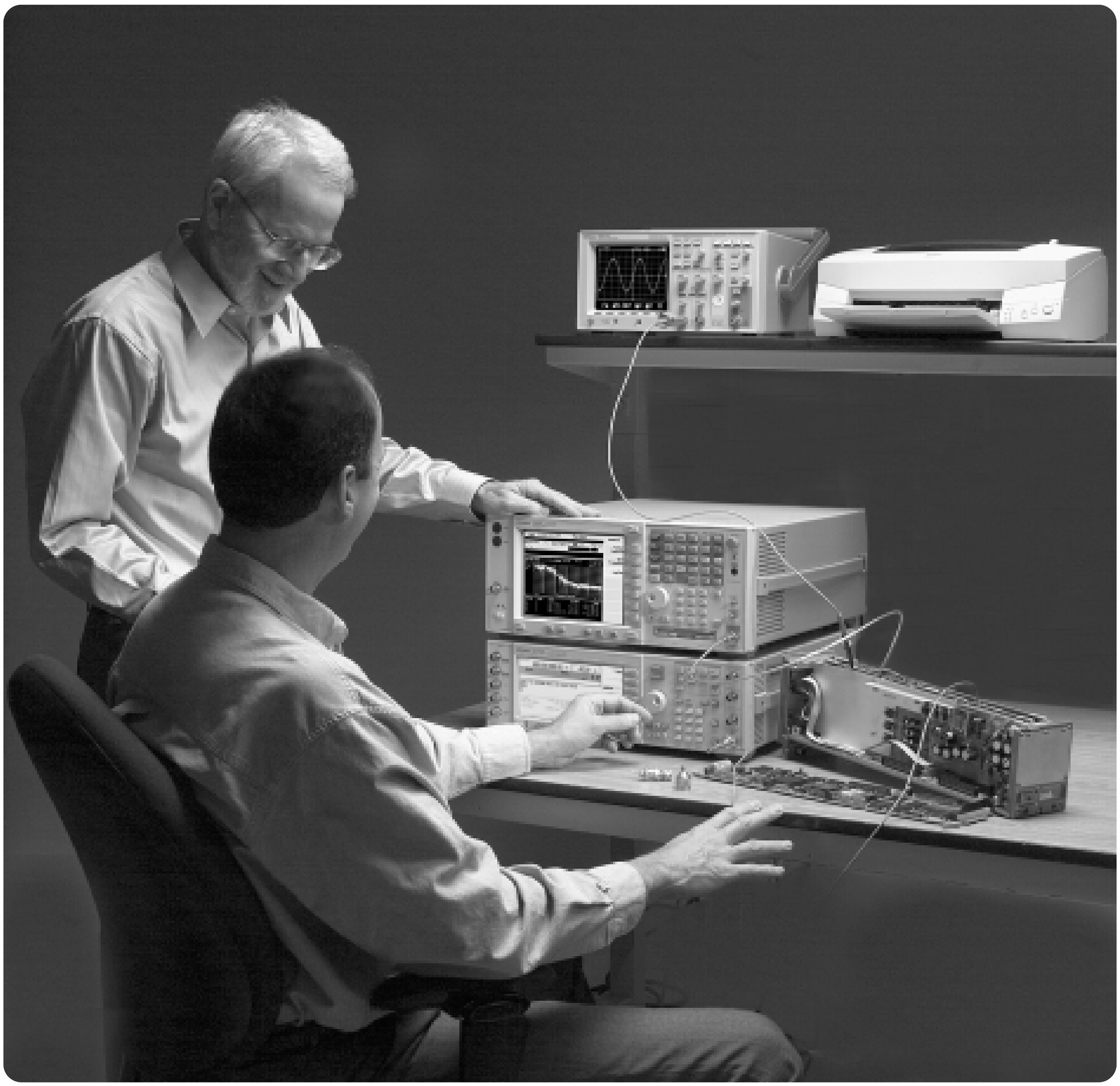
Agilent
Spectrum Analysis Basics

Application Note 150



Agilent Technologies

Table of Contents

Table of Contents — continued

Chapter 1
Introduction

This application note is intended to explain the fundamentals of swept-tuned, superheterodyne spectrum analyzers and discuss the latest advances in spectrum analyzer capabilities.

At the most basic level, the spectrum analyzer can be described as a frequency-selective, peak-responding voltmeter calibrated to display the rms value of a sine wave. It is important to understand that the spectrum analyzer is not a power meter, even though it can be used to display power directly. As long as we know some value of a sine wave (for example, peak or average) and know the resistance across which we measure this value, we can calibrate our voltmeter to indicate power. With the advent of digital technology, modern spectrum analyzers have been given many more capabilities. In this note, we shall describe the basic spectrum analyzer as well as the many additional capabilities made possible using digital technology and digital signal processing.

Frequency domain versus time domain

Before we get into the details of describing a spectrum analyzer, we might first ask ourselves: “Just what is a spectrum and why would we want to analyze it?” Our normal frame of reference is time. We note when certain events occur. This includes electrical events. We can use an oscilloscope to view the instantaneous value of a particular electrical event (or some other event converted to volts through an appropriate transducer) as a function of time. In other words, we use the oscilloscope to view the waveform of a signal in the time domain.

Fourier[1] theory tells us any time-domain electrical phenomenon is made up of one or more sine waves of appropriate frequency, amplitude, and phase. In other words, we can transform a time-domain signal into its frequency-domain equivalent. Measurements in the frequency domain tell us how much energy is present at each particular frequency. With proper filtering, a waveform such as in Figure 1-1 can be decomposed into separate sinusoidal waves, or spectral components, which we can then evaluate independently. Each sine wave is characterized by its amplitude and phase. If the signal that we wish to analyze is periodic, as in our case here, Fourier says that the constituent sine waves are separated in the frequency domain by 1/T, where T is the period of the signal[2].

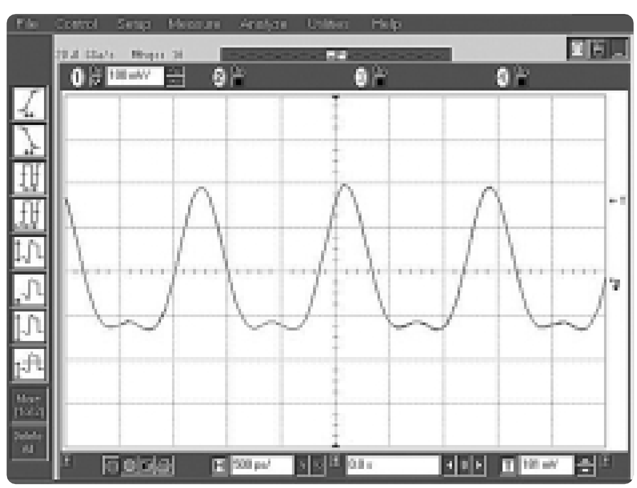


Figure 1-1. Complex time-domain signal

1. Jean Baptiste Joseph Fourier, 1768-1830. A French mathematician and physicist who discovered that periodic functions can be expanded into a series of sines and cosines.
2. If the time signal occurs only once, then T is infinite, and the frequency representation is a continuum of sine waves.

Some measurements require that we preserve complete information about the signal - frequency, amplitude and phase. This type of signal analysis is called *vector signal analysis*, which is discussed in *Application Note 150-15, Vector Signal Analysis Basics*. Modern spectrum analyzers are capable of performing a wide variety of vector signal measurements. However, another large group of measurements can be made without knowing the phase relationships among the sinusoidal components. This type of signal analysis is called *spectrum analysis*. Because spectrum analysis is simpler to understand, yet extremely useful, we will begin this application note by looking first at how spectrum analyzers perform spectrum analysis measurements, starting in Chapter 2.

Theoretically, to make the transformation from the time domain to the frequency domain, the signal must be evaluated over all time, that is, over ± infinity. However, in practice, we always use a finite time period when making a measurement. Fourier transformations can also be made from the frequency to the time domain. This case also theoretically requires the evaluation of all spectral components over frequencies to ± infinity. In reality, making measurements in a finite bandwidth that captures most of the signal energy produces acceptable results. When performing a Fourier transformation on frequency domain data, the phase of the individual components is indeed critical. For example, a square wave transformed to the frequency domain and back again could turn into a sawtooth wave if phase were not preserved.

What is a spectrum?

So what is a spectrum in the context of this discussion? A spectrum is a collection of sine waves that, when combined properly, produce the time-domain signal under examination. Figure 1-1 shows the waveform of a complex signal. Suppose that we were hoping to see a sine wave. Although the waveform certainly shows us that the signal is not a pure sinusoid, it does not give us a definitive indication of the reason why. Figure 1-2 shows our complex signal in both the time and frequency domains. The frequency-domain display plots the amplitude versus the frequency of each sine wave in the spectrum. As shown, the spectrum in this case comprises just two sine waves. We now know why our original waveform was not a pure sine wave. It contained a second sine wave, the second harmonic in this case. Does this mean we have no need to perform time-domain measurements? Not at all. The time domain is better for many measurements, and some can be made only in the time domain. For example, pure time-domain measurements include pulse rise and fall times, overshoot, and ringing.

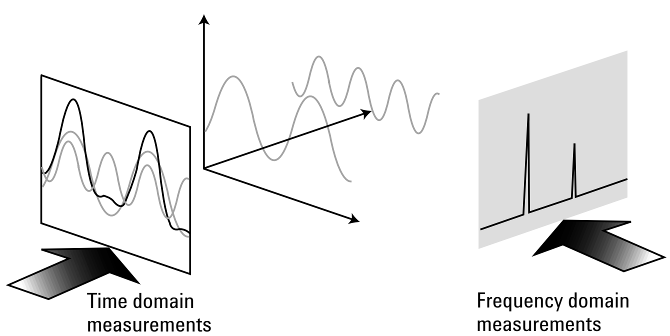


Figure 1-2. Relationship between time and frequency domain

Why measure spectra?

The frequency domain also has its measurement strengths. We have already seen in Figures 1-1 and 1-2 that the frequency domain is better for determining the harmonic content of a signal. People involved in wireless communications are extremely interested in out-of-band and spurious emissions. For example, cellular radio systems must be checked for harmonics of the carrier signal that might interfere with other systems operating at the same frequencies as the harmonics. Engineers and technicians are also very concerned about distortion of the message modulated onto a carrier. Third-order intermodulation (two tones of a complex signal modulating each other) can be particularly troublesome because the distortion components can fall within the band of interest and so will not be filtered away.

Spectrum monitoring is another important frequency-domain measurement activity. Government regulatory agencies allocate different frequencies for various radio services, such as broadcast television and radio, mobile phone systems, police and emergency communications, and a host of other applications. It is critical that each of these services operates at the assigned frequency and stays within the allocated channel bandwidth. Transmitters and other intentional radiators can often be required to operate at closely spaced adjacent frequencies. A key performance measure for the power amplifiers and other components used in these systems is the amount of signal energy that spills over into adjacent channels and causes interference.

Electromagnetic interference (EMI) is a term applied to unwanted emissions from both intentional and unintentional radiators. Here, the concern is that these unwanted emissions, either radiated or conducted (through the power lines or other interconnecting wires), might impair the operation of other systems. Almost anyone designing or manufacturing electrical or electronic products must test for emission levels versus frequency according to regulations set by various government agencies or industry-standard bodies. Figures 1-3 through 1-6 illustrate some of these measurements.

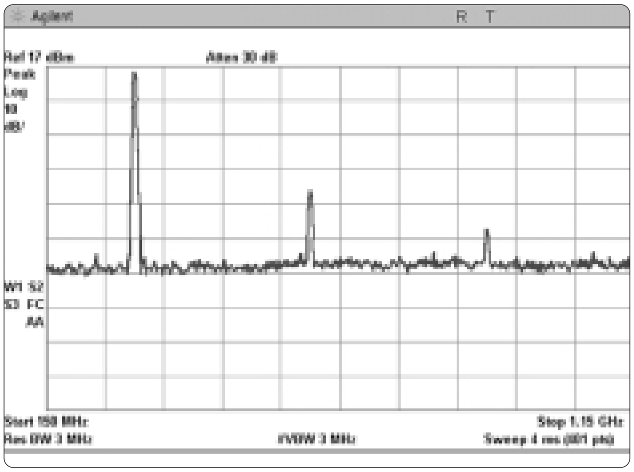


Figure 1-3. Harmonic distortion test of a transmitter

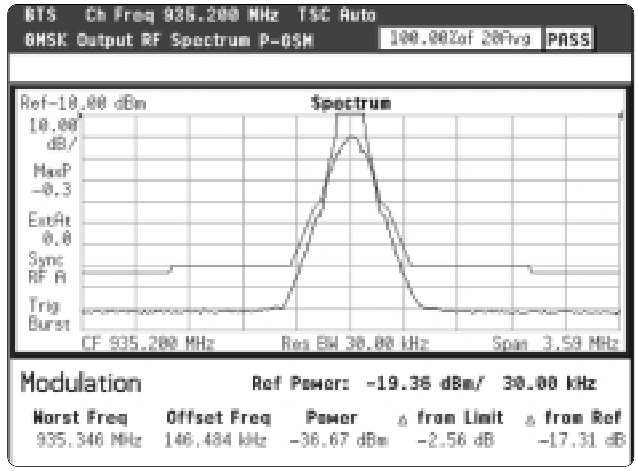


Figure 1-4. GSM radio signal and spectral mask showing limits of unwanted emissions

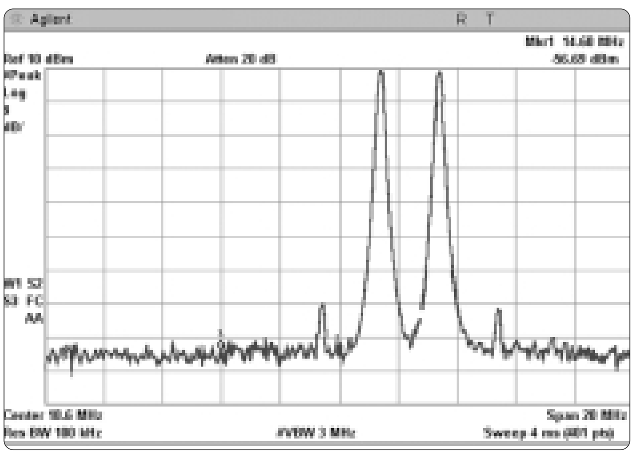


Figure 1- 5. Two-tone test on an RF power amplifier

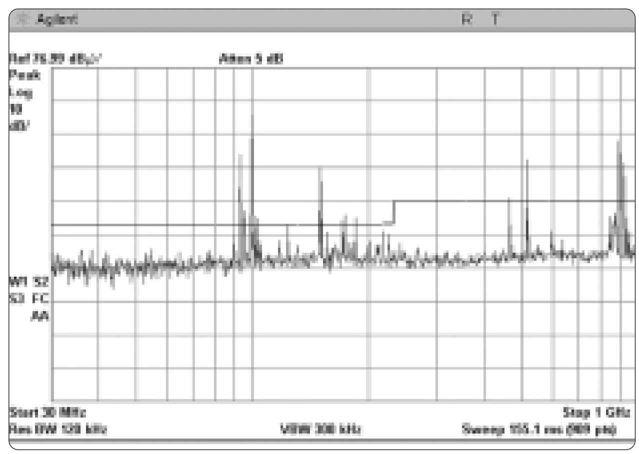


Figure 1-6. Radiated emissions plotted against CISPR11 limits as part of an EMI test

Types of measurements

Common spectrum analyzer measurements include frequency, power, modulation, distortion, and noise. Understanding the spectral content of a signal is important, especially in systems with limited bandwidth. Transmitted power is another key measurement. Too little power may mean the signal cannot reach its intended destination. Too much power may drain batteries rapidly, create distortion, and cause excessively high operating temperatures.

Measuring the quality of the modulation is important for making sure a system is working properly and that the information is being correctly transmitted by the system. Tests such as modulation degree, sideband amplitude, modulation quality, and occupied bandwidth are examples of common analog modulation measurements. Digital modulation metrics include error vector magnitude (EVM), IQ imbalance, phase error versus time, and a variety of other measurements. For more information on these measurements, see *Application Note 150-15, Vector Signal Analysis Basics*.

In communications, measuring distortion is critical for both the receiver and transmitter. Excessive harmonic distortion at the output of a transmitter can interfere with other communication bands. The pre-amplification stages in a receiver must be free of intermodulation distortion to prevent signal crosstalk. An example is the intermodulation of cable TV carriers as they move down the trunk of the distribution system and distort other channels on the same cable. Common distortion measurements include intermodulation, harmonics, and spurious emissions.

Noise is often the signal you want to measure. Any active circuit or device will generate excess noise. Tests such as noise figure and signal-to-noise ratio (SNR) are important for characterizing the performance of a device and its contribution to overall system performance.

Types of signal analyzers

While we shall concentrate on the swept-tuned, superheterodyne spectrum analyzer in this note, there are several other signal analyzer architectures. An important non-superheterodyne type is the Fourier analyzer, which digitizes the time-domain signal and then uses digital signal processing (DSP) techniques to perform a fast Fourier transform (FFT) and display the signal in the frequency domain. One advantage of the FFT approach is its ability to characterize single-shot phenomena. Another is that phase as well as magnitude can be measured. However, Fourier analyzers do have some limitations relative to the superheterodyne spectrum analyzer, particularly in the areas of frequency range, sensitivity, and dynamic range. Fourier analyzers are typically used in baseband signal analysis applications up to 40 MHz.

Vector signal analyzers (VSAs) also digitize the time domain signal like Fourier analyzers, but extend the capabilities to the RF frequency range using downconverters in front of the digitizer. For example, the Agilent 89600 Series VSA offers various models available up to 6 GHz. They offer fast, high-resolution spectrum measurements, demodulation, and advanced time-domain analysis. They are especially useful for characterizing complex signals such as burst, transient or modulated signals used in communications, video, broadcast, sonar, and ultrasound imaging applications.

While we have defined *spectrum analysis* and *vector signal analysis* as distinct types, digital technology and digital signal processing are blurring that distinction. The critical factor is where the signal is digitized. Early on, when digitizers were limited to a few tens of kilohertz, only the video (baseband) signal of a spectrum analyzer was digitized. Since the video signal carried no phase information, only magnitude data could be displayed. But even this limited use of digital technology yielded significant advances: flicker-free displays of slow sweeps, display markers, different types of averaging, and data output to computers and printers.

Because the signals that people must analyze are becoming more complex, the latest generations of spectrum analyzers include many of the vector signal analysis capabilities previously found only in Fourier and vector signal analyzers. Analyzers may digitize the signal near the instrument's input, after some amplification, or after one or more downconverter stages. In any of these cases, relative phase as well as magnitude is preserved. In addition to the benefits noted above, true vector measurements can be made. Capabilities are then determined by the digital signal processing capability inherent in the analyzer's firmware or available as add-on software running either internally (measurement personalities) or externally (vector signal analysis software) on a computer connected to the analyzer. An example of this capability is shown in Figure 1-7. Note that the symbol points of a QPSK (quadrature phase shift keying) signal are displayed as clusters, rather than single points, indicating errors in the modulation of the signal under test.

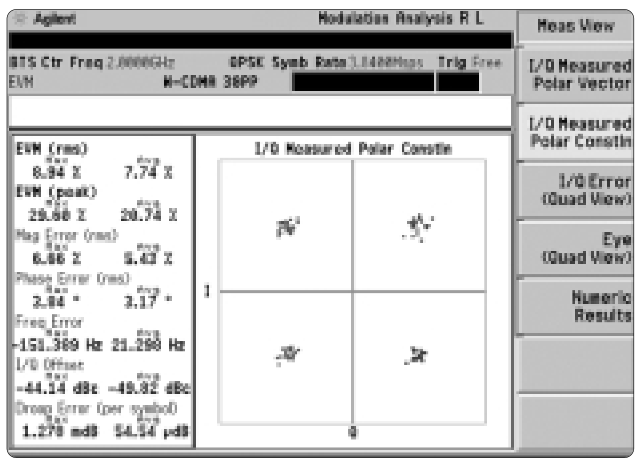


Figure 1-7. Modulation analysis of a QPSK signal measured with a spectrum analyzer

We hope that this application note gives you the insight into your particular spectrum analyzer and enables you to utilize this versatile instrument to its maximum potential.

Chapter 2
Spectrum Analyzer Fundamentals

This chapter will focus on the fundamental theory of how a spectrum analyzer works. While today's technology makes it possible to replace many analog circuits with modern digital implementations, it is very useful to understand classic spectrum analyzer architecture as a starting point in our discussion. In later chapters, we will look at the capabilities and advantages that digital circuitry brings to spectrum analysis. Chapter 3 will discuss digital architectures used in modern spectrum analyzers.

