INTERMODULATION MEASUREMENTS

NONLINEAR DISTORTION MEASUREMENT

TRANSIENT INTERMODULATION MEASUREMENT

Virtually all electronic circuits and systems exhibit nonlinear input-output transfer characteristic. Mixers, frequency multipliers, modulators and square-law detectors represent examples of intentional class members while linear power amplifiers, active filters and microwave transmitters, in which nonlinearity represents an undesirable deviation of the system from ideal, or linear, operation, are examples of unintentional members.

Whenever a number of signals of differing frequencies pass through a nonlinear device energy is transferred to frequencies that are sums and differences of the original frequencies. These are the intermodulation products (IMPs). In such cases, the instantaneous level of one signal may effectively modulate the level of another signal, hence the name intermodulation. In a transmitting system, the results of excessive intermodulation are unwanted signals that may cause interference. In a receiver, internally-generated intermodulation can hinder reception of the desired signals. It is interesting to note that the ear's cochlea has a similar non-linear response and produces sums and differences of the input frequencies in the same way particularly with loud sounds [1].

It has also been found that passive components, normally considered to be linear, can also generate IMPs. A variety of situations can arise in which nonlinear resistance junctions can be formed at metallic mating surfaces. Such junctions may result from salt or chemical depositions or from corrosion. The result is sometimes known as the "rusty bolt effect" because rusted bolts in structures have been known to exhibit such nonlinearities. This phenomenon is referred to as passive intermodulation (PIM). Sources of PIM include waveguides, irectional couplers, duplexers and antennas [2, 6].

Intermodulation may also occur at the amplifier-loudspeaker interface [7], or in general due to the nonlinear interaction between the input signal of a two-port and a signal injected to the output port and propagating into the input via a feedback network [8]. Externally-induced transmitter intermodulation, also known as reverse intermodulation, back intermodulation, and antenna-induced intermodulation, is the mixing of a carrier frequency with one or more interferring signals in a transmitter's final stage [9]. Moreover, lack of screening of open-wire transmission lines can result in significant coupling to adjacent lines frequently giving rise to intermodulation products [10]. Furthermore, intermodulation may arise when an array of receiving antennas is illuminated with a transient impulsive electromagnetic plane wave [11].

In discussing the sources of IMPs it is convenient to divide nonlinear mechanisms yielding IMPs into two principal forms. The first is due a nonlinear amplitude input/output characteristic (AM/AM), which causes amplitude compression with increasing input amplitude. The second mechanism occurs because of the variation of phase shift through the device, or the system, as the input amplitude is changed (AM/PM).

Depending on the signal characteristics, sources of IMPs can be divided into two categories: a. Static nonlinearity, depending solely on the amplitude of the signal, and b. Dynamic nonlinearity, depending not only on the amplitude but also on the time properties or frequency composition of the signal.

Static nonlinearities usually encountered in electronic circuits and systems can be classified into clipping, crossover and soft nonlinearities [12] as shown in Fig. 1. Among the hard nonlinearities of clipping (which is significant near maximum input amplitudes) and crossover (mostly significant at small input amplitudes), the soft nonlinearity is usually the most important in the transfer characteristic of an electronic circuit. If the frequency content or the time properties of the input signal affect the transfer characteristic of the circuit or the system, the resulting nonlinearities may be called dynamic. Intermodulation products resulting from dynamic nonlinearities are referred to as transient intermodulation (TIM), slew induced distortion (SID) or dynamic intermodulation distortion (DIM) [13, 16]

SIMPLE INTERMODULATION THEORY

IMPs occur when two or more signals exist simultaneously in a nonlinear environment. In general, if N signals, with frequencies f_1 to f_N are combined in a static nonlinearity, the output will contain spectral components at frequencies given by

$$n = 1 \sum k_n \ f_n$$

where k_n is a positive integer, a negative integer, or zero, and $n = 1 \sum |k_n|$ is the order of the IMP. Even with a small number of input signals, N, a very large number of IMPs are generated. Fortunately, not all products are equally troublesome. Depending on the system involved, some of these IMPs can be neglected since they will be filtered out at some point. For example, most of the communication systems operate over a limited frequency band. Thus, IMPs falling out of the band will be attenuated. Moreover, amplitudes of the IMPs generally decrease with the order of the products and high order products can often be neglected. Low order intermodulation components such as the second-order component $f_m - f_n$ and $f_m + f_n$ the third-order components occuring at frequencies $2 f_m - f_n$ and $f_m + f_n - f_q$ are usually the most troublesome, having the largest magnitudes and/or lying close to the originating frequencies, making their removal by filtering practically difficult. However, a salient characteristic of PIM, as distinguished from the conventional IM counterpart, discussed in the preceding, is that the PIMs causing trouble are of a high order, say eleventh through the twenty first.

Analysis of non-linear systems differs from that of linear systems in several respects: there is no single analytical approach which is generally applicable (such as Fourier or Laplace transforms in linear systems); closed-form ana-

J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*. Copyright © 2007 John Wiley & Sons, Inc.

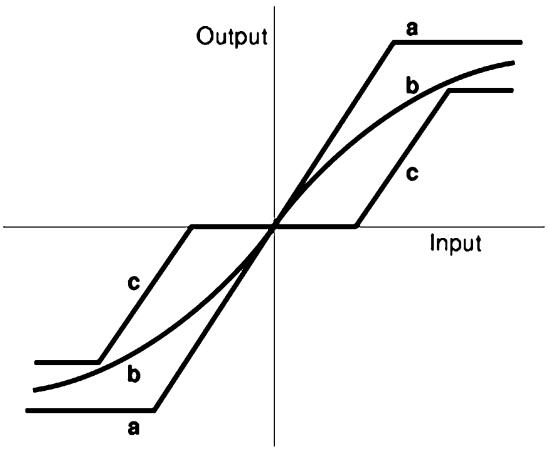

Figure 1. Different types of static nonlinearities: (a) clipping, (b) soft, (c) crossover.

lytical solutions of non-linear equations are not ordinarily possible; and there is rarely sufficient information available to enable a set of equations which accurately model the system to be derived. These factors preclude the exact analytical determination of non-linear effects, such as IMPs, in the general case. In order to get anything done at all it is usually necessary to make various simplifying assumptions and then to use an approximate model which will provide results of acceptable accuracy for the problem in hand.

A simple approach, therefore, is to use frequency domain techniques which provide a separate solution for each frequency present in the output. In general, such methods are: (a) centered around a description of the non-linear mechanism by a continuous function type of characteristic; for example a polynomial or a Fourier-series representation of the output in terms of the input, and (b) based on the simplifying assumption that this characteristic does not vary with frequency; that is a memoryless characteristic.

Memoryless nonlinear circuits are oftenly modeled with a power series of the form

$$V_{out} = n = 0 \sum k_n V_i^n \tag{1}$$

The first coefficient, k_O, represents the DC offset in the circuit. The second coefficient, k_1, is the gain of the circuit associated with linear circuit theory. The remaining coefficients, k_2 and above, represent the nonlinear behavior of the circuit. If the circuit were completely linear, all of the coefficients except k_1 would be zero.

The model can be simplified by ignoring the terms that come after the k_3 term. For soft nonlinearities, the size of k_n decreases rapidly as n gets larger. For many applications the reduced model of Eq. (2) is sufficient, since the second-order and third-order effects dominate. However, there are many devices, circuits and systems presenting difficulties for the polynomial approximation.

$$V_{out} = k_o + k_1 V_i + k_2 V_i^2 + k_3 V_i^3 \tag{2}$$

Assuming that the input signal is a two-tone of the form

$$V_i = V_1 \cos\omega_1 t + V_2 \cos\omega_2 t \tag{3}$$

then combining Eqs. (2) and (3), yields

$$\begin{aligned} V_{out} &= a_o + b_1\cos\omega_1 t + c_1\cos\omega_2 t \\ &+ b_2\cos 2\omega_1 t + c_2\cos 2\omega_2 t + b_3\cos(\omega_1+\omega_2)t + c_3\cos(\omega_1-\omega_2)t \\ &+ b_4\cos 3\omega_1 t + c_4\cos 3\omega_2 t + b_5(\cos(2\omega_1+\omega_2)t + \cos(2\omega_1-\omega_2)t) \\ &+ c_5(\cos(2\omega_2+\omega_1)t + \cos(2\omega_2-\omega_1)t) \end{aligned} \tag{4}$$

where

$$a_o = k_o + \frac{k_2}{2}(V_1^2 + V_2^2)$$

$$b_1 = k_1 V_1 + \frac{3}{4}k_3 V_1^3 + \frac{3}{2}k_3 V_1 V_2^2$$

$$c_1 = k_1 V_2 + \frac{3}{4}k_3 V_2^3 + \frac{3}{2}k_3 V_1^2 V_2$$

$$b_2 = \frac{1}{2}k_2 V_1^2$$

$$c_2 = \frac{1}{2}k_2 V_2^2$$

$$b_3 = c_3 = k_2 V_1 V_2$$

$$b_4 = \frac{1}{4}k_3 V_1^3$$

$$c_4 = \frac{1}{4}k_3 V_2^3$$

$$b_5 = \frac{3}{4}k_3 V_1^2 V_2$$

$$c_5 = \frac{3}{4}k_3 V_1 V_2^2$$

For equal amplitude input tones, Eq. (4) shows that the second-order terms, of amplitudes b_2, c_2, b_3 and c_3 will be increased 2 dB in amplitude when input tones are increased by 1 dB. The third-order terms, of amplitudes and b_4, c_4, b_5, are increased by 3 dB in amplitude when the input tones are increased by 1 dB.

While equation (1) is adequate, and widely used, to predict the intermodulation performance of a wide range of devices, circuits and systems, sometimes it cannot be

used. Examples include, but are not restricted to, prediction of spectral regrowth in digital communication systems, transient intermodulation and frequency-dependent nonlinearities, and passive intermodulation.

SPECTRAL REGROWTH

When a modulated signal passes through a nonlinear device, its bandwidth is broadened by odd-order nonlinearities. This phenomenon, called spectral regrowth or spectral regeneration, is a result of mixing products (intermodulation) between the individual frequency components of the spectrum [17]. The spectral regrowth can be classified in the two following categories: (1) in band intermodulations and (2) out band intermodulations. The first cannot be eliminated by linear filtering and they are responsible for the signal-to-noise ratio degradation and, consequently, for the bit error rate (BER) degradation in digital communication systems. The second generates the interference between adjacent channels and they can be filtered out at the nonlinear device output with certain output power penalty that is caused by the filter insertion losses. This spectral regrowth causes adjacent channel interference (ACI) which is measured by the adjacent channel power ratio (ACPR).

The ACPR is the power in the main channel divided by the power in the lower plus upper adjacent channels. Considering just the lower channel yields $ACPR_{LOWER}$ and the upper channel alone yields $ACPR_{UPPER}$. Analog cellular radio uses frequency or phase modulation, and the ACPR is adequately characterized by intermodulation distortion of discrete tones. Typically, third-order intermodulation product (IMP3) generation, in a two-tone test, is adequate to describe spectral regrowth. Thus, distortion in analog radio is accurately modeled using discrete tone steady-state simulation. Digital radio, however, uses complex modulation, and adjacent channel distortion has little relationship to intermodulation in a two-tone test [18],[19]. A modulated input signal applied to radio-frequency (RF) electronics in digital radio is a sophisticated waveform resulting from coding, filtering, and quadrature generation. Neither can it be represented by a small number of discrete tones (or frequencies), nor can the waveform be represented in a simple analytic form. Thus, in digital radio, ACPR is more difficult to predict than one- or two-tone responses since it depends not only on the intrinsic nonlinear behavior of the device (e.g. amplifier), but also on the encoding method (i.e. the statistics of the input stream) and the modulation format being used. The only way the input stream can conveniently and accurately be represented is by its statistics, and transforming these using an appropriate behavioral model provides accurate and efficient modeling of ACPR [20]. While in reference [20] the input signal is assumed Gaussian, digital communication signals are often far from being Gaussian. In reference [21] the input is assumed stationary but not necessarily Gaussian.

ACPR is, therefore, defined differently in the various wireless standards. The main difference being the way in which adjacent channel power affects the performance of another wireless receiver for which the offending signal is cochannel interference [20]. In general the ACPR can be defined as [20]

$$ACPR = \frac{f_3 \int S(f) d f}{f_1 \int S(f) d f} \tag{5}$$

where $S(f)$ is the power spectral density (PSD) of a signal whose channel allocation is between frequencies f_1 and f_2, and its adjacent channel occupies frequencies between f_3 and f_4. Regulatory authorities impose strict constraints on ACPR and accurate methods of its determination are of particular interest to those involved in wireless system design.

SIMPLE TRANSIENT INTERMODULATION THEORY

To illustrate how TIM distortion arises, consider a differential amplifier with negative feedback applied between the output and the inverting input and a voltage step applied to the noninverting input. If the open-loop gain of the amplifier were flat and the time delay through it were zero, the voltage step would instantaneously propagate undistorted through the amplifier, back through the feedback loop, and into the inverting input. There it would be subtracted from the input signal, and the difference signal, which is a voltage step occuring at the same time that the input voltage does, would be amplified by the amplifier. However, this is not the case when the open-loop gain of the amplifier is not flat and the time delay through it is not zero. When the voltage step occurs, the limited high-frequency response of the amplifier prevents the appearance of a signal at the amplifier output terminal until the internal capacitors of the amplifier can charge or discharge. This causes the momentary absence of a feedback signal at the inverting input to the amplifier, possibly causing the amplifier to severely overload until the feedback signal arrives.

