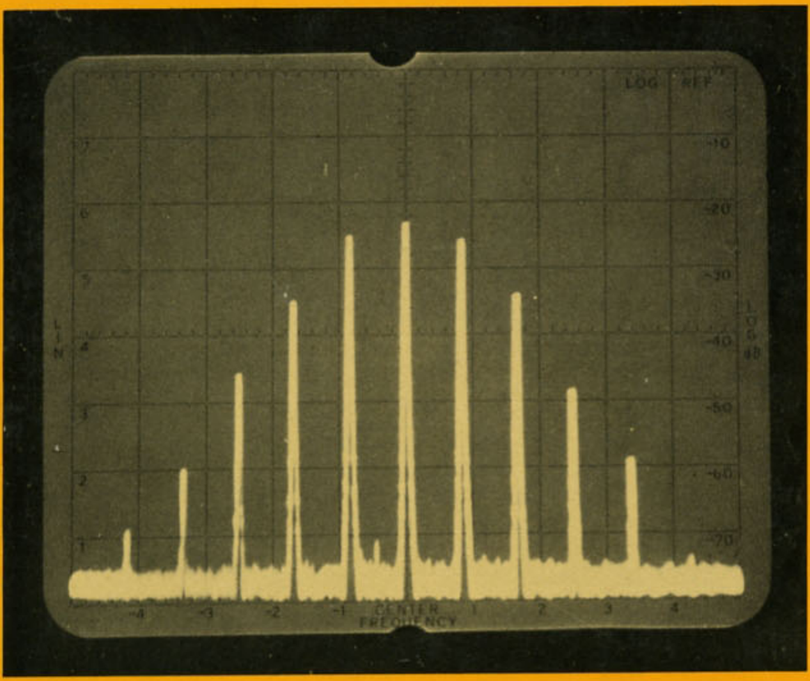


Spectrum Analyzer Series

APPLICATION NOTE 150

SPECTRUM ANALYSIS....
Spectrum Analyzer
Basics

HEWLETT hp PACKARD

April 1974

APPLICATION NOTE 150

SPECTRUM ANALYSIS

Spectrum Analyzer

Basics

Printed April 1974

HEWLETT hp PACKARD

CONTENTS

Rev. 7/15/75

CHAPTER 1

INTRODUCTION TO SPECTRUM ANALYSIS

General

The traditional way of observing electrical signals is to view them in the time domain using an oscilloscope. The time domain is used to recover relative timing and phase information which is needed to characterize electric circuit behavior. However, not all circuits can be uniquely characterized from just time domain information. Circuit elements such as amplifiers, oscillators, mixers, modulators, detectors and filters are best characterized by their frequency response information. This frequency information is best obtained by viewing electrical signals in the frequency domain. One instrument which displays the frequency domain is the spectrum analyzer.

What Is the Frequency Domain?

The frequency domain is a graphical representation of signal amplitude as a function of frequency. Figure 1.1 illustrates the relationship between the time domain and the frequency domain. In the time domain, all frequency components of a signal are seen summed together. In the frequency domain, complex signals (i.e., signals composed of more than one frequency) are separated into their frequency components, and the power level at each frequency is displayed.

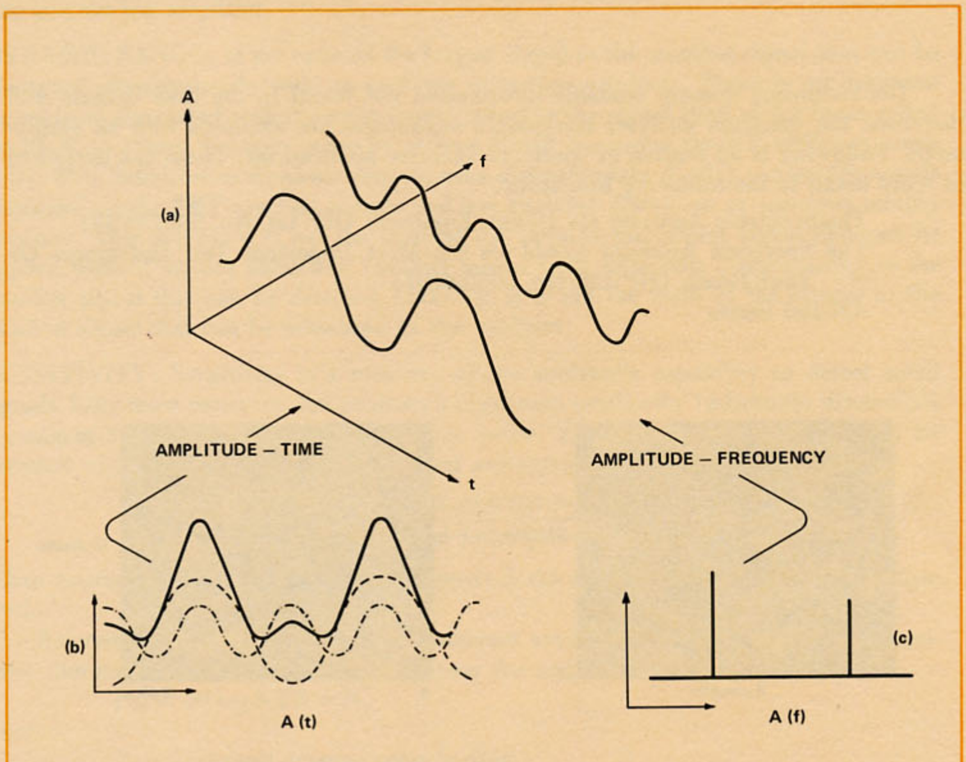


Figure 1.1 Relationship between the time and frequency domains: (a) Three-dimensional coordinates showing time, frequency, and amplitude. The addition of a fundamental and its second harmonic is shown as an example. (b) View seen in the time domain. On an oscilloscope, only the composite $f_1 + 2F_1$ would be seen. (c) View seen in the frequency domain. The components of the composite signal are clearly seen here.

What Is a Spectrum Analyzer?

To display the frequency domain requires a device that can discriminate between frequencies while measuring the power level at each. The spectrum analyzer was designed for this purpose. A spectrum analyzer is an instrument which graphically displays voltage or power as a function of frequency on a CRT (cathode ray tube). It can be used to analyze signals in the frequency domain.

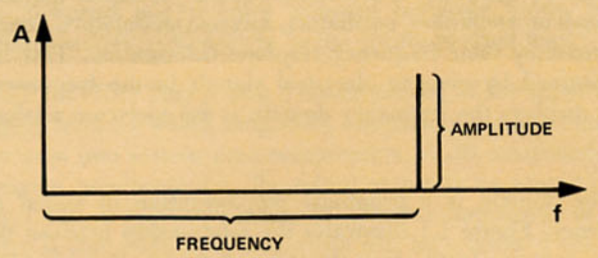


Figure 1.2 Frequency domain display of a CW signal.

Spectrum Analyzer Applications

The frequency domain contains information not found in the time domain and therefore, the spectrum analyzer has certain advantages not available with an oscilloscope. Following is an outline of spectrum analyzer applications. These are described in more detail in the following brochures:

The Spectrum Analyzer for Design Engineers (HP Lit. No. 5952-0932)
The Spectrum Analyzer Could Be the Most Important Test Instrument On Your Bench (HP Lit. No. 5952-9201)
AN-150 Series

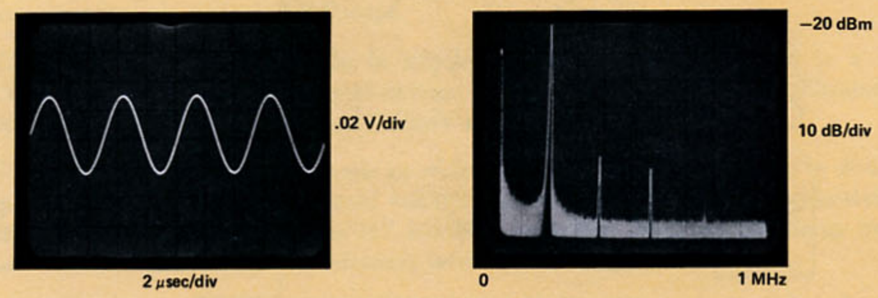


Figure 1.3

1. The analyzer is more sensitive to low level distortion than a scope. The sine wave in Figure 1.3 looks good in the time domain, but in the frequency domain, harmonic distortion can be seen.

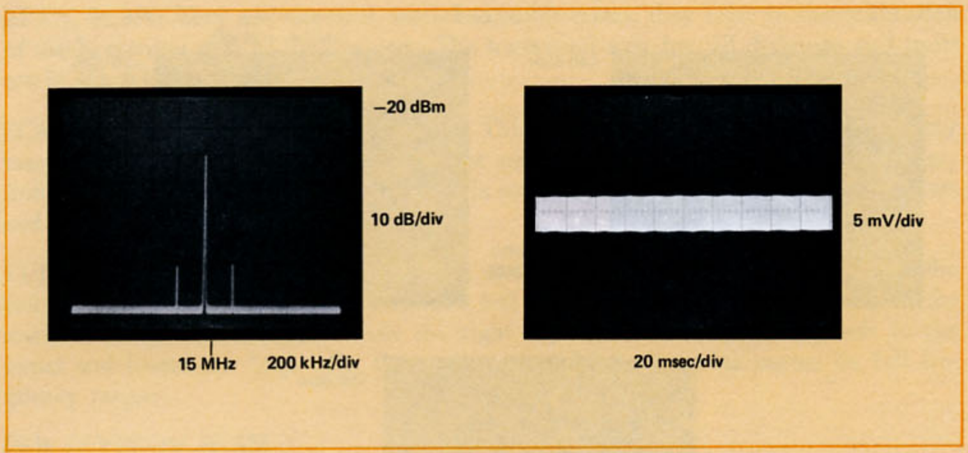


Figure 1.4

2. The sensitivity and wide dynamic range of the analyzer is useful for measuring low-level modulation. In Figure 1.4 the 2% AM can be measured easily with the spectrum analyzer where it can barely be seen with an oscilloscope.

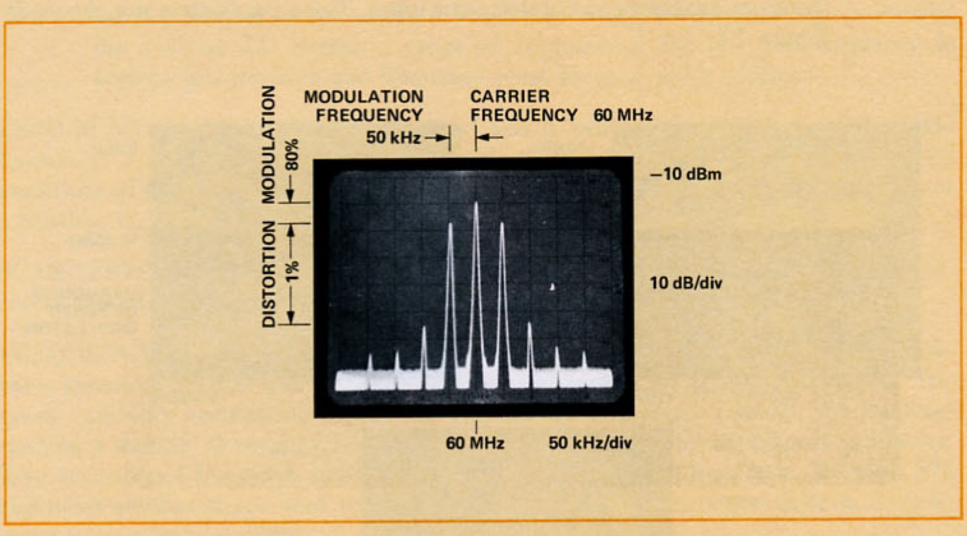


Figure 1.5

3. The spectrum analyzer can be used to measure AM, FM and pulsed RF. Figure 1.5 shows how the analyzer can be used to measure carrier frequency, modulation frequency, modulation level, and modulation distortion.

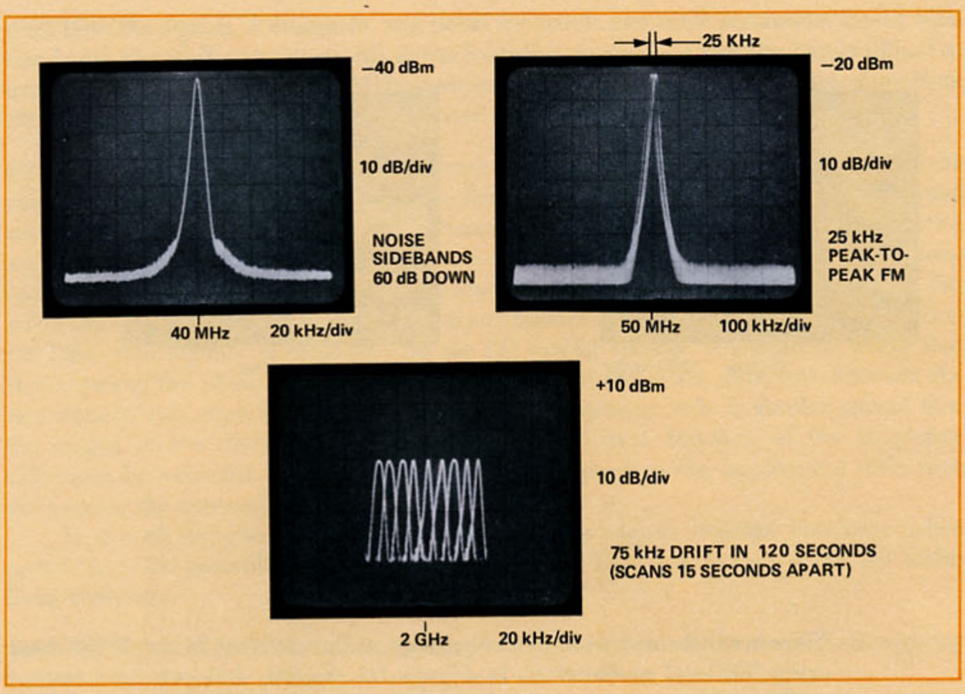


Figure 1.6

4. The spectrum analyzer can be used to measure long and short term stability. Parameters such as noise sidebands on an oscillator, residual FM of a source and frequency drift during warm-up can be measured using the spectrum analyzer's calibrated scan times. These parameters are shown in Figure 1.6.

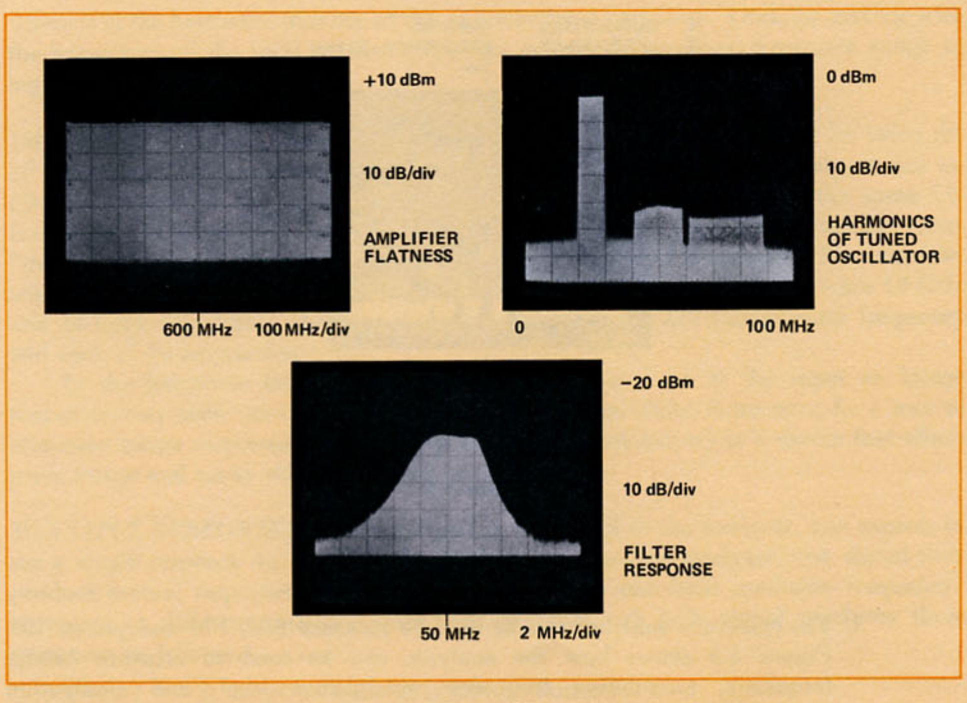


Figure 1.7

5. The swept frequency response of a filter or amplifier and the swept distortion measurement of a tuned oscillator are examples of swept frequency measurements possible with a spectrum analyzer, as shown in Figure 1.7. These measurements are simplified by using a variable persistence display or a tracking generator which will be described later.

