Super-Heterodyne Signal Analyzers

Description and Applications

Contents

1. Signal Analysis Background

Introduction

Signal analyzers encompass test and measurement receivers known as spectrum analyzers and vector signal analyzers. The signal analyzer is to the frequency domain what the oscilloscope is to the time domain: a general purpose test instrument that can measure and display electrical signals. Like oscilloscopes, signal analyzers come in varying grades of performance and feature set offerings. For the most part, bandwidth distinguishes one oscilloscope from another. However, signal analyzers can vary in tuning range, bandwidth, dynamic range and measurement accuracy (both frequency and amplitude). The signal analyzer performance criterion or feature set offering needed is dependent on the measurement requirement. In addition, signal analyzers vary by their internal architectures. Some architectures are better suited than others for a particular measurement requirement. Lastly, performance alone does not guarantee success in the measurement. Like any precision instrument, knowing how to optimize the settings goes a long way toward enabling the best performance capability of the signal analyzer.

This application note serves as an introduction to the architecture and proper use of one particular signal analyzer structure: the super-heterodyne signal analyzer. Subjects covered in this application note:

- Link between time domain and frequency domain signal analysis
- Super-heterodyne principle: how the mixing process creates wanted and unwanted responses
- Architectural differences of various super-heterodyne signal analyzers
- RF Chain signal processing: how the gain control elements affect the measurement
- IF Chain signal processing: digital signal processing for detection, bandwidth setting and averaging
- Dynamic range optimization
- Amplitude accuracy
- Spectrum monitoring applications

Time and Frequency Domain Representations of Signals

Signal analyzers are a frequency domain class of test instrumentation. However, before introducing the signal analyzer, it is illustrative to review the more familiar time domain representation of signals.

Suppose the signal of interest is a sinusoid given by the equation

$$s(t) = Asin(2\pi f_o t + \theta)$$

Equation 1-1.

Time, t, is the independent variable, whereas frequency, f_o is a fixed constant value. Everything inside the bracket () is the instantaneous phase. For simplicity, equate the instantaneous phase to x:$x = (2\pi f_o t + \theta)$. As time, t, advances so too does the instantaneous phase, x. When x=π/2, sin(x) is at its maximum and when x = 3π/2, sin (x) is at its minimum value. For x = 0, π, and 2π, sin(x) = 0. This process repeats itself with a period, $T = 1/f_o$ as shown in **Figure 1-1.**

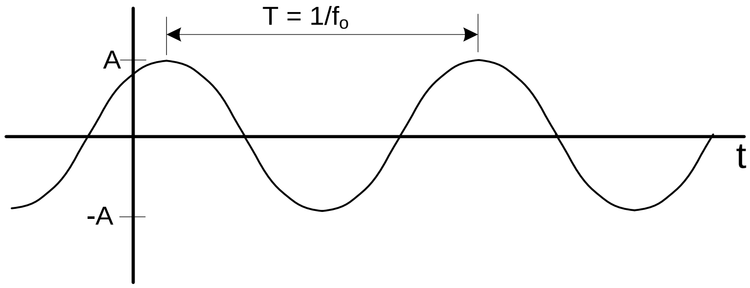

Figure 1-1. Time Domain Representation of a Sinusoid

The sinusoid represented in the time domain is a very familiar to any user of an oscilloscope. The sinusoid has further significance in that a *Fourier series expansion* of any periodic signal is a sum of sine and cosine terms [1]:

$$x(t) = a_0 + 2\sum_{n=1}^{\infty}\left[a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)\right]$$

Equation 1-2.

For example, a square wave represented as a Fourier series expansion has the form [2]:

$$x(t) = \frac{4}{\pi}\sum_{n\ odd}^{\infty}\frac{1}{n}\,sin\left(\frac{2\pi nt}{T}\right)$$

Equation1-3.

As shown in **Figure 1-2**, one can see how a square wave begins to form as the odd harmonics add together.

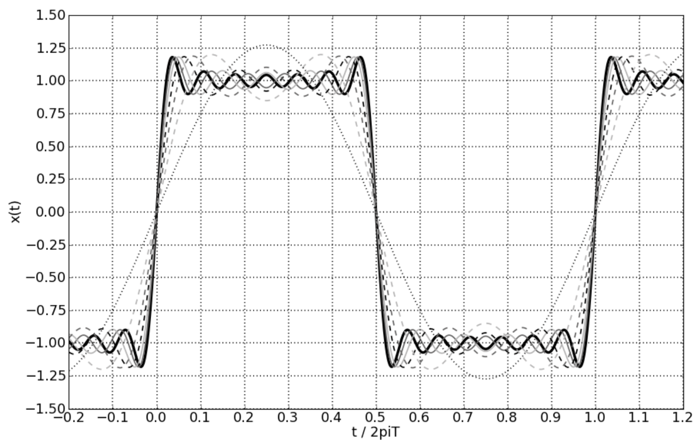

Figure 1-2. Square Wave Fourier Series Expansion

The ability to deconstruct any periodic signal into a series of sinusoids plays an important role in making the leap between the time domain and the frequency domain. Another link between the time and frequency domains is the Fourier transform, itself. The Fourier transform operates on non-periodic time domain signals and is given by **Equation 1-4.**

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi ft} df$$
$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt$$

Equation 1-4.

X(f) is the frequency domain representation of the time domain signal, *x(t)*. Some example transform pairs are shown in **Figure 1-3**. A sinusoid in the time domain translates to a pair of impulses in the frequency domain **Figure 1-3a**. The locations of the impulses are +/- the frequency value of the sinusoid, f_o. A single rectangular pulse in the time domain translates to a sinc function in the frequency domain **Figure 1-3b**. Finally, the pairing of the Fourier series expansion and the Fourier transform is evident in the pulse train. In the frequency domain, this pulse train is a series of impulse functions (sinusoids) in which the envelope of these discrete sinusoids results in the sinc function **Figure 1-3c**.

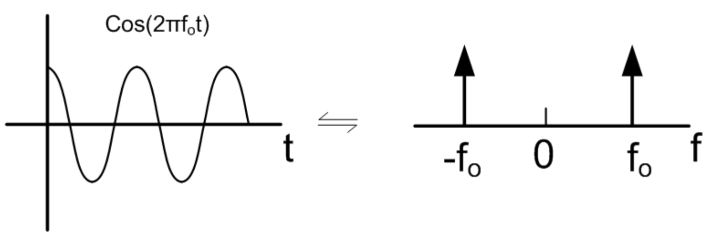

Figure 1-3a

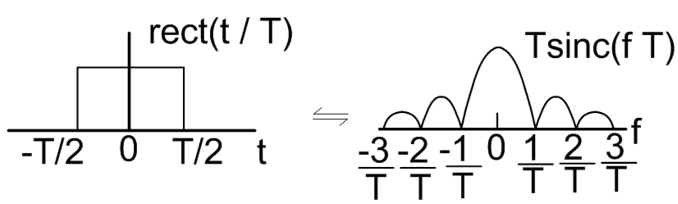

Figure 1-3b

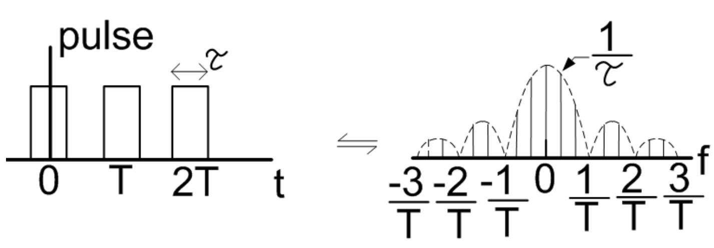

Figure 1-3c

Figure 1-3. Fourier Transform Pairs for Cosine, Single Rectangular Pulse, Periodic Pulse Train

By applying the Fourier transform to each sinusoid component of the square wave represented in **Equation 1-3,** the frequency domain view clearly shows the odd harmonics. **Figure 1-4** shows the time and frequency domain representation of the square wave.

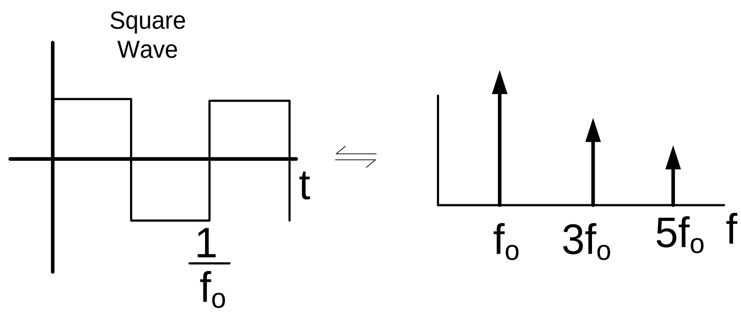

Figure 1-4. Fourier Transform Pair for the Square Wave

Why View Signals in the Frequency Domain?

Just because the Fourier transform allows a signal to be viewed in the frequency domain, why is this necessary? The answer is that examining the signal in the frequency domain allows some added insight in that is just not possible in the time domain. As an example, **Figure 1-5** shows a sinusoid with harmonic perturbations added. Both the time domain and frequency domain views are shown.

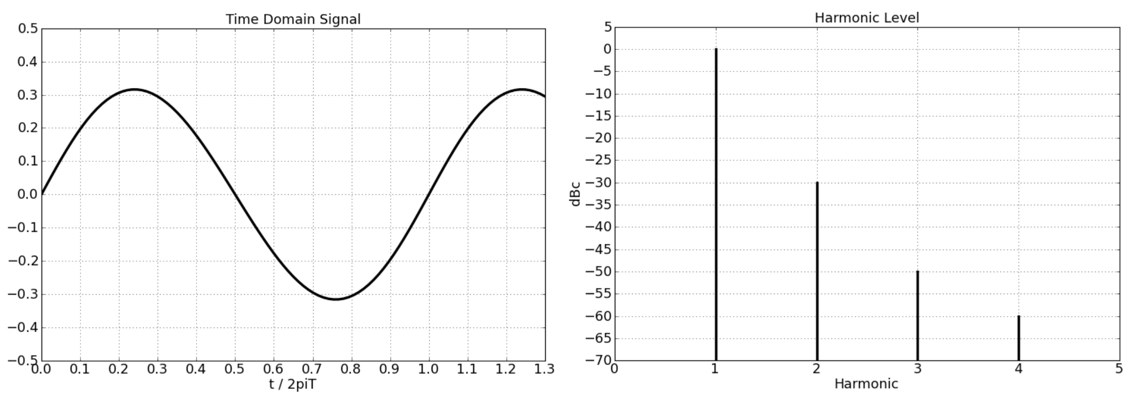

Figure 1-5. Sinusoid with Harmonics

When viewed in the time domain, the signal in **Figure 1-5** appears to be an undistorted sinusoid. It is only in the frequency domain that relatively large harmonics are apparent. In many situations, a -30 dBc 2^{nd} harmonic, which can only be deciphered in the frequency domain, may be unacceptable.

Another situation where frequency domain information is crucial is in diagnosing circuit and system problems. **Figure 1-6** shows the time domain view of a heavily distorted signal. The time domain view only gives the user confirmation that there is in fact a heavily distorted waveform. No information is readily apparent on what could be causing this distortion.

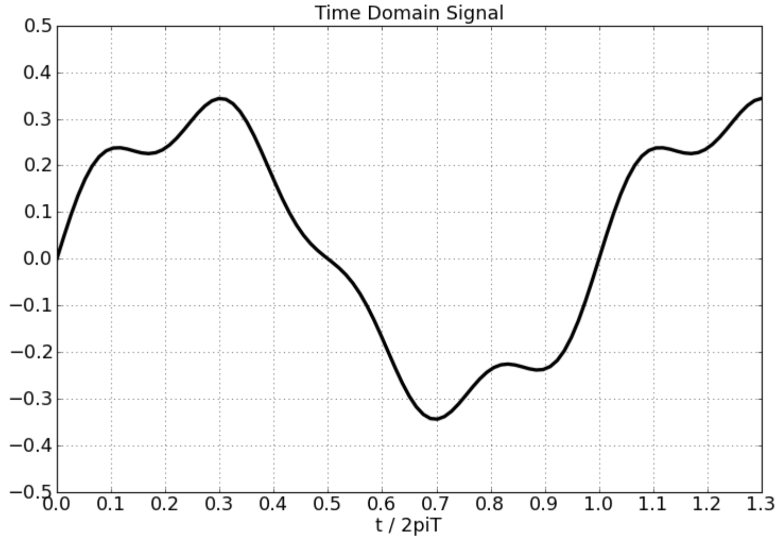

Figure 1-6. Heavily Distorted Signal in the Time Domain

Only in the frequency domain, shown in **Figure 1-7**, is it apparent that most of the problem is a result of large 3rd and 4th harmonic levels.

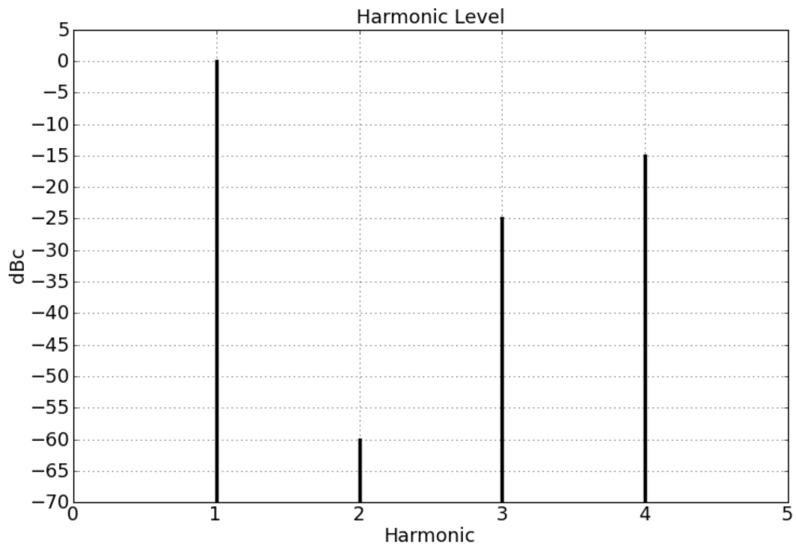

Figure 1-7. Frequency Domain View of the Signal in Figure 1-6

Another situation where the frequency domain might give more insight is in the examination of jitter. Normally associated with digital signals, jitter results from phase noise causing time crossing fluctuations in the rising and falling edges of the digital signal. Jitter in the time domain is shown in the top graph of **Figure 1-8**. Viewing the digital signal in the frequency domain shows the phase noise associated with the jitter. The phase noise as represented in the bottom graph in **Figure 1-8** has a frequency dependent amplitude pedestal.

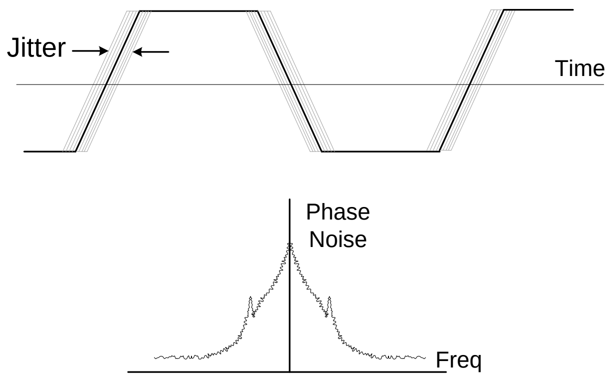

Figure 1-8. Jitter and Phase Noise.

Examining the phase noise pedestal at particular frequency offsets allows better diagnosis of issues leading to large value jitter problems. For example, in **Figure 1-8** the phase noise viewed in the frequency domain shows evidence of a spurious signal. This unwanted spur could be the source of the poor jitter performance. Whether the spur is due to power supply noise, fan vibrations, etc., the frequency domain view gives the user information that is just not possible in the time domain view.

Using the Oscilloscope for Frequency Domain Analysis

Discrete Fourier transform (DFT) and its more efficient implementation, the fast-Fourier transform (FFT), use sampled data to make the transformation between time and frequency domains. Modern oscilloscopes mostly operate on sampled data, so why not just use the oscilloscope and perform an FFT and present the data in the frequency domain? In fact, many oscilloscopes have software that does precisely this operation.

The answer to the question of why not use an oscilloscope lies in the performance requirements of the measurement. Frequency, bandwidth, and dynamic range are the main considerations.

Until recent years the bandwidth of the oscilloscope did not approach RF and microwave frequencies. Hence, a signal analyzer was the only solution for frequency domain examination of signals in the RF and Microwave range. At present, oscilloscopes are available with greater than 30 GHz bandwidth, making bandwidth not as strong of a differentiator between signal analyzers and oscilloscopes. However, price would need to be considered as signal analyzers can be the more cost effective solution for measurements at higher frequencies.

The big differentiator between signal analyzers and oscilloscopes is dynamic range, specifically the spurious-free dynamic range. This is the dynamic range specification needed for the measurement of distortion such as harmonics. Sampled data systems all require Analog-to-Digital Converters (ADC), sometimes referred to as digitizers. Spurious free dynamic range is a function of the digitizer's amplitude resolution, operating frequency, and the sample rate. Because the signal analyzer can frequency shift the RF signal down to a lower digitizer input frequency, the signal analyzer's digitizer can operate at a lower sample rate. Availability of higher resolution digitizers as well as simply operating at a lower digitizer input frequency gives an advantage to the signal analyzer over the oscilloscope in terms of its dynamic range performance. Typically, the differences in dynamic range between the signal analyzer and the oscilloscope can exceed 50 dB.

Oscilloscopes do not normally specify spurious-free dynamic range, whereas signal analyzers highlight this performance specification. This demonstrates the strengths and weaknesses of these two platforms. The oscilloscope is optimized for wide bandwidth to accurately characterize fast slew-rate signals in the time domain, whereas the signal analyzer is optimized to achieve as high a dynamic range as possible for the measurement of signal distortion.

Spectrum Analysis versus Vector Signal Analysis

As previously mentioned, the signal analyzer broadly defines the class of instrumentation known as spectrum analyzers and vector signal analyzers. In the early days of signal analyzers when only narrowband analog modulation formats existed, the spectrum analyzer was the only category. With the advent of digital modulation formats, their relatively wide modulation bandwidths brought forth the need for the vector signal analyzer. The key differentiator between the spectrum analyzer and the vector signal analyzer is the measurement bandwidth.

However, the terms spectrum analyzer and vector signal analyzer, which depict the label of the test instrument, are not always consistent from vendor to vendor. To make matters even more confusing, some spectrum analyzers are capable of making vector signal analyzer type measurements. And some so called vector signal analyzers have spectrum analyzer type traits. The following sections will define both the spectrum analyzer and the vector signal analyzer architectural differences. Because of the inconsistency in nomenclature, the user should rely more on the specifications rather than the label of the instrument.

Rather than dwell on the terms spectrum analyzer and vector signal analyzer, which are labels given by the instrument manufacturer, it is perhaps more instructive to define the measurements. *Vector signal analysis* in the context of this document refers to the measurement of the entire bandwidth of the modulated signal. This type of analysis requires that the analog IF bandwidth of the test receiver be at least as wide as the modulation bandwidth of the signal being measured. The top of **Figure 1-9** shows how the shaded region depicting the test receiver's analog IF bandwidth encompasses the entire signal. Digital modulation metrics such as error vector magnitude (EVM) and complementary cumulative distribution function (CCFD) require capturing the entire modulation bandwidth.

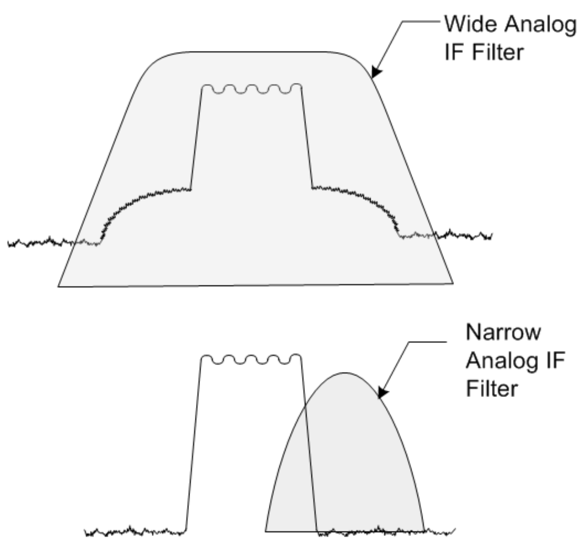

**Figure 1-9. Top: Vector Signal Analysis.
Bottom: Spectrum Analysis**

Spectrum analysis will be used in this document to mean the measurement of the signal using the test receiver's narrow analog IF bandwidths. This is depicted in the bottom portion of **Figure 1-9**. Later sections in this document will explain why this is true, but limiting the IF bandwidth allows for higher dynamic range performance. Measurements such as adjacent channel leakage ratio (ACLR), intermodulation distortion, and harmonic distortion are greatly enhanced with a relatively narrow receiver IF bandwidth.

Super-Heterodyne versus Direct Conversion Architectures

This introductory section concludes with a mention of two major categories used in signal analyzers. The super-heterodyne receiver architecture is the subject of the rest of this document. The other architecture now being used in signal analyzers is the direct conversion receiver, which also goes by the names homodyne receiver and zero IF (ZIF) receiver. This document will not cover the direct conversion receiver in detail; however, a brief mention of this receiver will be given to highlight its difference with the super-heterodyne receiver.

Figure 1-10 shows the basic structure of the direct conversion receiver. A single local oscillator is used to shift the incoming RF signal down to baseband. Baseband contains two paths, I-path and Q-path, corresponding to In-phase and Quadrature paths. Each path is then digitized separately.

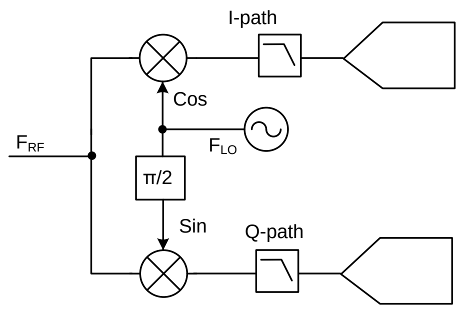

Figure 1-10. Direct Conversion Receiver.

The direct conversion receiver has benefits over the super-heterodyne receiver in terms of bandwidth and compactness, such as only one local oscillator is needed and there are fewer requirements on the RF path filtering. However, the super-heterodyne receiver, in general, is capable of more dynamic range than the direct conversion receiver.

2. Super-Heterodyne Principle

Brief History of the Super-Heterodyne Receiver

The heterodyne principle was coined by its inventor, Reginald Fessenden in 1901 [4]. The term has the Greek roots, heteros, meaning "other" and dynamis, meaning "force". The heteros part refers to the translation to another frequency and the dynamis part refers to the apparent amplification of the detected signal during the heterodyne process. This first instantiation did not resemble modern day receivers; rather, it used two antennas to receive two RF signals. When combined in an envelope detector, these two signals created a beat note: a signal at the difference frequency. This was an audio frequency beat note.

Little progress in this receiver was made until Edwin Howard Armstrong in 1918 was able to develop the idea of using the frequency conversion of higher frequency signals down to the range of the then common heterodyne receiver. He was able to observe that the modulation on a signal did not alter during this frequency conversion process. The super part of super-heterodyne refers to super-sonic, meaning that the heterodyne process was extended above the audio frequency range.

Frequency Shift Property

One property of the Fourier transform is the shift theorem which states that if X(f) is the Fourier transform of x(t), then the Fourier transform of x(t) multiplied by a complex exponential, $e^{j2\pi f_c t}$ results in the frequency domain signal shifted in frequency by the amount -fc [3].

$$\mathcal{F}\{x(t)e^{j2\pi f_c t}\} = X(f - f_c)$$

Equation 2-1

Manipulating Euler's Identity allows a sinusoid to be represented by a pair of complex exponentials [3]. For example, the cosine signal representation is shown in **Equation 2-2**.

$$\cos(2\pi f_c t) = \frac{e^{j2\pi f_c t} + e^{-j2\pi f_c t}}{2}$$

Equation 2-2

By multiplying a time domain function, x(t), with a sinusoid and applying the Fourier transform shift theorem, the result is shown in **Equation 2-3**.

$$\mathcal{F}\left\{x(t)\left[\frac{e^{j2\pi f_c t} + e^{-j2\pi f_c t}}{2}\right]\right\} = \frac{1}{2}[X(f - f_c) + X(f + f_c)]$$

Equation 2-3

In the context of the Fourier transform, **Equation 2-3** is known as the modulation theorem [3]. Multiplying a signal by a sinusoid whose frequency is f_c results in two copies, each with amplitude multiplier of ½, being shifted by $\pm f_c$.

In receivers, the mixer is the device that performs the frequency conversion process. The block diagram for the mixer is shown in **Figure 2-1**.

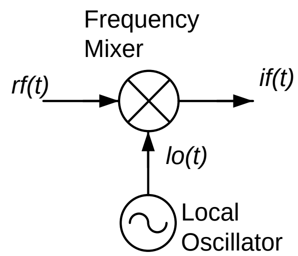

Figure 2-1. Frequency Conversion Process

The local oscillator (LO) is a sinusoid signal that can be tuned in frequency. To a first order approximation, the mixer performs straight time domain multiplication:

$$if(t) = rf(t) \text{ x } lo(t) = rf(t) \text{ x } Cos(2\pi f_{LO} t)$$

Equation 2-4

Applying **Equation 2-3** to **Equation 2-4** results in the frequency domain representation:

$$IF(f) = \frac{1}{2}[RF(f - f_{LO}) + RF(f + f_{LO})]$$

Equation 2-5

Figure 2-2 shows how the frequency mixing appears in the frequency domain. The signal *RF(f)* centered at DC before mixing appears at the IF port as two instances of itself centered at frequencies $-f_{LO}$ and $+f_{LO}$

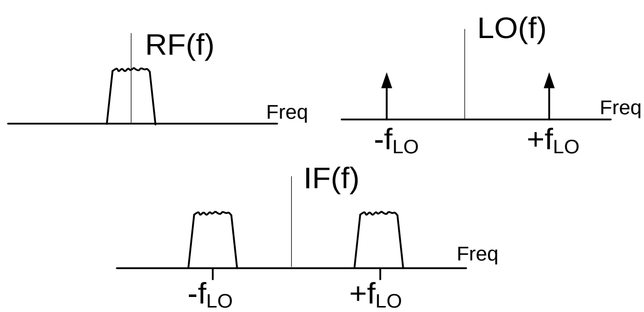

Figure 2-2. Modulation Theorem Applied to RF Signal Mixing

Note that the modulation content in the signal is preserved throughout this process.

Frequency Shift Property Applied to the Super-heterodyne Receiver

Now to see how the frequency shift property can be extended to the super-heterodyne receiver. The fundamental building blocks of the super-heterodyne receiver are shown in **Figure 2-3.**

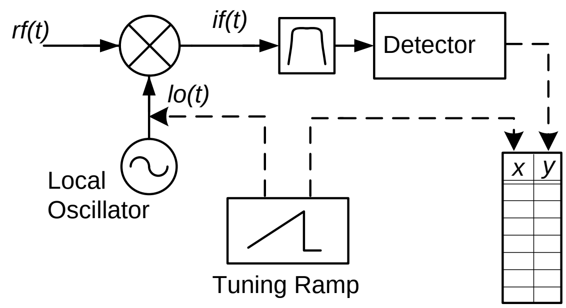

Figure 2-3. Basic Super-heterodyne Receiver

The full super-heterodyne structure adds some IF filtering and some means of converting the IF signal to magnitude and phase data by use of the block labeled *Detector*. The detected IF is converted to a digital value and recorded as y-data. A numeric tuning value used to control the LO frequency is also stored as x-data. The recorded x,y pairs corresponding to frequency / amplitude or frequency / phase data can be used directly to display the frequency domain spectrum of the signal or this data can be used for further signal processing.

Figure 2-4 shows the frequency shifting process in a super-heterodyne receiver.

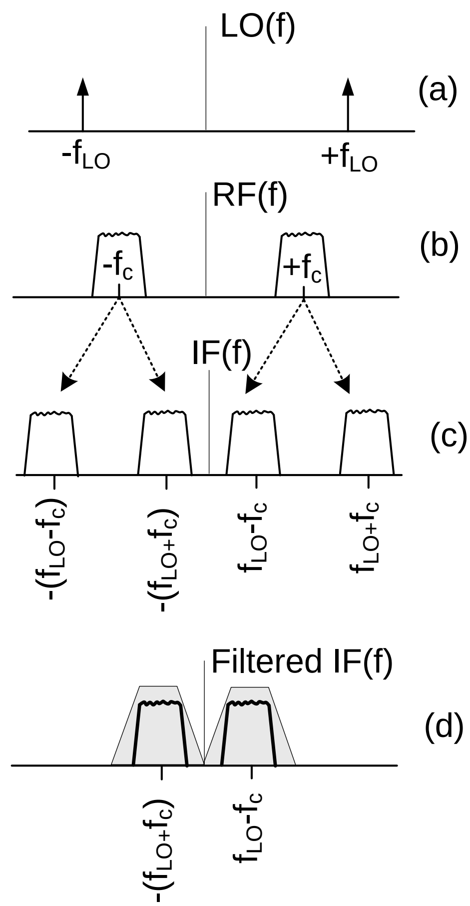

Figure 2-4. Mixing Process in a Super-heterodyne Receiver

Instead being centered at DC, the RF input signal, in general, is centered at some carrier frequency, f_c. The mixer functionally multiplies the LO signal with the RF signal. At the IF port of the mixer, the positive frequency component of the RF signal splits into two copies centered at $f_{LO} - f_c$ and $f_{LO} + f_c$. The negative frequency portion of the RF signal similarly splits into two copies at $-(f_{LO} - f_c)$ and $-(f_{LO} + f_c)$. After IF filtering, only the copies centered at $\pm(f_{LO} - f_c)$ are retained.

The negative frequency content is important for signal processing of complex signals, which is the case for the direct conversion receiver. However, for the analog portion of the super-heterodyne receiver, the signals are real, meaning that there is even magnitude symmetry and odd phase symmetry about DC as shown in the following:

$$|X(f)| = |X(-f)|$$

and

$$phase\big(X(f)\big) = -phase((X(-f))$$

For analyzing the analog portion of the super-heterodyne receiver, concentrating on only the positive portion of the frequency spectrum does not lose any information about the signal.

Moving to a Non-ideal Receiver

The previous sections rely on idealized components to preserve the frequency shift process in a super-heterodyne receiver. However, in actual receivers the components are far from ideal in terms of added noise and distortion to the measurement. Maintaining best measurement performance for dynamic range and measurement accuracy (both amplitude and frequency) requires a careful optimization of the system parameters. But, before discussing how to optimize system parameters, knowledge of the sources of the distortion and noise will be presented.

Mixing Process

Figure 2-5 once again reveals the basic pieces required to understand the frequency conversion (or mixing) process. Compared to **Figure 2-1**, a bandpass filter centered at the desired IF frequency has been added at the output port of the mixer.

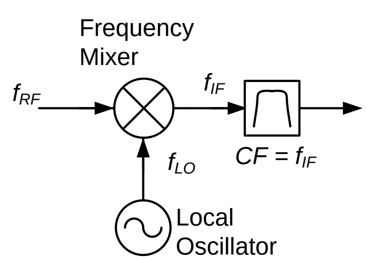

Figure 2-5. Basic Frequency Mixer

The *mixing equation*, shown in **Equation 2-6,** is a simplification of the frequency shift property.

$$Mxf_{RF} + Nxf_{LO} = f_{IF}$$

where, M and N = 0, ±1, ±2, ±3, ...

Equation 2-6

The RF signal "mixes" or frequency shifts to an IF signal using a sinusoid LO signal. Whenever the mixing equation is satisfied, an IF signal passes through the bandpass filter and is recorded as a response. But, there are an infinite number of solutions that satisfy the mixing equation, resulting in an infinite number of possible wanted and unwanted signals converting to IF.

Normally, the receiver is calibrated to be accurate for only a single pair of M, N values. All other M,N combinations that cause a response to fall inside the bandpass filter are unwanted and are known as spurious responses, or spurs for short. Many of these spurs are given specific names such as image responses. There are far more spurs than wanted responses and by understanding the spur mechanisms the user can avoid having the receiver's spur mask the signal being measured.

First, the mixing process for desired mixing responses will be examined. Suppose the signal analyzer is calibrated to respond to M,N = -1,1 and let the IF be centered at 100 MHz. Assume that a fixed frequency sinusoid, which is referred to a continuous wave (CW) signal, at a frequency of 1 GHz is present at the RF input port. The mixing equation simplifies to:

$$-f_{RF} + f_{LO} = f_{IF}$$

Equation 2-7

Solving **Equation 2-7** shows that at an LO frequency of 1100 MHz, the mixing equation is solved, resulting in an IF signal at a frequency of 100 MHz:

$$-1000 \text{ MHz} + f_{LO} = 100 \text{ MHz}$$
$$-1000 \text{ MHz} + 1100 \text{ MHz} = 100 \text{ MHz}$$

The spectrum view of this mixing process is shown in **Figure 2-6**.

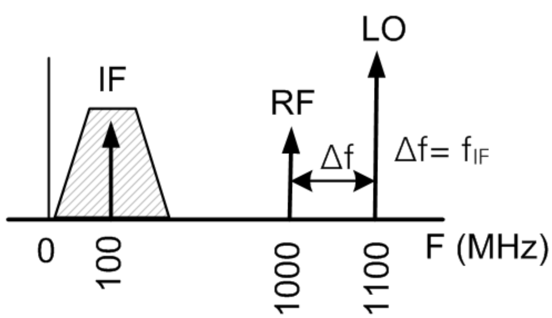

Figure 2-6. Spectrum Showing the Mixing Process for a Wanted Mixing Product

In a signal analyzer, often times the data is gathered over a range of frequencies. The LO is either stepped or swept in frequency. With a fixed frequency RF signal, the resulting IF, after applying the mixing equation, will also be swept or stepped. Furthermore, the shape of the IF filter will be traced out in the measured amplitude data. The process of stepping the LO over a range of frequencies with a fixed frequency RF input signal is shown in **Figure 2-7.**

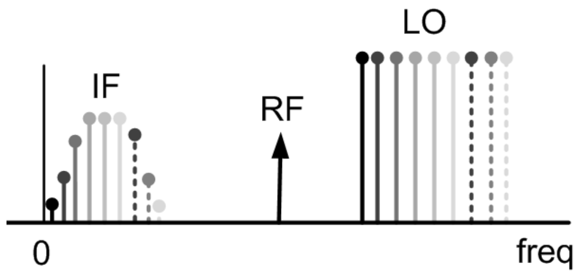

Figure 2-7. Mixing Process with Swept LO

Image Responses

Previously, the wanted M,N = -1,1 mixing product was analyzed. But the mixer will also respond to the M,N = 1,-1 product as well. Using the above example where the LO was tuned to 1100 MHz in order to convert a 1000 MHz RF signal down to an IF of 100 MHz, suppose another RF signal was present at 1200 MHz. In this case with M,N = 1,-1 the mixing equation simplifies to:

$$f_{RF} - f_{LO} = f_{IF}$$

or, 1200 MHz – 1100 MHz = 100 MHz

Equation 2-8

Figure 2-7 shows the spectrum with both the wanted RF at 1000 MHz mixing down to the 100 MHz IF as well as the unwanted RF signal at 1200 MHz also mixing down to the 100 MHz IF.

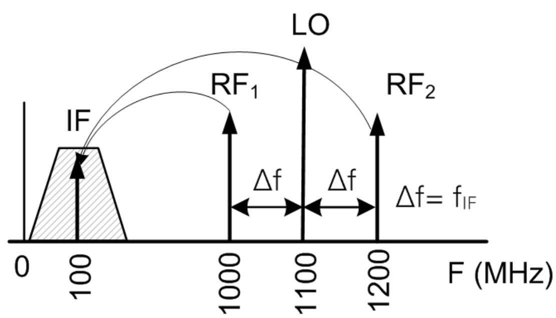

Figure 2-7. Desired Signal and its Image Response.

The unwanted spur at 1200 MHz is known as the image response. The image response, as evident from **Figure 2-7**, falls two times the IF frequency away from the desired response.

In **Figure 2-7,** the LO frequency is fixed and RF either one-IF below or one-IF above can cause a response. **Figure 2-8** shows another situation where the image response causes non-deterministic results. In this situation, the LO is tuning in order to create a swept spectrum display. In this example, a single RF signal at 1 GHz is present at the input port.

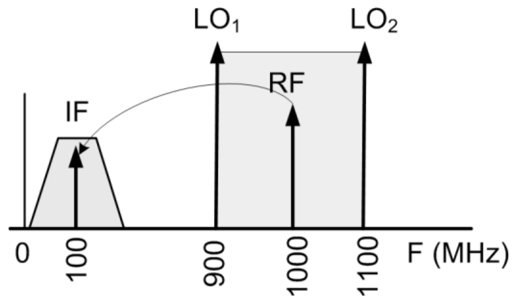

Figure 2-8. Single RF Causing an Image Response

When the signal analyzer is tuned to 800 MHz, the corresponding LO frequency is 900 MHz. The RF and LO signal mix using M,N = 1,-1 to create an IF response at 100 MHz. As the LO frequency is stepped, eventually it reaches 1100 MHz. This time the RF and LO mix using M,N = -1,1 once again creating a response at 100 MHz centered IF. The spectrum display of this scenario is shown in **Figure 2-9**.

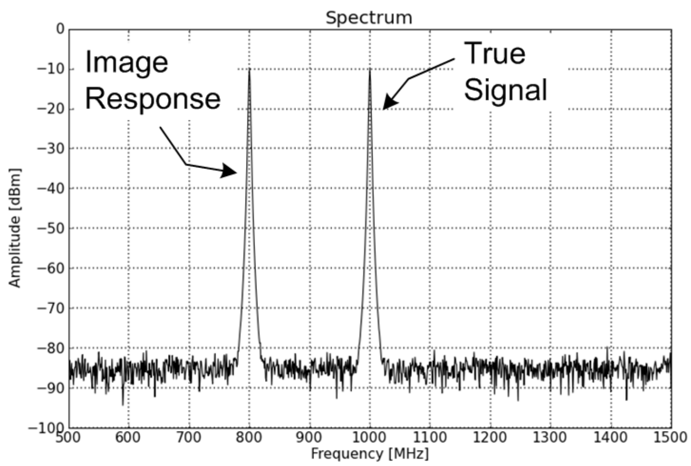

Figure 2-9. Spectrum of Single RF Causing Image

The two responses are separated by two IFs (200 MHz) in frequency and also note that their amplitudes are nearly the same. Remember, only a single tone at 1 GHz is input, yet two responses are present on the display. The 800 MHz displayed response, the image, is not real; however, the operator would not be able to decipher the image from the true response.

In signal analyzers with multiple frequency conversion stages, there will be images associated with each stage. For each stage, images will be present and will be spaced away from the desired signal by twice the IF frequency of that stage. Most often, the LO and IF frequencies in the later frequency conversion stages are fixed, so predicting the image frequency is not as difficult as with frequency conversion stages where the LO frequency is variable. The amount of suppression of the image signal amplitude is termed i*mage rejection* and applies to all the mixer stages in the system.

