ROHDE & SCHWARZ

Fundamentals of Spectrum Analysis using Spectrum Analyzer FSA

Info 001 138
Fundamentals FSA
E-1

Fundamentals of Spectrum Analysis using Spectrum Analyzer FSA

With the Spectrum Analyzer FSA, Rohde & Schwarz has set a milestone in the field of spectrum analysis [1]. The FSA is a versatile, easy-to-operate and convenient measuring instrument for a broad range of applications in the communications lab.

This application note is intended to aid the reader in getting acquainted with the methods of spectrum analysis. To understand the functions of an instrument as complex as the FSA, it is essential that an introduction be given into the theoretical background of spectrum analysis. It can also be useful for the experienced user of spectrum analyzers to recall related theoretical knowledge that will help him to eliminate routine measurement errors [2; 3; 4].

For this reason, this application note covers the fundamentals of spectrum analysis. Further brochures dealing with the solution of special measurement problems are in preparation.

CONTENTS

Analysis in time and frequency domain

A signal can be analyzed employing mathematical or measurement methods. Signals can be periodical waveforms or transients, and may also follow statistical laws. The CRT of an oscilloscope displays the amplitude of a signal as a function of time whereas a spectrum analyzer represents the same signal in the frequency domain. The two types of representations are totally different (Fig. 1).

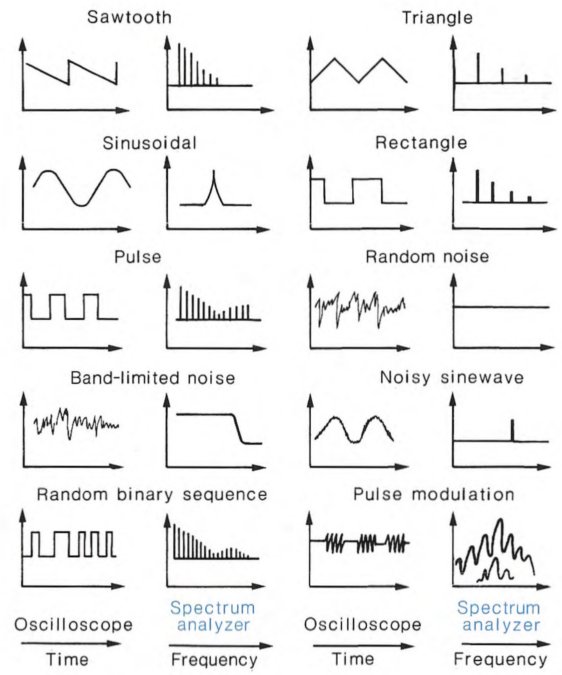


Fig. 1 Signal analysis with oscilloscope and spectrum analyzer

Time and frequency can be correlated mathematically by performing a Fourier transformation. Every signal varying with time has a characteristic spectrum in the frequency domain.

Fourier's theorem is as follows:

$$f(t) = 0.5 A_0 + \sum_{n=1}^{\infty} A_n \sin(n\omega_0 t) + \sum_{n=1}^{\infty} B_n \cos(n\omega_0 t) \qquad (1)$$

The values of Fourier coefficients A_n and B_n depend on the signal's waveform f (t) and can be calculated from:

$$A_0 = 2/T_0 \int_0^{T_0} f(t)\,dt \qquad (2)$$

$$A_n = 2/T_0 \int_0^{T_0} f(t) \cos(n\omega_0 t)\,dt \qquad (3)$$

$$B_n = 2/T_0 \int_0^{T_0} f(t) \sin(n\omega_0 t)\,dt \qquad (4)$$

where:

$A_0/2$	DC component
f(t)	time function
n	order of harmonic oscillation
T_0	period
$\omega_0 = 2\pi/T_0$	circular frequency

Normally, oscilloscopes operate with linear amplitude scaling. The variations of a signal as a function of time should be displayed with minimum distortion. For complex signals, this type of representation becomes very difficult and leads to substantial errors. Periodic signals with frequency components whose amplitudes are much below the amplitude of the fundamental (eg slightly modulated signals or harmonics) cannot be completely evaluated in most cases.

Modulation depth and harmonics contents can only be determined by cumbersome mathematical conversions. With spectrum analysis, complex signals can be evaluated much faster and easier in the frequency domain. By assigning the correct amplitude value to each spectral line, the spectrum analyzer provides significantly more detailed information on spectral distribution, harmonics contents, bandwidth and absolute signal level than an oscilloscope. With respect to frequency range, resolution, accuracy and dynamic range, the spectrum analyzer is far superior to an oscilloscope.

Fig. 2 shows the superposition of a periodic signal with sinusoidal components, resulting in a distorted waveform. It cannot be seen from the waveform whether or not the 4th or 5th harmonic contribute to the signal amplitude. In the amplitude/frequency domain, the signal is converted to its frequency spectrum which is then displayed.

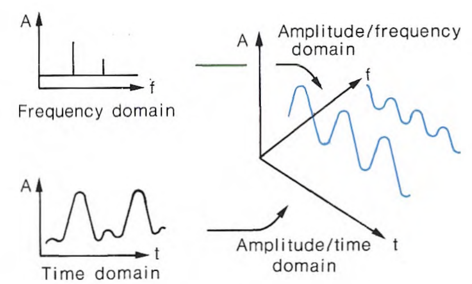


Fig. 2 Display of signal in time and frequency domains

The transformation of the sine signal shown in Fig. 3 into its frequency spectrum permits determination of its harmonics (Fig. 4) which is not possible with representation in the time domain.

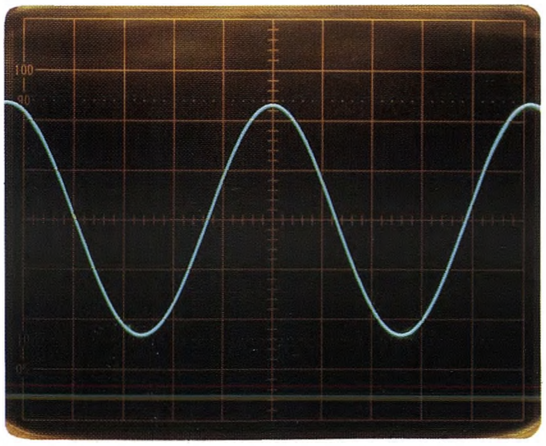


Fig. 3 Sine signal in time domain (frequency 1 MHz; level −20 dBm; time deflection 0.2 μs/div; amplitude setting 10 mV/div)

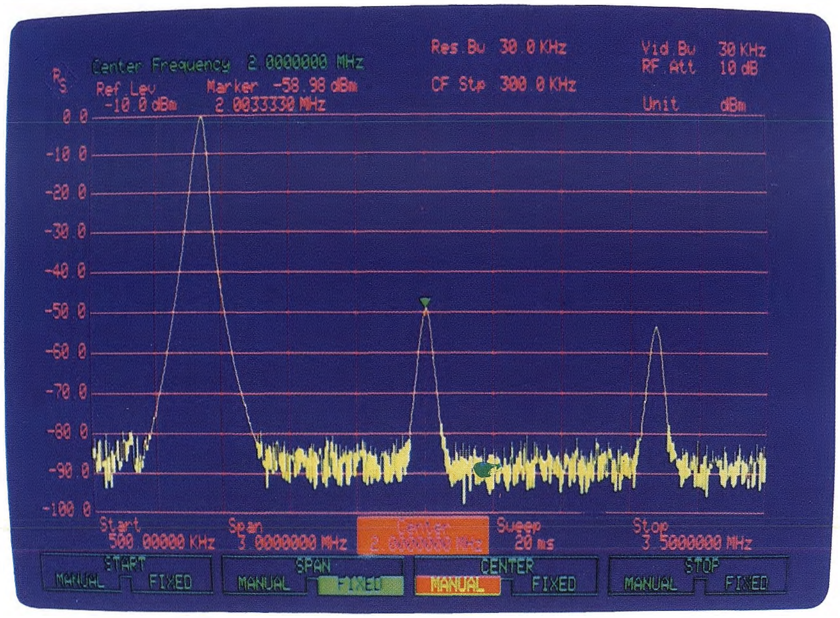


Fig. 4 Sine signal of Fig. 3 in frequency domain, displayed on Spectrum Analyzer FSA

Types of spectrum analyzers

Spectrum analyzers should be capable of analyzing both low-frequency and microwave signals. This results in the requirement for a very large frequency coverage that can be divided into three ranges:

- ▷ **AF range** up to approx. 1 MHz,
- ▷ **RF range** up to approx. 2 GHz,
- ▷ **Microwave range** 2 GHz to 40 GHz (or above)

The first range up to approx. 1 MHz includes signals encountered in low-frequency electronics, acoustics and mechanics. The second range from 100 Hz to 2 GHz is assigned to high-frequency and the third range between 2 and 40 GHz to microwave engineering. Depending on the frequency range and the individual application, measurements require different characteristics. Therefore, this section will briefly discuss spectrum analysis techniques already implemented.

FFT (Fast Fourier Transformation) analysis

Analyzers employing the discrete Fourier transformation method use highspeed ADCs (analog-to-digital converters) featuring high resolution in conjunction with high-speed microprocessor and memory circuits. The signal is sampled in the time domain, digitized and processed to determine its magnitude and phase. The accuracy of the results depends on the quality of the analog input stage and of the ADC as well.

FFT spectrum analyzers (Fig. 5) operate in real time and provide rapid results. With semiconductor devices available today, processing speeds can be realized which allow discrete FFT analysis of signals with frequencies up to 1 MHz. Considering the continuous progress of semiconductor technology, extension of the frequency range towards higher signal frequencies are foreseeable in the future.

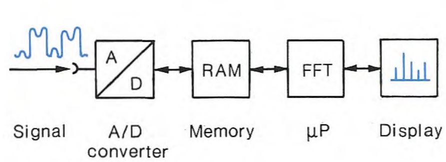


Fig. 5 Block diagram of typical FFT spectrum analyzer

Also non-periodic signals can be analyzed with FFT spectrum analyzers providing very good frequency and amplitude resolutions. The dynamic range is, however, limited to approx. 85 dB. FFT spectrum analyzers are used in acoustics to analyze switching noise such as impulse and crackling noise, mechanical resonances of loudspeaker enclosures as well as for measuring the mechanical stability of magnetic tape units and record players, where mechanical instabilities can lead to deterioration of the sound reproduction quality.

Real-time analysis

Real-time spectrum analyzers consist of a number of frequency-selective indicators connected in parallel. The principle of parallel sampling is illustrated by Fig. 6. The resolution filter should approach the ideal rectangular transfer function (Fig. 7) to avoid crosstalk caused by filter overlapping as well as multiple response or spectrum gaps. However, in practical applications this idealized transfer function cannot be implemented.

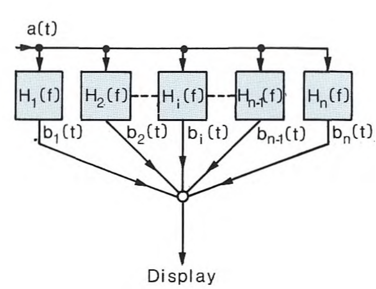


Fig. 6 Principle of parallel sampling with n bandpass filters whose output signals are displayed

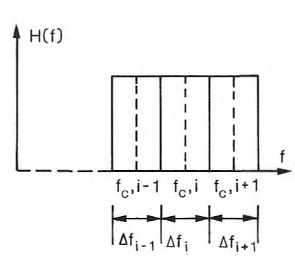


Fig. 7 Ideal (rectangular) filter characteristics for parallel sampling

The number of discrete spectral components that can be displayed is equal to the number of filters. The resolution of this analyzer type is determined by the filter bandwidths. Higher resolutions can be obtained by increasing the number of filters. The number of filters is, however, limited due to economical considerations.

The application range of real-time spectrum analyzers for measurements of signals up to 100 kHz is limited by the complexity of the equipment. As far as dynamic range and amplitude resolution are concerned, real-time analysis is superior to FFT analysis.

When bandpass filters with variable centre frequency f_c (Fig. 8) are used, the absolute filter bandwidth increases with increasing centre frequency (at constant relative bandwidth $\Delta f / f_c$). With spectrum analyzers, a constant absolute filter bandwidth (Δf) is required in most cases.

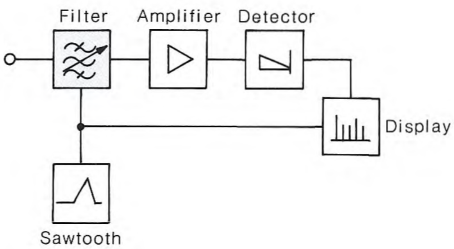


Fig. 8 Spectrum analyzer with tunable filter

Analysis with direct conversion

A tunable oscillator can be realized more easily for a given range than a filter with constant characteristics operating within the same frequency range. The tunable oscillator, as well as the signal to be analyzed, drives a mixer stage. The horizontal sweep of the display is synchronized with the tuning process (Fig. 9). All signal components are applied to the display section through a lowpass filter and an amplifier without being attenuated. This configuration has the disadvantage that each signal to be analyzed generates a beat-frequency marker having a width of 2 times f_{LP} (with f_{LP} being the 3-dB-cut-off frequency of the lowpass filter). A signal is thus being displayed with twice the bandwidth of the actual lowpass filter.