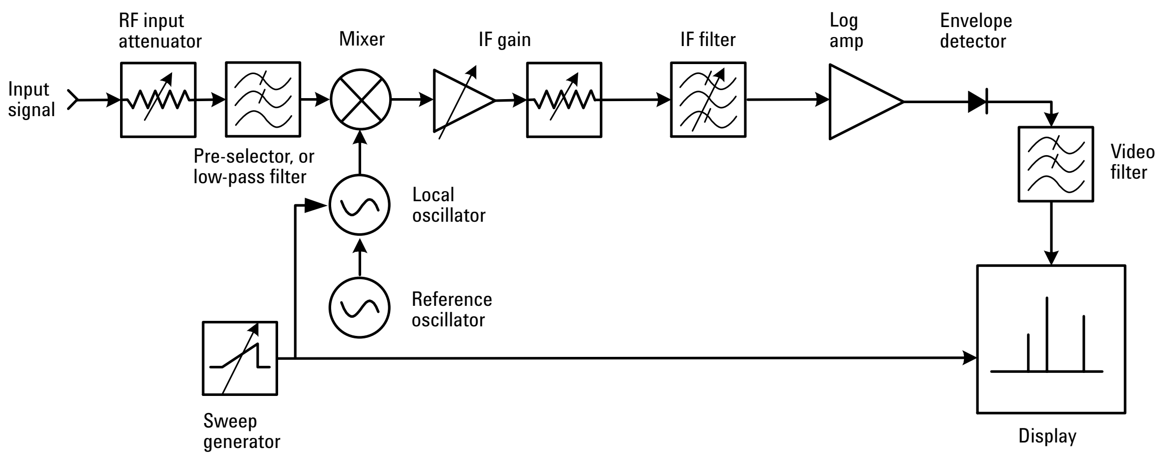


Figure 2-1. Block diagram of a classic superheterodyne spectrum analyzer

Figure 2-1 is a simplified block diagram of a superheterodyne spectrum analyzer. Heterodyne means to mix; that is, to translate frequency. And super refers to super-audio frequencies, or frequencies above the audio range. Referring to the block diagram in Figure 2-1, we see that an input signal passes through an attenuator, then through a low-pass filter (later we shall see why the filter is here) to a mixer, where it mixes with a signal from the local oscillator (LO). Because the mixer is a non-linear device, its output includes not only the two original signals, but also their harmonics and the sums and differences of the original frequencies and their harmonics. If any of the mixed signals falls within the passband of the intermediate-frequency (IF) filter, it is further processed (amplified and perhaps compressed on a logarithmic scale). It is essentially rectified by the envelope detector, digitized, and displayed. A ramp generator creates the horizontal movement across the display from left to right. The ramp also tunes the LO so that its frequency change is in proportion to the ramp voltage.

If you are familiar with superheterodyne AM radios, the type that receive ordinary AM broadcast signals, you will note a strong similarity between them and the block diagram of Figure 2-1. The differences are that the output of a spectrum analyzer is a display instead of a speaker, and the local oscillator is tuned electronically rather than by a front-panel knob.

Since the output of a spectrum analyzer is an X-Y trace on a display, let's see what information we get from it. The display is mapped on a grid (graticule) with ten major horizontal divisions and generally ten major vertical divisions. The horizontal axis is linearly calibrated in frequency that increases from left to right. Setting the frequency is a two-step process. First we adjust the frequency at the centerline of the graticule with the center frequency control. Then we adjust the frequency range (span) across the full ten divisions with the Frequency Span control. These controls are independent, so if we change the center frequency, we do not alter the frequency span. Alternatively, we can set the start and stop frequencies instead of setting center frequency and span. In either case, we can determine the absolute frequency of any signal displayed and the relative frequency difference between any two signals.

The vertical axis is calibrated in amplitude. We have the choice of a linear scale calibrated in volts or a logarithmic scale calibrated in dB. The log scale is used far more often than the linear scale because it has a much wider usable range. The log scale allows signals as far apart in amplitude as 70 to 100 dB (voltage ratios of 3200 to 100,000 and power ratios of 10,000,000 to 10,000,000,000) to be displayed simultaneously. On the other hand, the linear scale is usable for signals differing by no more than 20 to 30 dB (voltage ratios of 10 to 32). In either case, we give the top line of the graticule, the reference level, an absolute value through calibration techniques[1] and use the scaling per division to assign values to other locations on the graticule. Therefore, we can measure either the absolute value of a signal or the relative amplitude difference between any two signals.

Scale calibration, both frequency and amplitude, is shown by annotation written onto the display. Figure 2-2 shows the display of a typical analyzer. Now, let's turn our attention back to Figure 2-1.

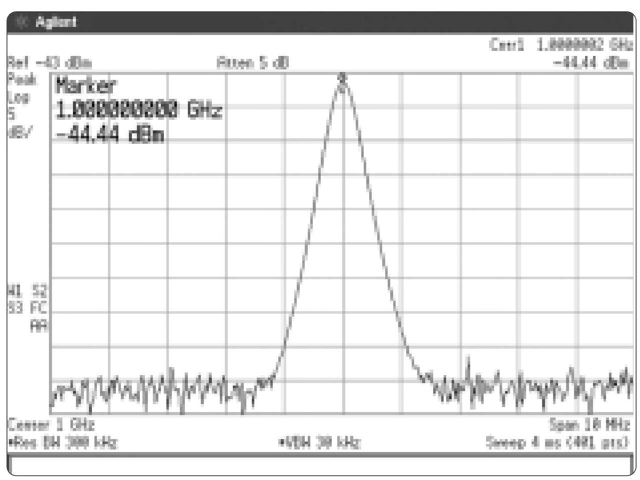


Figure 2-2. Typical spectrum analyzer display with control settings

1. See Chapter 4, "Amplitude and Frequency Accuracy."

RF attenuator

The first part of our analyzer is the RF input attenuator. Its purpose is to ensure the signal enters the mixer at the optimum level to prevent overload, gain compression, and distortion. Because attenuation is a protective circuit for the analyzer, it is usually set automatically, based on the reference level. However, manual selection of attenuation is also available in steps of 10, 5, 2, or even 1 dB. The diagram below is an example of an attenuator circuit with a maximum attenuation of 70 dB in increments of 2 dB. The blocking capacitor is used to prevent the analyzer from being damaged by a DC signal or a DC offset of the signal. Unfortunately, it also attenuates low frequency signals and increases the minimum useable start frequency of the analyzer to 100 Hz for some analyzers, 9 kHz for others.

In some analyzers, an amplitude reference signal can be connected as shown in Figure 2-3. It provides a precise frequency and amplitude signal, used by the analyzer to periodically self-calibrate.

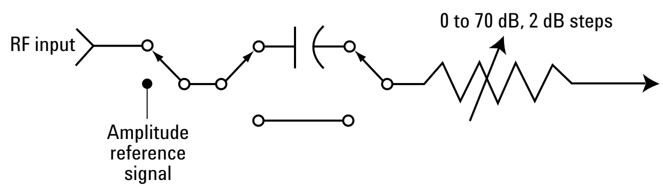


Figure 2-3. RF input attenuator circuitry

Low-pass filter or preselector

The low-pass filter blocks high frequency signals from reaching the mixer. This prevents out-of-band signals from mixing with the local oscillator and creating unwanted responses at the IF. Microwave spectrum analyzers replace the low-pass filter with a preselector, which is a tunable filter that rejects all frequencies except those that we currently wish to view. In Chapter 7, we will go into more detail about the operation and purpose of filtering the input.

Tuning the analyzer

We need to know how to tune our spectrum analyzer to the desired frequency range. Tuning is a function of the center frequency of the IF filter, the frequency range of the LO, and the range of frequencies allowed to reach the mixer from the outside world (allowed to pass through the low-pass filter). Of all the mixing products emerging from the mixer, the two with the greatest amplitudes, and therefore the most desirable, are those created from the sum of the LO and input signal and from the difference between the LO and input signal. If we can arrange things so that the signal we wish to examine is either above or below the LO frequency by the IF, then one of the desired mixing products will fall within the pass-band of the IF filter and be detected to create an amplitude response on the display.

We need to pick an LO frequency and an IF that will create an analyzer with the desired tuning range. Let's assume that we want a tuning range from 0 to 3 GHz. We then need to choose the IF frequency. Let's try a 1 GHz IF. Since this frequency is within our desired tuning range, we could have an input signal at 1 GHz. Since the output of a mixer also includes the original input signals, an input signal at 1 GHz would give us a constant output from the mixer at the IF. The 1 GHz signal would thus pass through the system and give us a constant amplitude response on the display regardless of the tuning of the LO. The result would be a hole in the frequency range at which we could not properly examine signals because the amplitude response would be independent of the LO frequency. Therefore, a 1 GHz IF will not work.

So we shall choose, instead, an IF that is above the highest frequency to which we wish to tune. In Agilent spectrum analyzers that can tune to 3 GHz, the IF chosen is about 3.9 GHz. Remember that we want to tune from 0 Hz to 3 GHz. (Actually from some low frequency because we cannot view a 0 Hz signal with this architecture.) If we start the LO at the IF (LO minus IF = 0 Hz) and tune it upward from there to 3 GHz above the IF, then we can cover the tuning range with the LO minus IF mixing product. Using this information, we can generate a tuning equation:

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

where f_{sig} = signal frequency
f_{LO} = local oscillator frequency, and
f_{IF} = intermediate frequency (IF)

If we wanted to determine the LO frequency needed to tune the analyzer to a low-, mid-, or high-frequency signal (say, 1 kHz, 1.5 GHz, or 3 GHz), we would first restate the tuning equation in terms of f_{LO}:

$$f_{LO} = f_{sig} + f_{IF}$$

Then we would plug in the numbers for the signal and IF in the tuning equation[2]:

f_{LO} = 1 kHz + 3.9 GHz = 3.900001 GHz,
f_{LO} = 1.5 GHz + 3.9 GHz = 5.4 GHz, or
f_{LO} = 3 GHz; + 3.9 GHz = 6.9 GHz.

2. In the text, we shall round off some of the frequency values for simplicity, although the exact values are shown in the figures.

Figure 2-4 illustrates analyzer tuning. In this figure, f_{LO} is not quite high enough to cause the $f_{LO} - f_{sig}$ mixing product to fall in the IF passband, so there is no response on the display. If we adjust the ramp generator to tune the LO higher, however, this mixing product will fall in the IF passband at some point on the ramp (sweep), and we shall see a response on the display.

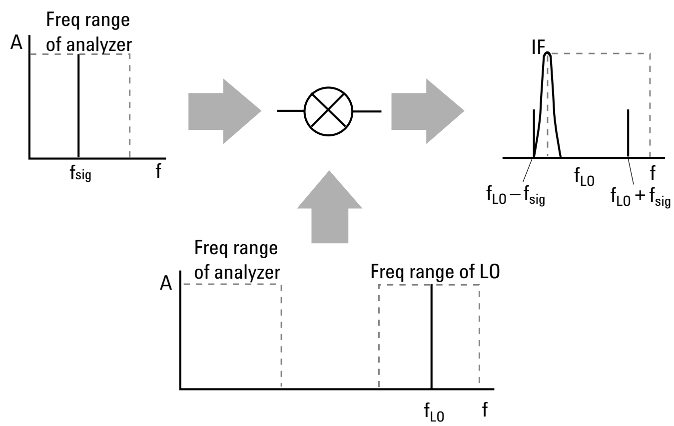


Figure 2-4. The LO must be tuned to $f_{IF} + f_{sig}$ to produce a response on the display

Since the ramp generator controls both the horizontal position of the trace on the display and the LO frequency, we can now calibrate the horizontal axis of the display in terms of the input signal frequency.

We are not quite through with the tuning yet. What happens if the frequency of the input signal is 8.2 GHz? As the LO tunes through its 3.9 to 7.0 GHz range, it reaches a frequency (4.3 GHz) at which it is the IF away from the 8.2 GHz input signal. At this frequency we have a mixing product that is equal to the IF, creating a response on the display. In other words, the tuning equation could just as easily have been:

$$f_{sig} = f_{LO} + f_{IF}$$

This equation says that the architecture of Figure 2-1 could also result in a tuning range from 7.8 to 10.9 GHz, but only if we allow signals in that range to reach the mixer. The job of the input low-pass filter in Figure 2-1 is to prevent these higher frequencies from getting to the mixer. We also want to keep signals at the intermediate frequency itself from reaching the mixer, as previously described, so the low-pass filter must do a good job of attenuating signals at 3.9 GHz, as well as in the range from 7.8 to 10.9 GHz.

In summary, we can say that for a single-band RF spectrum analyzer, we would choose an IF above the highest frequency of the tuning range. We would make the LO tunable from the IF to the IF plus the upper limit of the tuning range and include a low-pass filter in front of the mixer that cuts off below the IF.

To separate closely spaced signals (see "Resolving signals" later in this chapter), some spectrum analyzers have IF bandwidths as narrow as 1 kHz; others, 10 Hz; still others, 1 Hz. Such narrow filters are difficult to achieve at a center frequency of 3.9 GHz. So we must add additional mixing stages, typically two to four stages, to down-convert from the first to the final IF. Figure 2-5 shows a possible IF chain based on the architecture of a typical spectrum analyzer. The full tuning equation for this analyzer is:

$$f_{sig} = f_{LO1} - (f_{LO2} + f_{LO3} + f_{final\ IF})$$

However,

$$f_{LO2} + f_{LO3} + f_{final\ IF}$$
$$= 3.6\text{ GHz} + 300\text{ MHz} + 21.4\text{ MHz}$$
$$= 3.9214\text{ GHz, the first IF}$$

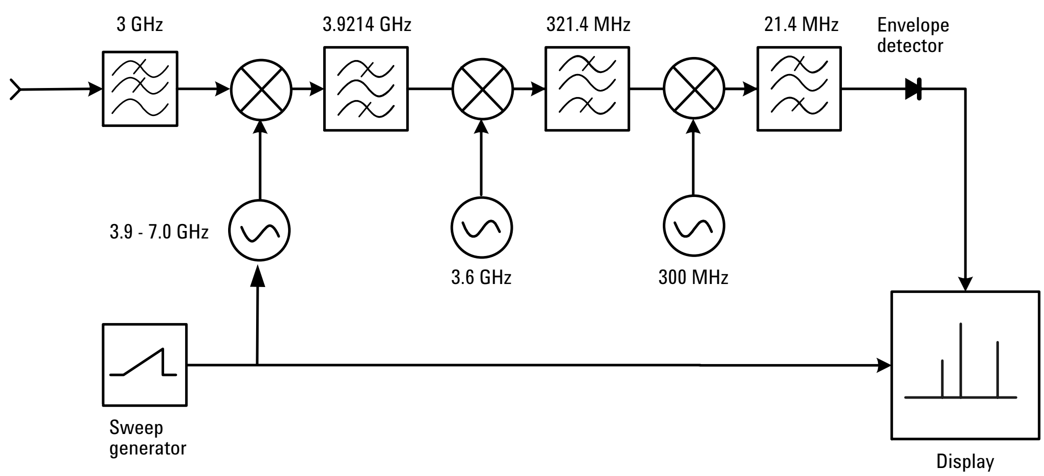


Figure 2-5. Most spectrum analyzers use two to four mixing steps to reach the final IF

So simplifying the tuning equation by using just the first IF leads us to the same answers. Although only passive filters are shown in the diagram, the actual implementation includes amplification in the narrower IF stages. The final IF section contains additional components, such as logarithmic amplifiers or analog to digital converters, depending on the design of the particular analyzer.

Most RF spectrum analyzers allow an LO frequency as low as, and even below, the first IF. Because there is finite isolation between the LO and IF ports of the mixer, the LO appears at the mixer output. When the LO equals the IF, the LO signal itself is processed by the system and appears as a response on the display, as if it were an input signal at 0 Hz. This response, called LO feedthrough, can mask very low frequency signals, so not all analyzers allow the display range to include 0 Hz.

IF gain

Referring back to Figure 2-1, we see the next component of the block diagram is a variable gain amplifier. It is used to adjust the vertical position of signals on the display without affecting the signal level at the input mixer. When the IF gain is changed, the value of the reference level is changed accordingly to retain the correct indicated value for the displayed signals. Generally, we do not want the reference level to change when we change the input attenuator, so the settings of the input attenuator and the IF gain are coupled together. A change in input attenuation will automatically change the IF gain to offset the effect of the change in input attenuation, thereby keeping the signal at a constant position on the display.

Resolving signals

After the IF gain amplifier, we find the IF section which consists of the analog and/or digital resolution bandwidth (RBW) filters.

Analog filters

Frequency resolution is the ability of a spectrum analyzer to separate two input sinusoids into distinct responses. Fourier tells us that a sine wave signal only has energy at one frequency, so we shouldn't have any resolution problems. Two signals, no matter how close in frequency, should appear as two lines on the display. But a closer look at our superheterodyne receiver shows why signal responses have a definite width on the display. The output of a mixer includes the sum and difference products plus the two original signals (input and LO). A bandpass filter determines the intermediate frequency, and this filter selects the desired mixing product and rejects all other signals. Because the input signal is fixed and the local oscillator is swept, the products from the mixer are also swept. If a mixing product happens to sweep past the IF, the characteristic shape of the bandpass filter is traced on the display. See Figure 2-6. The narrowest filter in the chain determines the overall displayed bandwidth, and in the architecture of Figure 2-5, this filter is in the 21.4 MHz IF.

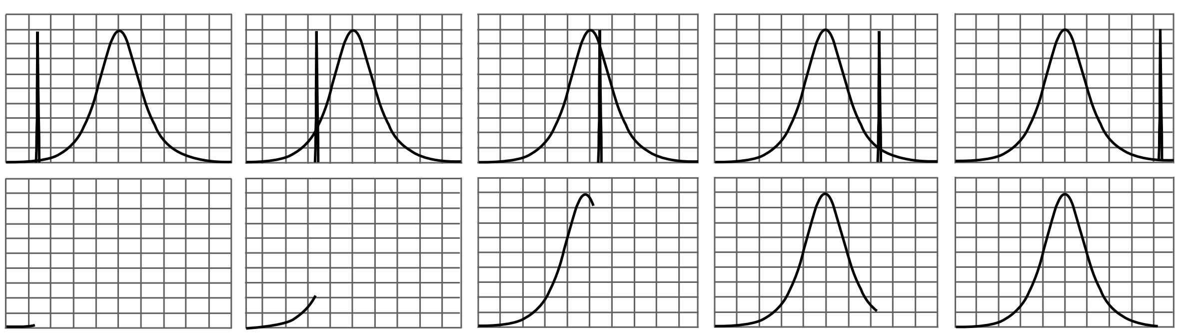


Figure 2-6. As a mixing product sweeps past the IF filter, the filter shape is traced on the display

So two signals must be far enough apart, or else the traces they make will fall on top of each other and look like only one response. Fortunately, spectrum analyzers have selectable resolution (IF) filters, so it is usually possible to select one narrow enough to resolve closely spaced signals.

Agilent data sheets describe the ability to resolve signals by listing the 3 dB bandwidths of the available IF filters. This number tells us how close together equal-amplitude sinusoids can be and still be resolved. In this case, there will be about a 3 dB dip between the two peaks traced out by these signals. See Figure 2-7. The signals can be closer together before their traces merge completely, but the 3 dB bandwidth is a good rule of thumb for resolution of equal-amplitude signals[3].

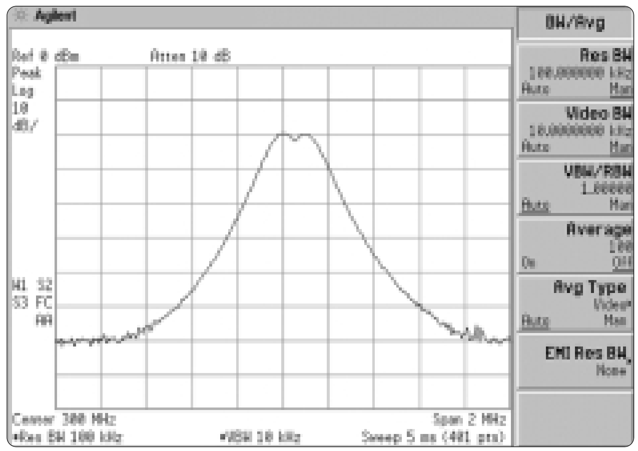


Figure 2-7. Two equal-amplitude sinusoids separated by the 3 dB BW of the selected IF filter can be resolved

More often than not we are dealing with sinusoids that are not equal in amplitude. The smaller sinusoid can actually be lost under the skirt of the response traced out by the larger. This effect is illustrated in Figure 2-8. The top trace looks like a single signal, but in fact represents two signals: one at 300 MHz (0 dBm) and another at 300.005 MHz (–30 dBm). The lower trace shows the display after the 300 MHz signal is removed.