IF Subharmonics

Another spur mechanism results from signals mixing to sub-multiples of the IF frequency. The mixing equation, shown in **Equation 2-6,** can be modified to include IF subharmonics by adding the qualifier, Q:

$$M\text{x}f_{RF} + N\text{x}f_{LO} = Q\text{x}f_{IF}$$

where, M and N = ±1, ±2, ±3, ...
and Q = 1/2, 1/3, . . .

Equation 2-9

The RF and LO signals can combine in a way to create so called subharmonic signals at the output of the mixer that are at frequencies of $f_{IF}/2$, $f_{IF}/3$, etc. If not adequately filtered, the harmonics of these subharmonic signals may fall at the desired IF frequency. Harmonics can be generated in nonlinear stages in the IF such as an IF Amplifier. **Figure 2-10** shows this process.

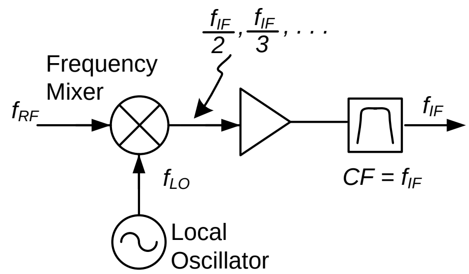

Figure 2-10. Subharmonics in the IF

Take the example where f_{IF} = 100 MHz, and f_{LO} = 1100 MHz. In this case, an RF signal at 1050 MHz will mix with the LO to create a 50 MHz response at the IF port. A second harmonic of an IF amplifier is 100 MHz, which propagates through the IF filter and records as a response. **Figure 2-11** shows the spectrum for the scenario described here.

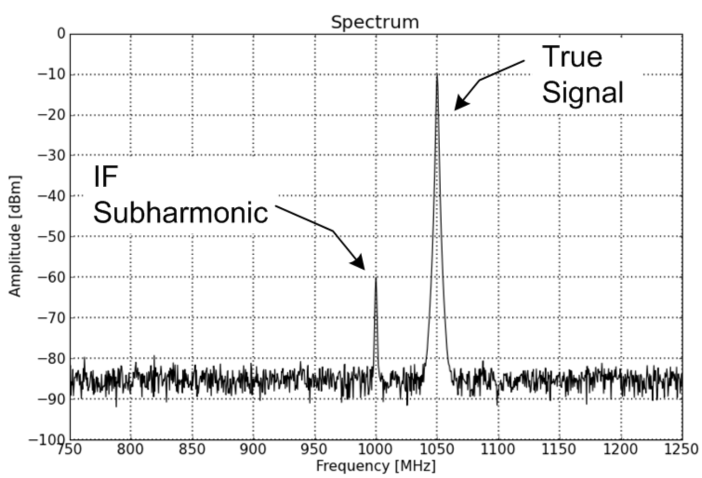

Figure 2-11. Spectrum Showing IF Subharmonic at a Displayed Frequency of 1000 MHz

In this example, the true signal is at 1050 MHz. However, when the analyzer is tuned to 1000 MHz, this 1050 MHz signal mixes to the IF such that a subharmonic is created. Harmonic distortion of the IF chain multiplies the subharmonic to the IF where it is displayed as an unwanted spur.

General M,N Spurs

Considering other M,N combinations within the context of the mixing equation **2-6**, many other spurious responses can be generated. Careful design of the receiver can try to minimize their amplitudes or better yet filtering can be added to remove these spurs completely. However, the most careful design will still have conditions that can lead to the general spurious response.

Luckily, the amplitude of the spurious signal falls as a function of the spur order as shown in the following:

Spur Order = |M| + |N|

Higher order spurs, in general, have lower amplitudes than lower order spurs. This at least bounds the problem of infinite number of possible spurs, since the vast majority will fall below the noise floor of the signal analyzer and will never be seen.

Consider the case of the 2,-1 spur created from a 600 MHz signal at the RF port. Again, the IF is centered at 100 MHz. When the signal analyzer is tuned to 1000 MHz, the corresponding LO frequency is 1100 MHz. The mixer equation with M,N = 2,-1 is satisfied as shown by the following:

$$2\text{x}f_{RF} - f_{LO} = f_{IF}$$
$$2\text{x}600\text{ MHz} - 1100\text{ MHz} = 100\text{ MHz}$$

Figure 2-11 shows the spectrum of this scenario.

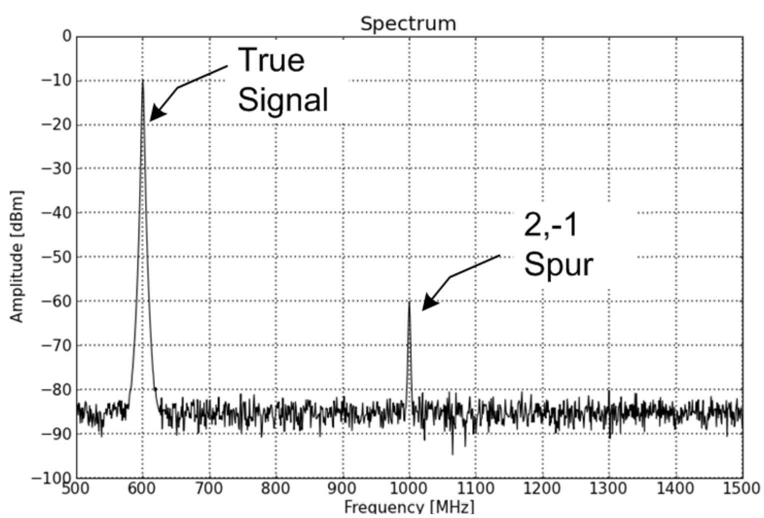

Figure 2-12. Spectrum Showing the 2,-1 Spur of a 1000 MHz Input Signal

When the signal analyzer is tuned to 600 MHz, the true response is displayed. However, when the analyzer is tuned to 1000 MHz, the 600 MHz signal generates an M,N = 2,-1 spur that is displayed at 1000 MHz.

There are also some redundant spurs. For example, consider the frequencies used to generate the IF subharmonic distortion in **Figure 2-11**: f_{RF} = 1050 MHz, f_{LO} = 1100 MHz, and f_{IF} = 100 MHz. The LO frequency corresponds to a tune frequency of 1000 MHz. The M,N = -2,2 spur is:

$$-2\text{x}f_{RF} + 2\text{x}f_{LO} = f_{IF}$$
$$-2\text{x}1050\text{ MHz} + 2\text{x}1100\text{ MHz} = 100\text{ MHz}$$

The spectrum of this situation is exactly as depicted in **Figure 2-11**. So, a single RF signal is now responsible for multiple spur mechanisms.

IF Feedthrough

When the frequency at the RF port equals the IF port frequency, the RF signal can leak through the mixer, circumventing the mixing process. In other words, independent of the LO frequency, the RF signal is present at the IF port. **Figure 2-13** shows this spurious mechanism in the block diagram.

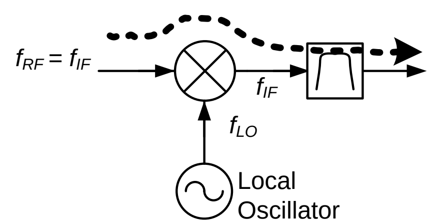

Figure 2-13. Block Diagram View of IF Feedthrough

The term given to this type of spur is IF feedthrough. Traditionally, this spur was also known as baseline lift as it manifested in a dramatic increase in the apparent displayed noise floor.

Using 100 MHz as the IF frequency, if the RF frequency is 100 MHz, no matter where the LO frequency is tuned, the spur will be present.

Signal analyzer structures put in place filtering to remove as much of the IF feedthrough as possible. However, filters do not have infinite rejection and some of the IF feedthrough does leak into the final IF. The performance specification given to the amount of suppression of the IF feedthrough signal is termed *IF rejection*.

LO Feedthrough

Using the mixer equation, when the tune frequency is set to 0 Hz, the LO frequency is set to the IF frequency. Using the IF and LO frequencies from the previous examples, the following shows this:

$$-f_{RF} + f_{LO} = f_{IF}$$
$$0 \text{ MHz} + 100 \text{ MHz} = 100 \text{ MHz}$$
$$\text{or, } f_{LO} = f_{IF}$$

When this situation occurs, the LO signal can leak though the mixer and appear at the IF. Since these frequencies line up, the result is a spurious response displayed at DC. **Figure 2-13** shows this leakage path on the simplified signal analyzer block diagram.

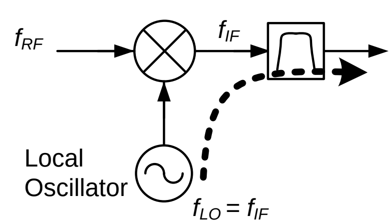

Figure 2-13. Block Diagram Showing LO Feedthrough

Figure 2-14 shows how LO leakage is manifested on the spectrum display as a response at a tune frequency of zero Hz.

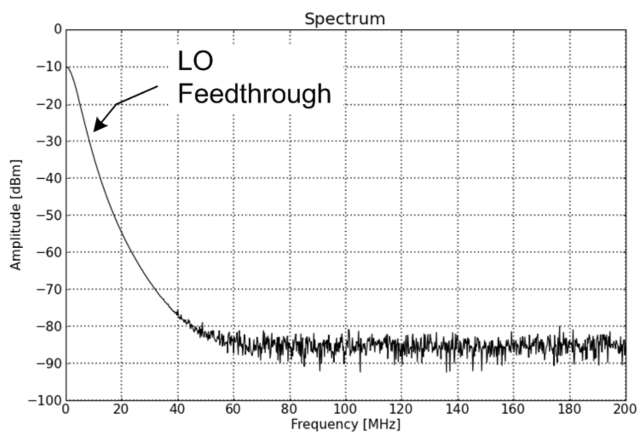

Figure 2-14. Spectrum Showing LO Feedthrough

LO Emissions

LO emissions is the term used to describe the LO signal leaking back through to the RF port of the mixer and showing up as an output signal at the RF input port of the receiver. This leakage path is shown in **Figure 2-15**.

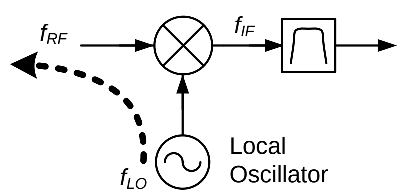

Figure 2-15. Block Diagram Showing LO Emissions.

Although not a spurious response per se, a signal being emitted from the input port of a receiver is normally an unexpected occurrence. In some measurement situations, this unexpected signal going into the output port of the device under test (DUT) may be undesirable. For spectrum monitoring where an antenna is connected to the input port, this leakage signal has now made a beacon out of the signal analyzer.

Residual Responses

Besides the LO, there are other sources of spurious signals in the signal analyzer. Reference oscillators for the LO phase lock loop (PLL), switching power supplies, and calibration signals can all generate signals. These signals can find various paths to the final IF, either through leakage in the signal paths or even coupling between control and power supply routing traces. Once this signal falls into the final IF, a response is recorded. **Figure 2-16** shows one potential leakage path: one of the LO PLL reference oscillator signals.

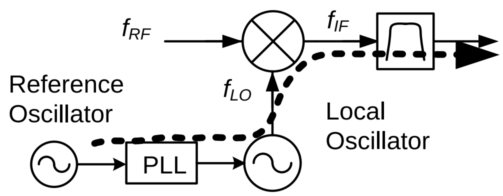

Figure 2-16. Residual Responses

These internally generated responses can occur when there is no signal present at the RF input port of the signal analyzer. These signals are termed *residual responses* and are usually specified with the RF input port terminated in the signal analyzer's characteristic impedance (normally 50 Ohms).

Cautionary Note on Nomenclature

Some of the spurious response terms used in the context of the super-heterodyne receiver are also used in direct conversion receivers, yet have different meanings. Image response and LO feedthrough are two of the terms that are used in both architectures, yet are defined differently. For the direct conversion receiver **Figure 1-10** the I and Q paths must be matched in amplitude response and their relative phases must be separated 90 degrees. Unbalance between relative amplitudes and phases between these two path results in spurious responses. The IF of the direct conversion receiver is split into two halves. The LO can leak through and cause a spurious response at the center of the IF. This is termed LO feedthru. A signal present in one half of the IF can result in an amplitude suppressed version of itself in the other half of the IF. This is termed image response. One should be aware that these terms have different meanings for the two different receiver architectures and care should be applied when comparing instrument specifications between these receiver architectures.

3. Super-Heterodyne Signal Analyzer Structures

Section 2 discussed the mixing process and many of its associated spurious response mechanisms. In this section, some of the more popular super-heterodyne systems will be shown. In these systems, attempts are made to minimize the impact of spurious responses. This section will concentrate on the frequency of the signal as it progresses though the signal analyzer, and later sections will concentrate on the amplitude of the signal.

Single Conversion Stage Structure

The most basic structure is the so called *single stage downconverter*. **Figure 3-1** shows the block diagram of this structure.

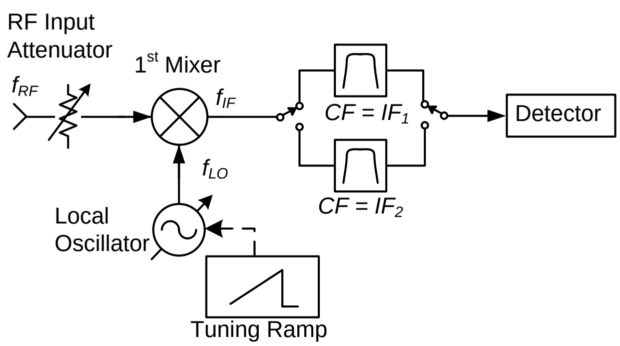

Figure 3-1. Single Stage Downconverter Block Diagram.

In the single stage downconverter, only one LO and one mixer stage are used to convert the RF input signal to the final IF. The final IF frequency is, in most cases, not at baseband, meaning the final IF is bandpass filtered rather than lowpass filtered. The purpose of the multiple filters in the IF is to suppress the IF feedthrough response. If IF_1 is the default path, at the tune frequency that equals the IF_1 frequency, then IF_2 path is selected.

The one major spur that is not suppressed is the image. Due to the lack of image signal suppression, the intent for this structure is not for general purpose applications. This structure is especially not appropriate for over-the-air (antenna connected) measurements where the spectrum contains unknown signals. However, this receiver is quite well suited for manufacturing test applications. In these applications, the signal under test is known in frequency and the connection to the DUT is normally though a shielded cabled which greatly attenuates the reception of unknown signal. Both the known frequency nature and the shielded connection make deciphering the signal analyzer spurious responses from the true response very predictable.

The advantage of this structure is its compactness and cost. Requiring a single LO and a single mixer results in a relatively simple design which tends to drive down both size and cost. In some applications, multiple downconverters are required and often phase coherence must be maintained. One way of ensuring phase coherence, meaning all receivers have known and predictable relative phase versus frequency behavior, is to share the LO where one downconverter is the master (LO source) and the downstream downconverters are the slaves (consumers of the LO). In such systems, routing only one LO signal via coaxial cables between downconverters is much simpler and less costly than systems with multiple LOs and mixer stages.

Multiple Conversion Lowband Structure

The multiple conversion structure attempts to overcome the lack of image rejection present in the single stage structure. The multiple stage structure is broadly divided into two different types: lowband and highband.

Figure 3-2 shows the block diagram of the lowband multiple stage super-heterodyne receiver. This architecture is designed to accept RF input frequencies ranging from near DC to RF. Typical maximum RF frequencies are 3 GHz to 7 GHz. True DC is impossible as the double balanced mixer used for the first conversion stage cannot accept more than a few hundred kilovolts without risk of damage. However, operation down to a few Hz is possible as long as the DC component is small.

The first LO has variable frequency that is configured to up convert RF signals to a fixed frequency first IF. The remaining LOs are fixed in frequency and are used with the later frequency conversion stages to progressively down convert the first IF to a final IF that can efficiently be detected.

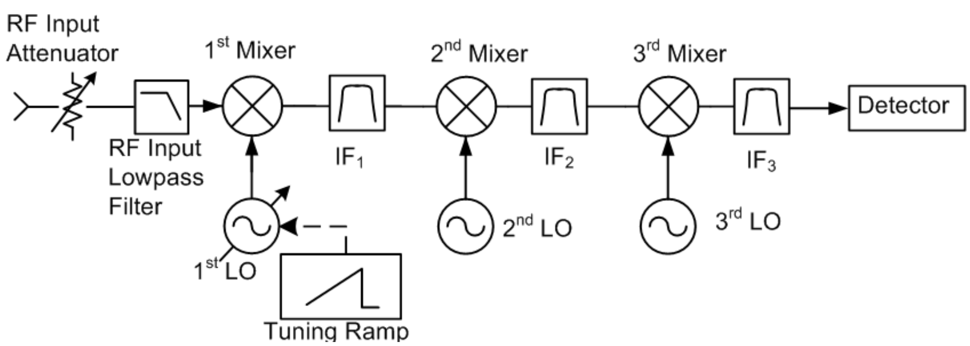

Figure 3-2 Triple Conversion Lowband Structure Block Diagram.

The main features of the lowband structure are:

- High image response suppression
- Low IF feedthrough response
- Low LO emissions

The first IF placed higher than the maximum input RF frequency enables all of these features. The first converter mixing equation governing this structure is shown in **Equation 3-1.**

$$-f_{RF} + f_{LO} = f_{IF}$$

Equation 3-1

With a high first IF, **Equation 3-1** implies that the LO frequency range is above both the IF and the maximum RF input frequency. **Figure 3-3** shows the relationship between the RF, LO and IF frequency ranges.

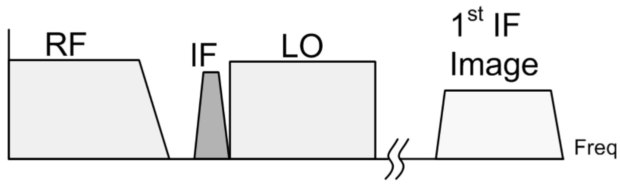

Figure 3-3. Spectrum of Lowband Frequencies

As an example, if the input RF frequency range is near DC to 3.6 GHz, and the IF is 4.6 GHz, the resulting LO frequency range is 4.6 GHz to 8.2 GHz.

Recall that the image responses are separated in frequency by twice the IF. Using the above frequency range values, the image of the first IF ranges from 9.2 GHz to 12.8 GHz.

In **Figure 3-2**, the lowpass filter in the RF input path is responsible for reducing the spurious signal content resulting from the following three mechanisms:

1. First IF image. The images are pushed far enough out in frequency in relation to the input RF frequency range that the lowpass filter can easily attenuate signals in the image frequency range.
2. First IF feedthrough. The challenge for the lowpass filter is to minimize the attenuation of the RF signals in the RF input range and to provide a great deal of attenuation at the first IF frequency. The ratio between the first IF frequency and the maximum RF input frequency governs the lowpass filter complexity.
3. First LO emissions. Since the first LO frequency range is above the RF input frequency range, the LO signals fall into the stopband of the input lowpass filter and cannot pass on out through to the RF input port.

Why are at least three frequency conversion stages required? The answer is that the first IF frequency is too high to directly down convert to the final IF in one hop. At a 4.6 GHz first IF, using the above example, it is too difficult to attenuate the most challenging M,N spur: the 2,2 spur (actually M,N = -2,2 or 2,-2) generated in the latter mixer stages. This spur is present at one-half the IF frequency away from the input frequency. With low final IF frequencies, this places an unusually high burden on the complexity of filters possessing higher center frequencies. Multiple conversion stages are required to progressively mix the high first IF frequency down to the low final IF frequency.

The filters in the first and second IFs are carefully designed to mostly attenuate the challenging 2,2 spur as well as images and LO emissions. Inter-stage LO emissions can lead to residual responses: the LO and harmonics of the LO from a following stage can leak into the preceding stage. These LO harmonics can mix, possibly converting to one of the IF frequencies. Once converted to one of the IFs, this becomes an unwanted spurious response.

Multiple Conversion Highband Structure

Extending the lowband idea to higher input frequencies requires ever higher first IF and first LO frequencies. Challenges with filtering, dynamic range, first LO phase noise and above all cost become relevant at higher frequencies. The highband structure is normally used in conjunction with the lowband structure. The highband's minimum frequency starts at the lowband's maximum frequency.

The block diagram of the highband structure is shown in **Figure 3-4**.

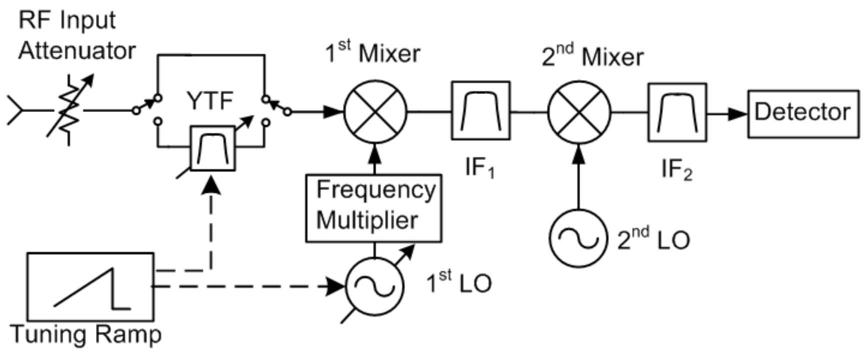

Figure 3-4. Highband Structure Block Diagram.

The main difference between the lowband and the highband structure is the tunable bandpass filter in the RF input path. At present the most common tunable filter type used in signal analyzers is the YIG tuned filter (YTF). YIG is a ferromagnetic material whose resonant frequency is directly proportional to an applied DC magnetic field [[5]]. The YTF exploits the properties of YIG creating a tunable bandpass filter. YTFs can tune from a few GHz to greater than 50 GHz with bandwidths ranging from a few 10's of MHz to a few 100's of MHz.

Because of the relatively narrow bandwidths, the YTF is very effective at attenuating images of the first IF and the very challenging 2,2 spurs. Instead of requiring a very high first IF frequency as in the case of the lowband structure,

the highband structure uses a relatively low first IF frequency. Typical first IF frequencies of highband structures in signal analyzers range from 300 MHz to 800 MHz, which is always below the minimum tune frequency of the highband path.

Figure 3-5 shows the spectrum of a highband structure with a first IF of 600 MHz. In this case, the signal analyzer is tuned to 4 GHz. The YTF also tunes to the center frequency of the signal analyzer, allowing a 4 GHz signal to pass. The YTF frequency response is represented by the dotted line in **Figure 3-5**. A signal at twice the IF (2 x 600 MHz) below the center frequency, would be considered an image if unfiltered. However, the narrow bandwidth of the YTF quite easily filters out this potential image response.

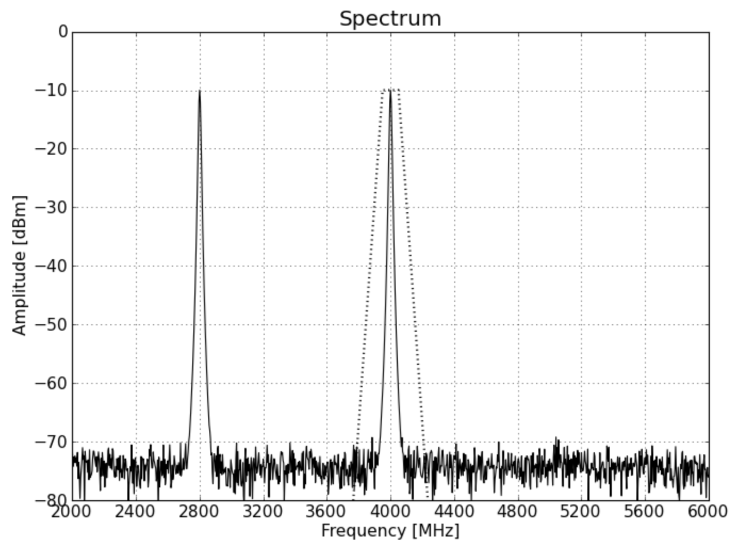

Figure 3-5. Image Suppression in the Highband Structure.

An added benefit of the highband structure is fewer conversion stages. The first IF frequency is low enough that normally only two conversions are required. Most often, the first IF frequency of the highband structure matches the second IF frequency of the lowband structure, allowing sharing of the final conversion stages. Further benefits provided by the YTF are attenuation of both the LO emissions and IF feedthrough.

However, the YTF is not ideal in all measurement circumstances. The YTF uses open loop tuning, meaning that the center frequency of the bandpass response is prone to tuning drift. This center frequency instability translates to relative amplitude and phase shift when the YTF is tuned to a constant frequency. For vector signal analysis, the lack of phase and amplitude stability due to the YTF normally precludes its use for the analysis of digital modulation metrics.

Another problem with the YTF may be the bandwidth. At lower tune frequencies, the YTF bandwidth is usually limited to 100 MHz, sometimes less. Modern communication formats have bandwidths that may be wider than the signal being measured.

To overcome the YTF's degradation of digital modulation analysis performance, the YTF, can be bypassed in some structures. However, once the YTF is bypassed the signal analyzer is now exposed to many of the spur mechanisms that the YTF prevents: images of the first IF, first IF feedthrough and first LO emissions. In a controlled environment, these spurs may not be an issue. Using the YTF path to decipher actual signals from spur signals before applying the bypass can at least reduce the chances that the measured signal is in fact not a spurious response.

Multiple Conversion Block Converter Structure

The block converter is one more structure; however, this structure is not commonly seen in most signal analyzer structures. The block converter's block diagram is shown in **Figure 3-6**.

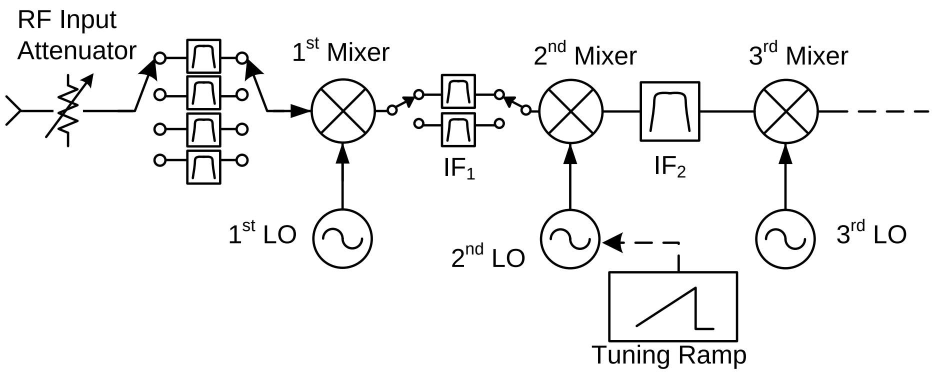

Figure 3-6. Block Diagram of the Block Converter Structure.

The main feature of this structure is that the first LO is fixed in frequency and the first IF and second LO frequency is variable. Often a lowband structure will be constructed and then a block converter will be placed in front of the lowband structure. The block converter, which usually operates at high RF and microwave frequencies, downconverts a segment, or block, of spectrum to the lowband frequency range.

Final IF Frequency Selection

Almost all modern signal analyzers rely on the Analog-to-Digital Converter (ADC) to digitize the final IF signal. Choosing the IF frequency and determining the parameters of the final IF filter are completely dependent on the ADC's sample clock rate. A very detailed analysis of the ADC can be found in reference [[6]]. Only the ADC as it relates to the signal analyzer is discussed here. **Figure 3-7** shows the relevant pieces of the ADC system.

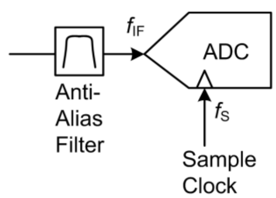

Figure 3-7. Analog-to-Digital Converter Components

The anti-alias filter (AAF) limits the bandwidth of the input signal before it enters the ADC. The filtered signal is sampled for a very short duration once every $1/f_S$ seconds, where f_S is the sample clock frequency. This sampling is depicted in **Figure 3-8**.

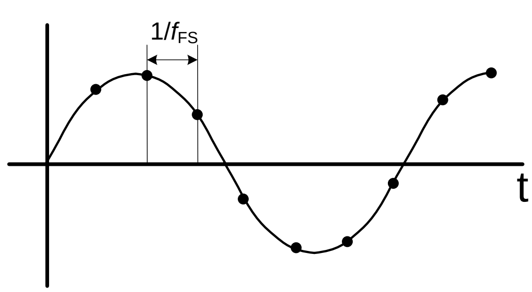

Figure 3-8. Sampling in the Time Domain

The *Nyquist theorem* states that in order to accurately reconstruct a signal using digital sampling, the sample rate must be two times the frequency of the input signal [7]. As long as the bandwidth of the input signal can be constrained, the Nyquist theorem can be violated. This is referred to as bandpass sampling. Bandpass sampling is best explained in the frequency domain. See **Figure 3-9**.

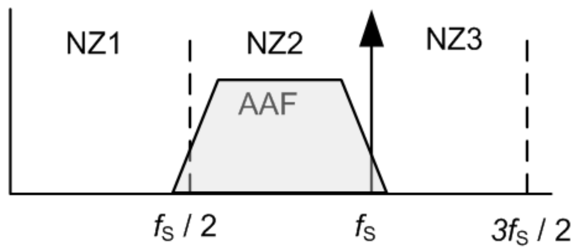

Figure 3-9. Bandpass Sampling in the Frequency Domain

Nyquist zones (NZ) are segments of the frequency spectrum, each with a range that spans one-half the ADC clock sample rate. As long as the input signal is band limited to be within one Nyquist zone, then aliasing will not occur. Aliasing can be explained in terms of the frequency domain representation of sampling. After sampling, the signals in any Nyquist zone appear in Nyquist zone one (NZ1) as shown in **Figure 3-10**.

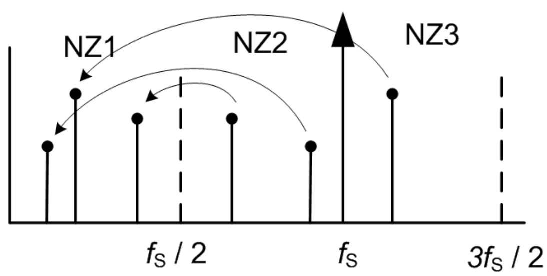

Figure 3-10. Signals Sampling Down to Baseband.

Nyquist zone one is often referred to as baseband. Failure to constrain the input signal frequencies to one Nyquist zone leads to ambiguity in the baseband. For example, if NZ2 is defined as the valid Nyquist zone, a signal present in NZ3 can appears baseband as well as the valid signals in NZ2. This creates ambiguity in trying to decipher the valid signal that appear in NZ1 from the invalid signals. The anti-alias filter as shown in **Figure 3-9** performs the task of constraining the input signal to one Nyquist zone NZ2 in the example shown in **Figure 3-9**.

Fifty percent of the ADC sample rate is the theoretical maximum bandwidth, however ractical constraints on analog AAF filter design limit the maximum bandwidth to approximately 40% of the ADC sample rate.

Variable Bandwidth Final IF Filters

The distinction between vector signal analysis and spectrum analysis was introduced in section 1. These terms, in the context of this document, refer to whether the final IF bandwidth is wide enough to capture the entire modulation bandwidth of the measured signal. For spectrum analysis, the final IF is intentionally narrow banded to enhance dynamic range performance. This subject is explained further in section 6.

Figure 3-11 shows a possible implementation of the narrowband IF filters in the signal analyzer chain.

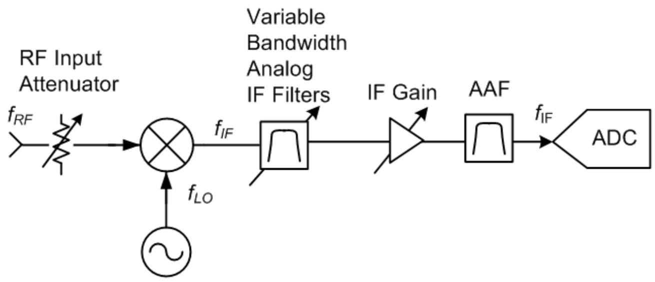

Figure 3-11. Signal Analyzer Block Diagram Highlighting the Analog IF Filters

The variable bandwidth filters can either be a bank of fixed tuned filters that are switched or it can be a single filter with tuning elements. In either case, ideally these filters are placed in the signal chain such that as many devices as possible are prevented from being exposed to the fundamental signals when the distortion components are being measured.

Some signal analyzers do not possess narrow band IF filters. These signal analyzers are only capable of making vector signal analysis measurements.

4. RF Chain Signal Processing

In this section, an analysis of the signal analyzer elements used in controlling the signal amplitude is presented. This is the RF chain signal processing, which does not include the processing of the final IF signal. The next section will discuss the IF signal processing. As part of the RF chain, LO phase noise will also be discussed in this section.

Amplitude Representation in Signal Analyzers

Even though the amplitude units for the signal analyzer can be in terms of power, the signal analyzer is actually a voltage measuring instrument. Post processing converts measured voltage to a variety of amplitude scales, some power based and some voltage based.

Power units by convention, are based on the mean squared value of a sinusoid as shown in the following:

$$P = \frac{V_{RMS}^2}{Z_o} \, Watts,$$

where V_{RMS} is the root mean square voltage of a sinusoid and Z_o is the nominal input impedance of the signal analyzer (typically 50 Ohms).

Figure 4-1 shows how the root mean square value of a sinusoid is related to the peak amplitude by the following:

$$V_{RMS} = {V_{peak}}/{\sqrt{2}}\,.$$

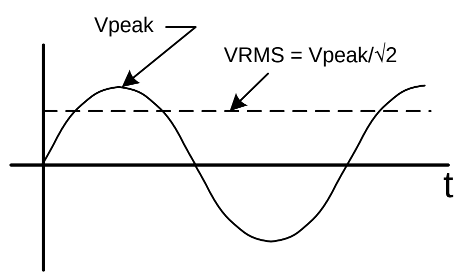

Figure 4-1. Sinusoid RMS Value

Power is expressed in decibels on a log scale, exploiting the compressive nature of logarithms. Further, the most common normalization is the milliwatt, resulting in log units of dBm, dB relative to 1 milliwat. One milliwatt corresponds to 0 dBm and one watt corresponds to +30 dBm For a signal analyzer whose nominal RF port impedance is Z_o, the conversion from V_{RMS} to log power in units of dBm is given in **Equation 4-1**.

$$P\,[dBm] = 10\mathrm{x}log_{10}\left[\frac{V_{RMS}^2}{Zo * 0.001}\right]$$

Equation 4-1

Other amplitude units may be dBuV (dB relative to a microvolt) and dBmV (dB relative to a millivolt) as shown in **Equation 4-2**.

dBuV = dBm + 90 + 10log(Zo); for Zo = 50 Ohms, dBuV = dBm +107
dBmV = dBuV – 60

Equation 4-2

The amplitude units differ between oscilloscopes and signal analyzers. Oscilloscopes tend to have high impedance inputs so as to have minimal effect on the voltage being measured. Signal analyzers are intended to be power measurement devices. Furthermore, the impedance of the signal analyzer is intended to match the source impedance of the device being tested. This is normally 50 Ohms for most RF applications and 75 Ohms for cable television (CATV) devices.

RF/IF Path Amplitude Control Elements

Figure 4-2 shows a very simplified block diagram of the super-heterodyne signal analyzer highlighting the amplitude control elements. In most signal analyzers there are many stages of variable gain IF amplification, but here they are all represented with one element.

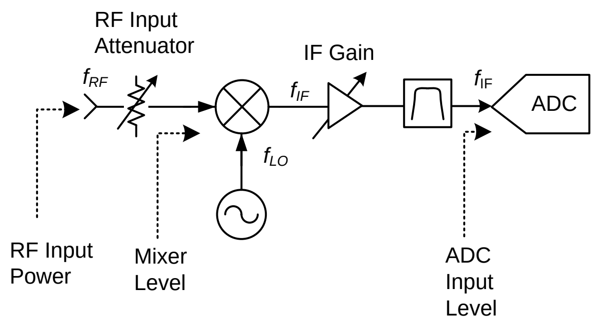

Figure 4-2. RF/IF Signal Path Amplitude Control

RF input power is the total signal power expressed in dBm at the input port of the signal analyzer. If the signal is a CW tone, then this is RMS power of the signal. If the signal contains modulation, then this power is computed by integrating average power over the modulation bandwidth. For multiple signals, this is the average power summation of all signals.

The *RF input attenuator* can be directly controlled by the operator. The amount of attenuation varies between 0 dB and some maximum value (ranging between 50 dB and 70 dB). Resolution varies between 1 dB and 10 dB.

Mixer level is related to the input power of the first mixer by:

Mixer Level = RF Input Power – RF Input Attenuator setting

For instance, if the RF input power is -10 dBm and the RF input attenuator setting is 10 dB, the mixer level is -20 dBm. This term is not to be taken too literally as it does not include the effects of frequency response between the RF input port and the mixer. The true power incident on the first mixer may be slightly off from -20 dBm due to frequency response, but the true power is not the definition of mixer level.

ADC input level is the average power expressed in dBm of a CW tone at the input of the ADC or digitizer. *IF output power level* is also a term used to express this value.

Reference Level and Gain Setting Equations

Reference level refers to the maximum power level of an input CW tone that can be measured without overdriving the IF backend of the signal analyzer. Reference level is also the name given to the vertical axis control of the signal analyzer display. In most signal analyzers, the reference level is the top graticule on the vertical display axis. The user has direct control of this reference level setting. When a CW signal at the RF input has an amplitude equal to the reference level setting, the signal analyzer's internal gains are configured such that the signal level at the ADC is maximized. Letting the input signal power level exceed the reference level value runs the risk of gain compressing component in the RF frontend. Gain compressing the frontend leads to more distortion and degradation of amplitude accuracy. The other risk at letting the signal level exceed the reference level setting is that the ADC can be overdriven. Overdriving the ADC results in distortion so dramatic that no information of the signal can be recovered. Most signal analyzers do not have warning mechanisms for frontends being overdriven. However, an ADC being overdriven often results in a warning such as "IF Overload."

The gain setting parameters that the user can directly control are as follows:

- Reference level
- RF input attenuation
- Maximum mixer level
- IF output level, which is the same as the ADC input level

Maximum mixer level defaults to a value such that the mixer level for any given reference level and input attenuator combination is at least 10 dB to 15 dB below the 1 dB gain compression level of the RF frontend. IF output level is normally set 3 dB to 10 dB below the full scale input level of the ADC. Most often there are default values for the maximum mixer level and IF output level that are optimum. Exposing these parameters allows the experienced user to experiment with measurement specific custom settings.

IF gain as shown in **Figure 4-2** is not directly controllable by the user; however, it is indirectly controlled via the reference level setting. The gain setting equations are constructed such that the displayed amplitude represents the CW signal amplitude at the RF input port. When the RF attenuator is *auto-coupled* to the reference level, the gain equations are as follows:

RF Input Attenuation setting = Reference Level – Maximum Mixer Level

Signal Analyzer Gain = IF Output Level – Reference Level

IF Gain = Signal Analyzer Gain + RF Attenuation Setting

Often the RF attenuator under auto-coupled conditions has a minimum value (5 dB to 10 dB) so that the user does not inadvertently allow the attenuation to become low enough such that there is risk of damaging the front end with too high input power. For manual RF attenuator control, the Signal Analyzer Gain and IF Gain equations still apply.