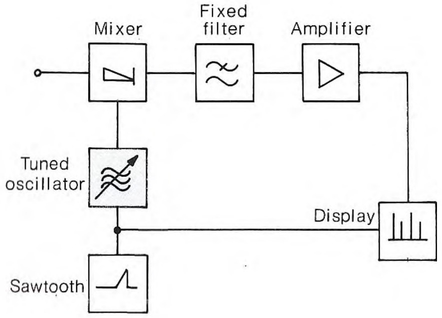


Fig. 9 Analysis with direct conversion

Superheterodyne analysis

Almost all modern spectrum analyzers, including FSA, employ the superheterodyne principle. The signal applied to the spectrum analyzer input is mixed with the output signal of a tunable oscillator having a frequency of f_{LO} (t). The bandpass filter features a constant centre frequency and a constant, switch-selectable bandwidth (Δf). This filter provides a signal path for the frequency component of the time function f_{in} (t) to be analyzed, for which the following relationship applies:

$$f_{in}(t) = f_{LO}(t) \pm f_{IF} \qquad (5)$$

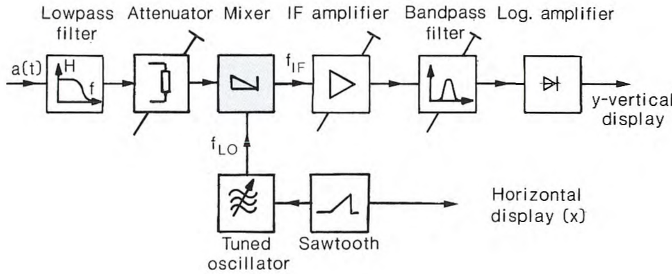


Fig. 10 Simplified block diagram of spectrum analyzer with sequential sampling (superheterodyne principle)

By conversion into a fixed intermediate frequency (f_{IF}), the disadvantages of the direct conversion method can be overcome. A sawtooth generator tunes the oscillator and drives the horizontal deflection circuitry of the display. Also, with this method, the display is ambiguous because two different frequencies are both converted into the IF signal during the mixing process. The resolution is determined by the characteristic of the IF filter and does not depend on the frequency of the input signal. The IF filter can be designed to achieve the required selectivity.

In contrast to spectrum analyzers with tunable input filters, the incoming signal is tuned to the IF analyzer filter by conversion. Since superheterodyne spectrum analyzers should operate up to frequencies of 2 GHz, several IF stages are required. With intermediate frequencies in the range from 3 to 25 MHz, filter bandwidths from 10 Hz to 3 MHz are easily obtained. The bandwidth of the FSA can be tuned between 6 Hz and 3 MHz in steps of 1 – 3 – 10 as well as quasi-continuously.

Design and operation of a spectrum analyzer (superheterodyne principle)

Fig. 11 shows the block diagram of a superheterodyne spectrum analyzer. Additional conversion stages are not shown since they are not essential to explain the spectrum analyzer's principle of operation. By special filter design, a better image frequency rejection (see up/down conversion, image frequencies) can be obtained, direct reception of the IF signal can be avoided, and in the other direction, re-radiation of the oscillator signal can be prevented at the input of the spectrum analyzer.

The signal to be analyzed is applied to the input filter (2) via the input attenuator (1) which is switchable in 10-dB steps in most cases. This filter performs several functions:

▷ it prevents multiple reception of a signal (image reception),

▷ it prevents direct reception of the IF signal (IF breakthrough),

▷ it prevents the oscillator signal from being fed back to the input of the spectrum analyzer.

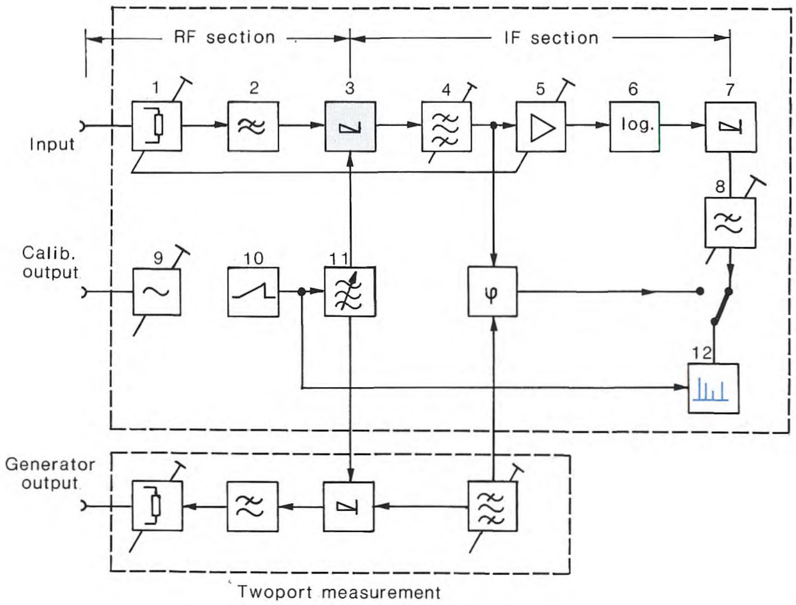


Fig. 11 Block diagram of spectrum analyzer (numbers explained in text)

The input mixer (3) converting the input signals in conjunction with the tunable oscillator (11) determines the frequency-dependent amplitude error and the dynamic characteristics of the spectrum analyzer. The oscillator (11) should supply a stable, high-purity signal of exactly known frequency. It is tuned by a sawtooth generator (10) whose output amplitude determines the frequency span of the analysis and whose frequency determines the sweep time. The (RF) calibration source (9), with an exactly defined output level, allows accurate level calibration of the spectrum analyzer.

The IF filter (4) determines the characteristics of the analysis. A switchable amplifier (5) and the input attenuator (1) permit variation of the overall gain. The level range that can be displayed is determined by the logarithmic amplifier (6). Subsequent to the IF detector (7), noise can be averaged by a switchable video filter (8). The level versus frequency is indicated by the display (12).

With a minimum of effort, the spectrum analyzer can be expanded to operate as a two-port test set with tracking generator (scalar measurements). If this generator is available, the phase shift of a device under test, connected between the generator output and the spectrum analyzer input, can also be measured. The results of phase measurements can be used to calculate the impedance or group delay values (complex measurements).

RF input section

The RF input section consists of the input attenuator, the input filter, the mixer, and the superheterodyne oscillator. The usable frequency range, the absolute sensitivity, and the unambiguity of the display depend on the design concept of the RF input section. The upper frequency limit of the usable range is determined by the frequency response of the mixer, the sweep range of the oscillator, and by the cut-off frequency of the input lowpass filter. The internal frequency limit of 0 Hz cannot be displayed since the oscillator frequency would then be identical to the intermediate frequency, resulting in a representation of the IF filter transfer function (zero response). Another aspect for consideration of the lower frequency limit is the highpass filter at the input of the spectrum analyzer. The capacitor at the input prevents DC voltage from being applied to the mixer. The input impedance is 50 Ω. The impedance of 75 Ω commonly used in television systems can be transformed to 50 Ω by using matching pads resulting in minimum loss of sensitivity.

An ambiguous display may be caused by non-linearities of the mixer or by the mixing process. Due to limited linearity of the mixer, harmonics of the input signal are generated, and combination frequencies are generated for multiple input signals. When the distortion products (see non-linearities) are just indiscernible because of the presence of intrinsic noise, the optimum dynamic range is obtained, which can also be varied by choosing a maximum permissible input level. The maximum level is determined by rated power dissipation and by the attenuation provided by the input attenuator. For sinusoidal signals, continuous power is essential whereas the time-voltage product is of prime concern for pulse spectra. With pulse signals and large RF bandwidths, the mixer may unnoticeably be driven into saturation since the entire energy contained in the RF bandwidth is used to drive the mixer (Fig. 12).

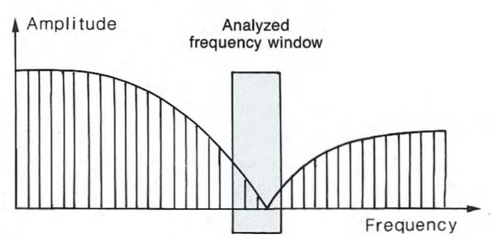


Fig. 12 Risk of overloading mixer with broadband signals

DESIGN AND OPERATION

When the passband of the analyzer filter coincides with a null in the spectrum, the sensitivity of the spectrum analyzer should be increased by reducing the input attenuation. This will, however, increase the risk of overloading the mixer.

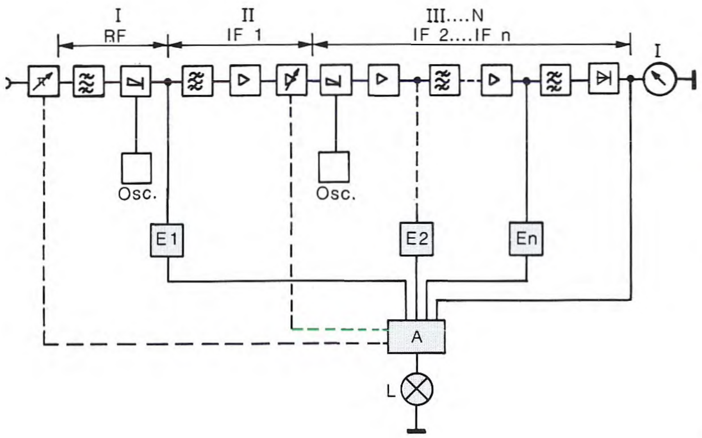


Fig. 13 Overload circuitry

Mixer overloading can be indicated with the aid of an overload detection circuitry. An RF receiver made up of several amplifier stages with limited dynamic range and with bandwidth decreasing from stage to stage requires overload indication if no automatic gain control is provided (Fig. 13). An indicator (L) can be driven via a common evaluation circuitry (A) by a number of overload detectors (E) at points of the analyzer circuitry where the bandwidth of subsequent amplifier stages is varied. The evaluation circuitry performs logical combinations of the outputs of the overload detectors to indicate an overload condition whenever any of the amplifier stages is driven into saturation. Fast response of the overload detectors is of great importance. The response time of each overload detector should be less than the time corresponding to the pulse bandwidth of the previous amplifier stage (6-dB-bandwidth corresponding to pulse bandwidth of a filter).

IF section

The IF section consists of the mixer, the IF filter, the IF preamplifier, a logarithmic amplifier, and a detector. The IF section amplifies the input signal to obtain the amplitude required for linear detection.

The IF gain is varied in steps, depending on the amplitude of the input signal. By using a logarithmic amplifier, an increased display range is obtained in contrast to linear display. When pulse or noisy signals are to be analyzed, correction factors have to be applied to the IF resolution filter, the logarithmic amplifier and the detector, as the characteristics of these components are subject to changes for non-sinusoidal signals. Both the logarithmic amplifier and the detector can exhibit non-linearities that have to be compensated. By substituting and calibrating the amplifier path with a calibration source that supplies an exactly defined signal level, amplitude errors can be reduced significantly. Frequency-independent errors of the logarithmic amplifier and the detector are taken into account during the calibration and do therefore not affect the measurements.

For low-noise measurements, the RF attenuation at the analyzer input should be reduced and the IF attenuation increased, since harmonics of interest would otherwise be buried in noise if high RF attenuation were selected. When low-distortion measurements are to be performed, the RF attenuation should be increased and the IF attenuation reduced to avoid driving the mixer into saturation.

Local oscillator

To convert the frequency of the input signal into frequencies within the bandwidth of the IF filter, the local oscillator is swept between limits f_{LO1} and f_{LO2} (Fig. 14).

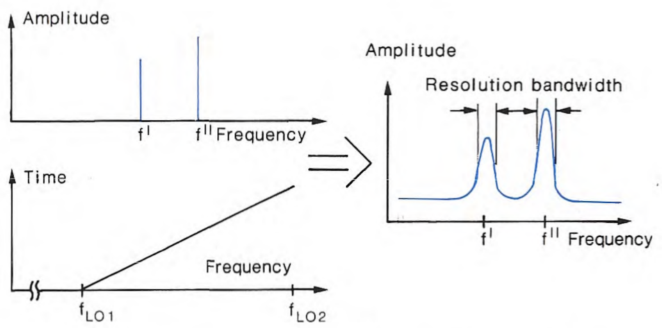


Fig. 14 Oscillator frequency and signal components versus time

With state-of-the-art spectrum analyzers, a frequency synthesizer is used involving phase-locking of the local oscillator to an internal reference oscillator. With an exactly defined intermediate frequency and the mathematical correlations applying to the synthesizer, a microprocessor can determine the frequency of the input signal. The Spectrum Analyzer FSA provides this spectral display with synthesizer accuracy. The frequency conversion shall now be explained by means of the example shown in Fig. 14. Spectral components f' and f'' are converted into the fixed intermediate frequency at a given time by mixing with the oscillator frequency varied with time, and will then be displayed at the corresponding positions within the spectrum. The oscillator output signal must have high spectral purity (see also phase noise) as variations of f_{LO} will result in variations of the intermediate frequency.

Long-term drift of an oscillator (i.e. frequency variations occurring over time periods between several seconds and days) is caused by aging effects that result in changes of the tuning voltage applied to free-running oscillators. With a phase-locked oscillator, these very slight drifts can only occur due to variations of the reference frequency to which the oscillator is synchronized. A high-quality 10-MHz crystal-stabilized oscillator will exhibit a temperature drift of 2×10^{-9} Hz/°C and a long-term drift of 5×10^{-10} over 24 hours.