If the input signal, to the differential amplifier, is formed of a sine wave superimposed on a square wave, the amplifier will exhibit the same response to the abrupt level changes in the square wave as it did to the voltage step discussed in the preceding. During the momentary absence of the feedback when the square wave changes level, the amplifier can either saturate or cutoff. If this occurs, the sine wave momentarily disappears from the signal at the output terminal of the amplifier, or it momentarily decreases in amplitude. This happens because the saturated or cutoff amplifier appears as a short circuit or open circuit, respectively, to the sine wave, and this component of the input signal is interrupted from the output signal. Thus, resulting in TIM [16].

A point to be noted is that if the term were understood literally, this would imply transients of both high and low frequencies and/or high or low operating levels. In other words, all transients. In actual practice, however, TIM occurs only for signals with simultaneous high level and high frequencies-not lower levels or lower frequencies. The key parameter of such signals is that they are characterized by high signal slopes, not just high frequencies or high levels. Neither high frequencies nor high levels in themselves necessarily result in distortion, unless their combination is such that a high effective signal slope is produced. TIM is actually generated when the signal slope approaches or

exceeds the amplifier slew rate. This can happen for either transient or steady-state signals. Thus, a more easily understood term to what actually happens would be one which relates both slew rate and signal slope. A more descriptive term to describe the mechanism would, therefore, be the Slew Induced Distortion (SID). Other descriptive variations of the terminology are "slew rate distortion" or "slewing distortion" [22].

Because of the complexity of the mechanism resulting in TIM, especially handling the frequency dependence of the amplifier nonlinearity and the incorporation of the feedback, equation (1) can not be used to predict the TIM performance of nonlinear devices and recourse to other analytical techniques, for example Volterra-series or harmonic balance analysis, would be inevitable.

VOLTERRA-SERIES AND HARMONIC BALANCE ANALYSIS

Volterra series describes a system with frequency-dependent nonlinearity in a way which is equivalent to the way Taylor series approximates an analytic function. Depending on the amplitude of the exciting signal, a nonlinear system can be described by a truncated Volterra series. Similar to the Taylor series representation, for very high amplitudes the Volterra series diverges. Volterra series describe the output of a nonlinear system as the sum of the response of a first-order operator, a second-order one, a third-order one and so on [23]. Every operator is described either in the time domain or in the frequency domain with a kind of transfer function, called a *Volterra kernel*.

In Volterra-series analysis the nonlinear circuit is treated purely as an AC problem. Assuming that none of the input signals are harmonically related, an iterative solution can be applied for circuits not operated under distortion saturation conditions. First the circuit is solved for the input signals. These results are then used to calculate the second-order distortion products and these are treated as generators at a different frequency to the input signals and the network again solved. This is then repeated for higher order distortion products. This leads to extremely fast calculation of distortion behavior. Simulation at higher power levels can be achieved by feeding back contributions from higher order distortion products [24]. The use of Volterra series to characterize the output as a function of the input [25], [26] can, therefore, provide closed-form expressions for all the distortion products of a frequency-dependent nonlinearity excited by a multisinusoidal signal.

However, techniques using Volterra series suffer from the disadvantage that a complex mathematical procedure is required to obtain a closed-form expression for the output amplitude associated with a single component of the output spectrum. Moreover, the problem of obtaining output products of orders higher than the third becomes prohibitively difficult unless it may be assumed that higher-order contributions vanish rapidly [27]. The Volterra series approach is, therefore, most applicable to mild non-linearities where low order Volterra kernels can adequately model the circuit behavior. With appropriate assumptions and simplifications, many useful features of the Volterra series technique can be used to find approximate expressions for TIM (SID). These are quite accurate for relatively small distortion conditions [28] and [29].

Alternatively, most RF and microwave circuit-analysis are based on the harmonic-balance analysis [30]. The Harmonic Balance technique works by processing the linear part of the circuit in the frequency domain and the nonlinear part in the time domain. Computation in the frequency domain is very fast and efficient especially for frequency selective components such as transmission lines and resonant circuits. Computations in the time-domain are followed by Fourier transform. Harmonic balance analysis can, therefore, handle intermodulation distortion provided that there are not too many excitation tones. In the harmonic balance technique an initial estimate is required for the final wave shape, and this is refined interactively during analysis. The harmonic balance method computes the response of a nonlinear circuit by iteration, and the final result is a list of numbers which do not indicate which nonlinearities in the circuit are mainly responsible for the observed nonlinear behavior. Hence such method is suitable for verification of circuits that have already been designed. This method does not present information from which designers can derive which circuit parameters or circuit elements they have to modify in order to obtain the required specifications [31]. While Volterra-series analysis can provide such information, it is applicable only to weak nonlinearities.

While viewed as a universal solution, and has been widely used, the harmonic balance analysis may be unnecessarily slow, cumbersome, and prone to subtle errors [32] especially for weak nonlinreaties or when a nonlinear device is excited by very small signals. Volterra-series analysis is generally more accurate than harmonic balance for these types of problems and it is several orders of magnitude faster than a harmonic-balance analysis [32]. Moreover, Volterra-series analysis integrates well with linear analysis tools, supporting simultaneous optimization of several parameters of the nonlinear system. Therefore, Volterra theory appears to be ideal tool for circuits and systems that are not strongly nonlinear, but have aspects of linear and nonlinear circuits [32]. However, Volterra-series analysis becomes very cumbersome above third-order products, and for products above fifth order, it loses most of its advantages over the harmonic balance analysis. The major disadvantage of Volterra series is the occasional difficulty in deciding whether the limitations to weakly nonlinear operation have been exceeded or not.

In fact Volterra-series analysis and the harmonic balance technique complement each other [32]. Thus, while the Volterra-series analysis works well in those cases where harmonic-balance works poorly, the harmonic-balance works well where the Volterra-series works poorly. Volterra-series analysis is, therefore, not appropriate for mixers, frequency multipliers, saturated power amplifiers and similar strongly driven and/or hard nonlinearities. Volterra-series analysis is suitable for small-signal amplifiers, phase shifters, attenuators and similar small signal and/or soft nonlinearities.

Another technique for analyzing nonlinear systems is the describing function. This approach can yield closed-

form expressions for a feedback system that contains an isolated static nonlinearity in the feedback loop [33]. Since it is not possible to map all nonlinear circuits and systems to such a feedback system, the describing function method has restricted applications.

PASSIVE INTERMODULATION (PIM)

While the concept of intermodulation in active devices such as amplifiers, filters and mixers is familiar and well documented, the effects of intermodulation in passive components such as directional couplers, cables, coaxial connectors, power splitters, antennas, and electromechanical and solid-state programmable attenuators are less familiar and less documented. More recently, evidence has emerged that PIM has an impact in other system equipment, such as amplifiers and extenders, fiber nodes and interface units [34]. Poor mechanical contact, dissimilar metals in direct contact, ferrous content in the conductors, debris within the connector, poor surface finish, corrosion, vibration, and temperature variations are among the many possible causes of PIM. The sources of PIM have been studied extensively; see [35–43] and the references cited therein. Similar to the intermodulation products in active devices, PIM is generated when two or more RF signals pass through RF passive devices having nonlinear characteristics [41],[42]. Generally the nonlinearities of RF passive devices consist of contact nonlinearity and material nonlinearity [43]. Contact nonlinearity refers to all metal contact nonlinearities causing nonlinear current-voltage behavior, such as tunneling effect, micro-discharge, and contact resistance. Material nonlinearity refers to the bulk material itself. Magneto-resistivity of the transmission line, thermal resistivity, and non-linear hystresis of ferromagnetic material are good examples [43]. PIM generation in RF passive devices is caused by the simultaneous appearance of one or more of these PIM sources, and the overall performance is often dominated by one principal PIM source [43]. In the case of antennas, PIM is generated not only by the same PIM sources as in general RF passive components but also by the external working environment, such as conducting metal materials.

Over the years equation (1) was used to describe the nonlinear current/voltage conduction characteristics of passive components; see for example references [37]–[39] and the references cited therein. While this approach results in simple expressions for the magnitudes of the harmonics and intermodulations products resulting from multisinusoidal excitations, it suffers from the following shortcomings.

In order to predict high order harmonic or intermodulation product magnitudes it is necessary to determine coefficients of terms of similar order in the polynomial. A prerequisite to obtaining the coefficients of the terms of high order polynomials, is the measurement of output products of the same order. For example, to obtain the coefficients of a fifth-order polynomial, it is necessary to measure the output fifth-order components. With increasing use of narrow band components in multicouplers used in base stations of mobile radio systems, it becomes difficult to determine high order coefficients in the nonlinear characteristic because the measured high order product amplitudes from which they are computed are influenced to an unknown extent by the system selectivity [44]. To overcome these problems, an exponential method has recently been used to predict the intermodulation arising from corrosion [45].

ALIAS-INTERMODULATION DISTORTION

Alias-intermodulation distortion is a relatively recently discovered form of distortion. It is a subtle but apparently very common form of distortion inherent in most digital recording/replay systems and is a result of two simultaneous nonlinear mechanisms [46]. Converting audio signals from the analog domain to the digital domain requires sampling and quantization. In the sampling process, the analog signal is sampled with a frequency equal to or greater than twice its bandwidth. By itself, this process is nonlinear and will produce a spectrum of frequencies mirrored around the sampling frequency. These samples are then quantized. In the quantization process, the exact value of the sampled analog signal is replaced by one value taken from a finite number of quantized values. Thus information consisting of a finite number of values is substituted for the possible infinite number of values of the analog signal. This quantization process is inherently nonlinear and if the processed signal is a multisinusoidal signal, then harmonics and intermodulation products will be produced. Thus, if the original analog signal is formed of a fundamental and its harmonics, and if the frequency of one or more of these harmonics is greater than half the sampling frequency, then because of the sampling process new frequencies will be generated within the bandwidth of the original signal. Usually, these new frequencies will result from the sampling of the higher harmonics that are close to half the sampling frequency. By themselves, these new frequencies may be considered not harmful as they usually occupy the high end of the audio spectrum. This phenomenon is referred to as aliasing distortion [46].

However, because of the inherent nonlinearity of the quantization process, these new frequencies will intermodulate with the audible components of the signal (usually occupying the lower end of the spectrum) and may produce new audible components. These new components are not harmonically related to the original signal and will, therefore, degrade the signal quality. This kind of distortion is called Aliasing Intermodulation Distortion (AID). This AID may be enhanced later on when the output of the digital to analog converter is applied to electronic amplifiers and/or electromechanical transducers. In fact the inherent nonlinearity of these components will inevitably result in new intermodulation components and thus further degradation of the quality of the signal.

INTERMODULATION CHARACTERIZATION

Although it is important to understand the origin of intermodulation and the engineering techniques for avoiding it, it is equally important to be able to characterize it objectively, preferably in a way that correlates well with the

subjective perception of the intermodulation. Being able to characterize an imperfection in this way is an important step toward eliminating it as a system performance degradation.

Several techniques for characterizing intermodulation distortion have been proposed. While some of these techniques measure the total intermodulation distortion, others distinguish between the various intermodulation products. The latter are much to be preferred, for subjective perception of intermodulation shows that equal amounts of total intermodulation disortion differ widely in their effect according to how the total is made up.□□

Depending on the signal characteristics, techniques for characterization of intermodulation distortion can be classified into two categories:□(a) Steady-state techniques, where characterization is performed on the assumption that the input to the system under consideration is a multisinusoidal signal, and □(b) Dynamic techniques, where characterization is performed on the assumption that the input to the system under consideration is formed of a sinusoidal signal superimposed on another signal characterized by rapid changes of state; for example a square wave or a sawtooth wave. While steady-state techniques can be used for characterizing both RF and audio systems, dynamic techniques are usually used for characterizing only audio systems.

Steady-State Techniques

The Intercept Point. Increasing the signal level at the input to a weakly nonlinear device will cause the IMPs to increase at the output (47). In fact, the increase in the amplitudes of the IMPs is faster than the increase in the output version of the input signal. For increasing fundamental input power, the fundamental output power increases in a linear manner, according to the gain or loss of the device. At some point, gain compression occurs and the fundamental output power no longer increases with input power. The output power of the second-order intermodulation products also increases with fundamental input power, but at a faster rate. Recall that, according to the simple intermodulation theory, the second-order intermodulation changes 2 dB per 1 dB of change in the fundamental. Similarly, the third-order intermodulation changes 3 dB per 1 dB of change in the fundamental. Thus, on a logarithmic scale, as shown in Fig. 2, the lines representing the second- and third-order intermodulation products have twice and three times, respectively, the slope of the fundamental line.□

If there was no gain compression, the fundamental input power could be increased until the second-order intermodulation would eventually catch up with it and the two output power levels would be equal. This point is referred to as the second-order intercept point (IP2). The third-order intermodulation product also increases faster than the fundamental, and those two lines will intersect at the third-order intercept point (IP3). Rarely can either of these two points be measured directly, due to the gain compression of the fundamental. Instead, the intercept points are extrapolated from measurements of the fundamental and intermodulation products at power levels below where gain compression occurs. The intercept points are usually specified in dBm and may refer either to the output or to the input; the two points will differ by the gain of the system under consideration. The second-order and third-order intercept points are figures of merit which are independent of the signal level. Therefore, the intermodulation performance of two different systems can be compared quite easily if their intercept points are known [47].□□

Using the intercept point it is easy to calculate the relative intermodulation level corresponding to a given input signal level. In fact, the difference between the level of the second-order intermodulation and the fundamental signal level is the same as the difference between the fundamental signal level and the intercept point. Thus, if the second-order intercept point is +15 dBm and the fundamental signal level is −10 dBm (both referred to the output of the device), the difference between these two values is 25 dB. Therefore, the second-order intermodulation products will be 25 dB below the fundamental, or −35 dBm. So the intercept point allows easy conversion between fundamental signal level and the intermodulation level. □□

The difference between the level of the third-order intermodulation products and the fundamental signal level is twice the difference between the fundamental signal level and the third-order intercept point. (Note that the second-order intercept point is not the same as the third-order intercept point.) Suppose that the third-order intercept point is +5 dBm and the fundamental signal is −25 dBm, both referred to the output of the device. The difference between the intercept point and the fundamental is 30 dB, so the third-order intermodulation products will be two times 30 dB down from the fundamental. The relative distortion level is −60 dB and the absolute power of the intermodulation products is −85 dBm.□□

It is important, however, to note that the preceding analyses assume that the second-order and the third-order intermodulation curves have slopes of 2dB/dB and 3 dB/dB respectively. Thus, theoretically, the intercept points are not functions of the input power level. If a power sweep is performed, it is expected that the intercept points will remain constant. The intercept points can, therefore, be calculated from measurements at only one power level. However, if the input signal exceeds a certain limit, the amplitudes of the output fundamentals and the resulting intermodulation products will start to saturate, and the intercept points will usually drop off, indicating an invalid measurement. It is essential to know this limit. It is particularly useful for high dynamic range circuits and systems with relatively low output powers where the intermodulation is low, but only for signals that are low enough. Expanding the model of Eq. (2) to include fourth- and fifth-order terms [48] can do this.

