Agilent Vector Signal Analysis Basics

Application Note
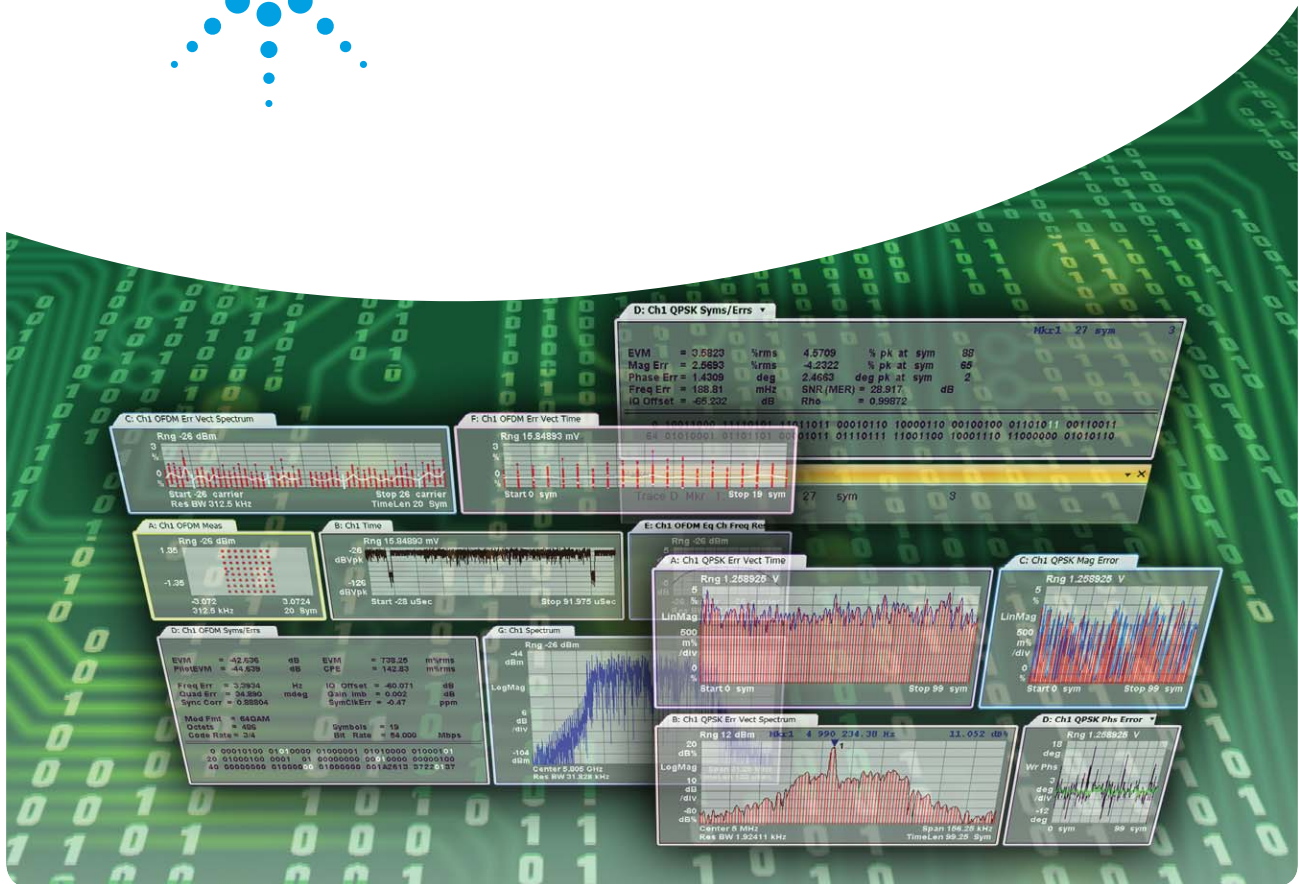

Agilent Technologies

Table of Contents

Try before you buy!

Download the 89600B software and use it free for 14 days to make measurements with your analysis hardware, or use our recorded demo signals by selecting **File> Recall> Recall Demo> QPSK>** on the software toolbar. Request your free trial license today:

www.agilent.com/find/89600B_trial

Vector Signal Analysis

This application note serves as a primer on vector signal analysis which, going forward, will be referred to as VSA. This section discusses VSA measurement concepts and theory of operation; the second section discusses vector-modulation analysis and, specifically, digital-modulation analysis.

Analog, swept-tuned spectrum analyzers use superheterodyne technology to cover wide frequency ranges; from audio, through microwave, to millimeter frequencies. Fast Fourier transform (FFT) analyzers use digital signal processing (DSP) to provide high-resolution spectrum and network analysis. Today's wide-bandwidth, vector-modulated (also called complex or digitally modulated), time-varying signals benefit greatly from the capabilities of FFT analysis and other DSP techniques. VSA provides fast, high-resolution spectrum measurements, demodulation, and advanced time-domain analysis. VSA is especially useful for characterizing complex signals such as burst, transient, or modulated signals used in communications, video, broadcast, radar, and software-defined radio applications.

Figure 1 shows a simplified VSA block diagram. VSA implements a very different measurement approach than traditional swept analysis; the analog IF section is replaced by a digital IF section incorporating FFT and digital signal processing algorithms. Traditional swept-tuned spectrum analysis is an analog system; VSA is fundamentally a digital system that uses digital data and mathematical algorithms to perform data analysis. VSA software accepts and analyzes digitized data from many measurement front ends, allowing you to troubleshoot throughout the system block diagram.
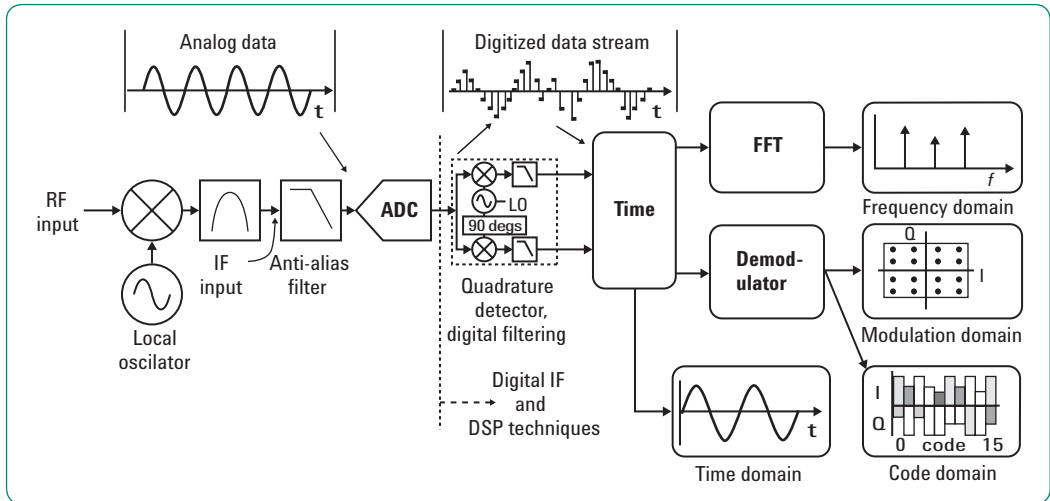

Figure 1. The vector signal analysis process requires a digitized analog input signal and then uses DSP technology process and provide data outputs; the FFT algorithm produces frequency domain results, the demodulation algorithms produce modulation and code domain results.

A significant characteristic of VSA is that it can measure and manipulate complex data, i.e. magnitude and phase information. In fact, it is called vector signal analysis because it takes complex input data, performs complex data analysis, and outputs complex data results that include both magnitude and phase information. Vector modulation analysis performs the basic functionality of a measurement receiver. You will learn about vector modulation and detection in the next section, Vector Modulation Analysis.

With the proper front end, VSA covers RF and microwave ranges, plus it provides additional modulation-domain analysis capability. These advancements are made possible through digital technologies such as analog-to-digital conversion and DSP that include digital intermediate frequency (IF) techniques and fast Fourier transform (FFT) analysis.

Because the signals that people must analyze are growing in complexity, the latest generations of signal analyzers have moved to a digital architecture and often include many vector signal analysis and modulation analysis capabilities. Some analyzers digitize the signal at the instrument input, after some amplification, or after one or more downconverter stages. In most of today's analyzers, phase, as well as magnitude, is preserved in order to perform true vector measurements. Other front ends, such as oscilloscopes and logic analyzers, digitize the entire signal, while also maintaining phase and magnitude information. VSA capability depends on the processing capability available to any of the front ends, either as an integrated measurement personality or as software running internally or on a computer connected to the front end.

VSA measurement advantages

Vector analysis measures dynamic signals and produces complex data results

VSA offers some distinct advantages over analog swept-tuned analysis. One major VSA advantage is its ability to better measure dynamic signals. Dynamic signals generally fall into one of two categories: time-varying or complex modulated. Time-varying are signals whose measured properties change during a measurement sweep (such as burst, gated, pulsed, or transient). Complex-modulated signals cannot be solely described in terms of simple AM, FM, or PM modulation, and include most of those used in digital communications, such as quadrature amplitude modulation (QAM).
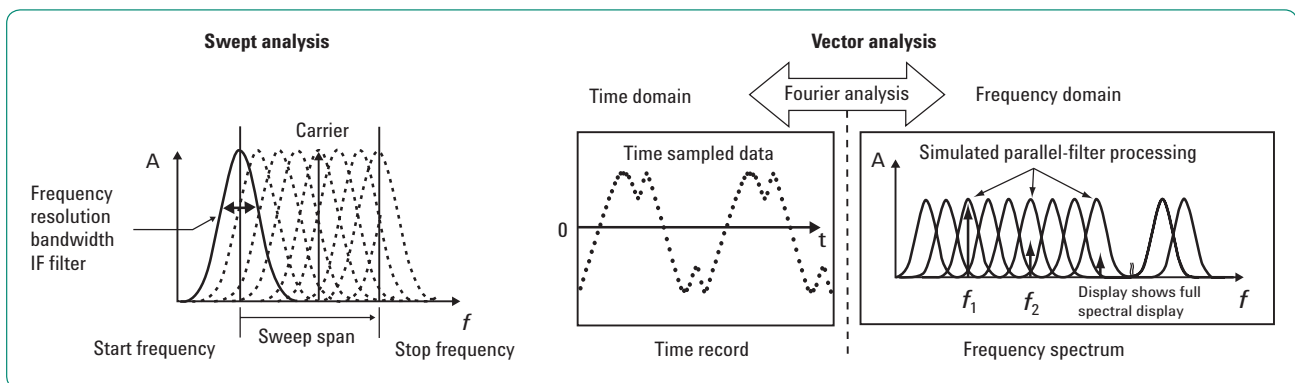

Figure 2. Swept-tuned analysis displays the instantaneous time response of a narrowband IF filter to the input signal. Vector analysis uses FFT analysis to transform a set of time domain samples into frequency domain spectra.

Traditional swept-spectrum analysis[1], in effect, sweeps a narrowband filter across a range of frequencies, sequentially measuring one frequency at a time. Unfortunately, sweeping the input works well for stable or repetitive signals, but will not accurately represent signals that change during the sweep. Also, this technique only provides scalar (magnitude only) information, though some other signal characteristics can be derived by further analysis of spectrum measurements.

The VSA measurement process simulates a parallel bank of filters and overcomes swept limitations by taking a "snapshot," or time-record, of the signal; then processing all frequencies simultaneously. For example, if the input is a transient signal, the entire signal event is captured (meaning all aspects of the signal at that moment in time are digitized and captured); then used by the FFT to compute the "instantaneous" complex spectra versus frequency. This process can be performed in real-time, that is, without missing any part of the input signal. For these reasons, VSA is sometimes referred to as "dynamic signal analysis" or "real-time signal analysis. The ability to track a fast-changing signal with VSA, however, is not unlimited. It depends on the computation capability available.

The VSA decreases measurement time
Parallel processing yields another potential advantage for high-resolution (narrow resolution bandwidth) measurements: faster measurement time. If you've used a swept-tuned spectrum analyzer before, you already know that narrow resolution bandwidth (RBW) measurements of small frequency spans can be very time-consuming. Swept-tuned analyzers sweep frequencies from point to point slowly enough to allow the analog resolution bandwidth filters to settle. By contrast, VSA measures across the entire frequency span at one time. However, there is analogous VSA settling time due to the digital filters and DSP. This means that the VSA sweep speed is limited by data collection and digital processing time rather than analog filters. But this time is usually negligible when compared to the settling time of analog filters. For certain narrow bandwidth measurements, VSA can complete a measurement up to 1000 times faster than conventional swept-tuned analysis.

In swept-tuned spectrum analysis, the physical bandwidth of the sweeping filter limits the frequency resolution. VSA doesn't have this limitation. VSA can resolve signals that are spaced less than 100 µHz apart. Typically, VSA resolution is limited by the signal and measurement front end's frequency stability, as well as by the amount of time you are willing to devote to the measurement. Increasing the resolution also increases the time it takes to measure the signal (the required time-record length).

Time-capture is a great tool for signal analysis and troubleshooting
Another feature that is extremely useful is the time-capture capability. This allows you to record actual signals in their entirety without gaps, and replay them later for any type of data analysis. All measurement features can be applied to the captured signal. For example, you could capture a transmitted digital communications signal and then perform both spectrum and vector-modulation analysis to measure signal quality or identify signal impairments.

---

1. For more information on spectrum analysis, see Agilent Application Note 150, Spectrum Analysis Basics, literature number 5952-0292EN.

DSP provides multi-domain measurements in one instrument

The use of digital signal processing (DSP) also yields additional benefits; it provides time, frequency, modulation, and code domain measurement analysis simultaneously. Having these capabilities increases the value to you and improves the quality of your measurements. VSA uses FFT analysis to provide easy and accurate views of both time and frequency domain data. The DSP provides vector modulation analysis, including both analog and digital modulation analysis. The analog demodulation algorithm produces AM, FM and PM demodulation results, similar to that of a modulation analyzer, allowing you to view amplitude, frequency, and phase profiles versus time. The digital demodulation algorithm performs a broad range of measurements on many digital communications standards (such as GSM, cdma2000®, WiMAX™, LTE, and more) and produces many useful measurement displays and signal-quality data.

VSA clearly provides important benefits and, when paired with the appropriate front end, can realize even greater benefits. For example, combine it with conventional analog swept-tuned analyzers to provide even higher frequency coverage and increased dynamic range capability. Used with oscilloscopes, VSA provides wide bandwidth analysis, while combining it with logic analyzers allows you to probe FPGAs and other digital baseband blocks of your radio.