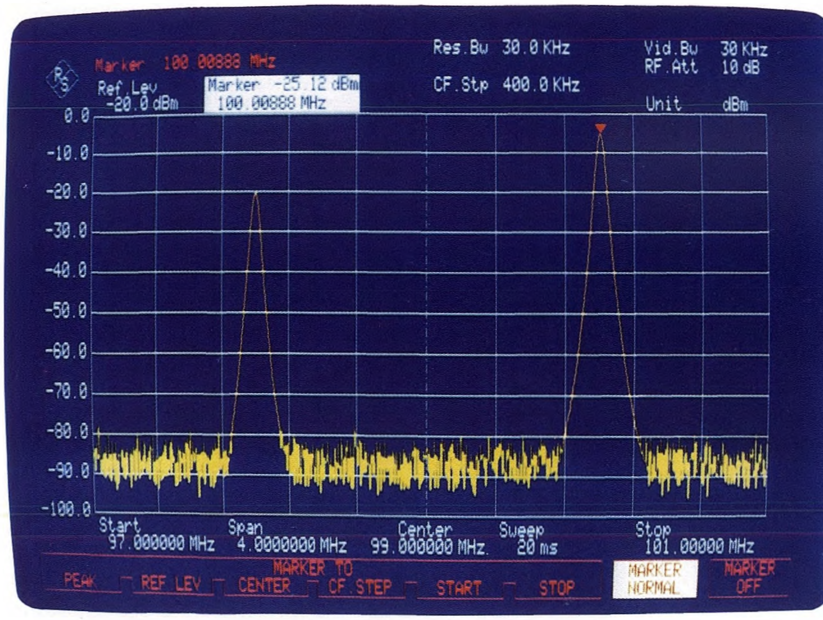


Fig. 15 Frequency conversion of Spectrum Analyzer FSA

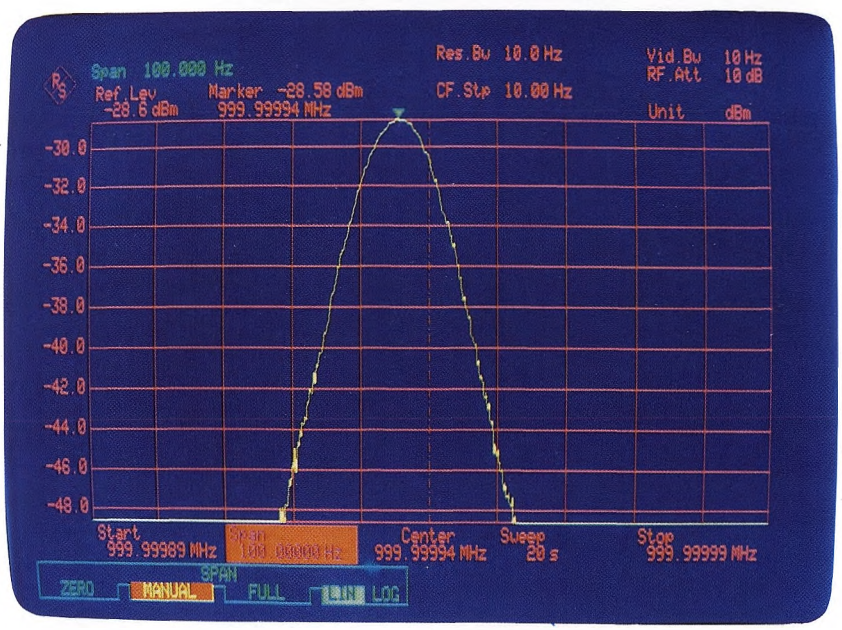


Fig. 16 Residual FM of sine signal

Also short-term drift (i.e. frequency variations occurring over time periods between several milliseconds and seconds) may be encountered with free-running oscillators. The drift is called spurious FM or residual FM, and is most often caused by noise effects in the oscillator circuitry, by noise induced on the tuning-voltage line, or by temperature effects (Fig. 16). Residual FM:

$$\Delta f_{FM} \sim \frac{1}{Q_B} \left(\frac{f_{LO2} - f_{LO1}}{U_2 - U_1} \right) \quad (6)$$

where:

Q_B = Q factor of oscillator circuitry
$f_{LO1,2}$ = upper/lower frequency limit of oscillator
$U_{1/2}$ = tuning voltage at upper/lower frequency limit

With phase-locked oscillators, the extremely high Q factor of the reference oscillator is also reflected in the local oscillator within the locking bandwidth of the phase-locked loop as ratio f_{REF}/f_{LO}, thus resulting in a significant reduction of short-term drift.

The minimum resolution bandwidth of a spectrum analyzer is determined by residual FM. The peak-to-peak value of residual FM should always be less than the minimum resolution bandwidth; otherwise no clearly defined display of the spectrum can be obtained.

Residual AM of the oscillator output signal does not affect the resolution bandwidth but is reflected in phase noise. This can be seen when measuring signals of very low levels when their frequencies are very close to the carrier frequency.

With spectrum analyzers including several superheterodyne conversion stages or local oscillators, the first oscillator is swept in most cases denoting these spectrum analyzers as “swept front-end spectrum analyzers”.

DESIGN AND OPERATION

Video filter

When analyzing sinusoidal signals, the video filter is used to reduce noise by averaging the signal applied to the display section. For sine-signal analysis, the video bandwidth should be selected to be about the same as the IF bandwidth, whereas it should be only between one third and one tenth of the IF bandwidth when analyzing noisy signals. Pulse signals should not be averaged due to their completely different characteristics. Pulse signals exhibit high peak amplitudes of brief duration and low average values (depending on their duty factor). To avoid measurement errors, the bandwidth should be 3 to 10 times the IF bandwidth in these cases (Fig. 17).

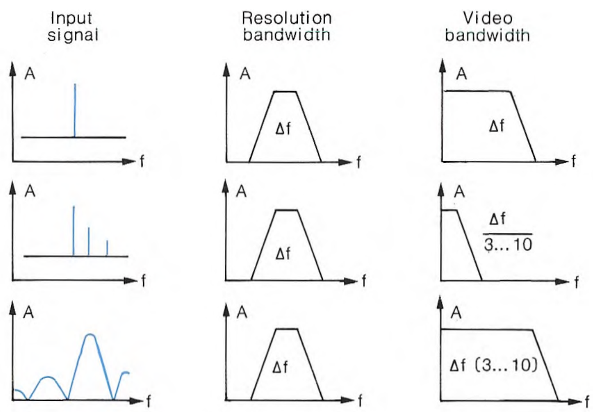


Fig. 17 Adjustment of video filter bandwidth

Display and display memory

The analog signal at the output of the video filter can directly be applied to the input of the vertical deflection circuitry of the spectrum analyzer. With advanced spectrum analyzers, the output signal from the video filter is digitized by an ADC. The digital data are then processed by µP firmware to allow processing by a graphics display controller and their visual representation on a monitor (Fig. 18).

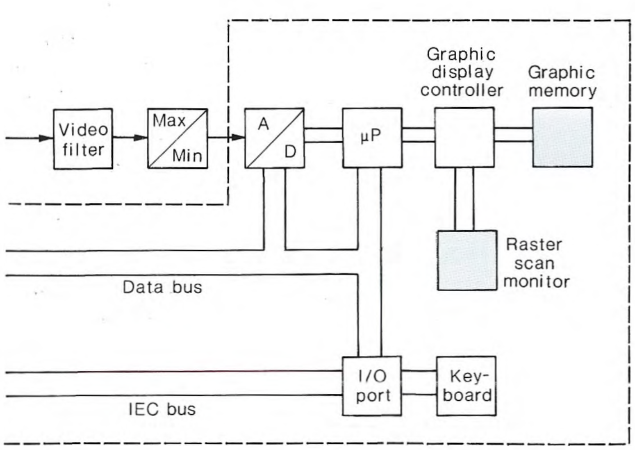


Fig. 18 Conversion of analog signal for display on monitor

By extensive use of digital techniques, the Spectrum Analyzer FSA is capable of providing sufficient resolution over a dynamic range between 0.1 and 110 dB. The measured values are displayed on a high-resolution colour monitor with 512 x 1024 pixels.

By using a digital display memory, the measurement capabilities of a spectrum analyzer are greatly expanded while simplifying the operation. After the first sweep, a complete spectrum is displayed that is updated at the end of each subsequent sweep period. This is an advantage for high-resolution measurements which require slow sweep displays. In addition, displays can be "frozen" and compared with other spectra allowing the representation of differences between current and stored signals (frequency deviation measurements on modulated signal sources and drift measurements). By digital averaging, the noise floor is reduced to a line which allows direct readout of the signal-to-noise ratio.

By displaying all important parameter settings on the CRT, the user is informed of reference level, centre frequency, span, resolution bandwidth, attenuator settings, etc., so that erroneous parameter settings can be corrected immediately. By connecting the FSA to a printer or plotter, a hardcopy of all parameter settings can be provided (Fig. 19).

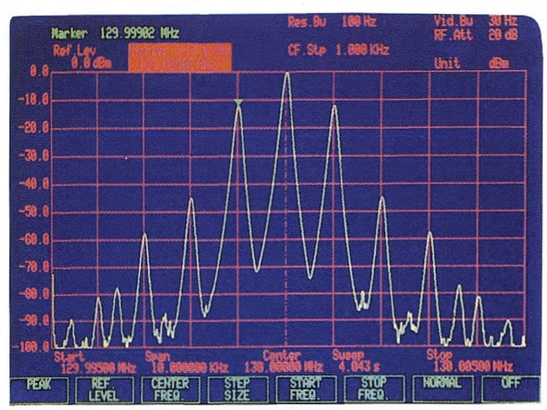


Fig. 19 Readout of all essential parameters on Spectrum Analyzer FSA

Mixing

Up/down conversion, image frequencies

The frequency of the input signal (f_{in}) is converted into a fixed intermediate frequency (f_{IF}) by mixing with frequency f_{LO} of the local oscillator. The intermediate frequency can be obtained by two distinct mixing processes:

first case: $f_{IF} = |f_{LO} - f_{in}|$ (7)

second case: $f_{IF} = f_{LO} + f_{in}$ (8)

The first case is called "down conversion" because the intermediate frequency is below the oscillator frequency. Consequently, the second case is called "up conversion". The term "mixing" should be understood as the generation and utilization of combination frequencies (f_{com}) that result from the non-linear characteristics of an electronic component which influence the voltages or currents of signals with frequencies f_1 and f_2 (in equation 9, m and n represent integer numbers):

$$f_{com} = |m f_1 \pm n f_2| \quad (9)$$

The ambiguity of the mixing process results in a clear disadvantage since a discrete value of the intermediate frequency (f_{IF}) can be obtained in two different ways from an input signal (f_{in}) and the oscillator frequency (f_{LO}). This shall be explained for the case of down conversion (Fig. 20):

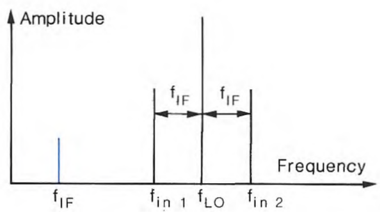


Fig. 20 Ambiguity of superheterodyne principle

$$f_{IF} = f_{LO} - f_{in1} \quad (10)$$

$$f_{IF} = f_{in2} - f_{LO} \quad (11)$$

For a given oscillator frequency (f_{LO}), the fixed intermediate frequency will result from both

$$f_{in1} = f_{LO} - f_{IF} \quad (12)$$

and

$$f_{in2} = f_{LO} + f_{IF} = f_{in1} + 2f_{IF} = f_{image} \quad (13)$$

f_{in2} is called the "image frequency" of f_{in1}. By considering the positions of f_{in1} and f_{in2} within the spectrum, this effect can easily be understood.

The attenuation required for the image frequency (f_{in2}), in respect to the "correct" reception frequency (f_{in1}), by a preselection circuit tuned to f_{in1} is called "signal-to-image-ratio". By providing sufficient RF preselection (i.e. attenuation of f_{in2}) that may, for instance, be implemented by connecting the lowpass filter of the RF input section **before** the mixer stage, an unambiguous display can be obtained.

For double superheterodyne conversion, the first intermediate frequency (f_{IF1}) should therefore be relatively high, providing a high signal-to-image ratio. The second IF filter (f_{IF2}) provides nearby selectivity. f_{IF2} should be relatively low to allow high IF gain and stability.

By mixing with harmonics of f_{LO}, further ambiguities may occur, which can also be eliminated with the aid of preselection.

Harmonics mixing

The Spectrum Analyzer FSA uses for mixing solely the fundamental of the oscillator. In the following, the mixing principle using harmonics is however described.

To expand the frequency range of a spectrum analyzer towards very high frequencies with a minimum of effort, the input signal is not only mixed with the fundamental of the local oscillator but also with its harmonics. Corresponding to the following equation

$$f_{IF} = |f_{in} - n f_{LO}| \quad (14)$$

several spectral lines of different amplitude will be displayed on the CRT (Fig. 21). It is up to the user to locate the spectral lines within the required frequency range. This problem can be solved by offsetting the local oscillator frequency by a certain magnitude (Δf). This results in frequency shift of the mixture product from the nth harmonic and the upper sideband by an amount $n \times \Delta f$. This allows the user to distinguish between the "correct" spectral lines and others of different n and sign.

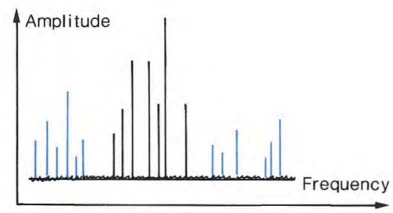


Fig. 21 Ambiguous display

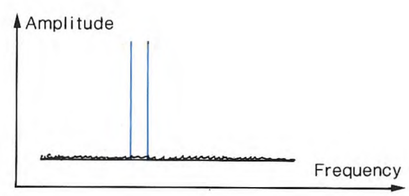


Fig. 22 Unambiguous display through use of preselector

Another suitable measure consists in connecting a tracking filter (preselector) before the mixer. This filter has a narrow bandpass and is tuned in parallel with the local oscillator so that its centre frequency coincides with the frequency displayed (Fig. 22). By using a tracking filter, the bandwidth of the spectrum analyzer's RF input section becomes narrow. The preselector is a YIG filter having a bandwidth of approx. 40 MHz that is connected before the mixer and tuned in parallel with the local oscillator (a YIG oscillator with an upper frequency limit of approx. 8 GHz) or with the selected harmonic of the oscillator, offset by the intermediate frequency. When a preselector is used, a lowpass filter at the input of the spectrum analyzer is not required since signals with frequencies that are not within the passband of the filter are largely, attenuated due to extreme impedance mismatch. Overloading of the mixer is avoided as the YIG preselector is driven into saturation by high-level signals.

Additional advantages offered by the use of a tracking filter are improved IF rejection and the reduction of RF leakage caused by the 1st local oscillator. An attenuator should, however, be connected before the input of the spectrum analyzer to avoid impedance mismatch of highly sensitive signal sources.

The use of a tracking filter results in additional errors (hysteresis) caused by its transmission loss. Another disadvantage is the reduced sensitivity resulting from the basic attenuation of approx. 6 dB. The frequency response can be improved by using an integrated preselector/mixer system.