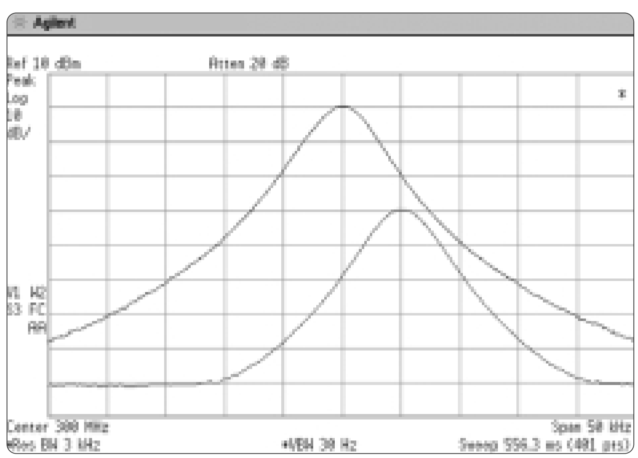


Figure 2-8. A low-level signal can be lost under skirt of the response to a larger signal

3. If you experiment with resolution on a spectrum analyzer using the normal (rosenfell) detector mode (See "Detector types" later in this chapter) use enough video filtering to create a smooth trace. Otherwise, there will be a smearing as the two signals interact. While the smeared trace certainly indicates the presence of more than one signal, it is difficult to determine the amplitudes of the individual signals. Spectrum analyzers with positive peak as their default detector mode may not show the smearing effect. You can observe the smearing by selecting the sample detector mode.

Another specification is listed for the resolution filters: bandwidth selectivity (or selectivity or shape factor). Bandwidth selectivity helps determine the resolving power for unequal sinusoids. For Agilent analyzers, bandwidth selectivity is generally specified as the ratio of the 60 dB bandwidth to the 3 dB bandwidth, as shown in Figure 2-9. The analog filters in Agilent analyzers are a four-pole, synchronously-tuned design, with a nearly Gaussian shape[4]. This type of filter exhibits a bandwidth selectivity of about 12.7:1.

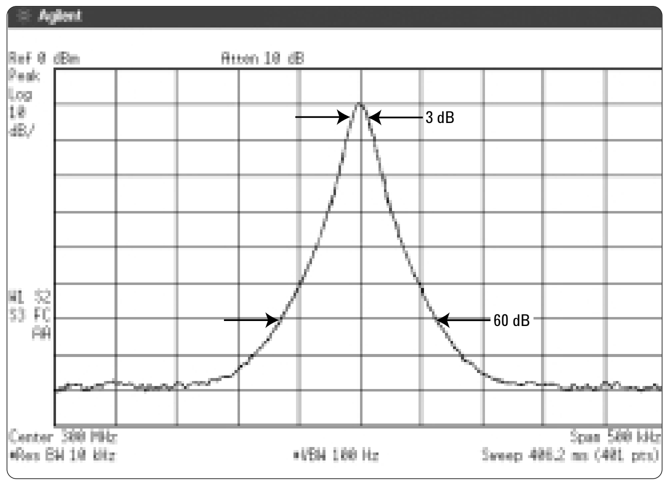


Figure 2-9. Bandwidth selectivity, ratio of 60 dB to 3 dB bandwidths

For example, what resolution bandwidth must we choose to resolve signals that differ by 4 kHz and 30 dB, assuming 12.7:1 bandwidth selectivity? Since we are concerned with rejection of the larger signal when the analyzer is tuned to the smaller signal, we need to consider not the full bandwidth, but the frequency difference from the filter center frequency to the skirt. To determine how far down the filter skirt is at a given offset, we use the following equation:

$$H(\Delta f) = -10(N)\log_{10}[(\Delta f/f_0)^2 + 1]$$

Where $H(\Delta f)$ is the filter skirt rejection in dB
N is the number of filter poles
Δf is the frequency offset from the center in Hz, and
f_0 is given by $\dfrac{RBW}{2\sqrt{2^{1/N}-1}}$

For our example, N=4 and Δf = 4000. Let's begin by trying the 3 kHz RBW filter. First, we compute f_0:

$$f_0 = \frac{3000}{2\sqrt{2^{1/4}-1}} = 3448.44$$

Now we can determine the filter rejection at a 4 kHz offset:

$$H(4000) = -10(4)\log_{10}[(4000/3448.44)^2 + 1]$$
$$= -14.8 \text{ dB}$$

This is not enough to allow us to see the smaller signal. Let's determine $H(\Delta f)$ again using a 1 kHz filter:

$$f_0 = \frac{1000}{2\sqrt{2^{1/4}-1}} = 1149.48$$

4. Some older spectrum analyzer models used five-pole filters for the narrowest resolution bandwidths to provide improved selectivity of about 10:1. Modern designs achieve even better bandwidth selectivity using digital IF filters.

This allows us to calculate the filter rejection:

$$H(4000) = -10(4)\log_{10}[(4000/1149.48)^2 + 1]$$
$$= -44.7 \text{ dB}$$

Thus, the 1 kHz resolution bandwidth filter does resolve the smaller signal. This is illustrated in Figure 2-10.

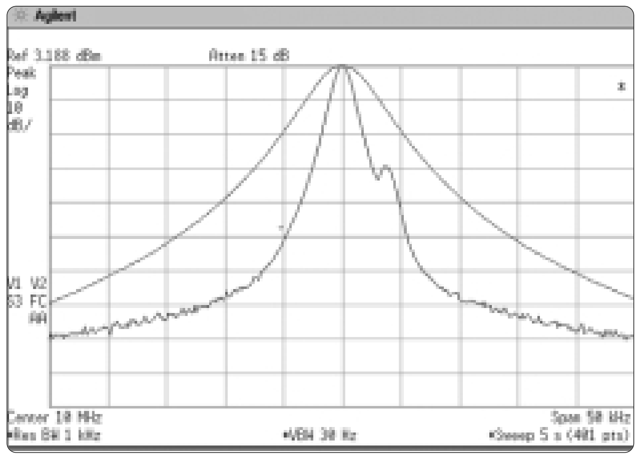


Figure 2-10. The 3 kHz filter (top trace) does not resolve smaller signal; reducing the resolution bandwidth to 1 kHz (bottom trace) does

Digital filters

Some spectrum analyzers use digital techniques to realize their resolution bandwidth filters. Digital filters can provide important benefits, such as dramatically improved bandwidth selectivity. The Agilent PSA Series spectrum analyzers implement all resolution bandwidths digitally. Other analyzers, such as the Agilent ESA-E Series, take a hybrid approach, using analog filters for the wider bandwidths and digital filters for bandwidths of 300 Hz and below. Refer to Chapter 3 for more information on digital filters.

Residual FM

Filter bandwidth is not the only factor that affects the resolution of a spectrum analyzer. The stability of the LOs in the analyzer, particularly the first LO, also affects resolution. The first LO is typically a YIG-tuned oscillator (tuning somewhere in the 3 to 7 GHz range). In early spectrum analyzer designs, these oscillators had residual FM of 1 kHz or more. This instability was transferred to any mixing products resulting from the LO and incoming signals, and it was not possible to determine whether the input signal or the LO was the source of this instability.

The minimum resolution bandwidth is determined, at least in part, by the stability of the first LO. Analyzers where no steps are taken to improve upon the inherent residual FM of the YIG oscillators typically have a minimum bandwidth of 1 kHz. However, modern analyzers have dramatically improved residual FM. For example, Agilent PSA Series analyzers have residual FM of 1 to 4 Hz and ESA Series analyzers have 2 to 8 Hz residual FM. This allows bandwidths as low as 1 Hz. So any instability we see on a spectrum analyzer today is due to the incoming signal.

Phase noise

Even though we may not be able to see the actual frequency jitter of a spectrum analyzer LO system, there is still a manifestation of the LO frequency or phase instability that can be observed. This is known as phase noise (sometimes called sideband noise). No oscillator is perfectly stable. All are frequency or phase modulated by random noise to some extent. As previously noted, any instability in the LO is transferred to any mixing products resulting from the LO and input signals. So the LO phase-noise modulation sidebands appear around any spectral component on the display that is far enough above the broadband noise floor of the system (Figure 2-11). The amplitude difference between a displayed spectral component and the phase noise is a function of the stability of the LO. The more stable the LO, the farther down the phase noise. The amplitude difference is also a function of the resolution bandwidth. If we reduce the resolution bandwidth by a factor of ten, the level of the displayed phase noise decreases by 10 dB[5].

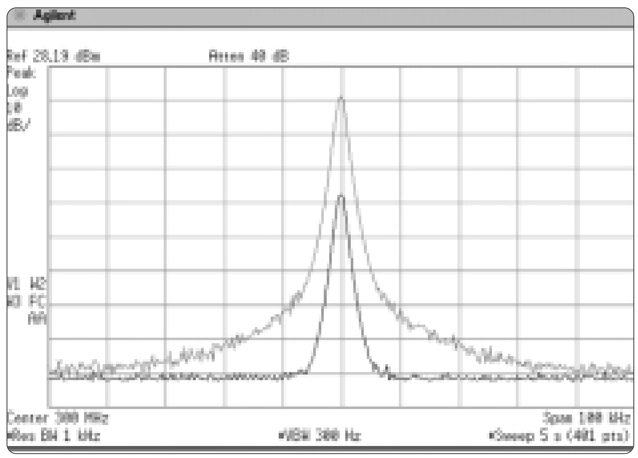


Figure 2-11. Phase noise is displayed only when a signal is displayed far enough above the system noise floor

The shape of the phase noise spectrum is a function of analyzer design, in particular, the sophistication of the phase lock loops employed to stabilized the LO. In some analyzers, the phase noise is a relatively flat pedestal out to the bandwidth of the stabilizing loop. In others, the phase noise may fall away as a function of frequency offset from the signal. Phase noise is specified in terms of dBc (dB relative to a carrier) and normalized to a 1 Hz noise power bandwidth. It is sometimes specified at specific frequency offsets. At other times, a curve is given to show the phase noise characteristics over a range of offsets.

Generally, we can see the inherent phase noise of a spectrum analyzer only in the narrower resolution filters, when it obscures the lower skirts of these filters. The use of the digital filters previously described does not change this effect. For wider filters, the phase noise is hidden under the filter skirt, just as in the case of two unequal sinusoids discussed earlier.

5. The effect is the same for the broadband noise floor (or any broadband noise signal). See Chapter 5, "Sensitivity and Noise."

Some modern spectrum analyzers allow the user to select different LO stabilization modes to optimize the phase noise for different measurement conditions. For example, the PSA Series spectrum analyzers offer three different modes:

- *Optimize phase noise for frequency offsets < 50 kHz from the carrier*
  In this mode, the LO phase noise is optimized for the area close in to the carrier at the expense of phase noise beyond 50 kHz offset.
- *Optimize phase noise for frequency offsets > 50 kHz from the carrier*
  This mode optimizes phase noise for offsets above 50 kHz away from the carrier, especially those from 70 kHz to 300 kHz. Closer offsets are compromised and the throughput of measurements is reduced.
- *Optimize LO for fast tuning*
  When this mode is selected, LO behavior compromises phase noise at all offsets from the carrier below approximately 2 MHz. This minimizes measurement time and allows the maximum measurement throughput when changing the center frequency or span.

The PSA spectrum analyzer phase noise optimization can also be set to auto mode, which automatically sets the instrument’s behavior to optimize speed or dynamic range for various operating conditions. When the span is ≥ 10.5 MHz or the RBW is > 200 kHz, the PSA selects fast tuning mode. For spans >141.4 kHz and RBWs > 9.1 kHz, the auto mode optimizes for offsets > 50 kHz. For all other cases, the spectrum analyzer optimizes for offsets < 50 kHz. These three modes are shown in Figure 2-12a.

The ESA spectrum analyzer uses a simpler optimization scheme than the PSA, offering two user-selectable modes, optimize for best phase noise and optimize LO for fast tuning, as well as an auto mode.

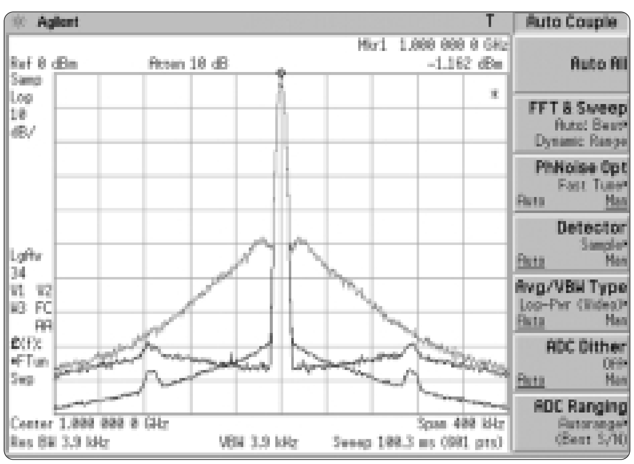


Figure 2-12a. Phase noise performance can be optimized for different measurement conditions

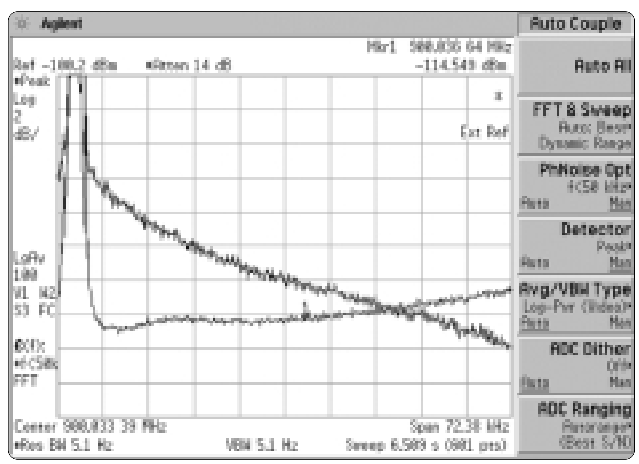


Figure 2-12b. Shows more detail of the 50 kHz carrier offset region

In any case, phase noise becomes the ultimate limitation in an analyzer's ability to resolve signals of unequal amplitude. As shown in Figure 2-13, we may have determined that we can resolve two signals based on the 3 dB bandwidth and selectivity, only to find that the phase noise covers up the smaller signal.

Sweep time

Analog resolution filters

If resolution were the only criterion on which we judged a spectrum analyzer, we might design our analyzer with the narrowest possible resolution (IF) filter and let it go at that. But resolution affects sweep time, and we care very much about sweep time. Sweep time directly affects how long it takes to complete a measurement.

Resolution comes into play because the IF filters are band-limited circuits that require finite times to charge and discharge. If the mixing products are swept through them too quickly, there will be a loss of displayed amplitude as shown in Figure 2-14. (See "Envelope detector," later in this chapter, for another approach to IF response time.) If we think about how long a mixing product stays in the passband of the IF filter, that time is directly proportional to bandwidth and inversely proportional to the sweep in Hz per unit time, or:

$$\text{Time in passband} = \frac{\text{RBW}}{\text{Span/ST}} = \frac{(\text{RBW})(\text{ST})}{\text{Span}}$$

where RBW = resolution bandwidth and
ST = sweep time.

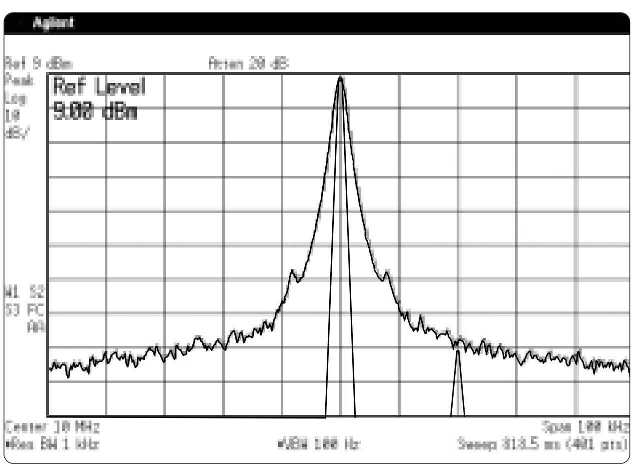


Figure 2-13. Phase noise can prevent resolution of unequal signals

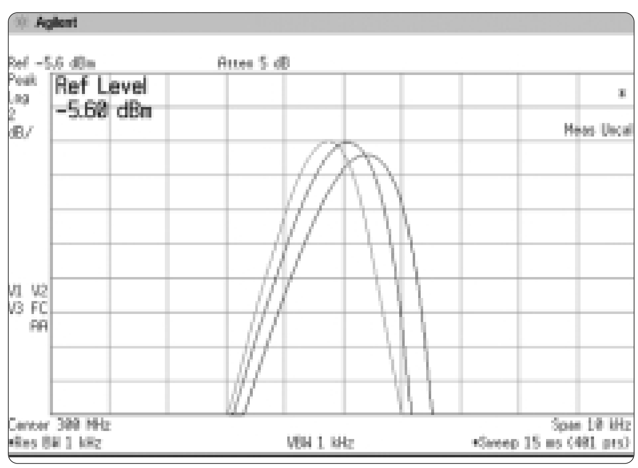


Figure 2-14. Sweeping an analyzer too fast causes a drop in displayed amplitude and a shift in indicated frequency

On the other hand, the rise time of a filter is inversely proportional to its bandwidth, and if we include a constant of proportionality, k, then:

$$\text{Rise time} = \frac{k}{\text{RBW}}$$

If we make the terms equal and solve for sweep time, we have:

$$\frac{k}{\text{RBW}} = \frac{(\text{RBW})(\text{ST})}{\text{Span}} \quad \text{or:}$$

$$\text{ST} = \frac{k\,(\text{Span})}{\text{RBW}^2}$$

The value of k is in the 2 to 3 range for the synchronously-tuned, near-Gaussian filters used in many Agilent analyzers.

The important message here is that a change in resolution has a dramatic effect on sweep time. Most Agilent analyzers provide values in a 1, 3, 10 sequence or in ratios roughly equaling the square root of 10. So sweep time is affected by a factor of about 10 with each step in resolution. Agilent PSA Series spectrum analyzers offer bandwidth steps of just 10% for an even better compromise among span, resolution, and sweep time.

Spectrum analyzers automatically couple sweep time to the span and resolution bandwidth settings. Sweep time is adjusted to maintain a calibrated display. If a sweep time longer than the maximum available is called for, the analyzer indicates that the display is uncalibrated with a “Meas Uncal” message in the upper-right part of the graticule. We are allowed to override the automatic setting and set sweep time manually if the need arises.

Digital resolution filters

The digital resolution filters used in Agilent spectrum analyzers have an effect on sweep time that is different from the effects we’ve just discussed for analog filters. For swept analysis, the speed of digitally implemented filters can show a 2 to 4 times improvement. FFT-based digital filters show an even greater difference. This difference occurs because the signal being analyzed is processed in frequency blocks, depending upon the particular analyzer. For example, if the frequency block was 1 kHz, then when we select a 10 Hz resolution bandwidth, the analyzer is in effect simultaneously processing the data in each 1 kHz block through 100 contiguous 10 Hz filters. If the digital processing were instantaneous, we would expect sweep time to be reduced by a factor of 100. In practice, the reduction factor is less, but is still significant. For more information on the advantages of digital processing, refer to Chapter 3.

Envelope detector[6]

Spectrum analyzers typically convert the IF signal to video[7] with an envelope detector. In its simplest form, an envelope detector consists of a diode, resistive load and low-pass filter, as shown in Figure 2-15. The output of the IF chain in this example, an amplitude modulated sine wave, is applied to the detector. The response of the detector follows the changes in the envelope of the IF signal, but not the instantaneous value of the IF sine wave itself.

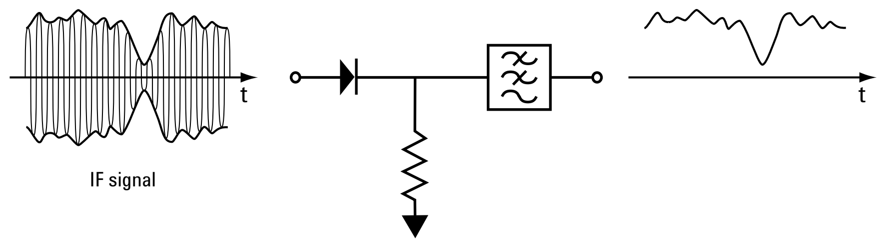


Figure 2-15. Envelope detector

For most measurements, we choose a resolution bandwidth narrow enough to resolve the individual spectral components of the input signal. If we fix the frequency of the LO so that our analyzer is tuned to one of the spectral components of the signal, the output of the IF is a steady sine wave with a constant peak value. The output of the envelope detector will then be a constant (dc) voltage, and there is no variation for the detector to follow.

However, there are times when we deliberately choose a resolution bandwidth wide enough to include two or more spectral components. At other times, we have no choice. The spectral components are closer in frequency than our narrowest bandwidth. Assuming only two spectral components within the passband, we have two sine waves interacting to create a beat note, and the envelope of the IF signal varies, as shown in Figure 2-16, as the phase between the two sine waves varies.

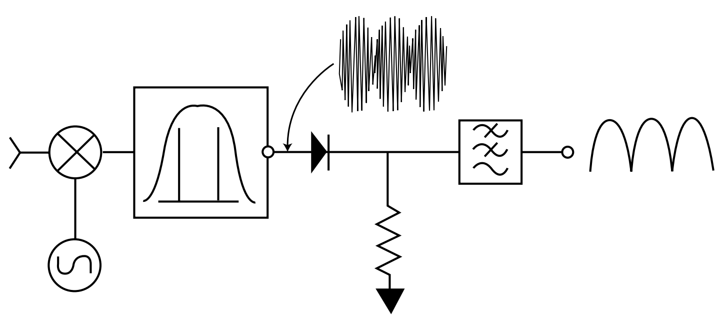


Figure 2-16. Output of the envelope detector follows the peaks of the IF signal

6. The envelope detector should not be confused with the display detectors. See "Detector types" later in this chapter. Additional information on envelope detectors can be found in Agilent *Application Note 1303, Spectrum Analyzer Measurements and Noise,* literature number 5966-4008E.
7. A signal whose frequency range extends from zero (dc) to some upper frequency determined by the circuit elements. Historically, spectrum analyzers with analog displays used this signal to drive the vertical deflection plates of the CRT directly. Hence it was known as the video signal.