Moreover, at the low power levels, the intercept points will start to change as the noise floor of the measuring instrument; usually a spectrum analyzer, is approached. Thus, indicating an invalid measurement. It is important, therefore, to look at the variation of the intercept-points as functions of power as this provides a good way of checking the valid measurement range.

Two-Tone Test. □□The two-tone test is extensively used in characterizing a wide range of devices. Magnetic tapes

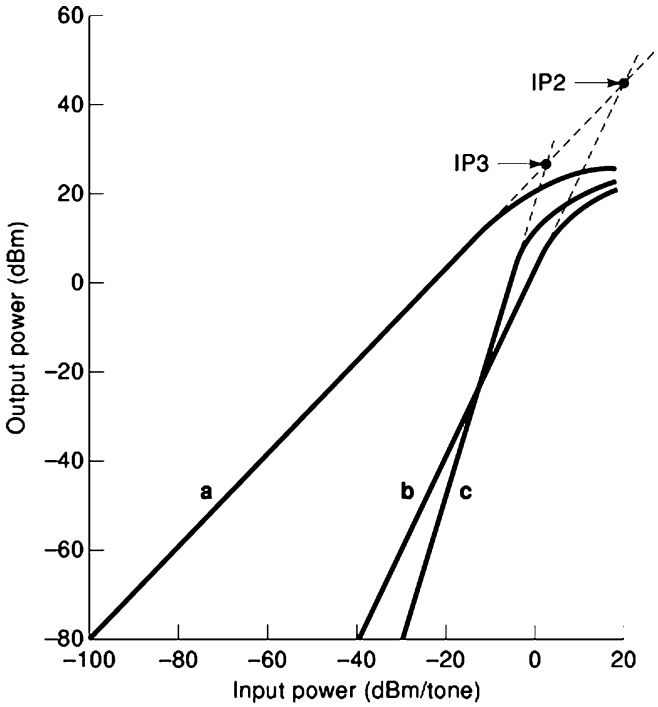

Figure 2. Third-order and second-order intercept points are determined by extending the fundamental, the second-, and the third-order intermodulation transfer function lines. (a) Fundamental transfer function, slope = 1; (b) Second-order intermodulation, slope = 2; (c) Third-order intermodulation, slope = 3. IP3, third-order intercept point; IP2, second-order intercept point.

[49], microwave and millimeter-wave diode detectors [50], analog-to-digital convetrers [51] and [52], gamma correctors [53] and electrical components such as resistors, capacitors, inductors, as well as contacts of switches, connectors, and relays [54] are few examples. The two-tone test is also used to characterize the performance of the basilar membrane of the cochlea [55].

The two-tone test can also be used to determine the transfer characteristic of a nonlinear device modelled by the polynomial approximation of Eq. (2). With the input formed of two properly selected frequencies ω_1 and ω_2, if the second-order and third-order intermodulation products are measured separately, then it is possible to find, from the measured data, the coefficients of the quadratic and cubic terms,k_2 and k_3 respectively, in the polynomial approximation of Eq. (2). If in addition, the IMPs are measured at two sets of values of ω_1 and ω_2, then it is possible to identify the dominant physical nonlinear process from the variation of IMPs with test frequencies (13).

The two-tone test can be used to determine the complex transfer characteristic of a nonlinear device exhibiting AM/AM nonlinearity only with fixed phase shift between the output and the input. In this case a complete set of measurement, for all the two-tone intermodulation products produced by the nonlinearity, at two different power levels is necessary [56]. If the device under consideration exhibits both AM/AM and AM/PM nonlinearities, then determination of a unique set of polynomial coefficients requires a complete set of intermodulation measurements at three different power levels [56]. The set obtained at the highest power level will decide the amplitude range within which the characterization will be valid.□□

Due to the basic assumption that the nonlinearities are represented by polynomials, high accuracy representation of the device characteristics will require the difficult accurate measurements of higher order intermodulation products, in addition to the increased complications and considerable efforts involved in the analysis [56]. Another difficulty from which this method suffers, arises from the necessity of measuring complete sets of two tone intermodulation products which are spread over a relatively wide frequency range and consequently may put stringent specifications on the measuring instruments and techniques if accurate measurements are to be achieved.

In the two-tone test use is made of one of the inband IMPs to describe a device, a circuit or a system nonlinearity. Measurements are made in or near the frequency range of interest. In this test, the input signal consists of two frequencies, ω_1 and ω_2 of equal amplitude and a fixed amount of frequency spacing. At the output of the circuit or the system under test the amplitudes of the third-order intermodulation products $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ are measured. The intermodulation distortion is defined as the ratio between the root sum square of the intermodulation products and the root sum square of the twin-tone amplitudes.□□

Unless a wave analyzer or a spectrum analyzer is available, the implementation of the two-tone test invariably require amplification of the whole output spectrum to get components and on a normalized value (100%). Then, ω_1 and ω_2 are suppressed, and the remaining components $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ are measured with an ac voltmeter or oscilloscope. Especially at audio frequencies, this approach requires steep filters, one set of filters for each set of ω_1 and ω_2. For the same reason $\omega_2 - \omega_1$ cannot be too low.

So it will never be a real narrow-band system. This narrow-band aspect is particularly important for higher frequencies, where equalizers, in the reproduction audio channel, may give unequal amplification of the components in the spectrum [57].

In the audio-frequency range, a new version of the two-tone test overcomes the above-mentioned disadvantages [57]. This is based upon the multiplication of the spectrum by itself. Thus, if a two-tone input signal is given by Eq. (3), with $V_1 = V_2 = V$, then, multiplying the input spectrum by itself-that is, squaring- yields

$$\begin{aligned} V_i^2 &= V^2(\cos\omega_1 t + \cos\omega_2 t)^2 \\ &= V^2(1 + \frac{1}{2}\cos 2\omega_1 t + \frac{1}{2}\cos 2\omega_2 t + \cos(\omega_1 + \omega_2)t \\ &+ \cos(\omega_1 - \omega_2)t) \end{aligned} \quad (6)$$

Assuming that the system under test is narrowband, Eq. (4) reduces to

$$\begin{aligned} V_{out} = b_1\cos\omega_1 t + c_1\cos\omega_2 t + b_5\cos(2\omega_1 - \omega_2)t \\ + c_5\cos(2\omega_2 - \omega_1)t \end{aligned} \quad (7)$$

Multiplying the output spectrum by itself-that is, squaring-yields

$$\begin{aligned} V_{out}^2 &= (b_1\cos\omega_1 t + c_1\cos\omega_2 t + b_5\cos(2\omega_1 - \omega_2)t + c_5\cos(2\omega_2 - \omega_1)t)^2 \quad (8) \\ &= \frac{1}{2}(b_1^2 + c_1^2 + b_5^2 + c_5^2) + (\frac{1}{2}b_1^2 + c_1 b_5)\cos 2\omega_1 t + (\frac{1}{2}c_1^2 + b_1 c_5)\cos 2\omega_2 t + \frac{1}{2}b_5^2\cos(4\omega_1 - 2\omega_2)t \\ &+ \frac{1}{2}c_5^2\cos(4\omega_2 - 2\omega_1)t + (b_1 c_1 + b_5 c_5)\cos(\omega_1 + \omega_2)t + (b_1 c_1 + b_1 b_5 + c_1 c_5)\cos(\omega_1 - \omega_2)t \\ &+ b_1 b_5\cos(3\omega_1 - \omega_2)t + c_1 c_5\cos(3\omega_2 - \omega_1)t + (b_1 c_5 - b_5 c_1)\cos(2\omega_1 - 2\omega_2)t + b_5 c_5\cos(3\omega_1 - 3\omega_2)t \end{aligned} \quad (9)$$

where

$$b_1 = c_1 = k_1 V + \frac{9}{4}k_3 V^3$$

and

$$b_5 = c_5 = \frac{3}{4}k_3 V^3$$

Inspection of the spectra of Eqs. (6) and (9) shows that:

1. Both spectra are split into two parts, a lower region and a higher region. All components in the lower region are not affected by the choice of ω_1 and ω_2. They are only affected by the difference in frequencies $\omega_2 - \omega_1$. This means that as far as $\omega_2 - \omega_1$ is kept constant, the lower region of the spectrum will not be affected if measurement is swept through the whole band.
2. If the distortion is not too high − that is, for $k_3 << k_1$ - then the dc component is essentially constant. Also, Eqs. (6) and (9) show that the amplitudes of the components at frequencies ω_1 and ω_2 are equal to the dc component. In fact this dc component represents the 100% level.
3. The amplitude of the components at frequencies $3\omega_2 - \omega_1$ and $3\omega_1 - \omega_2$ is given by $b_1 b_5 = c_1 c_5$. Thus if $b_1 = c_1$ are normalized to unity- that is, the dc component is unity- then the amplitude of these frequency components is directly proportional to the amplitude of the third-order IMPs at frequencies $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$.

Thus the IMP measurement reduces to Ref. [57]:

1. Squaring.
2. Keeping the dc level to a normalized value.
3. Filtering out the components at frequencies $2\omega_2 - \omega_1$ and $2\omega_1 - \omega_2$.
4. Measuring with an ac voltmeter.

While this approach offers simpler and more flexible instrumentation, its application is limited to low-distortion circuits and systems.

Three standard two-tone test methods are in common use when testing audio-frequency circuits and systems. These are the SMPTE, the CCIF, and the ICE intermodulation tests.

SMPTE Intermodulation Test. In the SMPTE (Society of Motion Picture and Television Engineers) test of intermodulation distortion, the system input is a combination of a large-amplitude low-frequency sine wave with a small-amplitude high-frequency sine wave [58]. Often the large-amplitude component is 80% of the rated maximum input amplitude, and its frequency ω_1 is either 50 Hz or 60 Hz. The small-amplitude component is often 20% of the rated maximum input amplitude and therefore falls within the small-signal regime of the system operation; its frequency ω_2 is often 7 kHz. The large sine wave excites nonlinearities in the system under test, and hence it modulates the small-signal transfer function. Because the two input components are independent, the response of the system under test (in the presence of the large component) is effectively the response of a linear time-varying network. The SMPTE intermodulation test quantifies nonlinearity by reference to the amplitude modulation of the small-amplitude high-frequency component as it appears at the output of the system under test.

CCIF Intermodulation Test. In the CCIF (International Telephonic Consultative Committee) test of intermodulation distortion, the input to the system under consideration is a mixture of two sinusoids, each 50% of full rated amplitude, one at $\omega_1 = 14$ kHz and the other at $\omega_2 = 15$ kHz [59]. Intermodulation distortion is quantified by reference to the amplitude of the resulting $\omega_2 - \omega_1 = 1$ kHz difference-frequency tone at the output.

The CCIF intermodulation test has been successfully used for measurement of high-frequency distortion in audio systems. But it is sensitive only to asymmetrical distortion mechanisms which produce even-order products. If a spectrum analyzer, or sharp cutoff filtering techniques, is used to look at the odd-order products, $2\omega_2 - \omega_1$ and

$2\omega_1 - \omega_2$, as well, which in this case lie at 16kHz and 13 kHz, then the test is also sensitive to symmetrical distortion mechanisms.

IEC Total Difference-Frequency Test. In the IEC (International Electrotechnical Commission) total difference-frequency distortion test, the input is a mixture of two sinusoids, each 50% of full rated amplitude, and angular frequencies ω_1 and ω_2 chosen such that $\omega_1 = 2\omega_o$ and $\omega_2 = 3\omega_o$. Intermodulation distortion is quantified by reference to the amplitude of the difference-frequency tone at ω_o. In the original proposal of the total difference-frequency test, ω_o was chosen to correspond to 5 kHz; this has been changed to 4 kHz [60].

The IEC intermodulation test is fully in-band and detects both even-order and odd-order nonlinearities. However, it does not distinguish between them. A modified version of this test using $\omega_1 = 7\omega_o$ and $\omega_2 = 5\omega_o$ results in second-order IMPs at $2\omega_o$ and $12\omega_o$, and it results in third-order IMPs at $3\omega_o$, $9\omega_o$, $17\omega_o$ and $19\omega_o$. Thus, all the IMPs are well-separated in frequency from each other and from harmonics of the test signals. Proper selection of ω_o results in ω_1 and ω_2, and at least one second-order and one third-order product fall within the bandwidth of the system under consideration [13].