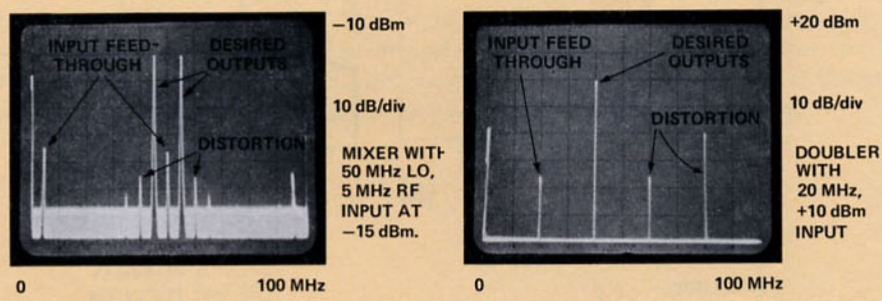


Figure 1.8

6. Frequency conversion devices can be easily characterized with a spectrum analyzer. Such parameters as conversion loss, isolation, and distortion are readily determined from the display.

Types of Spectrum Analyzers

There are two basic types of spectrum analyzers, swept-tuned and real-time analyzers. The swept-tuned analyzers are tuned by electrically sweeping them over their frequency range. Therefore, the frequency components of a spectrum are sampled sequentially in time.

This enables periodic and random signals to be displayed, but makes it impossible to display transient responses. Real-time analyzers, on the other hand, simultaneously display the amplitude of all signals in the frequency range of the analyzer; hence the name real-time. This preserves the time dependency between signals which permits phase information to be displayed. Real-time analyzers are capable of displaying transient responses as well as periodic and random signals.

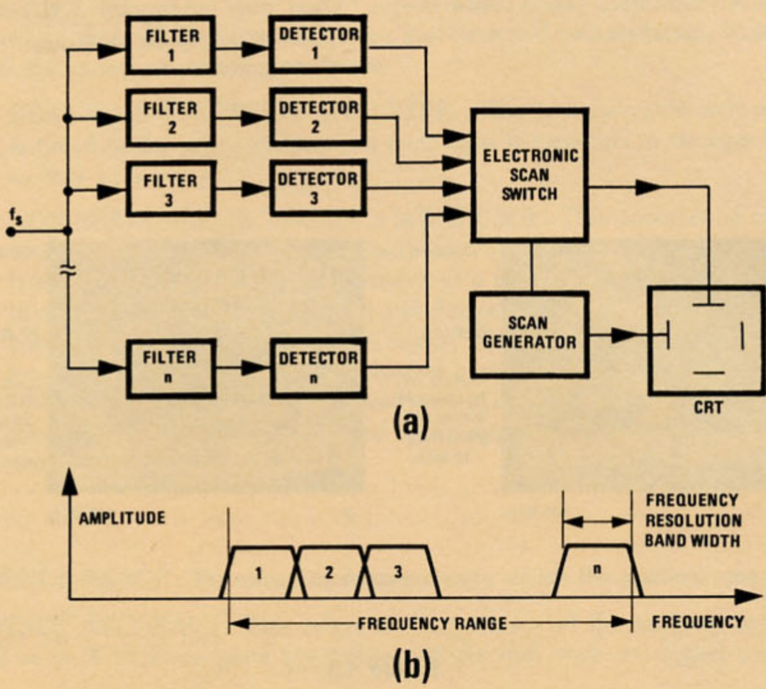


Figure 1.9 (a) Block diagram of a real-time spectrum analyzer. The electronic scan switch samples the filters fast enough to display the instantaneous response of the input signals. (b) The bandpass filters are staggered to permit continuous coverage of the frequency spectrum.

A real-time or multichannel analyzer is basically a set of staggered bandpass filters as shown in Figure 1.9. The composite amplitude of the signals within each filter passband is displayed as a function of the overall filter frequency range. The frequency range of the analyzer is limited by the number of filters and their bandwidth. This arrangement is usually very expensive because of the large numbers of filters required to cover a spectrum, and lacks flexibility in the frequency scale because of the fixed resolution of the filters. It is real time, however, and is better at analyzing low frequencies such as audio and sub-audio range. The bandpass of the filters can be made very narrow for good resolution without having to sacrifice sweep speed as in the swept tuned analyzers. The multichannel analyzer and Fourier Analyzer shown in Figure 1.10 are two widely used real-time analyzers. The Fourier analyzer performs the conversion from time domain to frequency domain by mathematical computation of the Fourier Transform.

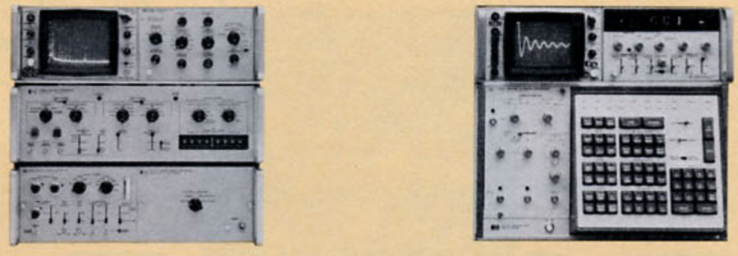


Figure 1.10 A multichannel and Fourier analyzer are shown on the left and right respectively. Both instruments are real-time spectrum analyzers.

Swept-Tuned Spectrum Analyzers

The swept-tuned analyzers are usually of the trf (tuned radio frequency) or superheterodyne type. A block diagram of a trf analyzer is shown in Figure 1.11. A trf analyzer consists of a bandpass filter whose center frequency is tunable over a desired frequency range, a detector to produce vertical deflection on a CRT, and a horizontal scan generator used to synchronize the tuned frequency to the CRT horizontal deflection. It is a simple, inexpensive analyzer with wide frequency coverage, but lacks resolution and sensitivity. Because trf analyzers have a swept filter they are limited in sweep width depending on the frequency range (usually one decade or less). The resolution is determined by the filter bandwidth, and, since tunable filters don't usually have constant bandwidth, is dependent on frequency.

The trf analyzer is relatively inexpensive and is often used for microwave applications due to the availability of broadband tuned filters. However, the use of tuned analyzers eliminates the real-time aspect of the display, and sweep rates (MHz/sec) must be consistent with the charge time of the tunable filter.

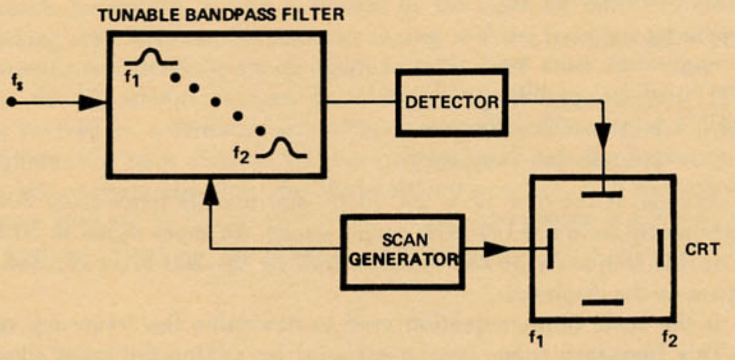


Figure 1.11 Block diagram of a swept trf spectrum analyzer. The frequency spectrum is sampled using a tuned bandpass filter.

The most common type of spectrum analyzer differs from the trf spectrum analyzers in that the spectrum is swept through a fixed bandpass filter instead of sweeping the filter through the spectrum. The block diagram of a swept superheterodyne spectrum analyzer is shown in Figure 1.12. The analyzer is basically a narrowband receiver which is electronically tuned in frequency by applying a saw-tooth voltage to the frequency control element of a voltage tuned local oscillator. This same saw-tooth voltage is simultaneously applied to the horizontal deflection plates of the CRT. The output from the receiver is synchronously applied to the vertical deflection plates of the CRT and a plot of amplitude versus frequency is displayed.

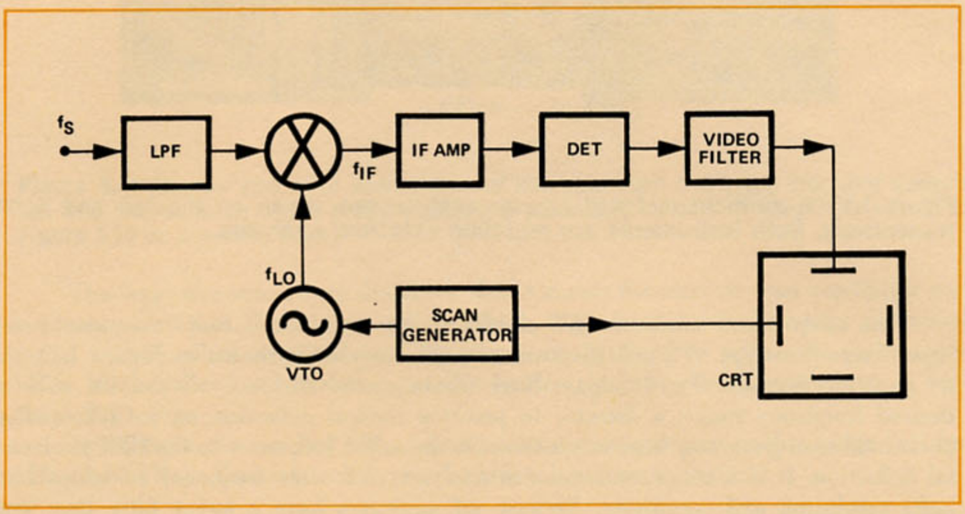


Figure 1.12 Block diagram of a swept superheterodyne spectrum analyzer. The input signal is mixed with a tuned LO frequency to produce a fixed IF which can be detected and displayed.

The analyzer is tuned through its frequency range by varying the voltage on the LO (local oscillator). The LO frequency is mixed with the input signal to produce an IF (intermediate frequency). When the frequency difference between the input signal and the LO frequency is equal to the IF frequency, then there is a response on the analyzer.
If $f_s < f_{LO}$
then

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

where

f_s = input signal frequency
f_{LO} = local oscillator frequency
f_{IF} = intermediate frequency

For example, if the first IF is 200 MHz, and the LO tunes from 200-310 MHz, the analyzer would have a 0-110 MHz tuning range. An input signal at 50 MHz would mix with an LO frequency of 250 MHz to produce the 200 MHz IF, and a response would appear on the display.

This is the basic tuning equation used to determine the frequency range of the analyzer. This frequency range can be extended by mixing the input signal with the harmonics of the LO frequency. Harmonic mixing is discussed in the next chapter.

The advantages of the superheterodyne technique are considerable. It obtains high sensitivity through the use of IF amplifiers, and many decades in frequency can be tuned. Also, the resolution can be varied by changing the bandwidth of the IF filters. However, the superheterodyne analyzer is not real-time, and, once again, sweep rates must be consistent with the IF filter charge time.

The superheterodyne approach is the most flexible, and this is what will be discussed in this note.

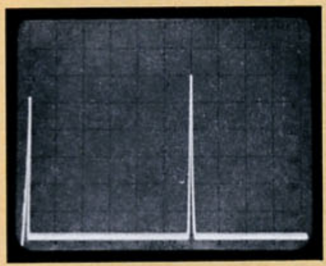


Figure 1.13 Display of a CW signal

Here is a spectrum analyzer display of a CW signal. The signal can be measured from the display.

The response at the left edge of the CRT is sometimes called the "zero frequency indicator" or local oscillator feedthrough. It occurs when the analyzer is tuned to zero frequency, and the local oscillator passes directly through IF creating a response on the CRT even when no input signal is present. (For zero frequency tuning, $F_{LO} = F_{IF}$). This effectively limits the lower tuning limit.

Spectrum Analyzer Requirements

To accurately display the frequency and amplitude of a signal on a spectrum analyzer, the analyzer itself must be properly calibrated. A spectrum analyzer properly designed for accurate frequency and amplitude measurements will satisfy these requirements.

General:

1. Flat frequency response (amplitude is independent of frequency)
2. All functions should be calibrated

Frequency:

1. Wide tuning range
2. Wide frequency display range
3. Stability
4. Resolution
5. Signal identification

Amplitude:

1. High sensitivity
2. Low internal distortion
3. Linear and Logarithmic display modes (voltage and dB)
4. Absolute measurement of amplitude

Frequency Measurements

Modern spectrum analyzers are calibrated in both frequency and amplitude for relative and absolute measurements. The frequency calibration is usually set by the tuning and scale factor controls. The frequency scale can be scanned in three different modes—full, per division, and zero scan. The full scan mode is used to locate signals because the widest frequency ranges are displayed in this mode. Sometimes an inverted marker below the baseline, as shown in Figure 1.14, is used to track the frequency dial on the front panel. When the marker is set under a signal, the frequency dial indicates the signal frequency.

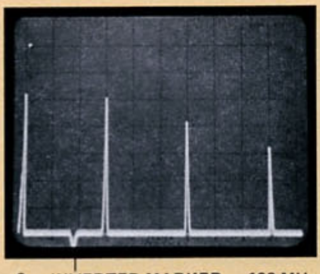


Figure 1.14 An inverted marker is used to indicate the frequency that the analyzer is tuned to. The marker is only present in the full scan mode.

The per division mode is used to zoom-in on a particular signal. In per division, the center frequency of the display is set by the Tuning control and the scale factor is set by the Frequency Span or Scan Width control. The signal in Figure 1.15 was picked out by the inverted marker and zoomed-in on.

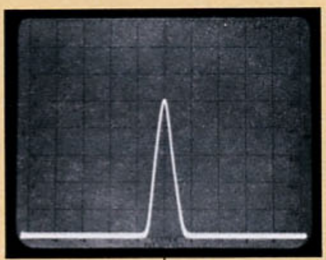


Figure 1.15 This signal was picked out with the inverted marker and zoomed-in on using the per-division scan mode.

In the zero scan mode, the analyzer acts as a fixed-tuned receiver with selectable bandwidths for recovering modulating signals or real-time monitoring of a single signal. AM or FM broadcasts can be heard by plugging headphones into the vertical output on the spectrum analyzer since, in this mode, the analyzer displays amplitude variations versus time at a single frequency. There usually is a built-in envelope detector for recovering AM. FM can be recovered by using the slope of the IF filter as a frequency discriminator. A time domain view of a modulating signal using zero scan is shown in Figure 1.16.

Absolute frequency measurements are usually made from the spectrum analyzer tuning dial.

Relative frequency measurements require a linear frequency scan. By measuring the relative separation of two signals on the display, the frequency difference can be

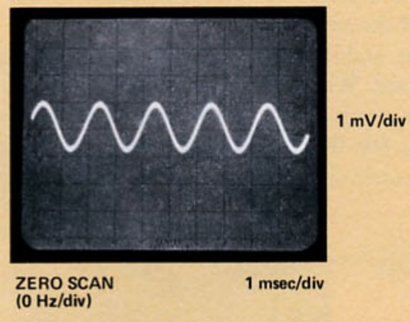


Figure 1.16 This is the modulating wave from an AM signal. The analyzer displays amplitude versus time of the tuned frequency in the zero scan mode.

determined. The linear scan of the spectrum analyzer also means an HP 8406A Frequency Comb Generator can be used to make more accurate absolute frequency measurements. By measuring on the analyzer's CRT display the frequency separation between a known harmonic from the Frequency Comb Generator and an unknown input signal, the frequency of the input signal can be determined. If the analyzer were not sweeping linearly, this would not be possible.

Stability

It is important that the spectrum analyzer be more stable than the signals being measured. The stability of the analyzer depends on the frequency stability of its local oscillators. Stability is usually characterized as either short term or long term. Residual FM is a measure of the short term stability which is usually specified in Hz peak-to-peak. Short term stability is also characterized by noise sidebands which are a measure of the analyzer's spectral purity. Noise sidebands are specified in terms of dB down and Hz away from a carrier in a specific bandwidth. An example of noise sidebands and residual FM is shown in Figure 1.17. Long term stability is characterized by the frequency drift of the analyzer's LO's. Frequency drift is a measure of how much the frequency changes during a specified time (i.e., Hz/min. or Hz/hr).