VSA measurement concepts and theory of operation

As mentioned earlier, VSA is fundamentally a digital system that uses DSP to perform spectrum analysis with FFTs, and uses demodulation algorithms to perform vector-modulation analysis. The FFT is a mathematical algorithm that operates on time-sampled data and provides time-to-frequency domain transformations. The analog signal must be digitized in the time-domain, then the FFT algorithm executes to compute the spectra. Conceptually, VSA implementation is simple and straightforward: acquire a digitized input signal, then compute the measurement results. See Figure 3. However, in practice, there are many factors that must be accounted for in order for the measurement to be meaningful and accurate. (For more information about FFT analysis, refer to the **Additional Resources** section at the end of this application note.)
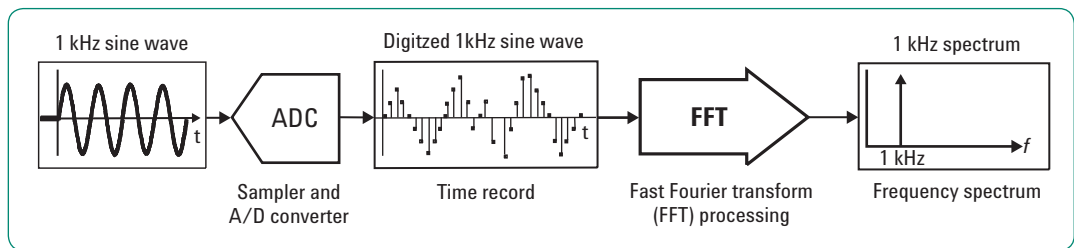

Figure 3. 1 kHz FFT analysis example: digitize time-domain signal and use FFT analysis to convert it to the frequency domain

6

If you are familiar with FFT analysis, you already know the FFT algorithm makes several assumptions about the signal it is processing. The algorithm doesn't check to verify the validity of these assumptions for a given input, and it will produce invalid results, unless you or the instrument validate the assumptions.

Figure 1 illustrates a general VSA system block diagram. Through DSP, many functions could occur at different places. Figure 4 shows a generalized diagram of the technique that Agilent uses. The VSA measurement process includes these fundamental stages:

Measurement front-end
1. Signal conditioning with frequency translation. Depending on the front-end hardware used, different signal conditioning steps may be needed and/or available.
2. Analog-to-digital conversion
3. Quadrature detection
VSA software
4. Digital filtering and resampling
5. Data windowing
6. FFT analysis (for vector modulation, blocks 5 and 6 are replaced with the demodulator block)

The first stage of the measurement process is called signal conditioning. This stage includes several important functions that condition and optimize the signal for the analog-to-digital conversion and FFT analysis. The first function is AC and DC coupling. This option is necessary if you need to remove unwanted DC biases in the measurement setup. Next, the signal is either amplified or attenuated for optimal signal level into the mixer. The mixer stage provides frequency translation, or RF-to-IF downconversion, and mixes the signal down to the final IF. This operation is the same as the superheterodyne function of swept-tuned analysis and extends FFT analysis capabilities through microwave. In practice, it may take several downconversion stages to reach the final IF frequency. Some signal analyzers provide external IF input capability; by providing your own IF, you can extend the VSA's upper frequency range to match a receiver you provide.
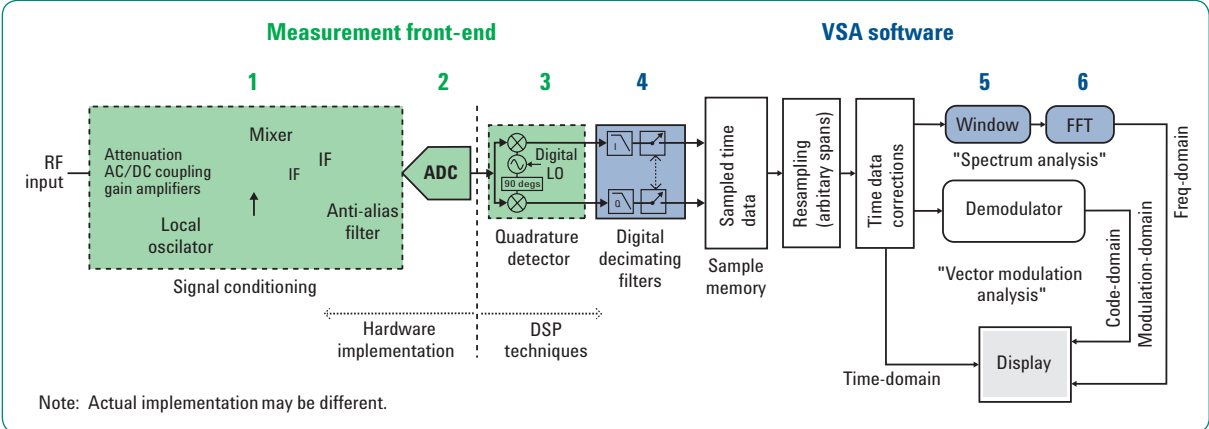

Figure 4. Simplified block diagram showing RF hardware front end and vector signal analysis software.

The final stage of the signal conditioning process is extremely important to a sampled system and FFT analysis: signal alias protection. Anti-alias filtering performs this function. If the VSA measurement does not have adequate protection from aliasing, it may show frequency components that are not part of the original signal. The sampling theorem states that if the signal is sampled at a rate greater than 2 times the highest significant frequency component present in the signal, the sampled signal can be reconstructed exactly. The minimum acceptable sample rate is called the Nyquist rate. Thus,

$$f_s > 2 \, (f_{max})$$

where f_s = sample rate
       f_{max} = highest frequency component

If the sampling theorem is violated, "aliasing" error products can result. Therefore, to prevent alias products for a given maximum frequency, there must not be significant signal energy above 1/2 the sample rate. Figure 5 shows a set of sample points, which fit two different waveforms. The higher-frequency waveform violates the sampling theorem. Unless an anti-alias filter is used, the two frequencies will be indistinguishable when processed digitally.

To prevent aliasing, two conditions must be satisfied:
1. The input signal to the digitizer/sampler must be band limited. In other words, there must exist a maximum frequency (f_{max}) above which no frequency components are present.
2. The input signal must be sampled at a rate that satisfies the sampling theorem.

The solution to the aliasing problem seems simple enough. First you select the maximum frequency (f_{max}) that the front end hardware will measure, then make sure the sampling frequency (f_s) is twice that frequency. This step satisfies condition number 2 and makes sure that the VSA software can accurately analyze the frequencies of interest. Next you insert a low-pass filter (an anti-aliasing filter) to remove all frequencies above f_{max}, thus ensuring that the measurement will exclude all frequencies except those you are interested in. This step satisfies condition number 1 and makes sure the signal is band limited.
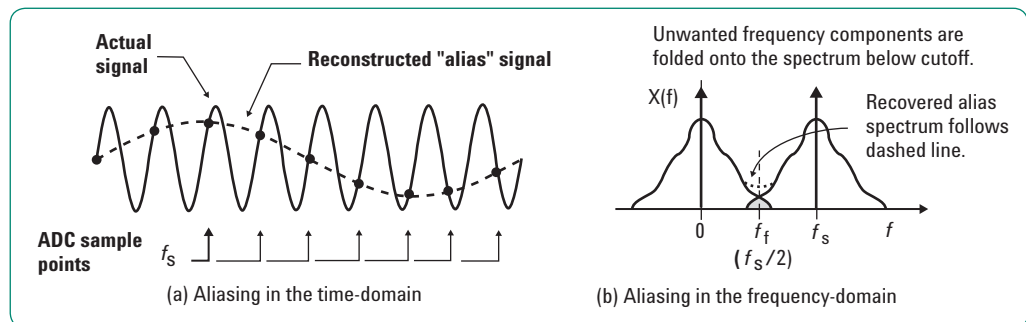

(a) Aliasing in the time-domain   (b) Aliasing in the frequency-domain

Figure 5. Aliasing products occur when the signal is undersampled. Undesirable frequencies appear under the alias of another (baseband) frequency

There are two factors that complicate this simple anti-aliasing solution. The first, and easiest to address, is that the anti-alias filter has a finite roll off rate. As shown in Figure 6, there is a transition band in practical filters between the passband and stopband. Frequencies within the transition band could produce alias frequencies. To avoid these alias products, the filter cutoff must be below the theoretical upper frequency limit of f_s divided by 2. An easy solution to this problem is to oversample (sample above the Nyquist rate). Make the sampling frequency slightly above 2 times f_{max} so that it is twice the frequency at which the stopband actually starts, not twice the frequency you are trying to measure. Many VSA implementations use a guard band to protect against displaying aliased frequency components. The FFT computes the spectral components out to 50% of f_s (equivalently $f_s/2$). A guard band, between approximately 40% to 50% of f_s (or $f_s/2.56$ to $f_s/2$), is not displayed because it may be corrupted by alias components. However, when the VSA software computes the inverse FFT, the signals in the guard band are used to provide the most accurate time-domain results. The high-roll-off-rate filter, combined with the guard band, suppresses potential aliasing components, attenuating them well below the measurement front end's noise floor.

The second complicating factor in alias protection (limited frequency resolution) is much harder to solve. First, an anti-alias filter that is designed for wide frequency spans (high sample rates) is not practical for measuring small resolution bandwidths for two reasons; it will require a substantial sample size (memory allocation) and a prohibitively large number of FFT computations (long measurement times). For example, at a 10 MHz sample rate, a 10 Hz resolution bandwidth measurement would require more than a 1 million point FFT, which translates into large memory usage and a long measurement time. This is unac-ceptable because the ability to measure small resolution bandwidths is one main advantage of VSA.

One way of increasing the frequency resolution is by reducing fs, but this is at the expense of reducing the upper-frequency limit of the FFT and ultimately the analyzer bandwidth. However, this is still a good approach because it allows you to have control over the measurement resolution and frequency range. As the sample rate is lowered, the cut-off frequency of the anti-alias filter must also be lowered, otherwise aliasing will occur. One possible solution would be to provide an anti-aliasing filter for every span, or a filter with selectable cutoff frequencies. Implementing this scheme with analog filters would be difficult and cost prohibitive, but it is possible to add additional anti-alias filters digitally through DSP.
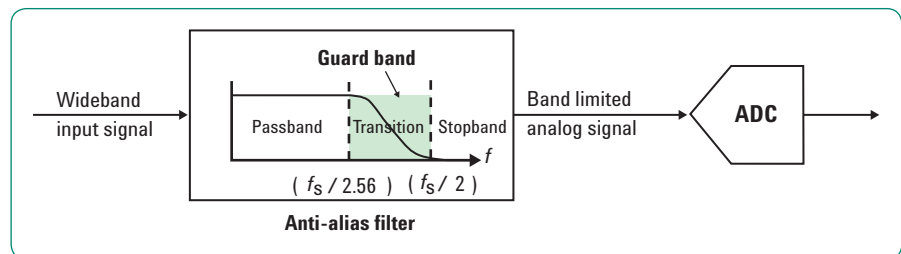

Figure 6. The anti-alias filter attenuates signals above $f_s/2$. A guard band between 40% to 50% of f_s is not displayed

9

Digital decimating filters and **resampling** algorithms provide the solution to the limited frequency resolution problem. This is the solution used in Agilent VSA software. Digital decimating filters and resampling perform the operations necessary to allow variable spans and resolution bandwidths. The digital decimating filters simultaneously decrease the sample rate and limit the bandwidth of the signal (providing alias protection). The sample rate into the digital filter is fs; the sample rate out of the filter is f_s/n, where "n" is the decimation factor and an integer value. Similarly, the bandwidth at the input filter is "BW," and the bandwidth at the output of the filter is "BW/n". Many implementations perform binary decimation (divide-by-2 sample rate reduction), which means that the sample rate is changed by integer powers of 2, in $1/(2^n)$ steps (1/2, 1/4, 1/8, ... etc.). Frequency spans that result from "divide by 2^n" are called cardinal spans. Measurements performed at cardinal spans are typically faster than measurements performed at arbitrary spans due to reduced DSP operations.

The decimating filters allow the sample rate and span to be changed by powers of two. To obtain an arbitrary span, the sample rate must be made infinitely adjustable. This is done by a resampling or interpolation filter, which follows the decimation filters. For more details regarding resampling and interpolation algorithms, refer to the **Additional Resources** section at the end of this application note.

Even though the digital and resampling filters provide alias protection while reducing the sample rate, the analog anti-alias filter is still required, since the digital and resampling filters are, themselves, a sampled system which must be protected from aliasing. The analog anti-alias filter protects the analysis at its widest frequency span with operation at f_s. The digital filters follow the analog filter and provide anti-alias protection for the narrower, user-specified spans.

Additional precautions must be taken, with regard to band-limit signals and to protect against anti-aliasing when using oscilloscopes as a front-end to VSA software. See the **Additional Resources** section at the end of this document for more information.

The next complication that limits the ability to analyze small resolution bandwidths is caused by a fundamental property of the FFT algorithm itself; the FFT is inherently a baseband transform. This means that the FFT frequency range starts from 0 Hz (or DC) and extends to some maximum frequency, f_s divided by 2. This can be a significant limitation in measurement situations where a small frequency band needs to be analyzed. For example, if the measurement front end has a sample rate of 10 MHz, the frequency range would be 0 Hz to 5 MHz ($f_s/2$). If the number of time samples (N) were 1024, the frequency resolution would be 9.8 kHz (f_s/N). This means that frequencies closer than 9.8 kHz could not be resolved.

As just mentioned, you can control the frequency span by varying the sample rate, but the resolution is still limited because the start frequency of the span is DC. The frequency resolution can be arbitrarily improved, but at the expense of a reduced maximum frequency. The solution to these limitations is a process called band selectable analysis, also known as zoom operation or zoom mode. Zoom operation allows you to reduce the frequency span while maintaining a constant center frequency. This is very useful because it allows you to analyze and view small frequency components away from 0 Hz. Zooming allows you to focus the measurement anywhere within the measurement front end's frequency range (Figure 7).