For small tuning ranges of the YIG oscillator (covering one octave), the entire frequency range (e.g. from 2 to 18 GHz) cannot be displayed simultaneously. The frequency range is therefore split into several overlapping frequency bands with each band covering a frequency range of 2:1. Fig. 23 shows an example of setting the frequency bands where the 1st IF is 510 MHz and the 1st local oscillator is swept between 2 and 4 GHz.

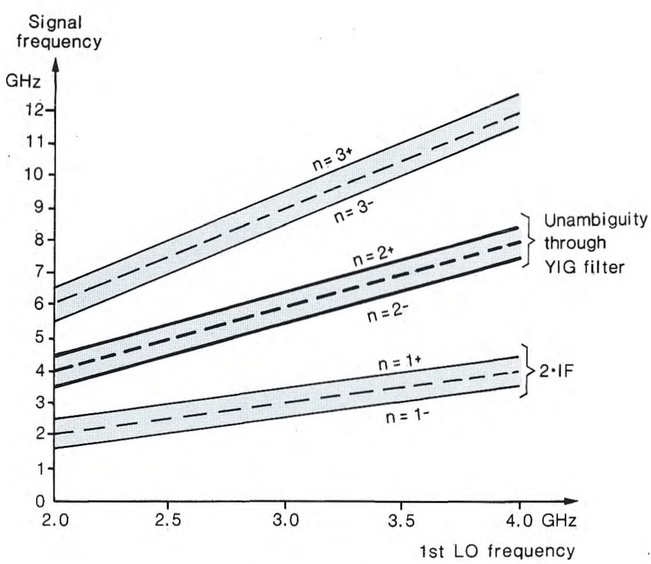


Fig. 23 Frequency bands (1st IF 510 MHz; 1st LO 2 to 4 GHz)

External waveguide mixers

The use of external waveguide mixing is provided for microwave analyzers and does therefore not apply to the Spectrum Analyzer FSA. Within the fundamentals of spectrum analysis, this method for frequencies above 18 GHz should however also be described. Above 18 GHz, waveguides must be used instead of coaxial cables. The first superheterodyne conversion is made using special waveguide mixer diodes (Fig. 24). The fundamental of the local oscillator signal (f_{LO}) is applied via the external mixer connection of the spectrum analyzer to the mixer diode together with a DC voltage defining the diode's operating point. Because of the non-linear characteristic of the diode, harmonics of f_{LO} are generated that are mixed with the signals applied to the waveguide. High-frequency mixture products are filtered out by a lowpass filter while "low-frequency" mixture products are fed back into the spectrum analyzer via its external mixer input. These mixture products are then filtered again by the IF bandpass filter and displayed on the monitor. Since preselection cannot be provided in most cases, several signals are generated within the IF bandwidth of the spectrum analyzer. By switching between two different oscillator frequencies, the wanted signal can be distinguished from spurious signals, but measurement and evalution are cumbersome and equivocal in the presence of multiple signals in the spectrum.

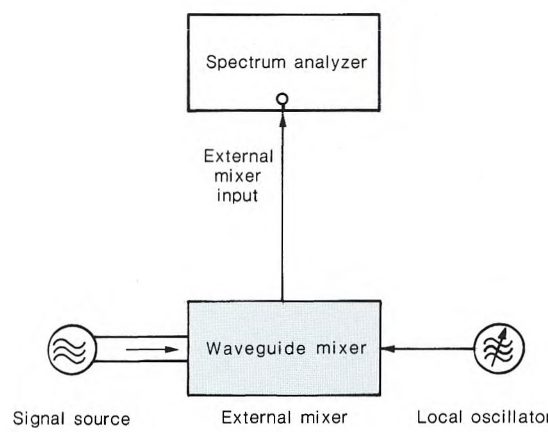


Fig. 24 Waveguide mixer

Non-linearities

Intermodulation

In some cases, a non-linear characteristic of electronic components such as semiconductor devices is desired when used for mixing and frequency multiplication. In other cases, the non-linear characteristic is, however, not desired, i.e. mixture products generated with two-signal measurements. Understanding of these principles is important for the user to ensure correct interpretation of analysis results. Since spectrum analyzers exhibit some non-linearity, manufacturers attempt to overcome this problem by employing state-of-the-art techniques.

By means of an example, driving a semiconductor diode with two different sine signals, this will be explained below. For low-level signals, the characteristic of a diode can be described by an exponential series:

$$i_D(t) = a_1 u(t) + a_2 u^2(t) + a_3 u^3(t) + a_4 u^4(t) \ldots \qquad (15)$$

When the diode is driven with very small signal levels, terms higher than the third term of the exponential series can be ignored.
The small-signal AC voltage, u(t), contains two components with frequencies f_1 and f_2 according to:

$$u(t) = \hat{U}_1 \sin(2\pi f_1 t) + \hat{U}_2 \sin(2\pi f_2 t) \qquad (16)$$

$\hat{U}_1$ and $\hat{U}_2$ include the possibility of amplitude modulation of both signals. By substitution of equation 16 in equation 15 and calculating the values up to and including the third term of equation 15, the results shown in the table below are obtained. It can be seen from this table how many additional spectral lines and mixture products can be generated by the semiconductor device although it is driven with only two signals of frequencies f_1 and f_2. In addition to the DC component and the fundamental, 2nd and 3rd order harmonics as well as 2nd and 3rd order difference frequencies and crossmodulation products will be present. The cubic term, a_3u^3 (equation 15) defines distortion products that may deteriorate the quality of the measurements.

DC component	$0.5\, a_2 (\hat{U}_1^2 + \hat{U}_2^2)$
Harmonics	$a_1 \hat{U}_1 \sin \omega_1 t$ $a_1 \hat{U}_2 \sin \omega_2 t$
Crossmodulation	$(a_3 0.75\, \hat{U}_1^3 + a_3 1.5\, \hat{U}_1 \hat{U}_2^2) \sin \omega_1 t$ $(a_3 0.75\, \hat{U}_2^3 + a_3 1.5\, \hat{U}_1^2 \hat{U}_2) \sin \omega_2 t$
2nd order harmonics	$-(a_2 0.5\, \hat{U}_1^2) \cos 2\omega_1 t$ $-(a_2 0.5\, \hat{U}_2^2) \cos 2\omega_2 t$
2nd order difference frequencies	$(a_2 \hat{U}_1 \hat{U}_2) \cos(\omega_1 - \omega_2)t$ $-(a_2 \hat{U}_1 \hat{U}_2) \cos(\omega_1 + \omega_2)t$
3rd order harmonics	$-(a_3 0.25\, \hat{U}_1^3) \sin 3\,\omega_1 t$ $-(a_3 0.25\, \hat{U}_2^3) \sin 3\,\omega_2 t$
3rd order difference frequencies	$-a_3 \hat{U}_1^2 \hat{U}_2 0.75 (\sin(2\omega_1 + \omega_2)t - \sin(2\omega_1 - \omega_2)t)$ $-a_3 \hat{U}_1 \hat{U}_2^2 0.75 (\sin(2\omega_2 + \omega_1)t - \sin(2\omega_2 - \omega_1)t)$

Table: Intermodulation products

Intermodulation (IM) is understood as the modulation of one signal by another. IM distortions can also be measured by applying three or more sinusoidal signals of different frequencies, and represent an essential factor that determines the quality of spectrum analyzer. This makes it obvious that components exhibiting square-law characteristics are required for many applications to avoid unwanted distortion products. Figures 25 to 28 show the spectra resulting from applying two sine signals of different frequencies to a diode (refer to Table).

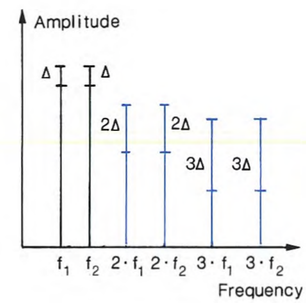


Fig. 25 Harmonics

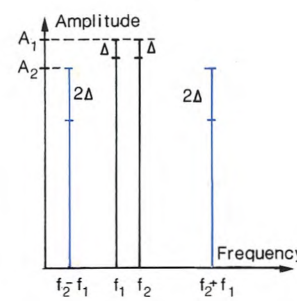


Fig. 26
2nd order difference frequencies

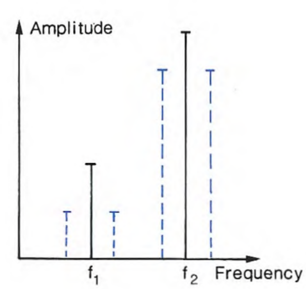


Fig. 27 Crossmodulation

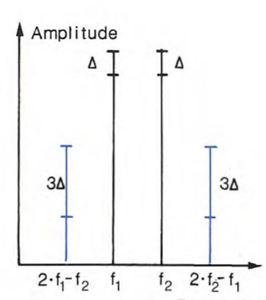


Fig. 28
3rd order difference frequencies

NON-LINEARITIES

Intercept point

The intercept point is a measure for distortions at the output of an amplifier, mixer, or similar device under test depending on the amplitude of the input signal. For a linear two-port network, a linear relationship exists between the levels of the input and the output signal of the same frequency. If this linearity is not maintained, additional signals will be present at the output of the two-port network. The levels of the output signals do not have a linear correlation with the level of the input signal (Table).

The intercept point is that "fictitious" value at which the input signal and the intermodulation products have the **same level.** The intercept point is a purely theoretical value since, in case of overloading, non-linearities and saturation effects occur in the device under test (Fig. 29).

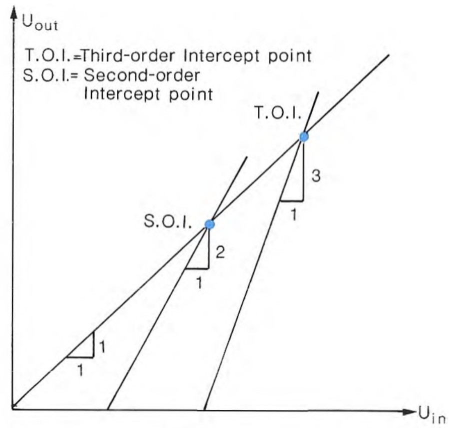


Fig. 29 Intercept point

Intermodulation distortions are amplitude-dependent and increase with higher signal levels. This increase is proportional to the order number of the intermodulation. 3rd order intermodulation results from a combination of the fundamental (1st harmonic) of one signal with the 2nd harmonic of another one. The order number of the intermodulation is calculated by adding the order number of the harmonics. If the amplitude of $\hat{U}_1$ and $\hat{U}_2$ (refer to intermodulation) is increased by magnitude $(\Delta U)^2$ whereas the amplitude of the intermodulation products (3rd order difference frequencies) is increased by magnitude $(\Delta U)^3$. In logarithmic scaling, this means an increase of $2 \times \Delta U$ (dB) and $3 \times \Delta U$ (dB), respectively.

Because of the excessive increase of the difference frequency levels, these will (mathematically) become identical to those of the input signals, so that the terms 2nd (SOI = second-order intercept) and 3rd order intercept points (TOI = third-order intercept) are used.
Correspondingly, an intercept point of nth order (NOI) can be calculated as well. The intercept points are mostly specified in dBm; their values should be as high as possible to obtain wide interference-free dynamic ranges.

With high-quality diode mixers, SOI values of up to +70 dBm can be achieved. Since this corresponds to a power of 10 kW into 50 Ω, it is obvious that the intercept point cannot be determined by practical measurements. The theoretical intercept point of nth order (NOI) can, however, be determined at far lower signal levels by employing mathematical methods.

Calculation of the intercept point

2nd order intercept point

$$\text{SOI} = U_1\text{dBm} + (U_1 - U_2)\ [\text{dB}] = (U_1 + \Delta U)\ \text{dBm} \quad (17)$$

3rd order intercept point

$$\text{TOI} = U_1\text{dBm} + \tfrac{1}{2}(U_1 - U_2)\ [\text{dB}] = (U_1 + \Delta U/2)\ \text{dBm} \quad (18)$$

nth order intercept point

$$\text{NOI} = U_1\text{dBm} + 1/(N-1)\,(U_1 - U_2)\ [\text{dB}] = (U_1 + \Delta U/(N-1))\ [\text{dBm}] \quad (19)$$

Example:
Comparison of specifications of two spectrum analyzers

Spectrum analyzer 1
When applying two signals of different frequencies having the same level of −30 dBm, the rejection of 3rd order difference frequencies is better than 70 dB.

Spectrum analyzer 2
When applying two signals of different frequencies having the same level of −40 dBm, the rejection of 3rd order difference frequencies is better than 100 dB.
Which spectrum analyzer is the better one?

Solution

Spectrum analyzer 1
3rd order intercept point = −30 + 70/2 = +5 dBm

Spectrum analyzer 2
3rd order intercept point = −40 + 100/2 = **+10 dBm**

It should be noted that:

▷ if an amplifier is connected before an active component, the intercept point value of the latter is impaired by the gain of the amplifier, and

▷ if an attenuator is connected before an active component, the intercept-point value of the latter is improved by the attenuation.

The latter method is employed very often to improve the linearity for wanted signals of sufficient amplitude.

Input/output intercept point
According to Fig. 30, all values (the TOI in this example) are most often measured at the output of an active component. The value determined in this way is called "output intercept point" (IP_{3out}). This value is, however, not commonly used since it depends on the gain of the active component and can therefore be manipulated to claim impressive specifications.

The input intercept point, IP_{3in} is determined after calculation of IP_{3out} and subtraction of the gain value. This input intercept point is a solid basis for comparisons and therefore used in most cases.

The ideal conditions represented in this brochure where an exponential series was previously used to describe the forward characteristic of a diode do not apply to all types of active components. FETs exhibit, for example, significant variations of the intercept points over their linear dynamic range. In practice, frequency dependence can furthermore be ascertained.