The IF gain also compensates for frontend frequency response. Recall that the signal analyzer tries to faithfully reflect the power of the input signal. However, the amplitude response of the components in front of the first mixer and the first mixer itself are not flat with frequency. Compensation for frequency response using IF gain will tend to

display the noise floor as an inverse of the frequency response curve. As shown in **Figure 4-3**, roll-off in the frontend frequency response results in roll-up of the noise floor after gain compensation in the IF.

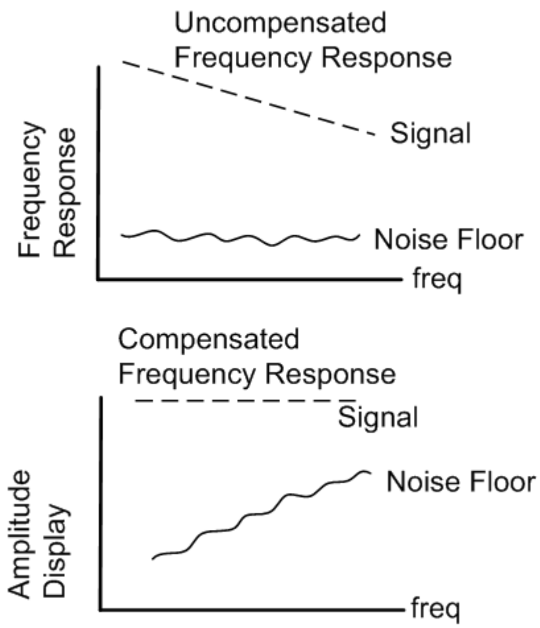

Figure 4-3. Frequency Response Compensation

Mixer Level Effect on Frontend Distortion

Regarding first mixer M,N spurs, lowering the incident power at the mixer changes the distortion and amplitude as a function of the mixing equation's RF signal multiplier, M. Every dB decrease in the mixer level decreases the spur amplitude by the absolute value of M in dB. For example, for a spur with an M multiplier whose absolute value is 2, a 10 dB drop in the mixer level lowers the spur amplitude by 20 dB. Conversely, an increase in mixer level raises the spur amplitude by |M| dB.

Mixer level is varied by changing the power level of the signal being measured or by changing the RF input attenuator setting. Changing the test signal's power is most often not very convenient, which leaves the RF input attenuation as the primary means of improving the signal analyzer's distortion performance.

Mixer Level Effect on Frontend Noise

The noise performance specification for the signal analyzer is the *average noise level*, or sometimes DANL for *displayed average noise level*. DANL is normally specified for an RF input attenuation value of 0 dB. As the RF input attenuation is increased, the signal-to-noise ratio (SNR) decreases. This is shown in **Figure 4-4**.

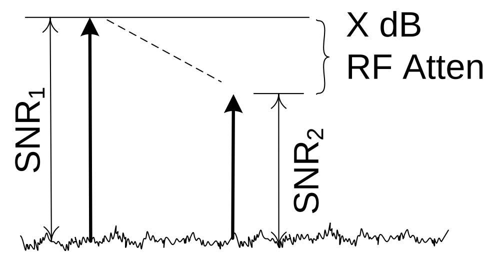

Figure 4-4. Signal-to-Noise Ratio at Two Different RF Input Attenuator Settings.

The signal to the left in **Figure 4-4** is with one RF input attenuator setting and the signal to the right is with X dB more RF input attenuation. At the input of the first mixer, no amplification has taken place, so the noise floor, no matter the RF input attenuator setting, is kTB or -174 dBm/Hz. Thermal noise, or kTB noise, is the noise power generated by a 50 Ohm resistor under the condition that this source resistor is attached to a matched impedance load. This noise power is normalized to a measurement bandwidth, B, of 1 Hz. With entirely passive loss elements in the signal chain, the noise power, assuming matched source/load conditions, can never drop below the kTB level.

Because the noise floor cannot drop below kTB, as the signal is attenuated, the SNR drops dB per dB of attenuation value. However, a constant noise floor is not what is recorded. The signal analyzer is calibrated to try to ensure that the measured signal level at the input port is accurately portrayed. When the RF attenuator is increased, the signal level at the input port does not change, only the signal level at the input of the first mixer. IF gain is increased to compensate for RF attenuation, which increases the displayed noise floor. **Figure 4-5** shows the two signals from **Figure 4-4** as they would appear on the display. The signal powers are now the same, but the noise floor associated with the right-hand signal is X dB higher than the noise associated with the signal on the left. Compensating IF gain accounts for both equal displayed signal levels and noise floor offsets between the two measurement conditions.

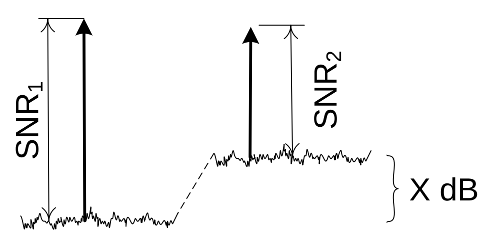

Figure 4-5. Display Adjusted for RF Input Attenuator Setting.

In regards to the analog frontend of the signal analyzer, increasing the RF input attenuation value lowers the amplitudes of the distortion components generated by the signal analyzer. However, increasing the RF input attenuation value also degrades the signal analyzer’s frontend noise performance. A balance must be made between the signal analyzer’s noise and distortion performance using the RF input attenuator setting. Section 6 is entirely dedicated to describing this tradeoff.

ADC Dynamic Range

The ADC used at the very end of the signal chain has a very different optimization for dynamic range. **Figure 4-6** shows the ADC SNR and spurious-free dynamic range (SFDR) versus ADC input level.

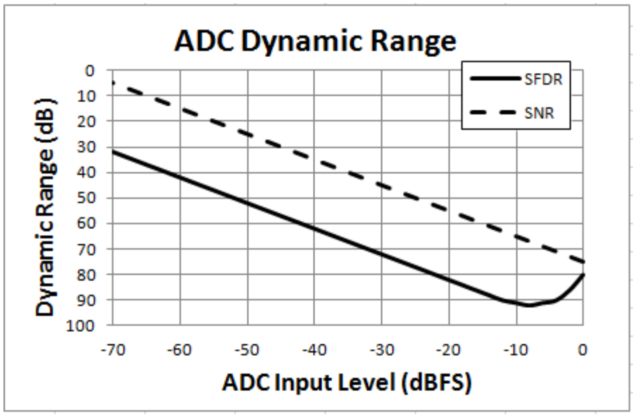

Figure 4-6. ADC SNR and SFDR versus ADC Input Level

For both ADC SNR and SFDR, these values have their highest values near the maximum ADC input power level. Maximum ADC input power level is termed f*ull scale input level*. The x-axis in **Figure 4-6** is dB relative to full scale input level (dBFS).

Consequently, two optimizations must take place in the signal analyzer. The RF input attenuator must be used to set the optimum mixer level for best tradeoff between frontend generated noise and distortion. The IF gain must then be set such that the signal at the end of the signal analyzer's IF chain is close to the ADC's full scale input level. More on this optimization technique in section 6.

Preamplifier

The preamplifier is placed in front of the first mixer in the signal analyzer. Most often, but not always, the preamplifier is placed after the RF input attenuator. **Figure 4-7** shows the block diagram of the signal analyzer with the preamplifier present. Note that the preamplifier path is either selected or bypassed.

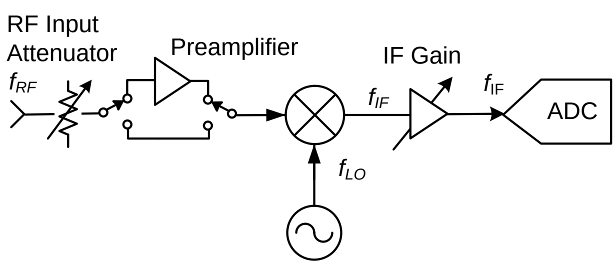

Figure 4-7. Preamplifier in the Signal Analyzer Block Diagram

The purpose of the preamplifier is to lower the noise floor of the signal analyzer. Caution must be exercised when enabling the preamplifier path as it does add gain in front of the first mixer. The maximum input power level must be lowered by the preamplifier gain value. Also, the preamplifier will generate harmonic distortion and intermodulation distortion. In most cases, this distortion degrades more than the improvement in noise floor, resulting in degraded signal analyzer performance. However, if the application relies on improved noise performance alone, enabling the preamplifier is very effective.

Explaining how the preamplifier improves system noise figure adds some insight for low noise measurements as well as gives some foundation in the event that the user may want to add external preamplification for further noise performance improvement.

Noise figure is the logarithmic representation of noise factor, *f* , as shown in **Equation 4-3.**

$$Noise\ Figure\ [dB] = 10 * log_{10}(f)$$

Equation 4-3

Noise factor is defined for a two-port device and is the ratio of input SNR to output SNR [8] as shown in **Equation 4-4**.

$$f = \frac{SNR_{in}}{SNR_{\ out}}$$

Equation 4-4

Using the nomenclature in **Figure 4-8**, the noise factor, shown in **Equation 4-4,** can be reduced to **Equation 4-5**.

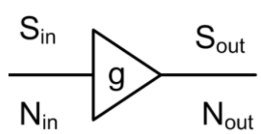

Figure 4-8. Terms Used to Describe Noise Figure.

$$f = \frac{S_{in}/N_{in}}{S_{out}/N_{out}} = \frac{N_{out}}{N_{in}g}$$

where S_{in} and S_{out} are the input and output signal powers respectively. N_{in} and N_{out} are the input and output noise powers respectively.

Equation 4-5

Gain, g, is defined as S_{out}/S_{in}. Input noise, N_{in}, is defined to be kTB. Normally, noise factor for a device is provided by the manufacturer or is a parameter that can be measured. Once noise factor is known, **Equation 4-5** can be solved for output noise.

The signal analyzer can be thought of as a two port device. The input is the RF input port and the output is the recorded amplitude data, which is often the display data. From this point of view, the signal analyzer can be considered to have a gain of 0 dB. The output amplitude is calibrated to equal the input amplitude. With G=1, which is the linear value corresponding to 0 dB, and expressing **Equation 4-5** in logarithmic terms, the signal analyzer noise figure is expressed as shown in **Equation 4-6.**

NF [dB] = *Output Noise Power* [dBm] – kTB

Equation 4-6

Output noise power is the average noise level specification given for the signal analyzer. Input noise is defined to be kTB or -174 dBm/Hz. So, as an example, if the average noise level specification for the signal analyzer is -154 dBm/Hz, the signal analyzer's noise figure equates to 20 dB.

To overcome the signal analyzer's noise figure, the preamplifier is placed in front of the signal analyzer as shown in **Figure 4-9**.

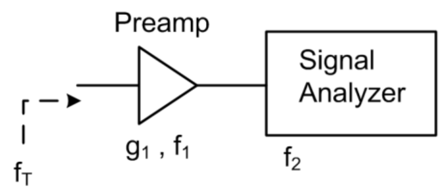

Figure 4-9. Cascaded Noise Figure

The system's total noise factor, f_T, follows the cascade noise figure equation [8]:

$$f_T = f_1 + \frac{f_2 - 1}{g_1}$$

Equation 4-7

If the gain of the preamplifier is sufficiently larger than the noise figure of the signal analyzer, as shown in **Equation 4-7,** the system noise figure approaches the noise figure of the preamplifier alone. Two important parameters define the preamplifier: low noise figure and high gain. In dB terms, to be effective, the gain of the preamplifier should be at least as large as the noise figure of the signal analyzer.

Phase Noise

As mentioned in section 1, phase noise is the frequency domain representation of jitter or phase fluctuation versus time. The signal analyzer contributes phase noise by means of its LO signals. Oscillators used in signal analyzers have an electronic signal to control the frequency of oscillation. Inevitably, noise will be present on the tune control which in turn implants phase modulation on the oscillator signal.

In the frequency domain, phase noise on an LO signal appears as shown in **Figure 4-10**.

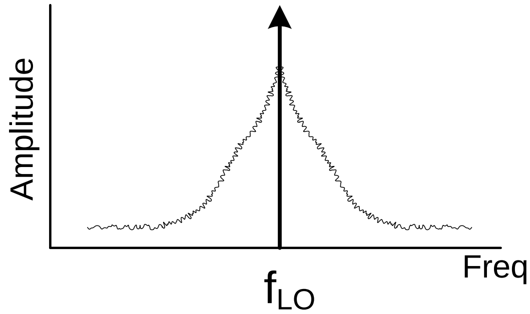

Figure 4-10. Phase Noise on an LO Signal.

A double-sideband pedestal of noise surrounds the LO carrier signal. Most often the phase noise power increases for frequencies close to the carrier.

When specifying phase noise, the phase noise is normally the single-sideband noise pedestal plotted versus frequency offset from the carrier. Single-sideband noise is shown in **Figure 4-11**.

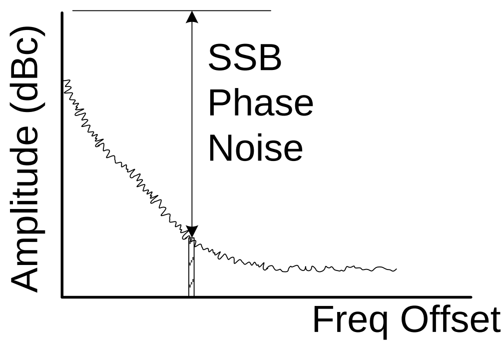

Figure 4-11. Single-sideband Phase Noise versus Offset Frequency.

Because of the variation with offset frequency, phase noise is specified at a particular offset frequency. Phase noise behaves like broadband noise in that phase noise power is a function of the measurement bandwidth. Phase noise is normalized to a 1 Hz measurement bandwidth resulting in units of dBc/Hz.

The frequency shift property of the Fourier transform, as shown in **Equation 2-1**, applies to phase noise. With phase noise superimposed on the LO signal, the LO's phase noise will implant on the IF signal after frequency mixing. If there is phase noise on the RF input signal from the DUT, the LO phase noise will add to the DUT's phase noise during the frequency mixing process.

The signal analyzer's phase noise performance determines whether or not certain measurements can be made. One measurement example is shown in **Figure 4-11**.

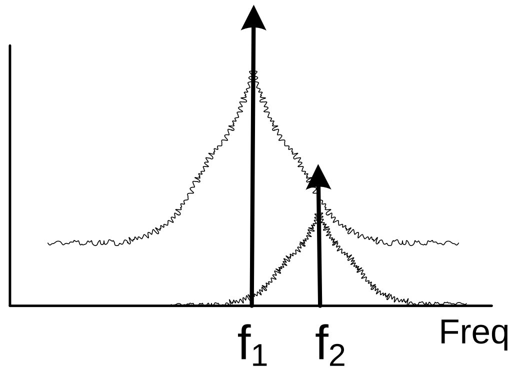

Figure 4-11. Phase Noise Masking.

In this example, a small amplitude signal at frequency f_2, is located close in frequency to a large amplitude signal located at a frequency of f1. The large amplitude signal's phase noise pedestal potentially hides the small signal amplitude. This problem is especially prevalent in the measure of two-tone intermodulation distortion when the two tones are closely spaced in frequency. The resulting intermodulation distortion products, which can be quite small, can easily be buried in the phase noise pedestal surrounding the fundamental tones.

Another measurement where the signal analyzer's phase noise is important is in the measurement of the DUT's phase noise. This situation is shown in **Figure 4-12**.

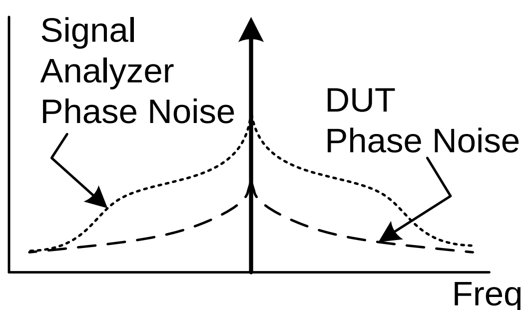

Figure 4-12. Signal Analyzer and DUT Phase Noise.

Often the signal analyzer is used to measure phase noise of the RF input signal from a DUT. If the signal analyzer's phase noise is larger than the DUT's, then this measurement is impossible.

The signal analyzer's phase noise at wide offset frequencies (greater than 1 MHz) is important in the measurement of adjacent channel power leakage ratio (ACLR) of digitally modulated communications signals. Section 6 will show how far offset phase noise can limit ACLR measurement performance.

5. IF Chain Signal Processing

IF Chain Signal Processing refers to the signal conditioning steps after the signal is digitized by the ADC. This includes realization of the *resolution bandwidth filters*, trace averaging, implementation of trace detector types and also FFT processing to convert the signal into frequency spectrum components. For spectrum analysis, only the magnitude versus frequency information is of interest. For vector signal analysis, both magnitude and phase are of interest. This section will concentrate on spectrum analysis IF chain signal processing.

Nearly all modern signal analyzers digitize the final IF signal and use digital signal processing (DSP) routines to perform IF chain signal processing. The DSP is performed in either dedicated integrated circuits (ASICs) or field-programmable gate arrays (FPGAs). Software on the host central processing unit (CPU) often augments the ASIC and FPGA processing. In some cases the entire signal processing is software oriented.

The digital signal processing algorithms, regardless of where the computation takes place, tries to mimic what was once done with analog hardware components. First, a brief review of the analog hardware implementation will be presented. Then, steps for mapping the analog signal processing components to the DSP will be discussed.

Analog IF Signal Processing

Figure 5-1 shows the block diagram of the IF signal processing section of pure analog signal analyzers. The main components are: the *resolution bandwidth filters, logarithmic amplifier, envelope detector,* and the *video bandwidth filter.*

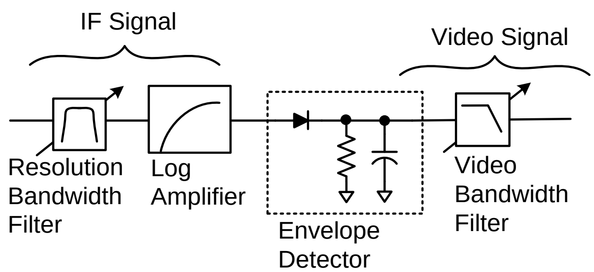

Figure 5-1. Analog IF Block Diagram

Resolution Bandwidth Filter

The *resolution bandwidth (RBW) filter*, in the context of the analog IF signal processing chain, is a variable bandwidth bandpass filter whose center frequency is the frontend's final IF. In most traditional topologies, the bandwidth is adjustable in a 1:3:5 progression, repeating once a decade: for instance 100 Hz, 300 Hz, 500 Hz, 1 kHz, 3 kHz, and so on. Bandwidths range from less than 100 Hz to greater than 1 MHz.

The RBW filter provides the means of distinguishing signals closely spaced in frequency. **Figure 5-2** demonstrates the resolving capability of the RBW filter.

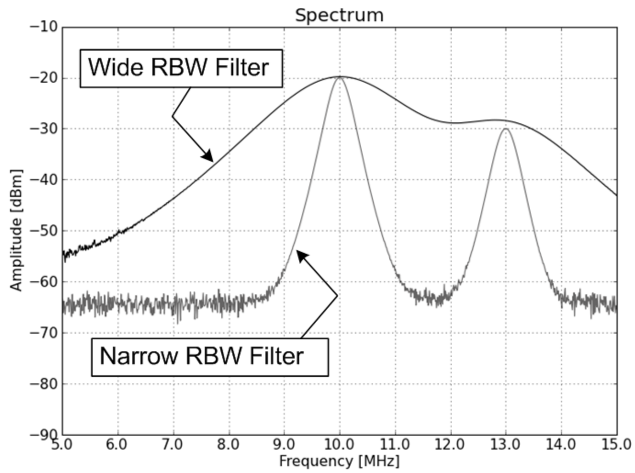

Figure 5-2. Resolving Power versus RBW Setting.

In **Figure 5-2**, with the narrow RBW filter setting, the two signals can clearly be distinguished. However, the wide RBW setting does not provide enough resolving power to separate these two signals.

Another feature of the RBW filter is its influence on the noise floor. In section 4, it was remarked that the thermal noise power is given by the equation N_o = kTB. The 'B' part is the measurement bandwidth. Since, the RBW filter is normally the narrowest filter in the entire signal analyzer, the 'B' can be regarded as being the bandwidth of the RBW filter. On a log scale, the noise power using two different RBW settings is given by:

$$N_{O2} - N_{O1} = 10\log\left(\frac{RBW_2}{RBW_1}\right)$$

where RBW is the 3 dB BW of the RBW filter

Equation 5-1

Equation 5-1 demonstrates that a decade change in the RBW setting results in a 10 dB change in the noise floor. **Figure 5-3** shows the influence that the RBW setting has on the noise floor.

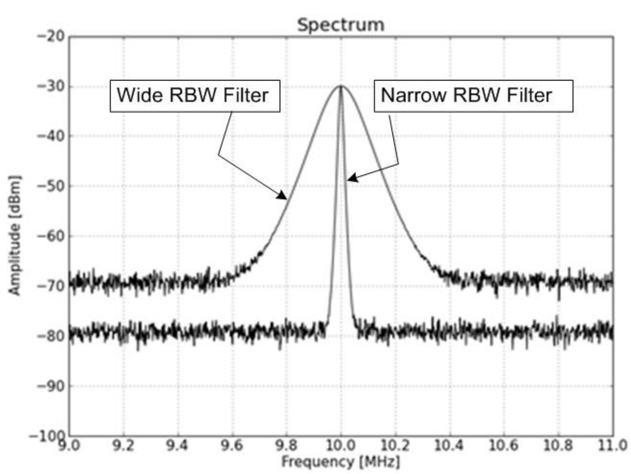

Figure 5-3. Noise Floor versus RBW Setting.

As demonstrated in **Figure 5-3**, the noise floor changes with RBW setting, but the amplitude of a CW signal does not. Therefore, signal-to-noise ratio is a function of RBW setting when measuring CW and narrow bandwidth signals.

Equation 4-6 gives the relationship between the signal analyzer's average noise level specification and the actual noise floor. Average noise level in units of dBm/Hz uses a 1 Hz RBW setting for this measurement. The average noise level generalized to any RBW setting follows **Equation 5-2**.

Signal Analyzer Noise Floor [dBm] = *Average Noise Level specification* [dBm/Hz]+ 10log(RBW) [dB]

Equation 5-2

For instance, if the average noise level specification for the signal analyzer is -150 dBm/Hz, the signal analyzer's noise floor in a 1 kHz RBW setting is:

Noise Floor [dBm] = -150 [dBm/Hz] + 10log(1 kHz)
= -120 dBm

Logarithmic Amplifier

Traditional signal analyzers commonly had a default display vertical axis calibrated in log power, and usually referenced to 1 mW [dBm]. Further, the default range was commonly 90 dB to 100 dB. Using **Equation 4-1**, if one were to try to view the voltages of two signals separated in amplitude by 100 dB, the voltage ratio would be 100,000:1. This voltage ratio would prove impractical for viewing on a single display. To accommodate viewing of large dynamic range on a single display some sort of signal amplitude compression is required. The l*ogarithmic amplifier* fulfils this need.

The *logarithmic amplifier*, or log amp, operates on the IF signal, whose voltage transfer function is logarithmic: Vout =A* log(Vin), where A is a scaling factor. **Figure 5-4** shows an approximation to the log amp transfer function. From **Figure 5-4**, the compression that the log amp provides is clearly evident. A miniscule change on the input voltage translates to many dBs of change in the output. Simultaneously viewing two signals 100 dB or more apart in amplitude is possible with the log amp.

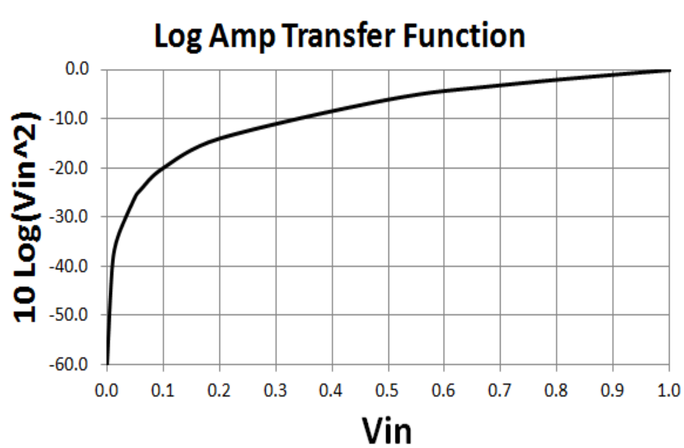

Figure 5-4. Log Amp Transfer Function.

Envelope Detector

Consider a stationary CW signal at the input of the signal analyzer shown in **Figure 2-7**. As the LO is swept, the resulting signal at the IF of the mixer is a sinusoid whose frequency is also swept. A swept frequency signal is sometimes called a chirp signal. As this chirped IF signal passes through the RBW filters, the amplitude of this signal has the shape of the RBW's frequency response. **Figure 5-5** depicts the IF signal as it exits the RBW filter.

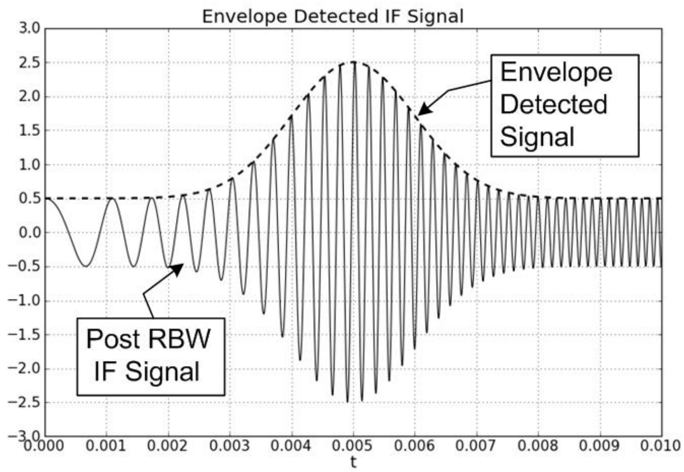

Figure 5-5. Time Domain View of IF Signal After the Resolution Bandwidth Filters.

The *envelope detector* is a simple diode rectifier, whose output traces only the peaks of the IF signal and is represented by the dashed line in **Figure 5-5**. Once detected, the waveform is referred to as the video signal. The detected signal's maximum frequency is no longer that of the IF signal, but rather the modulation riding on top of the IF signal.

Why convert to video? One answer is that is that traditional cathode ray tube (CRT) displays, which were driven directly by the video signal, could only respond to low frequency waveforms. More importantly, however, is that the information in a signal is its modulation characteristics. The central IF signal is merely a carrier for the modulation waveform. Stripping away this carrier still retains the useful information about the signal.

Video Bandwidth Filter

The other IF filter in the analog IF signal processing chain is the *video bandwidth (VBW) filter*. The VBW filter is a lowpass filter with user controlled bandwidth. As the name implies, this filter operates on the video signal, which is the modulation waveform, not the IF waveform. This distinction is important. Filtering the IF signal, which is the function of the resolution bandwidth filter, lowers the overall noise floor. The VBW filter reduces the variance of the noise. Another way of expressing it is that the VBW decreases the noisiness of the displayed noise.

Figure 5-6 demonstrates the effect of the VBW filter on the envelope detected trace. The lighter trace depicts the trace with a relatively wide VBW setting, whereas the dark trace represents the same signal with a narrower VBW setting.

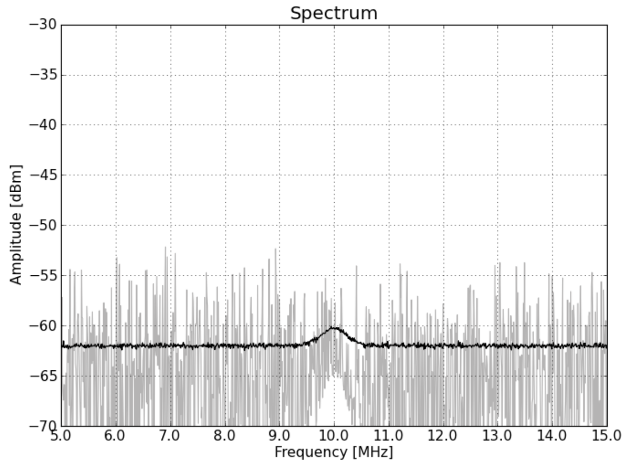

Figure 5-6. Signal with Two Different Video Bandwidth Filter Settings.

Note how the average value of the noise floor itself does not shift with VBW setting, just the fluctuation of noise. With a signal close to the noise floor, as shown in **Figure 5-6**, the signal is undetectable with the wide VBW setting. It is only with the trace smoothing provided by the narrower VBW setting can the near-noise signal be deciphered.

The VBW filter is implemented as a single pole resistor-capacitor (RC) lowpass filter whose corner frequency, or bandwidth, can be selected by the user.

Sweep Speed Considerations

Narrowing either or both the RBW filter setting and/or the VBW filter setting is not without consequences. Both filters have rather dramatic influence on sweep speed. In the time domain, the settling time of a filter grows longer as its bandwidth is decreased. When the IF signal's frequency matches the center frequency of the RBW filter, the RBW responds as if it were hit with an impulse in the time domain. A narrower RBW filter setting takes longer to respond to this impulse. Decreasing sweep time allows the filter to respond and settle to a final amplitude value.

The VBW filter has similar settling constraints in that the sweep speed of the signal analyzer must be slowed down to allow the filter response to settle.

Viewing Modulation

Figure 5-7 shows the frequency domain spectrum of a signal with amplitude modulation (AM)

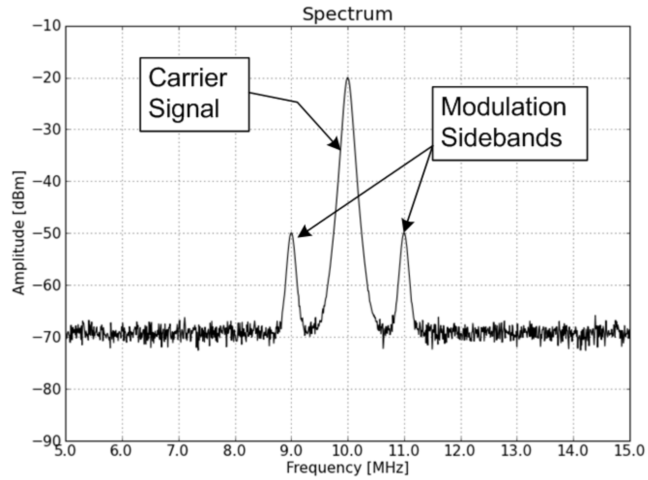

Figure 5-7. Amplitude Modulation in the Frequency Domain.

To resolve the AM sidebands in the frequency domain, the RBW filter setting must be narrower than the AM signal's modulation frequency. However, the signal analyzer can also be used to analyze the modulated signal in a time domain view. This is done with the signal analyzer's frequency span set to zero with the center frequency set to the signal's carrier frequency 10 MHz in the case of **Figure 5-7**. Zero span implies that the LO is not sweeping, but rather fixed at a constant frequency.

To view the modulation, the RBW filter setting must be wide enough to capture both the carrier and the modulation sidebands. In **Figure 5-7**, the AM rate is 1 MHz, so to capture both sidebands, the RBW setting would need to be at least 2 MHz. **Figure 5-8** shows the AM signal when the signal analyzer is set to zero span and the RBW filter setting is much wider than the modulation bandwidth of the test signal.

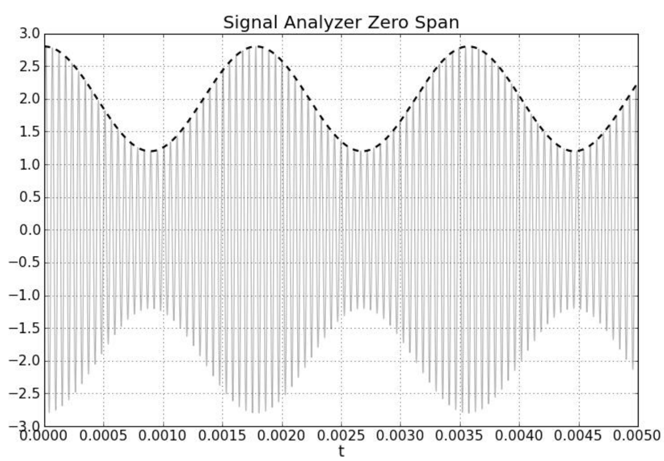

Figure 5-8. Amplitude Modulation in the Time Domain.

Note that the x-axis is time rather than frequency when in zero span.

The high frequency sine wave (light colored trace) in **Figure 5-8** is the carrier. The dashed line represents the modulation signal -- a sinusoid in this case. Envelope detection strips away the carrier, leaving just the modulation.

Detector Modes

At some point between the all analog and all digital IF evolution, the ADC was introduced to the analog IF to digitize the video signal (early ADCs did not have enough bandwidth to digitize the IF signal). Once implemented, a few features related to signal detection were introduced. The digitized video signal is quantized into a finite number of frequency and amplitude points – these are the data points eventually displayed. The data represented on these y-axis points is the subject of the detection modes. Some of these modes are as follows:

- Max peak: the maximum amplitude that occurs between the discrete frequency points, sometimes also referred to as the *peak detector*. With a relatively narrow RBW filter setting, the peak of the signal may occur between the discrete data points. Without the max peak detector, the signal data could potentially be lost.
- Min peak: sometimes also referred to as the *pit detector*. Same function as the peak detector, but records the minimum signal amplitude between the discrete frequency points.
- Sample: the amplitude data at the discrete frequency data point is recorded. This mode is the appropriate one for the measurement of noise.

Figure 5-9 shows the relationship between the continuous data and the detector modes.

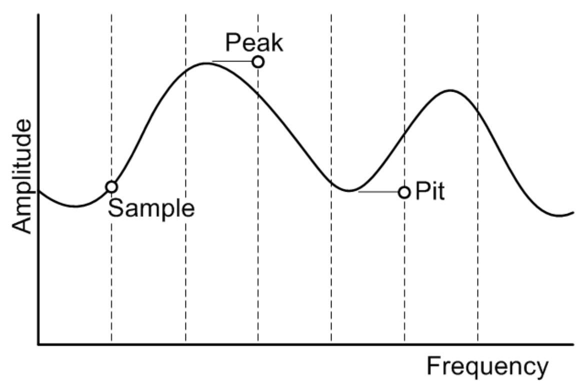

Figure 5-9. Analog IF Detector Modes

Challenges with the All Analog IF

Size and power are two downsides to the analog IF signal processing chain. The circuitry required to realize all the components is not small and certainly not power efficient.

However, the biggest mark against the analog IF is the amplitude accuracy. Among some of amplitude error terms associated with the analog IF are as follows:

- Log fidelity: the transfer function from Vin to Log(Vin) is not perfect. The wider the range of the amplitude scale, the larger this error.
- Bandwidth switching: changing either the RBW or the VBW filter settings results in imperfect relative amplitude accuracy between filter settings.
- Noise floor: the actual 3 dB bandwidth of the RBW filter has inaccuracies compared to the desired bandwidth setting. Since noise power is a function of the bandwidth, errors in the true bandwidth yield errors in the displayed noise floor.

- Temperature stability: holding all the analog circuitry stable over temperature is not an easy problem. No matter how well designed, the analog signal chain almost always yields amplitude error versus operating temperature.

IF Signal Processing with the Digital IF

Analog signal processing has largely been replaced with digital signal processing (DSP) algorithms. DSP techniques can now perform all the necessary signal conditioning of the IF signal. As mentioned earlier, the algorithms can either be implemented in ASIC, FPGA or in software; however, the fundamental DSP algorithms apply no matter which physical device is used. Due to the predictability of DSP algorithms, nearly all of the amplitude uncertainty error terms associated with the analog IF signal chain are simply not present in the digital IF implementation.

The DSP algorithms do vary amongst the different signal analyzer vendors. There are even differences amongst different products within the same vendor. This discussion will concentrate on the DSP routines used by National Instruments and even more specifically the routines associated with the NI-RFSA driver software used by many signal analyzers at National Instruments.

Another distinction is whether or not the end use is for spectrum analysis or vector signal analysis. Spectrum analysis uses magnitude versus frequency data, requiring a conversion from the time domain IF data to the frequency domain data. Vector signal analysis often uses only time domain data; however, the data is complex, containing magnitude and phase. Spectrum analysis is the main focus of this discussion.

Signal Processing Chain: Digital Hardware

In **Figure 5-10**, the block diagram of the signal processing chain portion that resides in the digitizer is shown. For more information, refer to Reference [9]. In the digitizer, the ADC samples the analog IF signal. As discussed in section 3, sampling signals that fall in any Nyquist zone appear as if they are in Nyquist zone one, the baseband. At the end of the processing, the frequency axis will be re-scaled to the appropriate analog IF Nyquist zone, but at this point in the processing chain, the digitized data is in the baseband as shown in **Figure 5-10**.

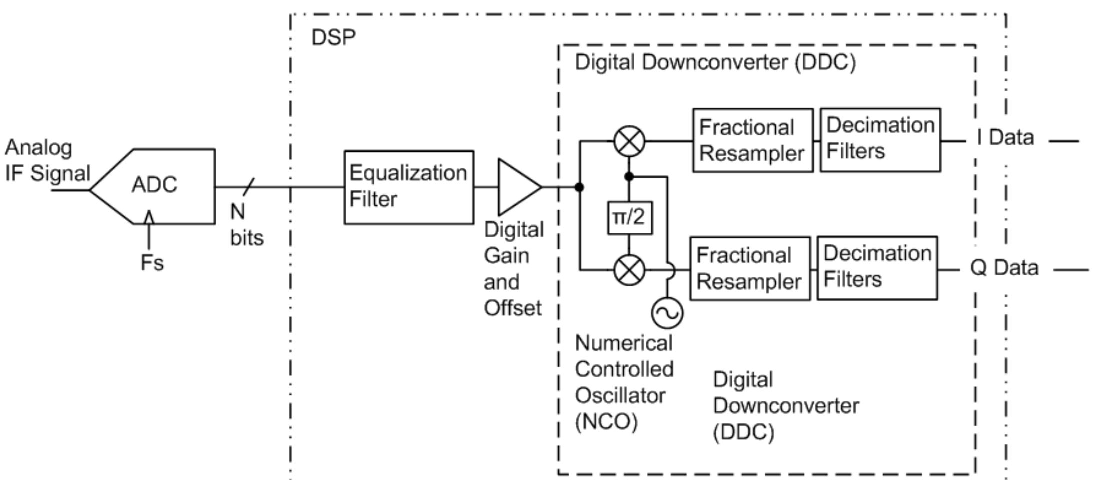

Figure 5-10. Digital Downconverter Block Diagram

The first block in the digital signal processing chain is the *equalization filter*. This is implemented as a finite impulse response (FIR) filter whose filtering coefficients, or taps, can be reconfigured for each separate data acquisition. The purpose of the equalization filter is to remove the magnitude and phase response of the analog frontend components. The equalization is over one Nyquist zone, which implies that this filter removes magnitude and phase frequency

response over the IF bandwidth of the signal analyzer. Factory calibration data and/or signal analyzer self-calibration data corresponding to IF channel magnitude and phase is stored in the signal analyzer's memory. The signal analyzer's driver software uses this measured data to compute the FIR filter coefficients and applies them to the equalization filter.