Zoom operation is a process of digital quadrature mixing, digital filtering, and decimating/resampling. The frequency span of interest is mixed with a complex sinusoid at the zoom span center frequency (f_z), which causes that frequency span to be mixed down to baseband. The signals are filtered and decimated/resampled for the specified span, and all out-of-band frequencies removed. This is the band-converted signal at IF (or baseband) and is sometimes referred to as "zoom time" or "IF time." That is, it is the time domain representation of a signal as it would appear in the IF section of a receiver. Zoom measurements are discussed further in the **Time-domain display** section near the end of this chapter.
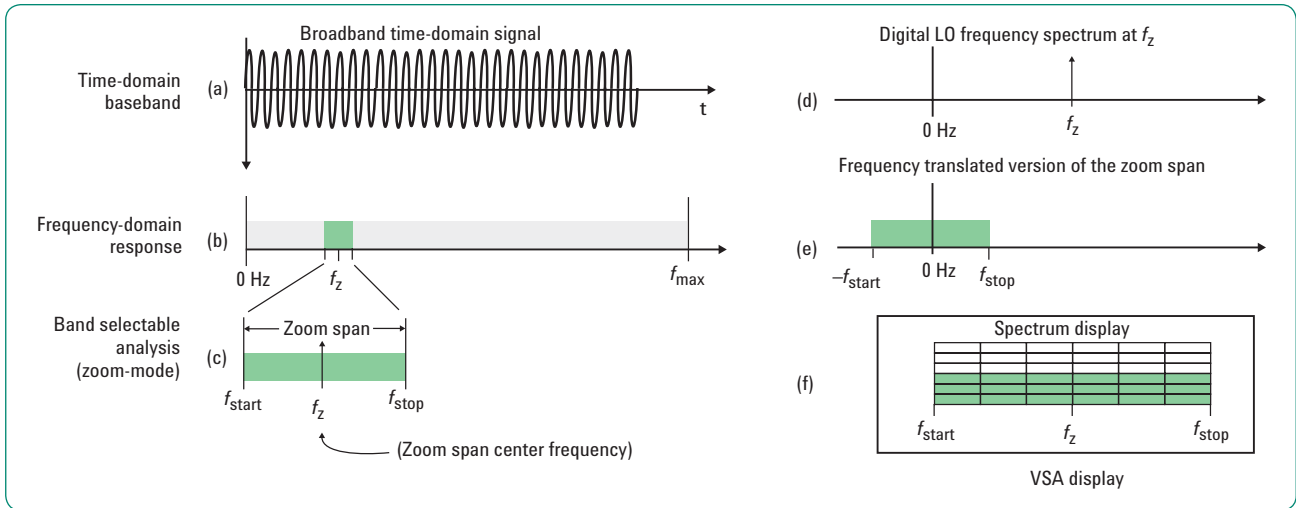

Figure 7. Band-selectable analysis (or zoom mode): (a) measured broadband signal, (b) spectrum of the measured signal, (c) selected zoom span and center frequency, (d) digital LO spectrum (@ zoom center frequency), (e) frequency span mixed down to baseband, (f) display spectrum annotation is adjusted to show the correct span and center frequency

Sample memory

The output of the digital decimating filters represents a bandlimited, digital version of the analog input signal in time-domain. This digital data stream is captured in sample memory (Figure 4). The sample memory is a circular FIFO (first in, first out) buffer that collects individual data samples into blocks of data called time records, to be used by the DSP for further data processing. The amount of time required to fill the time record is analogous to the initial settling time in parallel-filter analysis. The time data collected in sample memory is the fundamental data used to produce all measurement results, whether in the frequency domain, time domain, or modulation domain.

Time domain data corrections

To provide more accurate data results, VSA software may implement time data correction capability through an equalization filter. In vector analysis, the time data's accuracy is very important. Not only is it the basis for all of the demodulation measurements, but it is also used directly for measurements such as instantaneous power as a function of time. Correcting the time data is the last step in creating a nearly ideal bandlimiting signal path. While the digital filters and resampling algorithms provide for arbitrary bandwidths (sample rates and spans), the time-domain corrections determine the final passband characteristic of the signal path. Time-domain corrections would be unnecessary if the analog and digital signal paths could be made ideal. Time-domain corrections function as an equalization filter to compensate for passband imperfections. These imperfections come from many sources. The IF filters in the RF section, the

analog anti-aliasing filter, the decimating filters, and the resampling filters all contribute to passband ripple and phase nonlinearities within the selected span.

The design of the equalization filter begins by extracting information about the analog signal path from the self-calibration data based on the measurement front end's configuration. This is the data used to produce the frequency-domain correction output display. Once the analog correction vector has been computed, it is modified to include the effects of the decimating and resampling filters. The final frequency response computations cannot be performed until after you have selected the span, because that determines the number of decimating filter stages and resampling ratio. The composite correction vector serves as the basis for the design of the digital equalization filter that is applied to the time data.

Data windowing – leakage and resolution bandwidth

The FFT assumes that the signal it is processing is periodic from time record to time record. However most signals are not periodic in the time record and a discontinuity between time records will occur. Therefore, this FFT assumption is not valid for most measurements, so it must be assumed that a discontinuity exists. If the signal is not periodic in the time record, the FFT will not estimate the frequency components accurately. The resultant effect is called leakage and has the affect of spreading the energy from a single frequency over a broad range of frequencies. Analog swept-tuned signal analysis will produce similar amplitude and spreading errors when the sweep speed is too fast for the filter's bandwidth.

Data windowing is the usual solution to the leakage problem. The FFT is not the cause of the error; the FFT is generating an "exact" spectrum for the signal in the time record. It is the non-periodic signal characteristics between time records that cause the error. Data windowing uses a window function to modify the time-domain data by forcing it to become periodic in the time record. In effect, it forces the waveform to zero at both ends of the time record. This is accomplished by multiplying the time record by a weighted window function. Windowing distorts the data in the time domain to improve accuracy in the frequency domain. See Figure 8.
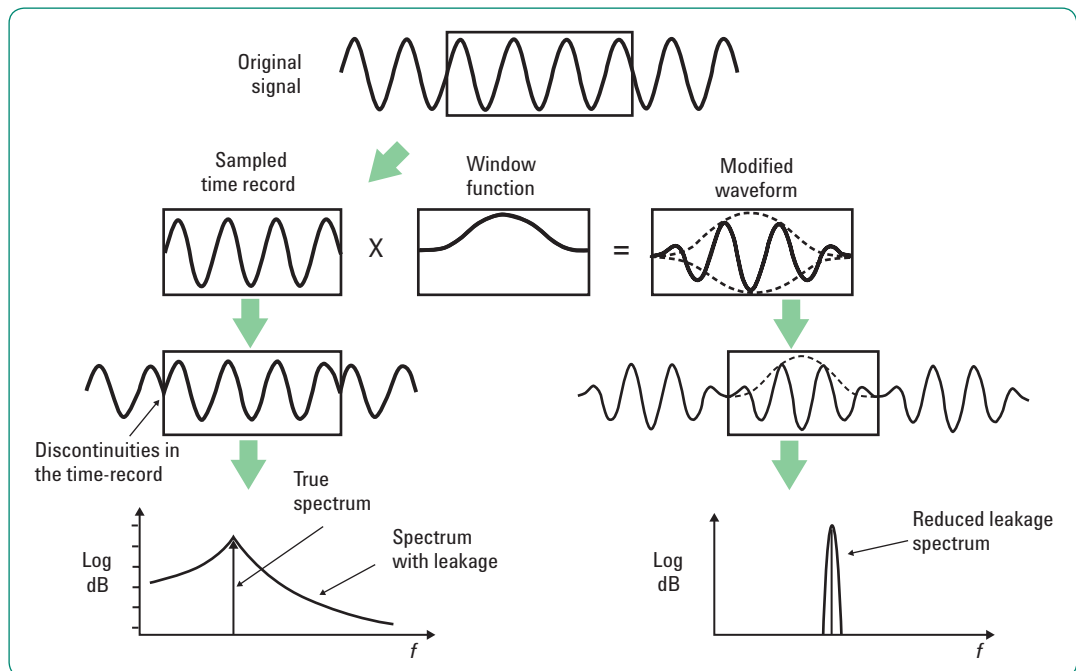

Figure 8. Window functions reduce the leakage errors in the frequency domain by modifying the time domain waveform

12

The Agilent 89600B VSA automatically selects the appropriate window filter based on assumptions of the user's priorities, derived from the selected measurement type. However, if you want to manually change the window type, you can select from several built-in window types. Each window function, and the associated RBW filter shape, offers particular advantages and disadvantages. A particular window type may trade off improved amplitude accuracy and reduced "leakage" at the cost of reduced frequency resolution. Because each window type produces different measurement results (just how different depends on the characteristics of the input signal and how you trigger on it), you should carefully select a window type appropriate for the measurement you are trying to make. Table 1 summarizes four common window types and their uses.

Table 1. Common window types and uses

Window	Common uses
Uniform (rectangular, boxcar)	Transient and self-windowing data
Hanning	General purpose
Gaussian top	High dynamic range
Flat top	High amplitude accuracy

The window filter contributes to the resolution bandwidth
In traditional swept-tuned analysis, the final IF filter determines the resolution bandwidth. In FFT analysis, the window type determines the resolution bandwidth filter shape. And the window type, along with the time-record length, determines the width of the resolution bandwidth filter. Therefore, for a given window type, a change in resolution bandwidth will directly affect the time-record length. Conversely, any change to time-record length will cause a change in resolution bandwidth as shown in the following formula:

$$RBW = \text{normalized ENBW}/T$$

where
$$ENBW = \text{equivalent noise bandwidth}$$
$$RBW = \text{resolution bandwidth}$$
$$T = \text{time-record length}$$

Equivalent noise bandwidth (ENBW) is a figure of merit that compares the window filter to an ideal, rectangular filter. It is the equivalent bandwidth of a rectangular filter that passes the same amount (power) of white noise as the window. Table 1-2 lists the normalized ENBW values for several window types. To compute the ENBW, divide the normalized ENBW by the time-record length. For example, a Hanning window with a 0.5 second time-record length would have an ENBW of 3 Hz (1.5 Hz-s/0.5 s).

Table 2. Normalized ENBW values

Window type	Normalized ENBW
Flat Top	3.819 Hz-s
Gaussian top	2.215 Hz-s
Hanning	1.500 Hz-s
Uniform	1.000 Hz-s

Fast Fourier transform (FFT) analysis

The signal is now ready for applying the FFT algorithm. The FFT is a record-oriented algorithm and operates on sampled data in a special way. Rather than acting on each data sample as the ADC converts it, the FFT waits until a number of samples (N) have been obtained (called a time record), then transforms the complete block. See Figure 9. In other words, a time record N samples long, is the input to the FFT, and the frequency spectrum N samples long, is the output.

The FFT's speed comes from the symmetry or repeated sample values that fall out of restricting N to powers of 2. A typical record length for FFT analysis is 1024 (2^{10}) sample points. The frequency spectrum produced by the FFT is symmetrical about the sample frequency $f_s/2$ (this value is called the folding frequency, f_f). Thus, the first half of the output record contains redundant information, so only the second half is retained, sample points 0 thru N/2. This implies that the effective length of the output record is (N/2) + 1. You must add 1 to N/2 because the FFT includes the zero line, producing outputs from 0 Hz thru N/2 Hz inclusive. *These are complex data points that contain both magnitude and phase information.*

In theory, the output of the FFT algorithm is (N/2) +1 frequency points, extending from 0 Hz to f_f. In practice however, a guard band is used for alias protection, so not all of these points are normally displayed. As mentioned earlier, a guard band (between approximately 40% to 50% of f_s) is not displayed because it may be corrupted by alias components. For example, for a record length of 2048 samples, which produces 1025 unique complex frequency points, only 801 may actually be displayed.
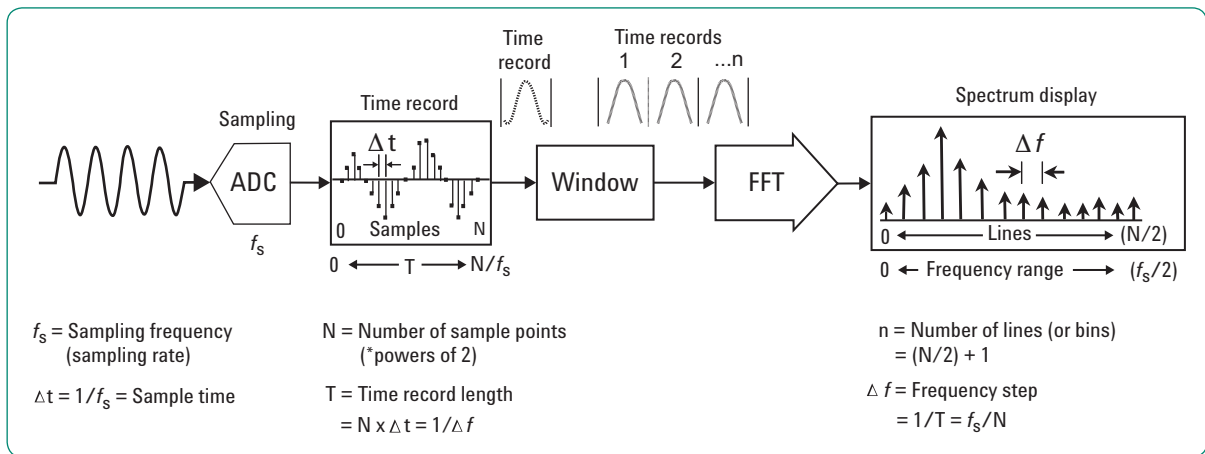

Figure 9. Basic FFT relationships

These frequency domain points are called lines, or bins, and are usually numbered from 0 to N/2. These bins are equivalent to the individual filter/detector outputs in bank-of-filters analysis. Bin 0 contains the DC level present in the input signal and is referred to as the DC bin. The bins are equally spaced in frequency, with the frequency step (Δf) being the reciprocal of the measurement time-record length (T). Thus, $\Delta f = 1/T$. The length of the time record (T) can be determined from the sample rate (f_s) and the number of sample points (N) in the time record as follows: $T = N/f_s$. The frequency (f_n) associated with each bin is given by:

$$f_n = nf_s/N$$
where   n is the bin number

The last bin contains the highest frequency, $f_s/2$. Therefore, the frequency range of an FFT is 0 Hz to $f_s/2$. Note that the highest FFT range is not f_{max}, which is the FFT algorithm's, upper-frequency limit, and may not be the same as the highest bin frequency.

Real-time bandwidth

Because FFT analysis cannot compute a valid frequency-domain result until at least one time record is acquired, the time-record length determines how long an initial measurement will take. For example, a 400-line measurement using a 1 kHz span requires a 400 ms time record; a 3200-line measurement requires a 3.2 s time record. This amount of data acquisition time is independent of the FFT's compute engine's processing speed.

After the time record has been captured, processing speed does become an issue. The amount of time it takes to compute the FFT, format, and display the data results, determines the processing speed and display update rate. Processing speed can be important for two reasons. First, higher processing speeds can translate to decreased overall measurement time. Second, the ability to measure dynamic signals is a function of the processing speed. The performance indicator is the real-time bandwidth (RTBW), which is the maximum frequency span that can be continuously processed without missing any event in the input signal.
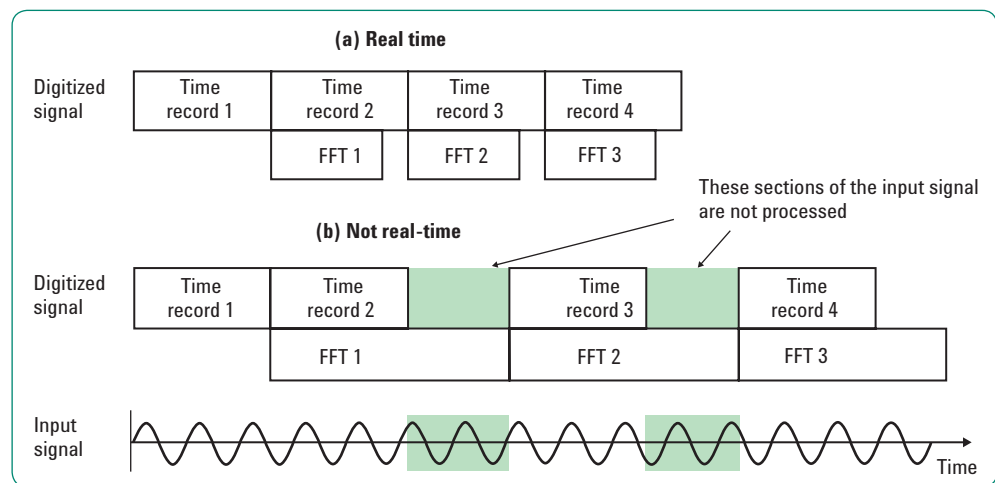

Figure 10. (a) Processing is "real-time" when the FFT processing time is ≤ the time-record length; no data is lost. (b) Input data is missed if the FFT processing time is greater than the time-record length.