The width of the resolution (IF) filter determines the maximum rate at which the envelope of the IF signal can change. This bandwidth determines how far apart two input sinusoids can be so that after the mixing process they will both be within the filter at the same time. Let's assume a 21.4 MHz final IF and a 100 kHz bandwidth. Two input signals separated by 100 kHz would produce mixing products of 21.35 and 21.45 MHz and would meet the criterion. See Figure 2-16. The detector must be able to follow the changes in the envelope created by these two signals but not the 21.4 MHz IF signal itself.

The envelope detector is what makes the spectrum analyzer a voltmeter. Let's duplicate the situation above and have two equal-amplitude signals in the passband of the IF at the same time. A power meter would indicate a power level 3 dB above either signal, that is, the total power of the two. Assume that the two signals are close enough so that, with the analyzer tuned half way between them, there is negligible attenuation due to the roll-off of the filter[8]. Then the analyzer display will vary between a value that is twice the voltage of either (6 dB greater) and zero (minus infinity on the log scale). We must remember that the two signals are sine waves (vectors) at different frequencies, and so they continually change in phase with respect to each other. At some time they add exactly in phase; at another, exactly out of phase.

So the envelope detector follows the changing amplitude values of the peaks of the signal from the IF chain but not the instantaneous values, resulting in the loss of phase information. This gives the analyzer its voltmeter characteristics.

Digitally implemented resolution bandwidths do not have an analog envelope detector. Instead, the digital processing computes the root sum of the squares of the I and Q data, which is mathematically equivalent to an envelope detector. For more information on digital architecture, refer to Chapter 3.

Displays

Up until the mid-1970s, spectrum analyzers were purely analog. The displayed trace presented a continuous indication of the signal envelope, and no information was lost. However, analog displays had drawbacks. The major problem was in handling the long sweep times required for narrow resolution bandwidths. In the extreme case, the display became a spot that moved slowly across the cathode ray tube (CRT), with no real trace on the display. So a meaningful display was not possible with the longer sweep times.

Agilent Technologies (part of Hewlett-Packard at the time) pioneered a variable-persistence storage CRT in which we could adjust the fade rate of the display. When properly adjusted, the old trace would just fade out at the point where the new trace was updating the display. This display was continuous, had no flicker, and avoided confusing overwrites. It worked quite well, but the intensity and the fade rate had to be readjusted for each new measurement situation. When digital circuitry became affordable in the mid-1970s, it was quickly put to use in spectrum analyzers. Once a trace had been digitized and put into memory, it was permanently available for display. It became an easy matter to update the display at a flicker-free rate without blooming or fading. The data in memory was updated at the sweep rate, and since the contents of memory were written to the display at a flicker-free rate, we could follow the updating as the analyzer swept through its selected frequency span just as we could with analog systems.

8. For this discussion, we assume that the filter is perfectly rectangular.

Detector types

With digital displays, we had to decide what value should be displayed for each display data point. No matter how many data points we use across the display, each point must represent what has occurred over some frequency range and, although we usually do not think in terms of time when dealing with a spectrum analyzer, over some time interval.

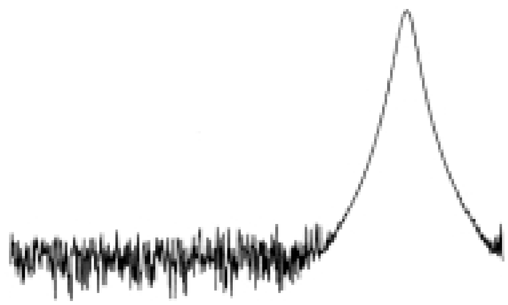


Figure 2-17. When digitizing an analog signal, what value should be displayed at each point?

It is as if the data for each interval is thrown into a bucket and we apply whatever math is necessary to extract the desired bit of information from our input signal. This datum is put into memory and written to the display. This provides great flexibility. Here we will discuss six different detector types.

In Figure 2-18, each bucket contains data from a span and time frame that is determined by these equations:

Frequency:	bucket width = span/(trace points - 1)
Time:	bucket width = sweep time/(trace points - 1)

The sampling rates are different for various instruments, but greater accuracy is obtained from decreasing the span and/or increasing the sweep time since the number of samples per bucket will increase in either case. Even in analyzers with digital IFs, sample rates and interpolation behaviors are designed to be the equivalent of continuous-time processing.

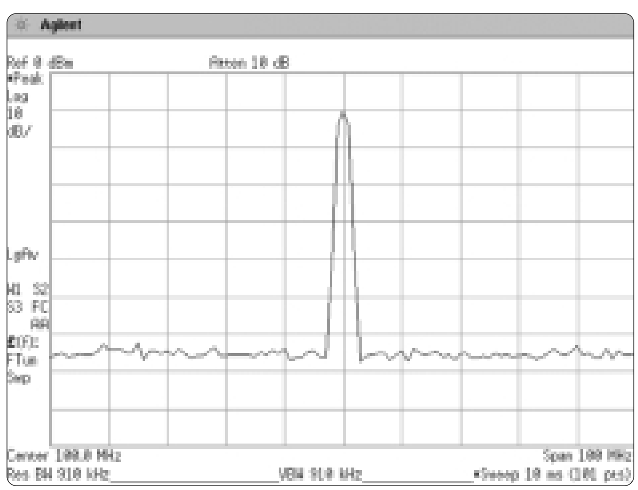


Figure 2-18. Each of the 101 trace points (buckets) covers a 1 MHz frequency span and a 0.1 millisecond time span

The “bucket” concept is important, as it will help us differentiate the six detector types:

- Sample
- Positive peak (also simply called peak)
- Negative peak
- Normal
- Average
- Quasi-peak

The first 3 detectors, *sample, peak,* and *negative peak* are easily understood and visually represented in Figure 2-19. *Normal, average,* and *quasi-peak* are more complex and will be discussed later.

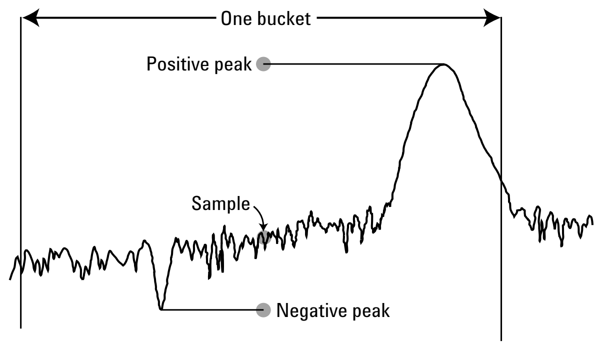


Figure 2-19. Trace point saved in memory is based on detector type algorithm

Let’s return to the question of how to display an analog system as faithfully as possible using digital techniques. Let’s imagine the situation illustrated in Figure 2-17. We have a display that contains only noise and a single CW signal.

Sample detection

As a first method, let us simply select the data point as the instantaneous level at the center of each bucket (see Figure 2-19). This is the *sample* detection mode. To give the trace a continuous look, we design a system that draws vectors between the points. Comparing Figure 2-17 with 2-20, it appears that we get a fairly reasonable display. Of course, the more points there are in the trace, the better the replication of the analog signal will be. The number of available display points can vary for different analyzers. On ESA and PSA Series spectrum analyzers, the number of display points for frequency domain traces can be set from a minimum of 101 points to a maximum of 8192 points. As shown in figure 2-21, more points do indeed get us closer to the analog signal.

Figure 2-20. Sample display mode using ten points to display the signal of Figure 2-17

Figure 2-21. More points produce a display closer to an analog display

While the *sample* detection mode does a good job of indicating the randomness of noise, it is not a good mode for analyzing sinusoidal signals. If we were to look at a 100 MHz comb on an Agilent ESA E4407B, we might set it to span from 0 to 26.5 GHz. Even with 1,001 display points, each display point represents a span (bucket) of 26.5 MHz. This is far wider than the maximum 5 MHz resolution bandwidth.

As a result, the true amplitude of a comb tooth is shown only if its mixing product happens to fall at the center of the IF when the sample is taken. Figure 2-22a shows a 5 GHz span with a 1 MHz bandwidth using *sample* detection. The comb teeth should be relatively equal in amplitude as shown in Figure 2-22b (using *peak* detection). Therefore, *sample* detection does not catch all the signals, nor does it necessarily reflect the true peak values of the displayed signals. When resolution bandwidth is narrower than the sample interval (i.e., the bucket width), the sample mode can give erroneous results.

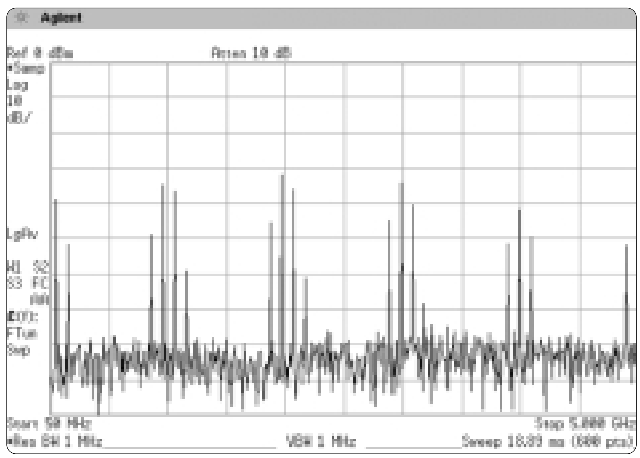


Figure 2-22a. A 5 GHz span of a 100 MHz comb in the sample display mode

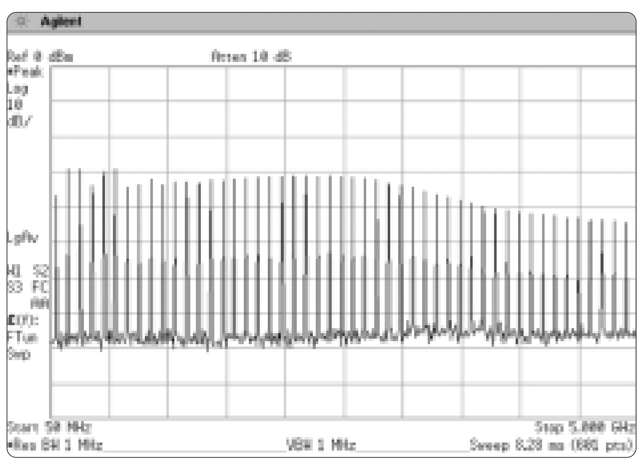


Figure 2-22b. The actual comb over a 500 MHz span using peak (positive) detection

Peak (positive) detection

One way to insure that all sinusoids are reported at their true amplitudes is to display the maximum value encountered in each bucket. This is the *positive peak* detection mode, or *peak*. This is illustrated in Figure 2-22b. *Peak* is the default mode offered on many spectrum analyzers because it ensures that no sinusoid is missed, regardless of the ratio between resolution bandwidth and bucket width. However, unlike *sample* mode, *peak* does not give a good representation of random noise because it only displays the maximum value in each bucket and ignores the true randomness of the noise. So spectrum analyzers that use *peak* detection as their primary mode generally also offer the *sample* mode as an alternative.

Negative peak detection

Negative peak detection displays the minimum value encountered in each bucket. It is generally available in most spectrum analyzers, though it is not used as often as other types of detection. Differentiating CW from impulsive signals in EMC testing is one application where *negative peak* detection is valuable. Later in this application note, we will see how negative peak detection is also used in signal identification routines when using external mixers for high frequency measurements.

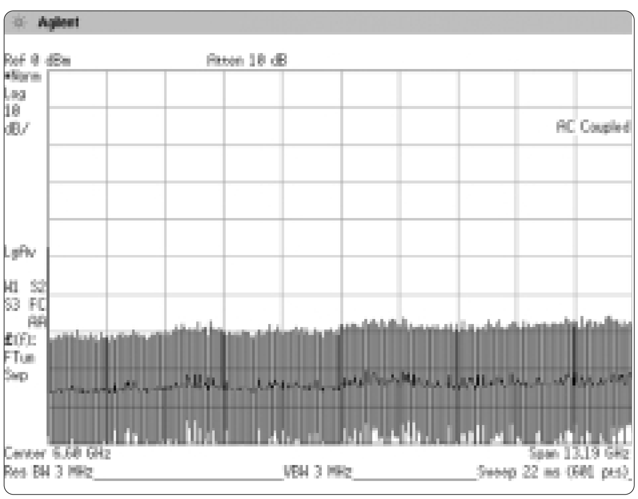


Figure 2-23a. Normal mode

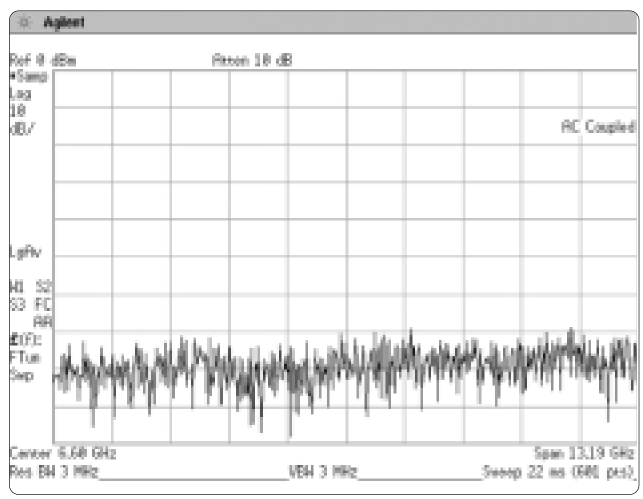


Figure 2-23b. Sample mode

Figure 2-23. Comparison of normal and sample display detection when measuring noise

Normal detection

To provide a better visual display of random noise than *peak* and yet avoid the missed-signal problem of the *sample* mode, the *normal* detection mode (informally known as rosenfell[9]) is offered on many spectrum analyzers. Should the signal both rise and fall, as determined by the positive peak and negative peak detectors, then the algorithm classifies the signal as noise. In that case, an odd-numbered data point displays the maximum value encountered during its bucket. And an even-numbered data point displays the minimum value encountered during its bucket. See Figure 2-25. *Normal* and *sample* modes are compared in Figures 2-23a and 2-23b.[10]

9. rosenfell is not a person's name but rather a description of the algorithm that tests to see if the signal rose and fell within the bucket represented by a given data point. It is also sometimes written as "rose'n'fell".
10. Because of its usefulness in measuring noise, the *sample* detector is usually used in "noise marker" applications. Similarly, the measurement of channel power and adjacent-channel power requires a detector type that gives results unbiased by *peak* detection. For analyzers without averaging detectors, *sample* detection is the best choice.

What happens when a sinusoidal signal is encountered? We know that as a mixing product is swept past the IF filter, an analyzer traces out the shape of the filter on the display. If the filter shape is spread over many display points, then we encounter a situation in which the displayed signal only rises as the mixing product approaches the center frequency of the filter and only falls as the mixing product moves away from the filter center frequency. In either of these cases, the pos-peak and neg-peak detectors sense an amplitude change in only one direction, and, according to the normal detection algorithm, the maximum value in each bucket is displayed. See Figure 2-24.

What happens when the resolution bandwidth is narrow, relative to a bucket? The signal will both rise and fall during the bucket. If the bucket happens to be an odd-numbered one, all is well. The maximum value encountered in the bucket is simply plotted as the next data point. However, if the bucket is even-numbered, then the minimum value in the bucket is plotted. Depending on the ratio of resolution bandwidth to bucket width, the minimum value can differ from the true peak value (the one we want displayed) by a little or a lot. In the extreme, when the bucket is much wider than the resolution bandwidth, the difference between the maximum and minimum values encountered in the bucket is the full difference between the peak signal value and the noise. This is true for the example in Figure 2-25. See bucket 6. The peak value of the previous bucket is always compared to that of the current bucket. The greater of the two values is displayed if the bucket number is odd as depicted in bucket 7. The signal peak actually occurs in bucket 6 but is not displayed until bucket 7.

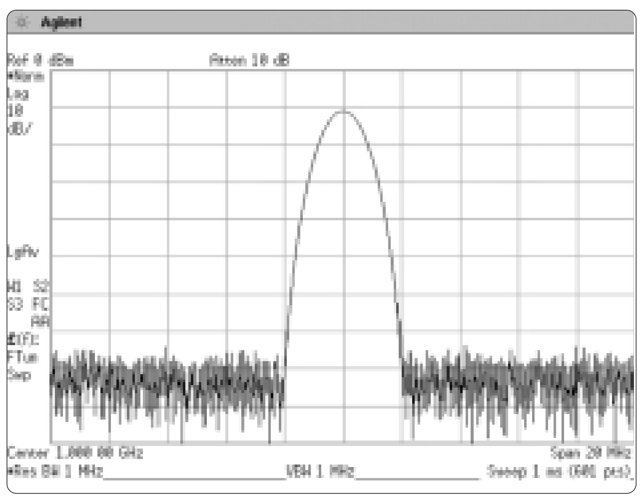


Figure 2-24. Normal detection displays maximum values in buckets where signal only rises or only falls

The *normal* detection algorithm:

> If the signal rises and falls within a bucket:
>
> > Even numbered buckets display the minimum (negative peak) value in the bucket. The maximum is remembered.
> >
> > Odd numbered buckets display the maximum (positive peak) value determined by comparing the current bucket peak with the previous (remembered) bucket peak.
>
> If the signal only rises or only falls within a bucket, the peak is displayed. See Figure 2-25.

This process may cause a maximum value to be displayed one data point too far to the right, but the offset is usually only a small percentage of the span. Some spectrum analyzers, such as the Agilent PSA Series, compensate for this potential effect by moving the LO start and stop frequencies.

Another type of error is where two peaks are displayed when only one actually exists. Figure 2-26 shows what might happen in such a case. The outline of the two peaks is displayed using peak detection with a wider RBW.

So *peak* detection is best for locating CW signals well out of the noise. *Sample* is best for looking at noise, and *normal* is best for viewing signals and noise.

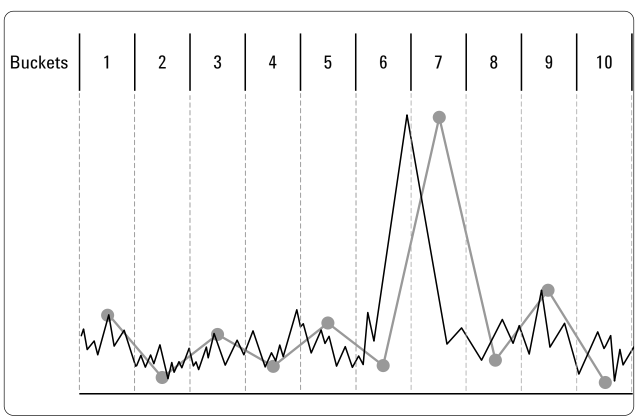


Figure 2-25. Trace points selected by the normal detection algorithm

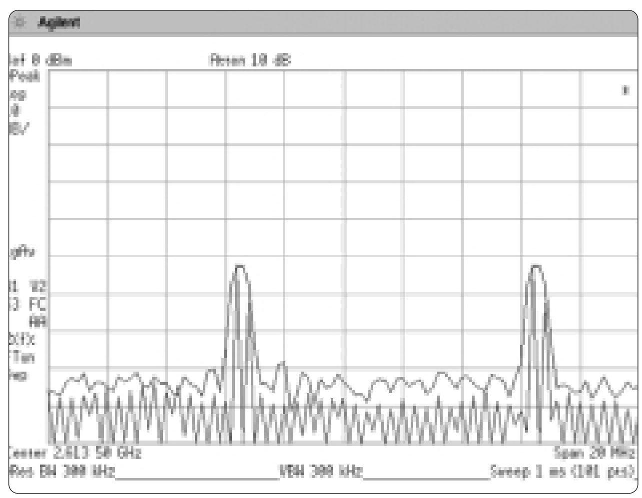


Figure 2-26. Normal detection shows two peaks when actually only one peak exists

Average detection

Although modern digital modulation schemes have noise-like characteristics, *sample* detection does not always provide us with the information we need. For instance, when taking a channel power measurement on a W-CDMA signal, integration of the rms values is required. This measurement involves summing power across a range of analyzer frequency buckets. *Sample* detection does not provide this.

While spectrum analyzers typically collect amplitude data many times in each bucket, sample detection keeps only one of those values and throws away the rest. On the other hand, an averaging detector uses all the data values collected within the time (and frequency) interval of a bucket. Once we have digitized the data, and knowing the circumstances under which they were digitized, we can manipulate the data in a variety of ways to achieve the desired results.