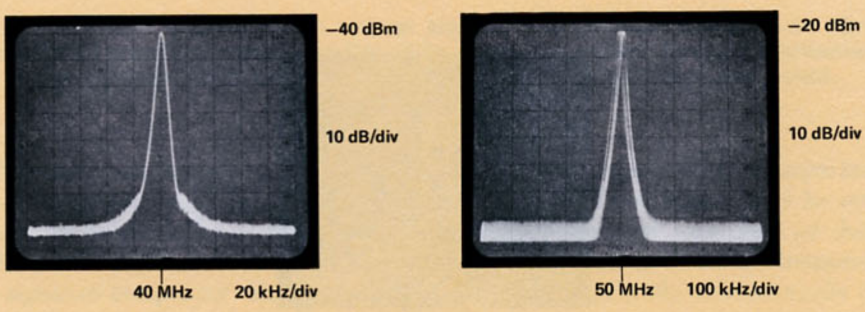


Figure 1.17 A signal with noise sidebands indicating poor spectral purity is shown at left. The residual FM of an oscillator is shown at the right.

Both short term and long term stability can be improved by phase locking the tuned LO to a tooth of a crystal generated frequency comb. The stability of the analyzer is then basically that of the reference crystal oscillator.

Phase locking (or stabilization) should occur automatically whenever a narrow frequency scan is selected. When the tuned local oscillator locks to the crystal reference, some display shift may result due to the pulling of the LO to the lock frequency. If there were no compensation for this, the signal could disappear from the display. In HP analyzers, this shift is compensated for, and no jumping in the display results.

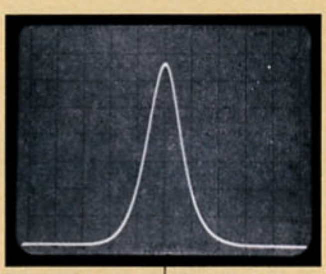


Figure 1.18 As filter is tuned by CW signal, its bandpass shape is traced out on the display.

Resolution

Before the frequency of a signal can be measured on a spectrum analyzer it must first be resolved. Resolving a signal means distinguishing it from its nearest neighbors. The resolution of the spectrum analyzer is limited by its narrowest IF bandwidth. For example, if the narrowest bandwidth is 1 kHz then the nearest any two signals can be and still be resolved is 1 kHz. This is because the analyzer traces out its own IF bandpass shape as it sweeps through a CW signal.

Since the resolution of the analyzer is limited by bandwidth, it seems that by reducing the IF bandwidth indefinitely, infinite resolution will be achieved. The fallacy here is that the usable IF bandwidth is limited by the stability (residual FM) of the analyzer. If the internal frequency deviation of the analyzer is 1 kHz, then the narrowest bandwidth that can be used to distinguish a single input signal is 1 kHz. Any narrower IF filter will result in more than one response or an intermittent response for a single input frequency as shown in Figure 1.19. A practical limitation exists on the IF bandwidth, as well, since narrow filters have long time constants and would require excessive scan time.

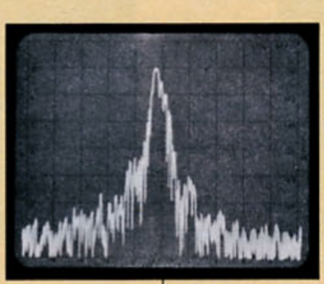


Figure 1.19 Jitter on a signal which is greater than the IF bandwidth.

The resolution of a spectrum analyzer is determined by its IF bandwidth. The IF bandwidth is usually the 3 dB bandwidth of the IF filter. The ratio of the 60 dB bandwidth (in Hz) to the 3 dB bandwidth (in Hz) is known as the shape factor of the filter. The smaller the shape factor, the greater is the analyzer's capability to resolve closely spaced signals of unequal amplitude. If the shape factor of a filter is 15:1, then two signals whose amplitude's differ by 60 dB must differ in frequency by 7.5 times the IF bandwidth before they can be distinguished separately. Otherwise, they will appear as one signal on the spectrum analyzer display. Figure 1.20 shows the bandwidth and shape factor of a typical Gaussian filter.

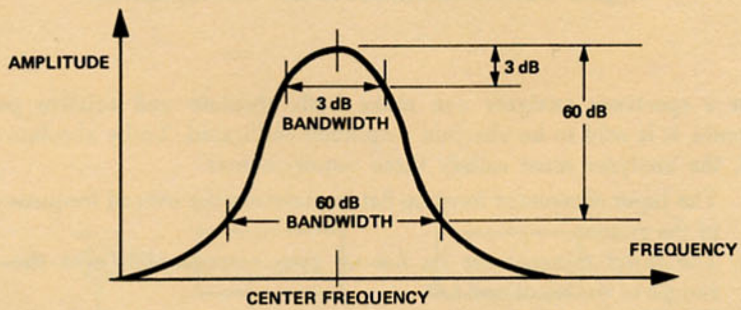


Figure 1.20 Typical Gaussian filter. Signals of equal amplitudes can just be resolved when they are separated by the 3 dB bandwidth. Unequal signals can be resolved if they are separated by greater than half the bandwidth at the amplitude difference between them.

There is a practical limitation on shape factor. The number of poles used in the IF filters will normally determine the shape factor of a synchronously tuned filter. Synchronously tuned (or Gaussian) filters are normally used because of their phase linearity. Shape factor could be improved through the use of stagger-tuned (or square topped) filters. However, such filters have phase discontinuities at their band edges, and, therefore, a ringing is produced when signals are rapidly swept through them as in a spectrum analyzer.

Sometimes IF filter shape factor is specified as the ratio of the 60 dB bandwidth to the 6 dB bandwidth, or the ratio of the 40 dB bandwidth to the 3 dB bandwidth. This can cause difficulty when comparing filters. For example, a 20:1 shape factor measured 60 dB/3 dB is roughly equivalent to a 10:1 shape factor measured 60 dB/6 dB. Even though the shape factor is smaller when specified as the ratio of 60 dB/6 dB, the resolving capability is the same. Therefore, the shape factor is useful as a means of determining filter sharpness only when the dB bandwidths used to determine shape factor are the same for the filters being compared.

The ability of a spectrum analyzer to resolve closely spaced signals of unequal amplitude is not a function of the IF filter shape factor only. Noise sidebands can also reduce the resolution. They appear above the skirt of the IF filter and reduce the offband rejection of the filter. This limits the resolution when measuring signals of unequal amplitude. Noise sidebands can be seen in Figure 1.21.

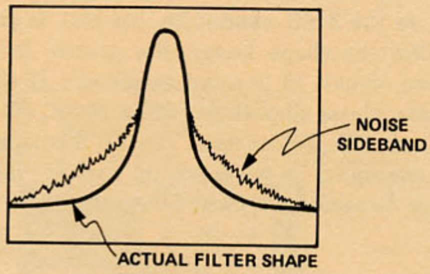


Figure 1.21 Noise sidebands can limit resolution.

Absolute Amplitude Calibration

When a spectrum analyzer can make both absolute and relative power level measurements, it is said to be absolute amplitude calibrated. To be absolute amplitude calibrated, the analyzer must satisfy these requirements:

1. The input attenuator must be flat to maintain the overall frequency response of the system.
2. The input mixer must be flat or gain compensated over the frequency ranges of the input and LO.
3. The IF attenuator must be accurate for proper amplitude display.
4. The log/linear amplifier must be extremely accurate.
5. The sweep times must be slow enough to allow the IF and video filters to respond fully, or an uncalibrated situation will result.
6. Uncalibrated situations should be indicated by an uncal warning light, or avoided by automatic adjustment of the sweep time for any setting of the video or IF bandwidth or frequency sweep width (See Figure 1.22).
7. Must indicate the absolute signal level represented by some point on the CRT for any control settings, and an amplitude internal reference must be supplied.

Figure 1.22 shows the effects of scanning too fast. When the scan time gets too fast for amplitude calibration, the displayed amplitude decreases and the apparent bandwidth increases. Consequently, frequency resolution gets worse.

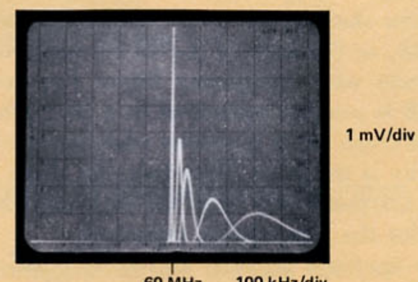


Figure 1.22 When scan time gets shorter and shorter, displayed amplitude gets less and less, and apparent bandwidth gets wider and wider.

Sensitivity

Sensitivity is a measure of the analyzer's ability to detect small signals. The maximum sensitivity of an analyzer is limited by its internally generated noise. This noise is basically of two types: thermal (or Johnson) and nonthermal noise. Thermal noise power can be expressed as: $P_N = kTB$

where: P_N = Noise power in watts

k = Boltzmann's Constant (1.38×10^{-23} joule/°K)

T = absolute temperature, °K

B = bandwidth of system in hertz

As seen from this equation, the noise level is directly proportional to bandwidth. Therefore, a decade decrease in bandwidth results in a 10 dB decrease in noise level and consequently 10 dB better sensitivity.

Nonthermal noise accounts for all noise produced within the analyzer that is not temperature dependent. Spurious emissions due to nonlinearities of active elements, impedance mismatch, etc. are sources of nonthermal noise. A figure of merit, or noise figure, is usually assigned to this nonthermal noise which when added to the thermal noise gives the total noise of the analyzer system. This system noise which is measured on the CRT, determines the maximum sensitivity of the spectrum analyzer. Because noise level changes with bandwidth, it is important, when comparing the sensitivity of two analyzers, to compare sensitivity specifications for equal bandwidths.

A spectrum analyzer sweeps over a wide frequency range, but is really a narrow band instrument. All of the signals that appear in the frequency range of the analyzer are converted to a single IF frequency which must pass through an IF filter; the detector sees only this noise at any time. Therefore, the noise displayed on the analyzer is only that which is contained in the IF passband. When measuring discrete signals, maximum sensitivity is obtained by using the narrowest IF bandwidth.[1]

Video Filtering

Measuring small signals can be difficult when they are approximately the same amplitude as the average internal noise level of the analyzer. To facilitate the measurement, it is best to use some degree of video filtering which is usually adjustable from the front panel of the analyzer. A video filter is a post-detection low pass filter which averages the internal noise of the analyzer. When the noise is averaged, the input signal is easily seen. The effects of video filtering are shown in Figure 1.23.

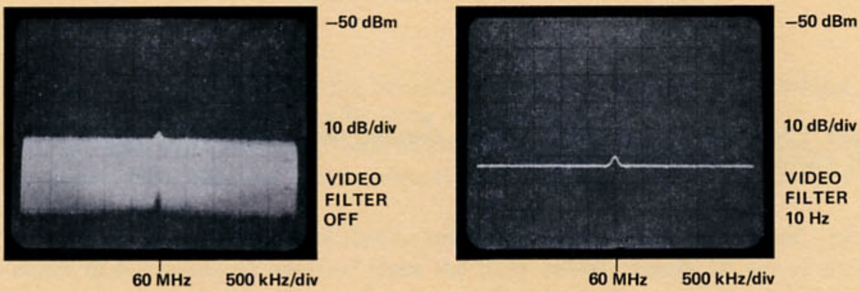


Figure 1.23 How video filter averages noise (right photo) for a better signal-to-noise ratio. No video filtering is used in the left photo.

[1] This is not true, however, for impulse signals such as phase coherent noise. The widest bandwidth gives maximum sensitivity for impulse signals. When measuring random signals such as random or white noise, sensitivity is independent of bandwidth because random input noise varies with bandwidth the same as internally generated noise. See AN 150-4.

Defining Spectrum Analyzer Sensitivity

Specifying sensitivity on a spectrum analyzer is somewhat arbitrary. One way of specifying sensitivity is to define it as the signal level when:

$$\text{signal power} = \text{average noise power}$$

This expression can be rewritten as:

$$\frac{S + N}{N} = 2$$

Where:

S = signal power
N = average noise power

The analyzer always measures signal plus noise. Therefore, when the input signal is equal to the internal noise level, the signal will appear 3 dB above the noise as in Figure 1.24. When the signal power is added to the average noise power, the power level on the CRT is doubled (increased by 3 dB) because the signal power = average noise power. This is the definition of sensitivity used in describing HP spectrum analyzers. Since a 3 dB difference between the signal level and the average noise level is discernible, it is possible to relate this definition to minimum discernible signal.

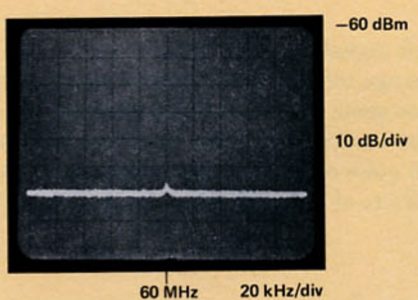


Figure 1.24 When signal power equals average noise power, the signal will appear 3 dB above the average noise level.

Input Signal Level

Figure 1.25 illustrates a typical range of signal levels that can be applied to the input of a spectrum analyzer with no input attenuation, and the analyzer's reaction to these signal levels.

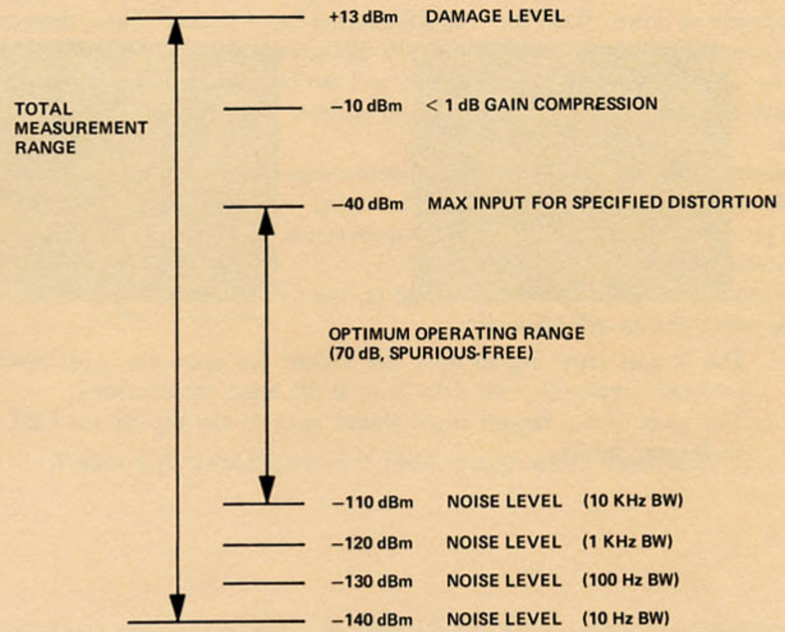


Figure 1.25 Typical spectrum analyzer display range.

The maximum input level to the spectrum analyzer is the damage level or burn-out level of the input circuit. This is typically +13 dBm for the input mixer and +30 dBm for the input attenuator.

Before reaching the damage level of the analyzer, the analyzer will begin to gain compress the input signal. This gain compression is not considered serious until it reaches 1 dB. The maximum input signal level which will always result in less than 1 dB gain compression is called the linear input level. Above 1 dB gain compression the analyzer is considered to be operating nonlinearly because the signal amplitude displayed on the CRT is not an accurate measure of the input signal level.

Whenever a signal is applied to the input of the analyzer, distortion products are produced within the analyzer itself. These distortion products are usually produced by the non-linear behavior of the input mixer. They are typically 70 dB below the input signal level for signal levels not exceeding −40 dBm at the input of the first mixer. To accommodate larger input signal levels, an attenuator is placed in the input circuit before the first mixer. The largest input signal that can be applied, at each setting of the input attenuator, while maintaining the internally generated distortion products below a certain level, is called the optimum input level of the analyzer. For example, a −20 dBm optimum level setting means that all analyzer distortion products are below −90 dBm on the CRT, i.e. down 70 dB. The signal is attenuated 20 dB before the first mixer because the input to the mixer must not exceed −40 dBm, or the analyzer distortion products may exceed the specified 70 dB range. This 70 dB distortion-free range is called the spurious-free dynamic range of the analyzer. The display dynamic range is defined as the ratio of the largest signal to the smallest signal that can be displayed simultaneously with no analyzer distortion products present.

Dynamic range requires several things, then. The display range must be adequate, no spurious or unidentified response can occur, and the sensitivity must be sufficient to eliminate noise from the displayed amplitude range.