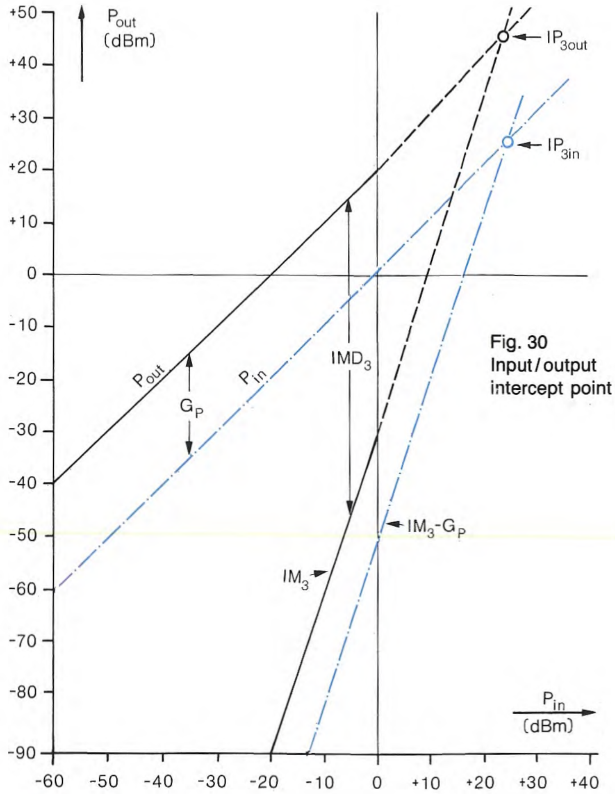
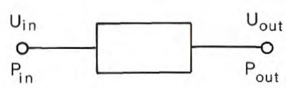


Fig. 30 Input/output intercept point

1-dB compression point

With increasing amplitude of the signal used to drive the input of a two-port network, the correlation between the output and the input signal becomes more and more non-linear (Fig. 31):

Fig. 31 Two-port signals



Voltage ratio = 20 log (U_{out}/U_{in}) [dB],
Level ratio = 10 log (P_{out}/P_{in}) [dB].

Level P_{in} at which the ratio of output power P_{out} to input power P_{in} is 1 dB less than gain V_{in}, is called the "1-dB compression point" (Fig. 32). As a rule of thumb, the 1-dB compression point is approx. 10 to 15 dB below IP_3. The non-linearities of a two-port network can be so complex, that the 1-dB compression point cannot be explicitly determined by solving a single equation.

For measurement applications, only the linear portion of the dynamic range is of any interest. The difference between the power values in the linear and the non-linear portion of the transfer function results from excessive increase of the harmonics contents.

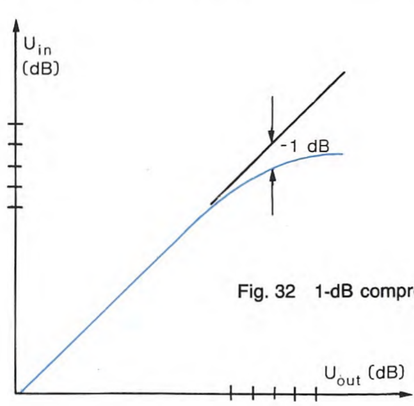


Fig. 32 1-dB compression point

Noise

Noise sources

Noise is understood as variations of signal amplitude versus time according to statistical laws. Noise that contains signals of all theoretically possible frequencies having identical amplitude is called "white noise". Noise can be generated within a circuitry or by external sources. Noise generated by external sources shall be discussed only briefly. In addition to the signal power ($P_{S,a}$) also some noise power ($P_{N,a}$) is, for instance, received by an antenna (Fig. 33). The value of $P_{N,a}$ depends on the frequency (f) and on the elevation angle of the antenna [5].

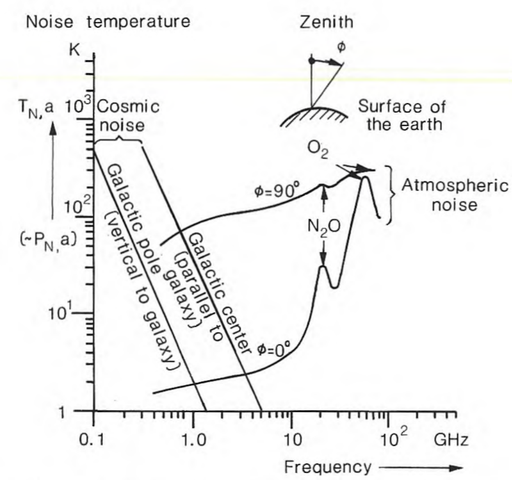


Fig. 33 Antenna noise temperature as a function of frequency (noise power as described in [5])

The noise power can also be caused by extra-terrestrial radiation sources that are mainly located within the Milky Way (galactic or cosmic noise). Noise is also generated by the absorption of electromagnetic radiation in the atmosphere and by the thermal noise of earth.

Noise sources within the circuitry generate a noise power ($P_{N,V}$), that is mostly caused by semiconductor noise and thermal noise of ohmic resistances as well as by charge-carrier flow. Thermal noise is generated by irregular movement of charge carriers in resistive materials. The average value of this current is zero whereas the rms value is not.

For a frequency span Δf, the noise current generated by the conductance element (G) which is at the absolute temperature (T) can be calculated as follows:

$$(i_{N,rms})^2 = 4kTG\Delta f \qquad (20)$$

The available noise power is then:

$$P_{N,V} = kT\Delta f \qquad (21)$$

where:

k = Boltzman's constant = 1.38×10^{-23} Ws/K
T = absolute temperature of the resistive element (K)
Δf = effective noise bandwidth (Hz)

The entire noise power is available at a noise-free load at T = 0 having the same conductance as the noise source.
Equation 21 applies to the following frequencies:

$$f[\mathrm{GHz}] << kT/h = 20.8\,T\,[\mathrm{K}] \qquad (22)$$

where:
$h = 6.62 \times 10^{-34}$ Js = Planck's constant

Current noise has an average value other than zero. The DC current through a vacuum tube varies statistically corresponding to the number of the electrons hitting the anode. The same applies for semiconductor junctions but not for bulk resistances (thermal noise). Shot noise, flicker noise, current distribution noise, or recombination noise provide that additional noise power.

Noise factor, noise figure

The most commonly used measure of sensitivity is the noise factor (F). If the wanted signal, P_{Sin}, and the noise power, P_{Nin}, are present at the input of a two-port network, the values of P_{Sout} and P_{Nout} can be measured at the output. The noise factor is then defined as follows:

$$F = \left(\frac{P_{Sin}}{P_{Nin}}\right) / \left(\frac{P_{Sout}}{P_{Nout}}\right) \qquad (23)$$

With S = signal and N = noise, the definition

$$F = (S/N)_{in}/(S/N)_{out} \qquad (24)$$

is equivalent to equation 23. The noise factor specifies by which factor the signal-to-noise (S/N) ratio is reduced by the two-port network. The relationship between noise figure and noise factor is as follows:

$$F[\mathrm{dB}] = 10 \log F \qquad (25)$$

The noise level defines the sensitivity of a spectrum analyzer (see phase noise) since only signals with amplitudes above the noise level can be displayed.
The determination of the average noise level is often difficult as the statistical variations are displayed on the spectrum analyzer as a "noise floor" extending over a width of several graticule units (Fig. 34). The lower the intrinsic noise of the spectrum analyzer, the higher its sensitivity. Compared with spectrum analyzers featuring a noise figure of approx. 25 dB, the FSA exhibits extremely low intrinsic noise of 12 dB (typical value). The average noise level can be determined by activating the video filter (Fig. 35) or by digital averaging.

In addition to specifying absolute sensitivity, another definition is often given: To make a signal still discernible, the signal power must be equal to the noise power. Thus the sensitivity can also be defined as:

$$\frac{\text{Signal power} + \text{Noise power}}{\text{Noise power}} = 2 \qquad (26)$$

This means that the amplitude caused by the input signal should be 3 dB above the noise. The obtainable sensitivity depends essentially on the operating bandwidth of the spectrum analyzer.

Equivalent noise bandwidth

The spectral noise figure is tailored to the reception and the transmission of infinitesimal frequency ranges (df). In practice, finite bandwidths are used both at the input and output. For a comparison of the signal and noise power densities (for a bandwidth of 1 Hz), either the frequency dependence of noise figure F (f) or, simplified, an average noise figure (F) based on the latter can be specified. This does, however, not reflect that the reductions of the signal-to-noise ratio can be above or below the value of F.

The concept of "equivalent noise bandwidth" is easily understood by considering an ideal, intrinsically noise-free, two-port network having a rectangular bandpass characteristic of bandwidth B_N that is symmetrical to the centre frequency, f_c, and which provides the same gain, such that the noise power will be the same at its output as the noise power at the output of a noise-generating two-port network.

With "coloured" input noise (i.e. noise with frequency-dependent amplitude), B_N depends not only on the transfer function of the transmission two-port network but also on the type of the input noise. The equivalent noise bandwidth is often (even in literature) incorrectly equated with the bandwidth of the wanted signal.

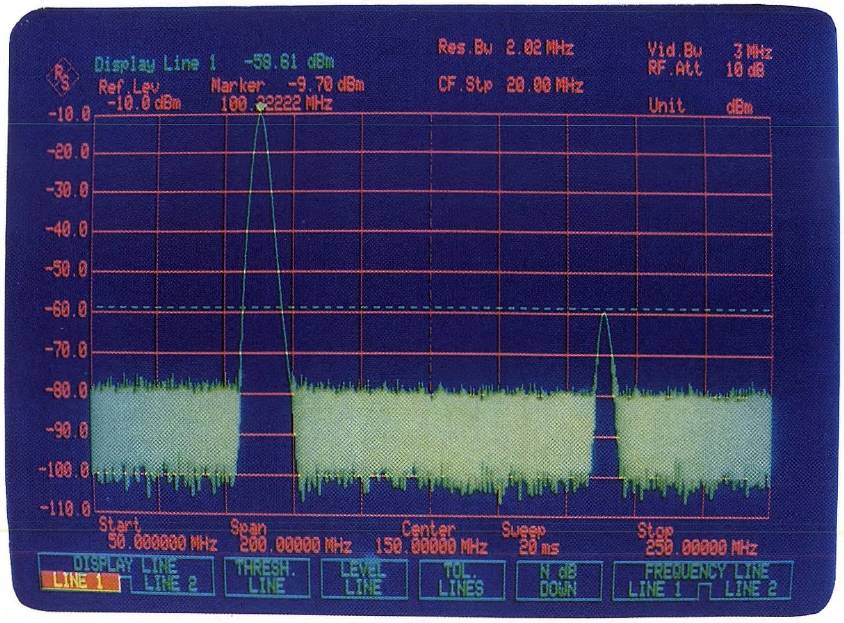


Fig. 34 With large video bandwidths, noise floor makes identification of small signals difficult or conceals such signals at all

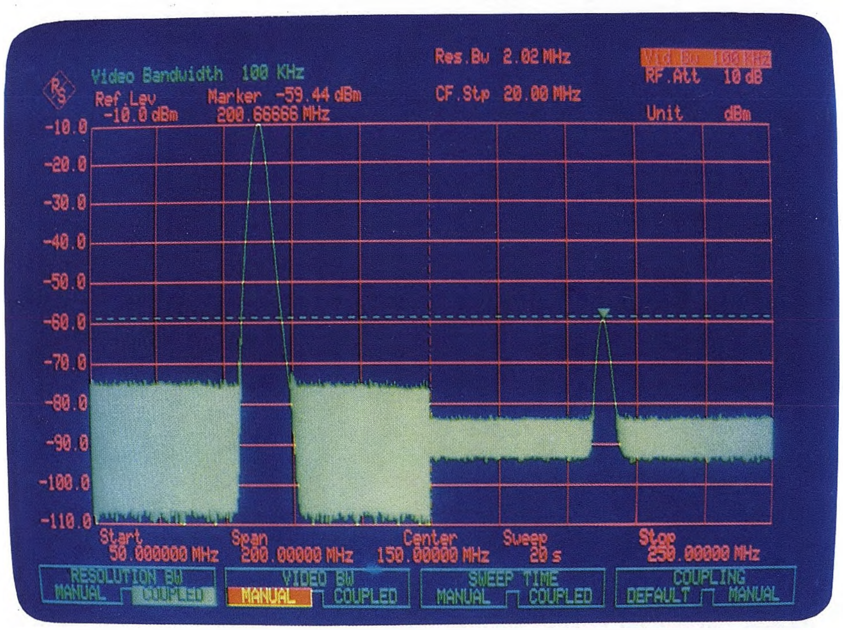


Fig. 35 Determination of noise level with different video filter settings

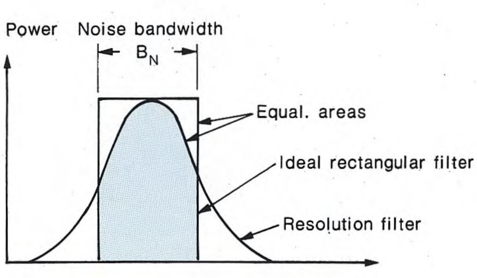


Fig. 36 Equivalent or effective noise bandwidth

According to [5], for the relationship between B_N and B_{SIG} and a simple tuned circuit, a correction factor of

$$\frac{B_N}{B_{SIG}} = \Pi/2 \qquad (27)$$

is obtained.

Depending on the transfer function, other correction factors are to be considered.

The method for the calculation of the effective noise bandwidth applied in equations 20 and 21 results from the equivalence of the areas below the curves in Fig. 36. For various filter types, mathematical approaches can be used since measurements of the effective noise bandwidth are rather elaborate.