Three-Tone Test. In this test, again, specific inband IMPs are selected to characterize the overall system nonlinearities [61]. The more even spectral distribution and flexibility, while still allowing discrete frequency evaluation, make this an attractive test for multifrequency systems such as communication and cable television systems. □□

In this test three equal-amplitude tones are applied to the input of the nonlinear system under consideration. Thus

$$V_i = V(\cos\omega_1 t + \cos\omega_2 t + \cos\omega_3 t) \tag{10}$$

Combining Eqs. (2) and (10), and using simple trigonometric identities, it is easy to show that the third-order term, $k_3 V_i^3$ will contribute, to the output spectrum, the following:

1. Three components at frequencies ω_1, ω_2 and ω_3 each with amplitude given by□

$$A_1 = \frac{15}{4} k_3 V^3 \tag{11}$$

2. Three components at frequencies $3\omega_1, 3\omega_2$ and $3\omega_3$ each with amplitude given by

$$A_3 = \frac{1}{4} k_3 V^3 \tag{12}$$

3. Twelve components at frequencies $2\omega_m \pm \omega_n, m, n = 1 - 3$ each with amplitude given by

$$A_{21} = \frac{3}{4} k_3 V^3 \tag{13}$$

4. Four components at frequencies $\omega_m \pm \omega_n \pm \omega_p, m, n, p = 1 - 3$ each with amplitude given by

$$A_{111} = \frac{3}{2} k_3 V^3 \tag{14}$$

Equations (13) and (14) show that an intermodulation product of frequency $\omega_m \pm \omega_n \pm \omega_p$ is 6 dB higher in level than an intermodulation product of frequency $2\omega_m \pm \omega_n$.

The intermodulation distortion is defined as the ratio between the amplitude of one of the intermodulation products of frequency $\omega_m \pm \omega_n \pm \omega_p$ to the amplitude of one of the three output tones.□□

In this test the choice of frequencies ω_1, ω_2 and ω_3 used to make the measurement is important. This is because a system's intermodulation performance may not be constant over its operating frequency range. □□

The three-tone test is widely used in characterizing the performance of RF amplifiers used in television boadcast transposers, where the vision carrier, color subcarrier and sound carrier frequency components interact in the presence of amplifier nonlinearities. If the three frequency components are represented as single frequencies - ω_v the vision carrier, ω_{sc} the color subcarrier and ω_s the sound carrier with amplitudes V_v, V_{sc} and V_s respectively, then the input signal can be expressed as

$$V_i = V_v \cos\omega_v t + V_{sc} \cos\omega_{sc} t + V_s \cos\omega_s t \tag{15}$$

Combining Eqs. (2) and (15), and using simple trigonometric identities, it is easy to show that the third-order term of Eq. (2) produces, among others, two in-band intermodulation components given by

$$V_{ip} = \frac{3}{2} k_3 V_v V_{sc} V_s \cos(\omega_v + \omega_s - \omega_{sc})t + \frac{3}{4} k_3 V_s V_{sc}^2 \cos(2\omega_{sc} - \omega_s)t \tag{16}$$

Intermodulation performance, of the transposer, is measured by taking it out of service and using the three-tone simulation of a composite video and sound signal, given by Eq. (15), as its input. The three levels and frequencies vary from system to system. Typical levels, below the peak synchronous pulse level, are $V_v = -6dB$, $V_{sc} = 17dB$ and $V_s = -10dB$. Under these conditions, the first term of Eq. (16) is the most visible, and the second term will be much lower in amplitude, typically 17 dB less. Using a spectrum analyzer, the relative amplitude of the major in-band intermodulation is measured and referenced to the level of peak synchronous pulse. Usually, the permissible level of the major in-band intermodulation component is −53 dB below the reference level. This three-tone test method is slow and requires spectrum analyzers with relatively large dynamic ranges. Moreover, it measures the system performance at one luminance level and one chrominance level. Thus, it does not test the system over its full operating range [62].□□

The inadequacy of the internationally accepted three-tone test method can be overcome, by using a modified color bar test signal [62]. The color bars are applied to the transposer via a test transmitter. The color bars and sound carrier therefore apply the three tones to the transposer, changing levels in rapid succession. With suitable processing, based on sampling the demodulated color bar signal for short intervals corresponding to a selected color, intermodulation levels can be measured simultaneously at seven different luminance levels and can be shown in histogram form [62].

Four-Tone Test. This test is commonly used in voice-frequency circuits. In it, the input consists of two pairs of

tones, thus approximating an input with Gaussian amplitude distribution, which is a more realistic representation of real-life signals. The frequencies of the input tones are selected to generate second- and third-order intermodulation products within the bandwidth of the device under test. These intermodulation products must be easily separated from input tones. A variety of frequencies can be selected for the input tones; a typical widely used set is described here. The first pair of tones is separated by 6 Hz ±1 Hz and centered at 860 Hz ±1 Hz, and the second pair of tones is separated by 16 Hz ±1 Hz and centered at 1380 Hz ±1 Hz. The four tones are of equal level within ±0.25 dB. Using Eq. 2, it is easy to show that the output of the device under test will contain six third-order intermodulation products in the range 1877 Hz to 1923 Hz, four second-order intermodulation products in the range 503 Hz to 537 Hz, and four second-order intermodulation products in the range 2223 Hz to 2257 Hz. Thus, the second-order products measured are the combination of two band-pass filters, one having a passband from 503 Hz to 537 Hz, and the other having a passband from 2223 Hz to 2257 Hz. The third-order products measured are the output of a bandpass filter having a passband from 1877 Hz to 1923 Hz.

Multitone Test. Although two-, three- and four-tone intermodulation distortion measurements present very useful information about the performance of a nonlinear system, these signals cannot simulate the final operation regime of the system under consideration. For example, RF power amplifiers used in modern telecommunications systems are expected to handle signals that are generally modeled as multitone spectra [63], [64]. On the other hand, the ideal testing signal in an audio system should be able to reveal the maximum amount of pertinent information about the nonlinear system under test. Such a signal should provide credible clues as to how the measured data can be linked to the perceived sound quality [65], [66]. Designers of audio as well as microwave systems are, therefore, seeking alternative techniques that can emulate closely the final operation regime of their systems. Although not a fully legitimate representative of real audio or microwave signals, the multitone signal nevertheless possesses some of their qualities. In fact the multitone test is generally able to provide realistic test conditions by approximating the frequency domain characteristic of a typical signal. Thus, the multitone test can provide much more meaningful information about the behavior of a nonlinear system than can be obtained by the standard two-, three-, and four-tone tests. However, generating a multitone test signal is nontrivial. The multitone test requires the design of low-crest-factor test signals [67].

Noise-Power Ratio (NPR) Test. In the NPR test, the input to the device under test is obtained from a white noise source which is band limited to the instantaneous frequency range of interest. This emulates a situation with many simultaneous input signals. Provided that none of the signals dominate, according to the central-limit theorem, the resulting voltage obtained when many uncorrelated signals are added will approach a Gaussian distribution. True white noise covers a frequency range of interest continuously, unlike discrete signals.□□

The NPR test measures the amount of intermodulation products power between two frequency ranges of white Gaussian noise. A white noise generator is used with its output frequency range limited by a bandpass filter according to the bandwidth of the device under test. A quiet channel is formed by a switchable bandreject filter, as shown in Fig. 3. Then, the resulting white noise signal is applied to the input of the device under test. At the output of the device under test is a receiver which is switch-tuned to the frequency of the bandreject filter used to produce the quiet channel.

The NPR test is widely used for evaluating the intermodulation poerformance of systems whose input signals spectrum distribution can be approximated by that of white noise. However, the NPR may be degraded by the noise floor of the system under test especially under very low loading conditions. It may also be degraded by the distortion products which are produced under high loading conditions [68].

Cross-Modulation. □□Cross-modulation occurs when modulation from a single unwanted modulated signal transfers itself across, and modulates the wanted signal. Cross-modulation is troublesome primarily if the desired signal is weak and is in adjacent to a strong unwanted signal. Even when the carrier of the strong unwanted signal is not passed through the system, the modulation on the undesired carrier will be transferred to the desired carrier. Cross modulation is, therefore, a special case of intermodulation. Recall that when the input to a nonlinear system is formed of a two-tone signal of the form of Eq. (3), then the amplitudes of the output components at frequencies ω_1 and ω_2 will be given by

$$b_1 = k_1 V_1 + \frac{3}{4} k_3 V_1^3 + \frac{3}{2} k_3 V_1 V_2^2 \tag{17}$$

and

$$c_1 = k_1 V_2 + \frac{3}{4} k_3 V_2^3 + \frac{3}{2} k_3 V_1^2 V_2 \tag{18}$$

respectively. Thus, the output obtained at each frequency ω_1 and ω_2, is dependent upon the amplitude of the signal component of the other frequency. If the amplitude of the wanted unmodulated carrier is V_1 and the instantaneous amplitude of the unwanted amplitude-modulated carrier is

$$V_2(t) = V_2(1 + m\cos\omega_m t) \tag{19}$$

then, using Eq. (17), the amplitude of the wanted carrier will be

$$b_1 = k_1 V_1 + \frac{3}{4} k_3 V_1^3 + \frac{3}{2} k_3 V_1 V_2^2 (1 + m\cos\omega_m t)^2 \tag{20}$$

For small values of m and with $k_3 << k_1$, Eq. (20) can be approximated by□

$$b_1 \cong k_1 V_1 + 3 k_3 V_1 V_2^2 m\cos\omega_m t \tag{21}$$

Thus the wanted carrier will be modulated by a modulation index

$$p = 3\frac{k_3}{k_1} V_2^2 m \tag{22}$$

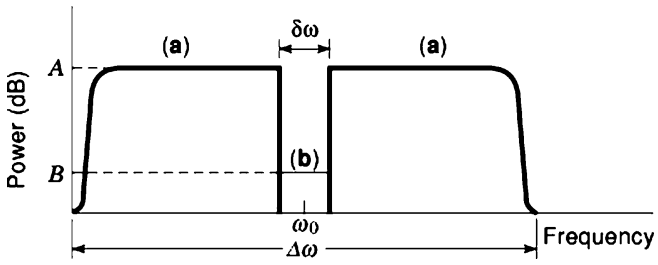

Figure 3. The output spectrum of a noise-power ratio measurement. (a) Injected noise. (b) Noise and intermodulation generated in the measurement bandwidth, $\delta\omega$, by the DUT. NPR = $A - B$.

The cross-modulation factor is then defined as

$$K = \frac{p}{m} \tag{23}$$

Thus, one frequency will be modulated by the modulation of the other frequency. Similar results can be obtained if the unwanted carrier is FM modulated.□□

Cross-modulation can be measured as the change in the amplitude of the wanted unmodulated carrier as a function of the amplitude of the unwanted unmodulated carrier. This is the procedure recommended by the NCTA (National Cable Television Association) standard cross-modulation measurement [69].

Alternatively, the cross-modulation can be measured using the definition of Eq. (23), that is measuring percentage modulation that appears on an unmodulated desired carrier due to the presence of an undesired modulated carrier, divided by the percentage modulation on the undesired carrier [70].

Cross-modulation can also be measured using two equal-amplitude carriers. The wanted carrier, ω_2 is unmodulated while the unwanted carrier, ω_1 is FM modulated. The output spectrum clearly shows the frequency deviation of the wanted carrier. Moreover, it can be shown that the frequency deviation of the intermodulation components, of the output spectrum, is larger than that of the original FM modulated unwanted carrier. For the intermodulation product of frequency $\alpha\omega_1 \pm \beta\omega_2$, the deviation will be multiplied by α. Thus, it may be easier to measure the cross-modulation by measuring the deviation of an intermodulation product rather than the deviation of the wanted unmodulated carrier [71].

Differential Gain. Differential gain (DG), a parameter of special interest in color-TV engineering, is conventionally defined as the difference in gain encountered by a low-level high-frequency sinusoid at two stated instantaneous amplitudes of a superimposed slowly varying sweep signal. In video signal transmission, the high frequency sinusoid represents the chromatic signal and the low frequency sinusoid represents the luminance signal. Corresponding to the theoretical conditions of the differential measurement, DG measurement is performed by a signal of the form of Eq. (3) with $\omega_2 >> \omega_1$ and $V_2 \to 0.0$ at $V_1 = 0.0$ and X [72]. Therefore, recalling that when the input to a nonlinear system is formed of a two-tone signal of the form of Eq. (3), then the amplitude of the output component at frequency ω_2 will be given by

$$c_1 = k_1 V_2 + \frac{3}{4} k_3 V_2^3 + \frac{3}{2} k_3 V_1^2 V_2 \tag{24}$$

Thus, the DG can be expressed as□

$$DG = 1 - \frac{k_1 + \frac{3}{4} k_3 V_2^2}{k_1 + \frac{3}{4} k_3 V_2^2 + \frac{3}{2} k_3 X^2} \tag{25}$$

DG can, therefore, be considered, to some extent, as a measure of the intermodulation performance of a system under test. □

Dynamic Range. Dynamic range can be defined as the amplitude range over which a circuit or a system can operate without performance degradation. The minimum amplitude is dictated by the input thermal noise and the noise contributed by the system. The maximum amplitude is dictated by the distortion mechanisms of the system under consideration. In general, the amount of tolerable distortion will depend on the type of the signals and the system under test. However, for the purpose of an objective definition the maximum amplitude will be considered the input signal level at which the intermodulation distortion is equal to the minimum amplitude [73]. The dynamic range can, therefore, be considered, to some extent, as a measure of the intermodulation performance of a system under test. □

A useful working definition of the dynamic range is that, it is (1) two-third of the difference in level between the noise floor and the intercept point in a 3 kHz bandwidth [74], or (2) the difference between the fundamental response input level and the third-order response input as measured along the noise floor (sometimes defined as 3 dB badwidth above the noise floor) in a 3 kHz bandwidth, as shown in Fig. 4. Reducing the bandwidth improves dynamic range because of the effect on noise.