The maximum dynamic range for a spectrum analyzer can be easily determined from its specifications. First check the distortion spec. For example, this might be "all spurious products down 70 dB for −40 dBm at the input mixer." Then, determine that adequate sensitivity exists. For example, 70 dB down from −40 dBm is −110 dBm. This is the level we must be able to detect, and the bandwidth required for this sensitivity must not be too narrow or it will be useless. Last, the display range must be adequate.

Notice that the spurious-free measurement range can be extended by reducing the level at the input mixer. For every 10 dB the signal level is reduced, the spurious products will go down at least 20 dB (a net improvement of 10 dB). The only limitation, then, is sensitivity.

To ensure a maximum dynamic range on the CRT display, check to see that the following requirements are satisfied.

1. The largest input signal does not exceed the optimum input level of the analyzer (typically −40 dBm with 0 dB input attenuation).
2. The peak of the largest input signal rests at the top of the CRT display (reference level).

Frequency Response

The frequency response of an analyzer is the amplitude linearity of the analyzer over its frequency range. If a spectrum analyzer is to display equal amplitudes for input signals of equal amplitude, independent of frequency, then the conversion (power) loss of the input mixer must not depend on frequency. If the voltage from the LO is too large compared to the input signal voltage then the conversion loss of the input mixer is frequency dependent and the frequency response of the system is nonlinear. For accurate amplitude measurements, a spectrum analyzer should be as flat as possible over its frequency range.

Flatness is usually the limiting factor in amplitude accuracy since it's extremely difficult to calibrate out. And, since the primary function of the spectrum analyzer is to compare signal levels at different frequencies, a lack of flatness can seriously limit its usefulness.

CHAPTER 2

EXTENDED FREQUENCY COVERAGE THROUGH HARMONIC MIXING

What Is Harmonic Mixing?

To understand harmonic mixing, it's useful to return to the basic block diagram of a spectrum analyzer input section:

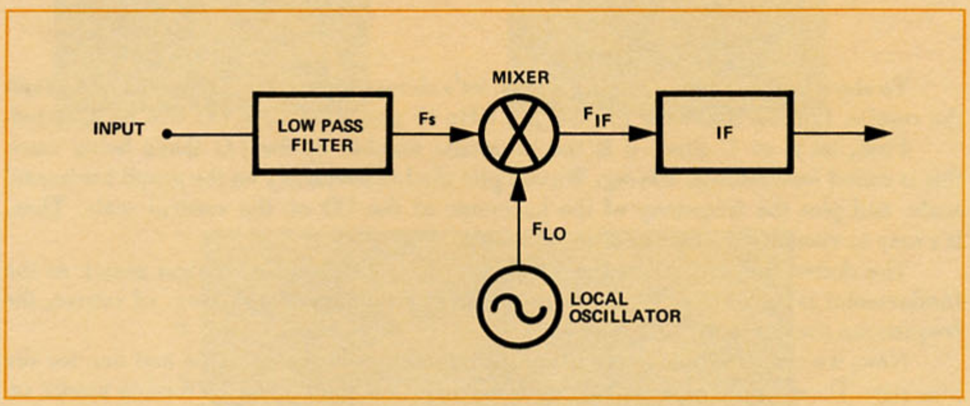


Figure 2.1 Spectrum analyzer Block Diagram.

F_{LO} = Frequency of local oscillator
F_S = Frequency of input signal
F_{IF} = First IF

A response will appear on the analyzer whenever $F_{LO} - F_S = F_{IF}$. This is the basic tuning equation for such a spectrum analyzer. In this case, the first IF and the local oscillator frequencies are *higher* than any input signal to be observed. For example, a 0-1500 MHz spectrum analyzer might be constructed by using a 2 GHz first IF and a 2-3.5 GHz local oscillator.

To extend the frequency coverage of the analyzer, it would be necessary to extend the local oscillator frequency range. However, for coverage into millimeter wave regions an extremely high local oscillator frequency would be required. Clearly, it would be impractical to build, say, a 40-80 GHz oscillator to allow 40 GHz coverage on a spectrum analyzer.

Stability, accuracy, and technology are all limiting factors on high frequency oscillator design.

Well, then, how can frequency be extended? One way would be to remove the low pass filter at the input. In this way the tuning equation could be modified to allow signals higher in frequency than the IF. That is, either equation below would describe the tuning of the analyzer.

$$F_S - F_{LO} = F_{IF}$$
$$F_{LO} - F_S = F_{IF}$$

Or, simplifying:

$$|F_S - F_{LO}| = F_{IF}$$

Then:

$$F_S - F_{LO} = \pm F_{IF}$$

and

$$F_S = F_{LO} \pm F_{IF}$$

Although this more general equation allows extended coverage, there are still limitations due to the available range of local oscillator frequencies.

Consider the effect of creating harmonics of the local oscillator in the input mixer.

This would allow high frequency local oscillator signals while using a lower frequency LO. For example, although a 2-4 GHz LO is used, the 10th harmonic goes from 20 to 40 GHz, thus satisfying the need for a high frequency LO. The tuning equation then becomes:

$$F_S = nF_{LO} \pm F_{IF}$$

This is the general tuning equation for a *harmonic mixing* spectrum analyzer. The input signal can mix with the fundamental or any of the harmonics of the local oscillator to produce the proper IF.

Tuning Range

To show how harmonic mixing works, let's choose some typical numbers and graph the results. For the following graphs, a 2 GHz IF and a 2-4 GHz LO will be assumed.

First, let n = 1. Since n is the harmonic number of the LO that's being used, this is called *fundamental mixing*. We can plot the LO frequency on the graph horizontal scale, and plot the frequency of the harmonic of the LO on the vertical scale. Then, it's easy to visualize the tuning of the analyzer.

The dotted line represents the harmonic of the LO (on the vertical scale) as the fundamental is tuned over its 2-4 GHz range. For fundamental mixing, of course, the frequencies are the same.

Now, we can draw one curve using the minus sign in the equation and one for the plus sign. These would represent signal frequencies for the 1^- and 1^+ mixing modes respectively. The number indicates the harmonic of the LO which is being used, and the plus or minus sign indicates the sign used in the basic tuning equation.

From the basic tuning equation we can see that the curves will always be separated from the dash line (LO frequency) by the IF (2 GHz). Then, there will be two curves for each harmonic of the LO, separated by twice the IF (or 4 GHz).

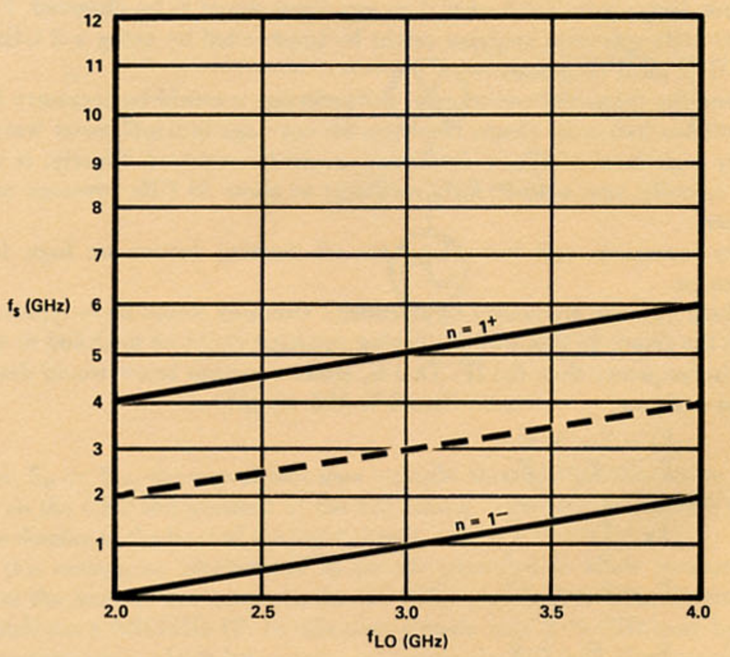


Figure 2.2 Fundamental Mixing.

Second Harmonic Mixing

Now we can readily determine the tuning curves for n = 2, that is, second harmonic mixing. First, we graph the frequency of the second harmonic of the LO as the dashed line. As the LO tunes from 2 to 4 GHz, the second harmonic tunes from 4 to 8 GHz.

The possible tuning frequencies can be represented by two curves, one 2 GHz above the LO and one 2 GHz below. As before, these would be called the 2^+ and 2^- mixing modes. Using second harmonic mixing, then, the analyzer covers the 2 to 10 GHz frequency range.

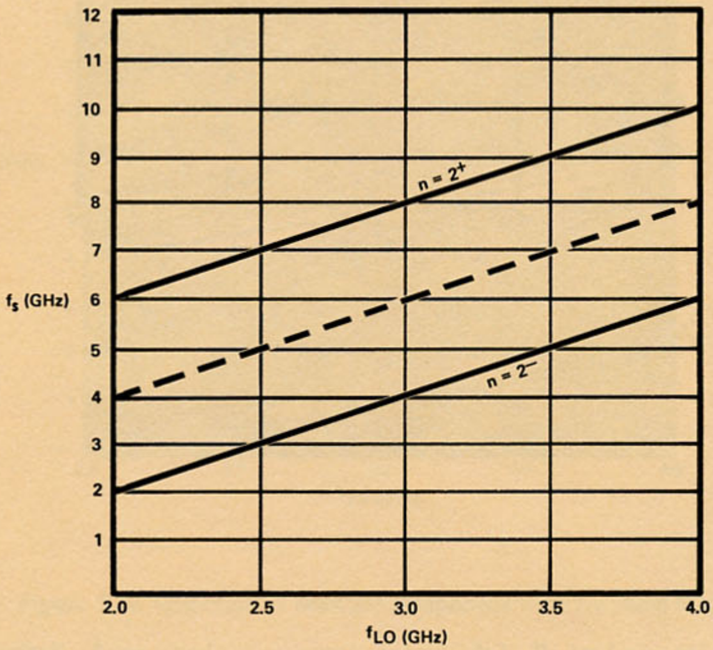


Figure 2.3 Second Harmonic Mixing.

Third Harmonic Mixing

Extending one step further, it's easy to see that the third harmonic of the LO is 6 to 12 GHz, and the tuning curves for third harmonic mixing are as shown.

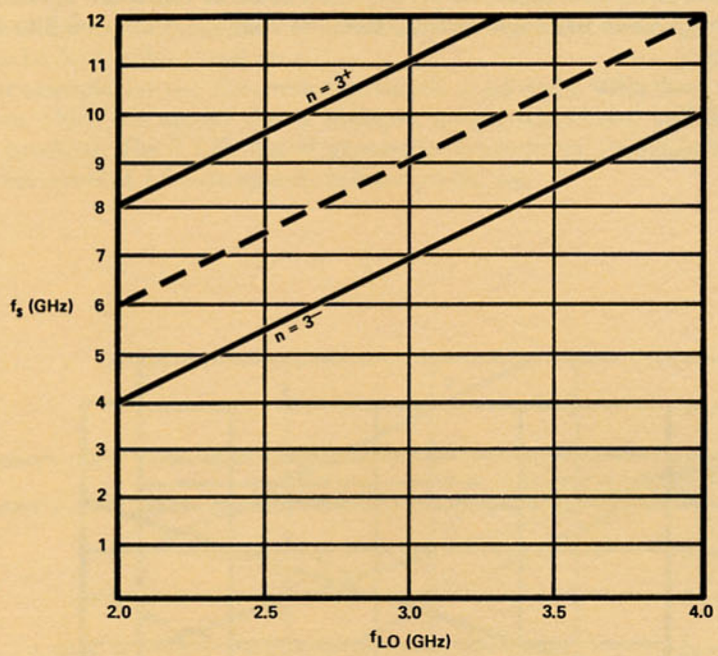


Figure 2.4 Third Harmonic Mixing.

Composite Tuning Curves

Now we can combine all of the tuning curves on a single graph. Notice that by using the first three harmonics of the LO we can measure input signals from 0-12 GHz. In modern analyzers, harmonics up through the tenth harmonic are used to extend the usable frequency range beyond 40 GHz.

The total scan on any frequency band (or harmonic mixing mode) depends on the harmonic of the LO in use. For example, on third harmonic mixing, a total scan of 6 GHz is possible (4 to 10 GHz), while on fundamental mixing, only a 2 GHz scan can be obtained. The frequency scan on any band is equal to the frequency scan of the LO times the harmonic number.

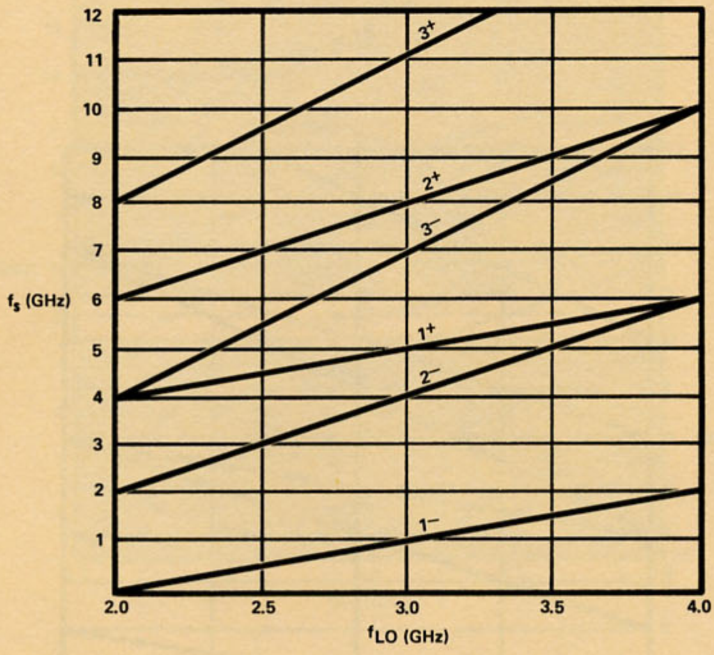


Figure 2.5 Composite Tuning Curves.

The horizontal (frequency) axis of the analyzer CRT is analogous to the local oscillator frequency since the CRT is swept by the same voltage ramp which tunes the LO. That is, on the display, signals which mix with a lower LO frequency will appear to the left of signals which mix with a higher LO frequency. To accurately identify the frequency of a signal on the CRT, then, it is important to know which harmonic mixing mode was used since there are several possible responses for each position on the CRT. In fact, a harmonic mixing analyzer with a "wide open" front end scans simultaneously through all its frequency bands.

Signal Identifier

A *signal identifier* is usually provided to allow identification of the proper mixing mode. The signal identifier shifts a local oscillator (other than the first) by a known amount. Then, the shift of the signal on the display will be proportional to the harmonic number for any mixing mode; the direction of the shift will depend on the plus or minus sign in the tuning equation. In this way, positive identification of any response can be obtained.

Unwanted Responses

Harmonic mixing, as we have seen, allows coverage to higher frequencies. Unfortunately, it also causes some other problems. These, however, can be solved by simple means.

The first of these is *image responses.*

> IMAGE RESPONSES — Response of the analyzer at one LO frequency to input signals of more than one frequency. A pair of image responses exists for each LO frequency and its harmonics.

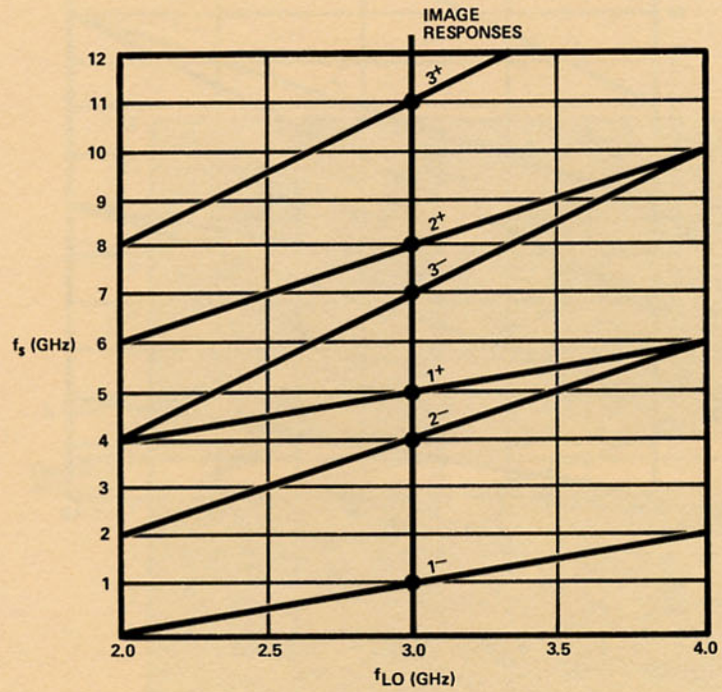


Figure 2.6 Image Responses.