The frontend will have some gain and offset errors mostly resulting from the finite amplitude step size of front end gain compensation circuitry. These offsets are compensated digitally on the equalized data.

Note that some of the frontend paths may not be able to be equalized due to a large amount of passband amplitude and phase ripple. Paths with SAW filers may fall into this category. In these cases, the number of FIR taps is too large for practical implementation in digital hardware. For magnitude only data, the equalization for these situations may take place entirely in software.

After the data is equalized and corrected for frontend gain offset, the data enters the d*igital downconverter* (DDC). The DDC uses a digital implementation of an IQ demodulator to convert the real data to complex I and Q data streams. A *numerically controlled oscillator* (NCO) as part of the IQ demodulator is used to correct for finite frontend frequency errors. These frequency errors can result from the frontend LOs having limited frequency resolution.

In many modulation formats, the data must be processed with specific data rates. These data rates most likely do not match the sample clock rate of the ADC. Instead of adjusting the ADC's clock rate, which can lead to unwanted spurs and noise degrading sample clock signal, the digitizers in most National Instruments signal analyzers use a *fractional resampler* in the DDC. The fractional resampler effectively changes the data rate of the I and Q data streams.

When the requested bandwidth is less than the entire Nyquist zone, it is beneficial to remove the unwanted spectrum digitally. This procedure is accomplished with the d*ecimation filters*. These filters not only filter out the unwanted
spectrum, but also reduce the data rate. Lower data rate leads to more efficient signal processing in the chain following the decimation filters. The Instantaneous Bandwidth property in the NI-RFSA driver software influences the decimation filter bandwidth.

For measurements requiring complex data, modulation analysis for example, the I and Q data streams are exported to software. For real data analysis, magnitude only spectrum analysis for example, only the I Data stream is exported to software.

Signal Processing Chain: Software

For magnitude only spectrum analysis, the flow diagram of the signal processing that NI currently uses is shown in **Figure 5-11**. The National Instruments implementation of the software signal processing is housed within the NI Spectral Measurement Toolkit

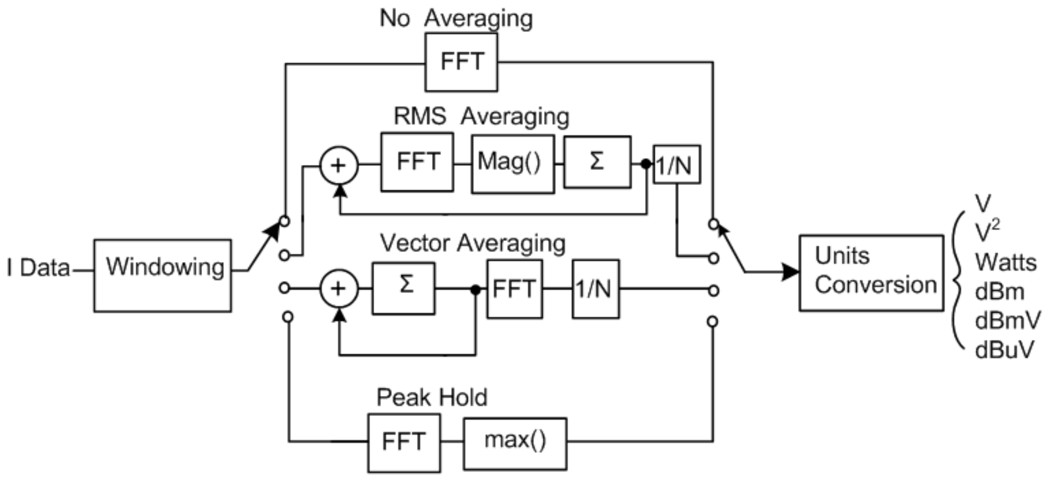

Figure 5-11. Spectrum Mode Signal Processing Block Diagram

Again, for magnitude only processing (spectrum analysis), only the real data from the hardware signal processing chain is required. This is the I Data digital bit stream. The blocks are: *windowing*, followed by *FFT/averaging*, and finally *amplitude unit conversion*. Each of these blocks will be discussed. But first, some background on FFT and spectral leakage is presented before leading into details on the windowing process.

Spectral Leakage in the FFT Process

The d*iscrete Fourier transform* (DFT) and the more computationally efficient *fast-Fourier transform* (FFT) operate on time domain data that has been sampled. The sampling process of converting the time domain signal, *x(t)*, into the sampled signal, *x(n)*, is shown in **Figure 5-12**:

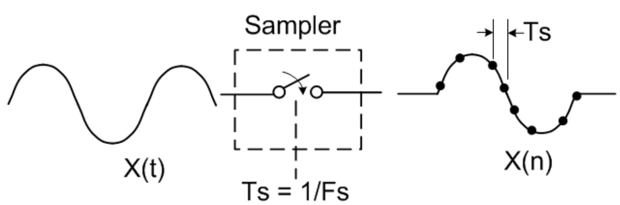

Figure 5-12. Sampling a Sinusoid Prior to FFT Processing.

The ADC acts as a momentary switch that outputs a sample of the incoming continuous time signal. The samples occur spaced at every 1/Fs in time, where Fs is the clock rate of the ADC. The sampling occurs over a finite period of time – this is the *data acquisition* duration. After the FFT processing, the frequency domain data also falls at discrete points. If there are N time domain samples, then there will be N frequency domain points. The frequency domain points are sometimes called *bins*.

The sampled signal appears as a pulsed signal where the samples have zero valued amplitude values outside the acquisition duration. From a signal processing perspective, the signal has rectangular pulse, or a rect(t), function superimposed on it. From **Figure 1-3**, a rectangular pulse in the time domain translates to a sinc function in the frequency domain.

For a sinusoid input signal, if its frequency is an integral multiple of Fs/N, then the frequency domain sinc function superimposed on the CW signal is as shown in **Figure 5-13**.

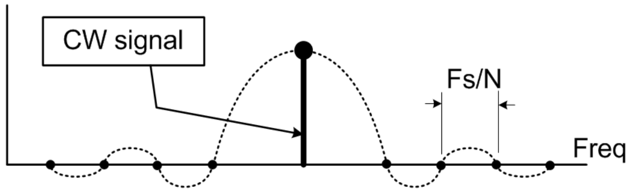

Figure 5-13. Spectrum of CW Signal <u>without</u> Spectral Leakage.

The sinc function is centered on the CW signal. The zero crossings of the sinc envelope fall on FFT frequency bins. Only the spectral line associated with the CW signal results.

Recall that all periodic signals can be deconstructed into a summation of sinusoids. Even if the signal is not periodic, FFT processing requires that the one data acquisition record be considered as periodic – that is, the FFT treats the one data acquisition as if it were replicated infinitely in time. So, the treatment of a single sinusoid can be generalized to any signal that is being processed with the FFT.

If the frequency of the sinusoid is restricted to an integral sub-multiples of the data clock rate (i.e. the signal frequency falls exactly on one of the FFT frequency bins), then the frequency domain is as expected: a single frequency domain spectral line positioned at the sinusoid's frequency. However, this is a very narrow restriction. In nearly all cases, the frequency of the input signal will not line up with one of the frequency bins. Once this occurs the sinc pulse, which is still centered on the frequency of the input signal, no longer has its zero crossings line up with multiples of Fs/N. This is depicted in **Figure 5-14**.

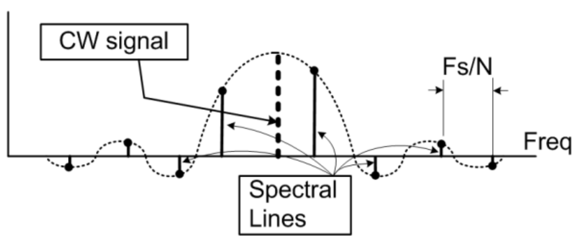

Figure 5-14. Spectrum of CW Signal *<u>with</u>* Spectral Leakage.

Now the zero crossings of the sinc function do not fall on FFT frequency bins. The result is not the single spectral line at the CW signal frequency, but rather a series of spectral lines that fall at the bin frequencies as shown. This phenomenon is known as spectral leakage.

The real damage of *off bin* sampling or spectral leakage is shown in **Figure 5-15**. A CW signal at approximately 50 MHz is shown. One version is centered on a frequency bin (dark line), the other is shifted in frequency so that it is no longer falling on a frequency bin (light line).

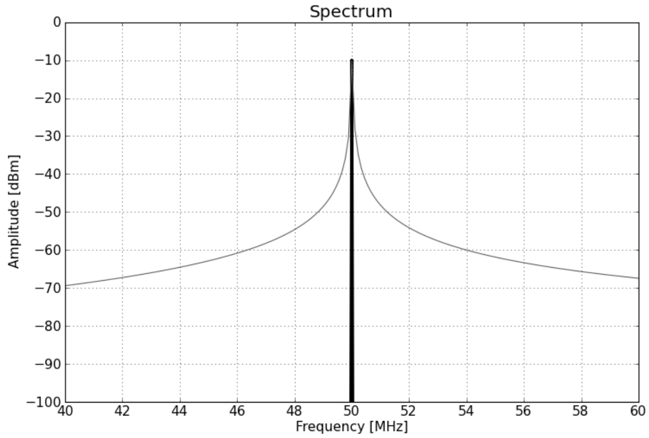

Figure 5-15. Signal On and Off Frequency Bin

The spectral leakage associated with this rectangular window causes a severe restriction on the dynamic range. **Figure 5-16** shows how a nearby signal that is 40 dB down from the central signal is nearly masked by the spectral leakage of the central signal.

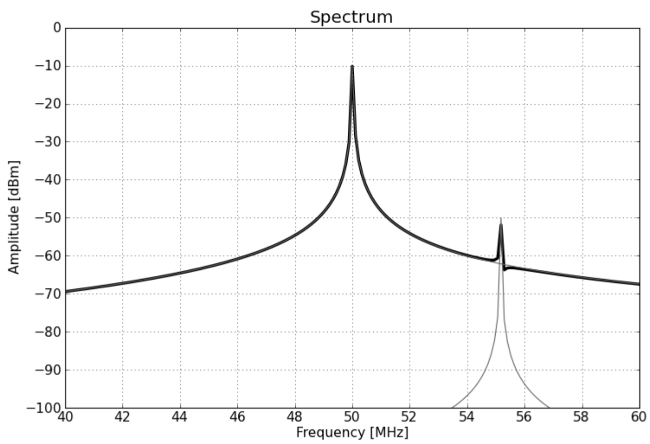

Figure 5-16. Spectral Leakage Masking a Nearby Signal.

The abruptness of the rectangular pulse edges in the time domain is what leads to the excessive leakage in the frequency domain. By multiplying a windowing function with the time domain signal before the FFT processing, the abruptness can at least be minimized [[11]]. **Figure 5-17** shows the time domain signal with (dark trace) and without (light trace) an applied windowing function (Hamming window in this case).

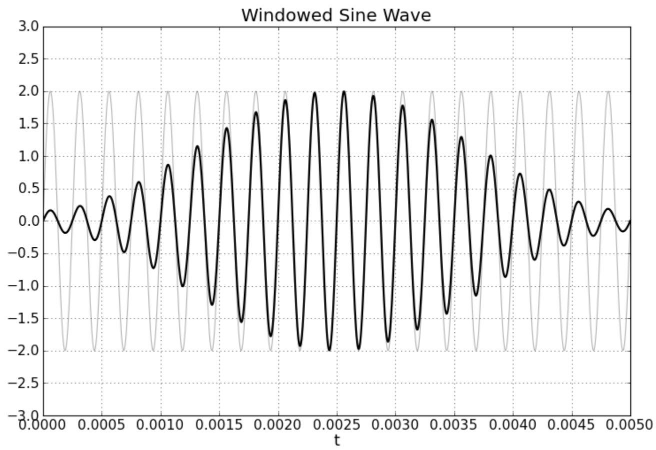

Figure 5-17. Signal Processed with Windowing Function in the Time Domain.

In the frequency domain, the effects of applying various windowing functions are shown in **Figure 5-18**. In this example, a signal has a frequency that does not line up with a frequency bin. The results with four different windowing functions are shown. The difference in the amount of spectral leakage varies greatly between the windowing functions.

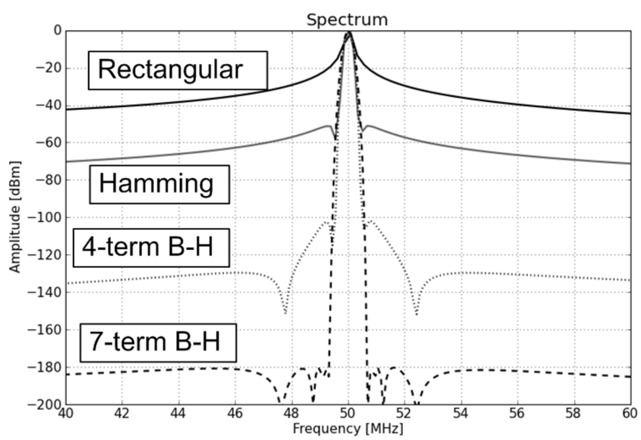

Figure 5-18. Spectral Leakage with Rectangular, Hamming, 4-Term Blackmann-Harris, and 7-Term Blackmann-Harris Windowing Functions.

Resolution Bandwidth Using Windowing Functions

Recall from the discussion of the analog IF signal processing that the RBW filter is used for frequency domain selectivity and SNR enhancement. The digital IF implementation used by National Instruments uses the windowing function to realize the RBW filter. The RBW bandwidth value is proportional to the number of time domain sample points. A higher number of sample points results in a narrower RBW bandwidth.

Figure 5-19 shows one example of the RBW realization using the Hamming windowing function.

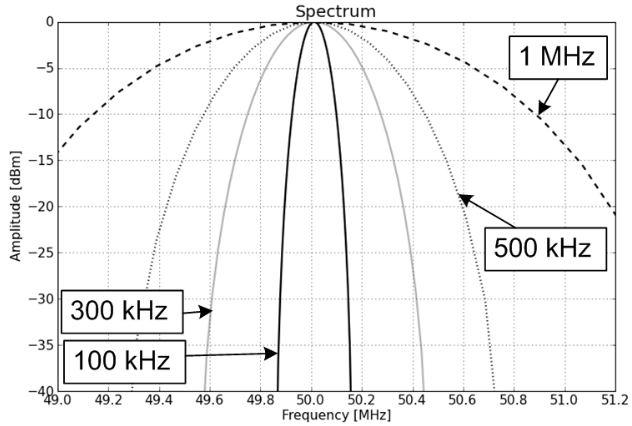

Figure 5-19. Hamming Windowing Function with Various Effective Resolution Bandwidths.

The number of time domain samples is adjusted to synthesize the different 3dB bandwidth settings; the bandwidth is inversely proportional to the number of samples. For example, the 500 kHz bandwidth filter requires twice the number of sample points as the 1 MHz bandwidth filter. Since the ADC has a constant sample rate, the data acquisition time required for the 500 kHz bandwidth is twice as long as that for the 1 MHz bandwidth.

Just as displayed noise floor is a function of RBW bandwidth for the analog only IF, so too is this behavior for digitally implemented RBW filters. Again signal selectivity and lower noise floor are reasons for choosing a narrower RBW filter.

Windowing Function Figures of Merit

There are dozens of windowing functions from which to choose. Some of the criteria for choosing the best windowing type for a particular measurement depend on: bin width, equivalent noise bandwidth scalloping loss, and side lobe attenuation. First, these terms are defined and then a summary table of the more popular windowing functions is given.

Bin Width

Bin width is the frequency resolution of the spectrum after FFT processing. If the number of time domain samples is held constant, then the bin width remains unchanged. **Figure 5-20** shows the spectrum of a CW signal with various windowing functions applied. For each window, the number of time domain samples is held constant and the control setting for the filter is specified using a constant bin width value.

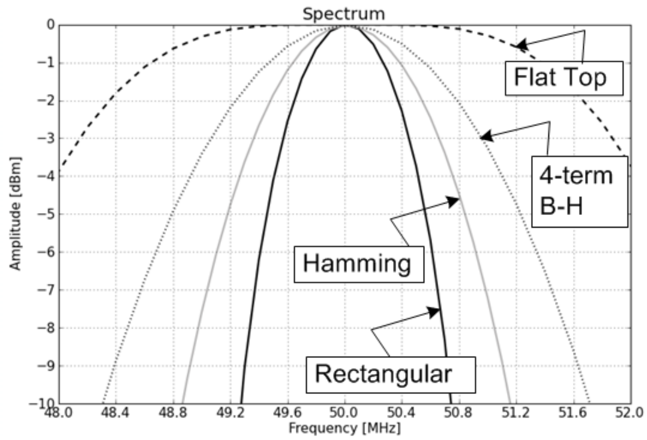

Figure 5-20. Windowing Functions Specified by Bin Width

Note that the 3 dB bandwidth associated with each windowing function varies. Between the widest and narrowest bandwidths of the windowing function in **Figure 5-20**, the ratio is 4.2 to 1. Holding the number of time domain samples constant for each windowing function, which implies a constant frequency domain bin width, does not result in constant filter bandwidth.

In **Figure 5-21**, the number of samples for each window is adjusted such that they all have the same 3 dB bandwidth.

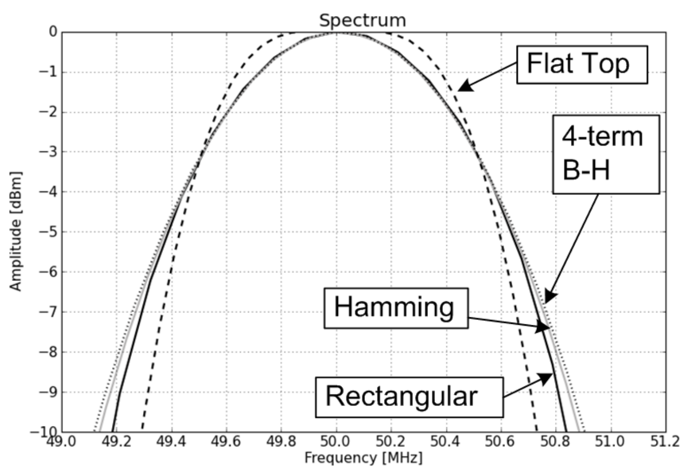

Figure 5-21. Windowing Functions Specified by 3 dB Bandwidth

Now the bandwidths are consistent, but the number of sample points for each windowing function is not constant. The 4.2:1 ratio of widest BW to narrowest BW in **Figure 5-20** becomes the ratio of sample points required to satisfy the consistent 3 dB BW requirement of **Figure 5-21**. Data acquisition time is directly proportional to the number of required sample points. The ADC is still sampling at the same rate for each windowing function, so in order to affect the number of sample points the acquisition time must be adjusted. So there is a tradeoff between selectivity of the windowing function and measurement time. Windows with lower selectivity require more sample points and hence longer acquisition time. For instance the Flat Top filter requires 4.2 times longer acquisition time over the rectangular filter to realize the same 3 dB bandwidth.

Equivalent Noise Bandwidth

Equivalent noise bandwidth (ENBW) is also referred to as e*ffective noise bandwidth.* Integrating the area under the curve of the magnitude squared of the windowing function in the frequency domain gives the *power gain*[11]. Multiplying the power gain by the noise power yields the filter's *noise power.* A perfect brick-wall filter (infinitely sloped skirts) has a noise power equal to its 3 dB bandwidth. However, the actual filter will have a slightly higher noise power value as there is more area under the curve than the brick-wall filter.

ENBW is defined as the area under the curve of the filter's frequency response divided by the frequency range used to compute the power gain measurement as shown in **Equation 5-1**.

$$ENBW = \frac{\int_{f_1}^{f_2} |X(f)|^2 \, df}{f_2 - f_1}$$

Equation 5-1

Figure 5-22 shows that if a perfect brick-wall filter is constructed such that is has the same noise power as the actual filter, the ENBW is the bandwidth of that brick-wall filter. ENBW is always higher than the 3 dB BW of the filter response.

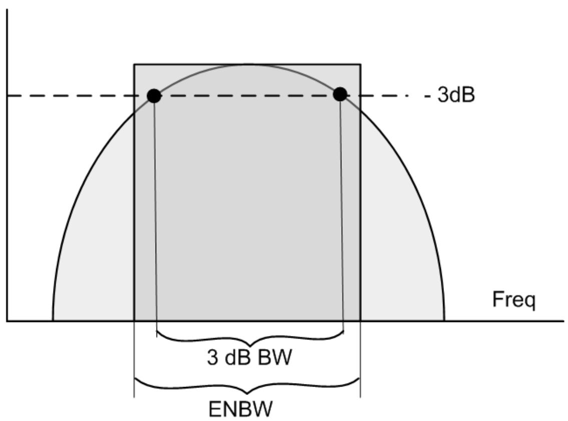

Figure 5-22. Equivalent Noise Bandwidth

What ENBW means to the user is that the noise floor of the signal analyzer is slightly affected by the choice of windowing function. The average noise specifications for the signal analyzer are usually given for a specific windowing function. If a different windowing function is used, then the noise floor can have an offset.

The windowing function bandwidth can be chosen in terms of ENBW instead of 3 dB BW. The number of samples is adjusted to ensure that the noise power of each windowing function is constant. **Figure 5-23** shows the filter response for the Hamming window when specified for 3 dB BW and ENBW.

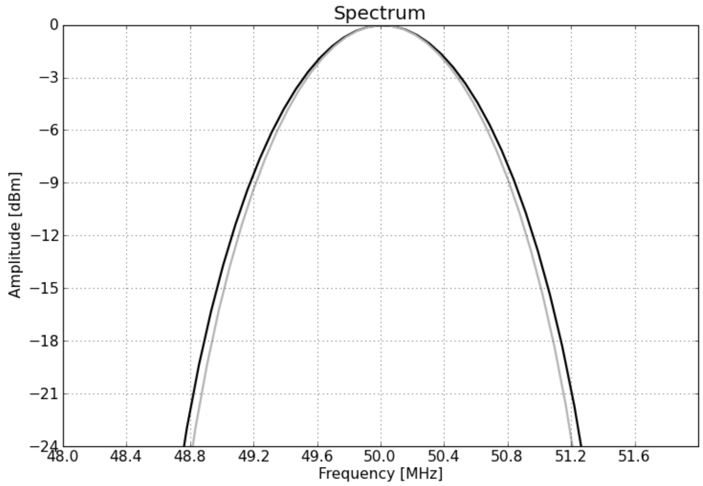

Figure 5-23. Hamming Window. Dark Trace = 3 dB BW Setting, Light Trace = ENBW Setting.

The ENBW filter is narrower than the 3 dB BW filter to compensate for the noise power gain.

Scalloping Loss

Signal frequencies off bin lead to amplitude errors as well as the generation of spectral leakage. The amplitude error is termed *scalloping loss* [[7]] and is depicted in **Figure 5-24**.

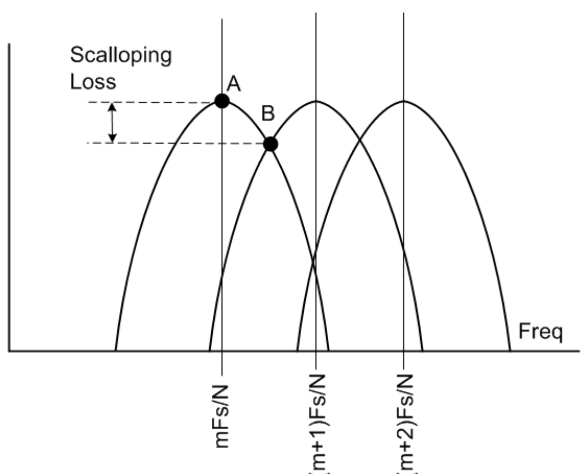

Figure 5-24. Scalloping Loss

If the signal frequency is positioned at point ‘A’ in **Figure 5-24**, the signal is on bin and there is no amplitude error. As the signal frequency drifts away from being on bin the amplitude will reach a minimum when the frequency is half way between bins (point ‘B’). The amplitude difference between maximum and minimum is the scalloping loss. The amount of scalloping loss varies amongst the various windowing functions and must be considered when selecting windowing type.

Sidelobe Attenuation

Figure 5-18 shows a dramatic difference in the stopband attenuation of a few different widowing functions. By zero padding the FFT (adding trailing zero amplitude points to the time domain sample waveform), more stopband detail of the windowing function is exposed. This is shown in **Figure 5-25** for the Hamming windowing function.

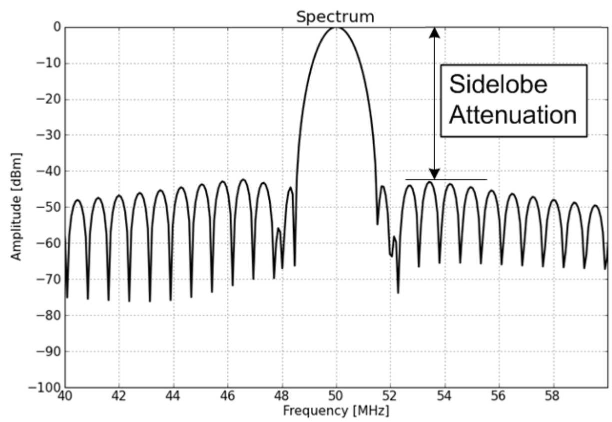

Figure 5-25. Hamming Widowing Function Sidelobe Attenuation

The amplitude difference from peak to one of the local maxima in the stopband is the sidelobe attenuation value. Again, the choice of windowing function dictates the amount of sidelobe attenuation.

Windowing Function Comparison

Table 5-1 shows some figures of merit for a few windowing functions.

3 dB BW is defined as the number of bins required to realize a RBW filter's 3 dB bandwidth. The number of bins is scaled to a 1 Hz BW. Use these as a relative ranking on the number of sample points required. For example, the Hanning window uses 1.451/0.886 = 1.63 times more sample points as compared to the rectangular window.

ENBW is also defined in terms of bins for a 1 Hz bandwidth.

Noise Power Gain is the dB difference in the noise power when specifying the filter in terms of its 3 dB bandwidth versus it ENBW, that is this is the noise power error when using the 3 dB BW specification.

Sidelobe Attenuation is the attenuation beginning 10 bins away from the center. Sometimes the closest sidelobe is used for this benchmark; however, this may be overly pessimistic for most signal analysis use cases.

Windowing Function	3 dB BW (bins)	ENBW (bins)	Noise Power Gain (dB)	Scalloping Loss (dB)	Sidelobe Attenuation (dB)
Rectangular	0.886	1.0	0.519	-3.905	30.3
Hanning	1.451	1.511	0.174	-1.396	71.1
Hamming	1.309	1.370	0.195	-1.727	47.3
Blackmann	1.656	1.739	0.214	-1.077	78.9
4-term Blackmann-Harris	1.913	2.019	0.234	-0.809	99.2
7-term Blackmann-Harris	2.501	2.652	0.253	-0.475	173.7
Flat Top	3.751	3.797	0.052	-0.011	91.9

Table 5-1 Windowing Function Comparison

Lowest acquisition time	Rectangular
Highest acquisition time	Flat Top
Least Noise Power Error	Flat Top
Lowest Scalloping Loss	Flat Top
Highest stopband attenuation	7-term Blackmann-Harris
Best overall compromise of speed, scalloping loss, stopband attenuation and noise power error	4-term Blackmann-Harris

Trace Averaging

The next block in the signal processing chain of **Figure 5-10** is *trace averaging*. In the previous discussion, the windowing in the signal processing chain emulates the analog IF's RBW filters. Trace averaging emulates the VBW filter functionality as well as some of the detector functions.

Vector Averaging

This is also called c*oherent averaging* [[7]]. In vector averaging, the time domain data from a specified number of data acquisitions is averaged. The averaged date is then processed with the FFT. In this manner the noise floor drops by 20log(N), where N is the number a trace averages specified. The intended application is for the measurement of time triggered events. If the signal is triggered at the same amplitude and phase from trace to trace, the averaged signal amplitude remains constant, but the noise floor decreases. This improves the signal to noise ratio of the measurement.

Another implementation is to operate on the spectral components by first applying the FFT operation. The real part and the imaginary part of each spectral component are then averaged separately.

RMS Averaging

RMS averaging is also known as *incoherent averaging* as it does not use an input signal to trigger the data acquisitions. The data acquisitions are free running which allows for a more general use case of measurement signals. RMS averaging operates on the power spectrum of the data and consequently the phase information is lost in this averaging process. As indicated in **Figure 5-10**, the time domain data is first processed with the FFT. For each spectral component, X_i, the magnitude is computed using:

$$|X_i| = \sqrt{(X_i)(X_i^*)}$$

where X_i^* is the conjugate of X_i

Equation 5-2

RMS averaging closely resembles the analog IF's VBW filter functionality. The variance of the noise floor is reduced, but its average value is not affected.

Peak-Hold Averaging

Peak-hold averaging is not actually an averaging function. The user specifies a number of trace averages – these are independent data acquisitions. The peak-hold detector retains the maximum value at each frequency bin of all the data acquisitions. This mode is useful detecting transient events.

Averaging Mode Comparisons

Figure 5-26 show the comparison of the RMS, Vector, Peak-hold averaging functions. For each one, 100 trace averages are specified.

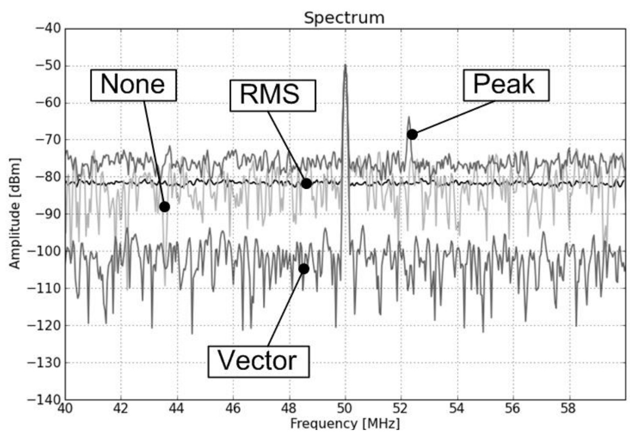

Figure 5-26. Averaging and Detection Modes

Video Bandwidth Filter Emulation

From the analog IF section of this document, the VBW filter uses a single pole lowpass structure. This filter operates on the video signal to reduce the variance of the random noise riding on the signal. The FFT-based signal analyzer does not implement a lowpass function, but rather maps the number of trace averages, using the RMS averaging mode to emulate the analog VBW's noise variance reduction capability [[12]].

Some measurement standards specify that the measurement be made with a particular VBW setting. Described here is the algorithm that maps the number of trace averages in a FFT-bases analyzer to the traditional analog IF RBW and VBW settings.

When the ratio of RBW to VBW is high, the approximation for the number of trace averages is proportional to the ratio of the equivalent noise bandwidths (ENBW) of the RBW to the VBW in the analog IF based signal analyzer. The analysis for this restriction is based on the demonstration that when detecting power using a nonlinear transformation, such as square-law detection, the spectral transformation from IF to baseband results in the narrowband power spectrum at IF convolved with itself, except at DC [12].

$$S'(f) = c\int_{-\infty}^{\infty} S(g)S(f-g)dg$$

where *S(f)* is the normalized power spectrum of RBW filter and *c* is a constant.

Equation 5-3

The ENBW of the RBW filter is then shown in **Equation 5-4.**

$$ENBW_R = \frac{\int_0^{\infty} S'(f)df}{S'(0)}$$

Equation 5-4

For a synchronously-tuned RBW implementation, the magnitude response of the N-order filter is shown in **Equation 5-5**.

$$|H(f)| = \left[\frac{1}{[(2(f-f_c)/a]^{2}+1}\right]^{(N/2)}$$

Equation 5-5

$$a = B/\sqrt{2^{1/N}-1}$$

Equation 5-6

where,
B = 3 dB BW of the filter
f_c = the RBW filter center frequency

When white noise is passed through the filter, the normalized power spectrum is shown in **Equation 5-7**.

$$S(f) = \left[\frac{1}{[(2(f-f_c)/a]^{2}+1}\right]^{(N^2/4)}$$

Equation 5-7

For a 4th order RBW filter, $ENBW_R$ = 0.842 x B = 0.842 x RBW setting, the ENBW of a single pole lowpass filter is (π/2)VBW or 1.571VBW. So for VBW << RBW, the number of trace averages, N, is shown in **Equation 5-7**.

$$N \sim ENBW_R / ENBW_V = 0.536 \times (RBW/VBW)$$

Equation 5-8

However if the VBW bandwidth is not much less than RBW bandwidth, then the VBW has less affect on trace averaging in which case the number of trace averages should decline to 1. **Equation 5-9** approximates the transition from VBW << RBW to VBW > RBW:

$$\text{Num Trace Averages} = \left[\left(k \frac{\text{RBW}}{\text{VBW}} \right)^p \right]^{1/p} \qquad (5\text{-}9)$$

where k =0.536 and p = 1.275.

Equation 5-9

Equation 5-8 is plotted in **Figure 5-27**.

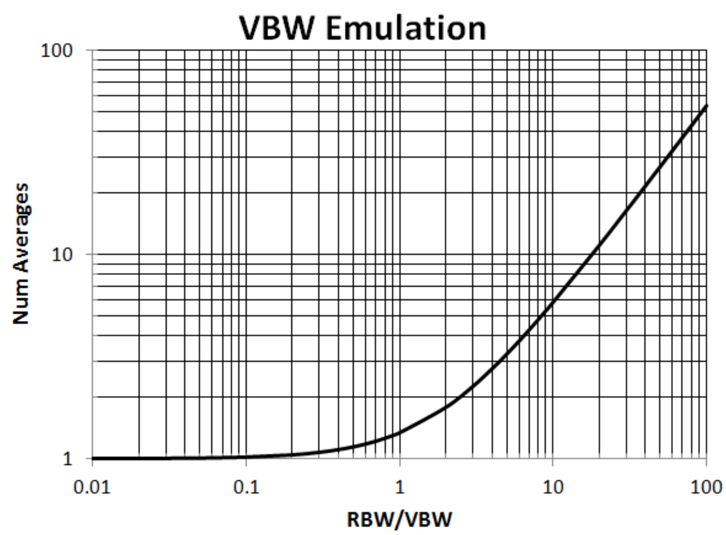

Figure 5-27. Number Trace Averages vs. RBW/VBW Ratio

Unit Conversion

The final block in the signal processing chain of **Figure 5-11** is the *unit conversion* stage. This allows the amplitude of the spectrum data to be recorded in several different formats including: Volts, $(\text{Volts})^2$, Watts (computed using user controlled impedance setting), dBm **Equation 4-1**, and dBuV or dBmV **Equation 4-2**.

Log Mode Trace Averaging

In section 4 of this document, the thermal noise (sometimes called kTB noise) is present at the input of the signal analyzer. The signal analyzer then adds internally generated noise. For both cases, the noise over the measurement range is considered to be additive white Gaussian noise (AWGN). White refers to the noise having a flat frequency response, and Gaussian refers to the probability distribution function (pdf) of the noise voltage having normal distribution as shown in **Figure 5-28**.

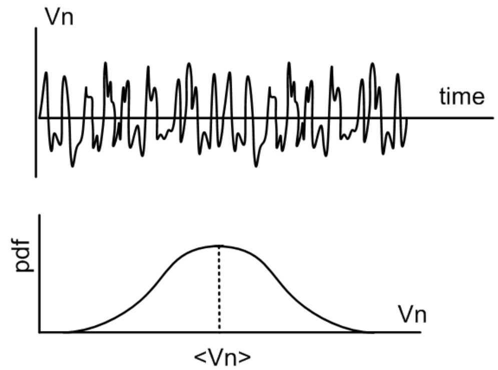

Figure 5-28. Thermal Noise Probability Distribution Function

With RMS averaging, the variance, or spread in the pdf curve, reduces without the average value of the noise voltage changing. This is depicted in **Figure 5-29**.

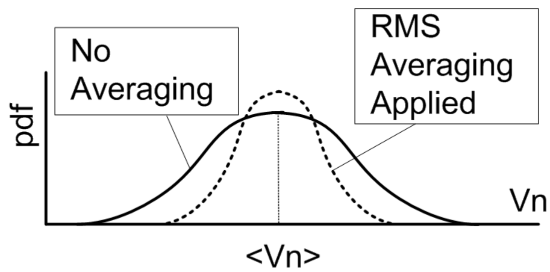

Figure 5-29. Variance Reduction with RMS Averaging

If one uses the ENBW setting for specifying the RBW bandwidth, then the displayed power after RMS averaging is the true representation of signals with any kind of statistics. If the 3 dB BW setting is used for specifying the RBW bandwidth, then the noise power gain corresponding to the selected windowing function must be added. For instance, if a 7-term Blackmann-Harris windowing function is used, then add 0.253 dB to the measured RMS averaged noise power.

The traditional spectrum analyzer with the analog IF **Figure 5-1** does not use RMS averaging. The signal is first sent through the log amplifier and then the envelope detector before the VBW filter effectively averages the signal. The act of envelope detection on a log scaled noise voltage alters the statistics of the noise probability distribution function as depicted in **Figure 5-30**.

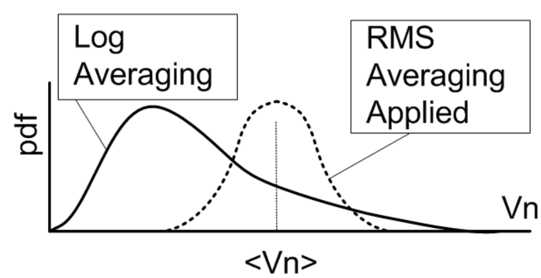

Figure 5-30. PDF with RMS and Log Averaging

The log averaging lowers the mean of the noise power by 2.51 dB [13]. The amplitude of a CW signal is not altered by either RMS averaging or log averaging, so for the measurement of CW signals, the SNR is a true 2.51 dB improvement using log averaging over RMS averaging.

In traditional spectrum analyzers with synchronously tuned RBW filters, the noise power gain is approximately 0.5 dB [14]. Thus the displayed average noise level, which effectively uses the 3 dB BW setting for the RBW shows an approximate -2.0 dB noise power from the true RMS noise level. **Figure 5-31** shows the SNR of a CW signal using RMS and Log averaging.

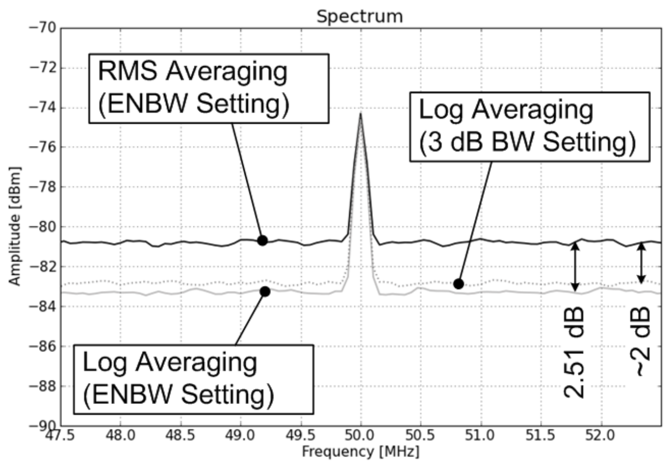

Figure 5-31. SNR with RMS Averaging and Log Averaging Modes.