Because the power level at which distortion becomes intolerable varies with signal type and application, a generic definition has evolved. The upper limit of a network's power span is the level at which the power of one IM product of a specified order is equal to the network's noise floor. The ratio of the noise-floor power to the upper-limit signal power is referred to as the network's dynamic range (DR). Thus the DR can be determined from [75]

$$DR_n = \frac{n-1}{n}[IP_{n,in} - MDS] \tag{26}$$

where DR_n is the dynamic range in decibles, n is the order, IP_{in} is the input intercept power in dBm, and MDS is the

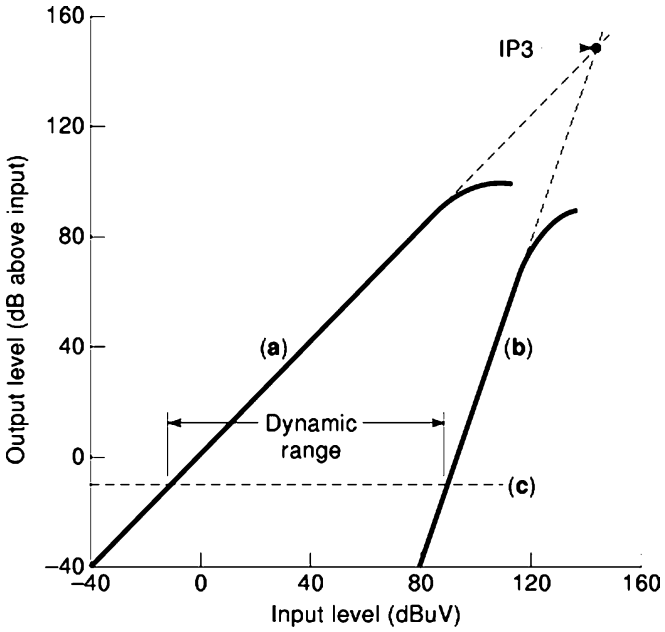

Figure 4. The dynamic range is the difference between the fundamental response input level and the third-order response input as measured along the noise floor. (a) Fundamental response. (b) Third-order intermodulation response. (c) Noise floor.

minimum detectable signal power in dBm.

Alternatively, in receiver circuits the spurious free dynamic range (SFDR) and the intermodulation free dynamic range (IFDR) are widely used to quantify the capability of the receiver to listen to a weak station, without disturbance from an intermodulation product generated by strong stations on other frequencies. The SFDR and the IFDR are in fact measures of how strong two signals can be before the level of their intermodulation products can reach the noise-floor of the receiver. The SFDR, or the IFDR, is defined as the difference in dB between the power levels of the third order intermodulation IM_3 (assuming that there is only a third-order nonlinearity) and the carrier when the IM_3 power level equals the noise floor at a given noise bandwidth. It can be expressed as [76]

$$SFDR = \frac{2}{3}[IIP_3 - EIN - 10\log_{10}(NBW)] \qquad (27)$$

where IIP_3 is the third-order input intercept point, EIN in (dB/Hz) is the equivalent input noise and NBW (in Hz) is the noise bandwidth.

Adjacent- and Co-Channel Power Ratio Tests. In modern telecommunication circuits signals comprising one or more modulated carriers are handled. The characterization of the intermodulation performance of such circuits can not, therefore, be performed using two-tone and three-tone input signals. A combination of equally spaced tones; in practice more than about ten sinusoids [77], with constant power and correlated or uncorrelated phases would be more appropriate [78].

Because of the nonlinearity of the device-under-test, intermodulation products will be generated. These intermodulation products can be classified as adjacent channel distortion; when their frequencies are located to the right or to the left of the original spectrum, or co-channel distortion; when their frequencies are located exactly over the original spectrum. The adjacent channel power ratio (ACPR) is defined as the ratio between the total linear output power and the total output power collected in the upper and lower adjacent channels [79]. The co-channel power ratio (CCPR) is defined as the ratio between total linear output power and total distortion power collected in the input bandwidth [79]. The intermodulation distortion ratio (IMR) is the ratio between the linear output power per tone and the output power of adjacent channel tones [79].

In fact the ACPR, CCPR and IMR distortion measurements are simple extensions to the two-tone intermodulation measurement [80]. However, it is important to firstly generate a very clean multi-tone signal. This can be easily achieved using the technique described in reference [81].

Dynamic Techniques

Sine-Square Test. This test employs a square wave as a type of signal characterized by rapid change of state [82]. In fact, this is a two-tone test where the signal consists of a 3.18 kHz square wave, which has been filtered with a simple one-pole, low-pass RC filter, at either 30 kHz or 100 kHz, and combined with a 15 kHz sine wave. The peak-to-peak amplitude ratio of the sine wave to the square wave is 1:4. The resulting square-wave signal component has a very high slope, which is in theory actually limited only by the low-pass filter. This test has the capability of stressing the amplifier to a high degree of nonlinearities related to signal slope and/or slew rate.

The output spectrum of the system under test is analyzed for the intermodulation products generated by nonlinear mixing of the sine and square waves. The rms sum of the intermodulation products relative to the amplitude of the 15 kHz sine wave is defined as the percentage distortion.

Because the test signal rate of change depends heavily on out-of-band (>20 kHz) harmonics of the square wave, this test can be led to give somewhat optimistic results for audio systems incorporating front-end low-pass filters [83]. Moreover, each component in the output spectrum has two contributory parts: (1) the dynamic intermodulation component caused by the rise-time portion of the square wave driving the amplifier to frequency dependent nonlinearity—that is, TIM—and (2) the static intermodulation component caused by the amplitude-dependent nonlinearity of the amplifier.

In order to separate the static intermodulation component, the sine-triangle test was proposed (82). The sine-triangle test is similar to the sine-square test with the square wave replaced by a triangular wave of equal peak-to-peak amplitude. This reduces drastically the rise-time, leaving only the intermodulation components caused by the static nonlinearities. However, both the sine-square and the sine-triangle tests do not uniquely separate the static and dynamic nonlinearities.

The high-pass-square-wave/sine (HP-SQ/S) test is a distinct modification of the sine-square and the sine-triangle tests [84]. The HP-SQ/S test is based on the sine-square test with the test signal further shaped by a single-pole RC highpass filter with a 3 dB rolloff frequency = 5 kHz. Thus sine-square test signal generators can be applied, requiring only an additional RC high-pass filter. The HP-SQ/S test signal better resembles real-life signals and acquires both static and dynamic nonlinear distortions simultaneously. Similar to the sine-square test, the distortion factor is defined as the amplitude ratio of the rms sum of the intermodulation products and referred to the amplitude of the sinusoid.

Sawtooth Wave Test. In this method a sawtooth wave is used as the signal that changes its state rapidly [85]. The signal is derived from inverting the phase of a 30 kHz sawtooth waveform with a 30 kHz/256 period as shown in Fig. 5. The signal therefore consists of two alternating series of signals, one series of instantaneously rising waveforms, the other of instantaneously falling signals.

Without reversal, application of the high-frequency asymmetrical sawtooth signal to a system under test causes the system's symmetrical and asymmetrical nonlinearities to generate a dc offset whose magnitude depends on the severity of the nonlinearity. The periodic polarity reversal merely "chops" this dc offset into an easily measured low-frequency ac signal. Thus, if the system under test is prone to TIM, then at the output of the low-pass filter a signal appears whose shape is rectangular. Each time the input signal reverses polarity, a rectangular output waveform appears that is due to the shift in average voltage or dc level. This output signal represents the amount of TIM in the system under test. TIM is calculated as

$$\text{TIM} = ((\text{amplitude(peak} - \text{to} - \text{peak)of sawtooth})/ (\text{amplitude(peak} - \text{to} - \text{peak)of square wave}))100\% \quad (28)$$

Because this test depends so heavily on extreme signal slope (and hence on out-of-band sawtooth harmonics), it may easily be misled in its assessment of TIM by systems incorporating low-pass filters [83].

Multitone Intermodulation Test (MIM). The MIM test is a variation of the CCIF intermodulation test in which two high-frequency tones spaced apart by a small frequency difference are applied to the system under test [83]. In order to retain the advantages of the CCIF test while incorporating sensitivity to symmetrical distortion mechanisms, a third tone, at frequency ω_3, is added. The three frequencies are chosen so that the resulting triple-beat product at $\omega_3 - \omega_2 - \omega_1$ is slightly below 1 kHz, while two of the tones produce a CCIF-like difference frequency product $\omega_2 - \omega_1$ at slightly above 1 kHz. Specifically, the three equal-amplitude tones are at frequencies 20, 10.05, and 9 kHz, resulting in a triple-beat product at 950 Hz and a difference-frequency product at 1050 Hz.

The difference-frequency and the triple-beat products are selected to lie so close to each other that both products can be passed through a relatively narrow band-pass filter centered about 1 kHz. The distortion percentage is defined as the value of the 950 Hz and 1050 Hz distortion products, measured together on an rms calibrated average responding ac voltmeter, referred to the rms value of the sine wave of the same peak-to-peak amplitude as the three-tone MIM test signal.

The MIM test enjoys the following attractive features:

1. Inexpensive instrumentation; no spectrum analyzers are required.
2. Simple measurement procedure.
3. Fully in-band stimulus and response.

The MIM test is not as stringent as the sine-square and sawtooth wave tests in terms of peak rate of change, and as a result it yields smaller TIM distortion percentages. However, because it does not resort to unrealistically high rates of change to stress the audio system under test, good subjective correlation can be expected.

INTERMODULATION MEASUREMENT

Measurement Equipment

Multi-Tone Tests. A block diagram of the system used for multi-tone intermodulation measurement is shown in Fig. 6. The multiple frequency source can be implemented from two or three synthesized sine/square/triangular wave generators. Amplifier/attenuator pairs can be added at the output of each generator. Bandpass filters can also be added to suppress the harmonic contents at the output of each generator. For RF measurements, harmonic suppression and isolation between different generators is achieved by using amplifier/circulator combinations and cavity resonators [86]. The synthesized sources are combined using hybrids or combiners of adequate isolation. Spectral purity at this point is crucial to the accuracy of the measurement. The multitone output is fed to the device under test (DUT).

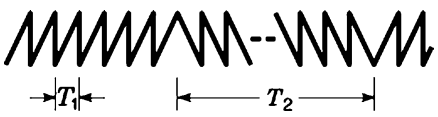

Figure 5. The input signal of a sawtooth wave test is derived from inverting the phase of a 30 kHz sawtooth waveform with a 30 kHz/256 period. T_1 = 1/30 kHz, $T_2 = 256T_1$.

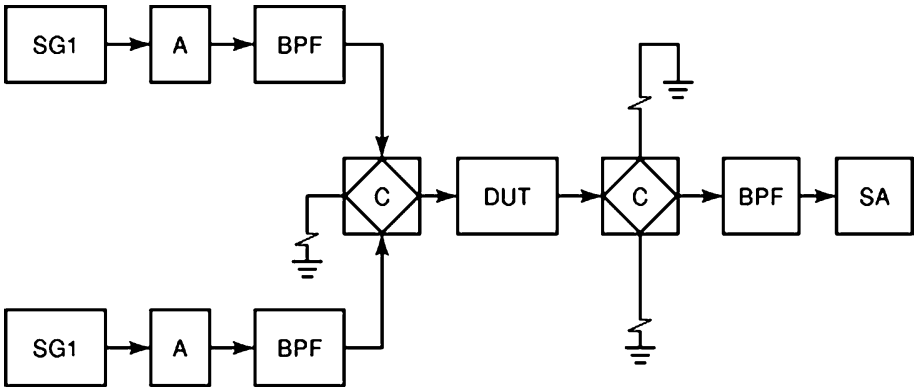

Figure 6. Block diagram of the two-tone test setup. Multitone tests require additional signal generators, combiners, amplifiers and bandpass filters. SG, signal generator; A, amplifier; BPF, bandpass filter; C, combiner; DUT, device under test; SA, spectrum analyzer.

The output of the DUT is fed to the spectrum analyzer. For RF measurements, the output of the DUT can be fed to directional couplers. The outputs of the directional couplers are fed to a television oscilloscope and/or a spectrum analyzer.

For audio-frequency measurements, resistive combiners are widely used for combining the outputs of two, or more, signal generators. Figure 7 shows a number of widely used resistive combining networks.

Measurement Using a Microcomputer. Intermodulation can, also, be measured using a microcomputer [87]. The block diagram of this technique is shown in Fig. 8. This technique is based on measuring the single tone input-output characteristic of the DUT using a vector voltmeter. The output of the vector voltmeter is fed to a microcomputer which converts it into three digital data lines representing the input amplitude, the output amplitude and the phase lag between the input and output signals. After storing the data, the microcomputer increments the amplitude of the input signal. After storing all the necessary data, the microcomputer, using a stochastic method, calculates the amplitudes of the intermodulation components of the DUT. Although the procedure reported in [87] uses a stochastic method for calculating the amplitudes of the intermodulation components resulting from a two tone input signal, the same procedure can be applied to any number of input tones using different analytical techniques for modelling the nonlinear characteristics of the DUT.