For example, if we have input signals at 1 and 5 GHz, these will both mix with a 3 GHz LO frequency and produce a response at the same place on the CRT. So, there can be several input signals which can appear to be at the same frequency.

The signal identifier would allow proper identification of each response, but it would be better to only have the desired response appear. This can be accomplished with a bandpass filter. In this way, all but the desired signal would be rejected and not appear. In the figure below, a 5 GHz bandpass is used to allow viewing a 5 GHz signal; any 1 GHz present at the input would be rejected.

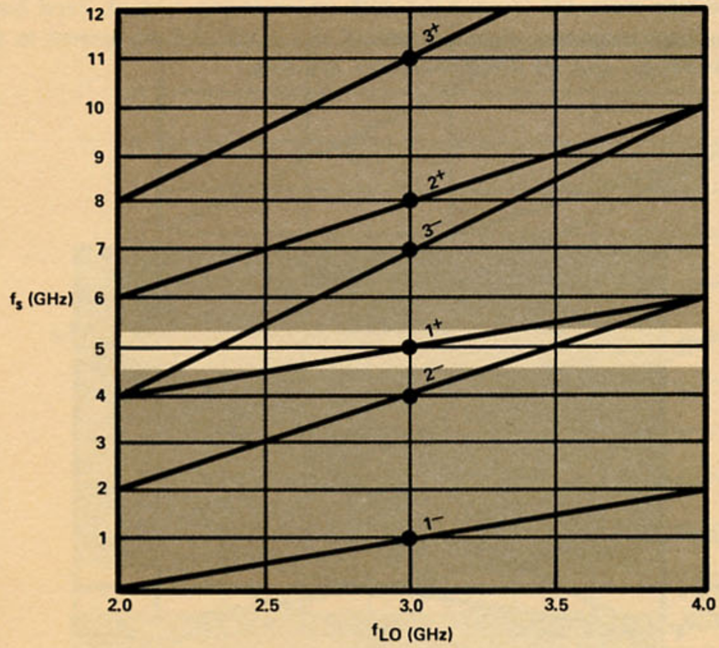


Figure 2.7 Rejection of Image Responses with Bandpass Filter.

Multiple Responses

MULTIPLE RESPONSES — Responses of the analyzer at more than one LO frequency to an input signal at a single frequency.

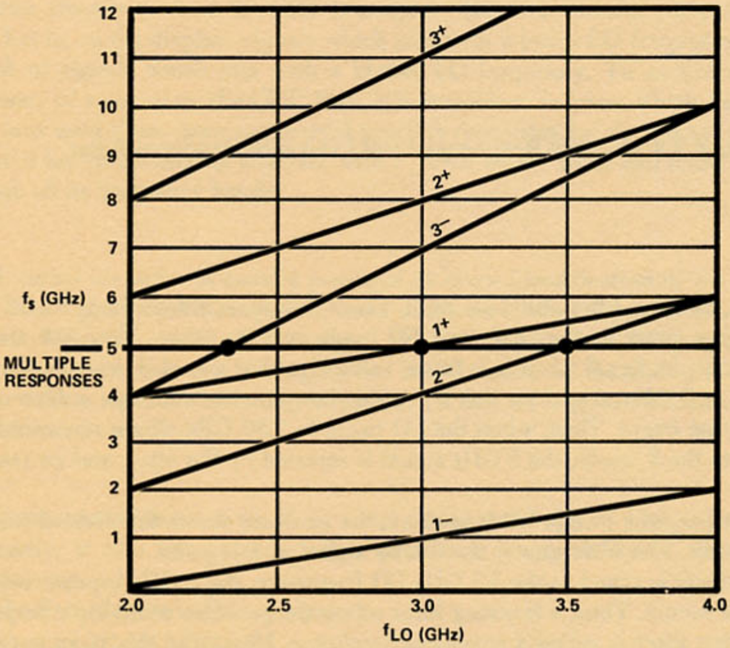


Figure 2.8 Multiple Responses.

This means, for example, that a 5 GHz signal could appear in three places on the CRT corresponding to the three LO frequencies with which it can mix. A fixed bandpass will not cure multiple responses since the desired signal will still be present as the LO is tuned over its range, and all three responses are seen.

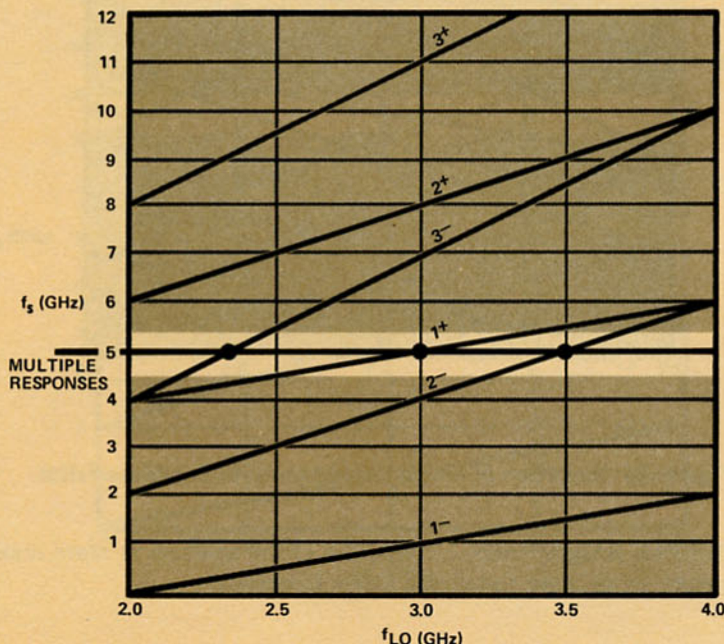


Figure 2.9 Multiple Responses Exist Even with Bandpass Filter.

Tracking Filters

Suppose a tracking filter could be built, i.e., a bandpass filter which would track the desired tuning curve as the analyzer's LO tunes over its range. Now, the analyzer will only see an input signal when it is tuned to receive that signal. For example, if we wished to view a 5 GHz signal on the 1^+ mixing mode, we could use a filter which tracked this tuning curve. Then, when the LO tunes to 2.33 GHz where we would have a response due to the 3^- mode, the 5 GHz signal is rejected by the filter, and no response appears.

Tuning further, to a 3 GHz LO frequency, the analyzer shows the desired response to the 5 GHz input. The filter passes the 5 GHz signal at this point.

As the analyzer is tuned to the 3.5 GHz LO frequency, the 5 GHz input is rejected, and no response occurs. Thus, a tracking filter solves the problem of multiple responses.

This tracking filter is called a *tracking preselector*. Normally, this is an automatic device which tracks the desired tuning curve when the appropriate frequency band is selected.

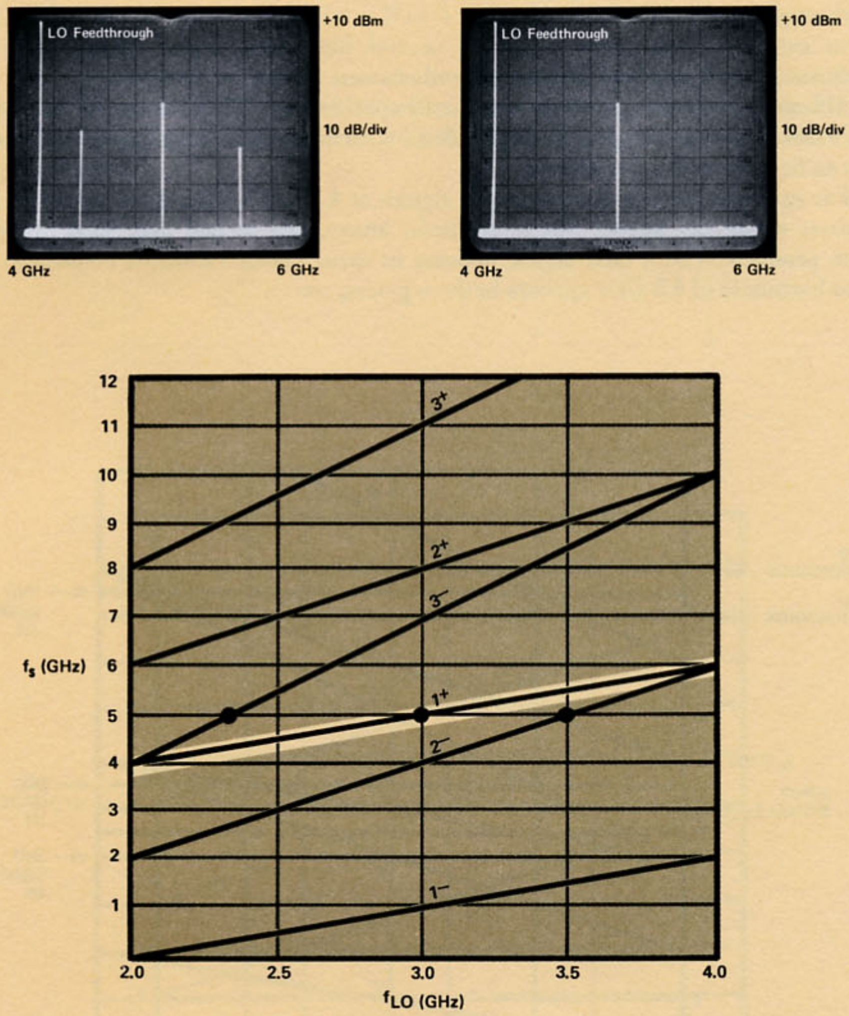


Figure 2.10 Rejection of Multiple Responses with Tracking Filter.

The preselector also eliminates any image responses from the display. Note that the preselector is most effective near the middle of any band where the tuning curves are spaced farthest apart.

Also, the use of the preselector means that wide scans over the full LO range can be made easily since all signals appearing on screen are from the selected band. So, for example, on fourth harmonic mixing, you can scan 8 GHz, etc.

Spurious Responses

Another problem encountered with a spectrum analyzer is spurious responses.

SPURIOUS RESPONSES — Responses of the analyzer to input signals which are created due to non-linear behavior of the spectrum analyzer.

An example of a spurious response is harmonic distortion. If a 4 GHz signal (with no distortion) is used as in input, the input mixer of the spectrum analyzer may generate some second harmonic distortion at 8 GHz. This would appear on the CRT and be a spurious response.

In extreme cases, there might be several high level input signals, and their harmonics plus all sum and difference products can appear on screen.

Although distortion products are usually specified as >70 dB down for some input level, usable dynamic range can be extended by further eliminating spurious responses. This can be done with a preselector.

For example, for two distortion-free signals at 4.5 and 5 GHz, there are 18 possible responses which can appear due to multiple, image, and second and third order distortion products. The 5 GHz signal appears in three places (multiple responses), the second harmonic of 4.5 GHz appears in three places, etc.

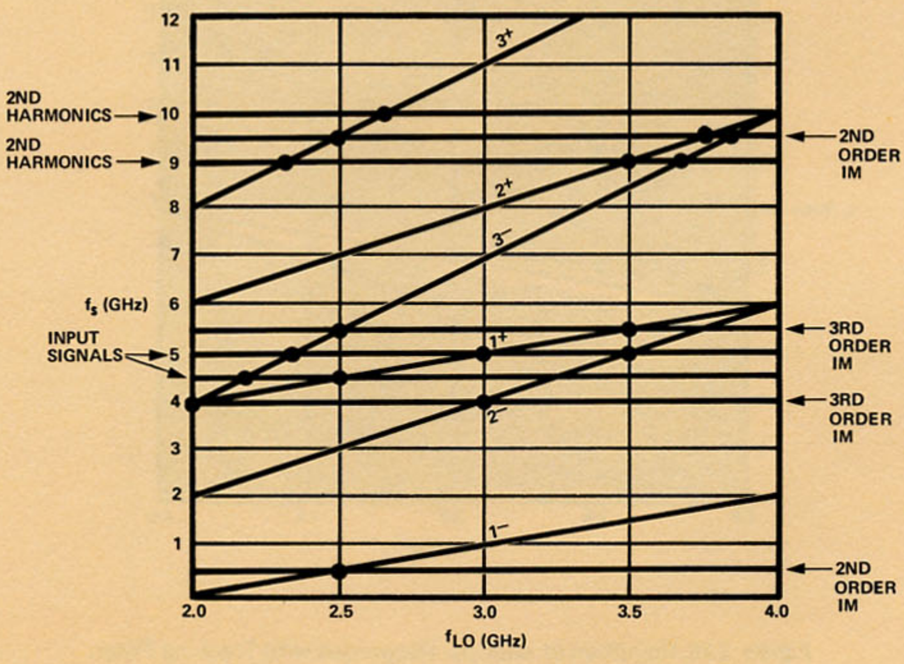


Figure 2.11 Spurious Responses to 4.5 and 5.0 GHz Inputs.

If a preselector is used, however, there are responses to only the two desired input signals. This is because each signal is passed into the mixer only at the precise time the analyzer is tuned to receive it. So, when the analyzer is tuned to the 9 GHz harmonic of 4.5 GHz, the 4.5 GHz is not present at the mixer to generate distortion. Only signals actually present are displayed and only on the selected band. A distortion-free range of greater than 100 dB is achievable with a preselector.

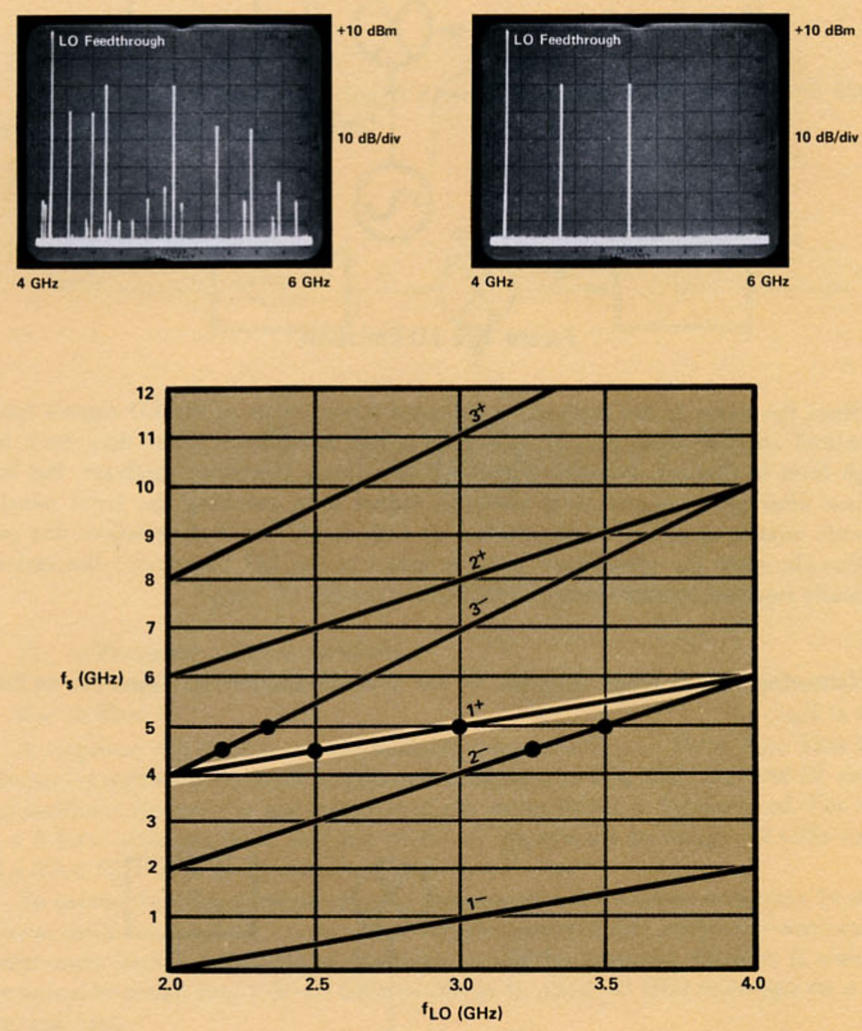


Figure 2.12 Elimination of Spurious Responses with Tracking Filter.

LO Emission

Once again, the input section of a harmonic mixing analyzer will help illustrate a point.