RTBW is the frequency span at which the FFT processing time equals the time-record length. There is no gap between the end of one time record and the start of the next. See Figure 10. If you increase the span past the real-time bandwidth, the record length becomes shorter than the FFT processing time, the time records are no longer contiguous, and some data will be missed. This is the usual case for RF measurements. Note, however, that time-captured data is real time, as all time samples flow directly to the available memory, with no lost data.

Time-domain display

VSA lets you view and analyze time-domain data. The displayed time-domain data may look similar to an oscilloscope display, but you need to be aware that the data you're viewing may be quite different. The time-domain display shows the time-data just before FFT processing. See Figure 4. VSA can provide two measurement modes, baseband and zoom. Depending on the measurement mode, the time-domain data you are viewing will be very different.

Baseband mode provides time data similar to what you would view on a digital oscilloscope. Like the traditional digital signal oscilloscope (DSO), VSA provides real-valued time data referenced to 0 time and 0 Hz (DC). However, the trace may appear distorted with VSA, especially at high frequencies. This is because VSA occurs at a sample rate chosen to optimize FFT analysis, which, at the highest frequencies may only be 2 or 3 samples per period; great for the FFT, but not so good for viewing. In contrast, the DSO is optimized for time-domain analysis and usually oversamples the input. In addition, a DSO may provide additional signal reconstruction processing that enable the DSO to display a better time-domain representation of the actual input signal. Also, at maximum span, some signals (particularly square waves and transients) may appear to have excess distortion or ringing because of the abrupt frequency cut-off of the anti-alias filter. In this sense, DSOs are optimized for sample rate and time-domain viewing, not power accuracy and dynamic range.

In **zoom** (or **band selectable**) mode, you are viewing the time waveform after it has been mixed and quadrature detected. Specifically, the time data you are viewing is the product of many steps, including, potentially, analog down conversion, IF filtering, digital quadrature mixing, and digital filtering/resampling, based on the specified center frequency and span. The result is a band-limited complex waveform that contains real and imaginary components and, in most cases, it looks different from what you would see on an oscilloscope display. This may be valuable information, depending on the intended use. For example, this could be interpreted as "IF time," the time-domain signal that would be measured with an oscilloscope probing the signal in the IF section of a receiver.

The digital LO and quadrature detection algorithms perform the **zoom** measurement function. In zoomed measurements, the selected frequency span is mixed down to baseband at the specified center frequency (f_{center}). To accomplish this, first the digital LO frequency is assigned the f_{center} value. Then the input signal is quadrature detected; it is multiplied or mixed with the sine and cosine (quadrature) of the center frequency of the measurement span. The result is a complex (real and imaginary) time-domain waveform that is now referenced to f_{center}, while the phase is still relative to the zero time trigger. Remember, the products of the mixing process are the sum and difference frequencies (signal $-f_{center}$ and signal $+f_{center}$). So the data is further processed by the low-pass filters to select only the difference frequencies. If the carrier frequency ($f_{carrier}$) is equal to f_{center}, the modulation results are the positive and negative frequency sidebands centered about 0 Hz. However, the spectrum displays are annotated to show the correct center frequency and sideband frequency values.

Figure 11 shows a 13.5 MHz sinewave measured in both baseband and zoom mode. The span for both measurements is 36 MHz and both start at 0 Hz. The number of frequency points is set to 401. The left-hand time trace shows a sinewave at its true period of approximately 74 ns (1/13.5 MHz). The right-hand time trace shows a sinewave with a period of 222.2 ns (1/4.5 MHz). The 4.5 MHz sinewave is the difference between the 18 MHz VSA center frequency of the VSA algorithm and the 13.5 MHz input signal.
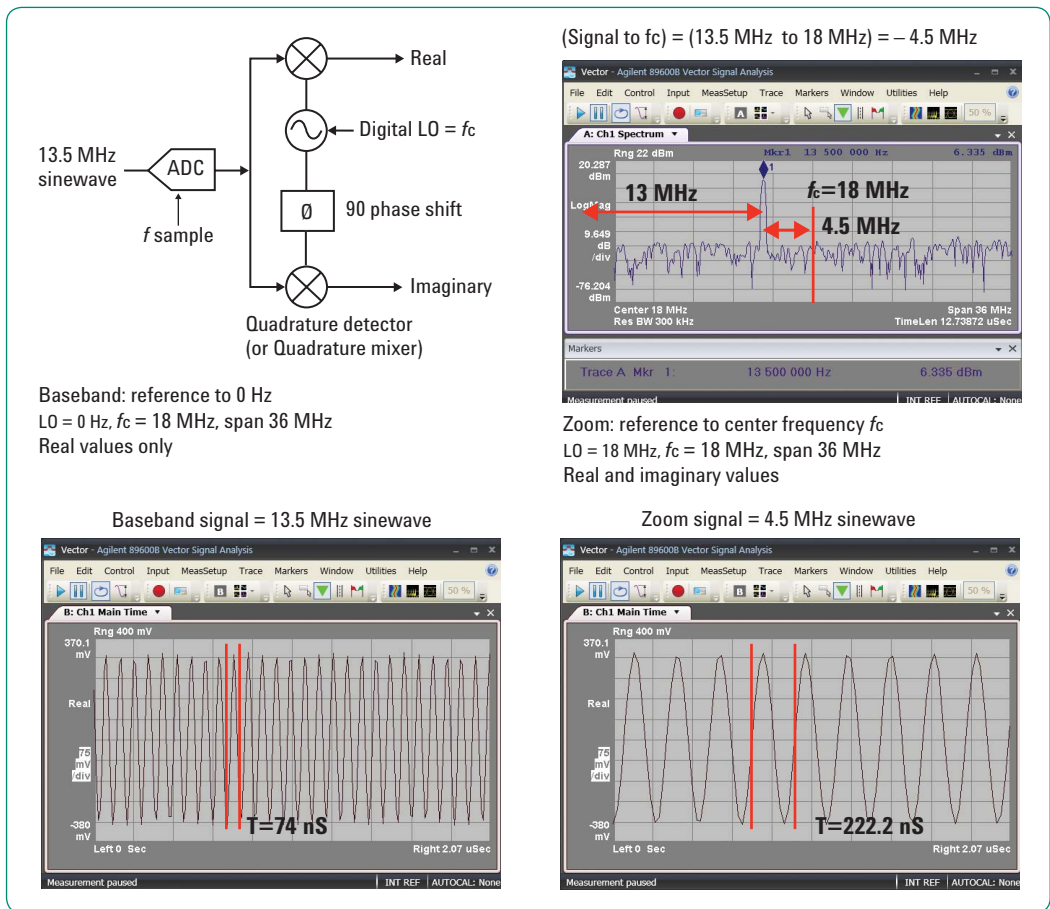

Figure 11. Baseband and zoom time data

17

Summary

This section presented a primer on the theory of operation and measurement concepts for vector signal analysis (VSA). We went though the system block diagram and described each function as it related to the FFT measurement process. You learned that the implementation used by VSA is quite different from conventional analog, swept-tuned signal analysis. VSA is primarily a digital system incorporating an all-digital IF, DSP, and FFT analysis. You learned that VSA is a test and measurement solution providing time-domain, frequency-domain, modulation-domain and code-domain signal analysis capabilities.

This chapter described the spectrum analysis capabilities of VSA, implemented using FFT analysis. The fundamentals of FFT measurement theory and analysis were presented. The vector analysis measurement concepts and demodulation block, which include digital and analog modulation analysis, are described in the section on Vector Modulation Analysis.

Vector Modulation Analysis

Introduction

The first section of this application note was essentially a primer on vector signal analysis (VSA) and discussed VSA measurement concepts and theory of operation. It also described the frequency-domain, spectrum analysis measurement capability of VSA, implemented through fast Fourier transform (FFT) analysis. This chapter describes the vector-modulation analysis and digital-modulation analysis measurement capability of VSA. Some swept-tuned spectrum analyzers can also provide digital-modulation analysis by incorporating additional digital radio personality software. However, VSA typically provides more measurement flexibility in terms of modulation formats and demodulation algorithm configuration, and provides a larger set of data results and display traces. The basic digital-modulation analysis concepts described in this chapter can also apply to swept-tuned analyzers that have the additional digital-modulation analysis software.

The real power of VSA is its ability to measure and analyze vector modulated and digitally modulated signals. Vector-modulation analysis refers to measuring measure complex signals, signals that have a real and imaginary component.

An important measurement tool that vector-modulation analysis provides is analog-modulation analysis. For example, the Agilent 89600B VSA software provides analog-modulation analysis and produces AM, FM, and PM demodulation results, similar to what a modulation analyzer would produce, allowing you to view amplitude, frequency, and phase profiles versus time. These additional analog-demodulation capabilities can be used to troubleshoot particular problems in a digital communications transmitter. For example, phase demodulation is often used to troubleshoot instability at a particular LO frequency.

Since digital communications systems use complex signals (I-Q waveforms), vector-modulation analysis is required to measure digitally-modulated signals. But vector-modulation analysis is not enough to measure today's complicated digitally-modulated signals. You also need digital-modulation analysis. Digital-modulation analysis is needed to demodulate the RF modulated carrier signal into its complex components (the I-Q waveforms) so you can apply the numerical and visual tools to help quickly identify and quantify impairments to the I-Q waveforms. Digital-modulation analysis detects and recovers digital data bits.

Digital demodulation also provides modulation quality measurements. The technique used in Agilent VSA (described later in this section) can expose very subtle signal variations, which translate into signal quality information not available from traditional modulation quality measurement methods. Various display formats and capabilities are used to view the baseband signal characteristics and analyze modulation quality. VSA offers traditional display formats such as I-Q vector, **constellation**, **eye**, and **trellis** diagrams. The symbol/error summary table shows the actual recovered bits and valuable error data, such as error vector magnitude (EVM), magnitude error, phase error, frequency error, rho, and I-Q offset error. Other display formats, such as **magnitude/phase error versus time**, **magnitude/phase error versus frequency**, or **equalization** allow you to make frequency response and group delay measurements or see code-domain results. This is only a representative list of available display formats and capabilities. Those available are dependent upon analysis capability and the type of digital-modulation format being measured.

VSA, with digital modulation provides measurement support for multiple digital communication standards, such as GSM, EDGE, W-CDMA, and cdma2000, as well as other digitally-modulated formats such as LTE, WLAN, and WiMAX, including MIMO signals. These signals are much more complex than the simple ones we will examine here. Measurements are possible on continuous or burst carriers (such as TDMA), and you can make measurements at baseband, IF, and RF locations throughout a digital communications system block diagram. There is no need for external filtering, coherent carrier signals, or symbol clock timing signals. The general-purpose design of the digital demodulation algorithms in Agilent VSA also allows you to measure non-standard formats, allowing you to change user-definable digital parameters for customized test and analysis purposes.

The remainder of this chapter contains additional concepts to help you better understand vector-modulation analysis, digital-modulation analysis, and analog-modulation analysis.

Vector modulation and digital modulation overview

Let's begin our discussion by reviewing vector modulation and digital modulation. As a reminder, while the terms modulators and demodulators may imply hardware, with software-based VSA, it's really DSP-based software that is performing the modulation/demodulation. Digital modulation is a term used in radio, satellite, and terrestrial communications to refer to modulation in which digital states are represented by the relative phase and/or amplitude of the carrier. Although we talk about digital modulation, you should remember that the modulation is not digital, but truly analog. Modulation is the amplitude, frequency, or phase modification of the carrier in direct proportion to the amplitude of the modulating (baseband) signal. See Figure 12. In digital modulation, it is the baseband modulating signal, not the modulation process, that is in digital form.
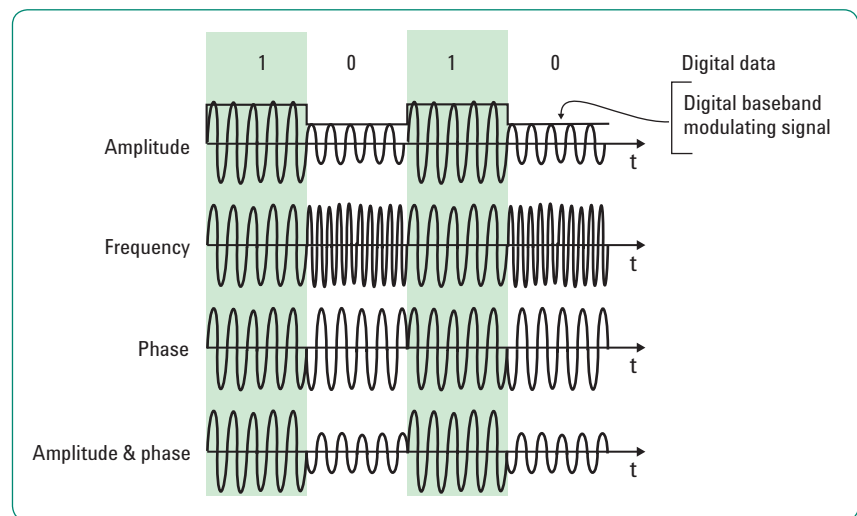

Figure 12. In digital modulation, the information is contained in the relative phase, frequency, or amplitude of the carrier

Depending on the particular application, digital modulation may modify amplitude, frequency, and phase simultaneously or separately. This type of modulation could be accomplished using conventional analog modulation schemes like amplitude modulation (AM), frequency modulation (FM), or phase modulation (PM). However, in practical systems, vector modulation (also called complex or I-Q modulation) is used instead. Vector modulation is a very powerful scheme because it can be used to generate any arbitrary carrier phase and magnitude. In this scheme, the baseband digital information is separated into two independent components: the I (In-phase) and Q (Quadrature) components. These I and Q components are then combined to form the baseband modulating signal. The most important characteristic of I and Q components is that they are independent components (orthogonal). You'll learn more about I and Q components and why digital systems use them in the following discussion.
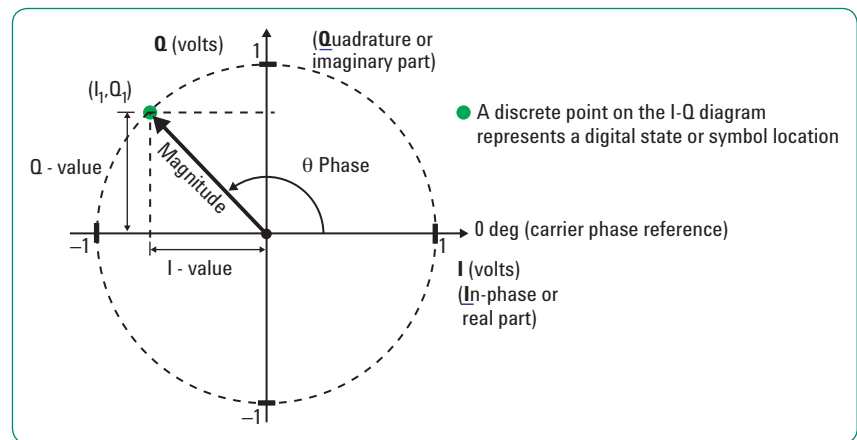

Figure 13. Digital modulation I-Q diagram

An easy way to understand and view digital modulation is with the I-Q or vector diagram shown in Figure 13. In most digital communication systems, the frequency of the carrier is fixed so only phase and magnitude need to be considered. The unmodulated carrier is the phase and frequency reference, and the modulated signal is interpreted relative to the carrier. The phase and magnitude can be represented in polar or vector coordinates as a discrete point in the I-Q plane. See Figure 13. I represents the in-phase (phase reference) component and Q represents the quadrature (90° out of phase) component. You can also represent this discrete point by vector addition of a specific magnitude of in-phase carrier with a specific magnitude of quadrature carrier. This is the principle of I-Q modulation.