Some spectrum analyzers refer to the averaging detector as an rms detector when it averages power (based on the root mean square of voltage). Agilent PSA and ESA Series analyzers have an *average* detector that can average the power, voltage, or log of the signal by including a separate control to select the averaging type:

Power (rms) averaging averages rms levels, by taking the square root of the sum of the squares of the voltage data measured during the bucket interval, divided by the characteristic input impedance of the spectrum analyzer, normally 50 ohms. Power averaging calculates the true average power, and is best for measuring the power of complex signals.

Voltage averaging averages the linear voltage data of the envelope signal measured during the bucket interval. It is often used in EMI testing for measuring narrowband signals (this will be discussed further in the next section). Voltage averaging is also useful for observing rise and fall behavior of AM or pulse-modulated signals such as radar and TDMA transmitters.

Log-power (video) averaging averages the logarithmic amplitude values (dB) of the envelope signal measured during the bucket interval. Log power averaging is best for observing sinusoidal signals, especially those near noise.[11]

Thus, using the average detector with the averaging type set to power provides true average power based upon rms voltage, while the average detector with the averaging type set to voltage acts as a general-purpose average detector. The average detector with the averaging type set to log has no other equivalent.

Average detection is an improvement over using sample detection for the determination of power. Sample detection requires multiple sweeps to collect enough data points to give us accurate average power information. Average detection changes channel power measurements from being a summation over a range of buckets into integration over the time interval representing a range of frequencies in a swept analyzer. In a fast Fourier transfer (FFT) analyzer[12], the summation used for channel power measurements changes from being a summation over display buckets to being a summation over FFT bins. In both swept and FFT cases, the integration captures all the power information available, rather than just that which is sampled by the sample detector. As a result, the average detector has a lower variance result for the same measurement time. In swept analysis, it also allows the convenience of reducing variance simply by extending the sweep time.

11. See Chapter 5, "Sensitivity and Noise."
12. Refer to Chapter 3 for more information on the FFT analyzers. They perform math computations on many buckets simultaneously, which improves the measurement speed.

EMI detectors: average and quasi-peak detection

An important application of *average* detection is for characterizing devices for electromagnetic interference (EMI). In this case, voltage averaging, as described in the previous section, is used for measurement of narrowband signals that might be masked by the presence of broadband impulsive noise. The average detection used in EMI instruments takes an envelope-detected signal and passes it through a low-pass filter with a bandwidth much less than the RBW. The filter integrates (averages) the higher frequency components such as noise. To perform this type of detection in an older spectrum analyzer that doesn't have a built-in voltage averaging detector function, set the analyzer in linear mode and select a video filter with a cut-off frequency below the lowest PRF of the measured signal.

Quasi-peak detectors (QPD) are also used in EMI testing. QPD is a weighted form of peak detection. The measured value of the QPD drops as the repetition rate of the measured signal decreases. Thus, an impulsive signal with a given peak amplitude and a 10 Hz pulse repetition rate will have a lower quasi-peak value than a signal with the same peak amplitude but having a 1 kHz repetition rate. This signal weighting is accomplished by circuitry with specific charge, discharge, and display time constants defined by CISPR[13].

QPD is a way of measuring and quantifying the "annoyance factor" of a signal. Imagine listening to a radio station suffering from interference. If you hear an occasional "pop" caused by noise once every few seconds, you can still listen to the program without too much trouble. However, if that same amplitude pop occurs 60 times per second, it becomes extremely annoying, making the radio program intolerable to listen to.

Averaging processes

There are several processes in a spectrum analyzer that smooth the variations in the envelope-detected amplitude. The first method, average detection, was discussed previously. Two other methods, *video filtering* and *trace averaging*, are discussed next.[14]

13. CISPR, the International Special Committee on Radio Interference, was established in 1934 by a group of international organizations to address radio interference. CISPR is a non-governmental group composed of National Committees of the International Electrotechnical Commission (IEC), as well as numerous international organizations. CISPR's recommended standards generally form the basis for statutory EMC requirements adopted by governmental regulatory agencies around the world.
14. A fourth method, called a noise marker, is discussed in Chapter 5, *"Sensitivity and Noise"*. A more detailed discussion can be found in *Application Note 1303, Spectrum Analyzer Measurements and Noise,* literature number 5966-4008E.

Video filtering

Discerning signals close to the noise is not just a problem when performing EMC tests. Spectrum analyzers display signals plus their own internal noise, as shown in Figure 2-27. To reduce the effect of noise on the displayed signal amplitude, we often smooth or average the display, as shown in Figure 2-28. Spectrum analyzers include a variable video filter for this purpose. The video filter is a low-pass filter that comes after the envelope detector and determines the bandwidth of the video signal that will later be digitized to yield amplitude data. The cutoff frequency of the video filter can be reduced to the point where it becomes smaller than the bandwidth of the selected resolution bandwidth (IF) filter. When this occurs, the video system can no longer follow the more rapid variations of the envelope of the signal(s) passing through the IF chain. The result is an averaging or smoothing of the displayed signal.

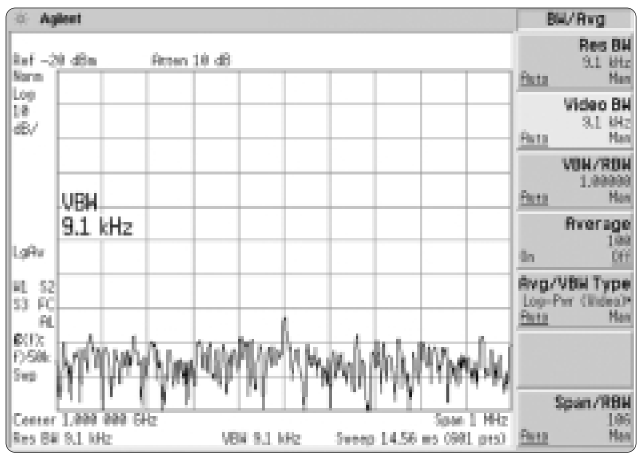


Figure 2-27. Spectrum analyzers display signal plus noise

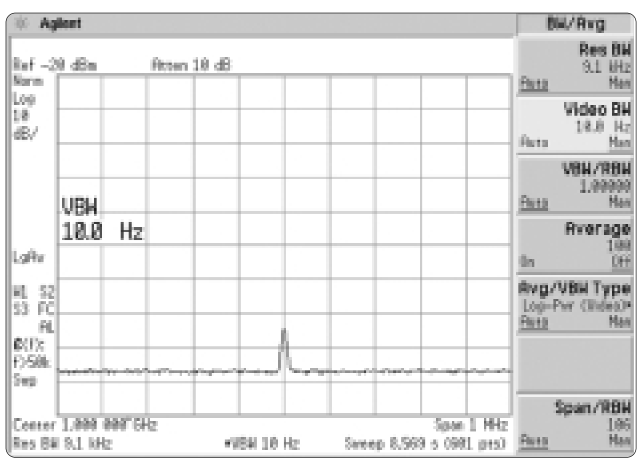


Figure 2-28. Display of figure 2-27 after full smoothing

The effect is most noticeable in measuring noise, particularly when a wide resolution bandwidth is used. As we reduce the video bandwidth, the peak-to-peak variations of the noise are reduced. As Figure 2-29 shows, the degree of reduction (degree of averaging or smoothing) is a function of the ratio of the video to resolution bandwidths. At ratios of 0.01 or less, the smoothing is very good. At higher ratios, the smoothing is not so good. The video filter does not affect any part of the trace that is already smooth (for example, a sinusoid displayed well out of the noise).

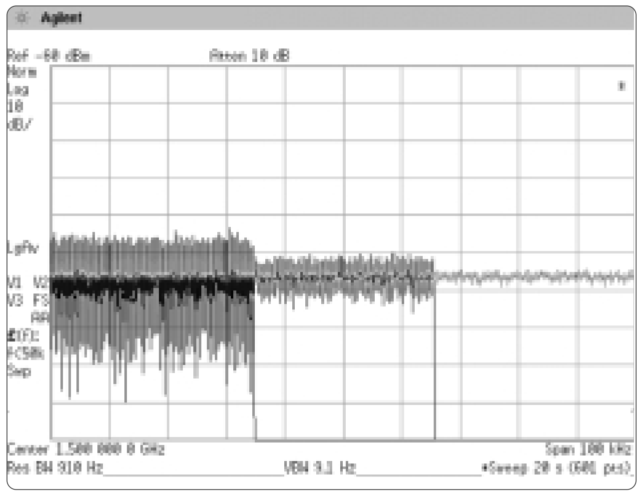


Figure 2-29. Smoothing effect of VBW-to-RBW ratios of 3:1, 1:10, and 1:100

If we set the analyzer to *positive peak* detection mode, we notice two things: First, if VBW > RBW, then changing the resolution bandwidth does not make much difference in the peak-to-peak fluctuations of the noise. Second, if VBW < RBW, then changing the video bandwidth seems to affect the noise level. The fluctuations do not change much because the analyzer is displaying only the peak values of the noise. However, the noise level appears to change with video bandwidth because the averaging (smoothing) changes, thereby changing the peak values of the smoothed noise envelope. See Figure 2-30a. When we select *average* detection, we see the average noise level remains constant. See Figure 2-30b.

Because the video filter has its own response time, the sweep time increases approximately inversely with video bandwidth when the VBW is less than the resolution bandwidth. The sweep time can therefore be described by this equation:

$$ST \approx \frac{k(\text{Span})}{(\text{RBW})(\text{VBW})}$$

The analyzer sets the sweep time automatically to account for video bandwidth as well as span and resolution bandwidth.

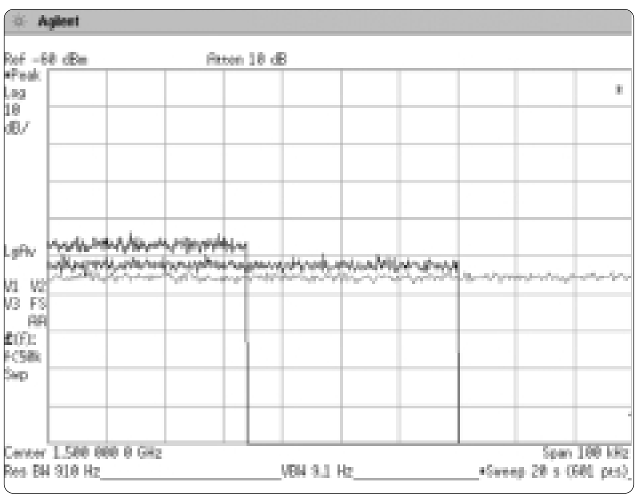


Figure 2-30a. Positive peak detection mode; reducing video bandwidth lowers peak noise but not average noise

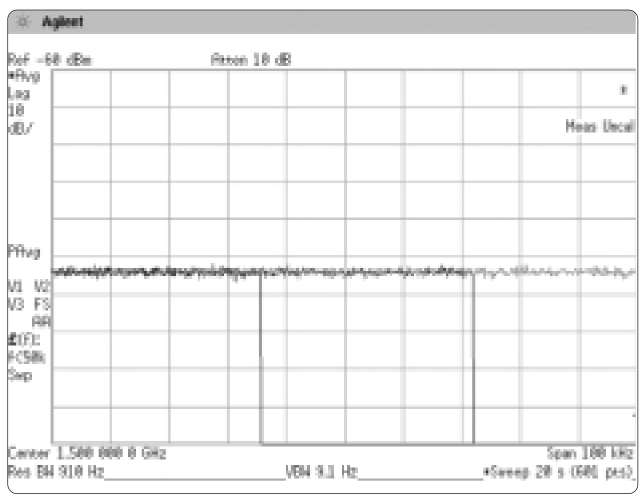


Figure 2-30b. Average detection mode; noise level remains constant, regardless of VBW-to-RBW ratios (3:1, 1:10, and 1:100)

Trace Averaging

Digital displays offer another choice for smoothing the display: trace averaging. This is a completely different process than that performed using the *average* detector. In this case, averaging is accomplished over two or more sweeps on a point-by-point basis. At each display point, the new value is averaged in with the previously averaged data:

$$A_{avg} = \left(\frac{n-1}{n}\right) A_{prior\ avg} + \left(\frac{1}{n}\right) A_n$$

where A_{avg} = new average value
$A_{prior\ avg}$ = average from prior sweep
A_n= measured value on current sweep
n = number of current sweep

Thus, the display gradually converges to an average over a number of sweeps. As with video filtering, we can select the degree of averaging or smoothing. We do this by setting the number of sweeps over which the averaging occurs. Figure 2-31 shows trace averaging for different numbers of sweeps. While trace averaging has no effect on sweep time, the time to reach a given degree of averaging is about the same as with video filtering because of the number of sweeps required.

In many cases, it does not matter which form of display smoothing we pick. If the signal is noise or a low-level sinusoid very close to the noise, we get the same results with either video filtering or trace averaging. However, there is a distinct difference between the two. Video filtering performs averaging in real time. That is, we see the full effect of the averaging or smoothing at each point on the display as the sweep progresses. Each point is averaged only once, for a time of about 1/VBW on each sweep. Trace averaging, on the other hand, requires multiple sweeps to achieve the full degree of averaging, and the averaging at each point takes place over the full time period needed to complete the multiple sweeps.

As a result, we can get significantly different results from the two averaging methods on certain signals. For example, a signal with a spectrum that changes with time can yield a different average on each sweep when we use video filtering. However, if we choose trace averaging over many sweeps, we will get a value much closer to the true average. See Figures 2-32a and b.

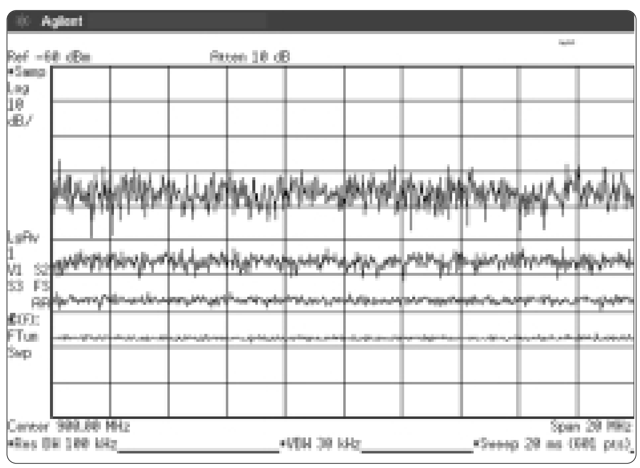


Figure 2-31. Trace averaging for 1, 5, 20, and 100 sweeps, top to bottom (trace position offset for each set of sweeps)

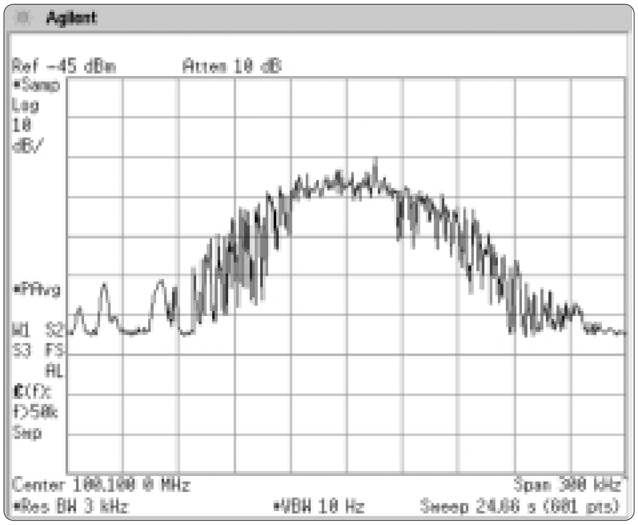


Figure 2-32a. Video filtering

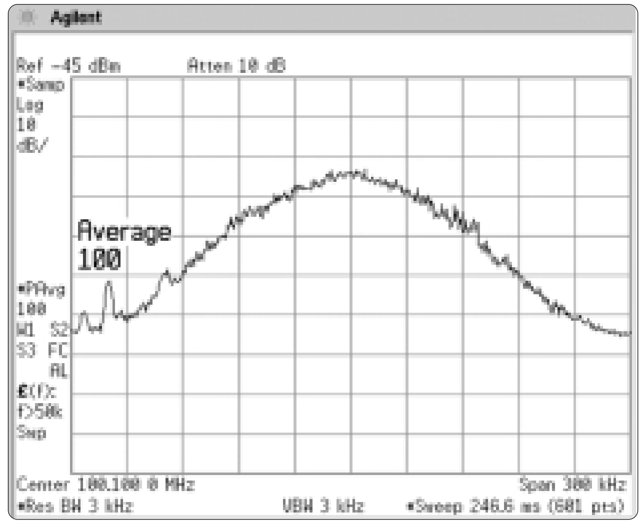


Figure 2-32b. Trace averaging

Figure 2-32. Video filtering and trace averaging yield different results on FM broadcast signal

Time gating

Time-gated spectrum analysis allows you to obtain spectral information about signals occupying the same part of the frequency spectrum that are separated in the time domain. Using an external trigger signal to coordinate the separation of these signals, you can perform the following operations:

- Measure any one of several signals separated in time; for example, you can separate the spectra of two radios time-sharing a single frequency
- Measure the spectrum of a signal in one time slot of a TDMA system
- Exclude the spectrum of interfering signals, such as periodic pulse edge transients that exist for only a limited time

Why time gating is needed

Traditional frequency-domain spectrum analysis provides only limited information for certain signals. Examples of these difficult-to-analyze signals include the following signal types:

- Pulsed RF
- Time multiplexed
- Time domain multiple access (TDMA)
- Interleaved or intermittent
- Burst modulated

In some cases, time-gating capability enables you to perform measurements that would otherwise be very difficult, if not impossible. For example, consider Figure 2-33a, which shows a simplified digital mobile-radio signal in which two radios, #1 and #2, are time-sharing a single frequency channel. Each radio transmits a single 1 ms burst, and then shuts off while the other radio transmits for 1 ms. The challenge is to measure the unique frequency spectrum of each transmitter.

Unfortunately, a traditional spectrum analyzer cannot do that. It simply shows the combined spectrum, as seen in Figure 2-33b. Using the time-gate capability and an external trigger signal, you can see the spectrum of just radio #1 (or radio #2 if you wished) and identify it as the source of the spurious signal shown, as in Figure 2-33c.

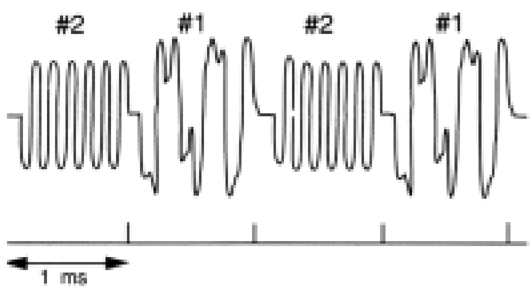


Figure 2-33a. Simplified digital mobile-radio signal in time domain

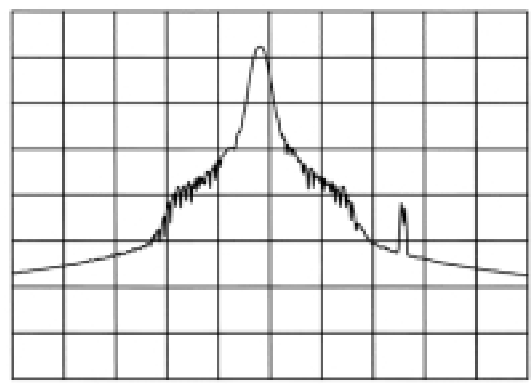


Figure 2-33c. Time-gated spectrum of signal #1 identifies it as the source of spurious emission

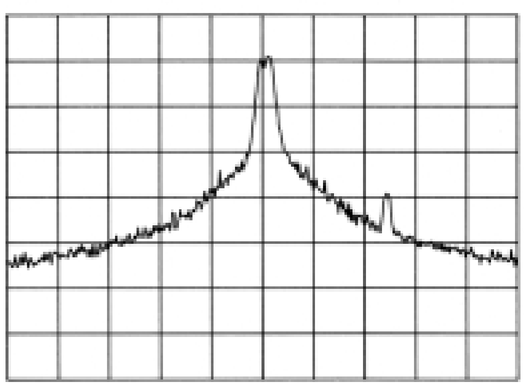


Figure 2-33b. Frequency spectrum of combined signals. Which radio produces the spurious emissions?

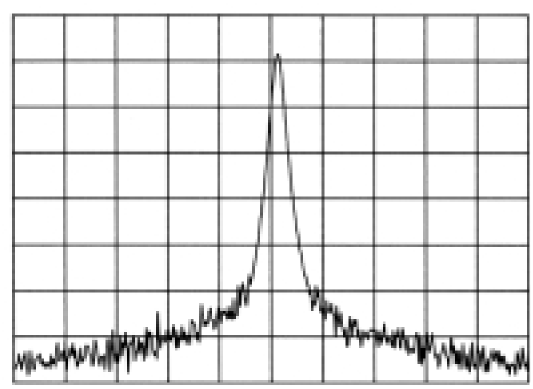


Figure 2-33d. Time-gated spectrum of signal #2 shows it is free of spurious emissions

Time gating can be achieved using three different methods that will be discussed below. However, there are certain basic concepts of time gating that apply to any implementation. In particular, you must have, or be able to set, the following four items:

- An externally supplied gate trigger signal
- The gate control, or trigger mode (edge, or level)
- The *gate delay* setting, which determines how long after the trigger signal the gate actually becomes active and the signal is observed
- The *gate length* setting, which determines how long the gate is on and the signal is observed

Controlling these parameters will allow us to look at the spectrum of the signal during a desired portion of the time. If you are fortunate enough to have a gating signal that is only true during the period of interest, then you can use level gating as shown in Figure 2-34. However, in many cases the gating signal will not perfectly coincide with the time we want to measure the spectrum. Therefore, a more flexible approach is to use edge triggering in conjunction with a specified gate delay and gate length to precisely define the time period in which to measure the signal.

