the above coefficients from tabulated measured data. The inclusion of frequency-dependence of the time delay terms and the use of complex coefficients in the model allows a very wide range of nonlinearities to be characterised. This is therefore potentially a powerful modelling method, assuming that the relevant coefficients can be obtained.

References

1. De Carvalho, N. B., and J. C. Pedro, "Compact formulas to relate ACPR and NPR to two-tone IMR and IP3," *Microwave Journal*, December 1999, pp. 70–84.
2. Pedro, J. C., and N. B. De Carvalho, "On the use of multi-tone techniques for assessing RF components' intermodulation distortion," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 47, No. 12, December 1999, pp. 2393–2402.
3. Schetzen, M, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley and Sons, New York, 1980.
4. Carvalho, N. B. and J. C. Pedro, "Multi-tone intermodulation distortion performance of 3rd order microwave circuits," *IEEE Symposium on Microwave Theory and Techniques*, Anaheim, USA, Vol. 2, June 1999, pp. 763–766.
5. Boyd, S., "Multitone signals with low crest factor," *IEEE Trans. on Circuits and Systems*, Vol. 33, October 1986, pp. 1018–1022.
6. Bauml, R. W., R. F. H. Fischer and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *IEE Electronics Letters*, Vol. 32, October 1996, pp. 2056–2057.
7. Rudin, W., "Some theorems on fourier coefficients," *Proc. of the American Mathematics Society*, Vol. 10, December 1959, pp. 855–859.
8. Greestein, L. J., and P. J. Fitzgerald, "Phasing multitone signals to minimize peak factors," *IEEE Trans. on Communications*, Vol. 29, July 1981, pp. 1072–1074.
9. Golay, M. J. E., "Complementary series," *IRE Trans. on Information Theory*, Vol. 7, April 1961, pp. 82–87.
10. Popovic, B. M., "Synthesis of power efficient multitone signals with flat amplitude spectrum," *IEEE Trans. on Communications*, Vol. 39, July 1991, pp. 1031–1033.
11. Tellambura, C., "Use of m-sequences for OFDM peak-to-average power ratio reduction," *IEE Electronics Letters*, Vol. 33, July 1997, pp. 1293–1294.

12. Barker, R. H., "Group synchronizing of binary digital systems," *Communication Theory*, 1953, pp. 273–287.
13. Golomb, S. W., and R. A. Scholtz, "Generalized Barker sequences," *IEEE Trans. on Information Theory*, Vol. 4, October 1965, pp. 533–537.
14. Newman, D. J., "An L^1 extrimal problem for polynomials," *Proc. of the American Mathematics Society*, Vol. 16, December 1965, pp. 1287–1290.
15. Schroeder, M. R., "Synthesis of low peak-factor signals and binary sequences with low autocorrelation," *IEEE Trans. on Information Theory*, 1970, pp. 85–89.
16. Jones, A. E., T. A. Wilkinson, and S. K. Barton, "Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission systems," *IEE Electronics Letters*, Vol. 30, December 1994, pp. 2098–2099.
17. Muller, S. H., and J. B. Huber, "OFDM with reduced peak-to-mean power ratio by optimum combination of partial transmit sequences," *IEE Electronics Letters*, Vol. 33, February 1997, pp. 368–369.
18. Choi, B-J, E-L Kuan, and L. Hanzo, "Crest factor study of MC-CDMA and OFDM," *Proc. IEEE Vehicular Technology Conference, Fall'99*, Amsterdam, The Netherlands, Vol. 1, 19–23 September 1999, pp. 233–237.
19. Hardy J. K., *High Frequency Circuit Design*, Chapter 1, Reston Publishing Company, Reston, Virginia, USA, 1979.
20. Katz, A., "TWTA Linearization," *Microwave Journal*, April 1996.
21. Krauss, H. L., C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, Chapter 12, John Wiley & Sons, New York, USA, 1980.
22. Minkoff, J. B., "Intermodulation noise in solid-state power amplifiers for wideband signal transmission," *Proc. AIAA 9th Communications Satellite System Conference*, March 1982.
23. Minkoff, J. B., "Wideband operation of nonlinear solid-state power amplifiers—comparisons of calculations and measurements," *Bell System Technical Journal*, Vol. 63, No. 2, February 1984, pp. 231–248.
24. Kuo, Y., "Noise loading analysis of a memoryless nonlinearity characterised by a Taylor series of finite order," *IEEE Trans. on Instrumentation and Measurement*, Vol. IM-22, September 1973, pp. 246–249.
25. Larkin, R., "Multiple signal intermodulation and stability considerations in the use of linear repeaters," *Proc. IEEE 41st Vehicular Technology Conference*, St. Louis, Missouri, USA, May 1991, pp. 747–752.
26. Shimbo, O., "Effects of Intermodulation, AM–PM conversion and additive noise in multicarrier TWT systems," *IEEE Proc.*, Vol. 59, February 1971, pp. 230–238.
27. Sechi, F., "Linearised class-B transistor amplifiers," *IEEE Journal of Solid-State Circuits*, Vol. SC-11, No. 2, 1976, pp. 264–270.
28. Saleh, A. A. M., "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. on Communications*, Vol. COM-29, November 1981, pp. 1715–1720.
29. Saleh, A. A. M., "Intermodulation analysis of FDMA satellite systems employing compensated and uncompensated TWT's," *IEEE Trans. on Communications*, Vol. COM-30, No. 5, May 1982, pp. 1233–1242.
30. Saleh, A. A. M., and J. Salz, "Adaptive linearization of power amplifiers in digital radio systems," *The Bell System Technical Journal*, Vol. 62, No. 4, April 1983, pp. 1019–1033.
31. Jeruchim, M., P. Balaban, and K. Shanmugan, *Simulation of Communications Systems*, New York: Plenum Press, 1992, Chapter 2.

32. Blum, R., and M. Jeruchim, "Modeling nonlinear amplifiers for communication simulation," *Proc. IEEE International Conference on Communications*, Boston, USA, June 1989, pp. 1468–1472.
33. Narayanan, S., "Transistor distortion analysis using Volterra series representation," *Bell System Technical Journal*, Vol. 46, May/June 1967, pp. 991–1024.
34. Law, C., and C. Aitchison, "Prediction of wide-band power performance of MESFET distributed amplifiers using the Volterra series representation," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-34, December 1986, pp. 1308–1317.
35. Jeruchim, M., P. Balaban, and K. Shanmugan, *Simulation of Communications Systems*, New York: Plenum Press, 1992, Chapter 2.
36. Steer M., and P. Khan, "An algebraic formula for the output of a system with large-signal, multifrequency excitation," *Proc. of the IEEE*, Vol. 71, January 1983, pp. 177–179.
37. Rhyne, G., M. Steer, and B. Bates, "Frequency-domain nonlinear circuit analysis using generalized power series," *IEEE Trans. on Microwave Theory and Techniques*, Vol. MTT-36, February 1988, pp. 379–387.
38. Steer, M., P. Khan, and R. Tucker, "Relationship between Volterra series and generalised power series," *Proc. of the IEEE*, Vol. 71, December 1983, pp. 1453–1454.