By forcing the carrier to one of several predetermined positions in the I-Q plane, you can then transmit encoded information. Each position or state (or transitions between the states in some systems) represents a certain bit pattern that can be decoded at the receiver. The mapping of the states or symbols at each symbol timing instant (when the receiver interprets the signal) on the I-Q plane is referred to as a constellation diagram. See Figure 14. A symbol represents a grouping of the digital data bits; they are symbolic of the digital words they represent. The number of bits contained in each symbol, or bits-per-symbol (bpsym), is determined by the modulation format. For example, binary phase shift keying (BPSK) uses 1 bpsym, quadrature phase shift keying (QPSK) uses 2 bpsym, and 8-state phase shift keying (8PSK) uses 3 bpsym. Theoretically, each state location on the constellation diagram should show as a single point, but a practical system suffers from various impairments and noise that cause a spreading of the states (a dispersal of dots around each state). Figure 14 shows the constellation or state diagram for a 16 QAM (16-state quadrature amplitude modulation) format; note that there are 16 possible state locations. This format takes four bits of serial data and encodes them as single amplitude/phase states, or symbols. In order to generate this modulation format, the I and Q carriers each need to take four different levels of amplitude, depending on the code being transmitted.
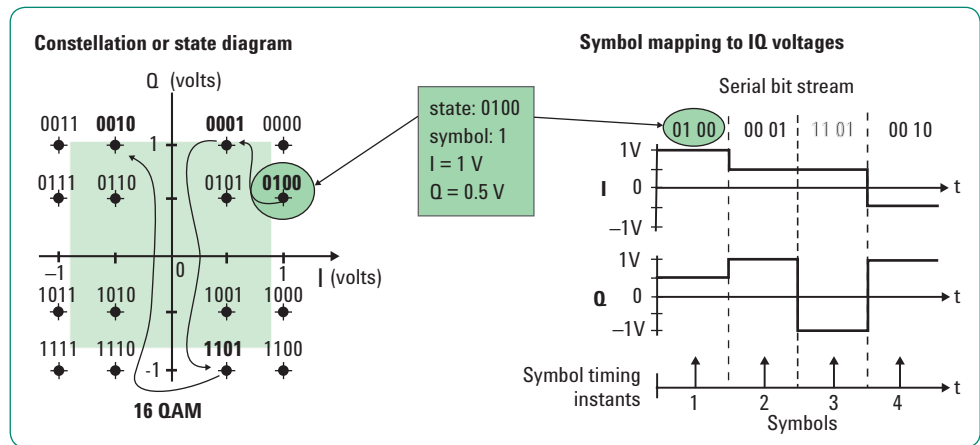
Figure 14. Each position, or state, in the constellation diagram represents a specific bit pattern (symbol) and symbol time

In digital modulation, the signal moves among a limited number of symbols or states. The rate at which the carrier moves between points in the constellation is called the symbol rate. The more constellation states that are used, the lower the required symbol rate for a given bit rate. The symbol rate is important because it tells you the bandwidth required to transmit the signal. The lower the symbol rate, the less bandwidth required for transmission. For example, the 16 QAM format, mentioned earlier, uses 4 bits per symbol. If the radio transmission rate is 16 Mbps, then the symbol rate = 16 (Mbps) divided by 4 bits, or 4 MHz. This provides a symbol rate that is one-fourth the bit rate and a more spectrally efficient transmission bandwidth (4 MHz versus 16 MHz). For more detailed information about digital modulation, see the **Additional Resources** section at the end of this application note.

I-Q modulation

In digital communications, I-Q modulation puts the encoded digital I and Q baseband information onto the carrier. See Figure 15. I-Q modulation generates signals in terms of I and Q components; fundamentally it is a hardware or software implementation of a rectangular to polar coordinate conversion.
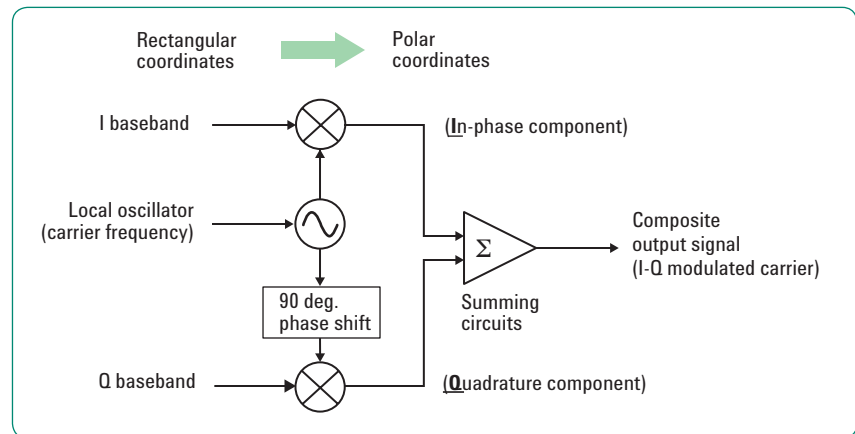

Figure 15. I-Q modulation

I-Q modulation receives the I and Q baseband signals as inputs and mixes them with the same local oscillator (LO). Remember, this may be a digital (software) LO. Thus, I and Q are both upconverted to the RF carrier frequency. The I information amplitude modulates the carrier producing the **in-phase** component. The Q information amplitude modulates a 90-degree (orthogonal) phase shifted version of the carrier producing the **quadrature** component. These two orthogonal modulated carrier signals are summed together producing the composite I-Q modulated carrier signal. The main advantage of I-Q modulation is the symmetric ease of combining independent signal components into a single, composite signal, and later splitting the composite signal into its independent component parts.

Signals that are separated by 90 degrees are known as being **orthogonal** to each other, or in **quadrature**. The quadrature relationship between I and Q signals means that these two signals are truly independent. They are two independent components of the same signal. While changes of the Q input certainly alter the composite output signal, they do not change the I component at all. Similarly, changes of the I input have no effect on the Q signal.

I-Q demodulation

As you can see in Figure 16, I-Q demodulation is a mirror image of the I-Q modulation shown in Figure 15. I-Q demodulation recovers the original I and Q baseband signals from a composite I-Q modulated input signal.
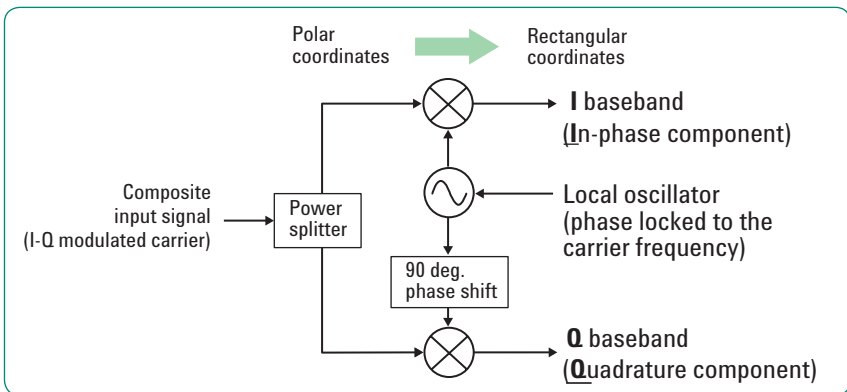

Figure 16. I-Q demodulation (or quadrature detection)

The first step in the demodulation process is to phase-lock the receiver LO to the transmitter carrier frequency. It is necessary that the receiver LO be phase-locked to the transmitter carrier (or mixer LO) to correctly recover the I and Q baseband components. Then, the I-Q modulated carrier is mixed with both an unshifted LO, and a 90 degree phase-shifted version of the LO, producing the original I and Q baseband signals or components. In software VSA, the 90 degree phase shifting can be done mathematically.

The I-Q demodulation process is fundamentally a polar to rectangular conversion. Normally, information cannot be plotted in a polar format and reinterpreted as rectangular values without doing a polar-to-rectangular conversion. See Figure 13. However, this conversion is exactly what is done by the in-phase and quadrature mixing processes performed by the I-Q demodulator.

Why use I and Q?

Digital modulation uses I and Q components because it provides a simple, efficient, and robust modulation method for generating, transmitting, and recovering digital data. Modulated signals in the I-Q domain provide many advantages:

1. The I-Q implementation provides a method to create complex signals (both phase and amplitude change). Instead of using phase modulation, which is nonlinear and difficult to do well, the I-Q modulator simply modulates the amplitude of the carrier and its quadrature in a linear manner. Mixers with wide modulation bandwidths and good linearity are readily available, as are baseband and IF software-based LOs. To produce a complex modulated signal, you only need to generate the baseband I and Q components of the signal. One key advantage of I-Q modulation is that the modulation algorithms can be used to generate a variety of modulations from digital formats to RF pulses, or even radar chirps, for example.

2. Demodulating the signal is also straightforward. Using I-Q demodulation, it is simple, at least in principle, to recover the baseband signals.

3. Looking at a signal in the I-Q plane often gives good insights about the signal. Effects like cross talk, data skew, compression, and AM-to-PM distortion, which are hard to visualize otherwise, are easy to see.

Digital RF communication system concepts

Figure 17 shows a generic, simplified block diagram of the basic architecture of a digital RF communications system that uses I-Q modulation. By understanding the fundamental concepts of this system, the operation of VSA with vector modulation analysis may also be understood. All sections of the communications transmitter and receiver can be measured and analyzed by VSA with vector modulation analysis. Also remember that even a software simulation of this block diagram can be analyzed with VSA, as VSA only requires time-sampled data.
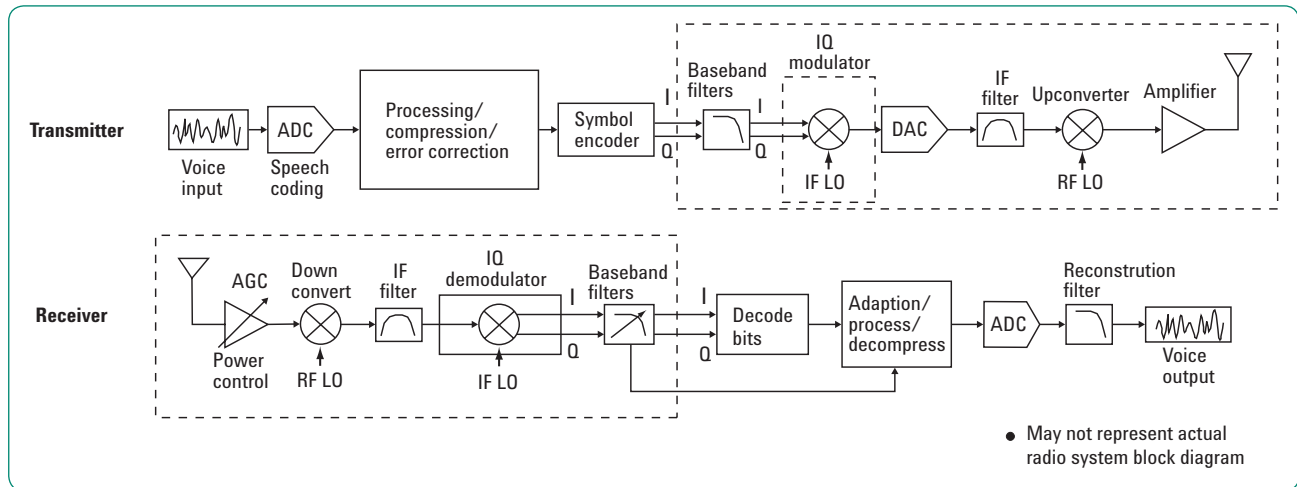

Figure 17. Digital RF communication system simplified block diagram. Note that ADCs and DACs may be located at different blocks.

Digital communication transmitter concepts

The communications transmitter begins with speech coding (assuming voice transmission) which is the process of quantizing the analog signal and converting it into digital data (digitization). Then, data compression is applied to minimize the data rate and increase spectral efficiency. Channel coding and interleaving are common techniques used to improve signal integrity by minimizing the effects of noise and interference. Extra bits are often sent for error correction, or as training sequences, for identification or equalization. These techniques can also make synchronization (finding the symbol clock) easier for the receiver. The symbol encoder translates the serial bit stream into the appropriate I and Q baseband signals, each corresponding to the symbol mapping on the I-Q plane for the specific system. The symbol clock represents the frequency and exact timing of the transmission of the individual symbols. At the symbol clock transitions, the transmitted carrier is at the correct I-Q (or magnitude/phase) value to represent a specific symbol (a specific point on the constellation). The time interval between individual symbols is the symbol clock period, the reciprocal is the symbol clock frequency. The symbol clock phase is correct when the symbol clock is aligned with the optimum instant to detect the symbols.

Once the I and Q baseband signals have been generated, they are filtered (band limited) to improve spectral efficiency. An unfiltered output of the digital radio modulator occupies a very wide bandwidth (theoretically, infinite). This is because the modulator is being driven by baseband I-Q square waves with fast transitions; fast transitions in the time domain equate to wide frequency spectra in the frequency domain. This is an unacceptable condition because it decreases the available spectrum to other users and causes signal interference to nearby users, called adjacent-channel-power interference. Baseband filtering solves the problem by limiting the spectrum and restricting interference with other channels. In effect, the filtering slows the fast transitions between states, thereby limiting the frequency spectrum. Filtering is not without tradeoffs, however; filtering also causes degradation to the signal and data transmission.