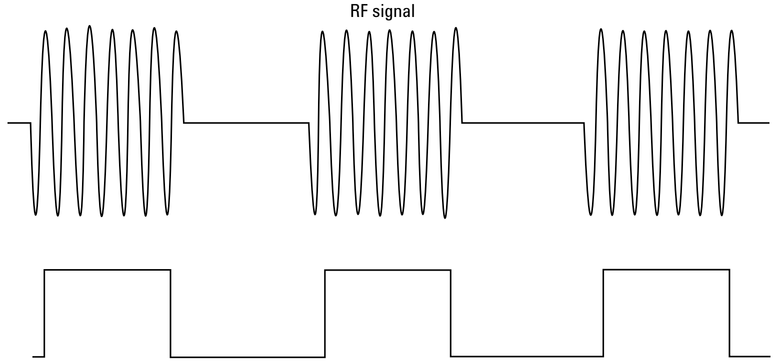


Figure 2-34. Level triggering: the spectrum analyzer only measures the frequency spectrum when gate trigger signal is above a certain level

Consider the GSM signal with eight time slots in Figure 2-35. Each burst is 0.577 ms and the full frame is 4.615 ms. We may be interested in the spectrum of the signal during a specific time slot. For the purposes of this example, let's assume that we are using only two of the eight available time slots, as shown in Figure 2-36. When we look at this signal in the frequency domain in Figure 2-37, we observe an unwanted spurious signal present in the spectrum. In order to troubleshoot the problem and find the source of this interfering signal, we need to determine the time slot in which it is occurring. If we wish to look at time slot 2, we set up the gate to trigger on the rising edge of burst 0, then specify a gate delay of 1.3 ms and a gate length of 0.3 ms, as shown in Figure 2-38. The gate delay assures that we only measure the spectrum of time slot 2 while the burst is fully on. Note that the gate delay value is carefully selected to avoid the rising edge of the burst, since we want to allow time for the RBW filtered signal to settle out before we make a measurement. Similarly, the gate length is chosen to avoid the falling edges of the burst. Figure 2-39 shows the spectrum of time slot 2, which reveals that the spurious signal is NOT caused by this burst.

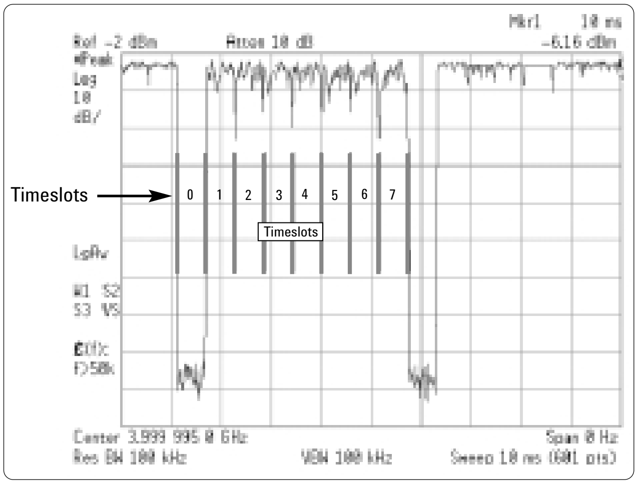


Figure 2-35. A TDMA format signal (in this case, GSM) with eight time slots

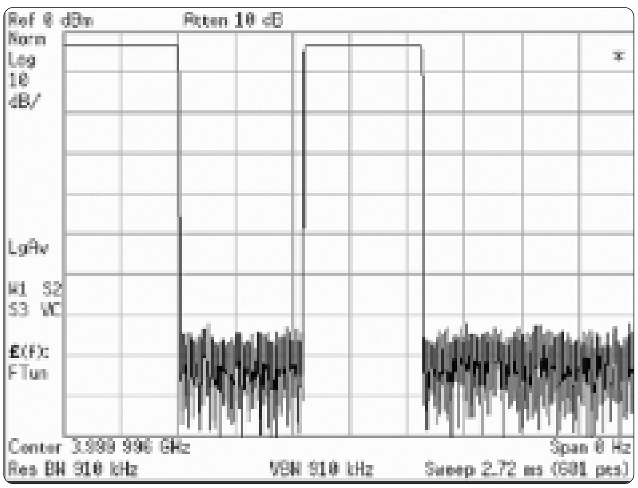


Figure 2-36. A zero span (time domain) view of the two time slots

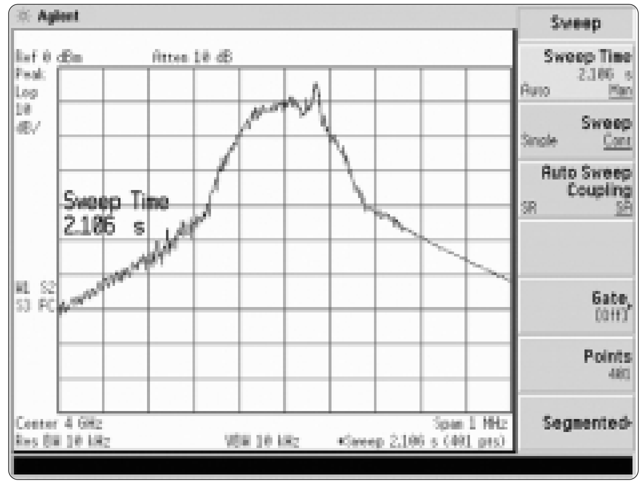


Figure 2-37. The signal in the frequency domain

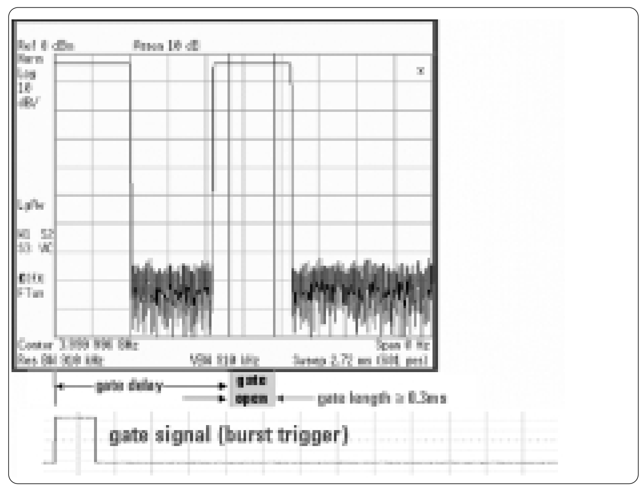


Figure 2-38. Time gating is used to look at the spectrum of time slot 2

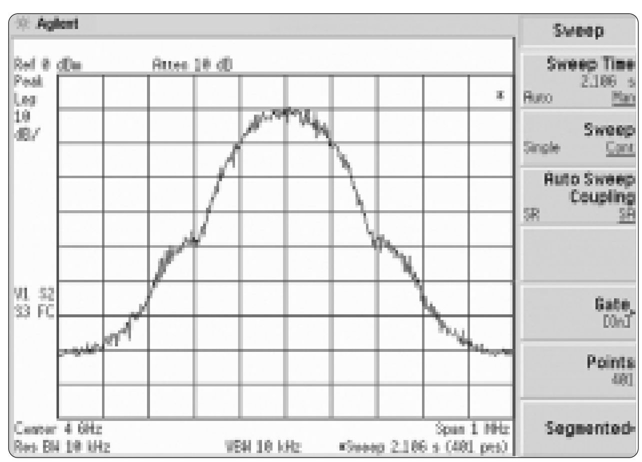


Figure 2-39. Spectrum of the pulse in time slot 2

There are three common methods used to perform time gating:

- Gated FFT
- Gated video
- Gated sweep

Gated FFT

Some spectrum analyzers, such as the Agilent PSA Series, have built-in FFT capabilities. In this mode, the data is acquired for an FFT starting at a chosen delay following a trigger. The IF signal is digitized and captured for a time period of 1.83 divided by resolution bandwidth. An FFT is computed based on this data acquisition and the results are displayed as the spectrum. Thus, the spectrum is that which existed at a particular time of known duration. This is the fastest gating technique whenever the span is not wider than the FFT maximum width, which for PSA is 10 MHz.

To get the maximum possible frequency resolution, choose the narrowest available RBW whose capture time fits within the time period of interest. That may not always be needed, however, and you could choose a wider RBW with a corresponding narrower gate length. The minimum usable RBW in gated FFT applications is always lower than the minimum usable RBW in other gating techniques, because the IF must fully settle during the burst in other techniques, which takes longer than 1.83 divided by RBW.

Gated video

Gated video is the analysis technique used in a number of spectrum analyzers, including the Agilent 8560, 8590 and ESA Series. In this case, the video voltage is switched off, or to "negative infinity decibels" during the time the gate is supposed to be in its "blocked" mode. The detector is set to *peak detection*. The sweep time must be set so that the gates occur at least once per display point, or bucket, so that the peak detector is able to see real data during that time interval. Otherwise, there will be trace points with no data, resulting in an incomplete spectrum. Therefore, the minimum sweep time is N display buckets times burst cycle time. For example, in GSM measurements, the full frame lasts 4.615 ms. For an ESA spectrum analyzer set to its default value of 401 display points, the minimum sweep time for GSM gated video measurements would be 401 times 4.615 ms or 1.85 s. Some TDMA formats have cycle times as large as 90 ms, resulting in long sweep times using the gated video technique.

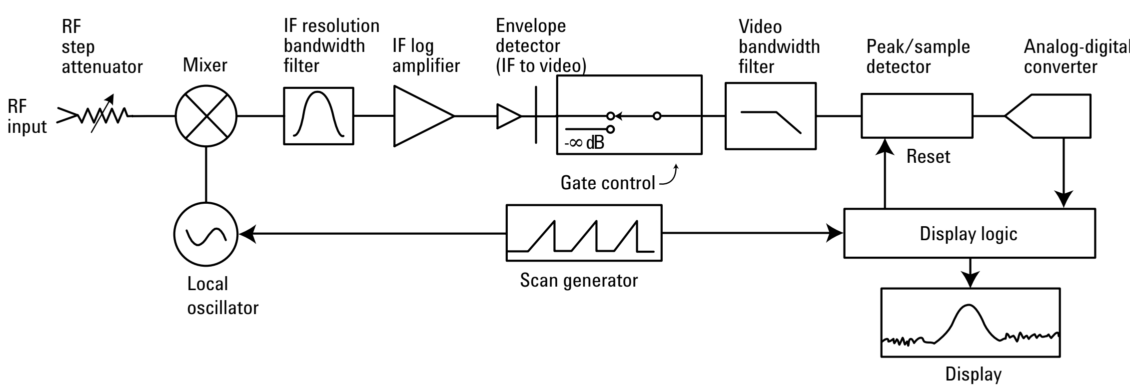


Figure 2-40. Block diagram of a spectrum analyzer with gated video

Gated sweep

Gated sweep, sometimes referred to as gated LO, is the final technique. In gated sweep mode, we control the voltage ramp produced by the scan generator to sweep the LO. This is shown in figure 2-41. When the gate is active, the LO ramps up in frequency like any spectrum analyzer. When the gate is blocked, the voltage out of the scan generator is frozen, and the LO stops rising in frequency. This technique can be much faster than gated video because multiple buckets can be measured during each burst. As an example, let's use the same GSM signal described in the gated video discussion earlier in this chapter. Using a PSA Series spectrum analyzer, a standard, non-gated, spectrum sweep over a 1 MHz span takes 14.6 ms, as shown in Figure 2-42. With a gate length of 0.3 ms, the spectrum analyzer sweep must be built up in 49 gate intervals (14.6 divided by 0.3), or. If the full frame of the GSM signal is 4.615 ms, then the total measurement time is 49 intervals times 4.615 ms = 226 ms. This represents a significant improvement in speed compared to the gated video technique which required 1.85 s for 401 data points. Gated sweep is available on the PSA Series spectrum analyzers.

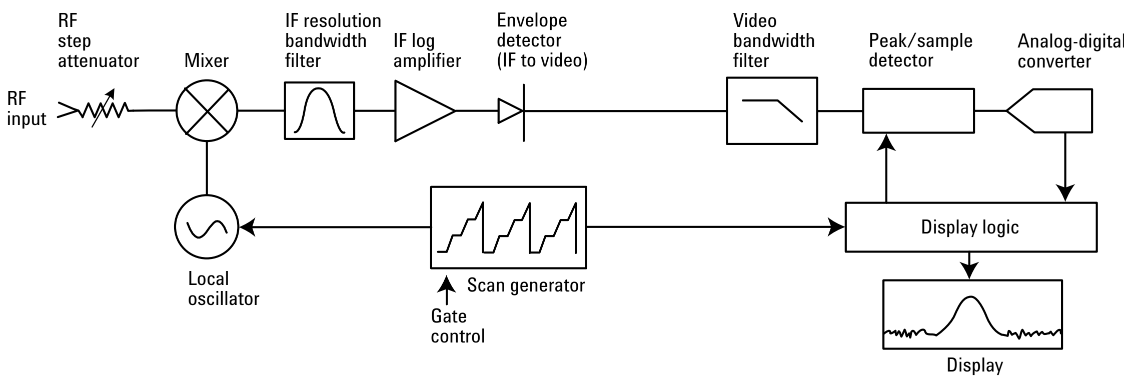


Figure 2-41. In gated sweep mode, the LO sweeps only during gate interval

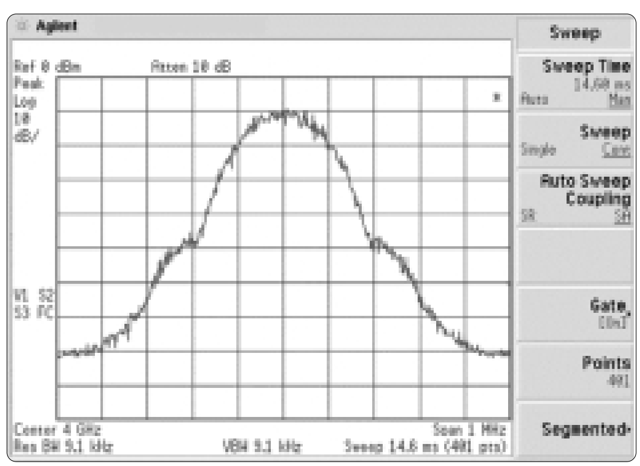


Figure 2-42. Spectrum of the GSM signal

Chapter 3
Digital IF Overview

Since the 1980's, one of the most profound areas of change in spectrum analysis has been the application of digital technology to replace portions of the instrument that had previously been implemented as analog circuits. With the availability of high-performance analog-to-digital converters, the latest spectrum analyzers digitize incoming signals much earlier in the signal path compared to spectrum analyzer designs of just a few years ago. The change has been most dramatic in the IF section of the spectrum analyzer. Digital IFs[1] have had a great impact on spectrum analyzer performance, with significant improvements in speed, accuracy, and the ability to measure complex signals through the use of advanced DSP techniques.

Digital filters

A partial implementation of digital IF circuitry is implemented in the Agilent ESA-E Series spectrum analyzers. While the 1 kHz and wider RBWs are implemented with traditional analog LC and crystal filters, the narrowest bandwidths (1 Hz to 300 Hz) are realized using digital techniques. As shown in Figure 3-1, the linear analog signal is mixed down to an 8.5 kHz IF and passed through a bandpass filter only 1 kHz wide. This IF signal is amplified, then sampled at an 11.3 kHz rate and digitized.

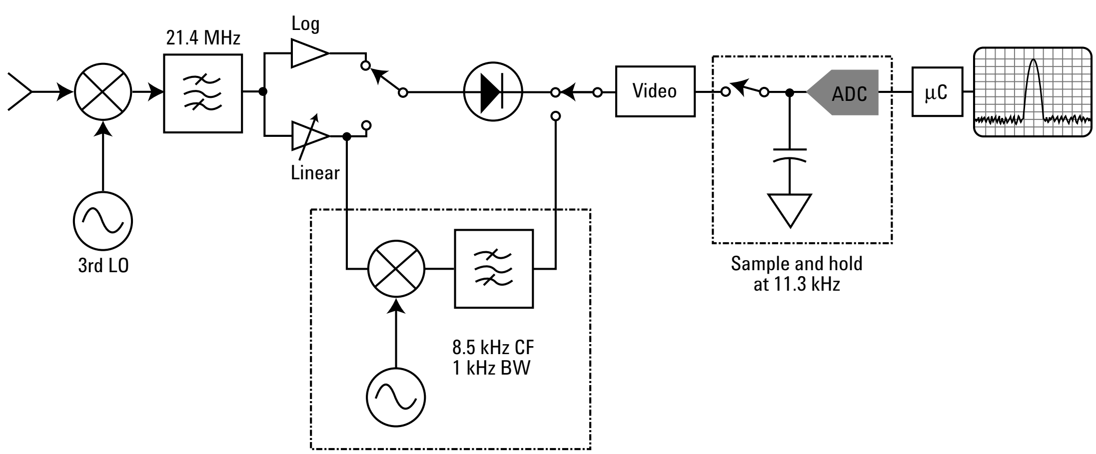


Figure 3-1. Digital implementation of 1, 3, 10, 30, 100, and 300 Hz resolution filters in ESA-E Series

Once in digital form, the signal is put through a fast Fourier transform algorithm. To transform the appropriate signal, the analyzer must be fixed-tuned (not sweeping). That is, the transform must be done on a time-domain signal. Thus the ESA-E Series analyzers step in 900 Hz increments, instead of sweeping continuously, when we select one of the digital resolution bandwidths. This stepped tuning can be seen on the display, which is updated in 900 Hz increments as the digital processing is completed.

As we shall see in a moment, other spectrum analyzers, such as the PSA Series, use an all-digital IF, implementing all resolution bandwidth filters digitally.

A key benefit of the digital processing done in these analyzers is a bandwidth selectivity of about 4:1. This selectivity is available on the narrowest filters, the ones we would be choosing to separate the most closely spaced signals.

1. Strictly speaking, once a signal has been digitized, it is no longer at an intermediate frequency, or IF. At that point, the signal is represented by digital data values. However, we use the term "digital IF" to describe the digital processing that replaces the analog IF processing found in traditional spectrum analyzers.

In Chapter 2, we did a filter skirt selectivity calculation for two signals spaced 4 kHz apart, using a 3 kHz analog filter. Let's repeat that calculation using digital filters. A good model of the selectivity of digital filters is a near-Gaussian model:

$$H(\Delta f) = -3.01 \text{ dB x} \left[\frac{\Delta f}{RBW/2}\right]^{\alpha}$$

where $H(\Delta f)$ is the filter skirt rejection in dB
Δf is the frequency offset from the center in Hz, and
α is a parameter that controls selectivity. $\alpha = 2$ for an ideal Gaussian filter. The swept RBW filters used in Agilent spectrum analyzers are based on a near-Gaussian model with an α value equal to 2.12, resulting in a selectivity ratio of 4.1:1.

Entering the values from our example into the equation, we get:

$$H(4 \text{ kHz}) = -3.01 \text{ dB x} \left[\frac{4000}{3000/2}\right]^{2.12}$$

$$= -24.1 \text{ dB}$$

At an offset of 4 kHz, the 3 kHz digital filter is down –24.1 dB compared to the analog filter which was only down –14.8 dB. Because of its superior selectivity, the digital filter can resolve more closely spaced signals.

The all-digital IF

The Agilent PSA Series spectrum analyzers have, for the first time, combined several digital techniques to achieve the all-digital IF. The all-digital IF brings a wealth of advantages to the user. The combination of FFT analysis for narrow spans and swept analysis for wider spans optimizes sweeps for the fastest possible measurements. Architecturally, the ADC is moved closer to the input port, a move made possible by improvements to the A-to-D converters and other digital hardware. Let's begin by taking a look at the block diagram of the all-digital IF in the PSA spectrum analyzer, as shown in Figure 3-2.