The noise level increases with rising filter bandwidth, namely by 10 dB for a bandwidth increase of 10. When comparing spectrum analyzers, it should be ensured that the sensitivities are referred to the same filter bandwidth since otherwise mathematical conversion will become necessary.

$$(S/N)_{effective} = (S/N)_{measured} + 10 \log \frac{\text{resolution bandwidth}}{\text{system bandwidth}} \qquad (28)$$

The correction factor of the signal-to-noise ratio for audio measurements (i.e. at a system bandwidth of 20 kHz) and at an IF filter bandwidth of 3 kHz is:

$$10 \log \frac{3\ \text{kHz}}{20\ \text{kHz}} = -8.2\ \text{dB} \qquad (29)$$

Absolute sensitivity

By determining the noise figure (see noise factor, noise figure), it is possible to determine the difference in dB between the intrinsic noise of the spectrum analyzer and the thermal noise level generated by a resistor. According to equation 21, the noise power of any resistor at the ambient temperature of +17° C, referred to a bandwidth of 1 Hz, is:

$$P_N/\Delta f = 1.38 \times 10^{-23} \times 290 = 4 \times 10^{-21}\ \text{W/Hz} \qquad (30)$$

At a bandwidth of 1 Hz, this yields

$$P_N = 10 \log 4 \times 10^{-18} = -174\ \text{dBm} \qquad (31)$$

This noise power corresponds to a voltage of 4.5×10^{-10} V into 50 Ω or to −67 dBµV. If the sensitivity of a spectrum analyzer is −140 dBm over a bandwidth of 10 Hz, the absolute sensitivity (S) can be calculated as follows:

$$S = -174\ \text{dBm/Hz} + 10 \log (\Delta f[\text{Hz}]/1\ \text{Hz})\ \text{dB} + F[\text{dB}] \qquad (32)$$

This yields a noise figure of

$$F = -140 + 174 - 10 = 24\ \text{dB} \qquad (33)$$

The theoretical absolute sensitivity of −174 dBm could only be achieved if the spectrum analyzer were not exhibiting intrinsic noise (having the ideal noise factor of F = 1). At a resolution bandwidth of 6 Hz, the FSA achieves an absolute sensitivity of better than −150 dBm (Fig. 37). This is equivalent to a noise figure of approx. 12 dB as calculated from equation 32 that has not been realized with other spectrum analyzers to date. Thanks to its extremely low intrinsic noise, the FSA is capable of displaying even low-level signals with amplitudes in the range between −150 and −145 dBm that would otherwise be completely buried in noise.

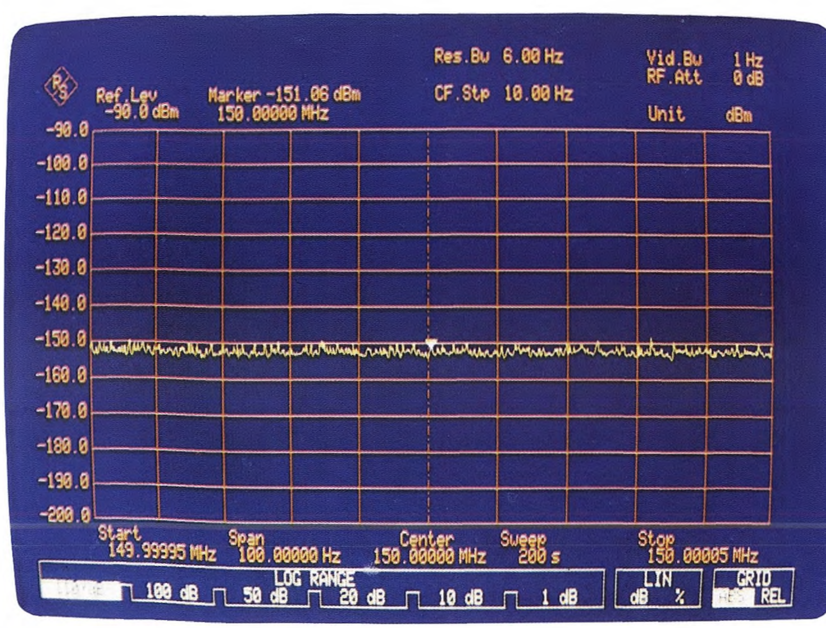


Fig. 37 Display of absolute sensitivity of Spectrum Analyzer FSA

Phase noise

The term "phase noise" describes all irregularities of the phase relationship of a signal, whether being low- or high-speed phenomena. All effects that widen the representation of an oscillator signal beyond the infinitely narrow line widths of the ideal oscillator are called phase-noise effects [6]. For free-running oscillators, the phase-noise amplitude close to the carrier frequency (offset by less than 1 kHz) is always very high (Fig. 38).

By employing phase synchronization, the phase noise close to the carrier frequency can be improved significantly, depending on the characteristics of the control loop. Within a certain spacing from the carrier frequency, the phase noise exhibited by a phase-locked oscillator can be worse than that of a free-running oscillator but will be reduced at larger spacing from the carrier frequency (Fig. 39).

Phase noise (W_{FM}) is much higher than amplitude noise (W_{AM}) in the vicinity of the resonance frequency, but will decrease more rapidly with increasing spacing from f_{LO}, so that both noise components will finally approach similar values as shown in Fig. 40.

Close to the carrier frequency, the dynamic range depends essentially on the spectral purity of the oscillator signal. The Spectrum Analyzer FSA features very high spectral purity of its oscillator signal measured at 1 kHz from the carrier frequency and with a bandwidth of 1 Hz. Its sideband noise is only −114 dBc whereas conventional spectrum analyzers have sideband noise levels between −85 dBc and −90 dBc when measured under identical conditions (Fig. 41).

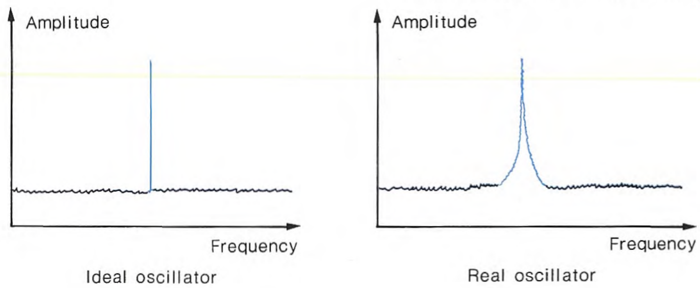


Fig. 38 Oscillator signal

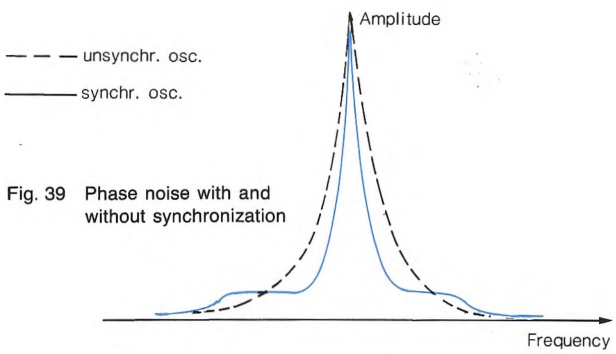


Fig. 39 Phase noise with and without synchronization

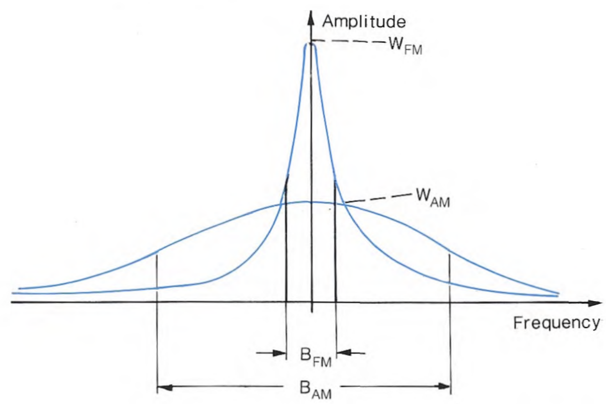


Fig. 40 AM-FM noise

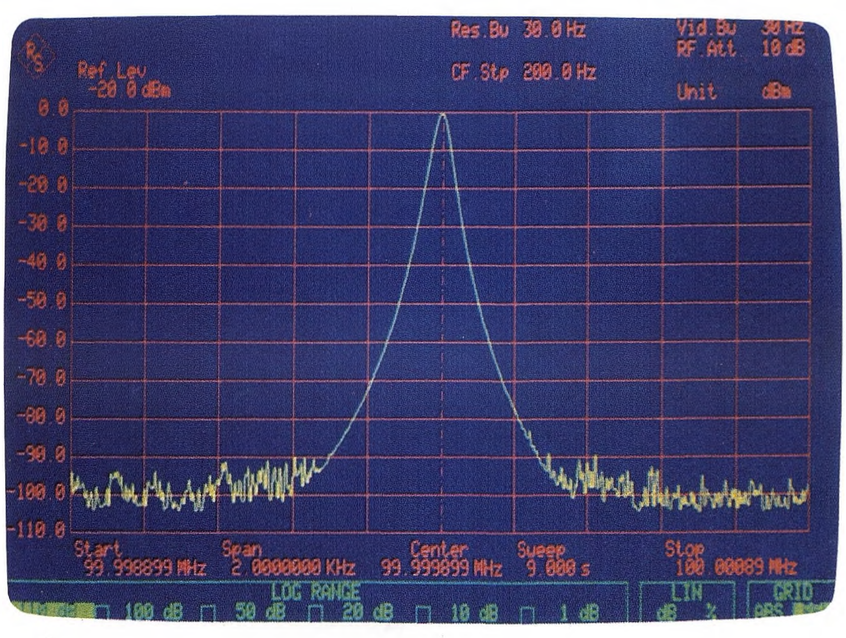


Fig. 41 Display of noise sidebands on Spectrum Analyzer FSA

DYNAMIC RANGE

Dynamic range

Measurement range

The measurement range of a spectrum analyzer is defined as the range of permissible input levels. Depending on the bandwidth selected, the smallest value that can be displayed is limited by the absolute sensitivity (see phase noise). The upper limit of the measurement range is set by the maximum permissible signal level applied to the input mixer (e.g. + 20 dBm at an absolute sensitivity of – 140 dBm). By using an input attenuator with a maximum power-handling capacity of 1 W (equal to + 30 dBm), the measurement range is 170 dB. However, a measurement range of 170 dBm cannot be displayed.

When multiple signals are applied, the optimum dynamic range is obtained at a signal level of e.g. – 40 dBm at the input of the mixer stage, thus avoiding mixture products of adjacent signals. If the power of the input signals does not exceed 1 W, a 70-dB input attenuator should be used to ensure a level of – 40 dB at the input of the mixer. To achieve the maximum signal-to-noise ratio with a signal applied to the mixer, the level is increased to e.g. – 20 dBm. This means that low RF attenuation and high IF gain are to be selected (Fig. 42).

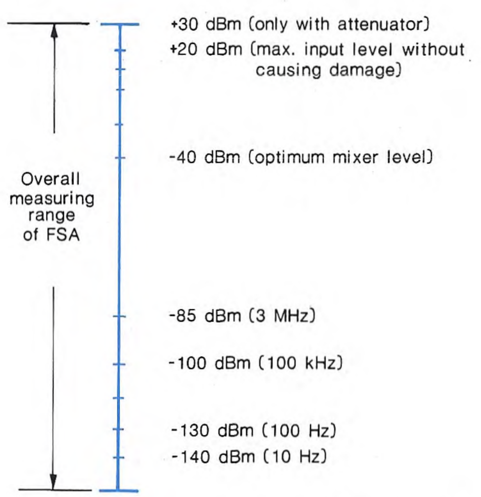


Fig. 42 Measurement range of spectrum analyzer as a function of resolution bandwidth

Display range

The vertical display can be calibrated with linear or logarithmic scaling. For a linear display, the graticule is divided into volts per division (similar to an oscilloscope) where the lowermost graticule line is used as the reference line. For a logarithmic display, the uppermost graticule line is used as the reference (LOG REF LEVEL).

Advanced spectrum analyzers can display a range of 100 dB. This means that two signals having levels of 100 mW and 10 pW can be displayed simultaneously and easily distinguished. The display range is limited by non-linearities of the logarithmic amplifier. The filter shape and phase noise must be taken into account as well. This leads to the definition of the "usable" range which is the part of the display range or linear dynamic range that can be used for measurements, i.e. that is free from intermodulation (IM), 3rd order distortions and noise.

Intermodulation-free range

The maximum dynamic range for low-level signals is defined by the noise figure and the resolution bandwidth (see also resolution bandwidth). For high-level signals, the dynamic range is limited by distortion products resulting from finite linearity of the RF input attenuator. Generally, the following definitions apply:

$$\text{Maximum dynamic range [dB]} = \frac{N-1}{N} (S\,[\text{dBm}] - NOI\,[\text{dBm}]) \quad (34)$$

where:

N = order number of distortion product

S = absolute sensitivity caused by intrinsic noise

NOI = nth order intercept point

Commonly, only 2nd and 3rd order distortion products will be of any interest. The maximum dynamic range can be calculated from equation 34.

2nd order

$$\text{Maximum dynamic range [dB]} = 0.5\,(S\,[\text{dBm}] - SOI\,[\text{dBm}]) \quad (35)$$

3rd order

$$\text{Maximum dynamic range [dB]} = 2/3\,(S\,[\text{dBm}] - TOI\,[\text{dBm}]) \quad (36)$$

For all input levels (IL) higher than the sum of absolute sensitivity and maximum dynamic range, a corresponding attenuation at the RF input is required. If the attenuation is excessive, the dynamic range is reduced for low-level signals. When the attenuation is too small, distortion products are generated. Generally, the following applies to RF attenuation:

$$\text{RF attenuation [dB]} = IL\,[\text{dBm}] - (N-1)/N\,(NOI\,[\text{dBm}]) - 1/N\,S\,[\text{dBm}] \quad (37)$$

A compromise is to be made with respect to the adjustment of the RF attenuation, since an attenuator is most often switchable in 10- or 5-dB steps. With the Spectrum Analyzer FSA, the RF attenuation can be varied in steps of 10, 5 and 1 dB.

The maximum dynamic range shall be determined by means of an example. Also the RF attenuation, the input level at the mixer and the absolute sensitivity will be calculated.