The signal degradation is due to reduction of spectral content and the overshoot and finite ringing caused by the filters time (impulse) response. By reducing the spectral content, information is lost and it may make reconstructing the signal difficult, or even impossible, at the receiver. The ringing response of the filter may last so long so that it affects symbols that follow, causing intersymbol interference (ISI). ISI is defined as the extraneous energy from prior and subsequent symbols that interferes with the current symbol such that the receiver misinterprets the symbol. Thus, selecting the best filter becomes a design compromise between spectral efficiency and minimizing ISI. There is a common, special class of filters used in digital communication design called Nyquist filters. Nyquist filters are an optimal filter choice because they maximize data rates, minimize ISI, and limit channel bandwidth requirements. You will learn more about filters later in this section. To improve the overall performance of the system, filtering is often shared, or split, between the transmitter and the receiver. In that case, the filters must be as closely matched as possible and correctly implemented, in both transmitter and receiver, to minimize ISI. Figure 17 only shows one baseband filter, but in reality, there are two; one each for the I and Q channel.

The filtered I and Q baseband signals are inputs to the I-Q modulator. The LO in the modulator may operate at an intermediate frequency (IF) or directly at the final radio frequency (RF). The output of the modulator is the composite of the two orthogonal I and Q signals at the IF (or RF). After modulation, the signal is upconverted to RF, if needed. Any undesirable frequencies are filtered out and the signal is applied to the output amplifier and transmitted.

Digital communications receiver concepts

The receiver is essentially an inverse implementation of the transmitter, but it is more complex to design. The receiver first downconverts the incoming RF signal to IF, then demodulates it. The ability to demodulate the signal and recover the original data is often difficult. The transmitted signal is often corrupted by such factors as atmospheric noise, competing signal interference, multipath, or fading.

The demodulation process involves these general stages: carrier frequency recovery (carrier lock), symbol clock recovery (symbol lock), signal decomposition to I and Q components (I-Q demodulation), I and Q symbol detection, bit decoding and de-interleaving (decode bits), decompressing (expansion to original bit stream), and finally, digital to analog conversion (if required).

The main difference between the transmitter and receiver is the need for carrier and symbol clock recovery. Both the symbol clock frequency and phase (or timing) must be correct in the receiver to demodulate the bits successfully and recover the transmitted information. For example, the symbol clock could be set to the correct frequency, but at the wrong phase. That is, if the symbol clock is aligned to the transitions between symbols, rather than the symbols themselves, demodulation will be unsuccessful.

A difficult task in receiver design is to create carrier and symbol clock recovery algorithms. Some clock recovery techniques include measuring the modulation amplitude variations, or in systems with pulsed carriers, the power turn-on event can be used. This task can also be made easier when channel coding in the transmitter provides training sequences or synchronization bits.

VSA digital modulation analysis concepts and theory of operation

VSA can be viewed as a software-based measuring receiver. It is really an I-Q receiver employing techniques similar to most digital radio receivers for decoding digital modulations. However, the difference is that VSA software is designed for high accuracy parametric measurement and display of modulation characteristics. Moreover, VSA is a measurement tool that can measure and analyze almost every aspect of a digital communications transmitter and receiver system.
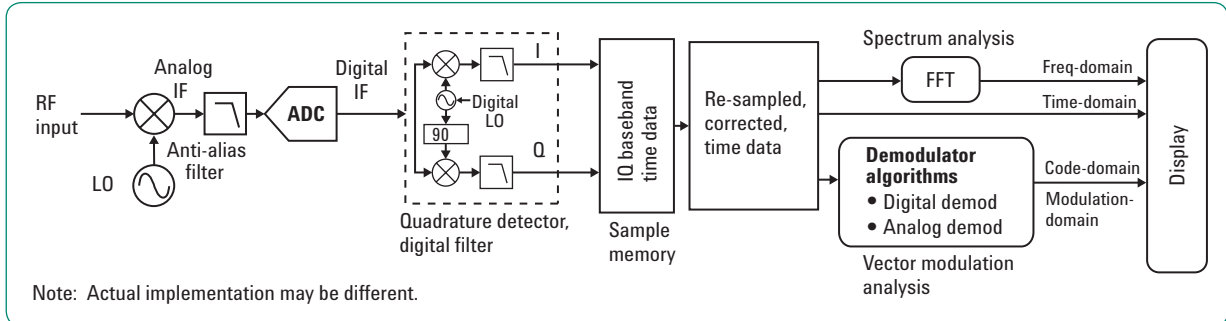

Figure 18. Simplified block diagram of a VSA measurement system with an RF front-end. For other front-ends, the VSA software will perform some functions not supported by the given front-end.

Figure 18 shows an Agilent 89600B VSA simplified system block diagram. You may notice that many of the system blocks are analogous to the digital communication receiver shown in Figure 17. The RF input signal is downconverted, through several stages of superheterodyne mixing, to an IF that can be accurately digitized by the ADC. This part of the process can be done by many different types of front ends. Some, like signal analyzers, will provide the RF signal detection and digitization of the IF. Others, like oscilloscopes and logic analyzers, will provide a completely digitized baseband signal. The required VSA input is digitized, time-sampled data. This digitized signal is then vector (quadrature) detected and digitally filtered; downconverted, if required, one last time to an I and Q baseband form (I-Q time data) and stored in RAM. From here, DSP algorithms demodulate the signal; recover the carrier and symbol clock and apply reconstructive filtering and decoding (recover the original bits). With this DSP software implementation, almost any modulation format can be demodulated.

VSA implementation is different from a radio receiver. VSA deals with the sampled signals on a block basis; the radio receiver processes data serially, in real time. When you supply the VSA software with radio receiver parameters, it synthesizes the receiver via processing in the DSP. It provides all the functions of a receiver, down to making analog waveforms. Because the signal has been virtually digitized, it can be post-processed and viewed in any of the time, frequency, or modulation domains.

VSA digital demodulation
The heart of the digital demodulation process in VSA is digital demodulation. Figure 19 shows a simplified block diagram of the digital demodulation algorithm implementation used in the Agilent 89600B. The digital demodulation algorithm provides general purpose demodulation that only requires a minimum of prior information about the input signal to perform demodulation, and can be used on a wide variety of modulation formats. The demodulator provides carrier lock, symbol clock recovery, and bit recovery (decodes the actual encoded 1s and 0s), and produces the I-Q measured waveforms. The demodulator also produces

ideal I-Q waveforms that are synthesized from the actual recovered bits (called I-Q reference waveforms). The I-Q measured and I-Q reference waveforms are subtracted to produce I-Q error waveforms. The I-Q error waveforms are analyzed to generate the modulation quality data results, which can be viewed in various data formats and display outputs.

The demodulation process begins by configuring the VSA's software-based demodulator. The demodulation algorithms must be configured to the specific digital modulation format to properly demodulate and analyze the signal. Most analysis options provide a set of standard presets, such as GSM, W-CDMA, cdma2000, or 802.11a/b/g, that automatically configure the demodulator for you. In those cases, you enter the center frequency, select a standard preset, and the software can demodulate the signal.

Flexible custom or user-definable demodulation

The Agilent 89600B VSA software gives you additional capabilities through a general purpose, user-definable demodulator. This allows you to customize the demodulator configuration for non-standard formats or troubleshooting purposes. The demodulation block diagram shown in Figure 19 shows the internal demodulation processes (enclosed in rectangular boxes) and the configuration parameters that you can set (enclosed by ovals or rounded rectangular boxes). The items enclosed by an oval identify configuration parameters that are required to define the demodulator for a measurement. The rounded rectangular boxes identify user-adjustable input parameters. At a minimum, the demodulation algorithm requires the modulation format (QPSK, FSK, and so forth), the symbol rate, the baseband filter type, and filter alpha/BT. This set of parameters is generally sufficient for the demodulator to lock to the signal and recover the symbols for many formats. Other formats, like custom OFDM require additional information specific to the format type.
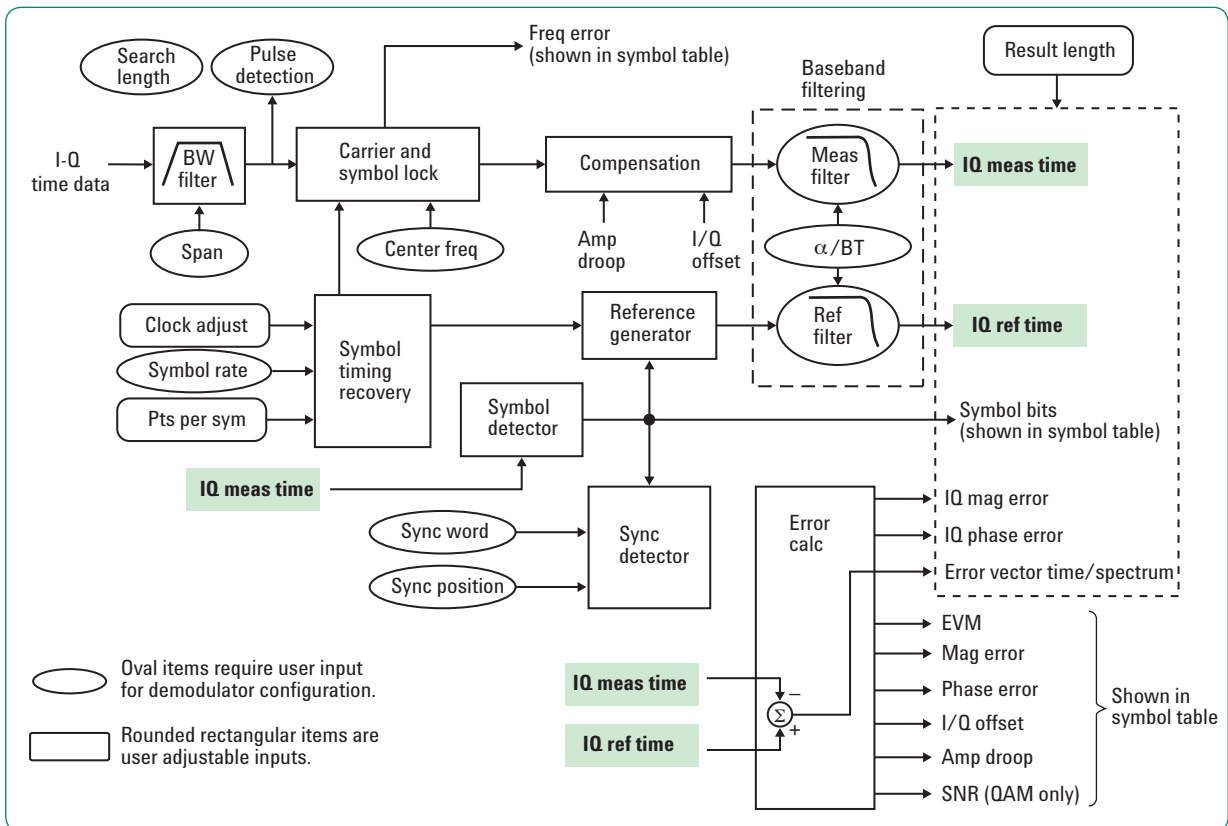
Figure 19. Digital block diagram

Digital demodulation baseband filtering

As mentioned earlier, baseband filtering is used in digital demodulation to limit bandwidth and reduce intersymbol interference. Also, just like the communication receiver, digital demodulation baseband filtering must be configured to match the system under test to accurately demodulate the signal. This requires that both the filter type (such as Nyquist or Gaussian) and filter bandwidth coefficients (such as **alpha** or **BT**) match, as well.

As shown in Figure 20, both the measured and reference I-Q waveforms have their own signal processing path and baseband filtering. The **I-Q measured** waveform must use baseband filtering that matches the receiver filtering of the system under test. This filter is called the measurement filter or **Meas Filter**. The **I-Q reference** waveform must use baseband filtering that matches the total system channel filtering, transmitter and receiver, of the system under test. This filter is called the reference filter or **Ref Filter**. The reference filter simulates the total channel filtering because it is used to synthesize the ideal I-Q signals that would be received by a "perfect" linear channel response. The demodulator must apply the total system channel filtering to accurately synthesize the reference I-Q waveform.

Selecting the correct filtering

In digital communications systems, baseband filtering may occur either at the transmitter or the receiver; or the filtering may be distributed between the transmitter and the receiver where half of the filtering is done in the transmitter and half is done in the receiver. This is an important concept that affects the filter type needed by the demodulator for the measured and reference I-Q waveforms. The **Meas Filter** of the VSA software represents baseband filtering in the system receiver and the **Ref Filter** represents baseband filtering in the entire system (total receiver and transmitter channel filtering).
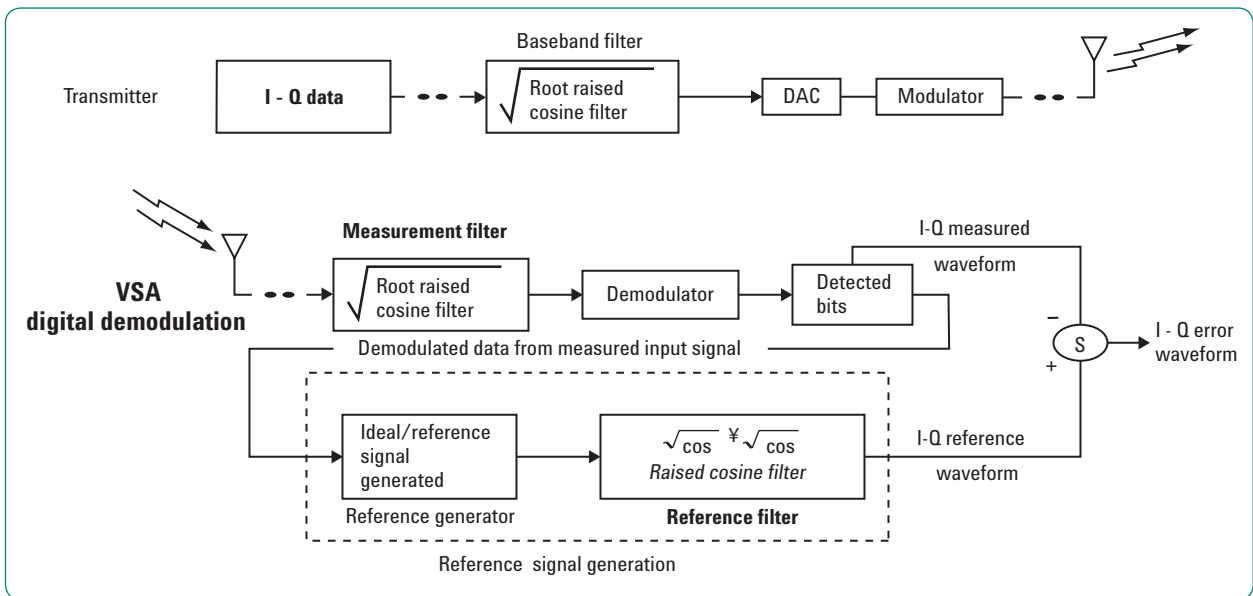

Figure 20. Selectable matching filters are used to represent filtering in transmitter and receiver. Detected bits are used, in conjunction with knowledge of modulation type and filtering, to determine ideal signal

Selecting correct demodulation baseband filtering may not be as straightforward and intuitive as you might think, especially for a distributed system. For example, consider a North American digital cellular (NADC) standard which uses distributed filtering; a root raised cosine filter is implemented in both the transmitter and receiver. See Figure 20. For the VSA demodulation, you would use a root raised cosine filter for the I-Q measured waveform (matching the receiver filtering in a system that includes a similar filter at the transmitter). For the I-Q reference waveform, you would use a raised cosine filter (matching the total system channel filtering). This is because the product of a sqrt (raised cosine) times a sqrt (raised cosine) is equal to a raised cosine filter.