Alternatively, microcomputers can be added to the measurement setup of Fig. 6 to:

1. Control the frequencies of the signal sources, especially in the millimeter wave length range where the difference in frequencies between the signal sources may be less than 0.001 of the base signal frequency [88].
2. Scan the base signal frequency over the measurement range of interest in predefined steps [89].
3. Correct the power from each source so that power delivery to the DUT will be the same across the whole frequency range scanned.
4. Read and calculate the parameters of interest during the measurements [90] and [91].

Noise-Power Ratio Test. Figure 9 shows a block diagram of a noise-power ratio test setup [68]. The setup consists of a white noise generator which applies an accurate level of white Gaussian noise power with known bandwidth (equals $\Delta\omega$ and centered around ω_o) to the DUT. The output of the DUT is measured with the bandreject filter out. When the bandreject filter, with bandwidth = $\delta\omega$ and centered around ω_o, is switched in, a narrow band of frequencies is attenuated by about 70 dB, and a quiet channel, of width $\delta\omega$ and centered around ω_o, is formed as shown in Fig. 3. At the output of the DUT, the noise power is measured in the quiet channel, using a bandpass filter with bandwidth $\delta\omega$ and centered around ω_o. This noise power is due to the thermal noise and the intermodulation introduced by the device under test (DUT). The NPR is the ratio between the noise power measured without the bandreject filter inserted before the DUT to that measured with the bandreject filter inserted. The white noise generator corrects the loading power level for the insertion loss of the bandreject filter.

Noise Floor and SFDR Test. Figure 10 shows a test setup for measurement of noise floor and the SFDR of a communication link [76]. To measure the noise floor of the communication link, the transmitter is switched off. Then the noises of the low-noise amplifier and the spectrum analyzer are measured. Switching the transmitter on increases the noise floor by the transmitter noise and therefore the dif-

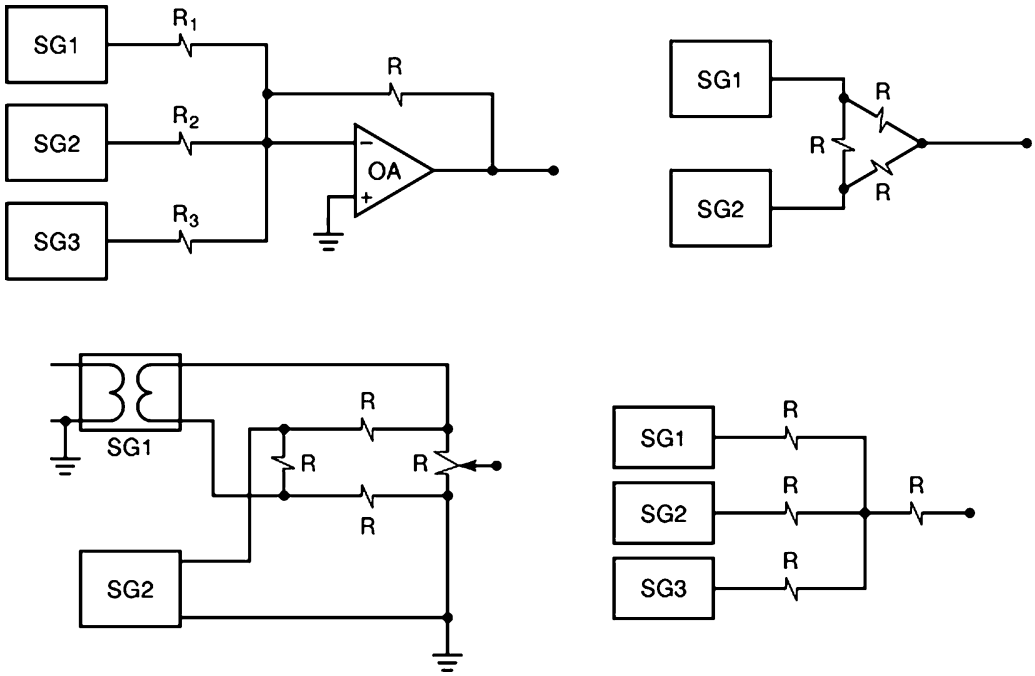

Figure 7. Different types of resistive combiners used in audio-frequency tests. SG, signal generator; OA, operational amplifier.

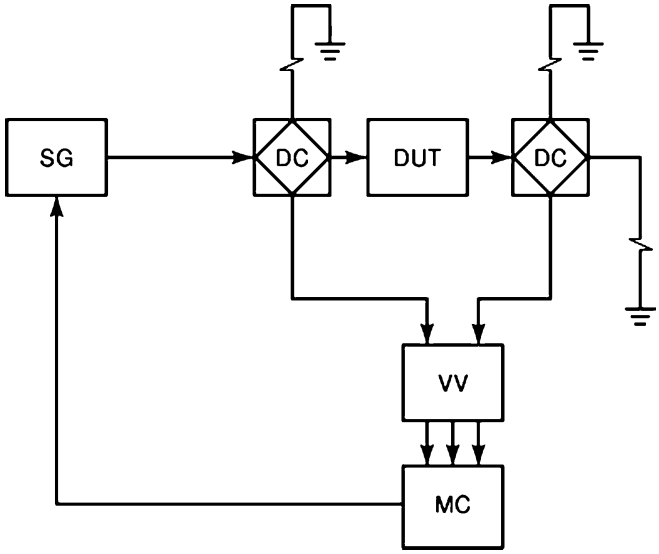

Figure 8. Block diagram of a microcomputer-based intermodulation measurement setup. SG, signal generator; DC, directional coupler; DUT, device under test; VV, vector voltmeter; MC, microcomputer.

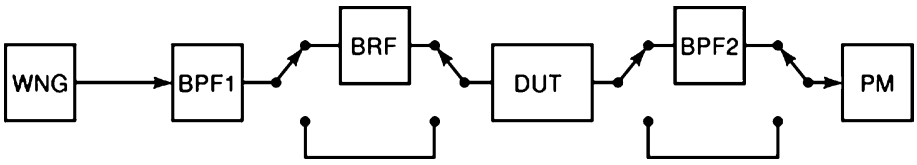

Figure 9. Block diagram of the noise-power ratio test setup. WNG, White noise generator; BPF1, bandpass filter with bandwidth $\Delta\omega$ centered around ω_0; BRF, bandreject filter with bandwidth $\delta\omega$ centered around ω_0; DUT, device under test; BPF2, band-pass filter with bandwidth $\delta\omega$ centered around ω_0; PM, power meter.

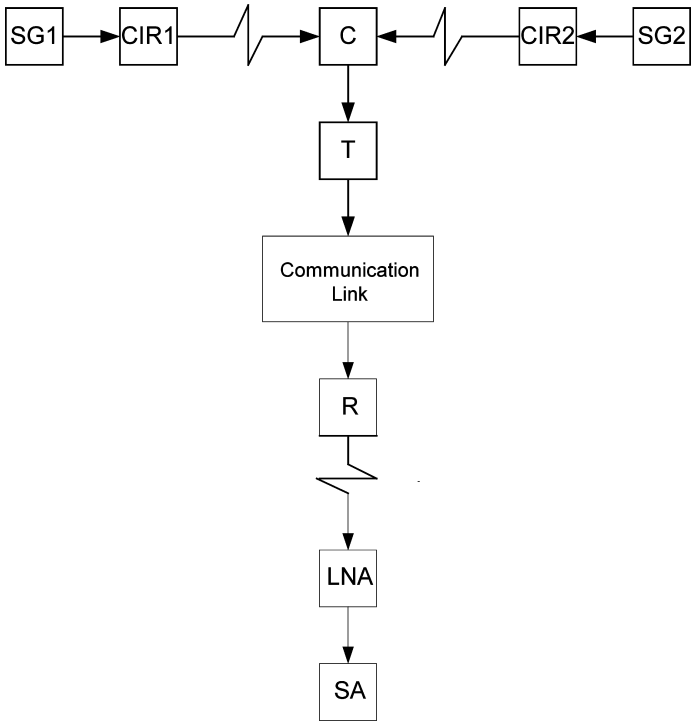

Figure 10. Setup for noise floor and SFDR measurement. SG, signal generator; CIR, circulator; C, combiner; T, transmitter; R, receiver; LNA, low-noise amplifier; SA, spectrum analyser.

ference between the two noise measurements is the noise generated by the transmitter.

To measure the SFDR the input power is decreased until the IM_3 level equals the noise floor. Recall that deceasing the input power by 1 dB decreases the IM_3 level by 3 dB. However, this is true only if the third-order nonlinearity is dominant. Higher order nonlinearities will contribute to the third-order intermodulation (IM_3) and in such cases the measured SFDR will be different from calculations obtained using equation (27).

Externally Induced Intermodulation Test. This is a two-tone test with one signal applied to the input and the other signal applied to the output [9]. A test setup is shown in Fig. 11. Two directional couplers are used to gauge both the forward-carrier power and the intermodulation product levels. Two more directional couplers are added to inject the interfering signal and to measure the actual injected value using the spectrum analyzer.

Measurement Accuracy

Multitone Tests. For accurate measurements of the intermodulation products, using multitone tests, it is essential to reduce, or remove, the nonlinear distortion originating in the signal sources and/or the measurement equipment. Measurement accuracy may, therefore, be affected by the purity of the signal sources, the linearity of the combiners and the performance of the spectrum analyzer.

Signal Sources. Measurement of the amplitudes of the intermodulation components requires the use of two or more signals. The frequencies of these signals must be non-commensurate. Otherwise, harmonics in one source might beat with the fundamental(s) of other signal(s) and interfere with the desired intermodulation components.

Ideally the signal generators would produce perfect sinusoids, but in reality all signals have imperfections. Of particular interest here is the spectral purity which is a measure of the inherent frequency stability of the signal. Perhaps the most common method to quantify the spectral purity of a signal generator is its phase noise [92]. In the time domain, the phase noise manifests itself as a jitter in the zero crossings of a sine wave. In the frequency domain, the phase noise appears as sidebands surrounding the original frequency. Thus, mixing with other frequencies, due to the nonlinearities of the device-under-test, would result in additional intermodulation products. It is, therefore, important to consider the intermodulation due to phase noise when calculating the intermodulation performance of the device- under-test [93].

Signal generators with automatic level control (ALC) may produce signals with unwanted modulation. The ALC is implemented by rectifying the output signal of the generator and feeding back the resulting dc voltage to drive an amplitude modulator. If a second signal is applied to the output of the signal generator, the detector will produce a signal at the difference in frequency between the two frequencies. This signal will modulate the generator's output. The frequency of the modulation sidebands will share the same spectral lines as the intermodulation products of interest. Isolating the signal generators and the combiners can minimize such effect. This can be achieved by ensuring that there is as much attenuation as possible between them.

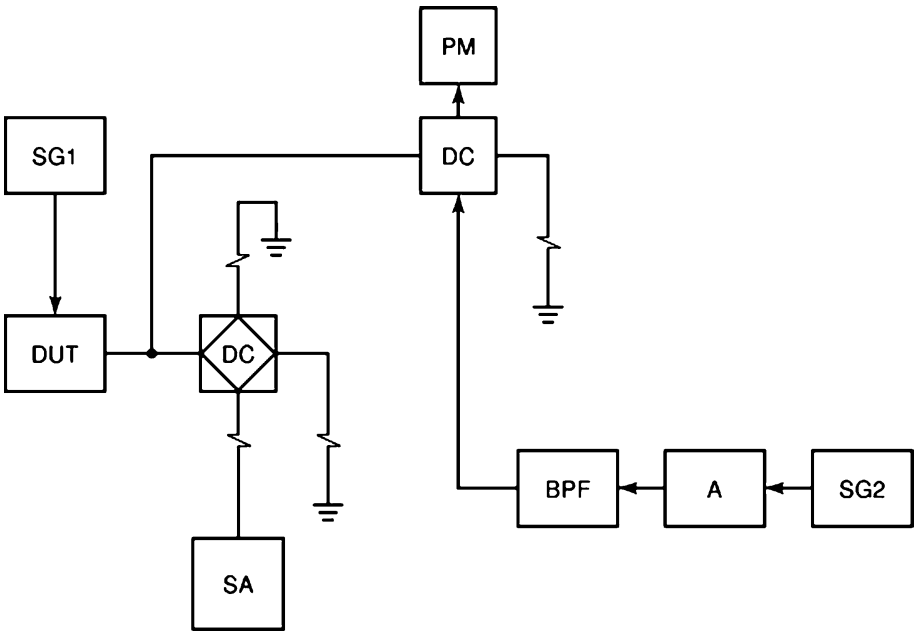

Figure 11. Measurement of externally-induced intermodulation can be performed by using two tones one injected at the input and one injected at the output of the DUT. SG, signal generator; DC, directional coupler; PM, power meter; SA, spectrum analyzer; BPF, bandpass filter; A, amplifier.

Combiners. Measurement of intermodulation products is performed by applying to the input of the circuit, or the system, under test a signal consisting of two, or more, different frequencies obtained from different signal generators. The outputs of the signal generators are, therefore, combined by a combiner. The combiner must provide sufficient isolation between the signal sources to reduce the possibility of producing intermodulation products before the combined input signal is applied to the circuit or the system under test. While resistive combiners are adequate for input signal levels up to few milli-volts, for larger voltage levels the use of power combiners may be inevitable [94]. Insertion of an attenuator in each arm of the combiner helps in minimizing the distortion components resulting from the interaction between the two signal sources. Such components, if generated, should be at least 80 dB below the fundamental components.

A simple test to determine whether adequate isolation has been achieved can be effected by introducing a variable attenuator between the signal source combiner and the DUT in Fig. 8. This is set to a low value during measurements but at set-up, when IMPs have been located on the spectrum analyzer, increasing the attenuation by 3 dB will result in a reduction in the observed IMP level. If this reduction is only 3 dB then it has to be assumed that the IMP observed has originated in the signal sources, not in the DUT. If, however, the reduction is 6 dB for a second-order IMP or 9 dB for a third-order [see Eq. (4)], then it is safe to assume that the IMP has originated in the DUT or the spectrum analyzer.