Example

The noise figure (F) of a spectrum analyzer was measured to be 20 dB at a bandwidth of 10 Hz. The 2nd order intercept point was calculated as +60 dB and IP_3 as +10 dB (equations 17 and 18). The input level (IL) is 1 mW (equal to 0 dBm).

By using equation 32, the absolute sensitivity is calculated as:

$$S = -174 + 20 + 10 = -144 \text{ dBm} \quad (38)$$

Using equation 35 or 36, the maximum dynamic range for IP_2 is:

$$\text{Dynamic range} = 0.5\,(+144+60) = 102 \text{ dB} \quad (39)$$

The following applies to IP_3:

$$\text{Dynamic range} = 0.667\,(+144+10) = 102.6 \text{ dB} \quad (40)$$

The RF attenuation can be calculated by using equation 37:

$$\text{RF attenuation } (IP_2) = 0-30 + 72 = 42 \text{ dB} \quad (41)$$

$$\text{RF attenuation } (IP_3) = 0-6.6 + 48 = 41.4 \text{ dB} \quad (42)$$

Because of the RF attenuation, the following levels are applied to the input of the mixer for the given signal level:

$$\text{2nd order:} \quad 0-42 \text{ dB} = -42 \text{ dBm} \quad (43)$$

$$\text{3rd order:} \quad 0-41.4 \text{ dB} = -41.4 \text{ dBm} \quad (44)$$

For the above example, a distortion-free dynamic display range as shown in Fig. 43 is obtained. Sine signals are assumed in this example.

For wideband noise signals, the lower limit of the dynamic range is defined by the absolute sensitivity, and its upper limit by the 1-dB compression point as well as by the 3-dB RF bandwidth. The permissible noise power per Hz is:

$$NP\,[\text{dBm/Hz}] = CP\,[\text{dBm}] - 10 \log (\text{RF bandwidth}/1 \text{ Hz})\,[\text{dB}] \quad (45)$$

where:

NP = noise power (dBm / 1 Hz)

CP = 1-dB compression point of input mixer

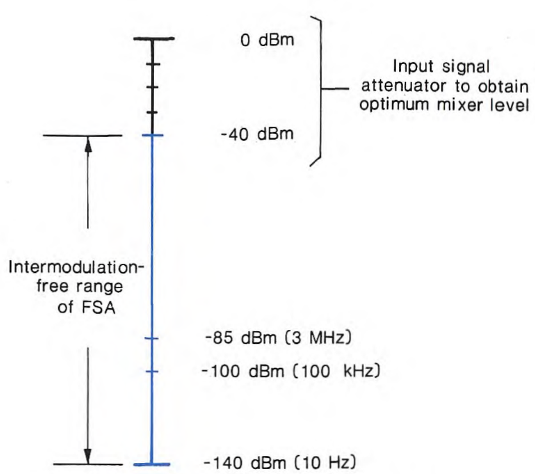


Fig. 43 Dynamic display range (distortion-free)

Example

For a spectrum analyzer, the maximum permissible input noise power shall be determined at a 3-dB RF bandwidth of 2 GHz and a 1-dB-compression point of −10 dBm. What power will be displayed for this noise power if the narrowest IF noise bandwidth is 10 Hz?

$$NP = -10 \text{ dBm} - 10 \log (2 \times 10^9/1) \text{ dB} = \mathbf{-103 \text{ dBm/Hz}}$$

$$\text{Display} = -103 \text{ dBm/Hz} + 10 \log (10/1) \text{ dB} = \mathbf{-93 \text{ dBm.}}$$

For wideband correlated signals, the absolute sensitivity is determined as the lower limit and converted into dBµV/MHz. The upper limit of the dymanic range is defined by the 1-dB compression point of the input mixer and by the 6-dB RF bandwidth.

The permissible spectral density per Hz is:

$$SD\,[\text{dB}\mu\text{V/MHz}] = CP\,[\text{dBm}] - 20 \log_{10} (\text{RF bandwidth}/(1 \text{ MHz}))\,[\text{dB}] \quad (46)$$

where:

SD = spectral density (dBµV / MHz)

CP = 1-dB compression point of the input mixer (dBm)

Example

For a spectrum analyzer, the maximum spectral density in dB (µV / MHz) shall be determined at a 6-dB RF bandwidth of 2 GHz and a 1-dB compression point of −10 dBm (=97 dBµV). What voltage-related value will be displayed on the spectrum analyzer for a spectral density of 100 kHz?

$$SD = 97 \text{ dB}\mu\text{V} - 20 \log_{10} (2 \times 10^3/1) \text{dB}, = \mathbf{+31 \text{ dB}\mu\text{V/MHz.}}$$

Displayed voltage-related value

$$= +31 \text{ dB}\mu\text{V/MHz} + 20 \log_{10}(0.1/1) \text{ dB} = \mathbf{+11 \text{ dB}\mu\text{V at 100 kHz}}$$

FILTER CHARACTERISTICS

Filter characteristics and signal resolution capability

Shape factor

The resolution capability of a spectrum analyzer is defined by the characteristics of the IF filter alone. Since rectangular filter shapes cannot be realized in practice, both the bandwidth and the shape of a filter must be defined. The filter bandwidth is either specified for an attenuation of 3 or 6 dB. The selectivity of a filter is specified by its shape factor that defines the slope of the filter skirt.

Definition of shape factor (SF) according to Fig. 44:

$$SF = \frac{\Delta f\ (60\ dB)}{\Delta f\ (3\ dB)} \qquad (47)$$

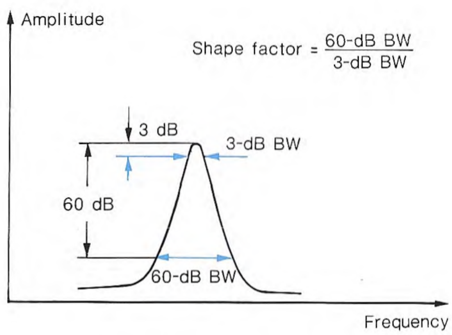


Fig. 44 Definition of shape factor

The smaller the shape factor, the steeper are the slopes of the filter skirt and the better is the selectivity. Better selectivity will, however, require sweep time T to be increased (see relationships between frequency span, sweep time and resolution bandwidth) since the settling time of a filter increases with greater skirt selectivity.

Five-pole Gaussian filters consisting of passive components have shape factors in the range from 10 to 12. The ideal, rectangular-shape filter has a shape factor of 1. With crystal filters, shape factors below 3 can be realized. However, these filters exhibit relatively poor settling characteristics.

Filter settling time

The pulse response of a filter depends on

▷ its pulse width and
▷ its characteristics.

Filters having almost rectangular transfer functions, i.e. small shape factors, contribute to better frequency resolution but result in increased measurement times since high-level input signals must be applied for a longer period of time before the signal amplitude at the filter output reaches its final value (see signal dwell time). Gaussian filters feature improved settling characteristics. The settling time is inversely proportional to the filter bandwidth and should be multiplied by a constant whose value depends on the special characteristics of the filter:

$$t_s = K / B \qquad (48)$$

where:
t_s = settling time
K = filter constant
B = 3-dB filter bandwidth

Signals of equal amplitudes

To be able to distinguish two signals of equal amplitude at adjacent frequencies, a 3-dB dip is defined (Fig. 45). This dip is reached close to the 6-dB bandwidth of the resolution filter. This absolute sensitivity still permits determination of the maximum levels of both signals.

Signals of very different amplitudes

The resolution capability of a spectrum analyzer for signals of very different amplitudes can be significantly improved by using filters with steeper slopes. Generally applicable definitions do not exist, as this characteristic of a spectrum analyzer depends on a variety of parameters. Very high resolution would be desirable so that two adjacent spectral components can be distinguished.

Fig. 46 shows the amplitude versus frequency characteristics of a filter at high and low resolution values where the filter is slowly tuned over the display range. For sufficiently different frequencies or narrow filter bandwidths, a dip is created by superposition of both curves. If the dip is clearly defined, amplitude and frequency of the lower-level signal can be determined with sufficient accuracy.

It is always recommended to use a filter whose bandwidth is narrow enough to result in separate spectra for both signals. As a matter of course, the filter bandwidth cannot be reduced to zero. The minimum filter bandwidth is limited by the relationship between sweep time T and dwell time t_d.

Signal dwell time

If the frequency is tuned over frequency range F (display width) of 3-dB bandwidth B during time T, a frequency component dwells within the filter bandwidth for time t_d. The relationship between these parameters is described by the following equation:

$$t_d = BT / F \qquad (49)$$

To allow for settling of the analyzer filter, a minimum dwell time (see equation 48) is to be maintained:

$$t_{d,min} = t_s = K / B \qquad (50)$$

If the sweep time is sufficiently long (i.e. $t_d > t_s$), the analyzer filter is settled completely, ensuring that the signal displayed reflects the static filter transfer function. If sweep time T is reduced, frequency span F is increased or analyzer bandwidth B is decreased so that the condition $t_d > t_s$ is not met, it can be seen that both the signal amplitudes and the resolution are reduced (Fig. 47).

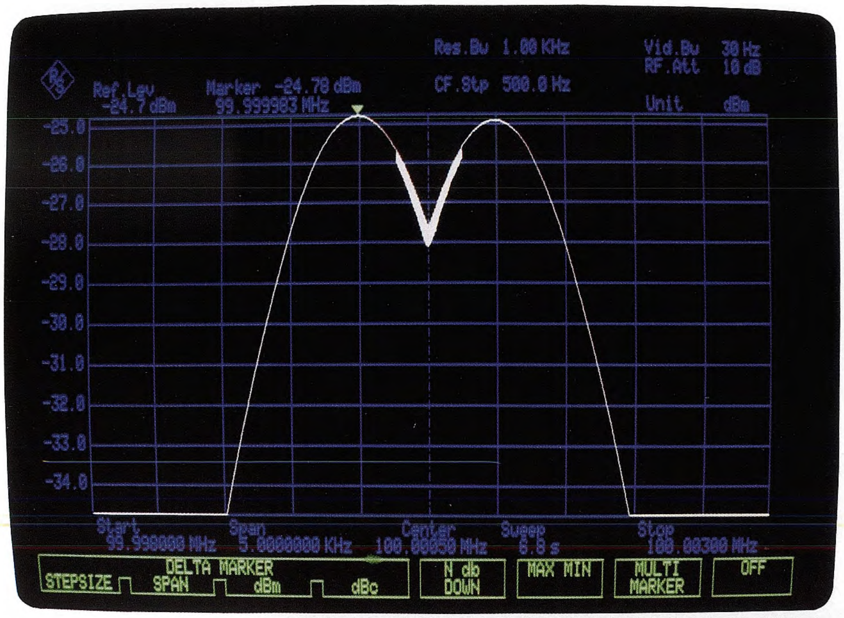


Fig. 45 Resolution capability for signals of equal amplitudes

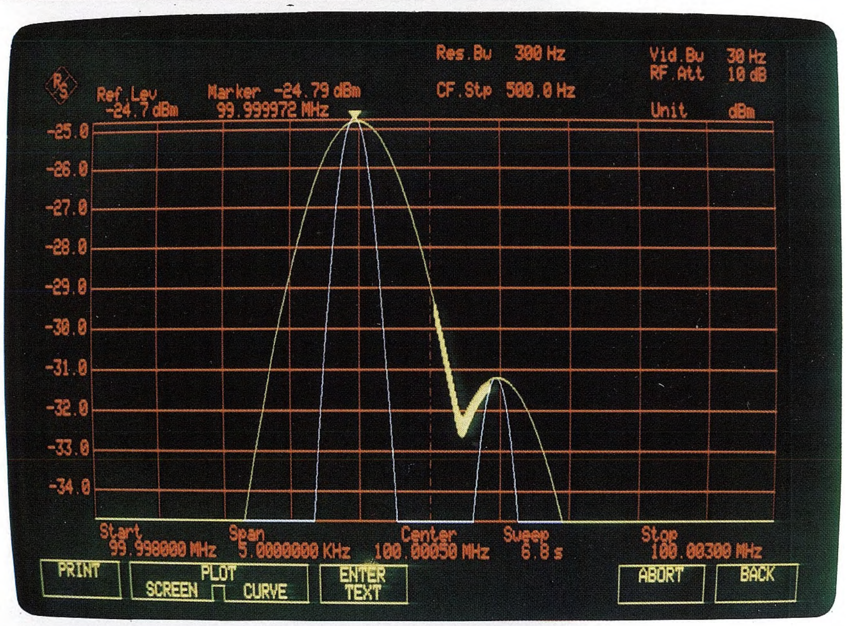


Fig. 46 Display of two signals of different amplitudes at different resolutions

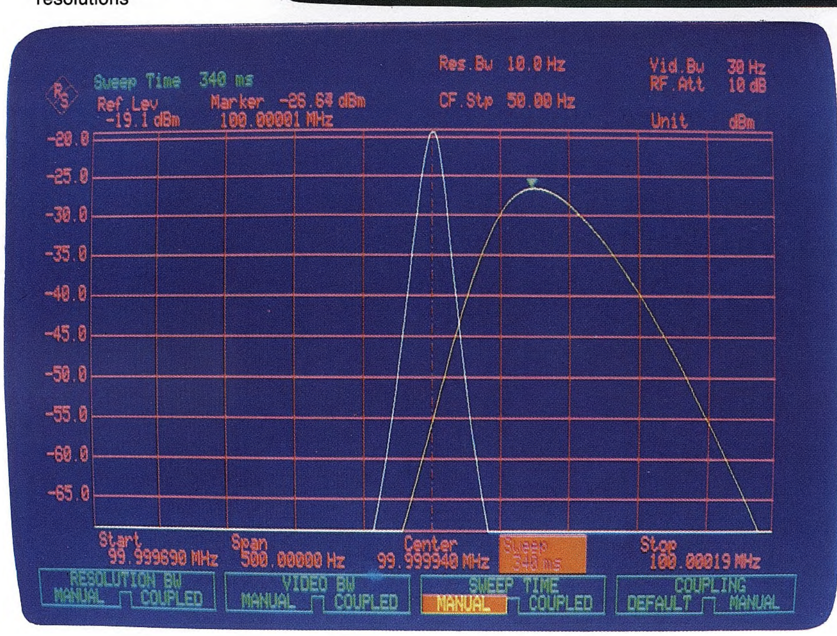


Fig. 47 Amplitude loss as a result of excessive sweep time

FILTER CHARACTERISTICS

The shape of the passband transfer function of the filter can also be of some importance. To allow coarse examination of the signals, a filter with relatively low skirt selectivity can be used. This permits reduction of the settling time as opposed to a filter of equal 3-dB bandwidth but higher skirt selectivity, which allows use of a shorter sweep time. If, however, the frequency range to be sampled should be examined as accurately as possible (especially with respect to very small amplitudes) close to spectral components of high amplitude, a filter with high skirt selectivity must be used (see signals of very different amplitudes).