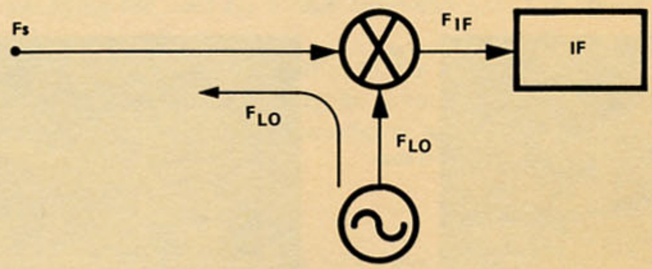


Figure 2.13 LO Emission.

Since the mixer is not typically a balanced mixer, some of the LO signals can be transmitted out the input port at a level of about a milliwatt. This may be a high enough level to damage a sensitive device. On a lower frequency analyzer, the input low pass filter prevents any local oscillator signal from reaching the front panel. A harmonic mixing analyzer, of course, has no such filter. Preselection solves this problem, though, since the preselector is never tuned to the LO frequency; the emission is typically rejected by more than 70 dB.

IF Feedthrough

Returning to the block diagram of the input, consider the effect of an input signal at F_{IF}.

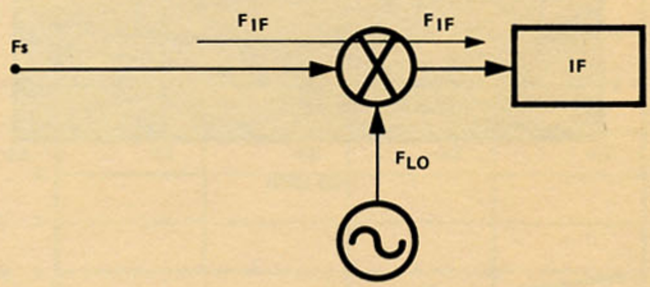


Figure 2.14 IF Feedthrough.

When $F_S = F_{IF}$, the signal can pass straight through the IF section and give a response on the CRT independent of the LO tuning. This would raise the entire baseline of the display to the level of the input signal.

For this reason, an alternate IF is usually provided to allow measurement of a signal near the first IF. For example, the HP 8555A has a 2.05 GHz IF except for two bands which use a 550 MHz IF. In this way, signals near 2.05 GHz may be measured.

If measurements at high frequencies are required while 2.05 is present at the input, a notch filter may be used to eliminate the undesired signal. Note, however, that the tracking preselector would also reject the undesired 2.05 GHz at the input when the analyzer is tuned to any other frequency.

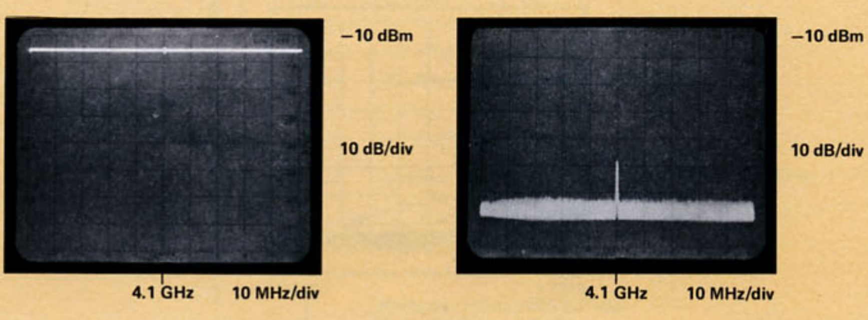


Figure 2.15 Elimination of IF Feedthrough with Preselector.

Using the Signal Identifier with a Preselector

When using a preselector, the signal identifier is not normally required. However, since the tracking filter has only finite rejection, there are cases where signal identification is mandatory to assure a valid measurement.

For example, assume a measurement of spurious outputs from a 5 GHz oscillator is to be made. The display of this signal would look like this if no preselector were used.

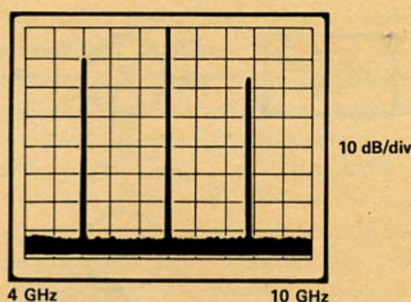


Figure 2.16 Multiple Responses to Strong 5 GHz Signal.

Now, if the preselector is used, we would expect only the desired response to appear. But since typical rejection of the preselectors commonly in use is only about 80 dB, we could see a display similar to this:

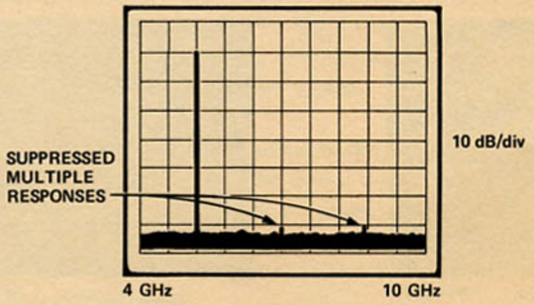


Figure 2.17 Feedthrough of Multiple Responses with Preselector Requires Signal Identifier for Positive Identification.

The only way to be certain that these responses are not spurious outputs would be to use the signal identifier to positively determine the mixing mode. Otherwise, a strong signal will appear as a small signal on each mixing mode and will hinder measurements where wide dynamic range is required.

Preselection for Low Frequency Analyzers

Tracking preselectors are not normally available for low frequency spectrum analyzers. There are some good reasons for this.

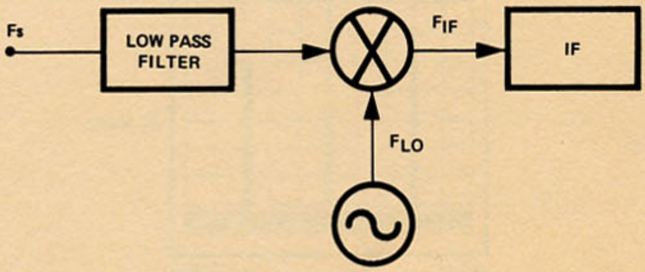


Figure 2.18 Low Frequency Analyzer Using Only 1^- mode.

By reviewing the basic block diagram for a low frequency analyzer, it's obvious that only the 1^- mixing mode is possible. That is, both F_{LO} and F_{IF} are higher than the cut off frequency of the low pass filter. This means there can be no multiple or image responses. Also, IF feedthrough and LO emission are both prevented by the low pass filter.

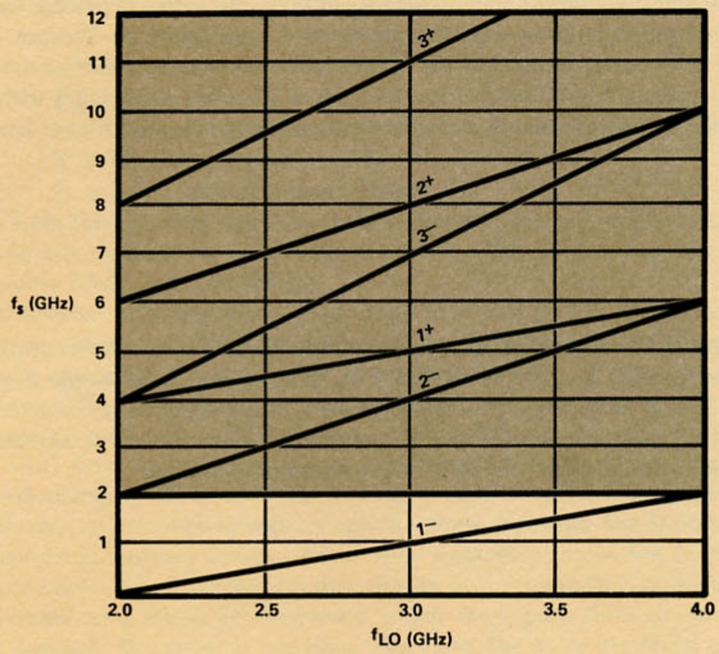


Figure 2.19 Only 1⁻ mode is possible with Low Pass Filter at Input.

Another limitation imposed is the frequency range covered by a low frequency analyzer may extend over several decades, making it extremely difficult to build tunable filters over the same range.

Thus, even though performance of the analyzer could be improved due to better rejection of spurious responses, preselectors are not normally used. However, fixed filters are used in some special applications to obtain the same effect.

Adverse Effects of Preselectors

Up to now we have discussed the benefits derived from the use of a tracking preselector. Obviously, there must also be some negative points.

Since the filter used must have some insertion loss, there will be a loss in sensitivity of the analyzer. Also, this loss is not constant with frequency so some degradation of overall flatness results.

Flatness can also be degraded due to mismatch between the preselector and the input mixer of the spectrum analyzer since the filter reflects nearly all power outside its passband. To correct this, attenuation is usually used between the preselector and mixer. However, this further degrades sensitivity.

A further reduction in flatness can result if the preselector does not accurately track the spectrum analyzer tuning.

Where maximum sensitivity and flatness are required, it is advantageous to remove the preselector from the circuit. But, for most applications, it is sufficient to remember a few simple rules: use at least 10 dB of input attenuation whenever possible, and check the tracking adjustments of the preselector periodically.

Spectrum Analyzer Gain versus Mixing Mode

As the spectrum analyzer uses higher harmonics of the LO, less oscillator power is available. This means that the conversion loss of the input mixer is greater for higher harmonics. The gain of the spectrum analyzer is compensated to account for this effect. Thus, it is necessary to identify the proper harmonic mixing mode before making any amplitude measurements. Also, since the gain of the analyzer changes with mixing mode, the displayed noise level will also change; sensitivity is less on higher harmonics.

Stability on Higher Harmonics

As the first LO is multiplied in frequency, any instability is also multiplied. For example, if the first LO has 30 Hz residual FM, then on fourth harmonic mixing, the FM would be 30 x 4, or 120 Hz. This effect can be a determining factor in selecting the minimum usable resolution bandwidth for any mixing mode. Excellent LO stability is necessary if high resolution is needed on higher harmonic mixing.

CHAPTER 3
TRACKING GENERATORS

The tracking generator is a special signal source whose RF output frequency tracks (follows) some other signal beyond the tracking generator itself. In conjunction with the spectrum analyzer, the tracking generator produces a signal whose frequency precisely tracks the spectrum analyzer tuning. Because of this feature, the two instruments combine to make a powerful and versatile measurement system. However, before going into the measurement capabilities of the system, look briefly at the simplified block diagram of a spectrum analyzer/tracking generator as shown in Figure 3.1. The spectrum analyzer portion is the same block diagram described in the previous sections.

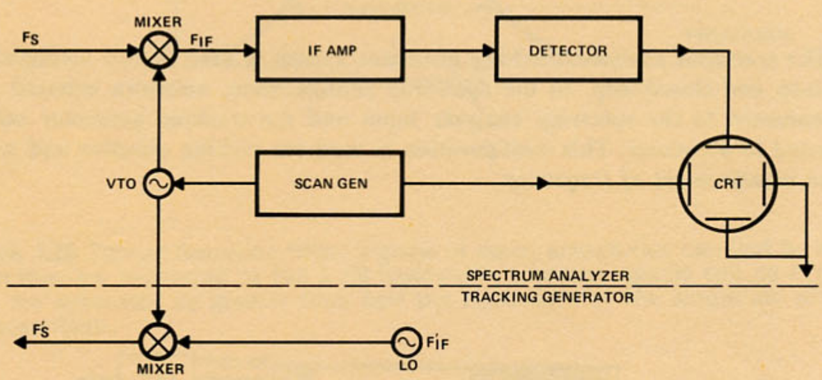


Figure 3.1 Simplified Spectrum Analyzer/Tracking Generator.

The tracking generator has a stable, fixed (F'_{IF}) local oscillator. If we make $F'_{IF} = F_{IF}$ then $F_S = F'_S$ where

$$F_s = F_{vto} - F_{IF} \text{ and } F'_s = F_{vto} - F'_{IF}.$$

Hence, the tracking generator frequency (F'_s) precisely tracks the spectrum analyzer tuning (F_s) since both are effectively tuned by the same VTO. This precision tracking exists in all analyzer scan modes. Thus, in full scan, the tracking generator output is a start-stop sweep, in per division scan the output is a ΔF sweep, and in zero scan the output is simply a CW signal.

The tracking generator signal is generated by synthesizing and mixing two or more oscillators. One oscillator is part of the tracking generator itself, the other oscillators are brought via an interface cable from the spectrum analyzer. Depending on the analyzer, the interface cable carries two or three LO's.

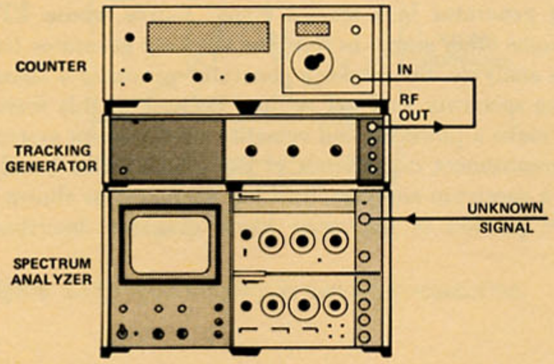


Figure 3.2 Open-Loop Configuration.

The spectrum analyzer/tracking generator system is used in two configurations: open-loop and closed-loop. In the open-loop configuration, unknown external signals are connected to the spectrum analyzer input and the tracking generator output is connected to a counter. This configuration is used for making *selective and sensitive precise measurement of frequency.*

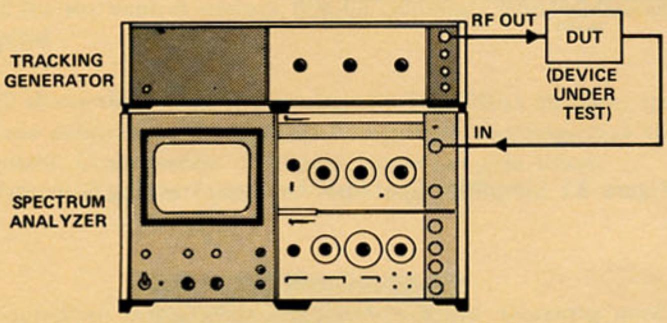


Figure 3.3 Closed-Loop Configuration.

In the closed-loop configuration, the tracking generator signal is fed into the device under test and the output of the device under test is connected to the analyzer input. This configuration is used for making *magnitude only swept transmission/reflection measurements;* i.e., insertion loss, return loss, SWR, and reflection coefficient. In the following pages, these two configurations will be discussed in some detail.

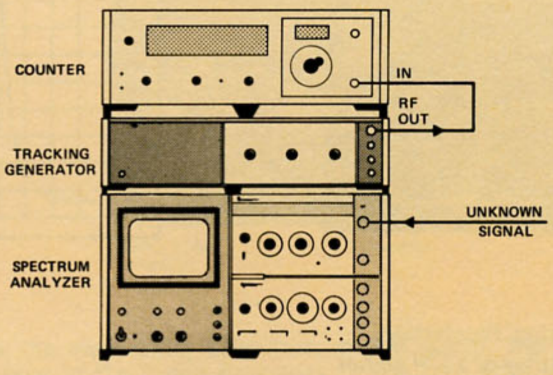


Figure 3.4 Precise Measurement of Frequency.

Figure 3.4 is an open-loop configuration where the spectrum analyzer/tracking generator is used as a highly selective, highly sensitive, precision frequency measurement system. The CRT photo shown in Figure 3.5 is an example of the measurement capability of this system.

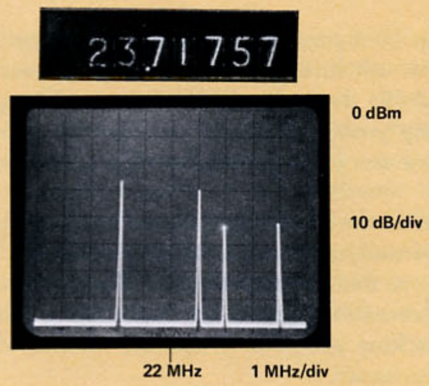
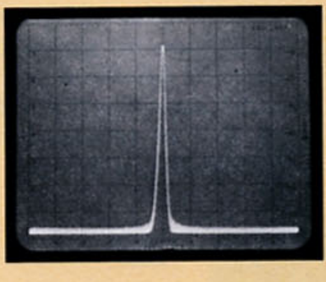


Figure 3.5 This CRT display shows several signals. The smallest signal (—44 dBm) was counted by placing a special marker at its peak. The marker forces a momentary pause in the analyzer scan, the pause allows the counter to trigger and count.