Alternatively, a technique which attenuates the parasitic intermodulation products which arise due to the interaction between the generators of the fundamental components, before the input of the spectrum analyzer, was described in Ref. [95]. A block diagram of the technique is shown in Fig. 12. The input to the system under test is formed by combining the outputs of two signal generators at frequencies ω_1 and ω_2 in the combiner. The hybrid combiner/splitter (HCS1) splits the combined signal into two branches with voltage-transfer ratio $a = \alpha$ and $b = \sqrt{1-\alpha^2}$ at the first and second outputs. Using Eq. (1), and assuming that the system under test and the compensator have identical nonlinear characteristics, the inputs of the second hybrid combiner/splitter (HCS2) can be expressed as□

$$V_a = n = 0 \sum k_n (\alpha V_i)^n \tag{29}$$

and

$$V_b = n = 0 \sum k_n (\sqrt{1-\alpha^2} V_i^n) \tag{30}$$

Using Eqs. (29) and (30), the output of the second hybrid combiner/splitter (HCS2), with voltage transfer ratio opposite in sign and equal to the reciprocal of that of HCS1, can be expressed as

$$V_{out} = n = 0 \sum - k_n (\sqrt{1-\alpha^2}(\alpha V_i)^n - \alpha(\sqrt{1-\alpha^2} V_i^n)) \tag{31}$$

According to Eq. (31), broadband compensation occurs for the linear components of the combined signal, with $n = 1$. Thus, all the linearly transformed spectral components are eliminated. This is also true for the intermodulation components which may result from the nonlinear interaction between the two signal generators. The output of HCS2 can, therefore, be applied directly to the spectrum analyzer.

This technique does not require complicated high-order selective filters and can attenuate the parasitic intermodulation components and the fundamental frequency components by about 50 dB over a wide range of frequencies

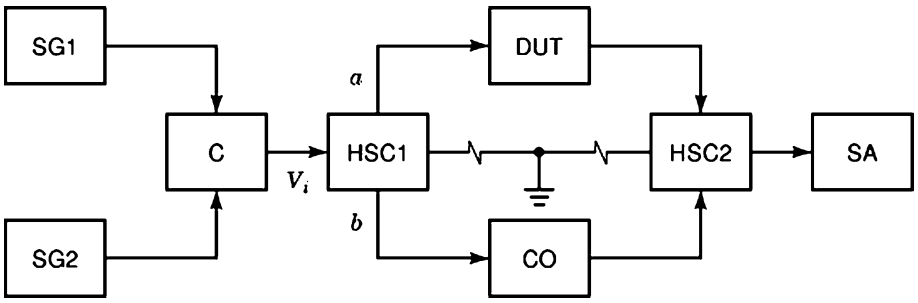

Figure 12. A technique for attenuating the intermodulation products arising due to the interaction between the signal generators of the fundamental components. SG, signal generator; C, combiner; HSC, hybrid splitter combiner; DUT, device under test; CO, compensator; SA, spectrum analyzer.

differing by 7 to 10 octaves. However, it requires a compensator with a nonlinear charactristic similar to that of the system under test.

Spectrum Analyzers. Spectrum analyzers are widely used in measuring the intermodulation performance of electronic circuits and systems. Internal circuits of the spectrum analyzers are, themselves, imperfect and will also produce distortion products [47]. The distortion performance of the analyzers is usually specified by the manufacturers, either directly or lumped into a dynamic range specification. The performance of the analyzer can be stretched, however, if the nature of these distortion products is understood.

Amplitudes of the distortion products, resulting from the internal circuits of the analyzer, can be reduced by reducing the signal levels at the analyzer's input. Thus, using internal and/or external attenuators can reduce the input signal levels to the analyzer and, hence reduce its distortion products and improve the intermodulation measurement range of the spectrum analyzer. However, reduced input levels to the analyzer means reduced signal-to-noise ratio, and the distortion component to be measured may be buried in the noise. While reducing the resolution bandwidth of the analyzer can reduce noise, this may lead to slower sweep rate. Thus, achieving an optimum dynamic range involves trade-offs between input signal levels and analyzer distortion. Usually, data sheets of analyzers will contain information about noise level in each resolution bandwidth and distortion products generated by the analyzer for each input level. Using these information the dynamic range of the analyzer for various input levels can be determined [96].

Whenever good selectivity, as well as sensitivity and dynamic range, are of prime importance, test receivers may be used in preference to spectrum analyzers [6].

Alternatively, if the frequencies of the intermodulation components of interest are sufficiently lower (or higher) than the fundamental frequencies, then lowpass (or highpass) filters can be used to remove the fundamental components which would give rise to other nonlinear distortion components in the spectrum analyzer. Attenuation factors of 80 dB or more, at frequencies outside the band of interest, are recommended. The insertion loss of the lowpass (or the highpass) filter should be as small as possible; 0.4 dB or less is recommended.

If the frequency of the intermodulation component of interest is not sufficiently higher (or lower) than the fundamental frequencies, then it would be necessary to have complicated multiple-section high-order filters with amplitude-frequency characteristics that are nearly rectangular. Such filters will change, to some extent, the amplitude of the intermodulation components and this will complicate the calculation of the intermodulation performance of the system under test. A method for compensating for a large fundamental component, thus allowing the measurement of small intermodulation components in its presence, was described in Ref. [97].

A block diagram of the compensation method is shown in Fig. 13. The input to the system under test is formed of one large amplitude signal at frequency ω_1 and one small amplitude signal at frequency ω_2 with $\omega_1 << \omega_2$. The output of the system under test contains fundamental components at frequencies ω_1 and ω_2, and intermodulation components at frequencies $\omega_2 \pm n\omega_1, n = 1, 2, ..., N$. In order to measure the small amplitude intermodulatiuon components it is necessary to avoid applying to the analyzer the fundamental component at frequency ω_2. This can be achieved as follows. The output of the system under test is fed to the band-reject filter BRF2 to suppress the fundamental component at ω_1. The output of the signal generator of frequency ω_2 is fed to the band-reject filter BRF1 to suppress any component at frequency ω_1 before reaching the phase-shifter through the combiner. The phase shifter compensates, at the frequency ω_2, the phase shift through the system under test

Ideally, the voltages, of frequency ω_2 at the inputs of the differential amplifier are equal. Thus, the output of the differential amplifier at frequency ω_2 is ideally zero. Practically the output voltage at ω_2 will be attenuated by 50–60 dB (6). The output of the differential amplifier, with suppressed fundamental component at frequency ω_2, can be applied to the spectrum analyzer.

This compensation technique, which entails additional filters and matching units, can be used only for broadband measurements with $\omega_1 << \omega_2$.

While spectrum analyzers using digital IF sections, may not suffer from the internally generated distortion, discussed in the preceding, they may suffer from the relatively low-level distortion products resulting from the analog-to-digital conversion. The amplitudes of these products is usually less sensitive to the amplitude of the signal compo-

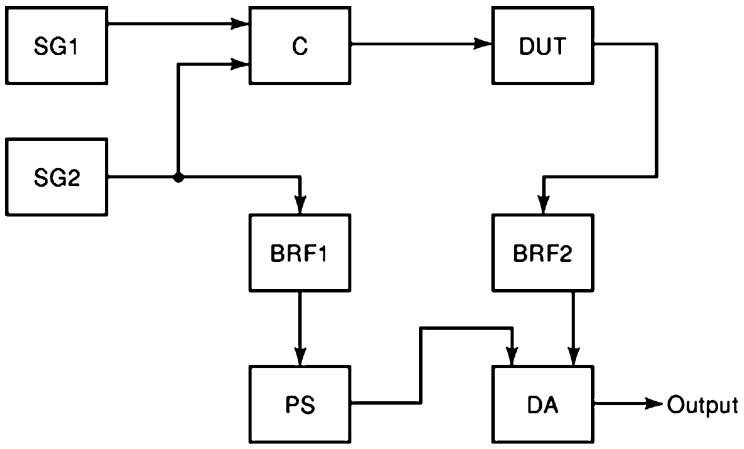

Figure 13. Compensation method for the measurement of small-amplitude intermodulation products in the presence of a large fundamental. SG, signal generator; C, combiner; DUT, device under test; BRF, bandreject filter; PS, phase shifter; DA, differential amplifer.

nents.

Noise-Power Ratio Test. The accuracy of the noise-power ratio (NPR) test is affected mainly by two factors: (1) the noise floor of the amplifier which will dominate under very low loading conditions, and (2) the distortion products produced under very high loading conditions. It is, therefore, recommended to sweep the loading between two prespecified start and stop levels. The NPR is measured at different levels and the largest measured value of NPR is considered as the worst case.

Microcomputer-Based Tests. Quantaization errors associated with the analog-to-digital conversion of the data in microcomputer-based intermodulation tests, must be taken into account. Measurement errors due to quantization are affected by the length of the binary digits and determine the dynamic range of operation [87].

BIBLIOGRAPHY

1. L.E. Kinsler, A.R. Frey, A.B. Coppens and J.V. Sanders, *Fundamentals of Acoustics*, John Wiley & Sons,pp. 267–268, 1982
2. K.Y. Eng and O.-C. Yue, High-order intermodulation effects in digital satellite channels, *IEEE Transactions on Aerospace and Electronic Systems*, Vol.**AES–17**,pp. 438–445, 1981
3. C.D. Bod, C.S. Guenzer and C.A. Carosella, Intermodulation generation by electron tunneling through aluminum-oxide films, *Proceedings of the IEEE*, Vol.**67**,pp. 1643–1652, 1979
4. W.H. Higa, Spurious signals generated by electron tunneling on large reflector antennas, *Proceedings of the IEEE*, Vol.**63**,pp. 306–313, 1975
5. P.L. Aspden and A.P. Anderson, Identification of passive intermodulation product generation in microwave reflecting surfaces, *IEE Proceedings-H*, Vol.**139**,pp. 337–342, 1992
6. P.L. Liu, A.D. Rawlins and D.W. Watts, Measurement of intermodulation products generated by structural components, *Electronics Letters*, Vol.**24**,pp. 1005–1007, 1988
7. M. Otala and J. Lammasniemi, Intermodulation at the amplifier-loudspeaker interface, *Wireless World*, Vol.**86**,pp. 45–47November 1980 andpp. 42–44,55, December 1980
8. E.M. Cherry and G.K. Cambrell, Output resistance and intermodulation distortion in feedback amplifiers, *Journal of the Audio Engineering Society*, Vol.**30**,pp. 178–191, 1982
9. E. Franke, Test setup gauges externally-induced transmitter IM, *Microwave & RF*, Vol.**32**, pp. 95–98,April 1993
10. W. Wharton, S. Metclafe and G.C. Platts, *Broadcast Transmission Engineering Practice*, Butterworth-Heinemann, Oxford, U.K., chapter 5, 1991.
11. J.M. Lindsey, L.S. Riggs and T.H. Shumpert, Intermodulation effects induced on parallel wires by transient excitation, *IEEE Transactions on Electromagnetic Compatibility*, Vol.**31**,pp. 218–222, 1989
12. M. Otala, Non-linear distortion in audio amplifiers, *Wireless World*, Vol.**83**, January,pp. 41–43, 1977
13. E.M. Cherry, Intermodulation distortion in audio amplifiers, *IREE Con. Intl.*, Australia,pp. 639–641, 1983
14. W.G. Jung, M.L. Stephens and C.C. Todd, An overview of SID and TIM-Part I, *Audio*, Vol.**63**, pp. 59–72,June 1979
15. R.R. Cordell, Another vie of TIM, *Audio*, Vol.**64**, pp. 38–49,February 1980
16. W.M. Leach, Transient IM distortion in power amplifiers, *Audio*, Vol.**59**, pp. 34–41,February 1975
17. S.A. Mass, Volterra analysis of spectral regrowth, *IEEE Microwave Guided Wave Letters*, Vol.**7**,pp. 192–193, 1997
18. J.F. Sevic, M.B. Steer, and A.M. Pavio, Nonlinear analysis methods for the simulation of digital wireless communication systems, *International Journal of Microwave Millimeter-wave Computer Aided Design*, Vol.**6**,pp. 197–216, 1996
19. J.F. Sevic, and M.B. Steer, Analysis of GaAs MESFET spectrum regeneration driven by a DQPSK modulated source, *IEEE International Microwave Symposium Digest*, pp. 1375–1378,June 1995
20. K.G. Gard, H.M. Gutierrez, and M.B. Steer, Characterization of spectral regrowth in microwave amplifiers based on the nonlinear transformation of a complex Gaussian process, *IEEE Transactions on Microwave Theory and Techniques*, Vol.**47**,pp. 1059–1069, 1999
21. G.T. Zhou, Analysis of spectral regrowth of weakly nonlinear amplifiers, *IEEE Communications Letters*, Vol.**4**,pp. 357–359, 2000
22. W.M. Leach, Suppression of slew rate and transient IM distortions in audio power amplifiers, *Journal of the Audio Engineering Society*, Vol.**25**,pp. 466–473, 1977