The -2.51 dB noise power bias is a predictable offset for AWGN, so for the measurement of CW signals, log averaging is very appropriate. However, for modulated signals where the statistics of the signal is not predictable, then log averaging does not accurately report the signal's power level. In the case of modulated signals, the VBW filter setting of the traditional spectrum analyzer needs to be set much wider than the RBW setting. This, however, offers no variance reduction in the signal. With the advent of digitizing the video signal, trace averaging using digital techniques made variance reduction possible without altering the mean value of the signals power. However, by using RMS averaging mode, the statistics of both the broadband thermal noise and the digitally modulated signal are unaltered, making RMS averaging the appropriate averaging mode for this class of signals.

Care should be used when comparing average noise level performance of signal analyzers. Even though modern signal analyzers with all digital IFs do not need to average on the envelope of the log data, some do this anyway using DSP means in order to add the apparent 2.51 dB improvement in noise level specifications. Unless the specification calls out the average noise level with RMS notation, one can assume that the noise is measured using log averaging. For a true comparison of the signal analyzer's noise figure or a true comparison of the SNR when measuring modulated signals, the noise specification should be converted to RMS.

6. Dynamic Range

Dynamic range is arguably one of the more import figures of merit for the signal analyzer. In section 2, the mechanisms that cause distortion in the signal analyzer were discussed. In section 4, the amplitude control of the signal analyzer was introduced. In this section, the discussion turns to how to optimize the signal analyzer to minimize both its distortion and noise. In the next section, the accuracy of the measurement will be presented.

We first need to define what is meant by dynamic range and what some of the more common measurements where dynamic range is important are. The dynamic range versus mixer level chart allows the user to determine how to optimize the signal analyzer settings for best dynamic range performance. Construction of the dynamic range vs. mixer level chart is first described, and then the use this chart for various measurements is discussed.

In section 1, the idea of signal analysis versus vector signal analysis was introduced. Recall that signal analysis uses narrowband analog IF filtering and vector signal analysis requires relatively wide analog IF bandwidth (bandwidth must be greater than the modulation bandwidth of the signal being measured). The signal analyzer optimization is different depending on whether the measurement is narrowband versus wideband. Optimizing the signal analyzer for each type of measurement is discussed next.

Dynamic Range Definitions

One of the most overloaded terms in RF is the term *dynamic range* itself. Without qualification dynamic range can be used to describe spurious-free dynamic range, harmonic distortion, signal-to-noise ratio, intermodulation distortion, gain compression, and more. But, in general, dynamic range usually refers to the ability to simultaneously display large and small amplitude signals in a single measurement.

How the signal analyzer corrupts the dynamic range of a general distortion measurement is shown in **Figure 6-1**.

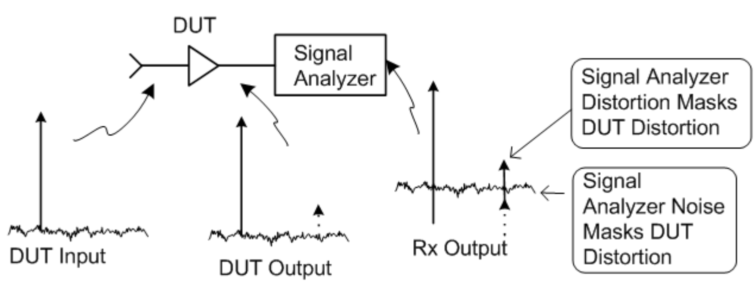

Figure 6-1. Signal Analyzer's Contribution to Distortion and Noise.

Consider a pure, distortion-free signal input to a DUT. The DUT will produce some distortion -- harmonic distortion in this case. Next the signal analyzer attempts to measure the DUT's distortion. The signal analyzer will create distortion of its own and these distortion products fall at the same frequencies as those of the DUT. Potentially, the amplitude of the signal analyzer's distortion could be larger than the DUT's. Additionally, noise generated by the signal analyzer could mask the DUT's distortion. With either signal analyzer generated distortion and/or noise, the effect is the same: the signal analyzer does not have enough dynamic range to measure the DUT's distortion.

For the context of this discussion, only the dynamic range of the signal analyzer is considered. The signal analyzer's dynamic range is what is trying to be maximized. The general definition of dynamic range for the signal analyzer is:

> *Signal analyzer dynamic range is the ratio of the input signal amplitude to the signal analyzer generated distortion or signal analyzer generated noise.*

This definition does not state anything about the amplitude display range, or about the maximum signal to the noise floor. This definition also does not state that the measurement of the signal and distortion be made with the same signal analyzer settings.

Maximizing the signal analyzer's dynamic range involves changing gain settings to trade off signal analyzer generated distortion and noise. When the distortion amplitude equals the noise floor, *maximum dynamic range* is achieved. **Figure 6-2** shows the conditions for maximum dynamic range.

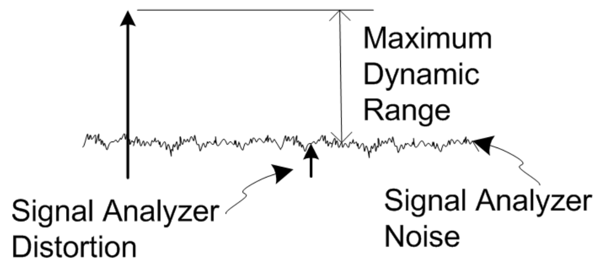

Figure 6-2. Maximum Dynamic Range

Now we turn to a listing of several distortion measurement examples.

Gain Compression

A linear device is normally a passive device which when driven with an input signal, its output signal has an amplitude that is a scaled version of the input signal's amplitude. This scaling is constant with applied power. A nonlinear device will have a scaling factor that can have different values depending on the applied signal level.

Many nonlinear devices, such as amplifiers and mixers, behave nearly linear when lower amplitude signals are applied. That is, these devices have near constant gain versus input power. **Figure 6-3** shows the graph output power vs. input power, both on a log scale, for an amplifier.

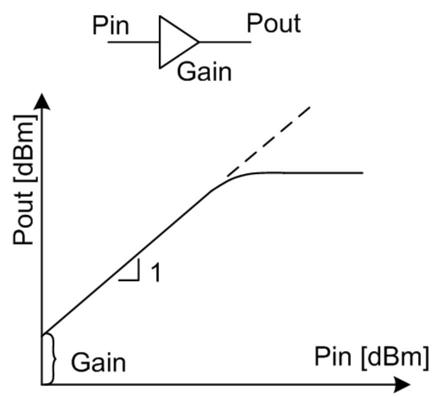

Figure 6-3. Pout versus Pin for Two Port Device

For lower input power levels, the slope of the curve is one, meaning that for every one dB increase in the power level of the signal at the amplifier's input port, the output power level increases by approximately one dB. The offset in this curve is the small-signal gain of the device.

Plotting the gain versus power at either the input or output, at some power level, the gain will begin to fall, or compress. The common figure of merit is the one dB gain compression level, which is the power level at which the small-signal gain drops by one dB. This level is termed P1dBm. The plot of gain versus power level is shown in **Figure 6-4**.

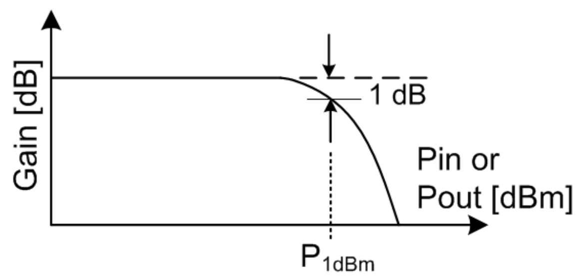

Figure 6-4. Gain Compression in a Two Port Device

Amplifiers normally specify P1dBm at the output and mixers normally specify P1dBm at the input.

The signal analyzer, largely comprised of amplifiers and mixers, also has a gain compression specification. Most often the specification conditions are with 0 dB RF input attenuation, which is a way of describing the mixer level. However, for default gain settings, one cannot directly measure the gain compression of the signal analyzer. Recall from section 4 the signal analyzer's internal gain is such that for an input signal whose power level is positioned at the reference level setting, the signal level at the final IF is close to the ADC's full scale input level. Measuring a signal whose amplitude is above the reference level runs the risk of over-driving the ADC. The maximum mixer level constrains such that the signal level at the first mixer is at least 10 dB to 15 dB below the signal analyzer's gain compression level. Measuring a signal with amplitudes nearing the signal analyzer's gain compression level results in an IF overload error, indicating that the ADC full scale input level is being exceeded.

For signal analysis, the danger exists when large amplitude signals fall outside the frequency span as depicted in **Figure 6-7**.

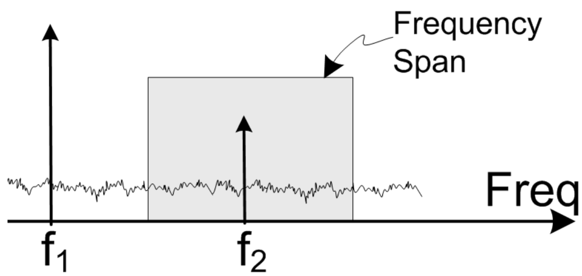

Figure 6-7. Gain Compression in a Signal Analyzer.

A large amplitude signal at frequency f_1 outside of the frequency span can drive the signal analyzer's frontend into gain compression without over-driving the ADC. Narrowband filters in the signal analyzer's IF chain can attenuate out of band signals. When driven into gain compression by out of band signals, the lower level signals that fall inside the viewing range, such as the signal at frequency f_2 in **Figure 6-7**, have an unpredictable amplitude error.

Another overlooked issue associated with signal analyzer gain compression is the measurement of digitally modulated signals. To a rough approximation, the displayed power spectral density of the modulated signal, whose modulation bandwidth equals BW_m, is $10Log(BW_m)$ below a CW signal whose power is the same at the average power of the digitally modulated signal. For instance, for a digitally modulated signal whose modulation bandwidth is 1 MHz, the signal level appears 60 dB below a CW tone whose power equals the average power of the modulated signal. However, this is only partly true since the displayed amplitude is also a function of RBW setting. **Figure 6-8** shows the relationship between the displayed digitally modulated signal and a CW signal of equal power levels. **Figure 6-8** is with a RBW setting of 1 Hz.

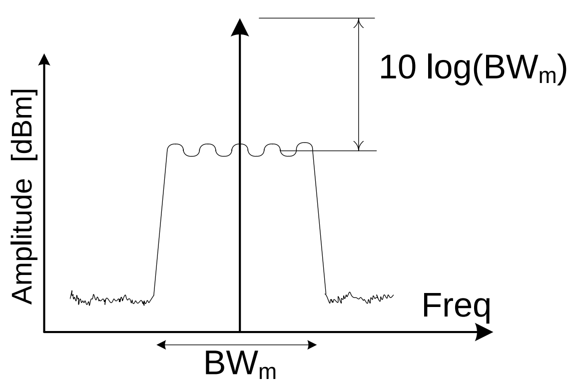

Figure 6-8. Power of a Modulated Signal

The signal analyzer gain compresses according to the total power at the first mixer. By measuring the power in band, a feature within the NI Spectral Measurement Toolkit, the total power of the modulated signal can be assessed. It is quite easy to have the modulated signal appear below the signal analyzer's reference level without registering an IF overload error, yet its total power exceeds the gain compression specification.

Many signal analyzers specify the gain compression using the *two-tone desensitization technique*. The two-tone desensitization measurement technique is a requirement for the signal analyzer as a signal tone technique as depicted for the amplifier example in **Figure 6-3** would cause an IF overload condition. **Figure 6-9** shows the spectrum for the two-tone desensitization technique.

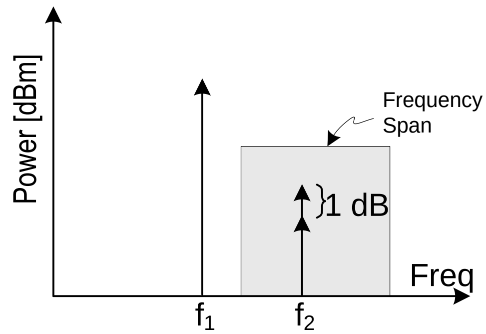

Figure 6-9. Two-tone Desensitization Measurement for Gain Compression

First a low level signal at frequency f_2 is placed within the frequency span of the signal analyzer. An out-of-band signal at frequency f_1, where the separation frequency is wider than the analog IF bandwidth of the signal analyzer, is configured such that its amplitude is well below the gain compression level. As the out-of-band signal amplitude is increased, the amplitude of the in-band signal at f_2 will fall by 1 dB at a certain power level. The power level of the out-of-band signal is considered the gain compression level of the signal analyzer. Because of attenuation due to IF filtering within the signal analyzer, the large amplitude tone will not overload the ADC.

Harmonic Distortion

Consider the signal analyzer when operating far below its gain compression level. Under this condition, the signal analyzer can be treated as a *weakly nonlinear device*. The output voltage as a function of input voltage for a weakly nonlinear device can be described with a power series expansion as shown in **Equation 6-1.**

$$V_{out}(t) = a_0 + a_1 V_{in} + a_2 V_{in}^2 + a_3 V_{in}^3 + \ldots$$

Equation 6-1

For the signal analyzer, V_{out} is used to form the power recorded for display and V_{in} is the voltage corresponding to the signal at the RF input port.

Consider the a sinusoidal input signal given by $V_{in} = Acos(\omega t)$. Inserting V_{in} into **Equation 6-1** and expanding terms results in **Equation 6-2**.

$$V_{out}(t) = a_1 Acos(\omega t) + a_2 A^2 \cos(2\omega t) + a_3 A^3 \cos(3\omega t) + \ldots$$

Equation 6-2

Plotting **Equation 6-2** on a spectrum with $\omega = 2\pi f_{o,}$ the result is the fundamental tone at frequency f_o, second harmonic at frequency $2f_{o,}$ third harmonic at frequency $3f_o$, and so on.

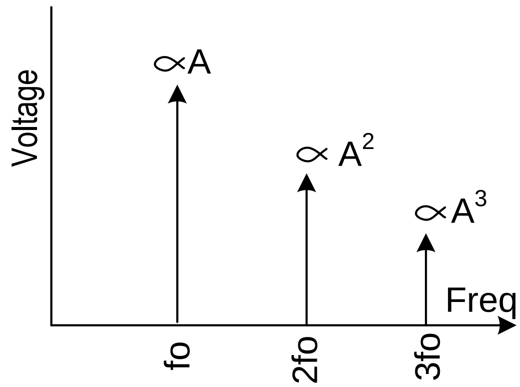

Figure 6-10. Spectrum Showing Harmonic Distortion.

The order of the distortion product stems from its amplitude dependence. The second harmonic has second order distortion because its amplitude is proportional to the square of the fundamental signal amplitude. Likewise, the third harmonic has third order distortion due to its amplitude being proportional to the cube of the fundamental signal's amplitude.

When plotted on the log scale, for every 1 dB change in the fundamental signal amplitude, the second order distortion product amplitude drops by 2 dB, the third order distortion product amplitude drops by 3 dB, and so on as shown in **Figure 6-11**.

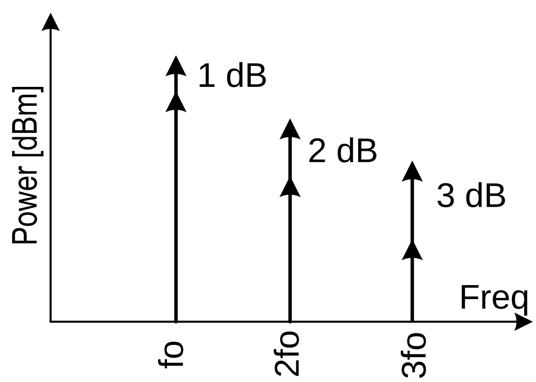

Figure 6-11. Harmonic Distortion on a Log Scale.

Second Harmonic Intercept

Second harmonic intercept (SHI) is another figure of merit for the signal analyzer. SHI is a single value with power units in dBm. With this single value, the user can determine the second harmonic distortion product amplitude for any given fundamental power level. **Figure 6-12** shows graphically the concept of SHI.

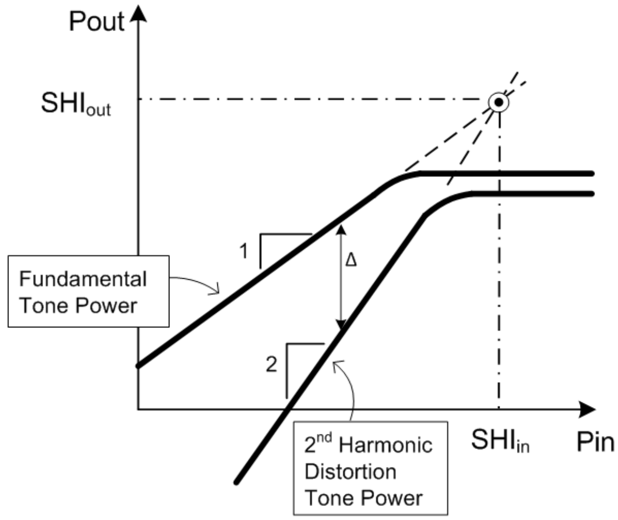

Figure 6-12. Second Harmonic Intercept

Both the fundamental tone power and 2nd harmonic distortion product power are plotted on the Pout vs. Pin graph. Below gain compression, the fundamental tone curve has a slope of one and the 2nd harmonic tone curve has a slope of two. If these two curves are extrapolated, they cross. This crossover point corresponds to the SHI value. Input SHI is the power at the input at the device where the lines cross and output SHI is the device's output power level where these lines cross. The difference between input and output SHI is the gain of the device. Of course, in reality these lines do not cross as the device goes into gain compression before crossover can occur.
Even though fictitious, SHI is still a very useful parameter for both comparing one signal analyzer's performance against another and for predicting distortion when operating in the linear region. By using simple geometry on the graph shown in **Figure 6-12**, an equation for SHI can be deduced as shown in the following:

$$SHI = Psig + \Delta \quad (dBm)$$

where *Psig* is the power of the fundamental tone and Δ = *Psig* – *Pharm*
Pharm is the power of the 2nd harmonic signal.

Equation 6-3

For output SHI, *Psig* is measured at the output of the device and for input SHI, *Psig* is measured at the input of the device. SHIout (dBm) =is equal to SHIin (dBm) plus Gain (dB). The signal analyzer has a gain of zero dB, and so SHIout is equal to SHIin. Furthermore, SHI for the signal analyzer is normally specified with 0 dB RF input attenuation.

For the signal analyzer, SHI is a given value. One can use this value to determine the harmonic level. For example, suppose the signal analyzer SHI specification is +40 dBm and the power at the input port with 0 dB RF input attenuation is -10 dBm. Plugging these values into **Equation 6-3** results in the following:

SHI = *Psig* + Δ
Δ = SHI – Psig
Δ = 40 – (-10) = 50 dB
Δ = *Psig* – *Pharm*
Pharm = *Psig* – Δ

Pharm = -10 – 50 = -60 dBm

This states that the harmonic level is -50 dBc, or -60 dBm.

Intermodulation Distortion

Intermodulation distortion (IMD) occurs when two or more signals are present at the input of a nonlinear device. Frequency mixing between these fundamental components will create distortion tones. The most common (IMD) is with two input tones of equal amplitude. Back to **Equation 6-1**, which shows the V_{out} vs. V_{in} relationship of a weakly nonlinear device, suppose V_{in} is the summation of two sinusoids shown in the following:

$$V_{in} = Acos(\omega_1 t) + Bcos(\omega_2 t)$$

the resulting output voltage is the following:

$$V_{out} = a_1\{Acos(\omega_1 t) + Bcos(\omega_2 t)\} + a_2\{A^2Bcos[(2\omega_1 - \omega_2)t]\} + a_2\{AB^2cos[(2\omega_2 - \omega_1)t]\} + \text{many other terms}$$

Equation 6-4

With $\omega_1 = 2\pi f_1$ and $\omega_2 = 2\pi f_2$, the two-tone IMD products closest in frequency to the fundamental tones are shown in **Figure 6-13**.

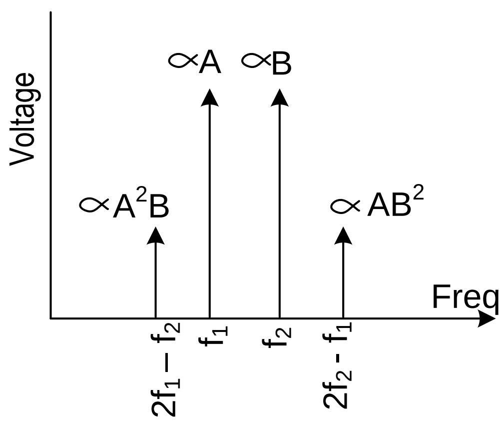

Figure 6-13. Third Order Intermodulation Distortion

These IMD products are particularly troublesome due to their proximity to the fundamental tones. If f_1 and f_2 are separated in frequency by Δf, the distortion products are also separated by Δf from their closest fundamental tones. Filtering these distortion products is nearly impossible when the fundaments tones are at high RF and the separation frequency is small.

If the fundamental tone powers both have amplitude of ‘A’, then it is easy to see that the distortion product amplitudes are proportional to A^3 making these third order. To reiterate, the nature of third order distortion is, as shown in **Figure 6-14** when the fundamental tone powers both change by 1 dB, that the IMD power levels change by 3 dB.

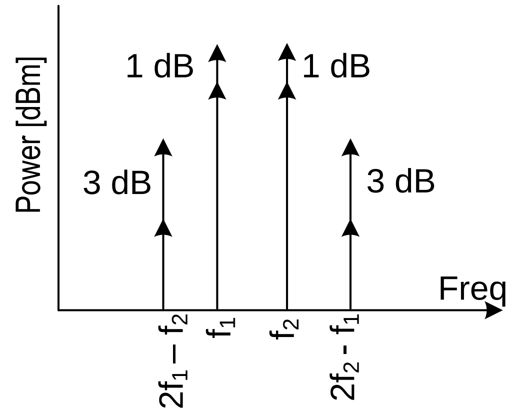

Figure 6-14. Third Order Intermodulation Distortion on a Log Scale

As with second harmonic distortion, the two-tone IMD also uses the concept of the intercept point as a figure of merit. The intercept point for two-tone intermodulation distortion goes by either third order intercept (TOI) or IP3. **Figure 6-15** shows the Pout vs. Pin relationship for third order IMD.

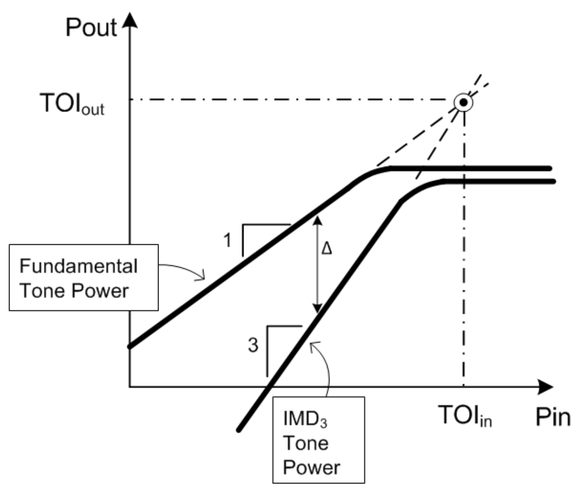

Figure 6-15. Third Order Intercept Point

For third order distortion, the IMD curve has a slope of three when operating in the linear region, which is below the gain compression level. It is assumed that the fundamental tone powers are both being adjusted to create the curve with the slope of one. Extrapolating these two curves beyond gain compression shows a similar crossover as with the second harmonic distortion graph of **Figure 6-11**. Input TOI or IP3 is the input power at the crossover and output TOI or IP3 is the output power at the crossover.

Amplifiers tend to specify output TOI or IP3 and mixers normally specify input TOI or IP3. As with SHI, TOIout (dBm) is equal to TOIin (dBm) plus Gain (dB). Again, signal analyzers have a gain of 0 dB, which makes TOIout equal to TOIin for the signal analyzer.

With some simple geometry on the fundamental and IMD curves in **Figure 6-15**, TOI equates to **Equation 6-5.**

$$TOI = Psig + \frac{\Delta}{2} \ (dBm)$$

where *Psig* is the power of each fundamental tone and Δ = *Psig* – P_{IDM}
P_{IMD} is the power of each IMD product.

Equation 6-5

For the signal analyzer, TOI is usually specified with the RF input attenuator set to 0 dB. To calculate the IMD power level, **Equation 6-5** can be rearranged into **Equation 6-6**.

$$P_{IMD} = 3P_{sig} - 2TOI$$

Equation 6-6

As an example, suppose the TOI specification for the signal analyzer is +20 dBm and the power of each fundamental tone is -10 dBm. The IMD power level is then as follows:

$$\begin{aligned} P_{IMD} &= 3\text{x}(-10\ dBm) - 2\text{x}(20\ dBm) \\ &= -50\ dBm \end{aligned}$$

Optimizing the Signal Analyzer's Dynamic Range Performance

In section 4, we presented a qualitative discussion on setting the signal analyzer's RF input attenuation and reference level settings in regards to the signal analyzer's noise and distortion. Now a more detailed analysis on how to optimize the signal analyzer's setting to maximize its dynamic range performance is presented. We will first analyze the optimization process for the measurement of CW signals for both narrowband and wideband IF modes. This will then help lead into the more challenging optimization for the measurement of digitally modulated signals.

Dynamic Range versus Mixer Level Chart

The *dynamic range vs. mixer level chart* is an often overlooked piece of information found in some signal analyzer data sheets. However, this tool's importance cannot be stressed enough for setting the signal analyzer's gain parameters for highest dynamic range. **Figure 6-16** shows an example dynamic range vs. mixer level chart taken from a signal analyzer's data sheet [[15]].

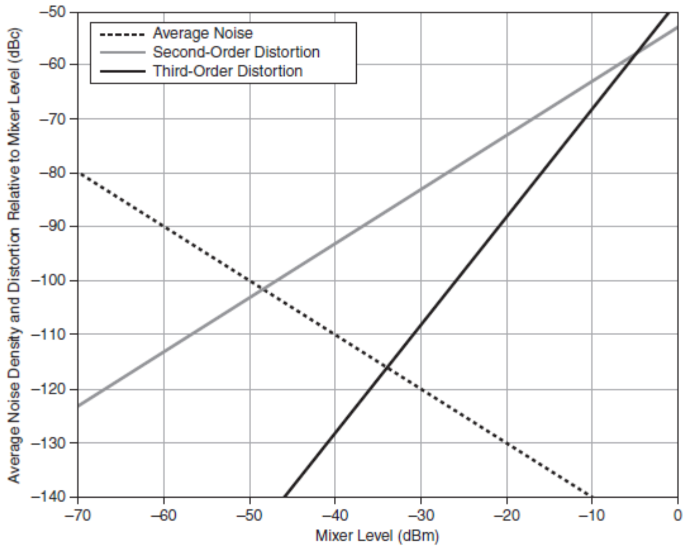

Figure 6-16. Example Dynamic Range Chart

Admittedly the axis labeling lacks clarity and most data sheets do not offer information on how to use this chart for dynamic range optimization. The next part of the discussion is centered on the construction of this chart. By showing how this chart is constructed, more insight in its use will be apparent. More importantly, some signal analyzer data sheets do not contain this chart. Being able to construct this chart with measured data will allow the user to know how to optimize the signal analyzer's dynamic range.

Figure 6-17 shows another view of the dynamic range chart with some clarity added.

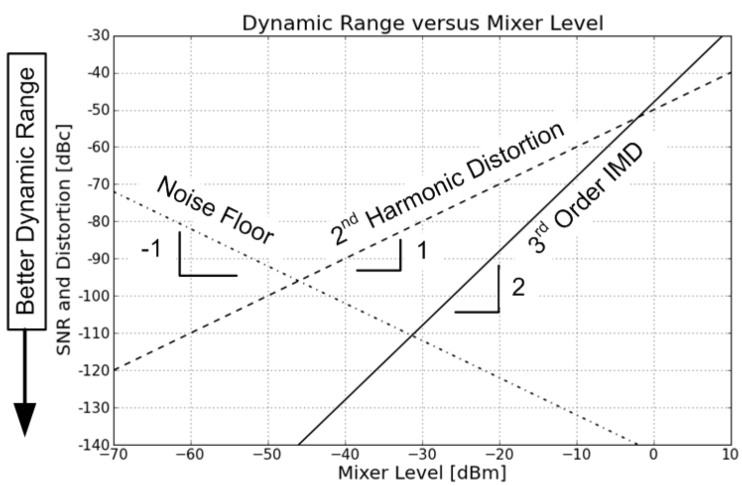

Figure 6-17. Dynamic Range Chart Definition

The three lines drawn are the *noise floor, 2nd harmonic distortion,* and 3rd *order intermodulation distortion*. When considering the noise floor curve, the y-axis serves as the noise power relative to the signal amplitude with units of dB relative to the carrier (dBc). This is signal-to-noise ratio with the sign inverted. When considering the distortion curves, now the y-axis represents the distortion amplitude relative to the fundamental signal level (the units are dBc). For both noise and distortions, the y-axis represents dynamic range with better dynamic range performance towards the bottom of the chart.

The x-axis is the mixer level. Why not use the signal analyzer's RF input power level for the x-axis? This is certainly possible; however, a separate chart for each RF input attenuation value would need to be constructed. By normalizing x-axis power to the input of the first mixer, one chart only needs consideration. Since changing the test signal's power level or the RF input attenuation setting affects the mixer level, using mixer level as the x-axis adds further flexibility in that the dynamic range chart is valid for any combination of power settings.

Noise Floor Curve on the Dynamic Range Chart

Assume there is no excess noise from the signal being measured. With this stipulation, the noise power at the input of the signal analyzer is kTB or -174 dBm/Hz for a 50 Ohm system. For passive components whose characteristic impedance is the same as the system (50 Ohm in this case), the noise floor cannot drop below kTB. Considering that only passive devices are present before the first mixer in the signal analyzer with the preamplifier bypassed, the noise floor at this point is kTB. Viewing the signals in **Figure 6-18**, with constant noise floor, for every one dB drop in the signal level, the SNR drops one dB.

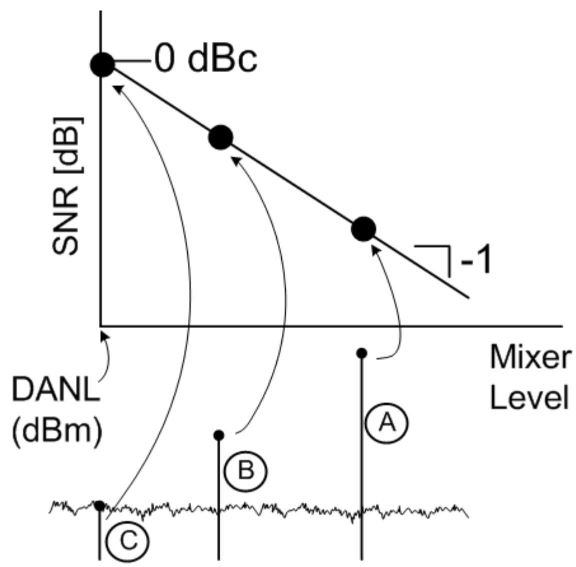

Figure 6-18. Mapping SNR to Dynamic Range Chart

Mapping the signals onto the dynamic range chart in **Figure 6-16**, signal A with its higher SNR relative to signal B falls lower on the y-axis. The dB per dB relationship with SNR and signal level gives the noise curve slope of -1. The signal analyzer will add noise to the measured results; however, the SNR to mixer level slope still stays the same value of -1.

This straight line curve is a y = mx + b equation to solve. The missing element is the offset, b. Consider when the signal level equals the noise floor, depicted by signal C in **Figure 6-18**. The SNR is by definition, 0 dBc. Solving for offset in the straight line equation shows b to equate to the value of the signal analyzer's DANL.

In **Figure 6-18**, the DANL in a 1 Hz RBW setting is -150 dBm. The equation of line 1 in **Figure 6-18** is:

$$y = -x - 150.$$

Figure 6-19 also shows the noise floor curve as a function of RBW setting. For the measurement of CW signals, the signal level is independent of the RBW setting, but from **Equation 5-2**, the noise varies by

10log(RBW ratio). In **Figure 6-19**, curve B is with an RBW setting of 1 kHz, giving a 30 dB offset in SNR over curve A which uses a 1 Hz RBW setting.

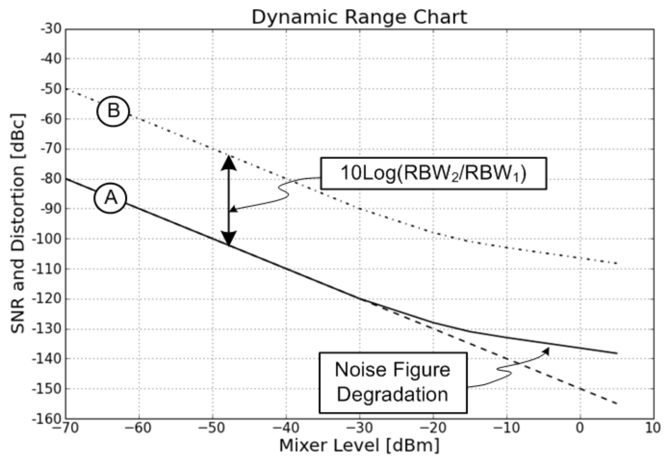

Figure 6-19. RBW Effect on Noise Curve

Figure 6-19 also shows some more realistic behavior of most signal analyzers in that as the mixer level increases, the system gain drops. This usually has an effect on noise figure degradation at these higher mixer levels.

Distortion Curves on the Dynamic Range Chart

In **Figure 6-11** we demonstrated that for a second order distortion product, every one dB change in the fundamental tone power, the distortion level changes two dB. The dynamic range chart plots relative distortion on the y-axis and **Figure 6-20** shows how to map absolute distortion level to relative distortion level.

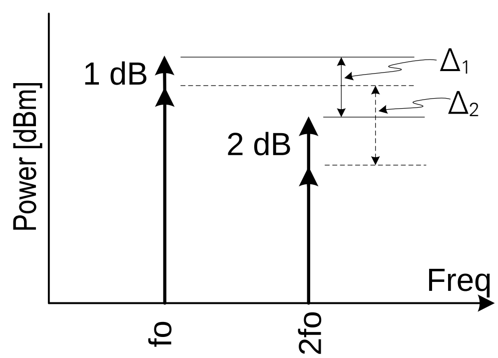

Figure 6-20. Relative 2nd Harmonic Distortion Level Change

With the fundamental a lower power setting, the relative 2nd harmonic power is Δ_2 dBc. Increasing the fundamental power by one dB, the relative 2nd harmonic power becomes Δ_1 dBc. The difference between Δ_2 and Δ_1 is +1 dB. The slope of the line in the dynamic range chart is the change in the relative harmonic level divided by the change in mixer level; hence, the slope is +1. In fact, the slope of any distortion line on the dynamic range chart is always one less than the order of the distortion.

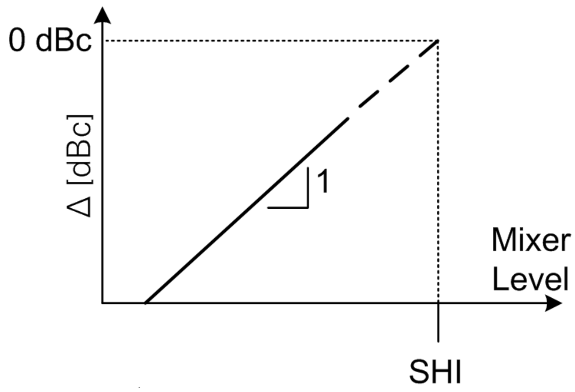

Figure 6-21. Mapping Second Harmonic Distortion to the Dynamic Range Chart

Again, there is a y = mx + b straight-line equation to solve where the offset, b, is unknown. From the definition of second harmonic intercept, when the mixer level is at the SHI value in dBm, the second harmonic is 0 dBc, yielding a value of –SHI for the offset, b. Even though this mixer level is beyond the gain compression level, it can still be used to draw the second harmonic distortion curve.

For third order IMD, a similar analysis reveals that for every one dB change in mixer level, the difference in relative distortion levels is two dB as shown in **Figure 6-22**.

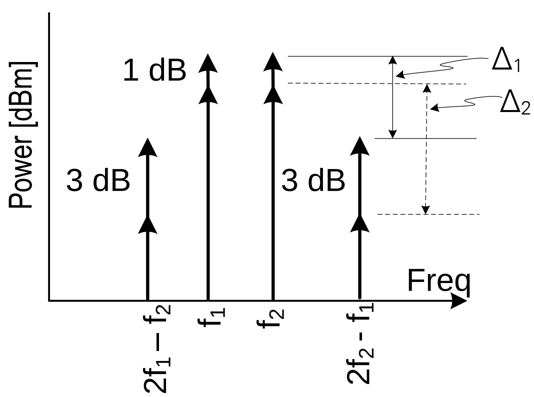

Figure 6-22. Relative 3rd Order IMD Level Change

The resulting slope of the IMD curve on the dynamic range is then +2. Like for the case of SHI, when the mixer level is at the TOI value, the IMD level is 0 dBc. Solving the y = mx + b strait-line equation shows that offset, b, equates to a value of -2xTOI.