Table 3 shows some commonly used filter types and examples of measurement and reference filter selections based on the transmitter filter type.

Table 3. Commonly used transmitter filter types

If the transmitter filter is:	The measure filter should be:	The reference filter should be:
Root raised cosine	Root raised cosine	Raised cosine
Raised cosine	None	Raised cosine
Gaussian	None	Gaussian
Any type	User defined	User defined

Filter alpha and BT bandwidth time product
Another filter parameter that must accurately represent the system under test is the filter bandwidth coefficient, specified as the filter **alpha** or **BT**. Each filter type will have a filter bandwidth coefficient associated with it; Nyquist filters use alpha and Gaussian filters use BT. The demodulator uses the same **alpha** or **BT** value for both the measurement filter and the reference filter.

There are many available filter types used in digital communications, depending on the intended application. Traditionally, the Nyquist (raised cosine) filter has been used because it minimizes ISI. Notice, in Figure 21, that the Nyquist filter impulse response has its peak amplitude at the symbol instant (t = 0) and is zero at all other surrounding symbol instants. That is, it crosses zero at integer multiples of the symbol period (1 divided by f_s). This means that Nyquist filtered symbols do not interfere with surrounding symbols (zero inter-symbol interference). Nyquist filters minimize ISI, however, there are applications when ISI is not the most important criteria. Two other commonly used filter types are the Gaussian and Chebyshev filters. The Gaussian filter does not have the best ISI characteristics, but it does have advantages in the balance of carrier power, occupied bandwidth, and symbol clock recovery. Gaussian filters are typically used in GSM (global system for mobile communications) wireless telephony systems. The Chebyshev filter has very sharp roll -off characteristics and is vital for reducing power leakage into adjacent channels. Chebyshev filters are often used in wireless telephony systems that use CDMA (code division multiple access) modulation schemes, like cdmaOne and cdma2000. There are many other filter types used in digital communications that are beyond the scope of this discussion (see the **Additional Resources** section at the end of this application note for more information).
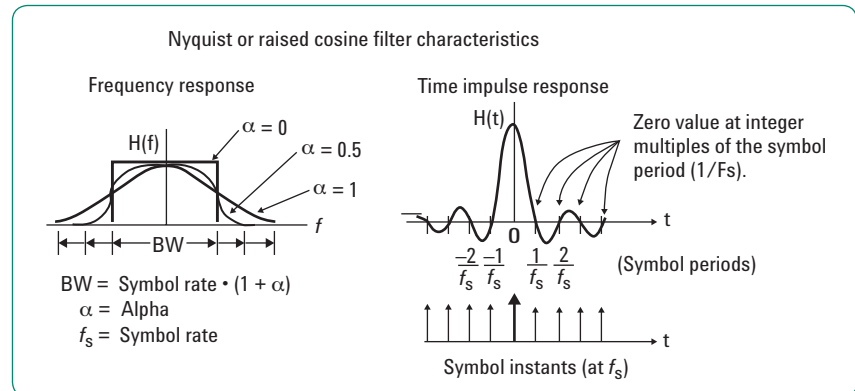

Figure 21. Nyquist filters minimize inter-symbol interference (ISI)

Alpha

Alpha describes the sharpness of the Nyquist (raised cosine) filter. See Figure 21. **Alpha** is also called the roll-off or excess bandwidth factor. A higher value for alpha increases the bandwidth that is used in excess of the theoretical minimum. Modulation theory states that the minimum bandwidth needed to transmit a signal is equal to one half the symbol rate. However, to realize this system bandwidth would require a perfect brick-wall (rectangular shaped) filter, where **alpha** equals 0 and the occupied bandwidth equals the symbol rate. But a brick-wall filter is not realizable, so in practice, communication systems typically use a filter **alpha** of 0.3. An **alpha** value of 0.3 means that it will use 30 % more occupied BW than the theoretical minimum. This value does provide a good compromise between spectral efficiency and minimum ISI. The occupied bandwidth for a given alpha is approximately equal to the sample rate times (1 + **alpha**).

BT, bandwidth time product

Filter BT, bandwidth time product, is the corresponding filter coefficient used for Gaussian filters and describes the sharpness of the filter. Gaussian filters typically use BT values between 0.3 to 0.5.

Demodulation analysis

Once a user provides the configuration inputs, the demodulator uses them and, through DSP, operates on the I-Q time data received in block format from the sample memory available for VSA. The VSA software can receive I-Q time data from external hardware (such as an Agilent X-Series signal analyzer or Infiniium Series oscilloscope) or from a recorded file. The demodulator uses the supplied center frequency and symbol rate to lock to the carrier and recover the symbol clock from the modulated carrier. Note that the demodulator reference clock does not need to be locked with the source clock. The demodulation algorithm automatically provides carrier and symbol lock; you do not need to supply an external source clock input. The signal then goes through a compensation process that applies gain and phase correction. The compensation data (such as amplitude droop and I-Q offset error data) is stored and can be viewed in the error summary table. Digital baseband filtering is then applied to recover the baseband I-Q waveforms (I-Q Meas Time data). The recovered I-Q waveforms are applied to a symbol detector that attempts to determine what symbols were transmitted, based upon the specified modulation format. From this block of symbols, the serial data bits (1s and 0s) are decoded and recovered. The reference generator uses the detected symbols in conjunction with the modulation format, the symbol rate, and the specified filtering, to synthesize an ideal set of I-Q reference baseband waveforms (I-Q Ref Time data). Finally, the measured I/Q waveforms and reference I-Q waveforms are compared to produce a host of error characteristics (deviation from perfect) such as phase error, magnitude error, and error vector magnitude (EVM).

I-Q measurement and I-Q reference signal

The quality of an I-Q modulated signal can be analyzed by comparing the measured signal to an ideal reference signal. See Figure 20. Demodulation produces two waveforms: an I-Q measured waveform and an I-Q reference waveform. The I-Q measured waveform is the demodulated baseband I-Q data for the measured input signal, also called IQ Meas Time. The I-Q reference waveform is the baseband I-Q data that would result after demodulating the input signal if the input signal were ideal (contained no errors), also called IQ Ref Time. The I-Q reference waveform is mathematically derived from the I-Q measured waveform recovered data bits, providing that the original data sequence can be recovered. The I-Q reference waveform generation begins by recovering the actual symbol bits from the demodulated I-Q measured waveform, and then reconstructing a sequence of ideal I and Q states. These states are then treated as ideal impulses and are baseband filtered according to the reference channel filtering, producing an ideal I-Q reference waveform. The quality of the input signal can then be analyzed by comparing the I-Q measured waveform to the I-Q reference waveform. Subtracting the reference waveform from the measured waveform provides the error vector waveform, or I-Q error waveform. This technique can expose very subtle signal variations, which translates into signal quality information not available from traditional modulation quality measurement methods.

Measurement concepts

Now that you have learned some digital modulation basics and VSA system theory, the next step is to learn about digital modulation analysis measurement results and trace displays used to troubleshoot systems. The following examples show the measurement results for a QPSK, quadrature phase shift keyed, modulated signal with a 50 ksym/s symbol rate and a root raised cosine baseband filter with alpha equal to 0.35. Quadrature means that the carrier signal shifts between phase states that are separated by 90 degrees. The signal shifts in increments of 90 degrees from 45 to 135, –45, or –135 degrees. QPSK has four available states. Each state is assigned a binary value from 0 to 3, which requires 2 bits per state, translating into 2 bits per symbol. Only two I and two Q values are needed to produce the four states, satisfying the 2 bits per state requirement.

Vector (or IQ) diagrams

The vector diagram, more commonly called an IQ diagram in digital modulation, shows the recovered complex, I-Q baseband signal at all moments in time. It shows the states and transitions between them as the signal moves from symbol to symbol in time. A vector drawn from the origin to a point on the vector diagram corresponds to the instantaneous voltage at that instant in time.

Figure 22a shows an example IQ diagram for the QPSK modulated signal mentioned earlier. The IQ diagram shows the 4 ideal state locations (indicated by the crosshairs) positioned at 45, 135, –45, and –135 degrees, the detected symbols, and the transitions between the symbols. The IQ diagram gives you the peak-to-average voltage ratios and can be used to determine amplifier compression among other things.
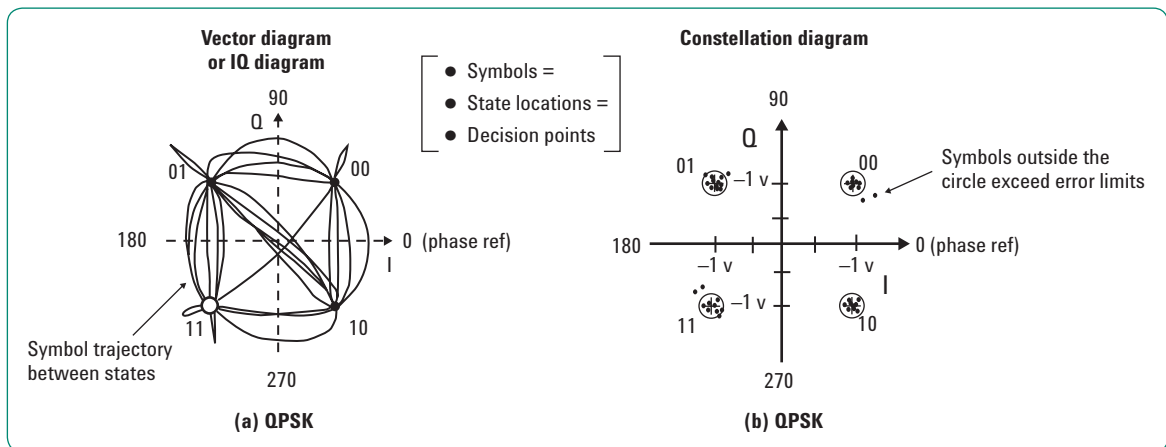

Figure 22. QPSK vector and constellation diagrams

Interpreting vector diagrams

When viewing signals in the I-Q plane, keep in mind that you are looking at the signal amplitude and phase relative to the carrier. The unmodulated carrier is the phase reference (0 degrees). In Figure 22, each of the detected symbols modulate the carrier at differing amplitudes and phases relative to the unmodulated carrier, but at the same frequency as the carrier. If the detected symbols were at different frequencies than the unmodulated carrier, they would appear to move as they continuously gain or lose phase relative to the unmodulated carrier. Also, digital modulation is a 3-dimensional measurement, modulation in terms of its I (in-phase) and Q (quadrature) components versus time. The IQ diagram is only a 2-dimensional display, so the time parameter, which cannot be seen, is perpendicular to the I-Q plane (or CRT screen).

Constellation Diagram

Figure 22b shows the constellation diagram for the same QPSK signal. The constellation diagram shows the carrier magnitude and phase synchronous with the symbol clock. The measurement points are commonly referred to as detection decision points and represent the detected symbols. It is similar to the I-Q diagram, without showing the transition traces between states.

Ideally all symbols should appear as a single point centered at the ideal state location (indicated by the crosshairs). The ideal state is where symbols occur if your signal is without error. However, signal impairments and other modulation errors cause deviations and the symbols will disperse and surround the ideal state location. The 89600B VSA allows you to put definable error limit circles around the ideal states. Showing the actual detected symbols versus the ideal states gives you a visual indication of the quality of your signal. Constellation diagrams help identify such signal impairments as amplitude imbalance, quadrature error, or phase noise.

Error vector magnitude (EVM)

The most widely used modulation quality metric in digital communications systems is error vector magnitude (EVM). The error vector is the vector difference at a given time between the ideal reference signal and the measured signal. See Figure 23. The error vector is a complex quantity that contains a magnitude and phase component. It is important not to confuse the magnitude of the error vector with the magnitude error, or the phase of the error vector with the phase error.

$$EVM\,[n] = \sqrt{I\,err[n]^2 + Q\,err[n]^2}$$

where [n] = measurement at the symbol time

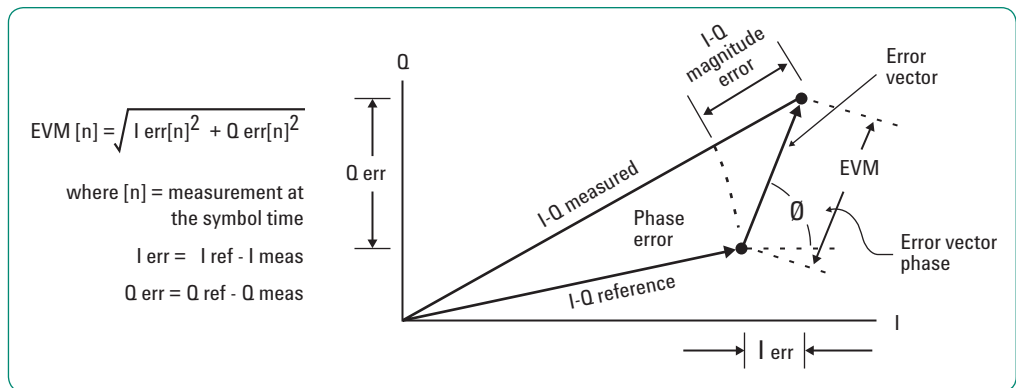I err = I ref - I meas

Q err = Q ref - Q meas

Figure 23. Error vector magnitude (EVM); difference between actual measured signal and ideal reference signal

EVM is defined as the root mean square (rms) of the error vector over time at the instants of the symbol clock transitions. By convention, EVM is usually normalized to either the amplitude of the outermost symbol or the square root of the average symbol power. The EVM measurement, also called relative constellation error (RCE) by some communications standards, and the various related data results are sensitive to any signal impairment that affects the magnitude and phase trajectory of a signal for any digital modulation format. This makes it an ideal measurement tool for troubleshooting communications system problems at baseband, IF, or RF sections of the radio.