Amplitude loss as a function of dwell time

With insufficient dwell time of the signal within the passband of the analyzer filter, an amplitude loss (ΔA) will result from incomplete settling of the filter. For Gaussian filters, the following relationship [2] applies:

$$\Delta A = \{1 + 0.195\,[F/(T B_{st}^2)]^2\}^{-0.25} \qquad (51)$$

where:

ΔA = amplitude loss
F = frequency span
T = sweep time
B_{st} = static resolution bandwidth

For constant frequency span and a given filter, Fig. 48 shows the amplitude loss as a function of sweep time.

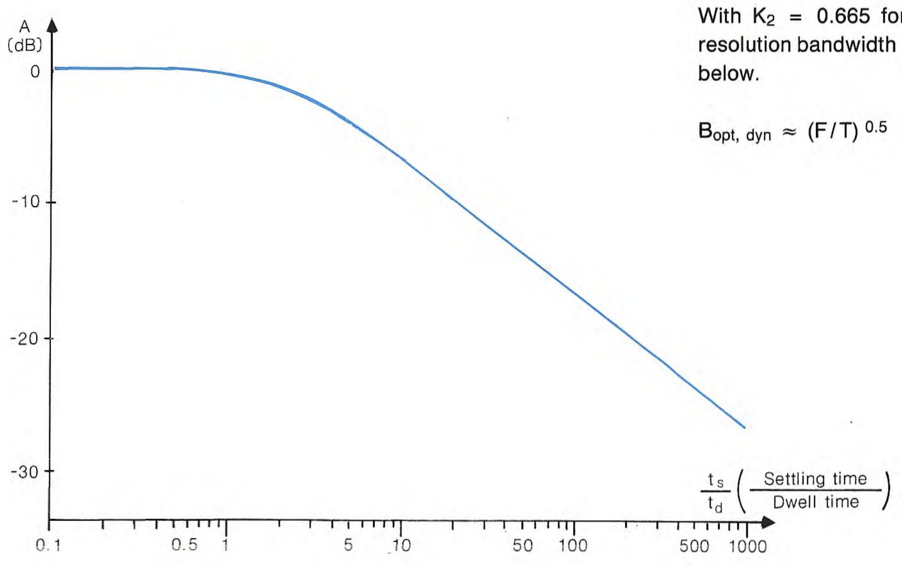


Fig. 48 Amplitude loss as a function of dwell time

Resolution bandwidth

In the previous sections, no distinction has been made between resolution and bandwidth so that these terms have been more or less interchangeable. The definitions of these two terms are not identical. **Resolution** refers to the signal or spectrum of signals displayed on the screen whereas **bandwidth** is related to the electrical characteristics of an amplifier contained in the spectrum analyzer. Both terms are correlated by sweep time and frequency span. In the following, the term "resolution" is to be understood as the dynamic resolution bandwidth while the term "bandwidth" means the static resolution bandwidth. If dwell time t_d during which the signal is within the passband of the filter is made sufficiently long (i.e. $t_d > t_s$), the dynamic resolution bandwidth will become equal to the static resolution bandwidth.

This is the case in equation 52 if $F << T B_{st}^2$.

$$B_{dyn} = B_{st}\{1 + K\,[F/(T B_{st}^2)]^2\}^{0.5} \qquad (52)$$

where:

B_{dyn} = dynamic resolution bandwidth
B_{st} = static resolution bandwidth
F = frequency span
T = sweep time
K_1 = filter constant (0.195 for Gaussian filters)

Fig. 49 shows the relationship between dynamic resolution bandwidth and dwell time at constant frequency span and constant static resolution bandwidth.

For a given sweep time and frequency span, the optimum dynamic resolution [3] bandwidth can be calculated by using equation 52.

$$B_{opt,\,dyn} = 1.41\,B_{opt,\,st} = 1.41 \times K_2\,(F/T)^{0.5} \qquad (53)$$

where:

$B_{opt,\,dyn}$ = optimum dynamic resolution bandwidth
$B_{opt,\,st}$ = optimum static resolution bandwidth
K_2 = filter constant (0.665 for Gaussian filters)

With K_2 = 0.665 for Gaussian filters, the optimum dynamic resolution bandwidth can be calculated approximatively as seen below.

$$B_{opt,\,dyn} \approx (F/T)^{0.5} \qquad (54)$$

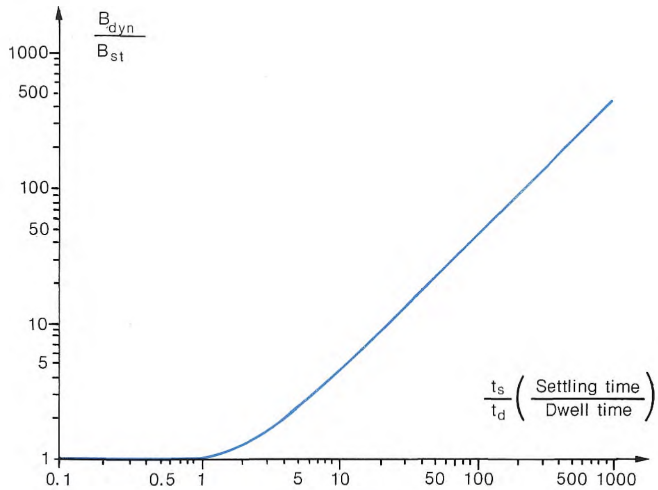


Fig. 49 Dynamic resolution bandwidth

Relationships between frequency span, sweep time and resolution bandwidth

Using the term for the dwell time from equation 50, the required minimum sweep time (T_{min}) can be calculated:

$$T_{min} = \text{const } F / B^2 \qquad (55)$$

The minimum sweep time increases with the square of the bandwidth reduction. If sweep time T is less than T_{min}, there is no full response from the filter. In this case, a spectral line with reduced amplitude (see amplitude loss as a function of dwell time) is displayed that becomes wider at higher frequencies. If the attack times are very short, spurious oscillations may occur.

Almost all spectrum analyzers feature mechanical coupling between the three parameters frequency span, bandwidth, and sweep time to facilitate operation, to obtain optimum operating conditions and to avoid measurement errors. With more advanced equipment, this coupling is implemented by microprocessor control. The coupling between the three parameters can be disabled for special applications but may result in uncalibrated level measurements which are identified by a message (e.g. UNCAL) appearing in the display or by illumination of a corresponding indicator light.

Frequency and level accuracy, calibration, tracking generator

Spectrum analyzers are used in applications where accuracy and time-related resolution, as well as the inherently low dynamic range of an oscilloscope, do not meet the requirements for high-precision signal analysis. For these applications, a spectrum analyzer should offer wide frequency range, high measurement accuracy, high sensitivity (see phase noise) and high intermodulation rejection (see non-linearities). The term "high measurement accuracy" is to be understood as high amplitude and frequency accuracy.

All components between the input connector and the IF amplifier of a spectrum analyzer can cause non-linear frequency response that results in amplitude errors. These components are the input attenuator, the cabling, the input filter, and the input mixer. A typical amplitude error for a 2-GHz spectrum analyzer is between ±1.5 and ±2 dB.

The high amplitude accuracy required is achieved by employing automatic calibration that is commonly used with high-quality level meters. An internally generated calibration signal with high amplitude stability, whose frequency is tuned synchronously with the sweep of the spectrum analyzer, is used as the amplitude reference. With the aid of the tracking generator, the frequency response is determined and then taken into account for calculation of the measurement results. By using a tracking generator and performing arithmetical corrections, frequency-dependent amplitude errors can be reduced to approx. ±0.5 dB.

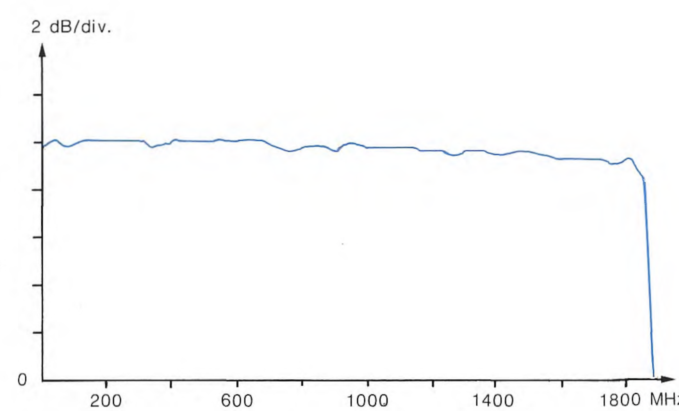


Fig. 50 Frequency response of superheterodyne frequency converter with lowpass cutoff frequency of 1.8 GHz

To achieve high frequency accuracy, a frequency synthesizer with extremely high spectral purity is used in the Spectrum Analyzer FSA to generate the tuning frequency. Frequency accuracy as well as short and long-term stability of the tuning frequency will depend only on the quality of the crystal used in the reference oscillator of the frequency synthesizer. The accuracy with which the frequency of a spectral line can be measured is, however, not only determined by the tuning frequency.

Additional errors are caused by deviations of the centre frequency of the selection filter connected into the signal path during a measurement. These errors can, for example, result from aging effects or temperature drift. By using an additional frequency synthesizer in the last frequency converter, the intermediate frequency can simply be offset by the deviation of the centre frequency.

For spectrum analyzers with synthesizer tuning, the specified frequency accuracy does not only apply to a preferred frequency, such as the centre frequency, but also to all measurement points along the entire frequency axis.

For frequency response and attenuation measurements of amplifiers, filters, etc., a tracking generator is synchronized with the local oscillator spectrum analyzer (Fig. 51). By phase-locking the tracking generator to the local oscillator, the output signal of the local oscillator is always in phase with the reference frequency or the frequency of the signal generated by the local oscillator. This is especially useful for phase measurements.

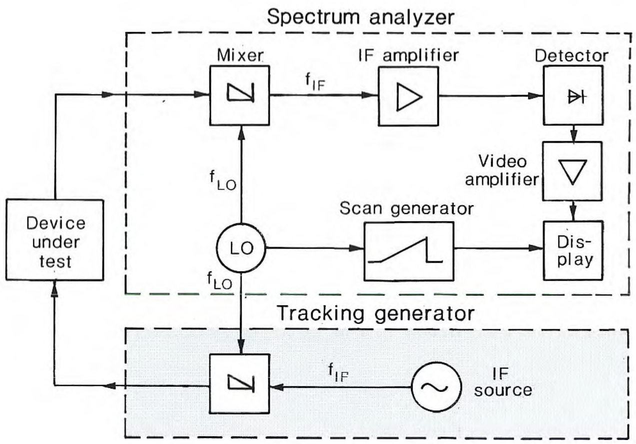


Fig. 51 Spectrum analyzer with tracking generator

The tracking generator is a signal source featuring high frequency stability. The output signal from the tracking generator is mixed with the signal supplied by the local oscillator. If this frequency is identical to the intermediate frequency of the spectrum analyzer, the frequency of the signal provided by the tracking generator is equal to the input frequency of the spectrum analyzer. This allows measurements of frequency response, gain or attenuation to be performed on two-port networks connected between the tracking generator and the spectrum analyzer.

REFERENCES

[1] Danzeisen, K.: Spectrum Analyzer FSA for 100 Hz to 1.8 GHz. News from Rohde & Schwarz (1987) No. 116, pp. 4 – 8.

[2] Danzeisen, K.: Spectrum analyzers – principles, realization; characteristics of search-tone analyzers – technical limits. Lecture held at Technical Institute Esslingen, February 1986, as part of course on high-frequency and microwave measurements.

[3] Engelson, M.: Understand Resolution for better Spectrum Analysis. Microwaves, December 1974, pp. 32 – 36.

[4] Löcherer, K. H.: RF transmission and reception. Lecture at University of Hanover, 1982.

[5] Löcherer, K. H.: Noise in electronic components. Lecture at University of Hanover, 1984.

[6] Meinke-Gundlach: Pocketbook of high-frequency measurements. Springer Verlag Berlin, 4th edition (volume 1, pp. 122 – 126, 140 – 141).

Spectrum Analyzer FSA

100 Hz to 1.8 (2) GHz

The Spectrum Analyzer FSA features not only low inherent noise, 10 dB better than comparable instruments, but it also sets new standards regarding phase noise, dynamic range, intermodulation, level and frequency accuracy. A state-of-the-art microprocessor concept ensures measurement versatility combined with ease of operation.

IEC 625 Bus
IEEE 488

- Amplitude-measurement range −145 dBm (−150 dBm typ.) to +30 dBm
- Phase noise <−114 dBc/Hz at 1 kHz from carrier
- Resolution bandwidths 6 Hz to 3 MHz quasicontinuously variable
- 9-inch colour monitor
- Optional tracking generator for 100 Hz to 1.8 GHz with built-in attenuator

Please ask for the data sheet
Spectrum Analyzer FSA

ROHDE & SCHWARZ

ROHDE & SCHWARZ GmbH & Co. KG. · D-8000 München 80 · Mühldorfstr. 15 · Tel. (089) 41 29-0 Int. + 49 89 41 29-0 · Telex 5 23 703
Printed in the Fed. Rep. of Germany · Subject to change · Data without tolerances: order of magnitude only
887 (Bv)