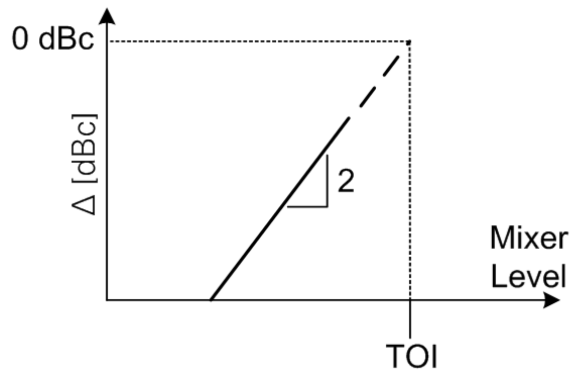

Figure 6-23. Mapping 3rd Order IMD to the Dynamic Range Chart

Figure 6-24 shows the second harmonic distortion and the third order IMD curves. In this example, SHI equals +30 dBm and TOI equals +20 dBm. The shaded region corresponds to power levels above the maximum mixer level where the IF overload error may occur. However, this graph shows how the two curves anchored at 0 dBc with fictitious mixer level values equal to either SHI or TOI

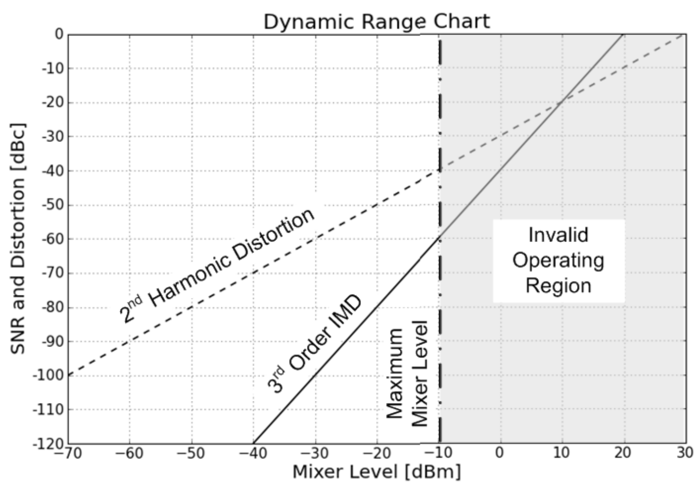

Figure 6-24. Distortion Curves on the Dynamic Range Chart

Phase Noise Curve on the Dynamic Range Chart

Phase noise when plotted on the dynamic range chart is the value at one particular offset frequency from the carrier. As shown in **Figure 6-25**, the single sideband phase noise value is a function of offset frequency,

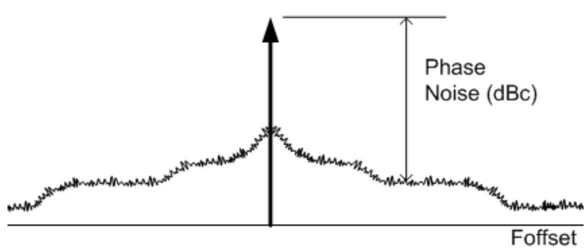

Figure 6-25. Phase Noise Pedestal

However, the relative phase noise power is constant as a function of carrier power level. Phase noise specification most often is normalized to a 1 Hz RBW value. Add 10Log(RBW) to the phase noise specified in dBc/Hz to arrive at the actual dBc value of phase noise.

On the dynamic range chart, phase noise appears as a horizontal line as shown in **Figure 6-26**.

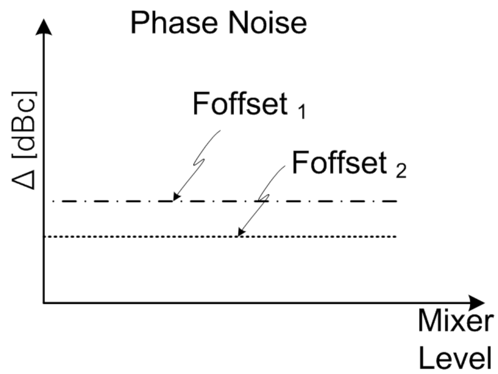

Figure 6-26. Phase Noise on the Dynamic Range Chart

Phase noise values at two different frequency offsets are shown.

Complete Dynamic Range Chart

Including the thermal noise, 2nd harmonic distortion, 3rd order intermodulation distortion and phase noise completes the dynamic range chart as shown in **Figure 6-27**.

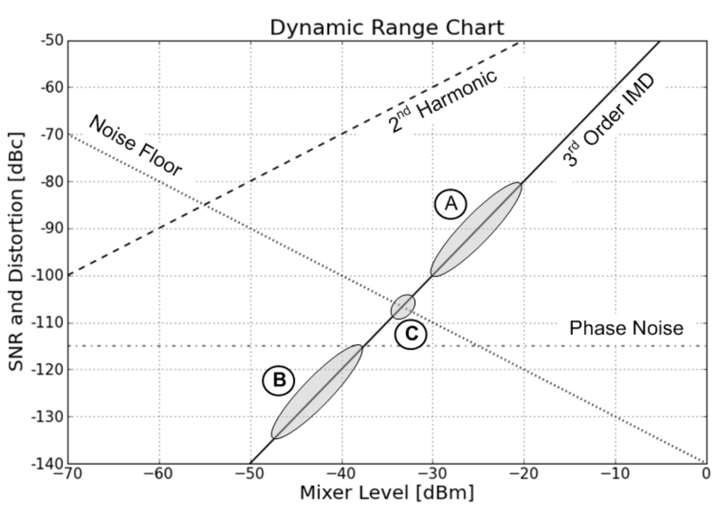

Figure 6-27. Complete Dynamic Range vs. Mixer Level Chart

The *noise floor* sets a hard line for this chart. Above the noise floor line, the signal can be viewed, but below the line, the signals are buried in noise. **Figure 6-28** demonstrates this concept. The letters correspond to the regions shown in **Figure 6-27**.

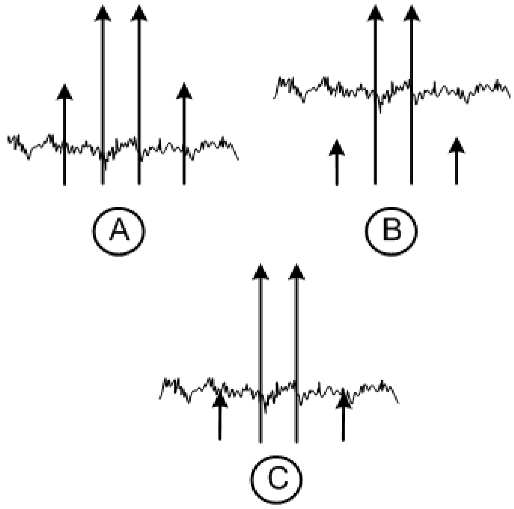

Figure 6-28. 3rd Order IMD at Various Mixer Levels.

When the mixer level is set for operating in region A, the distortion products are above the noise floor. The dynamic range of the signal analyzer is distortion limited. When the mixer level is set for operating in region B, the distortion products fall below the noise floor. In this case the dynamic range is noise limited. If the mixer level is set for region C, the distortion and noise levels are the same. The dynamic range is at its maximum when operating in region C.

When the mixer level is adjusted for maximum dynamic range, this value is termed the optimum mixer level. As shown in **Figure 6-29**, there is an optimum mixer level for 2nd harmonic distortion and one for 3rd order IMD.

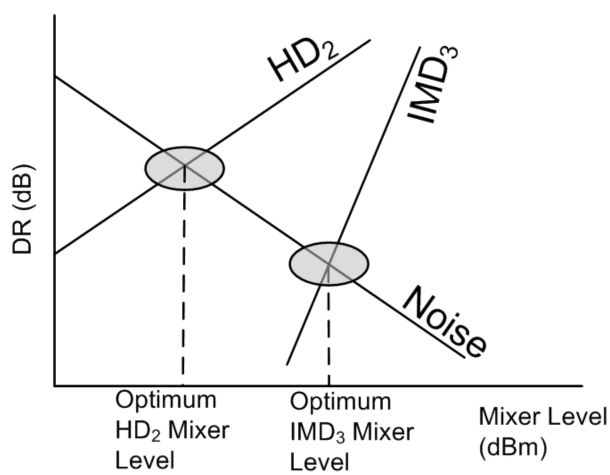

Figure 6-29. Optimum Mixer Levels

Some geometry on the curves shown in **Figure 6-29** yields the following equations for maximum dynamic range and the corresponding optimum mixer level.

3rd Order IMD:

Maximum dynamic range: $DR_{max} = (2/3)[TOI – DANL]$

Equation 6-7

Optimum mixer level: MLopt = (1/3)[2TOI + DANL]

Equation 6-8

2nd Order Harmonic Distortion:

Maximum dynamic range: DR_{max} = (1/2)[SHI – DANL]

Equation 6-9

Optimum mixer level: MLopt = (1/2)[SHI + DANL]

Equation 6-10

The following is a worked example for **Figure 6-27**:

RBW = 10 Hz
DANL = -150 dBm/Hz
DANL = -140 dBm
TOI = +20 dBm
DR_{max} = (2/3)[20 – (-140)]
= 106.7 dB
MLopt = (1/3)[2(20) + (-140)]
= -33.3 dBm

A Few Observations Regarding the Dynamic Range Chart

One of the first mental barriers to overcome is the seemingly low mixer level for best dynamic range performance. Users who are familiar with digitizers know that for best dynamic range, the signal should be set as high as possible (see **Figure 4-6** for more information on this topic). But, the dynamic range chart as we have defined it so far, does not consider the ADC. It considers only the frontend.

From **Figure 6-27**, the RF input attenuator must be set quite high to achieve around -35 dBm for maximum 3rd order IMD dynamic range and -55 dBm for maximum 2nd harmonic distortion dynamic range. During these measurements, it is quite easy to assume that the excessive noise with such low mixer levels impedes good dynamic range performance. But, the chart clearly demonstrates that the signal analyzer's best dynamic range performance is at these seemingly low mixer levels.

Figure 6-30 shows that a 10 dB improvement in noise floor equates to a 5 dB improvement in 2nd harmonic dynamic range and 6.67 dB in 3rd order IMD dynamic range performance.

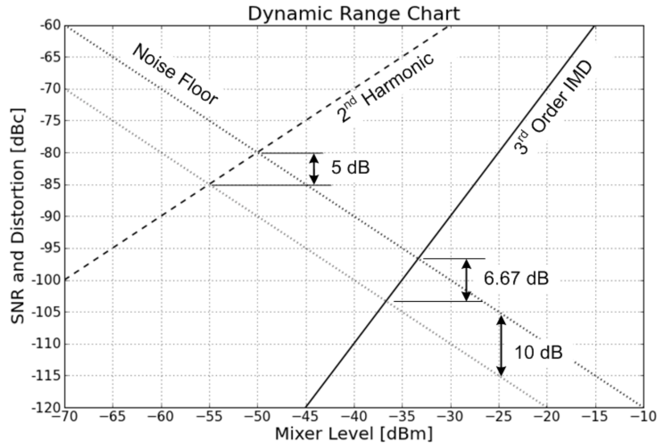

Figure 6-30. Noise Floor and the Dynamic Range Chart

Figure 6-31 shows two things: that higher TOI results in an IMD line that is displayed lower on the dynamic range chart and that a 10 dBm improvement in TOI lowers the IMD line 20 dB, but all this improvement still only yields a 6.67 dB improvement in dynamic range performance.

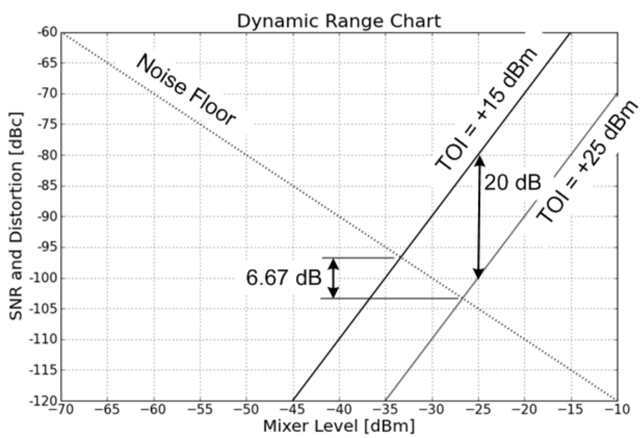

Figure 6-31. 3rd Order IMD and the Dynamic Range Chart

Preamplifier Dynamic Range

Does cascading a preamplifier in front of the signal analyzer help with dynamic range? From **Equation 4-6**, the preamplifier can certainly improve system noise figure. To answer the question of dynamic range, a numeric example is worked out. **Figure 6-32** shows the signal analyzer and preamplifier performance parameters for this example.

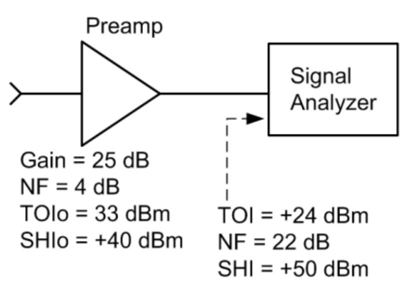

Figure 6-32. Preamp Cascaded with the Signal Analyzer

Equation 4-6 allows the cascaded preamplifier plus signal analyzer noise figure to be calculated. The following equations can be used to compute cascaded SHI and TOI:

$$SHI_{sys} = -20\log\left\{10^{(-SHI_1'/20)} + 10^{(-SHI_2'/20)}\right\}$$

Equation 6-11

$$TOI_{sys} = -10\log\left\{10^{(-TOI_1'/10)} + 10^{(-TOI_2'/10)}\right\}$$

Equation 6-12

The SHI' and TOI' are values referred to the system input by subtracting the system gain that precedes the device. For instance, SHI' for the signal analyzer is SHI of the signal analyzer minus the preamplifier gain or +50 – 25, or +25 dBm.

The results of the signal analyzer dynamic range and the cascaded preamplifier plus the signal analyzer dynamic range are summarized in **Table 6-1**.

Table 6-1. Preamp Dynamic Range Results

	Signal Analyzer	Preamp + Signal Analyzer	Change
TOI	+24 dBm	-1.5 dBm	-25.5 dB
SHI	+50 dBm	+12.6 dBm	-37.4 dB
Noise Figure	22 dB	4.8 dB	-17.2 dB
DANL	-152 dBm/Hz	-169.2 dBm/Hz	-17.2 dB
Max 3rd order dynamic range	117.3 dB	111.8 dB	-5.5 dB
Max 2nd harmonic dynamic range	101 dB	90.9 dB	-10.1 dB

Indeed, the system noise figure improves by 17.2 dB, but the system TOI degrades by 25.5 dB and system SHI degrades by 37.4 dB. The resulting dynamic range degradations for 3rd order IMD and 2nd harmonic are 5.5 dB and 10.1 dB respectively.

Figures 6-33 and 6-34 show the dynamic range charts for 3rd order IMD and 2nd harmonic distortion for the signal analyzer with and without the cascaded preamplifier. The x-axis requires some explanation. With the preamplifier in the system, the x-axis represents the power level at the input of the preamplifier. Note the degradation in dynamic range with the preamplifier on. Also note how much lower the power level must be with the preamplifier in the system for best dynamic range.

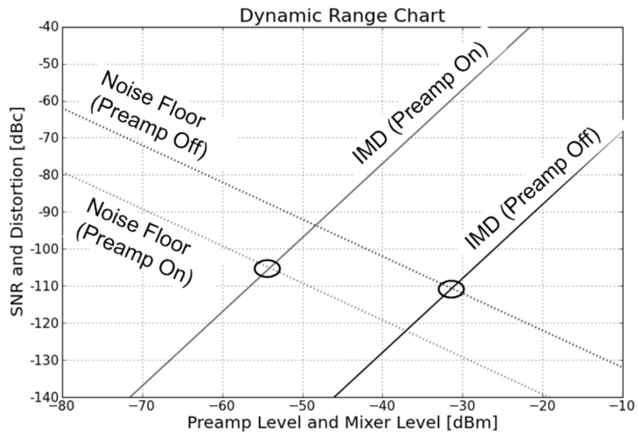

Figure 6-33. Preamp 3rd Order IMD and the Dynamic Range Chart

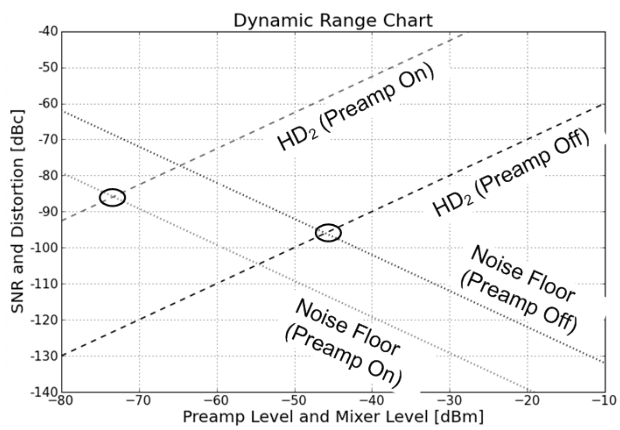

Figure 6-34. Preamp 2nd Harmonic Distortion and the Dynamic Range Chart

So the answer to the above question is no, the preamplifier does not improve dynamic range. The degradation in distortion is worse than the improvement in noise figure. Use the preamplifier for its intended purpose, which is to improve system sensitivity for the measurement of near noise signals. Do not use it in the presence of large amplitude signals that may produce distortion.

Near Noise Distortion Measurements

In reality, as the CW distortion signal approaches the noise floor the signal and noise combine. In the dynamic range chart, a true intersection to the distortion and noise curves does not actually occur. **Figure 6-35** demonstrates the effect of near-noise measurements.

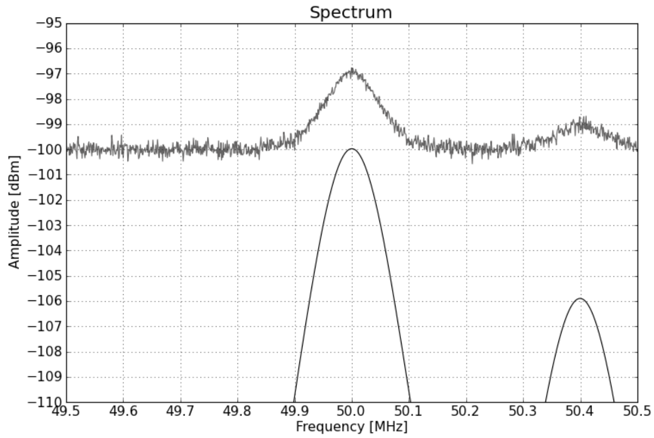

Figure 6-35. Near Noise Signals

The signal at 50 MHz in **Figure 6-35** has an amplitude equal to the noise floor. However, the noise floor shows a 3 dB addition. The dynamic range charts as drawn so far would not show this 3 dB noise bump. Even when the signal is below the noise floor, the noise floor shows a bump. At 50.4 MHz in **Figure 6-35**, the signal is 6 dB below the noise floor, yet the noise floor shows a 1 dB bump above this signal.

Figure 6-36 shows the combined signal plus noise near the intersection of the distortion and noise curves.

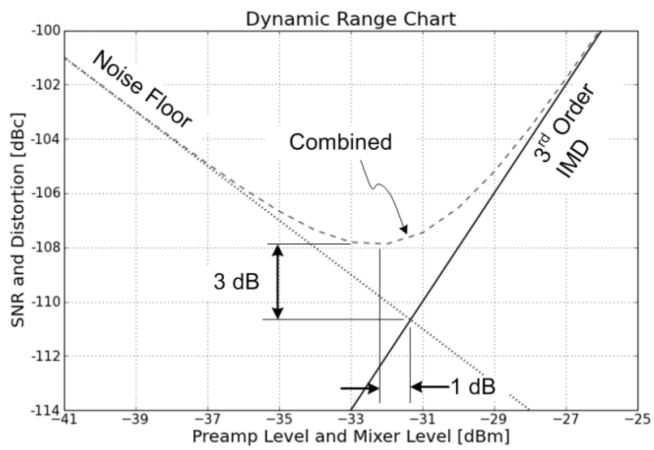

Figure 6-36. Near Noise Signals on the Dynamic Range Chart

Note that the signal plus noise adding in the manner described is only valid for RMS averaged noise. For 3^{rd} order IMD and noise, the maximum dynamic range occurs one dB lower in mixer level than the optimum mixer level described by **Equation 6-8**. For second order distortion, the optimum mixer level described by **Equation 6-10** is still valid when considering the distortion plus noise curve.

RF Input Attenuator Effect on Dynamic Range

The RF input attenuator affects dynamic range in two different ways, the first way being the range of RF input attenuation values. The highest performance signal analyzers have 70 dB RF input attenuation range. With +30 dBm maximum measureable input power, the 70 dB allows for a -40 dBm mixer level. Using the dynamic range chart of **Figure 6-27** as a guide, optimum mixer levels for 3^{rd} order IMD falls in the -30 dBm to -40 dBm mixer level range. RF input attenuator range is set in order to measure the highest measureable signals with best dynamic range.

The other way the RF input attenuator affects dynamic range is with its step resolution. The maximum dynamic range only occurs if the mixer level can hit the optimum value. With coarse RF input attenuator resolution, the optimum mixer level may never be achieved. **Figure 6-37** demonstrates the potential unachievable dynamic range as a function of RF input attenuator resolution.

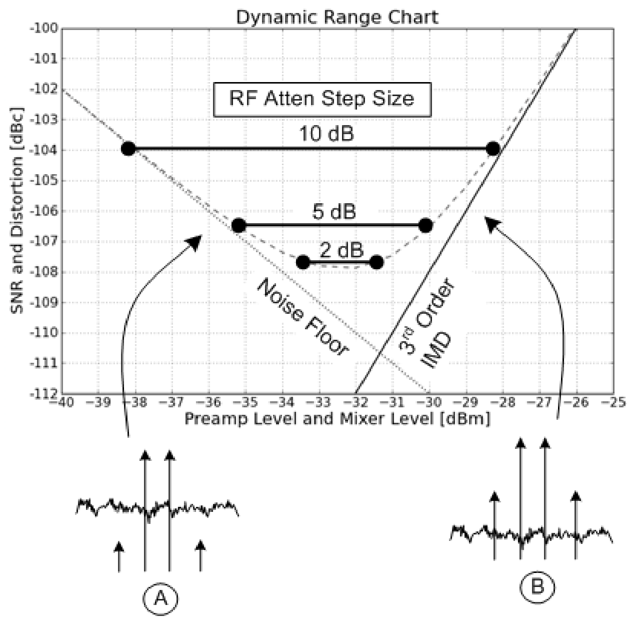

Figure 6-37. RF Input Attenuator Step Size vs. Dynamic Range

In worst case scenarios, the mixer level could bounce back and forth between noise limited dynamic range (graphic A in **Figure 6-37**) and distortion limited dynamic range (graphic B in **Figure 6-37**).

Table 6-2 summarizes the potential dynamic range that may be given up versus RF input attenuator step size.

Table 6-2. Potential Unavailable Dynamic Range vs. RF Input Attenuator Step Size

RF Atten. Step Size	Potential Unavailable Dynamic Range (Ideal)	Potential Unavailable Dynamic Range (Actual)
10 dB	6.7 dB	4.08 dB
5 dB	3.3 dB	1.29 dB
2 dB	1.3 dB	0.23 dB
1 dB	0.7 dB	0.06 dB

In **Table 6-**2 the i*deal* column uses the unrealistic noise and 3rd order IMD lines intersecting with no near-noise addition and the *Actual* column is with the noise and distortion curves adding, which is the true behavior of the signal analyzer. The *Ideal* column could also be used to determine the dynamic range change when operating in region A of **Figure 6-27**.

One can see that there is a diminishing return on dynamic range with smaller RF input attenuator step size. When considering optimization at the maximum dynamic range location, a two dB RF input attenuator resolution may be sufficient.

ADC Contribution to Dynamic Range

As briefly mentioned in section 4, the ADC has a very real impact on the signal analyzer's dynamic range. How the ADC contributes to the dynamic range performance of the signal analyzer is now presented.

Representative spurious-free dynamic range (SFDR) data for an ADC is shown in **Figure 6-38** [[16]].

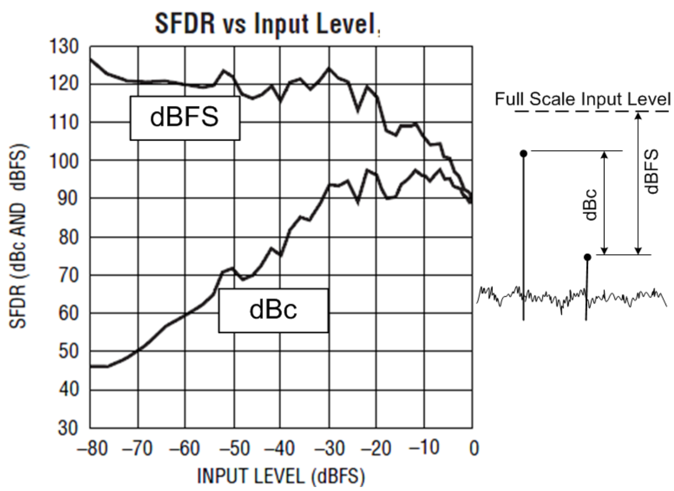

Figure 6-38. ADC Spurious-Free Dynamic Range

The two lines are different viewpoints of the same data. The spectrum graphic in **Figure 6-38** helps explain the meaning of the two curves. The *dBFS* curve is the distortion amplitude relative to the maximum ADC input amplitude (full scale input level). The *dBc* curve is the distortion amplitude relative to the fundamental signal. The *dBc* curve is more relevant to the mapping of the ADC distortion onto the dynamic range chart.

Studying this ADC SFDR data reveals that for the fundament signal amplitude near the full scale level that the distortion improves as the fundamental amplitude decreases. However, below -15 dBFS, the distortion amplitudes do not decrease with decreasing fundamental signal amplitude. Reference [[6]] describes the reasons for this behavior in the pipelined ADC topology which is primarily used for IF sampling in today's signal analyzer structures.

Figure 6-39 demonstrates how the two different behavioral regions of the ADC SFDR data relate to the dynamic range chart.

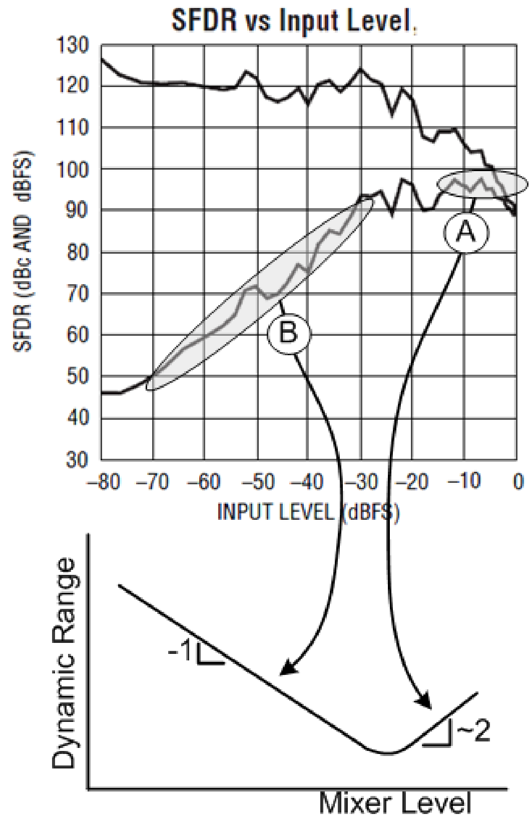

Figure 6-39. ADC Spurious-Free Dynamic Range Mapped to the Dynamic Range Chart

In region A, the dynamic range of the ADC improves with decreasing mixer level. Mixer level in this case is directly related to the ADC input power through the IF gain of the signal analyzer. It is assumed that there is enough IF gain to drive the signal to the ADC's full scale input level. The slope of the dynamic range curve in region A is approximately two for two-tone IMD. In region B, the distortion amplitudes stay fixed for any given fundamental signal level. The slope for the dynamic range curve is -1. This is the same slope as the noise curve, but keep in mind that it is ADC distortion that is being discussed here.

The ADC distortion looks pretty damaging to the dynamic range performance of the signal analyzer, and it can be if the signal analyzer settings are not controlled properly. It is here that we must consider vector signal analysis separately from spectrum analysis.

ADC Distortion Associated with Vector Signal Analysis

Recall that the term vector signal analysis refers to a relatively wide analog IF bandwidth where the signal and distortion components fall inside this bandwidth. In this case the SFDR of the ADC dominates distortion performance for multi-tone intermodulation distortion.

In **Figure 6-40**, the ADC distortion curve is shown with the frontend distortion. These curves are for two-tone IMD.

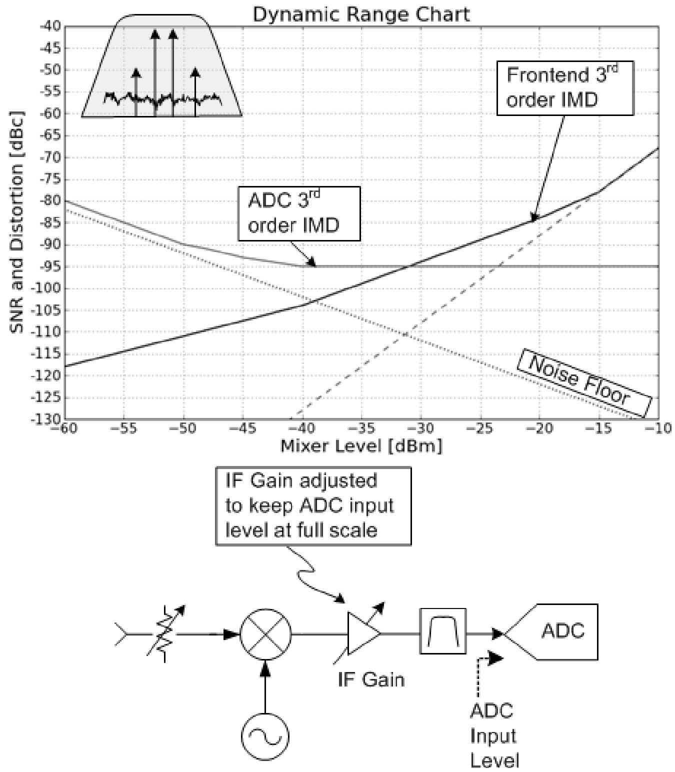

Figure 6-40. Dynamic Range Chart in VSA Mode. IF Gain Constantly Adjusted.

The overall system distortion will be the worst of the ADC and frontend distortion. It is broken out in **Figure 6-40** for discussion purposes. In this particular case, the IF gain is adjusted at each mixer level setting to ensure that the fundamental signal amplitudes are brought close to the ADC's full scale level. The IF gain is indirectly controlled with the reference level. When operating near full scale, the ADC's SFDR is maximized. At a certain mixer level, approximately -40 dBm in **Figure 6-40**, the IF gain maxes out. Once there is no more IF gain to offer, the ADC can no longer operate at full scale and the SFDR performance begins to drop.

Notice the frontend IMD performance in **Figure 6-40**. As the mixer level drops and more IF gain is requested, the final analog amplifiers now begin to dominate the frontend TOI performance. Each device contributes to the system TOI as shown in **Equation 6-13**.

$$TOI_{sys} = TOI_{device} - Gain_{before\ device}$$

Equation 6-13

As the system gain in front of the device increases, which is the case for low reference level settings, the device's TOI becomes a larger contributor to the overall system TOI.

Figure 6-41 shows what can happen when the user is not careful in adjusting IF gain (via the reference level setting) to ensure near full scale ADC operating level.

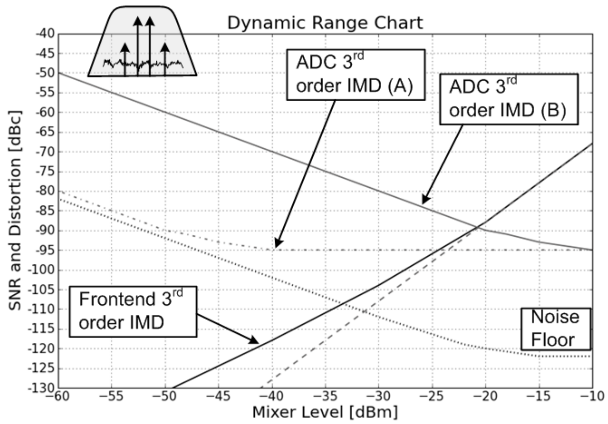

Figure 6-41. Dynamic Range Chart in VSA Mode. IF Gain not Adjusted.

The line labeled *ADC 3^{rd} order IMD (A)* is the case from **Figure 6-40** where IF gain adjustment takes place. However, the line labeled *ADC 3^{rd} order IMD (B)* is where the reference level setting is fixed at the highest mixer level value shown (-10 dBm in this case). Now the ADC SFDR completely dominates the signal analyzer's dynamic range performance.

Careful comparison between **Figure 6-40** and **Figure 6-41** shows that the frontend IMD degradation is not as severe when the IF gain is not constantly being adjusted. From **Equation 6-13**, less gain in front of the device equates to better system TOI performance. The noise floor at high mixer levels does start to degrade. Noise is opposite from distortion in that more system gain is required in front of any given device to make that device's noise contribution to the overall system noise as small as possible. ADCs have very large noise figures (> 30 dB) requiring a great deal of up front system gain to ensure the ADC's noise is not dominating the system noise performance. At higher mixer levels, the IF gain is lower and the ADC's noise starts to dominate. This, however, is not really a problem for system dynamic range performance as the distortion at high mixer levels dominates over noise.

To summarize in vector signal analysis mode the wide analog IF bandwidth exposes the ADC to the fundamental signals during the measurement of the distortion components. Because of this, the dynamic range of the signal analyzer is limited to the SFDR performance of the ADC. However, in many actual applications, the ADC cannot be made to operate at full scale input level, and then the true SFDR degrades.

ADC Distortion Associated with Spectrum Analysis

For spectrum analysis, the analog IF bandwidth is narrow enough to select only the fundamental signal or the distortion signal at any given time. **Figure 6-42** demonstrates the effect of narrowband filtering in the final IF of the signal analyzer.

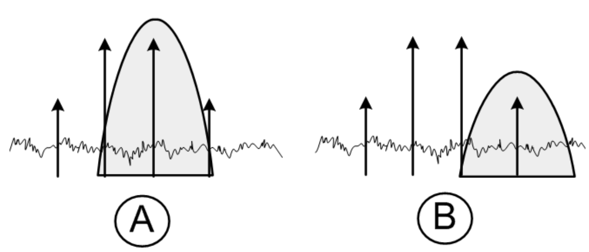

Figure 6-42. IMD Measurements using Spectrum Analysis Mode.

At point A in **Figure 6-42**, one of the fundamental signals in this two-tone IMD measurement is selected. At point B, the distortion product is being measured. The IF filter bandwidth is narrow enough that all signals, except the distortion tone being measured, are attenuated.

Figure 6-43 shows that the location of the analog IF filters should be placed before both the ADC and the final analog IF gain stages. In fact, placing the IF filters as far up the signal chain as possible is a goal. However, practical reasons of ensuring high enough filter quality factor, Q, for the narrowest bandwidths usually restricts placement of the IF filter in the final IF.

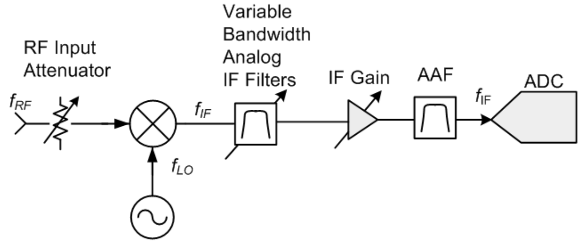

Figure 6-43. Final IF Amplifier and ADC Affected by Analog IF Filters

By not allowing the relatively large amplitude fundamental signal to reach both analog devices and the ADC, which sit subsequent to the narrowband IF filter, gives the signal analyzer several advantages over the wideband IF used in vector signal analysis mode:

- The chances of gain compression and worse, delivering signals whose amplitude are greater than the ADC's full scale input level, are diminished
- The effective TOI of the devices following the IF filter is dramatically improved
- The IF gain can be increased which makes the ADC noise figure contribution to the overall system noise much less

Of the advantages listed, the effective TOI improvement is the one that renders itself to a quantitative analysis.

Effective TOI Improvement for Analog Devices

Figure 6-44 is used to illustrate how the analog narrowband IF filters improve the effective TOI of an analog active device. A two-tone signal drives the input of a bandpass filter with 'R' dB of stopband rejection placed in front of the analog gain device.

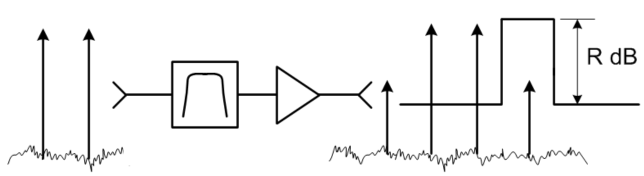

Figure 6-44. Effective TOI Enhancement with Filtering in Front of Analog Device

From **Equation 6-4**, the third order IMD products resulting from fundamental signals with amplitude, A, are described by **Equation 6-14.**

$$IMD_{lower} = kA^3 cos[(2\omega_1 - \omega_2)t]$$
$$IMD_{upper} = kA^3 cos[(2\omega_2 - \omega_1)t]$$

where k is a proportionality constant
let r be the linear representation of the filter rejection.

Equation 6-14

$$r = 10^{-R/20}$$

Equation 6-15

With the filter, the IMD signal equations become **Equation 6-16**.

$$IMD_{lower} = k(rA)^3 cos[(2\omega_1 - \omega_2)t]$$
$$IMD_{upper} = k(rA)^3 cos[(2\omega_2 - \omega_1)t]$$

Equation 6-16

The ratio of the amplitudes of the IMD products with the filter and without the filter is shown in **Equation 6-17**.

$$\frac{IMD\ amplitude\ with\ filter}{IMD\ amplitude\ without\ filter} = \frac{kr^3A^3}{kA^3} = r^3$$

Equation 6-17

On a log scale, the IMD amplitudes are reduced by *R dB*. Without filtering, TOI of the active device is given by **Equation 6-18**.

$$TOI_{device} = P_{fund} + \frac{P_{fund} - P_{IMD}}{2}$$

where P_{fund} is the power level of the fundamental signal in dBm and P_{IMD} is the power level of the distortion product.

Equation 6-18

Adding the filter gives an effective TOI shown in **Equation 6-19**.

$$TOI_{effecive} = P_{fund} + \frac{P_{fund} - (P_{IMD} - 3R)}{2}$$
$$TOI_{effective} = TOI_{device} + {}^{3}/_{2}\,R$$

Equation 6-19

Equation 6-19 demonstrates that by placing a filter in front of an analog device the effective TOI of the combination increases by 1.5 times the filter rejection in dB. This assumes both fundamental tones are attenuated by the same amount R. Actual filters normally do not have constant stopband rejection, so the results in equation may be conservative.

The effective TOI improvement given by **Equation 6-19** is quite startling. Designers of active devices go through heroic efforts to improve the TOI by a modest 2 or 3 dB. To yield a 30 dB improvement using very moderate 20 dB filter rejection is quite profound.

Effective TOI Improvement for ADCs

The act of placing a filter in front of the ADC also improves effective TOI. However, a closed form equation cannot be rendered. The analysis is ADC device dependent and involves careful study of the particular device's SFDR chart similar to that shown in **Figure 6-39**.

To illustrate the steps in determining the effective TOI of an ADC with a bandpass filter placed in front of it, **Figure 6-45a** and **Figure 6-45b** is used.

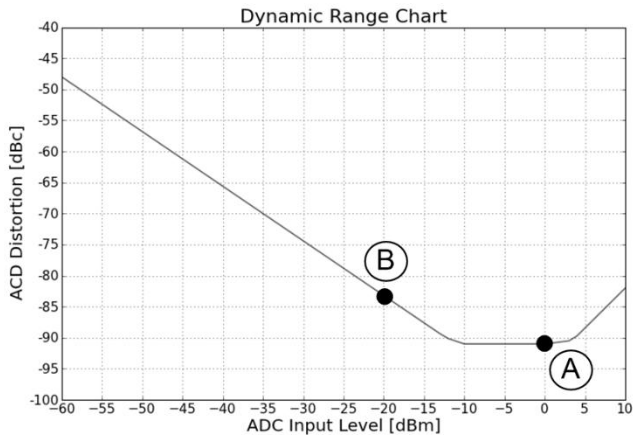

Figure 6-45a.