Figure 24 shows examples of the modulation quality measurements that were defined in Figure 23. The **error vector time** data (Trace A) is the computed error vector magnitude between corresponding symbol points in the I-Q measured and I-Q reference signals. The **error vector spectrum** data (Trace B) shows the spectrum of the **error vector time** data. In other words, the error vector time data is windowed and FFT'd to produce the **error vector spectrum** trace. This format can reveal the spectral content of the unwanted signals that drive the modulated carrier away from its ideal path. If those error components are deterministic, they will show up in the error vector spectrum trace as spectra. Measuring these spectra can give added insight into the nature and origin of these error signals. Often the frequency peaks will correspond to a signal from somewhere else in the block diagram. The **I-Q magnitude error** (Trace C) and **I-Q phase error** (Display D) traces show the error between the I-Q measured and the I-Q reference signals. **I-Q magnitude error** displays the magnitude error and **I-Q phase error** displays the phase error.

Figure 24. Trace A, error vector time, shows the error vector magnitude at the symbol points. Trace B, error vector spectrum, shows the frequency spectrum of the error vector time data. Trace C, IQ mag error, is the magnitude difference between the measured and reference IQ waveforms. Trace D, IQ phase error, is the phase difference between the measured and reference IQ waveforms.

Symbol table/error summary

The **symbol table/error summary** measurement results may be the most power-ful of the digital demodulation tools. Here, demodulated bits can be seen along with error statistics for all of the demodulated symbols. For example, reviewing the rms EVM value can quickly help you assess modulation accuracy. Other valu-able errors are also reported. Figure 25 shows the **symbol table/error summary** data for the QPSK signal used previously. The marker readout shows the values for the highlighted bits in the symbol table, which represent symbol number 3 and has a value of 2 (a binary value of "10"). Note that this agrees with the two bits per symbol required by QPSK modulation. The errors table shows the available statistical and error data for a QPSK modulated signal. Additional error values are available for other formats.

We have now covered some of the fundamental modulation measurements, but there are many more qualitative displays and quantitative measurements available. You can find more information about them in the **Additional Resources** section at the end of this application note.
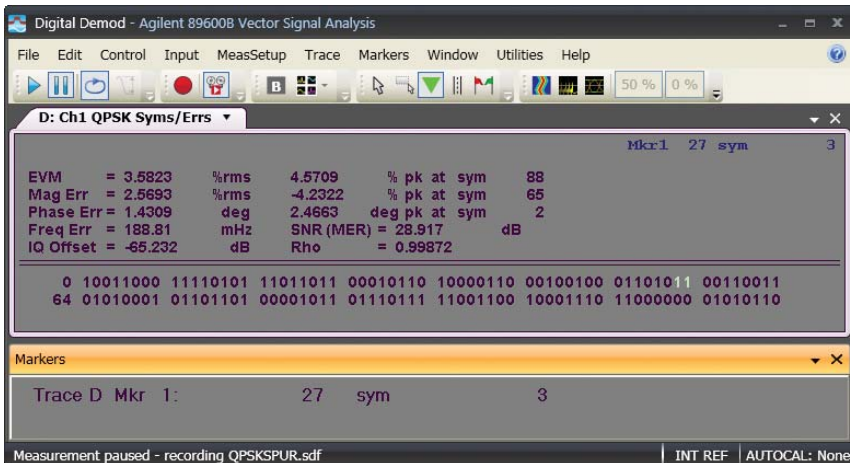

Figure 25. The symbol table/error summary data provides the demodulated bits and error statistics for all the demodulated.

Analog modulation analysis

Another important measurement tool that vector modulation analysis may provide is analog modulation analysis. Analog modulation analysis produces AM, FM, and PM demodulation results, similar to what modulation analysis would output, allowing you to view amplitude, frequency, and phase profiles versus time. These analog modulation analysis capabilities enhance the digital modulation analysis capabilities, already available in VSA, to provide a complete measurement solution to analyze digital communications systems. For example, you may use analog modulation to analyze intentional modulations like FSK (frequency shift keying), to analyze unintentional modulations like phase noise or AM-to-PM conversion, or to analyze singleshot signal parameters like frequency or phase settling, or pulse shaping.

Demodulating a signal may seem complicated, but the fundamental characteristics of vector modulation and the I-Q modulation process make demodulation straightforward. Recall that the vector or IQ diagram (Figures 13 and 22) shows the instantaneous magnitude and phase of the modulated carrier relative to the unmodulated carrier, and that it was very useful to show digital modulation characteristics. However, it also provides an easy way to view analog modulation signal characteristics, such as AM, PM, and FM modulations, as shown in Figure 26. An unmodulated continuous wave (CW) signal would simply show as a single stationary point with constant magnitude and phase. An AM signal has its trajectory along a fixed line through the origin and only the magnitude of the signal changes. An FM signal is shown as a circle centered at the origin as its trajectory, with the instantaneous frequency deviation given by the rate of change of its phase. PM looks similar to FM, although, the relative change in phase is the controlling parameter. In vector modulation, in which both the amplitude and phase can change simultaneously, the vector trajectory changes in both magnitude and phase.
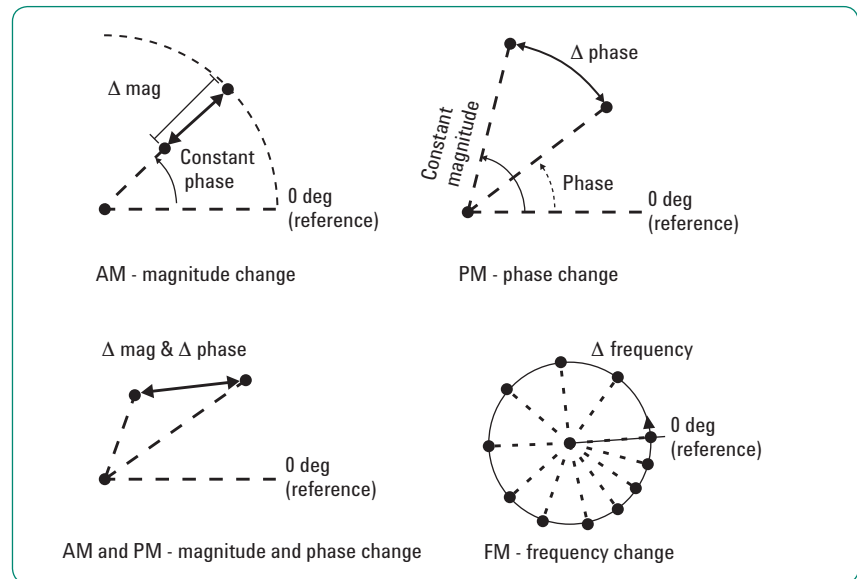

Figure 26. Analog modulation characteristics in the I-Q plane on the I-Q diagram.

Traditionally, to view the carrier modulation waveforms (modulation envelope), you would remove the carrier by "detecting" and displaying the resultant waveform on an oscilloscope in amplitude-versus-time display. However, instead of thinking in terms of detecting the modulation on the carrier, in vector modulation, it is helpful to think of viewing the modulation as what is left after the carrier has been "shifted" or "downconverted" to 0 Hz. An intuitive review of frequency shifting will show how amplitude and phase changes imparted on the carrier are "detected" using basic trigonometric identities.
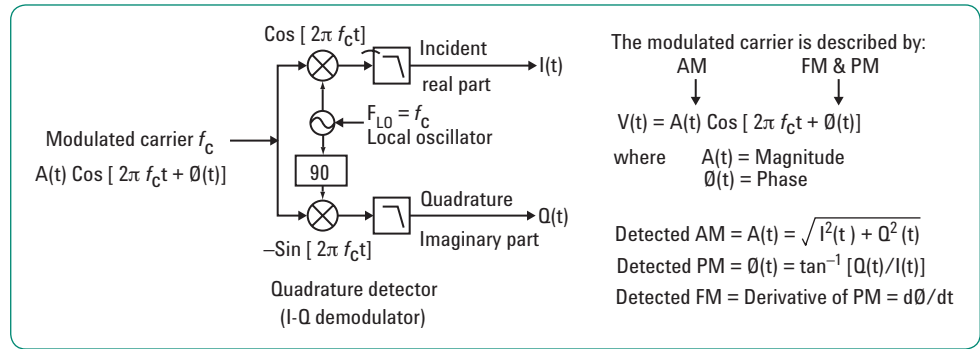

Figure 27. Modulation is measured by shifting the carrier frequency, fc, down to 0 Hz. When the digital LO frequency is equal to the modulated carrier frequency, the output of the quadrature detector, I(t) and Q(t) time-domain waveforms, is the modulation on the carrier.

As shown in Figure 27, the complex-modulated carrier signal, described by V(t) = A(t)Cos[2πfct + Ø(t)] and frequency, fc, are input to the quadrature mixer (or quadrature detector) for frequency translation. To recover the baseband modulating signal, the carrier is first down-converted to baseband (0 Hz) by setting the LO frequency equal to fc. Then the baseband signals are low-pass filtered so only the difference frequencies remain. This process produces real I(t) and imaginary Q(t) time-domain waveforms that are the magnitude and phase difference between the modulated carrier signal and the unmodulated LO signal, referenced to the carrier. This is the carrier modulation in the form of I(t) and Q(t) components (rectangular coordinates), not in terms of magnitude, A(t), and phase, Ø(t). You may recall that I-Q demodulation performs a polar to rectangular conversion. However, applying some math and simple trigonometric identities to the I(t) and Q(t) components allows us to extract the instantaneous amplitude, A(t), or phase, Ø(t), changes impressed on the carrier. This is the AM and PM modulation.

The amplitude modulation, AM, is the variation of the carrier magnitude, A(t), versus time and is extracted from I(t) and Q(t) by taking the square root of the sum of the squares of I(t) and Q(t).

$$AM = A(t) = sqrt[I2(t) + Q2(t)]$$

Phase modulation, PM, is the phase variation, Ø(t), versus time and is equal to the arctangent of [Q(t) / I(t)]. Frequency modulation, FM, is the derivative of the phase shift verses time, dØ/dt. That is, FM is the derivative of PM:

$$PM = Ø(t) = arctan[Q(t)/I(t)]$$

$$FM = derivative of the PM = (dØ/dt)$$

Starting with I(t) and Q(t) and applying basic trigonometric identities, we have done a first order demodulation of the carrier and are able to look at AM, PM, and FM modulation. In practice, VSA software uses sophisticated demodulation algorithms that incorporate frequency and phase error correction routines so that intentional and unintentional modulation can be accurately separated from the carrier. Analog demodulation completely isolates the PM and FM from the AM. Likewise, AM is completely isolated from the PM and FM.

Summary

This chapter presented the basics of vector or digital modulation techniques and digital modulation analysis as implemented in software-based vector signal analysis. We described digital (vector or I-Q) modulation and common digital modulation formats. VSA architecture is similar to a digital communication system so we described the radio receiver and transmitter block diagrams, noting the process of generating, transmitting, receiving and recovering the original digital information. We then looked at the VSA digital demodulation block and described each function and the related demodulation parameters necessary to properly demodulate your signal and make measurements. VSA, with vector modulation analysis, provides powerful and specialized measurement capability that allows you to characterize and troubleshoot throughout the entire block diagram of a digital system (system using I and Q signals). The information you learned in this chapter will help you make informed decisions regarding measurement instrumentation requirements and the tools necessary to analyze your communication system. VSA capability is available today as embedded radio personalities for signal analyzers, or as separate VSA software compatible with a wide range of measurement front ends and simulation software.

The scope of this application note is limited to the fundamental concepts. If you would like more in-depth information, you can review the suggested reference materials in the Additional Resources section at the end of this application note.

Additional Resources

Related Literature

Robert A Witte. *Spectrum and Network Measurements*. Atlanta, GA: Noble Publishing Corporation, 2001.

The Fundamentals of Signal Analysis, Application Note 243, literature number 5952-8898E.

Using Vector Modulation Analysis in the Integration, Troubleshooting and Design of Digital RF Communications Systems, Application Note 89400-8, literature number 5091-8687E.

Digital Modulation in Communications Systems — An Introduction, Application Note 1298, literature number 5965-7160E.

Testing and Troubleshooting Digital RF Communications Transmitter Designs, Application Note 1313, literature number 5968-3578E.

Testing and Troubleshooting Digital RF Communications Receiver Designs, Application Note 1314, literature number 5968-3579E.

Infiniium Oscilloscopes with 89600B VSA Software, Application Note, 5990-6819EN

InfiniiVision 6000/7000 Series Oscilloscopes with 89600B VSA Software, Application Note, 5990-6820EN

Making custom OFDM measurements using the Agilent 89600B VSA software with Option BHF, Application Note, literature number 5990-6824EN

Web

For additional information on using the 89600B VSA software, go to **www.agilent.com/find/89600B** and select **Technical Support** (tab) then **Manuals**

Agilent Email Updates

www.agilent.com/find/emailupdates
Get the latest information on the products and applications you select.

LXI

www.lxistandard.org
LAN eXtensions for Instruments puts the power of Ethernet and the Web inside your test systems. Agilent is a founding member of the LXI consortium.

Agilent Channel Partners

www.agilent.com/find/channelpartners
Get the best of both worlds: Agilent's measurement expertise and product breadth, combined with channel partner convenience.

WiMAX, Mobile WiMAX, and WiMAX Forum are trademarks of the WiMAX Forum.

cdma2000 is a US registered certification mark of the Telecommunications Industry Association.

Agilent
Advantage
Services

Agilent Advantage Services is committed to your success throughout your equipment's lifetime. We share measurement and service expertise to help you create the products that change our world. To keep you competitive, we continually invest in tools and processes that speed up calibration and repair, reduce your cost of ownership, and move us ahead of your development curve.

www.agilent.com/find/advantageservices

Agilent Electronic Measurement Group
KEMA Certified
ISO 9001:2008
Quality Management System

www.agilent.com/quality

www.agilent.com
www.agilent.com/find/
spectrumanalyzers

For more information on Agilent Technologies' products, applications or services, please contact your local Agilent office. The complete list is available at:

www.agilent.com/find/contactus

Americas

Canada	(877) 894 4414
Brazil	(11) 4197 3500
Mexico	01800 5064 800
United States	(800) 829 4444

Asia Pacific

Australia	1 800 629 485
China	800 810 0189
Hong Kong	800 938 693
India	1 800 112 929
Japan	0120 (421) 345
Korea	080 769 0800
Malaysia	1 800 888 848
Singapore	1 800 375 8100
Taiwan	0800 047 866
Other AP Countries	(65) 375 8100

Europe & Middle East

Belgium	32 (0) 2 404 93 40
Denmark	45 70 13 15 15
Finland	358 (0) 10 855 2100
France	0825 010 700*
	*0.125 €/minute
Germany	49 (0) 7031 464 6333
Ireland	1890 924 204
Israel	972-3-9288-504/544
Italy	39 02 92 60 8484
Netherlands	31 (0) 20 547 2111
Spain	34 (91) 631 3300
Sweden	0200-88 22 55
United Kingdom	44 (0) 118 9276201

For other unlisted Countries:
www.agilent.com/find/contactus
Revised: October 14, 2010

Product specifications and descriptions in this document subject to change without notice.

© Agilent Technologies, Inc. 2011
Printed in USA, February 17, 2011
5990-7451EN

Agilent Technologies