When unknown signals are connected to the analyzer input, they are separated into their frequency components and displayed on the CRT. These signals may be very small and undetectable by a counter or they may be large, but regardless of their level, the tracking generator output is enough to drive the counter. Thus, the frequency of any signal or component can be measured by simply *stopping the scan manually or electronically* at the peak of the displayed signal and counting the tracking generator output frequency as shown in Figure 3.6. This measurement would not have been possible if precision tracking did not exist between the analyzer tuning and the tracking generator frequency.



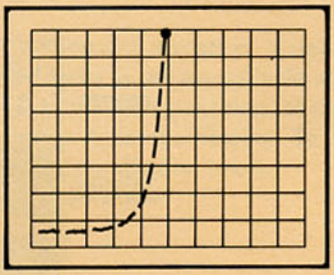


Figure 3.6 Counting Illustration.
A. This is the display of a CW signal.
B. To count the frequency shown in A, stop the scan manually or electronically at the peak of the signal and read the counter display.

The question that should come to mind at this point is, "Why can't we count the unknown signal by connecting it directly to the counter?" A partial answer to this question is yes, we can, but only for spectrally clean, stable, and high level signals. The complete answer will be apparent below when the system sensitivity and selectivity have been considered. But before doing this, let us establish that the indirect counting technique (counting the tracking generator frequency instead of the unknown signal itself) does not diminish the frequency accuracy of the system as a whole.

Frequency Accuracy

In the open-loop configuration, frequency accuracy is limited by two factors: (1) the tracking error which exists between the analyzer tuning and the generator output frequency, and (2) the residual FM of the tracking generator output frequency.

Since the tracking generator precisely tracks the analyzer tuning, tracking error is seldom worse than a few parts per million. Thus, the tracking error contribution is minimal.

The residual FM of the tracking generator is made up of the residual FM of the local oscillators in the analyzer and the tracking generator. These oscillators have very low residual FM, due to internal stabilization. Hence, the contribution of the residual FM of the tracking generator signal is also minimal.

Counting the tracking generator frequency instead of the unknown signal itself is potentially highly accurate.

Sensitivity

Any signal displayed on the analyzer CRT, from the analyzer burnout to the noise level, can be counted. This simply means that this is a counting system with a sensitivity identical to that of the analyzer.

The significance of the tracking generator in this application is that its output level is sufficient to drive a counter regardless of the level of the unknown signal, and therefore it effectively transfers the analyzer sensitivity to the counter. Furthermore, since the analyzer sensitivity is bandwidth dependent, we can increase sensitivity by narrowing the analyzer's IF bandwidth. Very small signals, on the order of tens of nanovolts (10^{-9}), can be displayed and their frequency counted. By contrast to a stand-alone counter, this sensitivity far exceeds available counter sensitivity and is equivalent to about 60 to 70 dB improvement. Figure 3.7 shows a signal at −110 dBm which was counted with the help of a tracking generator. There's no way to count this signal directly.

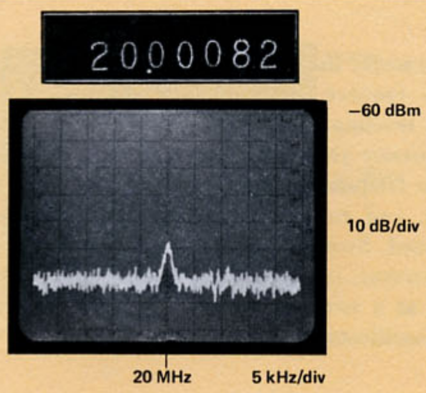


Figure 3.7 This signal is −110 dBm. It was counted using a spectrum analyzer/tracking generator.

Selectivity

The spectrum analyzer has several scan widths and several IF bandwidths, both ranging from very narrow to very broad. Adjusting the scan width controls the portion of the spectrum being scanned and displayed on the CRT; and adjusting the IF bandwidth controls the minimum separation required for resolving two closely spaced signals.

Another facet of interest is that the analyzer is amplitude-selective through its filter skirt. Signals differing in amplitude by as much as 70 dB and separated by some minimum frequency can be resolved and displayed simultaneously.

The spectrum analyzer/tracking generator makes a highly selective frequency measurement system as shown in Figure 3.8. The end limitation to selectivity is what can be resolved on the CRT. Here again, if a signal can be resolved, its frequency can be measured; simply stop the scan manually or electronically at the peak of the desired signal and read the counter display.

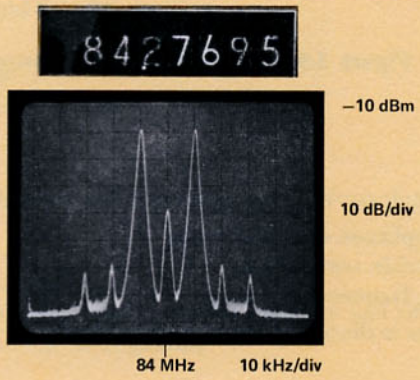


Figure 3.8 This is a suppressed carrier spectrum. The carrier frequency was counted as shown. Notice the selectivity of the spectrum analyzer/tracking generator system.

In summary, the fundamental signal, harmonics, sidebands (including FM and PM signals), and non-random spurious signals can be indirectly counted by adjusting the analyzer scan width and IF bandwidth. In this application, the counter effectively acquires the sensitivity and selectivity of the spectrum analyzer. The conventional counter, by contrast, is triggered by the highest level signal, thus affording little sensitivity and selectivity, if any.

Closed-Loop System: Swept Frequency Measurement (Device Characterization)

In this configuration, the spectrum analyzer/tracking generator becomes a self-contained, complete (source, detector, and display) swept frequency measurement system. An internal leveling loop in the tracking generator ensures a leveled output over the entire frequency range. The specific swept measurements that can be made with this system are frequency response (amplitude vs. frequency), magnitude only reflection coefficient, and return loss. From return loss or reflection coefficient, the SWR can be calculated. Swept phase and group delay measurements cannot be made with this system; however, it does make some unique contributions not made by other swept systems, such as a sweeper/network analyzer, a sweeper/spectrum analyzer, or a sweeper/detector/oscilloscope.

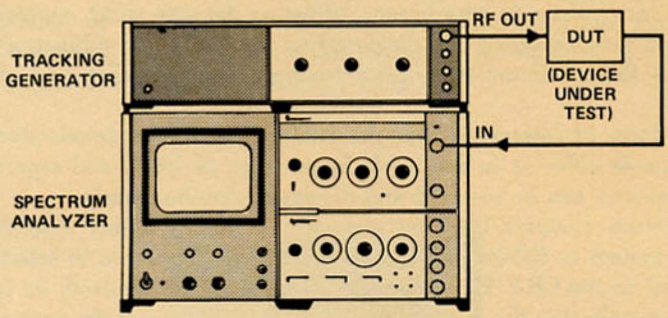


Figure 3.9 Swept Frequency Measurements.

The spectrum analyzer/tracking generator has three attractive features:

1. Large amplitude dynamic range
2. Narrow/wide scan widths
3. Excellent frequency accuracy

Each of these features is discussed separately on the following pages.

Amplitude Dynamic Range

The measurement dynamic range of the spectrum analyzer/tracking generator may exceed 120 dB. The key to this large dynamic range is the precision tracking which exists between the analyzer tuning and the generator RF output frequency. This precision tracking permits over-driving the mixer to the 1 dB compression level while keeping the analyzer distortion products off the CRT. Figure 3.10 shows the dynamic range capability of the spectrum analyzer/tracking generator system.

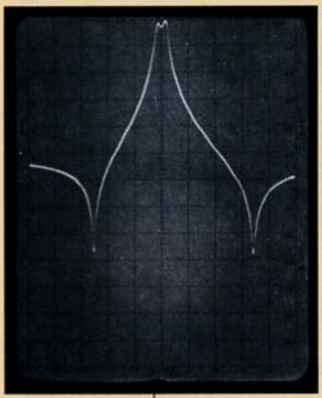


Figure 3.10 This CRT photograph shows the frequency response of a 278 kHz bandpass filter over a range greater than 85 dB.

Precision tracking means at every instant of time the generator fundamental frequency is in the center of the analyzer passband, and all generator harmonics, whether they are generated in the analyzer or are produced in the tracking generator itself, are outside the analyzer passband. Thus, only the tracking generator fundamental frequency is displayed on the analyzer's CRT. Second and third order harmonics and intermodulation products are clearly out of the analyzer tuning and, therefore, they are not seen. Thus, while these distortion products may exist in the measurement set-up, they are completely eliminated from the CRT display.

The 1 dB gain compression level is a point of convenience, but it is nonetheless considered the upper limit of the dynamic range. The lower limit, on the other hand, is dictated by the analyzer sensitivity which, as we know, is bandwidth dependent. The narrowest usable bandwidth in turn is limited by the tracking generator residual FM and any tracking drift between the analyzer tuning and the tracking generator signal.

The measurement dynamic range obtained with most other techniques does not exceed 60 to 80 dB, whether a network analyzer, a spectrum analyzer (with non-synchronous source) or a detector/oscilloscope is used for detection and display. The limitation differs with each display. For the network analyzer, the limitation is noise generated in the analyzer channels; for the spectrum analyzer, it is distortion created in the spectrum analyzer by the sweeper signal (since the sweeper is not synchronously tuned with analyzer), and for the detector/oscilloscope, it is the detector square law operation.

Narrow/Wide Scan Widths

Because of precision tracking, the tracking generator output frequency acquires the same scan capability as the spectrum analyzer. Hence, the analyzer's several calibrated scan widths which range from broadband to extremely narrow band are acquired by the tracking generator. However, it is the tracking generator's inherently low residual FM (<1 Hz below 110 MHz, <200 Hz above) which makes narrow scanning meaningful and usable. Figure 3.11 shows a 200 Hz scan and a 100 MHz scan using the spectrum analyzer/tracking generator system.

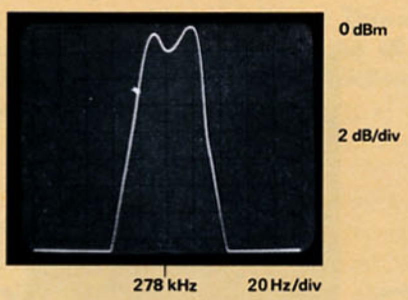


Figure 3.11 Narrow and wide scanwidths of the spectrum analyzer/ tracking generator. The narrow scan width photo shows the passband details of the filter shown in Figure 3.10 on previous page.

The large dynamic range discussed above makes the narrow scan capability particularly significant. Very narrow band, high "Q" devices can be swept frequency tested with excellent amplitude and frequency resolution. The wide scan width capability, on the other hand, makes it possible to view the frequency response over a wide frequency range.

By contrast, the residual FM of most conventional sweepers in this range is a few kHz (3-10 kHz); thus it precludes very narrow band measurements.

Frequency Accuracy

In the closed-loop configuration, the CRT displays the frequency response of the device under test. Frequency accuracy in this configuration means the ability to identify and measure the frequency of any point on the CRT display.

The amplitude dynamic range and the scan capability of the spectrum analyzer/ tracking generator make identification of any point on the display quite easy and unambiguous. Thus, we can identify the 3 dB points, 60 dB points, center frequency, etc. The frequency of any point on the display is measured by connecting a counter to the tracking generator output and stopping the scan (manually or electronically) at the point of interest. Figure 3.12 shows a 50 MHz bandpass filter. Its 3 dB and 60 dB bandwidths were counted.

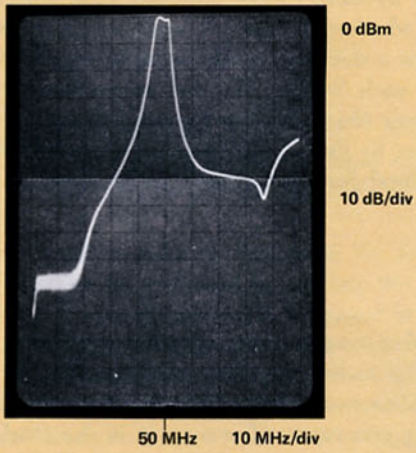


Figure 3.12 This CRT photograph shows the frequency response of a 50 MHz bandpass filter over a range greater than 90 dB.

The accuracy of this measurement is limited by the tracking generator residual FM plus the counter's own accuracy. Excluding the counter contribution, this accuracy ranges from 3 to 200 Hz for Hewlett-Packard spectrum analyzer/tracking generator systems. By comparison, the accuracy obtained from counting most conventional sweepers varies from 3 to 10 kHz. Here again, the limitation is the sweeper residual FM, which is usually several orders of magnitude greater than that of the tracking generator.

It should be pointed out that precision tracking does not affect frequency accuracy in the closed-loop configuration. To see why this is true, let us consider the case where tracking is deliberately degraded. Precision tracking places the tracking generator signal in the center of the IF bandwidth. Thus, when tracking is degraded, the tracking generator frequency shifts from the center of the IF bandwidth and the analyzer response simply decreases in amplitude, and a slight shift in the display results.

The CRT photograph in Figure 3.13 demonstrates this result. Curve A is the frequency response of a bandpass filter and is taken at maximum tracking. Curve B is for the same filter, but tracking has been deliberately degraded. Notice how in curve B the entire display has shifted downward, and that the center frequency, 3 dB points, 60 dB points, etc., are still preserved. This is true because of counting the frequency of the applied stimulus; i.e., the same response must always occur at the same stimulating frequency, even though the display itself may shift horizontally.

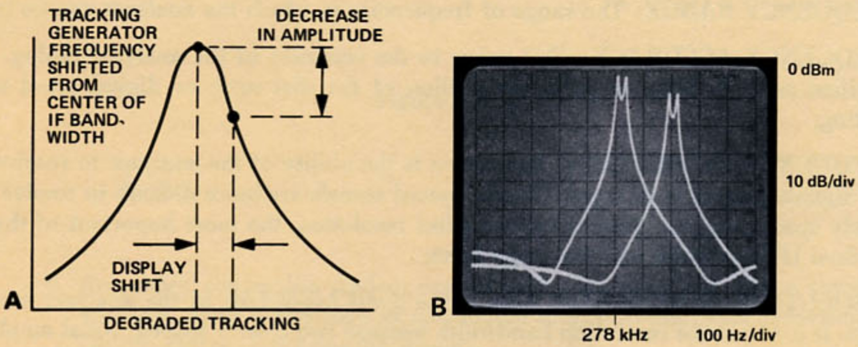


Figure 3.13 Degraded Tracking Effect.
A. The tracking generator frequency shifted from the center of the IF bandwidth.
B. CRT display with maximum and degraded tracking.

GLOSSARY OF SPECTRUM ANALYSIS TERMS

Spectrum analysis is defined as the study of energy distribution as a function of frequency. From this study comes valuable information for engineers designing a variety of systems and components; such as communication systems, control systems, radar and navigation systems and components for any of these systems.

SPECTRUM ANALYZER: The spectrum analyzer is an instrument designed to graphically present the energy distribution of a signal as a function of frequency. Typically, the display is a cathode ray tube (CRT) screen which shows amplitude vs. frequency. There are two types of spectrum analyzers; real-time and non-real-time. Non-real-time analyzers are known as scanning analyzers.

> REAL-TIME SPECTRUM ANALYZER: This is a spectrum analyzer which is tuned to the entire spectrum at once, thus it responds to changes in signals as they occur.
>
> NON-REAL-TIME (SCANNING) ANALYZER: This analyzer is only tuned to a single frequency at a given instant in time; so to analyze several signals, it sequentially scans through them one at a time. Because it must wait to tune to a particular frequency, it is not a real-time analyzer. The phenomenon under test must be repetitive, otherwise, it may not be detected. Scanning analyzers are usually the superheterodyne receiver type in which the first local oscillators (LO) or some IF is swept. HP analyzers are generally superheterodyne swept first LO, and all the terms defined below apply to our non-real-time spectrum analyzers.

FREQUENCY TERMS

FREQUENCY RANGE: The range of frequencies to which the analyzer can be tuned.

FREQUENCY ACCURACY: This refers to the accuracy of the analyzer tuning. It is specified as ± X MHz or parts per million of the dial scale or digital panel meter reading.

FREQUENCY RESOLUTION: Resolution is the ability of the analyzer to resolve two real signals present at its input. Closely spaced signals are more difficult to resolve than widely spaced signals. Several factors affect resolution, the most important of these is the final IF bandwidth and the shape factor.