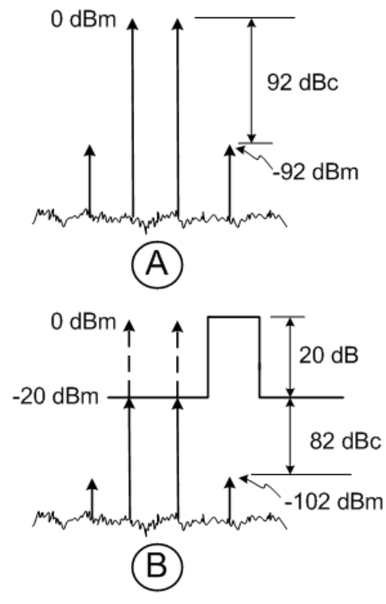

Figure 6-45b.
Effective TOI Enhancement with Filtering in Front of an ADC

In **Figure 6-45b**, the graphic labeled A, shows the fundamental signals placed at 0 dBm at the ADC input. **Figure 6-45a** shows the position of A on the dynamic range chart. **Figure 6-45a** is the ADC's SFDR data mapped to the dynamic range chart. The x-axis is the ADC input level in dBm with a full scale input level of +10 dBm. So 0 dBm in graphic A corresponds to -10 dBFS. The distortion product amplitudes are -92 dBc, or -92 dBm with a 0 dBm fundamental signal level.

In graphic B, the bandpass filter with constant stopband rejection of 20 dB is placed in front of the ADC. During the portion of the measurement where the IMD products are being measured, the fundamental tones are attenuated be the 20 dB filter stopband rejection. **Figure 6-45a** shows that at point B, the ADC distortion is approximately 82 dB, which places the distortion amplitudes at -102 dBm.

The distortion measurement is a two step process: first, measure the fundamental signal amplitude and then measure the IMD product amplitude. From a measurement prospective, the IMD amplitude is relative to the unfiltered amplitude of the fundamental signal. Without the filter, the dynamic range of the IMD measurement is 92 dB, which is the limit of the ADC's SFDR performance. With the filter, the dynamic range is roughly 10 dB better, which is an effective TOI improvement of 5 dB.

Empirical measurements bear out this roughly 10 dB improvement in dynamic range performance when using narrowband IF filtering compared to wideband filtering used for vector signal analysis measurements.

Reference Level Re-Ranging

Narrowband filtering certainly helps improve distortion performance of the ADC. However, using the filter alone still limits the overall signal analyzer's dynamic range. The dynamic range charts provided in this chapter show the frontend capable of greater than 110 dB of maximum third order dynamic range. Modest TOI of +18 dBm and DANL = -150 dBm/Hz (rms) yields 112 dB dynamic range using **Equation 6-7**. The filtered ADC still has 10 dB to 15 dB worse dynamic range than the frontend. To make the ADC's contribution to system level dynamic range small compared to the frontend, the extra step of so-called *reference level re-ranging* needs to take place.

Reference level re-ranging refers to using two different reference level settings, one for the measurement of the fundamental signals and one for the measurement of the distortion signals. The principle behind this relies on the fact that the ADC's distortion level measured in dB relative to full scale (dBFS) stays constant when operating 15 dB or so below full scale amplitude, shown in **Figure 6-38**.

Figure 6-46 illustrates an example of using reference level re-ranging to improve the effective dynamic range of the signal analyzer.

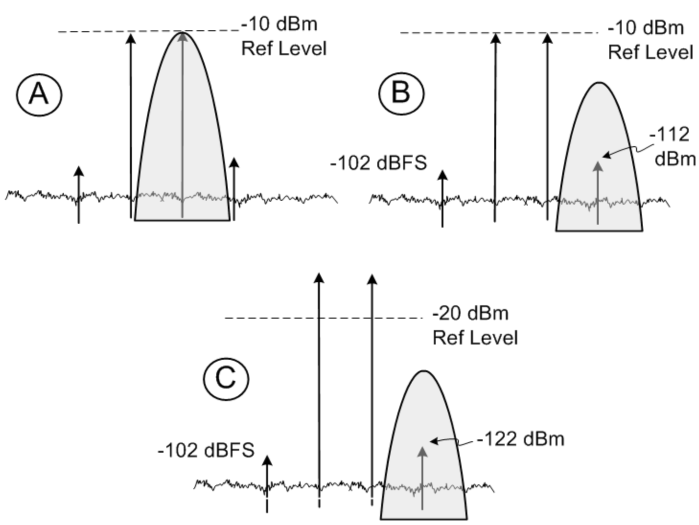

Figure 6-46. Re-Ranging the Reference Level to Improve ADC Effective TOI

Keep in mind that reference level is an indirect control of the IF gain. To be effective, most of the IF gain variation needs to take place in the stages that follow the narrowband analog IF filters.

In graphic A, the fundamental signal is being measured. The reference level is set so as to not overdrive the ADC's full scale input level. In graphic B, the distortion product is being measured without reference level re-ranging. In graphic C, the reference level has been re-ranged by 10 dB in this case. Because the large fundamental tones are attenuated by the filter's stopband rejection, the ADC does not overload, nor do the analog gain stages following the filter add distortion. Between graphics B and C, the ADC's distortion amplitude stays fixed at the same value relative to the ADC's full scale input level. However, the displayed amplitude of the distortion product is offset lower in value by the reference level change; 10 dB lower in this example.

Figure 6-47 summarizes the ADC performance for the various IF filtering modes. Wideband filtering, assuming that the reference level is constantly adjusted to bring the input signal at the ADC to the ADCs full scale input level results in system dynamic range limited to the performance of the ADC. This is usually in the 80 dB to 90 dB range. Using a narrowband filter for spectrum analysis mode results in approximately 10 dB improvement in dynamic range. As discussed, the best performance requires reference level re-ranging.

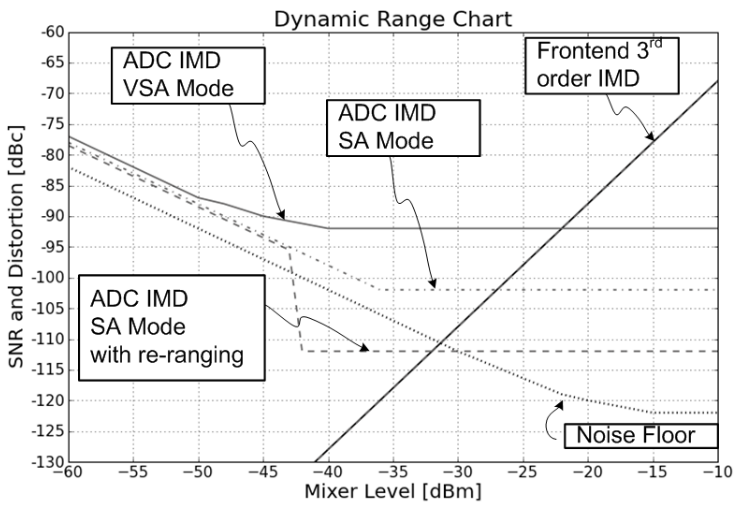

Figure 6-47. Summary of IF Filtering Effect on ADC IMD

The case for reference level re-ranging reverts back to the case of standard non re-ranging when there is no longer enough analog gain to support this mode. If the mixer level at the system, runs out of extra analog gain, falls below the optimum mixer level for maximum frontend dynamic range, then no loss in dynamic range performance is incurred.

Dynamic Range Considerations for Digitally Modulated Signals.

Nearly all of the discussion up to this point has involved only the analysis of CW signals. However, modern communications signals use digital modulation, which is relatively wideband. So much attention has been given to CW signals in this chapter that a question of relevance comes to mind. Several reasons for the detailed consideration of CW signals for the signal analyzer are as follows:

1. The signal analyzer, by tradition, is calibrated for the amplitude of a CW signal. This tradition will be hard to break.
2. For distortion and noise analysis, the digitally modulated signal can be broken down into CW sub-components.
3. Much of the design work for transmitters and receivers, even when their intended purpose is for digitally modulated signals, rely on CW characterization. Most often solid state devices such as mixer and amplifiers have only CW distortion performance specifications.

The dynamic range measurement considered in the following discussion is the measurement of spectral leakage into adjacent channels. **Figure 6-48** illustrates spectral re-growth generated by the signal analyzer. Suppose a distortion free digitally modulated signal is introduced at the input of the signal analyzer. Intermodulation distortion in the signal analyzer causes the shoulders seen in the adjacent channels. Primarily, this is third order distortion, but 5^{th} and sometimes 7^{th} order distortion components contribute to the spectral re-growth. Additionally, broadband noise can degrade SNR. The signal analyzer's phase noise is one more component that can become a limitation.

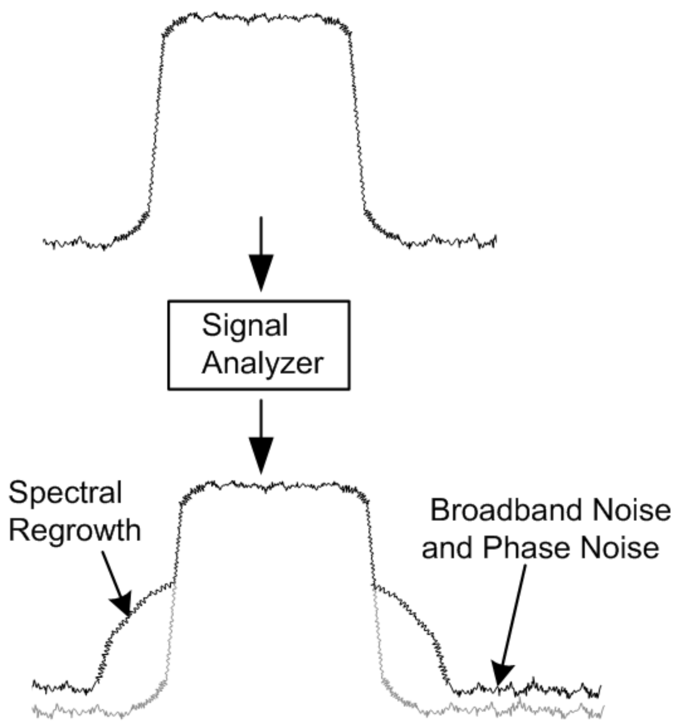

Figure 6-48. Spectral Re-growth Generation in the Signal Analyzer.

The measurement at hand is the characterization of a DUT. To ensure that the measured spectral regrowth is due to the DUT and not the signal analyzer, the signal analyzer's spectral re-growth must be made as small as possible.

One common metric for spectral re-growth is *adjacent channel power ratio* (ACPR) and a*djacent channel leakage ratio* (ACLR). In general, this measurement is the power in the main channel divided by the power in either the lower or upper adjacent channels. **Figure 6-49** shows the spectrum of this kind of measurement.

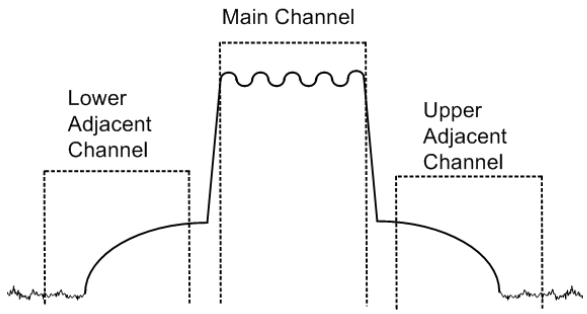

Figure 6-49. Defining Terms for the ACLR Measurement.

Different communication standards define their own unique methodologies in terms of the bandwidths used in integrated power in the adjacent channels. Some standards do not require integration for the power measurement, but rather define spot frequency measurements in clearly defined RBW and VBW filter settings. No matter the exact power ratio measurement methodology, the general ideas on signal analyzer generated noise and distortion associated with digitally modulated signal outlined in the following discussion are relevant.

Once again, whether the measurement uses wideband IF or narrowband IF, the dynamic range results and the measurement technique are different. Wideband in this case means that the IF bandwidth is wide enough to capture both the main channel and at least one of the adjacent channels. With wideband analysis (vector signal analysis) the intent is to rapidly measure both the main and the adjacent channels simultaneously. Narrowband IF in this case refers to the analog IF bandwidth significantly narrower than the main channel bandwidth. The intent for narrowband analysis (spectrum analysis) is that the main channel signal power is attenuated when measuring the adjacent channels. The concepts of wide and narrow band analysis are depicted in **Figure 6-50a and Figure 6-50b**.

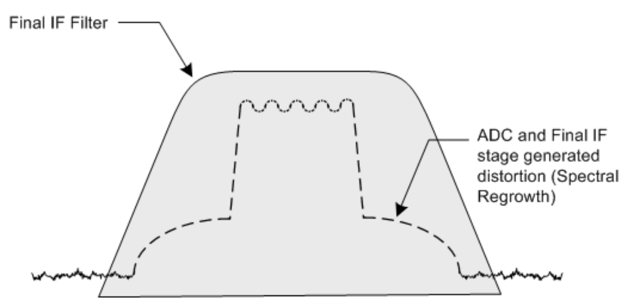

Figure 6-50a

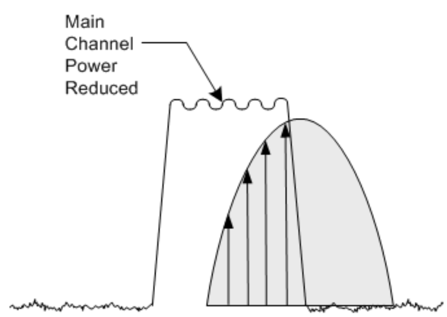

Figure 6-50b

Figure 6-50. Wide and Narrow IF Bandwidth Analysis on Modulated Signals

The goal here is to be able to represent the ACLR problem on the dynamic range chart for the same reasons as was done for the CW signal case, which is to determine maximum dynamic range and the associated optimum mixer level. For signal analyzers operating below gain compression, the analysis of internally generated distortion and noise can be approximated to a high degree of accuracy by breaking the digital signal down into a series of CW signal equivalents. First, the dynamic range chart's signal to noise ratio curve is analyzed.

SNR for Digitally Modulated Signal

In **Figure 6-8**, the idea behind the total integrated power of a signal with modulation was first presented. This partially describes the signal-to-noise ratio of the digitally modulated signal. **Figure 6-51** graphically describes SNR associated with digitally modulated signals.

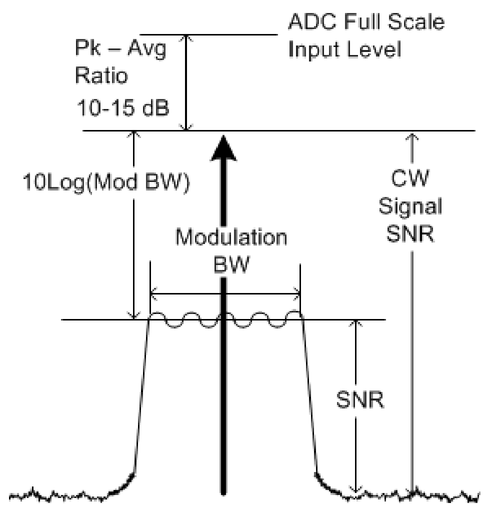

Figure 6-51. SNR of the Digitally Modulated Signal

As discussed earlier, the frontend components react according to total integrated power. The difference between a CW signal's power level and a digitally modulated signal's power level is approximately $10\log(BW_m)$, where BW_m is the modulation bandwidth. This can be significant: for BW_m = 1 MHz, this is a 60 dB difference; for a BW_m = 20 MHz, this is a 73 dB difference. For wide IF bandwidth, the crest factor or peak-to-average power ratio (PAPR) must be taken into account. Analog filtering will reduce the PAPR by a few dB, but to be safe it is wise to back off the mixer level by the PAPR of the signal.

After backing off by the PAPR and accounting for the modulation bandwidth, the SNR, as shown in **Figure 6-51**, is 75 dB to 90 dB lower than the SNR for an equivalent power CW signal.

For the CW signal, the SNR is a function of the resolution bandwidth filter setting. Not so for the digitally modulated signal. The signal itself is noise-like and of course the noise floor is noise-like. Both the signal's displayed power and the noise floor displayed power vary by 10log(RBW) as shown in **Figure 6-52**. SNR for the digitally modulated signal is independent of RBW setting.

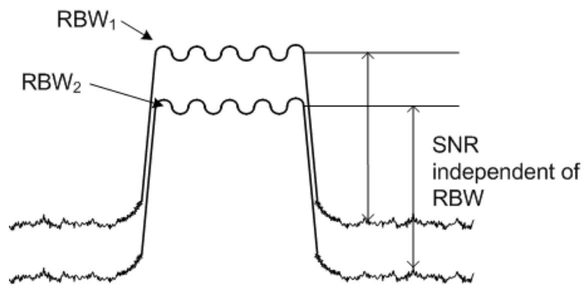

Figure 6-52. SNR for the Digitally Modulated Signal is Independent of RBW Setting.

As an example for the calculation of SNR, use the parameters of a W-CDMA signal: BW_m = 3.84 MHz and PAPR is approximately 12 dB. Also suppose the signal analyzer's maximum mixer level is -10 dBm and its displayed average noise level is -155 dBm/Hz. For a CW tone, the maximum SNR is as follows:

$$SNR_{CW,max} = Mixer\ Level_{max} - DANL$$

$$SNR_{CW,max} = -10 - (-155) = 145\ dB$$

The maximum SNR for the digitally modulated signal is as follows:

$$SNR_{dig,max} = SNR_{CW,max} - 10\log(BW_m) - PAPR$$
$$SNR_{dig,max} = 145 - 65.8 - 12 = 67.2\ dB$$

This is somewhat conservative as normally there is some headroom already built in to the system gain such that at maximum mixer level, the ADC is several dB away from being overdriven. What this means is that in many signal analyzers, the SNR of the digitally modulated signal does not need to be backed off by the full PAPR of the signal.

IMD for Digitally Modulated Signal

Normal operating mode for the signal analyzer places the maximum signal power below the signal analyzer's gain compression level. As in the case of CW signals, the signal analyzer can be considered a weakly nonlinear system. Give this constraint, the spectral re-growth stemming from intermodulation distortion for the digitally modulated signal can be approximated using the following analysis:

First divide the spectrum of the digitally modulated signal into a series of equal frequency slices as shown in **Figure 6-53**.

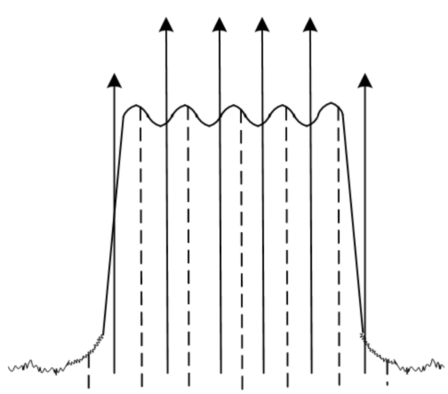

Figure 6-53. Digitally Modulated Signal Represented by CW Signal Equivalents.

For each frequency slice, compute the integrated power using as shown in **Equation 6-20.**

$$P_{slice} = 10\log\left(\sum_{i=f_a}^{f_b} 10^{X(f_i/10)}\Delta f\right)$$

where *X(f)* is the frequency domain representation of the signal, f_a and f_b are the frequency limits of the slice, and Δf is the frequency width of adjacent sample points.

Equation 6-20

Next, represent each frequency slice as a CW signal whose power level is P_{slice}. Place the CW signals frequency in the center of each associated frequency slice.

Using **Equation 6-6**, calculate the amplitude of the third order intermodulation distortion products. Use every possible combination of pairs the CW signals that represent the digitally modulated signal. The frequencies of each distortion product can be calculated by manipulating **Equation 6-4**. The distortion products will start falling at the same frequencies as other products. **Figure 6-54** shows this process.

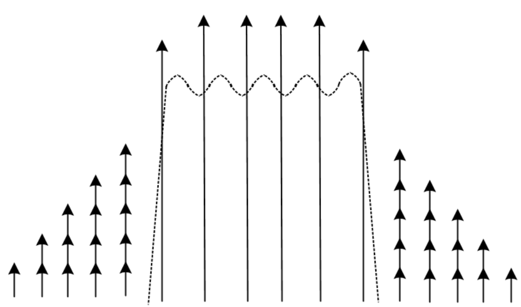

Figure 6-54. Spectral Regrowth from Two-tone IMD

Empirical evidence shows that the amplitudes of the multiple distortion products add as in-phase vectors. At each distortion component frequency the summed distortion power is calculated using **Equation 6-21.**

$$P_{IMD} = 20\log\left(\sum_{i} 10^{P_i/20}\right)$$

where P_i is the power in dBm of each individual distortion component.

Equation 6-21

For the case of W-CDMA, experimentally it has been found that the spectral re-growth power integrated over the adjacent channel can be approximated using **Equation 6-6**. The TOI value used is the signal analyzer's TOI with an experimentally determined offset. This offset is a function of PAPR, but is roughly 2 dB to 5 dB.

$$P_{IMD,adjacent} = 3P_{Mix\ Lvl} - 2(TOI_{CW} + TOI_{offset})$$

where $P_{Mix\ Lvl}$is the power of the digitally modulated signal at the signal analyzer's first mixer.

Equation 6-22

Suppose the TOI of the signal analyzer is +20 dBm, the signal power is 0 dBm, the RF input attenuator is 10 dB, and the TOI offset value is approximately 4 dB. The mixer level is -10 dBm. From **Equation 6-22**, the integrated spectral re-growth in the adjacent channel is -78 dBm, giving a relative value of -68 dB.

Phase Noise for Digitally Modulated Signal

Phase noise cannot be ignored for the measurement of adjacent channel spectral re-growth. **Figure 6-55** shows the process of determining the phase noise contribution of the digitally modulated signal in the signal analyzer.

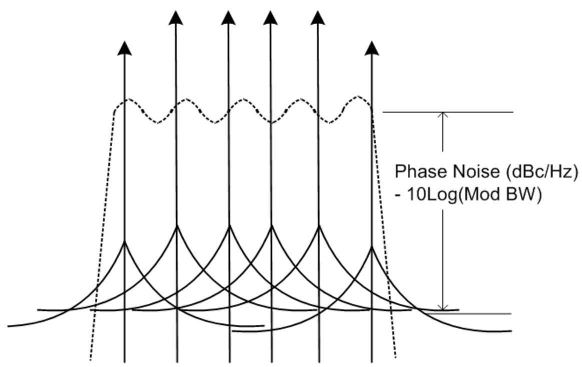

Figure 6-55. Signal Analyzer Phase Noise for the Digitally Modulated Signal.

Using the model of the digitally modulated signal, which uses CW signal representation, superimpose the signal analyzer's phase noise skirt onto each of the CW signals. At a single frequency in the adjacent channel, add the phase noise powers using **Equation 6-22**.

$$P_{Phase\ Noise} = 10\log\left(\sum_i 10^{P_i/10}\right)$$

where $P_{Phase\ Noise}$ is the summed phase noise power in a 1 Hz bandwidth and $P_i$is the phase noise power in dBm/Hz associated with each CW signal.

Equation 6-22

Quite often the phase noise pedestal of the signal analyzer is flat with frequency for offset frequencies beyond 1 MHz. For most wide bandwidth (> 1 MHz) modulation formats, this wider offset phase noise is the main contributor. Given a flat phase noise pedestal, the phase noise contribution to spectral re-growth is approximately the signal analyzer's far offset phase noise in dBc/Hz – 10log(BW_m).

Suppose the far offset phase noise of the signal analyzer is -150 dBc/Hz and, using W-CDMA as an example, the modulation bandwidth is 3.84 MHz. The relative spectral re-growth due to phase noise is roughly -84 dB.

ACLR using Vector Signal Analysis Mode

When using vector signal analysis mode, the mixer level and reference level need to be the same between the measurement of the main channel power and the adjacent channel power. The techniques outlined so far can be used for optimizing the RF input attenuator and reference level for best ACLR performance.

The dynamic range chart in **Figure 6-56** shows ACLR performance when the signal analyzer is configured for vector signal analysis mode.

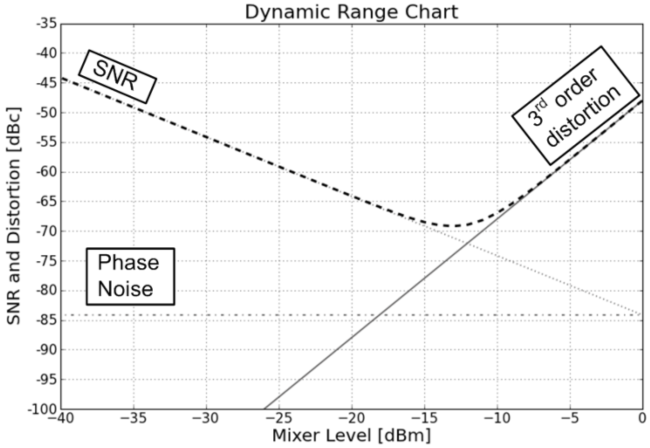

Figure 6-56. ACLR VSA Mode

In this example, the mixer level that yields best ACLR is around -13 dBm, yielding W-CDMA ACLR performance of around 70 dB.

ACLR using Spectrum Analysis Mode

In spectrum analysis mode, the IF bandwidth is set narrower than the signal's modulation bandwidth. During the measurement of the adjacent channel, the main channel signal is greatly attenuated before reaching the final IF amplifier stages and the ADC. This allows the reference level to be set lower, corresponding to higher IF gain. Not only is the distortion performance improved, but the overall SNR is better. The result is a 5 dB or more improvement in ACLR performance. **Figure 6-57** shows the dynamic range chart for W-CDMA ACLR when narrowband IF filters are used.

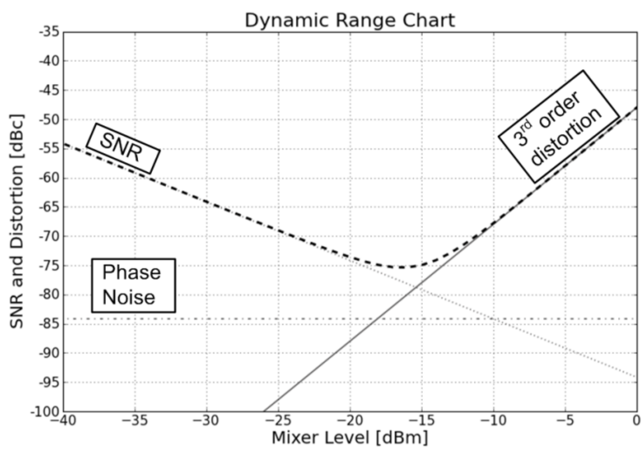

Figure 6-57. ACLR Spectrum Analysis Mode

Notice that the phase noise is no longer invisible to the measurement. The phase noise falls 5 dB below the intersection of the SNR and IMD curves. Phase noise in this example contributes approximately one dB to the ACLR degradation.

Summarizing the Dynamic Range Information

Generating these dynamic range graphs requires knowledge of the internal workings of the signal analyzer. At the very least it may require measurement of the signal analyzer's DANL and TOI beyond what is posted in the specification sheets. This section has mostly shown that optimizing the signal analyzer is not a trivial task. To extract the best dynamic range out of the signal analyzer, some work is involved. Simply letting the signal analyzer stay in its default attenuator and gain setting will result in only modest performance.

The major points of this dynamic range section are as follows:

- Mixer level is the most important item for achieving best dynamic range
- Wide bandwidth vector signal analysis mode dynamic range is limited by the ADC SFDR
- When in vector signal analysis mode, the reference level must be adjusted to the test signals power level. This sets the signal at the ADC to full scale, which results in the best SNR and SFDR performance.
- Narrow bandwidth spectrum analysis mode allows dynamic range beyond that of the ADC
- Re-ranging the reference level when measuring the distortion amplitude allows the best dynamic range
- ACLR measurements for digitally modulated signals follow the same techniques for optimizing dynamic range as for the measurement of CW signal IMD.

Follow the key concepts of adjusting mixer level using either the RF input attenuator or by adjusting the test signal amplitude. Then set the reference level to the power of the test signal. Those are guidelines; the rest is trial and error.

Attenuator Test

In the course of optimizing the signal analyzer for dynamic range, the attenuator test can be used as a check to determine whether the measured distortion is due to the test signal or due to the signal analyzer. When viewing the distortion amplitude, toggle the RF input attenuator setting. If the measured amplitude of the distortion amplitude does not change, then there is a high assurance that the distortion is entirely due to the test signal. However, if the distortion amplitude changes, then the signal analyzer is contributing to the measurement.

7. Amplitude Accuracy

In addition to possessing high dynamic range for the measurement of distortion, the signal analyzer also serves to accurately measure the signal amplitude. Attention now focuses on the uncertainty associated with the measurement of signal power.

Absolute power measurement refers to the single measurement of signal power. This entails recording the signal amplitude reported by the signal analyzer. Uncertainty is traceable back to some power standard, most commonly the National Institute of Standards and Technology (NIST).

Relative power measurement refers to two or more power measurements. For example, the harmonic measurement is a good example of a relative power measurement. Many of the absolute amplitude measurement uncertainty terms are common between the measurements and do not contribute to the relative measurement uncertainty. However, there are multiple measurements, so some of the uncertainties add up.

Absolute Amplitude Accuracy at the Calibration Reference Frequency

Many signal analyzers break up frequency response into two parts: one is sometimes termed *absolute amplitude accuracy at the reference frequency* and the other is termed the *frequency response*. Signal analyzers whose amplitude uncertainty is specified in this manner usually have an internal calibration signal. At the single point calibration frequency, an amplitude uncertainty is given.

Frequency Response

Often the frequency response is relative to the amplitude at the calibration reference frequency. **Figure 7-1** shows how frequency response is related to the absolute amplitude accuracy at the calibration reference frequency.

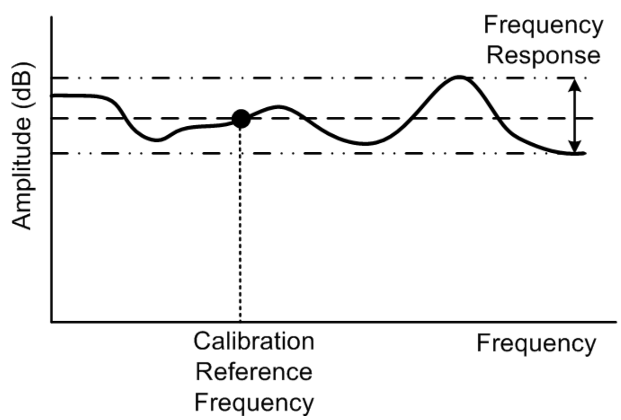

Figure 7-1. Frequency Response Relative to the Calibration Reference.

For single point power measurements at any given frequency, the frequency response uncertainty must be added to the absolute amplitude uncertainty. For relative power measurements involving the power ratio of multiple measurements at different frequencies, only the frequency response uncertainty applies.

IF Amplitude Response

Frequency response is measured at the center of the IF passband. Think of the IF passband being superimposed on the frequency response curve as suggested in **Figure 7-2**.

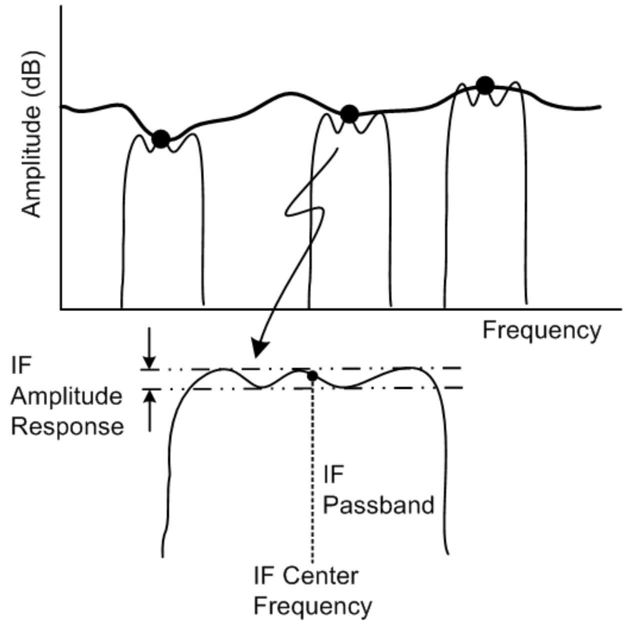

Figure 7-2. IF Amplitude Response

For signal analyzers that use FFT processing to create a spectrum display, the spectrum consists of multiple FFT segments cascaded together as shown in **Figure 7-3**.

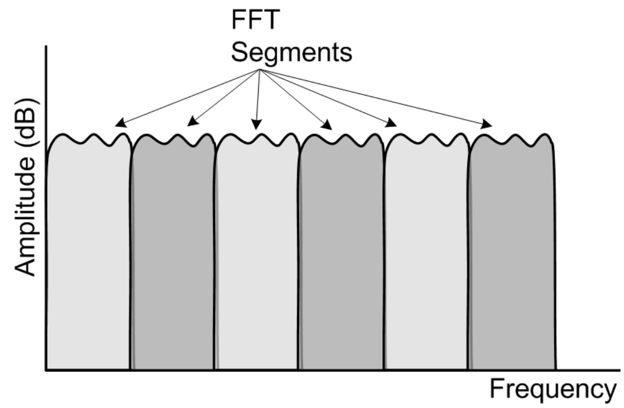

Figure 7-3. Spectrum Created from Multiple FFT Segments

When the measurement frequency is offset from the center of a particular FFT segment, the *IF amplitude response uncertainty* must be included. IF amplitude response is relative to the center of the IF passband as shown in **Figure 7-2**.

To eliminate the uncertainty associated with the IF amplitude response, set the frequency span to be narrow enough to ensure only one FFT segment is being displayed. This usually means that the span is narrower than the analog IF bandwidth. Then, adjust the center frequency setting such that the signal of interest is in the center of the display.

Signal Processing Amplitude Uncertainties

For signal analyzers that use FFT processing for displaying the spectrum, the scalloping loss mentioned in section 5 and summarized in **Table 5-1** must be added to the uncertainty. The amount of scalloping loss depends on the windowing function selected.

RF Input Attenuator Switching Uncertainty

The specification sheet for the signal analyzer may or may not include information for RF input attenuator switching uncertainty. In either case, the absolute amplitude accuracy at the calibration reference frequency and the frequency response are qualified for a particular RF input attenuation value. For the single frequency absolute amplitude measurement, it is best to use the RF input attenuation specified. However, for relative amplitude measurements, this uncertainty does not occur if the same RF input attenuator setting is used for all measurements.

The absolute measurement may need a different RF input attenuation value than what is used for the amplitude accuracy specifications. Likewise, the RF input attenuation value may need changing between measurements. If this is the case and the user wants to eliminate this error, then an RF substitution technique can be used. At the measurement frequency, set the RF input attenuator to value 1 and make an amplitude measurement. Then switch to RF input attenuator value 2 and measure the amplitude. The difference in amplitudes between these two measurements is the RF input attenuator switching error. The error can be subtracted from all subsequent measurements.

YIG Tuned Filter Amplitude Response

As mentioned in Section 3 of this document, the highband path in the signal analyzer may include a YIG tuned filter (YTF) as a preselector. Unfortunately, the frequency response when selecting the YTF can be compromised.

The YTF is an open loop device meaning that there is no feedback mechanism to ensure that the filter center frequency is precisely deterministic. When the signal analyzer is tuned to a particular highband frequency, the YTF more than likely will not be centered where it was during instrument calibration. Furthermore, within a measurement session the YTF’s center frequency is not repeatable. Tuning the signal analyzer from frequency A to frequency B and then back to frequency A again most likely results in a YTF center frequency offset.

The lack of YTF repeatability stems from several mechanisms. One is tuning hysteresis. The YTF center frequency is controlled by an electro-magnet which has its associated hysteresis: the magnetic flux strength at a particular current value differs depending on whether the current is being ramped up or being ramped down. The different magnetic flux translates to a YTF center frequency offset depending on which way the YTF is being tuned. Take for example the signal analyzer tuned to 10 GHz. If the previous center frequency was 5 GHz the YTF center frequency will be at one frequency. If the previous center frequency was at 15 GHz, then the YTF center frequency may be different. Efforts are made to reduce the hysteresis using a de-hysteresis pulse between each center frequency change.

Post tune drift is another repeatability mechanism. A fair amount of heat is generated in the electro-magnet in the YTF. The heat causes center frequency drift (the YTF metal housing expands with heat and this tends to concentrate the magnetic flux). The higher the tune frequency, the more magnet current and hence the more heat generated. In addition to the YTF center frequency drifting over time at elevated frequencies, the heat is retained when tuning to lower frequencies. At these lower frequencies, the YTF cooling off will cause center frequency drift.

Figure 7-4 shows the consequence of YTF center frequency drift and repeatability.

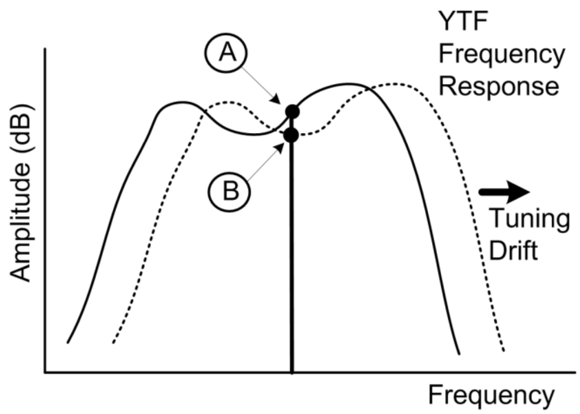

Figure 7-4. Frequency Response of the YTF

The frequency response over the bandwidth of the YTF is not flat. As the frequency changes due to either drift or repeatability, the measured amplitude will fluctuate.

The signal analyzer frequency response specifications reflect the extra amplitude uncertainty due to the YTF – there is usually a clear breakpoint in frequency response between lowband path and highband path.

Many signal analyzers have a path to bypass the YTF. **Figure 3-4** shows such a structure. Bypassing the YTF can lead to undesired spurious performance issues. However an RF substitution measurement can be used to reduce the YTF frequency response uncertainty. First use the YTF path to locate and measure the test signal's amplitude. With the YTF selected, the user can be assured that the signal displayed is the true signal and not an image. Next, select the YTF bypass path and measure the test signal's amplitude. When the YTF path is once again selected, the amplitude difference between the YTF and the YTF bypass path can be subtracted from subsequent measurements.

Near Noise Amplitude Error

Noise voltage is random in both amplitude and phase. Vector addition of the random noise voltage and a CW signal results in a combined amplitude that is still random. The method to combine amplitudes is through a power addition as shown in **Equation 7-1**.

$$P_{Signal+Noise} = 10\log\left(10^{P_{signal}/10} + 10^{P_{noise}/10}\right)$$

Equation 7-1

As demonstrated in **Figure 6-35**, as the signal amplitude approaches the noise floor of the signal analyzer, the signal power and the noise powers add. With sufficient trace averaging, this offset can be treated as a deterministic error, not an uncertainty. One can measure the noise floor adjacent to the signal, or if possible, by removing the signal. Then measure the displayed signal plus noise amplitude. The true signal power then can be computed.