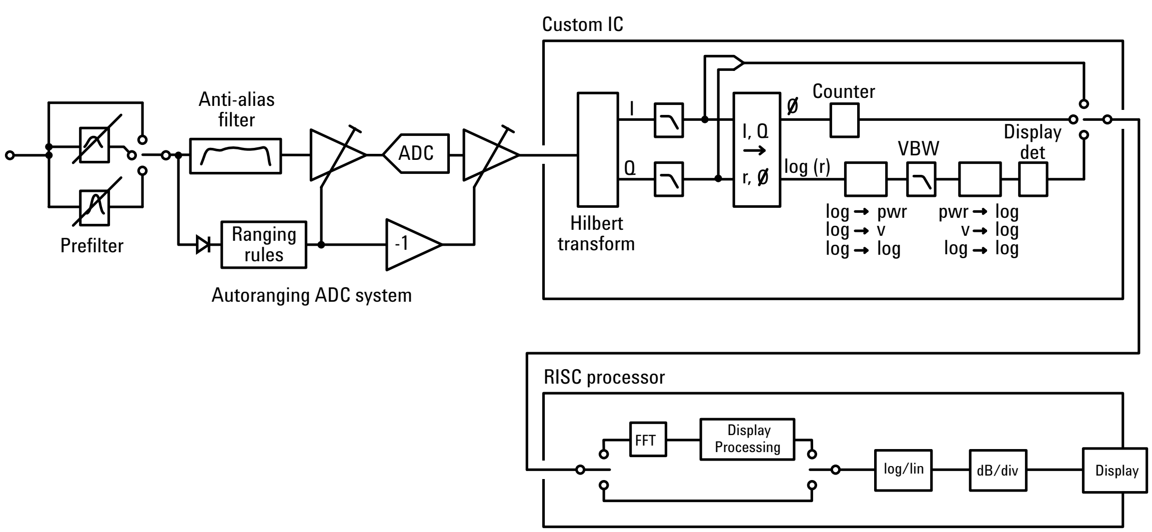


Figure 3-2. Block diagram of the all-digital IF in the Agilent PSA Series

In this case, all 160 resolution bandwidths are digitally implemented. However, there is some analog circuitry prior to the ADC, starting with several stages of down conversion, followed by a pair of single-pole prefilters (one an LC filter, the other crystal-based). A prefilter helps prevent succeeding stages from contributing third-order distortion in the same way a prefilter would in an analog IF. In addition, it enables dynamic range extension via autoranging. The output of the single-pole prefilter is routed to the autorange detector and the anti-alias filter.

As with any FFT-based IF architecture, the anti-alias filter is required to prevent aliasing (the folding of out-of-band signals into the ADC sampled data). This filter has many poles, and thus has substantial group delay. Even a very fast rising RF burst, downconverted to the IF frequency, will experience a delay of more than three cycles of the ADC clock (30 MHz) through the anti-alias filter. The delay allows time for an impending large signal to be recognized before it overloads the ADC. The logic circuitry controlling the autorange detector will decrease the gain in front of the ADC before a large signal reaches it, thus preventing clipping. If the signal envelope remains small for a long time, the autoranging circuit increases the gain, reducing the effective noise at the input. The digital gain after the ADC is also changed to compensate for the analog gain in front of it. The result is a “floating point” ADC with very wide dynamic range when autoranging is enabled in swept mode.

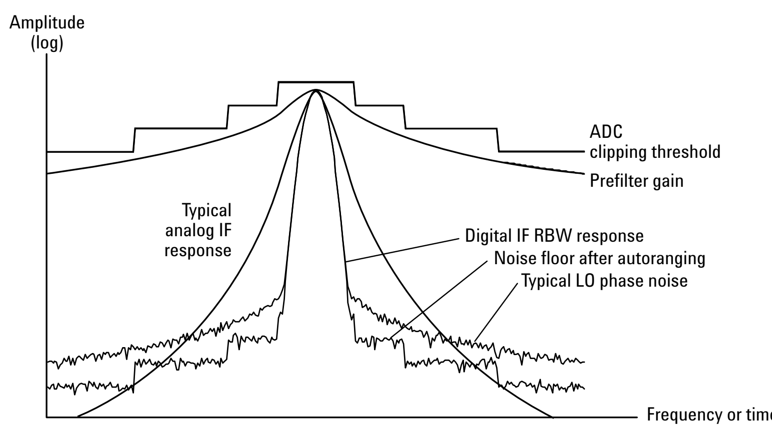


Figure 3-3. Autoranging keeps ADC noise close to carrier and lower than LO noise or RBW filter response

Figure 3-3 illustrates the sweeping behavior of the PSA analyzer. The single-pole prefilter allows the gain to be turned up high when the analyzer is tuned far from the carrier. As the carrier gets closer, the gain falls and the ADC quantization noise rises. The noise level will depend on the signal level frequency separation from the carrier, so it looks like a step-shaped phase noise. However, phase noise is different from this autoranging noise. Phase noise cannot be avoided in a spectrum analyzer. However, reducing the prefilter width can reduce autoranging noise at most frequency offsets from the carrier. Since the prefilter width is approximately 2.5 times the RBW, reducing the RBW reduces the autoranging noise.

Custom signal processing IC

Turning back to the block diagram of the digital IF (Figure 3-2), after the ADC gain has been set with analog gain and corrected with digital gain, a custom IC begins processing the samples. First, it splits the 30 MHz IF samples into I and Q pairs at half the rate (15 Mpairs/s). The I and Q pairs are given a high-frequency boost with a single-stage digital filter that has gain and phase approximately opposite to that of the single pole analog prefilter. Next, I and Q signals are low-pass filtered with a linear-phase filter with nearly ideal Gaussian response. Gaussian filters have always been used for swept spectrum analysis, because of their optimum compromise between frequency domain performance (shape factor) and time-domain performance (response to rapid sweeps). With the signal bandwidth now reduced, the I and Q pairs may be decimated and sent to the processor for FFT processing or demodulation. Although FFTs can be performed to cover a segment of frequency span up to the 10 MHz bandwidth of the anti-alias filter, even a narrower FFT span, such as 1 kHz, with a narrow RBW, such as 1 Hz, would require FFTs with 20 million data points. Using decimation for narrower spans, the number of data points needed to compute the FFT is greatly reduced, speeding up computations.

For swept analysis, the filtered I and Q pairs are converted to magnitude and phase pairs. For traditional swept analysis, the magnitude signal is video-bandwidth (VBW) filtered and samples are taken through the display detector circuit. The log/linear display selection and dB/division scaling occur in the processor, so that a trace may be displayed on any scale without remeasuring.

Additional video processing features

The VBW filter normally smoothes the log of the magnitude of the signal, but it has many additional features. It can convert the log magnitude to a voltage envelope before filtering, and convert it back for consistent behavior before display detection.

Filtering the magnitude on a linear voltage scale is desirable for observing pulsed-RF envelope shapes in zero span. The log-magnitude signal can also be converted to a power (magnitude squared) signal before filtering, and then converted back. Filtering the power allows the analyzer to give the same average response to signals with noise-like characteristics, such as digital communications signals, as to CW signals with the same rms voltage. An increasingly common measurement need is total power in a channel or across a frequency range. In such a measurement, the display points might represent the average power during the time the LO sweeps through that point. The VBW filter can be reconfigured into an accumulator to perform averaging on either a log, voltage or power scale.

Frequency counting

Swept spectrum analyzers usually have a frequency counter. This counter counts the zero crossings in the IF signal and offsets that count by the known frequency offsets from LOs in the rest of the conversion chain. If the count is allowed to run for a second, a resolution of 1 Hz is achievable.

Because of its digitally synthesized LOs and all-digital RBWs, the native frequency accuracy of the PSA Series analyzer is very good (0.1% of span). In addition, the PSA analyzer includes a frequency counter that observes not just zero crossings, but also the change in phase. Thus, it can resolve frequency to the tens of millihertz level in 0.1 second. With this design, the ability to resolve frequency changes is not limited by the spectrum analyzer, but rather is determined by the noisiness of the signal being counted.

More advantages of the all-digital IF

We have already discussed a number of features in the PSA Series: power/voltage/log video filtering, high-resolution frequency counting, log/linear switching of stored traces, excellent shape factors, an average-across-the display-point detector mode, 160 RBWs, and of course, FFT or swept processing. In spectrum analysis, the filtering action of RBW filters causes errors in frequency and amplitude measurements that are a function of the sweep rate. For a fixed level of these errors, the all-digital IF's linear phase RBW filters allow faster sweep rates than do analog filters. The digital implementation also allows well-known compensations to frequency and amplitude readout, permitting sweep rates typically twice as fast as older analyzers, and excellent performance at even four times the sweep speed.

The digitally implemented logarithmic amplification is very accurate. Typical errors of the entire analyzer are much smaller than the measurement uncertainty with which the manufacturer proves the log fidelity. The log fidelity is specified at ±0.07 dB for any level up to –20 dBm at the input mixer of the analyzer. The range of the log amp does not limit the log fidelity at low levels, as it would be in an analog IF; the range is only limited by noise around –155 dBm at the input mixer. Because of single-tone compression in upstream circuits at higher powers, the fidelity specification degrades to ±0.13 dB for signal levels up to –10 dBm at the input mixer. By comparison, analog log amps are usually specified with tolerances in the ±1 dB region.

Other IF-related accuracies are improved as well. The IF prefilter is analog and must be aligned like an analog filter, so it is subject to alignment errors. But it is much better than most analog filters. With only one stage to manufacture, that stage can be made much more stable than the 4- and 5-stage filters of analog IF-based spectrum analyzers. As a result, the gain variations between RBW filters is held to a specification of ±0.03 dB, ten times better than all-analog designs.

The accuracy of the IF bandwidth is determined by settability limitations in the digital part of the filtering and calibration uncertainties in the analog prefilter. Again, the prefilter is highly stable and contributes only 20 percent of the error that would exist with an RBW made of five such stages. As a result, most RBWs are within 2 percent of their stated bandwidth, compared to 10 to 20 percent specifications in analog-IF analyzers.

The most important purpose of bandwidth accuracy is minimizing the inaccuracy of channel power and similar measurements. The noise bandwidth of the RBW filters is known to much better specifications than the 2 percent setting tolerance, and noise markers and channel-power measurements are corrected to a tolerance of ±0.5 percent. Therefore, bandwidth uncertainties contribute only ±0.022 dB to the amplitude error of noise density and channel-power measurements.

Finally, with no analog reference-level-dependent gain stages, there is no "IF gain" error at all. The sum of all these improvements means that the all-digital IF makes a quantum improvement in spectrum analyzer accuracy. It also allows you to change analyzer settings without significantly impacting measurement uncertainty. We will cover this topic in more detail in the next chapter.

Chapter 4
Amplitude and Frequency Accuracy

Now that we can view our signal on the display screen, let's look at amplitude accuracy, or perhaps better, amplitude uncertainty. Most spectrum analyzers are specified in terms of both absolute and relative accuracy. However, relative performance affects both, so let's look at those factors affecting relative measurement uncertainty first.

Before we discuss these uncertainties, let's look again at the block diagram of an analog swept-tuned spectrum analyzer, shown in Figure 4-1, and see which components contribute to the uncertainties. Later in this chapter, we will see how a digital IF and various correction and calibration techniques can substantially reduce measurement uncertainty.

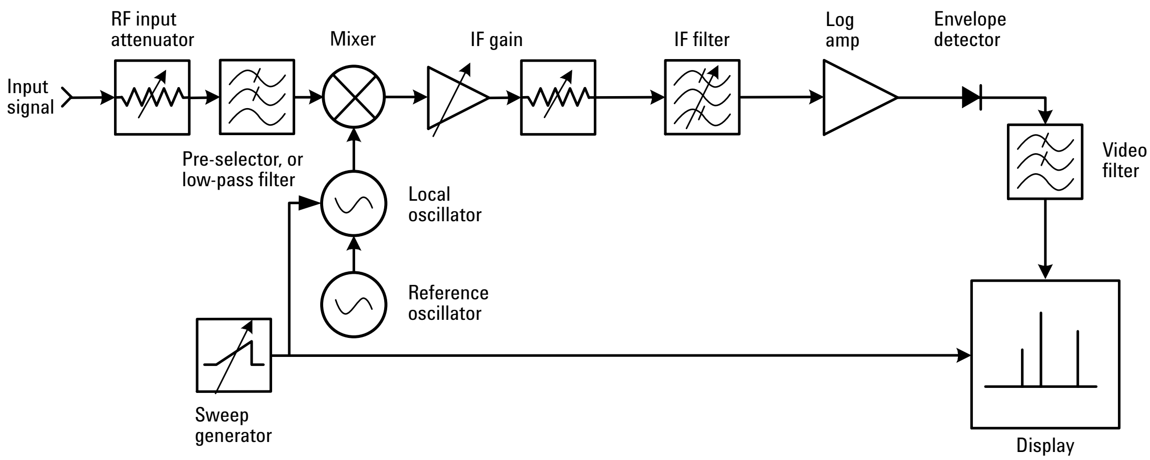


Figure 4-1. Spectrum analyzer block diagram

Components which contribute to uncertainty are:

- Input connector (mismatch)
- RF Input attenuator
- Mixer and input filter (flatness)
- IF gain/attenuation (reference level)
- RBW filters
- Display scale fidelity
- Calibrator (not shown)

An important factor in measurement uncertainty that is often overlooked is impedance mismatch. Analyzers do not have perfect input impedances, and signal sources do not have ideal output impedances. When a mismatch exists, the incident and reflected signal vectors may add constructively or destructively. Thus the signal received by the analyzer can be larger or smaller than the original signal. In most cases, uncertainty due to mismatch is relatively small. However, it should be noted that as spectrum analyzer amplitude accuracy has improved dramatically in recent years, mismatch uncertainty now constitutes a more significant part of the total measurement uncertainty. In any case, improving the match of either the source or analyzer reduces uncertainty[1].

1. For more information, see the Agilent *PSA Performance Spectrum Analyzer Series Amplitude Accuracy Product Note,* literature number 5980-3080EN.

The general expression used to calculate the maximum mismatch error in dB is:

$$\text{Error (dB)} = -20 \log[1 \pm |(\rho_{\text{analyzer}})(\rho_{\text{source}})|]$$

where ρ is the reflection coefficient

Spectrum analyzer data sheets typically specify the input voltage standing wave ratio (VSWR). Knowing the VSWR, we can calculate ρ with the following equation:

$$\rho = \frac{(\text{VSWR}-1)}{(\text{VSWR}+1)}$$

As an example, consider a spectrum analyzer with an input VSWR of 1.2 and a device under test (DUT) with a VSWR of 1.4 at its output port. The resulting mismatch error would be ±0.13 dB.

Since the analyzer's worst-case match occurs when its input attenuator is set to 0 dB, we should avoid the 0 dB setting if we can. Alternatively, we can attach a well-matched pad (attenuator) to the analyzer input and greatly reduce mismatch as a factor. Adding attenuation is a technique that works well to reduce measurement uncertainty when the signal we wish to measure is well above the noise. However, in cases where the signal-to-noise ratio is small (typically ≤7 dB), adding attenuation will increase measurement error because the noise power adds to the signal power, resulting in an erroneously high reading.

Let's turn our attention to the input attenuator. Some relative measurements are made with different attenuator settings. In these cases, we must consider the *input attenuation switching uncertainty*. Because an RF input attenuator must operate over the entire frequency range of the analyzer, its step accuracy varies with frequency. The attenuator also contributes to the overall frequency response. At 1 GHz, we expect the attenuator performance to be quite good; at 26 GHz, not as good.

The next component in the signal path is the input filter. Spectrum analyzers use a fixed low-pass filter in the low band and a tunable band pass filter called a preselector (we will discuss the preselector in more detail in Chapter 7) in the higher frequency bands. The low-pass filter has a better frequency response than the preselector and adds a small amount of uncertainty to the frequency response error. A preselector, usually a YIG-tuned filter, has a larger frequency response variation, ranging from 1.5 dB to 3 dB at millimeter-wave frequencies.

Following the input filter are the mixer and the local oscillator, both of which add to the *frequency response uncertainty*. Figure 4-2 illustrates what the frequency response might look like in one frequency band. Frequency response is usually specified as ± x dB relative to the midpoint between the extremes. The frequency response of a spectrum analyzer represents the overall system performance resulting from the flatness characteristics and interactions of individual components in the signal path up to and including the first mixer. Microwave spectrum analyzers use more than one frequency band to go above 3 GHz. This is done by using a higher harmonic of the local oscillator, which will be discussed in detail in Chapter 7. When making relative measurements between signals in different frequency bands, you must add the frequency response of each band to determine the overall frequency response uncertainty. In addition, some spectrum analyzers have a *band switching uncertainty* which must be added to the overall measurement uncertainty.

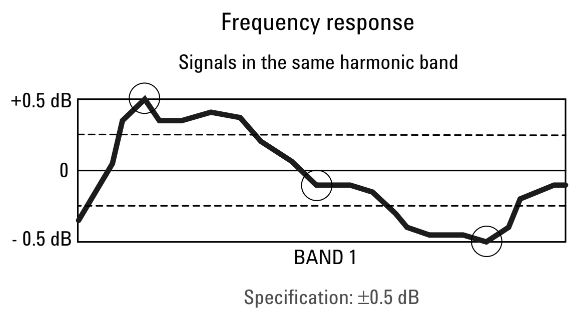


Figure 4-2. Relative frequency response in a single band

After the input signal is converted to an IF, it passes through the IF gain amplifier and IF attenuator which are adjusted to compensate for changes in the RF attenuator setting and mixer conversion loss. Input signal amplitudes are thus referenced to the top line of the graticule on the display, known as the reference level. The IF amplifier and attenuator only work at one frequency and, therefore, do not contribute to frequency response. However, there is always some amplitude uncertainty introduced by how accurately they can be set to a desired value. This uncertainty is known as *reference level accuracy*.

Another parameter that we might change during the course of a measurement is resolution bandwidth. Different filters have different insertion losses. Generally, we see the greatest difference when switching between LC filters (typically used for the wider resolution bandwidths) and crystal filters (used for narrow bandwidths). This results in *resolution bandwidth switching uncertainty*.

The most common way to display signals on a spectrum analyzer is to use a logarithmic amplitude scale, such as 10 dB per div or 1 dB per div. Therefore, the IF signal usually passes through a log amplifier. The gain characteristic of the log amplifier approximates a logarithmic curve. So any deviation from a perfect logarithmic response adds to the amplitude uncertainty. Similarly, when the spectrum analyzer is in linear mode, the linear amplifiers do not have a perfect linear response. This type of uncertainty is called *display scale fidelity*.

Relative uncertainty

When we make relative measurements on an incoming signal, we use either some part of the same signal or a different signal as a reference. For example, when we make second harmonic distortion measurements, we use the fundamental of the signal as our reference. Absolute values do not come into play; we are interested only in how the second harmonic differs in amplitude from the fundamental.

In a worst-case relative measurement scenario, the fundamental of the signal may occur at a point where the frequency response is highest, while the harmonic we wish to measure occurs at the point where the frequency response is the lowest. The opposite scenario is equally likely. Therefore, if our relative frequency response specification is ±0.5 dB as shown in Figure 4-2, then the total uncertainty would be twice that value, or ±1.0 dB.

Perhaps the two signals under test might be in different frequency bands of the spectrum analyzer. In that case, a rigorous analysis of the overall uncertainty must include the sum of the flatness uncertainties of the two frequency bands.

Other uncertainties might be irrelevant in a relative measurement, like the RBW switching uncertainty or reference level accuracy, which apply to both signals at the same time.

Absolute amplitude accuracy

Nearly all spectrum analyzers have a built-in calibration source which provides a known reference signal of specified amplitude and frequency. We then rely on the relative accuracy of the analyzer to translate the absolute calibration of the reference to other frequencies and amplitudes. Spectrum analyzers often have an *absolute frequency response* specification, where the zero point on the flatness curve is referenced to this calibration signal. Many Agilent spectrum analyzers use a 50 MHz reference signal. At this frequency, the specified absolute amplitude accuracy is extremely good: ±0.34 dB for the ESA-E Series and ±0.24 dB for the PSA Series analyzers.

It is best to consider all known uncertainties and then determine which ones can be ignored when doing a certain type of measurement. The range of values shown in Table 4-1 represents the specifications of a variety of different spectrum analyzers.

Some of the specifications, such as frequency response, are frequency-range dependent. A 3 GHz RF analyzer might have a frequency response of ±0.38 dB, while a microwave spectrum analyzer tuning in the 26 GHz range could have a frequency response of ±2.5 dB or higher. On the other hand, other sources of uncertainty, such as changing resolution bandwidths, apply equally to all frequencies.

Table 4-1. Representative values of amplitude uncertainty for common spectrum analyzers

Amplitude uncertainties (±dB)	
Relative	
RF attenuator switching uncertainty	0.18 to 0.7
Frequency response	0.38 to 2.5
Reference level accuracy (IF attenuator/gain change)	0.0 to 0.7
Resolution bandwidth switching uncertainty	0.03 to 1.0
Display scale fidelity	0.07 to 1.15
Absolute	
Calibrator accuracy	0.24 to 0.34

Improving overall uncertainty

When we look at total measurement uncertainty for the first time, we may well be concerned as we add up the uncertainty figures. The worst case view assumes that each source of uncertainty for your spectrum analyzer is at the maximum specified value, and that all are biased in the same direction at the same time. Since the sources of uncertainty can be considered independent variables, it is likely that some errors will be positive while others will be negative. Therefore, a common practice is to calculate the root sum of squares (RSS) error.

Regardless of whether we calculate the worst-case or RSS error, there are some things that we can do to improve the situation. First of all, we should know the specifications for our particular spectrum analyzer. These specifications may be good enough over the range in which we are making our measurement. If not, Table 4-1 suggests some opportunities to improve accuracy.