RESOLUTION BANDWIDTH: This is the 3 dB bandwidth of the analyzer final IF stage. It is called the resolution bandwidth because two closely spaced, equal amplitude signals are just resolved if they are separated by an amount equal to the analyzer 3 dB bandwidth. For example, a final IF bandwidth of 100 Hz just resolves two equal amplitude signals 100 Hz apart. However, if the two signals are less than 100 Hz apart, they will be within the IF bandwidth at the same time and, therefore, will appear as one signal.

OPTIMUM RESOLUTION BANDWIDTH: This is the bandwidth which provides the most convenient display commensurate with scan width and reasonable sweep time.

BANDWIDTH SHAPE FACTOR: Shape factor is a measure of how wide the final IF filter is below the 3 dB points. This factor shows the analyzer's capability to resolve closely spaced signals of unequal amplitudes.

Shape factor is the ratio of

$$\frac{\text{Bandwidth 60 dB down (in Hz)}}{\text{Bandwidth at 3 dB (in Hz)}}$$

The smaller this ratio the sharper is the filter and the more selective it is.

SPAN, SCAN: They are identical, interchangeable terms, they refer to the calibration of the horizontal axis of the analyzer. The horizontal axis has 10 divisions and units are in Hz, KHz or MHz.

SCAN WIDTH OR FREQUENCY SPAN: This refers to the portion of the frequency range which is displayed on the CRT. There are three Scan (Frequency Span) modes; Full, per division, and Zero. Per division is subdivided into three types; center, start and 0-10F.

FULL SCAN: In this mode, the horizontal axis of the CRT display corresponds to the full frequency range of the analyzer. The left edge graticule line corresponds to the analyzer lower end frequency while the right edge graticule line corresponds to the upper end frequency. Internally, the analyzer first LO is sweeping across its full frequency range.

PER DIVISION SCAN

A. CENTER SCAN: In this mode, the horizontal axis is symmetrical about the center frequency which is set by the analyzer tuning control. The total scan width is equal to 10 x scan width/division setting. Internally, the analyzer first LO sweeps symmetrically about the frequency selected by the tuning control.

B. START SCAN: In this mode, the horizontal axis Start frequency (instead of the Center frequency) is set by the tuning control. The total frequency span displayed is still 10 x frequency span/division setting. Internally, this mode is similar to a Start/Stop sweep mode in Sweep oscillators, i.e., the analyzer first LO sweeps a range of frequencies starting at the frequency set by the tuning control and stopping at the frequency set by the frequency span control.

C. 0-10F SCAN: In this mode, the start frequency is always 0, the end frequency is equal to 10F where F = the scan width/division setting. Internally, the analyzer LO sweeps a range of frequencies, the low end frequency is preset while the high end frequency is set by scan width control.

ZERO SCAN: In this mode, the horizontal axis is calibrated in real time at a single frequency. The analyzer is CW tuned in this mode, (i.e., 0 Hz/division) thus time variations of the signal amplitude are displayed on the CRT and the analyzer becomes a tunable, variable bandwidth receiver.

SCAN (FREQUENCY SPAN) LINEARITY: This refers to the accuracy of the horizontal axis of the display, i.e., the frequency error between two points on the CRT.

FREQUENCY STABILITY: This refers to the stability of the analyzer as a measurement system. Since the analyzer must be more stable than the signals it measures, frequency stability considerations become particularly important. Usually, the analyzer frequency stability is equal to the sum of the frequency stabilities of its local oscillators. Two measures of frequency stability are used; short term and long term. Residual FM and noise sidebands are used to measure short term stability while drift is used to measure long term stability.

RESIDUAL FM: Residual FM is a short term measure of the analyzer jitter or aberrations. It is usually measured in some bandwidth and is specified as so many Hz RMS, peak, peak-to-peak or average in one second interval. The analyzer's narrowest bandwidth and scan width are limited by the residual FM.

NOISE SIDEBANDS: Noise sidebands are a measure of the analyzer spectral purity, i.e., how spectrally clean the analyzer local oscillators are. This specification is significant because, for spectral purity measurements, the analyzer must be more spectrally pure than the signals it measures, and noise sidebands can limit resolution. Noise sidebands are specified as so many dB down from the carrier in a specific bandwidth. For example, the 8553B, .001 - 110 MHz spectrum analyzer, noise sidebands are more than 70 dB down, 50 kHz away from the carrier in a 1 kHz bandwidth.

STABILIZATION: To reduce the analyzer residual FM in narrow scans, one of its sweeping local oscillators (normally the first LO) is synchronized to a stable crystal oscillator. Two techniques are used, automatic phase control (APC, i.e., phase lock) and automatic frequency control (AFC, i.e., frequency lock). Fundamentally, these techniques are simple. In phase lock, the analyzer LO is mixed with a crystal signal (or harmonic), the output is fed to a phase discriminator and the output of the discriminator (the error signal) is summed with the LO tuning voltage. It should be noted that the output of the phase discriminator is a function of the phase difference between the two signals, this explains the name phase lock. Frequency lock is similar except that the output of the discriminator (correction signal) is a function of the frequency difference between the LO and crystal harmonic, therefore the smaller the difference the smaller the correction.

In a well designed stabilization technique, the signal remains on-screen when switching from unstabilized to stabilized operation. This is a feature of all HP stabilized analyzers.

HARMONIC MIXING: Harmonic mixing is a technique used to extend the frequency range of the spectrum analyzer. A response is displayed on the CRT when

$$F_s - F_{LO} = F_{IF} \text{ and } F_{LO} - F_s = F_{IF}$$

where F_s — frequency of an unknown signal

F_{LO} — frequency of the analyzer local oscillator

F_{IF} — first IF frequency of the analyzer

or, $F_s - F_{LO} = \pm F_{IF}$ so that $F_s = F_{LO} \pm F_{IF}$

A more general form of this formula is

$$F_s = nF_{LO} \pm F_{IF}$$

where n is the harmonic number of the analyzer local oscillator. Thus, by mixing with the harmonics of the local oscillator, we can extend the analyzer frequency range to higher and higher frequencies.

IMAGE RESPONSES: Image signals are a pair of frequencies separated by twice the IF frequency of the analyzer. When image signals are applied to the analyzer and no filtering is done in the input circuit the two image signals mix with the same LO frequency (or harmonic) and appear as one response on the CRT, hence the name "image responses." (They are mirror images around the LO frequency.) For example, with a 2 GHz IF frequency, 1 GHz and 5 GHz are image frequencies and so are 10 GHz and 14 GHz, i.e., a pair of image responses exists for the LO fundamental frequency and each of its harmonics.

In the harmonic mixing analyzer, no filtering is done at the input so image responses may occur and produce a misleading display. Thus, there must be a way to eliminate image responses. The tracking preselector (see below) is a device that eliminates image and many other unwanted products.

MULTIPLE RESPONSES: When a signal is connected to the analyzer, one expects to see a single response on the CRT. In the harmonic mixing analyzer, one signal may produce several responses by mixing with more than one local oscillator frequency/harmonic in different mixing modes. For example a 5 GHz signal produces three responses as follows:

$n = +1$ it mixes with the GHz LO fundamental
($5 - 3 \rightarrow 2$ GHz 1st IF)

n = −2 it mixes with the 2nd harmonic of 3.5 GHz
(2 x 3.5) − 5 → 2 GHz 1st IF

n = −3 it mixes with the 3rd harmonic of 2.33 GHz
(3 x 2.33) − 5 → 2 GHz 1st IF

A tracking preselector also eliminates multiple responses.

IF FEEDTHROUGH: In the harmonic mixing analyzer a signal equal in frequency to the 1st IF will lift the entire baseline regardless of the frequency control settings. This phenomenon is known as IF feedthrough. Lifting the baseline, of course, obscures the display and may completely submerge other signals present. Thus to eliminate any gap in the analyzer frequency coverage, an *alternate IF must be available* and the two IFs must not be harmonically related. A preselector also eliminates IF feedthrough.

PRESELECTOR: The preselector is a voltage tunable bandpass (typically 20-50 MHz) filter which tracks the analyzer tuning. It is used with the harmonic mixing analyzer to eliminate image, multiple and spurious responses.

The preselector generally uses an yittrium-iron-garnet (YIG) material. The YIG tuning properties extend typically from 1.8 - 18 GHz.

AMPLITUDE TERMS

DISPLAY RANGE: This refers to the difference (in dB) between the top line of the CRT and the lowest utilizable (calibrated) line. This range is limited by the screen size and the log amplifier.

DYNAMIC RANGE: Is the ratio of the largest signal to the smallest signal that can be displayed *simultaneously* with no analyzer distortion products. There is an optimum signal level associated with dynamic range, reducing the signal level at the input mixer by 10 dB increases dynamic range by at least 10 dB, assuming there is sufficient sensitivity. This definition of dynamic range means distortion-free display range. All signals displayed on the CRT are real signals and not analyzer distortion or spurious mixing products. This definition is not the same as the ratio of the largest signal that can be applied without serious distortion (usually means 1 dB amplitude compression) to the smallest signal that can be detected. Likewise, it is not the ratio of the largest to the smallest signal that can be measured by the analyzer.

SENSITIVITY: Sensitivity is a measure of the analyzer's capability to detect small signals. Like most receivers, the analyzer's maximum sensitivity (minimum discernible signal) is limited by the inherent average noise. Thus, an unknown signal can be detected when the signal power = inherent average noise power, or

$$\frac{\text{Signal power} + \text{inherent average noise power}}{\text{inherent average noise power}} = 2$$

When viewed on the CRT, this signal appears 3 dB above the inherent average noise level.

Referred to the analyzer input, the inherent average noise level has two components; thermal noise and noise resulting from the active elements. Thus,

$$P_{in} = 10 \log KTB + N_o$$

where:

P_{in} = inherent average noise in dBm
K = Boltzmann's constant—joules
T = Absolute temperature − °K
B = Equivalent IF bandwidth − Hz
N_o = The analyzer noise figure − the contribution of the active elements − dB

From this definition, we can see that the inherent average noise is bandwidth dependent, e.g., −110 dBm/10 KHz. Because of the power relationship above, a decade decrease in bandwidth results in 10 dB lower noise level and consequently 10 dB better sensitivity. When measuring inherent average noise level, a video filter with a bandwidth much less than the IF bandwidth should be inserted in the output circuit and the analyzer input should be terminated with its characteristic impedance.

OPTIMUM INPUT LEVEL: This refers to the signal level into the input mixer which ensures that the analyzer generated distortion products are below a certain level. For example, a —40 dBm optimum level into the mixer guarantees distortion products to be below −110 dBm, i.e. 70 dB down. The 70 dB range is the dynamic range defined above. Different mixers have different optimum levels, however, most HP spectrum analyzers specify —40 dBm to be the optimum input level.

LINEAR INPUT LEVEL: This is also a characteristic of the input mixer and it refers to an upper signal level for which amplitude compression is less than 1 dB. Above this level, the mixer goes into serious gain compression (1 dB compression equals a 12.2% error) and becomes non-linear, i.e., the mixer output is no longer proportional to the input signal level. Linear operation means amplitude compression is <1 dB, it does not mean the display is distortion-free. Distortion will be generated for input levels in the mixer above the optimum value.

MAXIMUM INPUT LEVEL: This is the damage level of the analyzer front-end. Two figures are usually given; one is the burn-out level of the mixer and the other is the damage level of the input attenuator. For example, −13 dBm may be the mixer burn-out level and +30 dBm the damage level for input attenuator.

AMPLITUDE ACCURACY: This refers to the error of the amplitude calibration of the analyzer. This error consists of several components; i.e., flatness, switching input attenuator, switching between bandwidths, IF gain of the log amplifier, amplitude display and the built-in calibrator. With IF substitution, this error can be reduced to the sum of the frequency response variations and the IF gain accuracy (log reference level control).

FLATNESS: Flatness is the term used to describe the frequency response variations of the spectrum analyzer. A ± 1 dB flatness means the peaks and the troughs of the frequency response of the analyzer are 2 dB apart.

AMPLITUDE DISPLAY ACCURACY: This is the accuracy of relative amplitude measurements made from the CRT scale.

LOG DISPLAY: A display mode in which the CRT vertical deflection is a logarithmic function of the input signal voltage. In this mode, the CRT is calibrated in dBm in most HP analyzers. dBμV and dBmV calibrations are also used.

LINEAR DISPLAY: A display mode in which the CRT vertical deflection is directly proportional to the RMS input signal voltage. In this mode, the CRT is calibrated in volts in most HP analyzers. In the HP economy analyzers, however, the linear mode has no scale calibration but the reference level is calibrated in dBm.

LOG REFERENCE LEVEL OR REFERENCE LEVEL: This refers to the absolute calibration of the top graticule line of the CRT. All HP analyzers are absolutely calibrated in amplitude, Log Reference Level is generally calibrated in dBm (0 dBm = 1 mW). Other calibrations such as volts, dBμV, dBmV are also used.

SPURIOUS RESPONSES: These are undesired responses that are generated in the analyzer and displayed on the CRT. There are two types of Spurious Responses; Harmonic and non-harmonic. Harmonic spurious responses are second, third, fourth, etc., harmonics of the input signal. Non-harmonic spurious responses are intermodulation and residual responses.

HARMONIC DISTORTION. When the input signal is of sufficient amplitude to drive the mixer into non-linear operation, *false* harmonics of the input signals may be generated in the mixer. These false harmonics are called harmonic distortion. Typically, second harmonic distortion is the largest component.

INTERMODULATION (IM) DISTORTION: When two strong and closely-spaced signals are fed into the analyzer input mixer, they drive the mixer into non-linear operation. This causes the two signals to interact and produce several mixing products. For example, if f_1 and f_2 are fed into the input mixer, then $2f_1 \pm f_2$, $2f_2 \pm f_1$, $2\ (f_1 \pm f_2)$ and $3\ (f_1 \pm f_2)$ etc. are some of the intermodulation distortion that may be seen on the CRT. $2\ f_1 - f_2$ and $2\ f_2 - f_1$ are called 3rd order IM products and are the most harmful distortion products. They are the most harmful because they are located close to f_1 and f_2; the rest of the IM products are farther removed from the spectrum of interest.

RESIDUAL RESPONSES: Responses which appear on the CRT with no input signal connected to the analyzer are called residual responses. These ever-present responses are distortion products which result when the local oscillator's fundamental or higher harmonics mix together to produce signals at one of the IF frequencies of the analyzer.

LO FEEDTHROUGH: In the scanning analyzer when the frequency of the first local oscillator is equal to the IF frequency, a response appears on the CRT. This response is known as the LO feedthrough. All HP analyzers are designed so that the LO feedthrough represents a zero frequency reference to calibrate the horizontal axis.

LO RADIATION: This is the energy detected at the front end of the analyzer. The first local oscillator feeds this energy out through the input mixer. In the absence of any input filters, the energy comes out the input connector and goes into anything connected to the analyzer.

MISCELLANEOUS

TRACKING GENERATOR: The tracking generator is a companion instrument to the spectrum analyzer; it (the tracking generator) produces a signal whose frequency precisely tracks the spectrum analyzer tuning. In addition to the standard analyzer applications, the spectrum analyzer/tracking generator is used to make two types of measurements; frequency response of unknown devices and selective/sensitive counting of unknown signal frequencies.

TRACKING ERROR: The finite difference between the tracking generator output frequency and the analyzer tuning is called tracking error. This error affects the frequency accuracy in the tracking generator measurements and, therefore, it should be made as small as possible.

SCAN TIME: The analyzer scan time is the total time it takes to make one full sweep. It is equal to 10 x scan time/division setting.

VIDEO SCAN TRIGGER: This is a triggering mode where the analyzer sweep is synchronized to the envelope of the RF input signal. It is used to recover modulation in the zero scan mode. In HP economy analyzers, this mode is called INT.

VIDEO FILTER (or DISPLAY SMOOTHING): The video filter is a post-detection averaging device which is used to average the noise present in the analyzer. Noise averaging filters are low pass filters whose bandwidth should be much narrower than the resolution bandwidth of the analyzer. However, since the video filter is located after the IF bandwidth, it does not affect frequency resolution.

In some measurements, e.g. AM, FM and pulsed RF demodulation, the video filter bandwidth should be equal to or greater than the IF bandwidth so as to pass the demodulated waveform without distortion. Thus, measurement flexibility requires the analyzer to have wideband as well as narrowband video filters and selection of video bandwidth should be accomplished easily and quickly.

IMPEDANCE: This is the input impedance of the analyzer. 50 Ω is the most commonly used, although 75 Ω is also widely used. 1 M Ω is used in the audio frequency range.

HEWLETT hp PACKARD

5952-1213

PRINTED IN U.S.A.