Figure 7-5 shows the terms used in the measurement of near noise signal amplitudes.

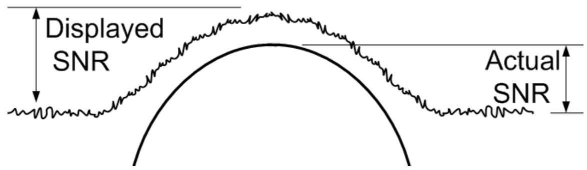

Figure 7-5. Near Noise Displayed SNR and Actual SNR.

Use **Equation 7-1** to compute the *actual SNR* from the measurement of the *displayed SNR* as shown in **Equation 7-2**.

$$Actual\ SNR = 10\log\left(10^{Displayed\ SNR/10} - 1\right)$$

Equation 7-2

The chart in **Figure 7-6** shows the curve generated by **Equation 7-2**.

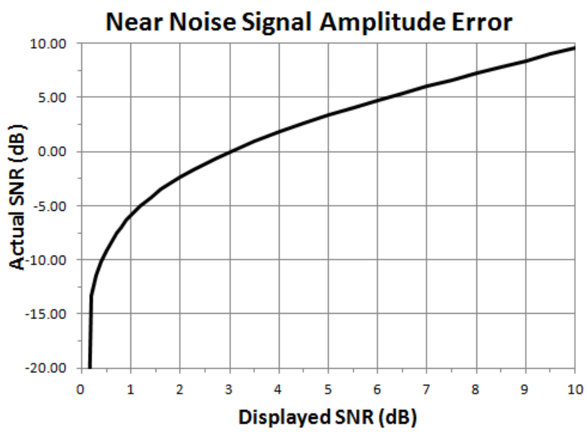

Figure 7-6. Signal + Noise versus Signal Level Relative to the Noise Floor.

Suppose the displayed SNR is 1.0 dB, then the actual SNR is -5.9 dB. This result implies that the signal is really almost 6 dB *below* the noise floor even though the apparent SNR is 1 dB. With knowledge of the noise floor power level and the actual SNR, the signal amplitude is readily apparent.

Of course, accounting for this error is time consuming. If the SNR is kept greater than 20 dB, then the error between the displayed SNR and the actual SNR is less than 0.05 dB.

Coherent Signal Addition

Coherent signal addition occurs when two or more non-random signals whose frequencies are the same combine. In the signal analyzer, this is most common in the measurement of distortion. The distortion component from the DUT can fall at the same frequency as the distortion from the signal analyzer. These two signals combine as voltages using vector addition. **Figure 7-7** shows that the resultant is the result of vector addition of the two signals.

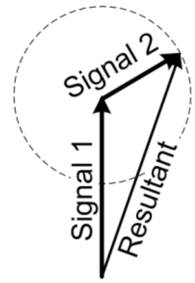

Figure 7-7. Coherent Addition of Two Signals

When the relative phases are the same, the two signals add constructively. At the other extreme, the signals with relative phases 180 degrees apart add destructively. The amplitude uncertainty due to two signals combining destructively is given by **Equation 7-3.**

$$Uncert = 20\log\left(1 \pm 10^{-\Delta/20}\right) \ [dB]$$

where Δ is the relative amplitudes of the two signals in dB

Equation 7-3

Equation 7-3 is plotted in **Figure 7-8**.

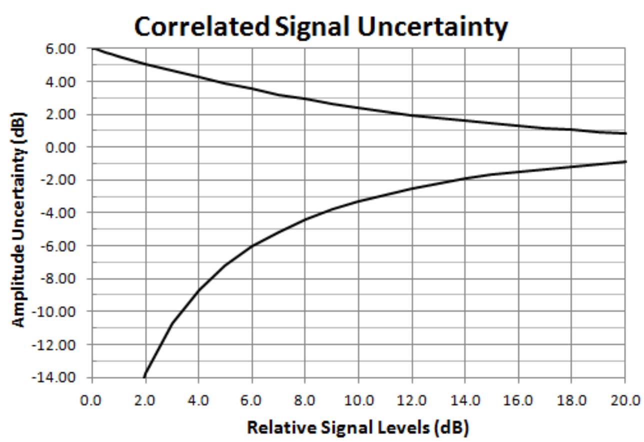

Figure 7-8. Amplitude Uncertainty versus Relative Signal Level for Coherent Signal Addition.

For example, if two signals at the same frequency have the same amplitude, the resulting combined signal could either range between 6 dB addition or completely cancel. For an uncertainty of less than ±1 dB, the relative signal amplitudes need to be at least 19 dB apart.

Figure 7-9 shows that for the dynamic range chart, when considering amplitude uncertainty, the range of available mixer levels is limited.

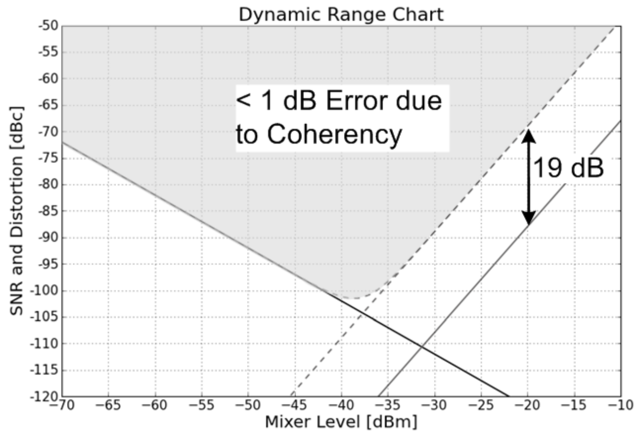

Figure 7-9. Dynamic Range Chart Show Range of Measurements for Uncertainty Due to Coherent Signal Addition is Less than 1 dB.

Mismatch Uncertainty

Signal analyzers can operate at high enough frequencies where only the average power is relevant. At very low frequencies, the phase of the voltage signal does not appreciably change as a function of cable length between the DUT and the signal analyzer. However, as the frequency increases, the assumption of constant phase does not hold true, in which case the amplitude of the voltage at the signal analyzer input does vary even for minute changes in cable length.

Average power, on the other hand, stays independent of cable length. Average power flow into the signal analyzer can be depicted as shown in **Figure 7-10**.

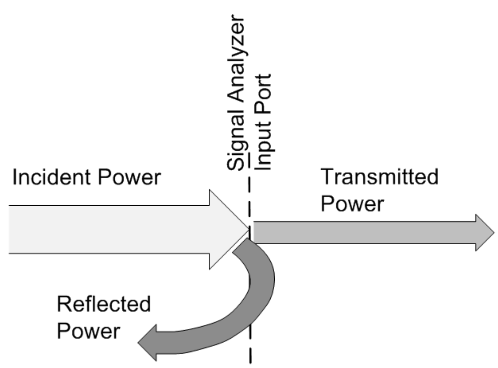

Figure 7-10. Power Flow into the Signal Analyzer.

Power flow is incident on the input of the signal analyzer. A portion of the signal is transmitted thorough to the internal path of the signal analyzer while the other portion gets reflected back to the DUT. Only the transmitted power is measured by the signal analyzer.

The term given to the ratio of the transmitted power to the incident power is *match loss*. Match loss is accounted for during the calibration process. However, what cannot be accounted for during the calibrations process is the re-reflected signal. **Figure 7-11** shows the flow graph of the re-reflected signal.

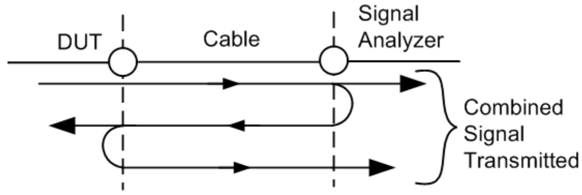

Figure 7-11. Reflected Signal Re-reflected Back to the Signal Analyzer.

The reflected signal exiting the signal analyzer's RF input port travels back to the DUT. At the DUT output, a portion of this signal re-reflects and travels once again to the signal analyzer. At the signal analyzer, the original transmitted signal and the re-reflected signals combine. This signal combination is a coherent addition of two signals. Unfortunately, the phases of the two signals are unknown, which creates an uncertainty. This uncertainty is termed *mismatch uncertainty*.

There are a few related terms used to qualify the input impedance of the signal analyzer. The majority of signal analyzers have a nominal impedance of 50 ohms. To calculate mismatch uncertainty, the *reflection coefficient* of both the signal analyzer input and DUT output must be known. *Voltage standing wave ratio* (VSWR) is normally the characteristic used to describe the signal analyzer's impedance. Most often *return loss* in dB is used to describe the DUT's output impedance. Use the following equations to arrive at reflection coefficients:

$$|\rho| = \frac{VSWR - 1}{VSWR + 1}$$

where ρ is the reflection coefficient

Equation 7-4

$$|\rho| = 10^{Ret\ Loss/20}$$

Equation 7-4a

$$Return\ Loss = 20\log(\rho)$$

Equation 7-4b

Once the signal analyzer and DUT terms are converted to reflection coefficients, the bounds of the mismatch uncertainty are shown in **Equation 7-6**.

$$Uncert = 20\log\left(1 \pm |\rho_{signal\ analyzer}||\rho_{DUT}|\right)$$

Equation 7-6

Figure 7-12 shows the relationship between VSRW and return loss.

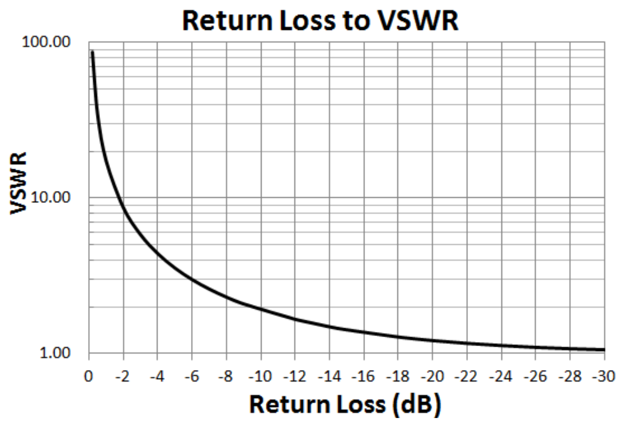

Figure 7-12. VSWR to Return Loss

As an example, suppose the signal analyzer has a VSWR specification of 1.5:1 and the DUT output return loss is -12 dB. The reflection coefficient for the signal analyzer is as follows:

$$p_{signal\ analyzer} = \frac{1.5 - 1}{1.5 + 1} = 0.2$$

The reflection coefficient for the DUT output is as follows:

$$\rho_{DUT} = 10^{-12/20} = 0.2512$$

The mismatch uncertainty is then the following:

$$Uncert = 20\log\big(1 \pm (0.2)(0.2512)\big) = +0.43,\ -0.45\ \text{dB}$$

Figure 7-13 shows a graph of the mismatch uncertainty using signal analyzer and DUT return losses as parameters.

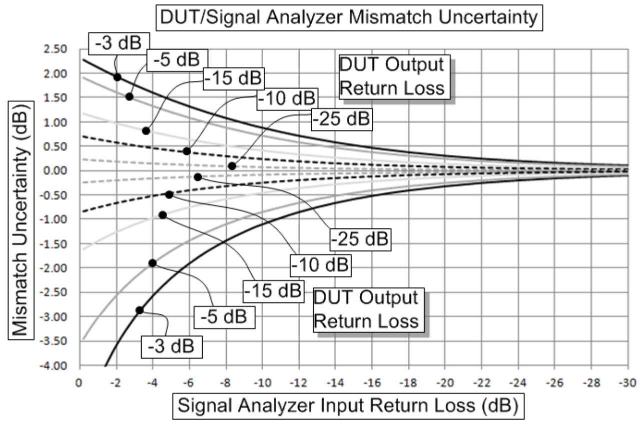

Figure 7-13. Mismatch Uncertainty. X-axis is Signal Analyzer Return Loss, Labeled Curves are the DUT Return Losses.

Most signal analyzer VSWR performance is a function of both frequency and RF input attenuator setting. In general, avoid using the lower value RF input attenuator settings for best VSWR performance.

Worst Case and Root Sum Squared Uncertainty

Based on the above descriptions, most uncertainty terms associated with amplitude measurements can be treated as statistically independent. When one uncertainty term does not depended on another uncertainty term, this independence is valid.

Worst case analysis, in which all the error terms are added, yields a very conservative overall measurement uncertainty. Empirical measurements using signal analyzers do not support the worst case error analysis.

A widely accepted method of combining uncertainty terms is the root-sum-of-squares (RSS) technique [[17]]. The RSS algorithm is described in **Equation 7-7**.

$$RSS = \sqrt{u_1^2 + u_2^2 + u_3^2 + \cdots}$$

where u_x are the individual uncertainty terms

Equation 7-7

The RSS technique gives a 95% confidence (2σ) that all measurements will have a combined uncertainty within the limits predicted by the equation.

Nearly all uncertainties in signal analyzer specification sheets have logarithmic units expressed in dB. One question is whether the uncertainty terms need to be converted to linear units (Watts) before applying the RSS equation. The log curve is nonlinear which may bias the mean of the individual error terms. **Figure 7-14** shows that for small value uncertainty terms, the slope of the linear power to log power curve is approximately linear. In these cases, the RSS equation, shown in **Equation** 7-7, can be used with terms expressed in dB.

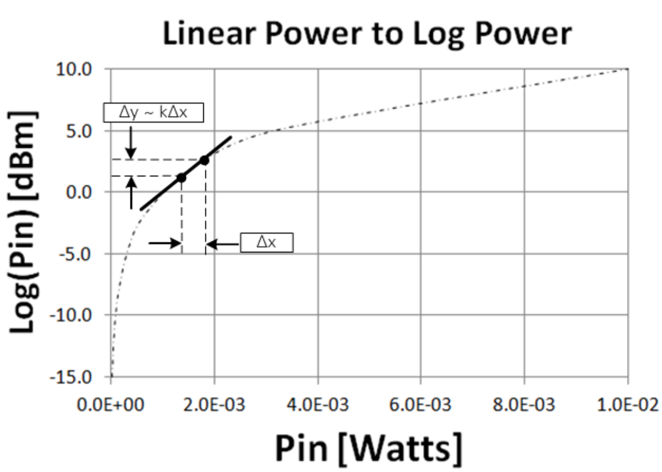

Figure 7-14. Small Value Amplitude Changes on the Log Scale are Approximately Linear.

To compute using linear terms, use the following equations:

Start with all uncertainties expressed in dB. Call these terms Δ_x. Then compute the linear term using the following equations:

$$u_x = 10^{\Delta x/10} - 1$$

Equation 7-8

Compute linear RSS as follows:

$$RSS_{lin} = \sqrt{\sum u_x^2}$$

Equation 7-9

Then convert to log as follows:

$$RSS_{log} = 10\log(1 + RSS_{lin})$$

Equation 7-10

Table 7-1 summarizes an experiment where six uncertainty terms, each of which has the same uncertainty value, are combined using worst case, RSS log and RSS linear.

Table 7-1. Worst Case vs. RSS Uncertainty Analysis

Individual Uncert.	Worst Case	RSS (log)	RSS (lin)	RSS Difference	Units
1 dB x 6	6	2.45	2.13	0.316	dB
0.5 dB x 6	3	1.22	1.14	0.089	dB
0.25 x 6	1.5	0.61	0.59	0.024	dB
0.1 x 6	0.6	0.24	0.24	0.004	dB

RSS Difference column is the dB error between RSS log where the uncertainties in dB are computed using **Equation 7-7** directly and RSS linear where the individual uncertainty terms are first converted to linear using **Equation 7-8**, **Equation 7-9**, and **Equation 7-10**.

One can see how conservative the worst case uncertainty analysis is compared to RSS analysis. Use to determine whether RSS log or RSS linear is appropriate. If all uncertainty terms are less than 0.5 dB, then the difference between RSS linear and RSS log is less than 0.1 dB. RSS linear is the more accurate method.

Power Meter Leveling

From an amplitude uncertainty perspective, the definition of an ideal source is one whose source impedance is exactly equal to the characteristic impedance of the system (normally 50 ohms or 75 ohms for CATV applications). Why is perfect source output impedance so important? Recall from **Figure 7-11**, that mismatch error occurs due to the re-reflected portion of the incident signal combining with the signal transmitted to the signal analyzer. If the re-reflected signal disappears, so does the mismatch error. **Figure 7-15** shows the block diagram of the system with an ideal (from an RF sense) source.

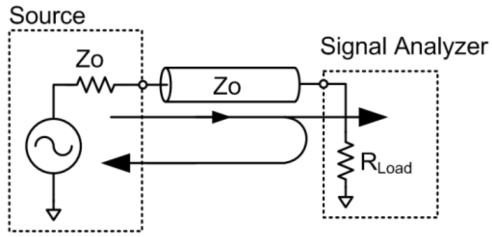

Figure 7-15. Ideal Source: Impedance Equals Characteristic Impedance of the System.

Comparing the signal flow of **Figure 7-15** with that of **Figure 7-11**, the re-reflected signal does not exist when the source match equals the system's characteristic impedance, Z_o. According to **Equation 7-6**, mismatch uncertainty drops to 0 dB if reflection coefficient, ρ_{DUT}, is zero. ρ_{DUT} in this example is the output impedance of the source. Reflection coefficient can be computed from **Equation 7-11**.

$$\rho = \frac{R - Z_0}{R + Z_o}$$

Equation 7-11

Let R be the output impedance of the source. When R equals characteristic impedance Z_o, the reflection coefficient is zero.

One method of creating an effectively perfect source match is to level the source as shown in **Figure 7-16**.

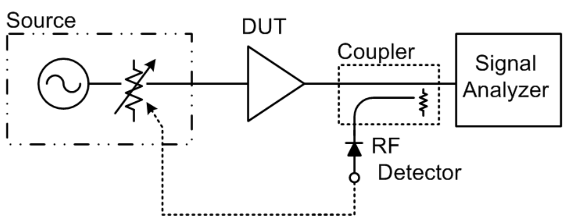

Figure 7-16. Source Leveling for Constant Impedance Source Impedance Using RF Detector.

A directional coupler siphons off a minute portion of the incident signal as it enters the RF input port of the signal analyzer. An envelope detector converts the coupled amplitude to a video signal that is then fed back to the source's amplitude control. The source creates a compensation amplitude component that, when combined with the signal analyzer's incident signal, results in a signal with flat frequency response as it enters the signal analyzer. **Figure 7-17** depicts the leveling technique.

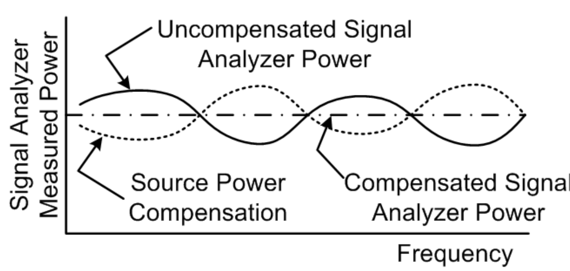

Figure 7-17. RF Detector Allows Source Power to Compensate for Signal Analyzer Mismatch.

A flat frequency response of the incident signal created using leveling effectively reproduces the ideal source that has an output impedance equal to the system characteristic impedance.

However, the leveling technique suffers from a few challenges. One is that the source may not have the ability to have its amplitude externally modulated. If the source is an over-the-air signal, this clearly cannot be amplitude controlled in this fashion. The coupler's frequency range can be limited, especially on the low end. The coupler introduces a few RF impairments, such as output match and isolation between coupler output port and the coupled arm. Finally, the envelope detector response is not perfectly flat with frequency and also not perfectly linear with power level.

The *power meter ratio* technique overcomes some of the deficiencies of the source leveling technique. This technique uses simultaneous measurements with the signal analyzer and a power meter. The power meter's inherently better amplitude accuracy can then be transferred to the signal analyzer measurement. The result, if proper care is given to the measurement technique, can yield effectively flat incident power into the signal analyzer. Flat incident power equates to no mismatch uncertainty term. However, if not properly implemented, the power meter ratio technique can introduce worse mismatch error than if nothing had been done.

Power splitters come in a few different topologies. Often times the labeling is not clear on the type of power splitter, so one must use the specifications as clues. The Wilkinson power splitter is lossless in the sense that no power (ideally) is dissipated within the power splitter. The input power is split evenly between the two output ports. Each port power is approximately 3 dB lower in power than the power at the input. The isolation between output ports can be quite large, exceeding 25 dB.

Figure 7-18 shows a setup with the Wilkinson power splitter driving a power meter with one output port and the signal analyzer with the other. The source is not limited to a true source; in general this is a DUT.

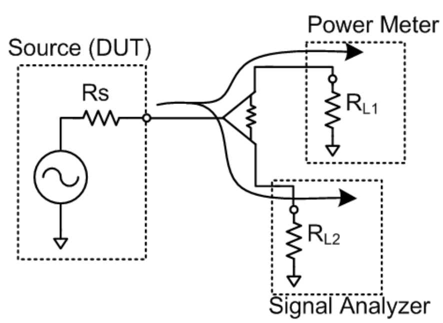

Figure 7-18. Wilkinson Power Splitter with Power Meter.

A perfectly reasonable assumption is that the DUT power evenly splits between the two output ports of the power splitter. The false assumption is that the power meter and the signal analyzer both measure the same power level. **Figure 7-19** shows measurement results with signal analyzer input return loss equal to -10 dB, source output match equal to -16 dB and power meter return loss equal to -25 dB. The power splitter is nearly ideal: 3 dB insertion loss, -30 dB return loss on all ports, and 25 dB isolation between output ports.

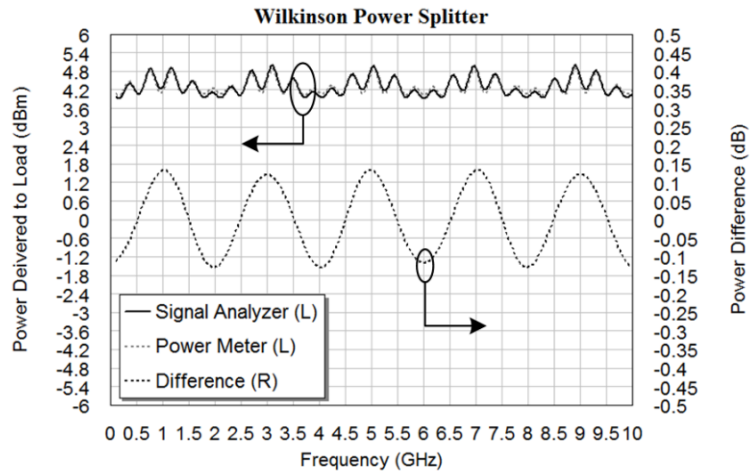

Figure 7-19. Power Meter Ratio Technique Using Wilkinson Power Splitter.

The *difference* signal represents the ratio of the power meter to signal analyzer. This difference signal is the amplitude error in the measurement that occurs when the power meter is used to determine the incident power into the signal analyzer. The uncertainty for this example is 0.3 dB (± 0.15 dB).

Wilkinson power splitters do not work at lower frequencies. They are often limited to operate over one or two octaves of frequency. To overcome the frequency range limitations, a resistive splitter is used. There are two flavors of resistive splitter: 3-resistor and 2-resistor. For both resistive splitters, power is dissipated within the splitter, which is the cost of broad frequency coverage.

Figure 7-20 shows the power meter ratio measurement using the 3-resistor power splitter.

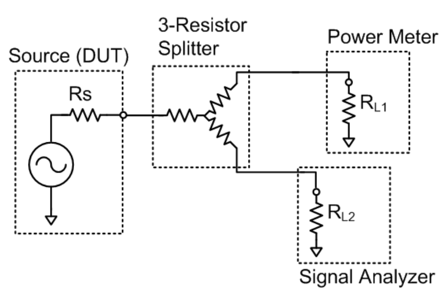

Figure 7-20. 3-Resistor Splitter.

The resistor values are Zo/3 or 16.67 Ohms each for a 50 Ohm characteristic impedance system. The insertion loss is roughly 6 dB for each port. The output port isolation is only 6 dB, which is a clear differentiation from the Wilkinson power splitter.

Figure 7-21 shows the results of using the power meter ratio technique with the 3-resistor power splitter. The source, signal analyzer, power meter load, and source matches are the same as used in the Wilkinson splitter example.

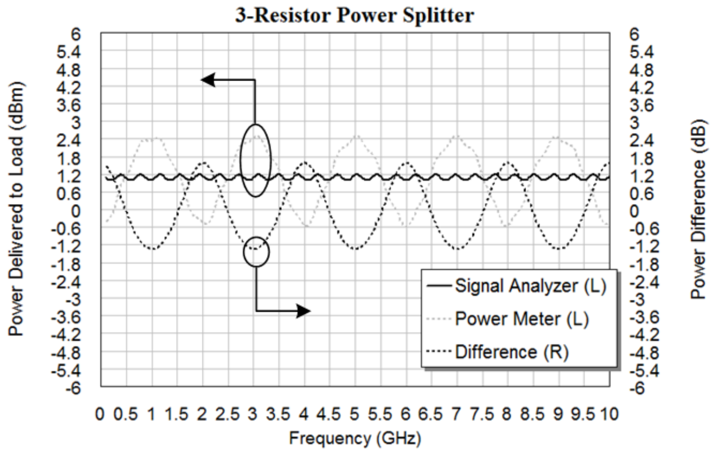

Figure 7-21. Power Meter Ratio Technique Using 3-Resistor Power Splitter.

Here the difference between the power meter and the signal analyzer is nearly 3 dB, clearly an abysmal result.
The 2-resistor power splitter has equal value resistors on the output arms, each with value equal to the characteristic impedance (50 Ohm most likely).

Figure 7-22 shows the connection for the power meter ratio technique using the 2-resistor power splitter.

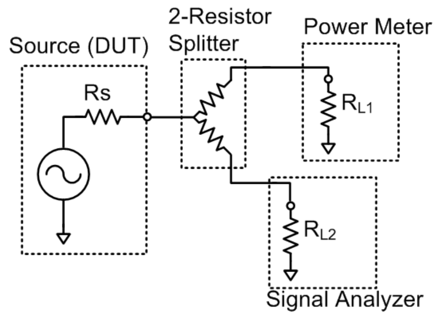

Figure 7-22. 2-Resistor Power Splitter.

Using the same load and source match values as in the previous examples, the measurement results for the 2-resistor power meter ratio technique using the 2-resistor power splitter are shown in **Figure 7-23**.

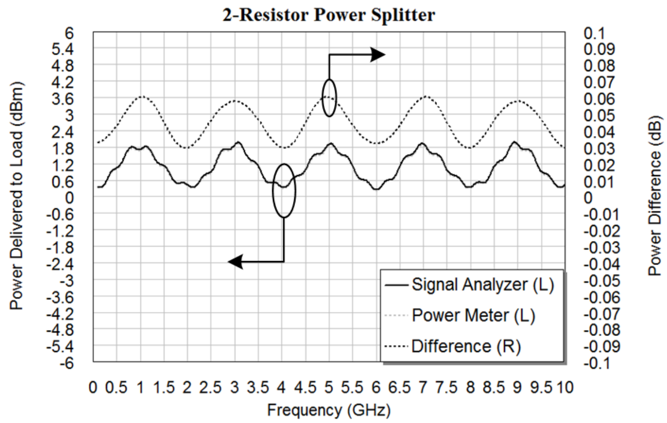

Figure 7-23. Power Meter Ratio Technique Using 2-Resistor Power Splitter.

Now the power meter and signal analyzer power measurements track to within ± 0.015 dB. Reference [[18]] explains the theory behind both the 3-resistor and 2-resistor power meter ratio technique. Some of the magic with the 2-resistor power splitter when used in this fashion is that the reflected signal from the signal analyzer appears at the power meter port. The power meter then can automatically account for this reflected power. By using the ratio of the power meter and the signal analyzer powers (remember, ratio when using amplitudes expressed on a log scale is subtraction) the result is flat incident power to the signal analyzer.

The methods to reduce the signal analyzer's mismatch uncertainty for amplitude measurements are summarized as follows:

- Wilkinson power splitter is marginal. Use it with care.
- 3-resistor power splitter is unacceptable. Avoid using it.
- 2-resistor power splitter is the best solution

The Wilkinson and 3-resistor power splitters have their place in this world, just not for the use of trying to drive down signal analyzer mismatch uncertainty.

8. Spectrum Monitoring Applications

Spectrum monitoring entails over-the-air measurements for the detection and surveillance of RF signals. The signal analyzer seems a natural fit for this application where not only does a magnitude versus frequency spectrum become useful for determining whether or not a signal of interest exists, but once detected, demodulation may be required to decode the signal. Probability of intercept (POI) becomes the chief metric for signal monitoring applications. Dynamic range and measurement speed feed into the POI are of merit. The dynamic range problem for signal monitoring applications differs somewhat from that of standard R&D and manufacturing test applications. This section highlights those differences and offers some considerations to enhance the dynamic range of the signal analyzer for spectrum monitoring applications.

The Spectrum Monitoring Dynamic Range Problem

For R&D and manufacturing tests, the signal analyzer is most often connected to the DUT using a shielded coaxial cable. In many cases, the signal being measured is known in frequency. This describes a relatively benign environment.

Spectrum monitoring most often involves connecting the signal analyzer to an antenna in order to receive signals over-the-air. Near pandemonium can result with a large number of unknown signals impinging on the signal analyzer. To compound matters worse, the signal of interest is often a weak signal in the presence of large amplitude interfering signals. **Figure 8-1** alludes to this situation.

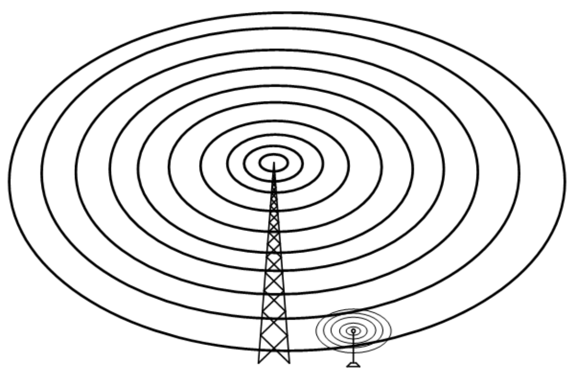

Figure 8-1. Weak Signal of Interest in Presence of Strong Interfering Signal.

In the frequency spectrum, the weak signal of interest/strong interfering signal appears as shown in **Figure 8-2**.

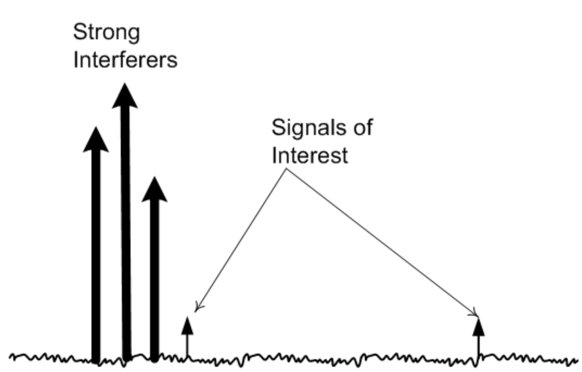

Figure 8-2. Frequency Spectrum of Low Amplitude Signals of Interest in the Presence of Large Amplitude Interferes.

Taking the spectrum depicted in **Figure 8-2** and applying it to the signal analyzer could result in the situation described in **Figure 8-3**.

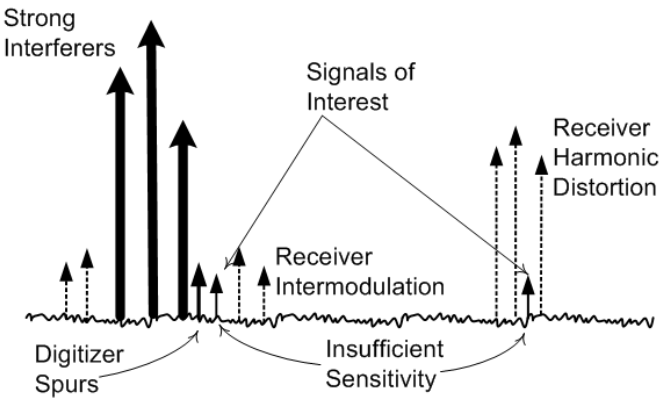

Figure 8-3. Distortion Created in the Signal Analyzer for the Spectrum Monitoring Situation.

Harmonics of the interfering signals can mask signals of interest. Intermodulation distortion created in the signal analyzer as a result of multiple interferes can mask signals of interest that fall near these IMD terms. Additionally, the signal analyzer's noise floor can bury the signal of interest in noise.

In section 6 the concept of the dynamic range chart helps to optimize the signal analyzer when measuring signals with sufficiently large amplitude. However, this is usually the wrong optimization for signal monitoring. Increasing the signal analyzer's RF input attenuation to drive down the signal analyzer's internally generated harmonic and intermodulation distortion has a severe impact on noise performance. Quite often, the SNR of the weak signal of interest cannot tolerate such a tradeoff in noise performance.

Spectrum Monitoring Receiver

As described in reference [[19]], a specialized receiver especially optimized for over-the-air spectrum monitoring may be warranted. **Figure 8-4** shows the additional blocks added to the signal analyzer that the spectrum monitoring receiver brings.

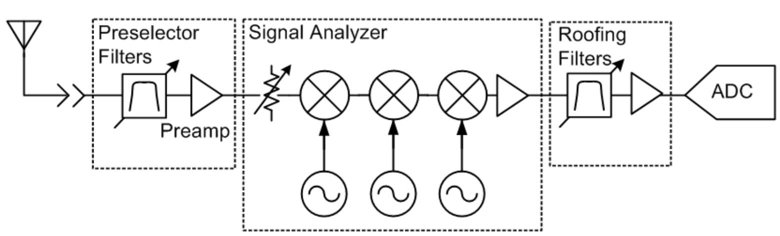

Figure 8-4. Spectrum Monitoring Receiver Block Diagram.

The primary block that the spectrum monitoring receiver adds to the signal analyzer is the preselector. The preselector can either be a tunable bandpass filter or a series switched filters placed at the RF input path. The primary function of these filters is to reduce harmonic distortion produced in the signal analyzer. And so, their bandwidths must be narrower than one octave (an octave is a doubling in frequency).

The location of the preamplifier is after the preselector filter. Recall from **Table 6-1**, a preamplifier added in front of the signal analyzer reduces dynamic range; distortion degrades faster than the improvement in noise performance. However, this is not necessarily true if the preamplifier is protected by the preselector filter. The insertion loss of the preselector filter does degrade noise figure performance as compared to the unprotected preamplifier, but the enhanced distortion performance far outweighs the noise figure degradation.

Figure 8-5 shows the harmonic distortion situation with and without preselection.

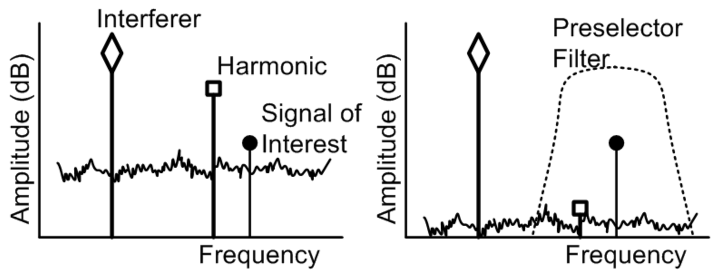

Figure 8-5. Harmonic Distortion with and without Preselection and Preamplification.

The left hand side of **Figure 8-5** shows the spectrum without preselection. The interferer generates harmonic distortion within the signal analyzer. If the harmonic distortion signal is near the signal of interest, deciphering the signal from the false distortion component becomes very difficult. POI drops appreciably if the signal analyzer's distortion gets confused for the true signal of interest.

The right hand spectrum shown in **Figure 8-5** is with the preselector included. First of all, the signal analyzer's harmonic distortion improves dramatically due to the filtering of the interfering signal. Next, with the large amplitude signals no longer impacting the signal analyzer, the preamplifier gain can be increased to improve the overall receiver's noise performance. This has the added benefit of discovering the weak signal of interest from the noise floor. Even for sufficiently large SNRs, by widening the resolution bandwidth filter setting, a faster scan of the spectrum can take place. With low system noise figure, a wider RBW setting is possible. Fast scanning also aids in higher POI.

One final component in the spectrum monitoring receiver as shown in **Figure 8-4** is the set of so called *roofing filters*. These are no more than a richer set of analog IF filters. For multiple interferers spaced too closely in frequency to be filtered adequately by the preselector filters, the roofing filters help protect the back end analog blocks and the digitizer from generating intermodulation distortion. These roofing filters behave according to the discussion related to **Figure 6-42**.

A wider range in bandwidth is offered with the roofing filters over the analog IF filters in the signal analyzer. This aids in a better tradeoff between selectivity and measurement speed.

9. Summary

Whether called spectrum analyzers or vector signal analyzers, signal analyzers that employ a super-heterodyne architecture have been described in this application note. Much of the discussion centered on the internal architecture of this type of signal analyzer. The motivation behind this application note is that a fuller understanding of how the signal analyzer works under the hood allows the user to optimize the measurement for the unique needs of a particular measurement.

References

[1] S. Haykin, Communication Systems, 2nd ed., 1983.

[2] http://mathworld.wolfram.com/ FourierSeriesSquareWave.html

[3] R. Bracewell, The Fourier Transform and Its Applications, 2nd ed., 1978.

[4] A. Douglas, “Who Invented the Superheterodyne?”, http://www.antiqueradios.com/superhet/

[5] J. Helszian, YIG Resonators and Filters, 1985.

[6] W. Kester, Analog-Digital Conversion, Analog Devices, 2004.

[7] R. Lyons, Understanding Digital Signal Processing, 2001

[8] G. Gonzalez, Microwave Transistor Amplifiers Analysis and Design, 1984

[9] NI PXIe-5622 Specifications, National Instruments Corporation.

[10]NI Spectral Measurements Toolkit User Guide, National Instruments Corporation.

[11] F. Harris, “On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform”, Proceedings of the IEEE, Vol. 66, No. 1, January 1978.

[12] “Emulating Video Bandwidth and Detectors in an FFT-based Vector Signal Analyzer”, National Instruments, 2012.

[13] A Moulthrop, M Muha, “Accurate Measurement of Signals Close to the Noise Floor on a Spectrum Analyzer”, IEEE Transactions on Microwave Theory and Techniques, Vol. 39, No 11, November 1991.

[14] “Spectrum Analysis, Application Note 150”, Hewlett-Packard Company, 1989.

[15] NI PXIe-5665 Specifications, National Instruments Corporation.

[16] LTC2208 Data Sheet, Linear Technology Corporation.

[17] P. Crisp, “Uncertainty Analysis for Laboratory Accreditation”, http://www.isobudgets.com/pdf/papers/05_Uncertainty_Analysis_for_Laboratory_Accreditation.pdf

[18] R. Johnson, “Understanding Microwave Power Splitters”, Microwave Journal, Dec. 1975.

[19] B. Avenell, “Maximizing Receiver Dynamic Range for Spectrum Monitoring”, Microwave Journal, Oct. 2012.