Before taking any data, we can step through a measurement to see if any controls can be left unchanged. We might find that the measurement can be made without changing the RF attenuator setting, resolution bandwidth, or reference level. If so, all uncertainties associated with changing these controls drop out. We may be able to trade off reference level accuracy against display fidelity, using whichever is more accurate and eliminating the other as an uncertainty factor. We can even get around frequency response if we are willing to go to the trouble of characterizing our particular analyzer[2]. This can be accomplished by using a power meter and comparing the reading of the spectrum analyzer at the desired frequencies with the reading of the power meter.

The same applies to the calibrator. If we have a more accurate calibrator, or one closer to the frequency of interest, we may wish to use that in lieu of the built-in calibrator. Finally, many analyzers available today have self-calibration routines. These routines generate error coefficients (for example, amplitude changes versus resolution bandwidth), that the analyzer later uses to correct measured data. As a result, these self-calibration routines allow us to make good amplitude measurements with a spectrum analyzer and give us more freedom to change controls during the course of a measurement.

Specifications, typical performance, and nominal values

When evaluating spectrum analyzer accuracy, it is very important to have a clear understanding of the many different values found on an analyzer data sheet. Agilent Technologies defines three classes of instrument performance data:

Specifications describe the performance of parameters covered by the product warranty over a temperature range of 0 to 55 °C (unless otherwise noted). Each instrument is tested to verify that it meets the specification, and takes into account the measurement uncertainty of the equipment used to test the instrument. 100% of the units tested will meet the specification.

Some test equipment manufacturers use a “2 sigma” or 95% confidence value for certain instrument specifications. When evaluating data sheet specifications for instruments from different manufacturers, it is important to make sure you are comparing like numbers in order to make an accurate comparison.

2. Should we do so, then mismatch may become a more significant error.

Typical performance describes additional product performance information that is not covered by the product warranty. It is performance beyond specification that 80% of the units exhibit with a 95% confidence level over the temperature range 20 to 30 °C. Typical performance does not include measurement uncertainty. During manufacture, all instruments are tested for typical performance parameters.

Nominal values indicate expected performance, or describe product performance that is useful in the application of the product, but is not covered by the product warranty. Nominal parameters generally are not tested during the manufacturing process.

The digital IF section

As described in the previous chapter, a digital IF architecture eliminates or minimizes many of the uncertainties experienced in analog spectrum analyzers. These include:

Reference level accuracy (IF gain uncertainty)

Spectrum analyzers with an all digital IF, such as the Agilent PSA Series, do not have IF gain that changes with reference level. Therefore, there is no IF gain uncertainty.

Display scale fidelity

A digital IF architecture does not include a log amplifier. Instead, the log function is performed mathematically, and traditional log fidelity uncertainty does not exist. However, other factors, such as RF compression (especially for input signals above –20 dBm), ADC range gain alignment accuracy, and ADC linearity (or quantization error) contribute to display scale uncertainty. The quantization error can be improved by the addition of noise which smoothes the average of the ADC transfer function. This added noise is called dither. While the dither improves linearity, it does slightly degrade the displayed average noise level. In the PSA Series, it is generally recommended that dither be used when the measured signal has a signal-to-noise ratio of greater than or equal to 10 dB. When the signal-to-noise ratio is under 10 dB, the degradations to accuracy of any single measurement (in other words, without averaging) that come from a higher noise floor are worse than the linearity problems solved by adding dither, so dither is best turned off.

RBW switching uncertainty

The digital IF in the PSA Series includes an analog prefilter set to 2.5 times the desired resolution bandwidth. This prefilter has some uncertainty in bandwidth, gain, and center frequency as a function of the RBW setting. The rest of the RBW filtering is done digitally in an ASIC in the digital IF section. Though the digital filters are not perfect, they are very repeatable, and some compensation is applied to minimize the error. This results in a tremendous overall improvement to the RBW switching uncertainty compared to analog implementations.

Examples

Let's look at some amplitude uncertainty examples for various measurements. Suppose we wish to measure a 1 GHz RF signal with an amplitude of –20 dBm. If we use an Agilent E4402B ESA-E Series spectrum analyzer with Atten = 10 dB, RBW = 1 kHz, VBW = 1 kHz, Span = 20 kHz, Ref level = –20 dBm, log scale, and coupled sweep time, and an ambient temperature of 20 to 30 °C, the specifications tell us that the absolute uncertainty equals ±0.54 dB plus the absolute frequency response. An E4440A PSA Series spectrum analyzer measuring the same signal using the same settings would have a specified uncertainty of ±0.24 dB plus the absolute frequency response. These values are summarized in Table 4-2.

Table 4-2. Amplitude uncertainties when measuring a 1 GHz signal

Source of uncertainty	Absolute uncertainty of 1 GHz, –20 dBm signal	
	E4402B	**E4440A**
Absolute amplitude accuracy	±0.54 dB	±0.24 dB
Frequency response	±0.46 dB	±0.38 dB
Total worst case uncertainty	±1.00 dB	±0.62 dB
Total RSS uncertainty	±0.69 dB	±0.44 dB
Typical uncertainty	±0.25 dB	±0.17 dB

At higher frequencies, the uncertainties get larger. In this example, we wish to measure a 10 GHz signal with an amplitude of –10 dBm. In addition, we also want to measure its second harmonic at 20 GHz. Assume the following measurement conditions: 0 to 55 °C, RBW = 300 kHz, Atten = 10 dB, Ref level = –10 dBm. In Table 4-3, we compare the absolute and relative amplitude uncertainty of two different Agilent spectrum analyzers, an 8563EC (analog IF) and an E4440A PSA (digital IF).

Table 4-3. Absolute and relative amplitude accuracy comparison (8563EC and E4440A PSA)

	Measurement of a 10 GHz signal at -10dBm			
Source of uncertainty	**Absolute uncertainty of fundamental at 10 GHz**		**Relative uncertainty of second harmonic at 20 GHz**	
	8563EC	**E4440A**	**8563EC**	**E4440A**
Calibrator	±0.3 dB	N/A	N/A	N/A
Absolute amplitude acc.	N/A	±0.24 dB	N/A	N/A
Attenuator	N/A	N/A	N/A	N/A
Frequency response	±2.9 dB	±2.0 dB	±(2.2 + 2.5) dB	±(2.0 + 2.0) dB
Band switching uncertainty	N/A	N/A	±1.0 dB	N/A
IF gain	N/A	N/A	N/A	N/A
RBW switching	N/A	N/A	N/A	N/A
Display scale fidelity	N/A	N/A	±0.85 dB	±0.13 dB
Total worst case uncertainty	±3.20 dB	±2.24 dB	±6.55 dB	±4.13 dB
Total RSS uncertainty	±2.91 dB	±2.01 dB	±3.17 dB	±2.62 dB
Typical uncertainty	±2.30 dB	±1.06 dB	±4.85 dB	±2.26 dB

Frequency accuracy

So far, we have focused almost exclusively on amplitude measurements. What about frequency measurements? Again, we can classify two broad categories, *absolute* and *relative* frequency measurements. Absolute measurements are used to measure the frequencies of specific signals. For example, we might want to measure a radio broadcast signal to verify that it is operating at its assigned frequency. Absolute measurements are also used to analyze undesired signals, such as when doing a spur search. Relative measurements, on the other hand, are useful to know how far apart spectral components are, or what the modulation frequency is.

Up until the late 1970s, absolute frequency uncertainty was measured in megahertz because the first LO was a high-frequency oscillator operating above the RF range of the analyzer, and there was no attempt to tie the LO to a more accurate reference oscillator. Today's LOs are synthesized to provide better accuracy. Absolute frequency uncertainty is often described under the frequency *readout accuracy specification* and refers to center frequency, start, stop, and marker frequencies.

With the introduction of the Agilent 8568A in 1977, counter-like frequency accuracy became available in a general-purpose spectrum analyzer and ovenized oscillators were used to reduce drift. Over the years, crystal reference oscillators with various forms of indirect synthesis have been added to analyzers in all cost ranges. The broadest definition of indirect synthesis is that the frequency of the oscillator in question is in some way determined by a reference oscillator. This includes techniques such as phase lock, frequency discrimination, and counter lock.

What we really care about is the effect these changes have had on frequency accuracy (and drift). A typical readout accuracy might be stated as follows:

±[(freq readout x freq ref error) + A% of span + B% of RBW + C Hz]

Note that we cannot determine an exact frequency error unless we know something about the frequency reference. In most cases we are given an annual aging rate, such as $\pm 1 \times 10^{-7}$ per year, though sometimes aging is given over a shorter period (for example, $\pm 5 \times 10^{-10}$ per day). In addition, we need to know when the oscillator was last adjusted and how close it was set to its nominal frequency (usually 10 MHz). Other factors that we often overlook when we think about frequency accuracy include how long the reference oscillator has been operating. Many oscillators take 24 to 72 hours to reach their specified drift rate. To minimize this effect, some spectrum analyzers continue to provide power to the reference oscillator as long as the instrument is plugged into the AC power line. In this case, the instrument is not really turned "off," but more properly is on "standby." We also need to consider the temperature stability, as it can be worse than the drift rate. In short, there are a number of factors to consider before we can determine frequency uncertainty.

In a factory setting, there is often an in-house frequency standard available that is traceable to a national standard. Most analyzers with internal reference oscillators allow you to use an external reference. The frequency reference error in the foregoing expression then becomes the error of the in-house standard.

When making relative measurements, span accuracy comes into play. For Agilent analyzers, span accuracy generally means the uncertainty in the indicated separation of any two spectral components on the display. For example, suppose span accuracy is 0.5% of span and we have two signals separated by two divisions in a 1 MHz span (100 kHz per division). The uncertainty of the signal separation would be 5 kHz. The uncertainty would be the same if we used delta markers and the delta reading would be 200 kHz. So we would measure 200 kHz ±5 kHz.

When making measurements in the field, we typically want to turn our analyzer on, complete our task, and move on as quickly as possible. It is helpful to know how the reference in our analyzer behaves under short warm up conditions. For example, the Agilent ESA-E Series of portable spectrum analyzers will meet published specifications after a five-minute warm up time.

Most analyzers include markers that can be put on a signal to give us absolute frequency, as well as amplitude. However, the indicated frequency of the marker is a function of the frequency calibration of the display, the location of the marker on the display, and the number of display points selected. Also, to get the best frequency accuracy we must be careful to place the marker exactly at the peak of the response to a spectral component. If we place the marker at some other point on the response, we will get a different frequency reading. For the best accuracy, we may narrow the span and resolution bandwidth to minimize their effects and to make it easier to place the marker at the peak of the response.

Many analyzers have marker modes that include internal counter schemes to eliminate the effects of span and resolution bandwidth on frequency accuracy. The counter does not count the input signal directly, but instead counts the IF signal and perhaps one or more of the LOs, and the processor computes the frequency of the input signal. A minimum signal-to-noise ratio is required to eliminate noise as a factor in the count. Counting the signal in the IF also eliminates the need to place the marker at the exact peak of the signal response on the display. If you are using this marker counter function, placement anywhere sufficiently out of the noise will do. Marker count accuracy might be stated as:

±[(marker freq x freq ref error) + counter resolution]

We must still deal with the frequency reference error as previously discussed. Counter resolution refers to the least significant digit in the counter readout, a factor here just as with any simple digital counter. Some analyzers allow the counter mode to be used with delta markers. In that case, the effects of counter resolution and the fixed frequency would be doubled.

Chapter 5
Sensitivity and Noise

Sensitivity

One of the primary uses of a spectrum analyzer is to search out and measure low-level signals. The limitation in these measurements is the noise generated within the spectrum analyzer itself. This noise, generated by the random electron motion in various circuit elements, is amplified by multiple gain stages in the analyzer and appears on the display as a noise signal. On a spectrum analyzer, this noise is commonly referred to as the *Displayed Average Noise Level, or DANL*[1]. While there are techniques to measure signals slightly below the DANL, this noise power ultimately limits our ability to make measurements of low-level signals.

Let's assume that a 50 ohm termination is attached to the spectrum analyzer input to prevent any unwanted signals from entering the analyzer. This passive termination generates a small amount of noise energy equal to kTB, where:

k = Boltzmann's constant (1.38 x 10^{-23} joule/°K)
T = temperature, in degrees Kelvin
B = bandwidth in which the noise is measured, in Hertz

Since the total noise power is a function of measurement bandwidth, the value is typically normalized to a 1 Hz bandwidth. Therefore, at room temperature, the noise power density is –174 dBm/Hz. When this noise reaches the first gain stage in the analyzer, the amplifier boosts the noise, plus adds some of its own. As the noise signal passes on through the system, it is typically high enough in amplitude that the noise generated in subsequent gain stages adds only a small amount to the total noise power. Note that the input attenuator and one or more mixers may be between the input connector of a spectrum analyzer and the first stage of gain, and all of these components generate noise. However, the noise that they generate is at or near the absolute minimum of –174 dBm/Hz, so they do not significantly affect the noise level input to, and amplified by, the first gain stage.

While the input attenuator, mixer, and other circuit elements between the input connector and first gain stage have little effect on the actual system noise, they do have a marked effect on the ability of an analyzer to display low-level signals because they attenuate the input signal. That is, they reduce the signal-to-noise ratio and so degrade sensitivity.

We can determine the DANL simply by noting the noise level indicated on the display when the spectrum analyzer input is terminated with a 50 ohm load. This level is the spectrum analyzer's own noise floor. Signals below this level are masked by the noise and cannot be seen. However, the DANL is not the actual noise level at the input, but rather the effective noise level. An analyzer display is calibrated to reflect the level of a signal at the analyzer input, so the displayed noise floor represents a fictitious, or effective noise floor at the input.

The actual noise level at the input is a function of the input signal. Indeed, noise is sometimes the signal of interest. Like any discrete signal, a noise signal is much easier to measure when it is well above the effective (displayed) noise floor. The effective input noise floor includes the losses caused by the input attenuator, mixer conversion loss, and other circuit elements prior to the first gain stage. We cannot do anything about the conversion loss of the mixers, but we can change the RF input attenuator. This enables us to control the input signal power to the first mixer and thus change the displayed signal-to-noise floor ratio. Clearly, we get the lowest DANL by selecting minimum (zero) RF attenuation.

1. Displayed average noise level is sometimes confused with the term "Sensitivity". While related, these terms have different meanings. Sensitivity is a measure of the minimum signal level that yields a defined signal-to-noise ratio (SNR) or bit error rate (BER). It is a common metric of radio receiver performance. Spectrum analyzer specifications are always given in terms of the DANL.

Because the input attenuator has no effect on the actual noise generated in the system, some early spectrum analyzers simply left the displayed noise at the same position on the display regardless of the input attenuator setting. That is, the IF gain remained constant. This being the case, the input attenuator affected the location of a true input signal on the display. As input attenuation was increased, further attenuating the input signal, the location of the signal on the display went down while the noise remained stationary.

Beginning in the late 1970s, spectrum analyzer designers took a different approach. In newer analyzers, an internal microprocessor changes the IF gain to offset changes in the input attenuator. Thus, signals present at the analyzer's input remain stationary on the display as we change the input attenuator, while the displayed noise moves up and down. In this case, the reference level remains unchanged. This is shown in Figure 5-1. As the attenuation increases from 5 to 15 to 25 dB, the displayed noise rises while the –30 dBm signal remains constant. In either case, we get the best signal-to-noise ratio by selecting minimum input attenuation.

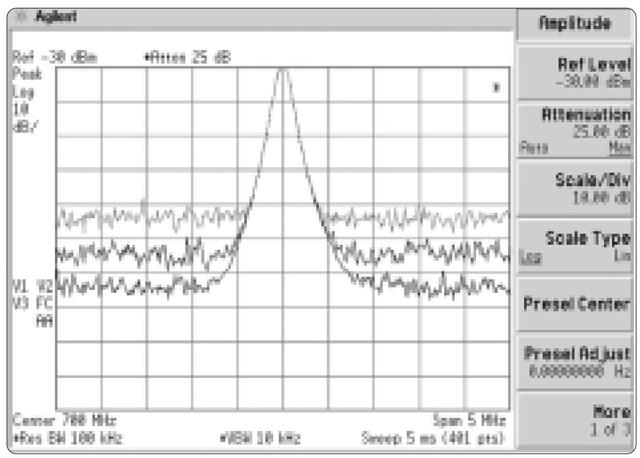


Figure 5-1. Reference level remains constant when changing input attenuation

Resolution bandwidth also affects signal-to-noise ratio, or sensitivity. The noise generated in the analyzer is random and has a constant amplitude over a wide frequency range. Since the resolution, or IF, bandwidth filters come after the first gain stage, the total noise power that passes through the filters is determined by the width of the filters. This noise signal is detected and ultimately reaches the display. The random nature of the noise signal causes the displayed level to vary as:

$$10 \log (BW_2/BW_1)$$

where BW_1 = starting resolution bandwidth
BW_2 = ending resolution bandwidth

So if we change the resolution bandwidth by a factor of 10, the displayed noise level changes by 10 dB, as shown in Figure 5-2. For continuous wave (CW) signals, we get best signal-to-noise ratio, or best sensitivity, using the minimum resolution bandwidth available in our spectrum analyzer[2].

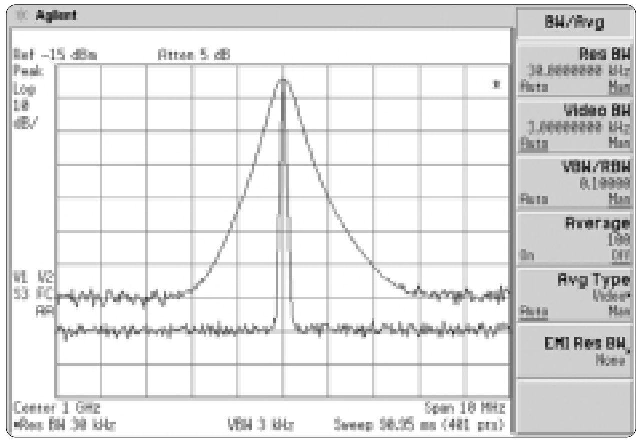


Figure 5-2. Displayed noise level changes as 10 log(BW_2/BW_1)

A spectrum analyzer displays signal plus noise, and a low signal-to-noise ratio makes the signal difficult to distinguish. We noted previously that the video filter can be used to reduce the amplitude fluctuations of noisy signals while at the same time having no effect on constant signals. Figure 5-3 shows how the video filter can improve our ability to discern low-level signals. It should be noted that the video filter does not affect the average noise level and so does not, by this definition, affect the sensitivity of an analyzer.

In summary, we get best sensitivity for narrowband signals by selecting the minimum resolution bandwidth and minimum input attenuation. These settings give us best signal-to-noise ratio. We can also select minimum video bandwidth to help us see a signal at or close to the noise level[3]. Of course, selecting narrow resolution and video bandwidths does lengthen the sweep time.

2. Broadband, pulsed signals can exhibit the opposite behavior, where the SNR increases as the bandwidth gets larger.
3. For the effect of noise on accuracy, see "Dynamic range versus measurement uncertainty" in Chapter 6.

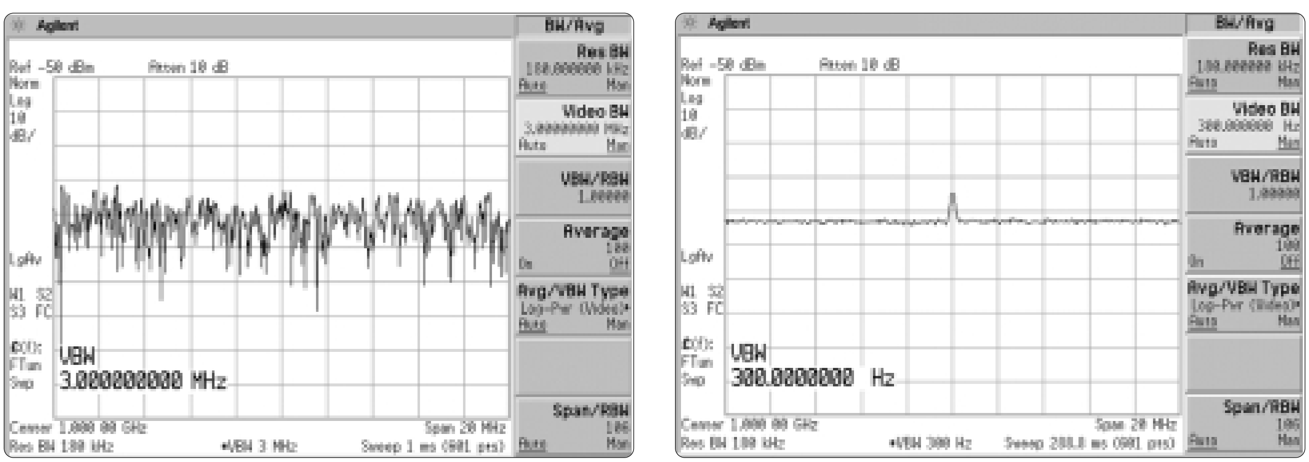


Figure 5-3. Video filtering makes low-level signals more discernable