23. M. Schetzen, *The Volterra and Wiener theories of nonlinear systems*, J. Wiley & Sons, 1980
24. E.V.D. Eijnde, J. Schoukers, Steady-State analysis of a periodically excited nonlinear system, *IEEE Transactions on Circuits and Systems*, Vol.**37**,pp. 232–242, 1990
25. S. Naryanan, Transistor distortion analysis using the Volterra series representation, *Bell System Technical Journal*, Vol.**46**,pp. 999–1024, 1967
26. D.D. Weitner and J.F. Spina, Sinusoidal analysis and Modeling of Weakly Nonlinear Circuits,Van Nostrand, New York, 1980
27. P. Harrop and T.A.C.M. Claasen, Modelling of an FET mixer, *Electronics Letters*, Vol.**14**,pp. 369–370, 1978
28. W.G. Jung, M.L. Stephens and C.C. Todd, An overview of SID and TIM-Part III, *Audio*, Vol.**63**, pp. 42–59,August 1979
29. M.T. Abuelma'atti, Prediction of the transient intermodulation performance of operational amplifiers, *International Journal of Electronics*, Vol.**55**,pp. 591–602, 1983
30. S.A. Mass, Nonlinear Microwave Circuits, Arech House, 1988
31. P. Wambacq and W. Sansen, *Distortion analysis of analog integrated circuits*, Kluwer Academic Publishers, Boston, 1998
32. S.A. Mass, Applying Volterra-series analysis, *Microwave & RF*, Vol.**38**,pp. 55–64, 1999
33. D. Atherton,Nonlinear control engineering-describing function analysis,VanNostrand-Reinhold,New York, 1975
34. S. Collins and K. Flynn, Intermodulation characteristics of ferrite-based directional couplers, *Microwave Journal*, Vol.**42**, pp. 122–130,November 1999
35. M. Bayrak and F.A. Benson, Intermodulation products from nonlinearities in transmission lines and connectors at microwave frequencies, *Proceedings of the IEE*, Vol.**122**,pp. 361–367, 1975
36. M.B. Amin and F.A. Benson, Nonlinear effects in coaxial cables at microwave frequencies, *Electronics Letters*, Vol.**13**,pp. 768–770, 1977
37. K.Y. Eng and O.C. Yue, High-order intermodulation effects in digital satellite channels, *IEEE Transactions on Aerospace and Electronic Systems*, Vol.**AES–17**,pp. 438–445, 1981
38. P.L. Aspden and A.P. Anderson, Identification of passive intermodulation product generation on microwave reflecting surfaces, *IEE Proceedings-H*, Vol.**139**,pp. 337–342, 1992
39. M. Lang, The intermodulation problem in mobile communications, *Microwave Journal*, Vol.**38**, pp. 20–28,May 1995
40. P.L. Lui, A.D. Rawlins, and D.W. Watts, Measurement of intermodulation products generated by structural components, *Electronics Letters*, Vol.**24**,pp. 1005–1007, 1988
41. B.G.M. Helme, Passive intermodulation of ICT components, *IEE Colloquium on Screening Effectiveness Measurements*, pp. 1/1–1/8, 1998
42. P.L. Lui and A.D. Rawlins, Passive nonlinearities in antenna systems, *IEE Colloquium on Passive Intermodulation Products in Antennas and Related Structures*, pp. 6/1–6/7, 1989
43. J.T. Kim, I.-K. Cho, M.Y. Jeong, and T.-G. Choy, Effects of external PIM sources on antenna PIM measurements, *ETRI Journal*, Vol.**24**, pp. 435–442,December 2002
44. J.G. Gardiner and H. Dincer, The measurement and characterisation of non-linear intercations among emissions from communal transmitting sites, *Proceedings of the Second International Conference on Radio Spectrum Conservation Techniques*,IEE Publication #224,pp. 39–43, 1983
45. M.T. Abuelma'atti, Prediction of passive intermodulation arising from corrosion, *IEE Proceedings- Science Measurememt and Technology*, Vol.**150**,pp. 30–34, 2003
46. R. Blake,Anti-alias filters: the invisible distortion mechanism in digital audio, Reprint 4966, 106th Audio Engineering Society Convention, Munich 1999
47. R.A. Witte, *Spectrum and Network Measurements*, Prentice Hall, New Jersey, Chapter 7, 1991
48. S. Hunziker and W. Baechtold, Simple model for fundamental intermodulation analysis of RF amplifiers and links, *Electronics Letters*, Vol.**32**,pp. 1826–1827, 1996
49. G.A.A.A. Hueber, B. Nijholt and H. Tendeloo, Twin-tone tape testing, *Journal of the Audio Engineering Society*, Vol.**24**,pp. 542–553, 1976
50. J. Li, R.G. Bosisio and K. Wu, A simple dual-tone calibration of diode detectors,IEEE Instrumentation and Measurement Technology Conference, Hamamatsu, Japan,pp. 276–279, 1994
51. J.D. Giacomini, Most ADC systems require intermodulation testing, *Electronic Design*, Vol.**40**, #17,pp. 57–65, 1992
52. M. Benkais, S.L. Masson, P. Marchegay, A/D converter characterization by spectral analysis in "dual-tone" mode, *IEEE Transactions on Instrumentation and Measurement*, Vol.**44**,pp. 940–944, 1995
53. B.D. Loughlin, Nonlinear amplitude relations and gamma correction,inK. Mcllwain andC. Dean, Eds., *Principles of Color Television*, New York: Wiley,pp. 200–256, 1956
54. M. Kanno and I. Minowa, Application of nonlinearity measuring method using two frequencies to electrical components, *IEEE Transactions on Instrumentation and Measurement*, Vol.**IM–34**,pp. 590–593, 1985
55. L. Robles, M.A. Ruggero and N.C. Rich, Two-tone distortion in the basilar membrane of the cochlea, *Nature*, Vol.**349**,pp. 413–414, 1991
56. T. Maseng, On the characterization of a bandpass nonlinearity by two-tone measurements, *IEEE Transactions on Communications*, Vol.**COM–26**,pp. 746–754, 1978
57. H. Roering, The twin-tone distortion meter: a new approach, *Journal of the Audio Engineering Society*, Vol.**31**,pp. 332–339, 1983
58. E.M. Cherry, Amplitude and phase intermodulation distortion, *Journal of the Audio Engineering Society*, Vol.**31**,pp. 298–303, 1983
59. H.H. Scott, Audible audio distortion, *Electronics*, Vol.**18**,pp. 126,January 1945
60. A.N. Thiele, Measurement of nonlinear distortion in a bandlimited system, *Journal of the Audio Engineering Society*, Vol.**31**,pp. 443–445, 1983
61. G.L. Heiter, Characterization of nonlinearities in microwave devices and systems, *IEEE Transactions on Microwave Theory and Techniques*, Vol.**MTT–21**,pp. 797–805, 1973
62. A.D. Broadhurst, P.F. Bouwer and A.L. Curle, Measuring television transposer intermodulation distortion, *IEEE Transactions on Broadcasting*, Vol.**34**,pp. 344–355, 1988
63. J.C. Pedro and N. B. de Carvalho, Analysis and measurement of multi-tone intermodulation distortion of microwave frequency converters, *IEEE International Symposium on Microwave Theory and Techniques*, 2001
64. J. Lerdworatawee and W. Namgoong, Revisiting spurious-free dynamic range of communication receivers, *IEEE Transactions on Circuits and Systems-I: Regular Papers*, Vol.**53**, 2006, pp. 937–943

65. E. Czerwinski, A. Voishvillo, S. Alexandrov and A. Terekhov, Multitone testing of sound system components – Some results and conclusions, Part I: History and theory, *Journal of the Audio Engineering Society*, Vol.**49**, 2001, pp. 1011–1048
66. E. Czerwinski, A. Voishvillo, S. Alexandrov and A. Terekhov, Multitone testing of sound system components – Some results and conclusions, Part II: Modeling and application, *Journal of the Audio Engineering Society*, Vol.**49**, 2001, pp. 1181–1192
67. A. Potchinkov, Low-crest-factor multitone test signals for audio testing, *Journal of the Audio Engineering Society*, Vol.**50**, 2002, pp. 681–694
68. B. Hessen-Schmidt, Test set speeds NPR measurements, *Microwaves & RF*, Vol.**33**,pp. 126–128,January 1994
69. B. Arnold, Third order intermodulation products in a CATV system, *IEEE Transactions on Cable Television*, Vol.**CATV–2**,pp. 67–79, 1977
70. O.A. Dogha and M.B. Das, Cross-modulation and intermodulation performance of MOS-FET's in tuned high-frequency amplifiers, *International Journal of Electronics*, Vol.**45**,pp. 307–320, 1978
71. J.H. Foster and W.E. Kunz, Intermodulation and crossmodulation in travelling-wave tubes, *Conference International Tubes pour Hyperfrequences*, Paris,pp. 75–79, 1964
72. "Differential phase and gain at work", Hewlett-Packard Application Note 175–1, 1975
73. J. Smith, *Modern communication circuits*, McGraw-Hill, N.Y., chapter 3, 1987
74. J. Dyer, The facts and figures of HF receiver performance, *Electronics World+Wireless World*, Vol.**99**,pp. 1026–1030, 1993
75. U.L. Rohde and D.P. Newkirk, RF/Microwave Circuit Design for Wireless Applications, John Wiley & Sons, New York, 2000
76. G. Steiner, W. Baechtold and S. Hunziker,Bidirectional single fibre links for base station remote antenna feeding, European Conference on Networks & Optical Communications, June 6–9, 2000, Stuttgart, Germany, 2000 [online] http://www.ifh.ee.ethz.ch/~erni/PDF_Paper/paper_NOC2000_PDF.pdf
77. R. Hajji, F. Beauregrd and F. Ghannouchi, Multitone power and intermodulation load-pull characterization of microwave transistors suitable for linear SSPA's design, *IEEE Transactions on Microwave Theory and Techniques*, Vol.**45**,pp. 1093–1099, 1997
78. N.B. Carvalho and J.C. Pedro, Multi-tone intermodulation distortion performance of 3rd order microwave circuits, *IEEE International Microwave Theory and Techniques Symposium Digest*, pp. 763–766, 1999
79. J.C. Pedro and N.B. Carvalho, On the use of multitone techniques for assessing RF components' intermodulation distortion, *IEEE Transactions on Microwave Theory and Techniques*, Vol.**47**,pp. 2393–2402., 1999
80. N.B. Carvalho and J.C. Pedro, Compact formulas to relate ACPR and NPR to two-tone IMR and IP3, *Mirowave Journal*, Vol.**42**, pp. 70–84,December 1999
81. Hajji, F. Beauregrd and F. Ghannouchi, Multi-tone transistor characterization for intermodulation and distortion analysis, *IEEE International Microwave Theory and Techniques Symposium Digest*, pp. 1691–1694, 1996
82. E. Leinonen, M. Otala, and J. Curl, Method for measuring transient intermodulation distortion (TIM), *Journal of the Audio Engineering Society*, Vol.**25**,pp. 170–177, 1977
83. R.R. Cordell, A fully in-band multitone test for transient internmodulation distortion, *Journal of the Audio Engineering Society*, Vol.**29**,pp. 578–586, 1981
84. P. Skritek, A combined measurement method for both dynamic intermodulation and static nonlinear distortions, *Journal of the Audio Engineering Society*, Vol.**35**,pp. 31–37, 1987
85. S. Takashi and S. Tanaka, A new method of measuring transient intermodulation distortion: Comparison with the conventional method, *Journal of the Audio Engineering Society*, Vol.**30**,pp. 10–16, 1982
86. G. Hamer, S. Kazeminejad and D.P. Howson,Test set for the measurement of IMDs at 900 MHz, IEE Colloquim on Passive Intermodulation Products in Antennas and Related Structures, IEE Digest no. 1989/94, London, 1989
87. T. Sasaki and H. Hataoka, Intermodulation measurement using a microcomputer, *IEEE Transactions on Instrumentation and Measurement*, Vol.**IM–30**,pp. 262–264, 1981
88. P.A. Morton, R.F. Ormondroyd, J.E. Bowers and M.S. Demokan, Large-signal harmonic and intermodulation distortions in wide-bandwidth GaInAsP semiconductor lasers, *IEEE Journal of Quantum Electronics*, Vol.**25**,pp. 1559–1567, 1989
89. S. Mukherjee, Vector measurement of nonlinear transfer function, *IEEE Transactions on Instrumentation and Measurement*, Vol.**44**,pp. 892–897, 1994
90. C. Tsironis, Two tone intermodulation measurements using a computer-controlled microwave tuner, *Microwave Journal*, Vol.**32**, pp. 161–163,October 1989
91. A.A.M. Saleh and M.F. Wazowicz, Efficeint, linear amplification of varying-envelope signals using FET's with parabolic transfer characteristics, *IEEE Transactions on Microwave Theory and Techniques*, Vol.**MTT–33**,pp. 703–710, 1985
92. B. Cheng, Signal generator spectral purity consideration in RF communications testing, *Microwave Journal*, Vol.**42**, pp. 22–32,December 1999
93. S. Ciccarelli, Predict receiver IM in the presence of LO phase noise, *Microwaves & RF*, Vol.**35**,pp. 86–90, 1996
94. A.M. Rudkin (Editor), *Electronic Test Equipment*, Granada, London, chapter 2, 1981
95. Yu. M. Bruk and V.V. Zakharenko, Broadband compensation for dynamic-range measurements by intermodulation, *Instruments and Experimental Techniques*, Vol.**36**, Part 1, #4,pp. 557–562, 1993
96. " *Spectrum Analyzer Series*", Hewlett-Packard Application Note 150–11, 1976
97. V.G. Frenkeland M.S. Shterengas, Auxiliary unit for a spectrum analyzer when measuring intermodulation distortion, *Measurement Techniques*, Vol.**32**, pp. 385–387,April 1989

MUHAMMAD TAHER ABUELMA'ATTI
King Fahd University of
Petrleum and Minerals, Box
203, Dhahran, Saudi Arabia,
31261