

Introduction to Network Analyzer Measurements

Fundamentals and Background

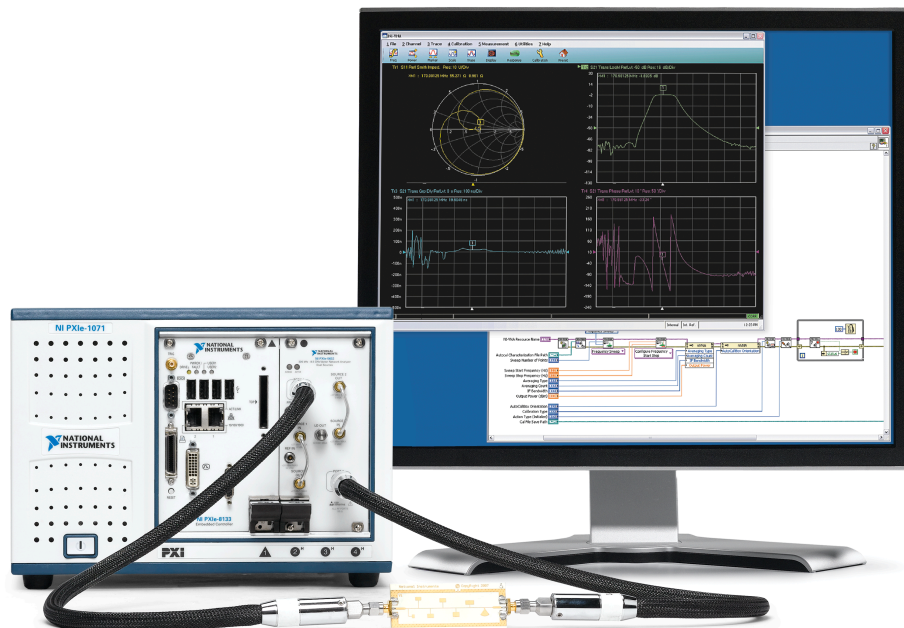


Table of Contents

1. Introduction to Network Analyzer Measurements.....	3
VNA Basics.....	3
Applications for Network Analyzers	5
History of Network Analysis.....	6
2. Understanding Wave Quantity Measurements	10
Measuring Wave Quantities	10
Introduction to S-parameters.....	12
3. Network Analyzer Architectures	15
Scalar Network Analyzers	15
Vector Network Analyzers.....	16
VNA Source	18
Signal-separation Hardware	19
Receivers	21
Common VNA Architectures	24
T/R Test Set	24
Full S-parameter Test Set.....	25
4. Vector Network Analyzer Calibration	29
Why Calibrate?	29
Sources of Error in VNA Measurements	30
Types of VNA Calibration	32
Modeling Two-Port Calibration.....	33
Two-Port Calibration Methods	36
SOLT Calibration.....	36
TRL Calibration	38
Tools for VNA Calibration	39
Calibration Considerations.....	41
Advanced Calibration Techniques	42
5. Glossary	43
6. References.....	44

1. Introduction to Network Analyzer Measurements

Before the network analyzer, determining the reflection coefficient of a circuit required you to manually calculate the phase of the reflection coefficient one frequency at a time on a piece of paper. This manual process required you to first find the maxima and minima of the standing voltage wave in a slotted line and repeat this process for all frequencies of interest. It is safe to say that if the vector network analyzer (VNA) had not been created, the high-frequency community would look very different today.

Today, the VNA characterizes high-frequency passive and active devices in their linear mode of operation by measuring their network parameters, called S-parameters, as a function of frequency. Over time, VNAs have been extended in hardware and in capability to also measure noise parameters and nonlinear characteristics, including compression, intermodulation, and hot S_{22} measurements¹. As a result, the VNA evolved into an instrument enabling multiple measurements with a single connection.

VNA Basics

In its simplest form, a network analyzer is an instrument used to measure impedance. At lower frequencies, you can measure impedance with relatively simple tools, including a sine wave generator, a volt meter, a current meter, and a calculator. Using these tools, you can measure the ratio between voltage and current and calculate the resulting impedance, as shown in **Figure 1.1**.

¹ Hot S_{22} measurements are S_{22} device measurements where drive power is provided to the device input while the device is excited at port 2, typically with an offset frequency.

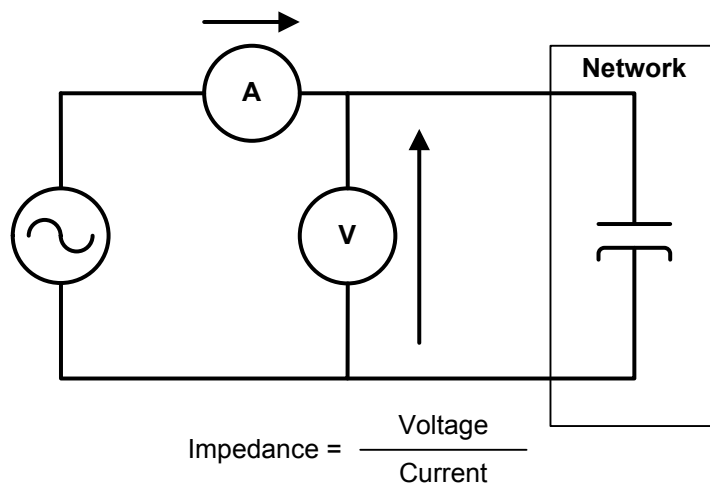


Figure 1.1. Determining Impedance

At RF and microwave frequencies, measurements of voltage and current become more complex. As a result, a VNA uses a more complex design to measure incident and reflected waves. In many ways, the VNA marries the principles of the basic impedance measurement with hardware appropriate for microwave frequencies.

When using a VNA to measure the impedance or the reflection factor, a sine generator stimulates the device under test (DUT). In addition, two receivers take the place of the combination of a volt meter and current meter. These receivers, with the help of signal separation hardware, characterize the response of the device by measuring the phase and amplitude of signals that are both incident to and reflected from the DUT. Finally, calibration capabilities are required to eliminate systematic errors and compute the appropriate ratios (similar to the impedance) necessary to produce S_{11} , which is one part of the S-parameters. The architecture of this simplified VNA is illustrated in **Figure 1.2**. To measure devices with more ports the architecture needs to be extended.

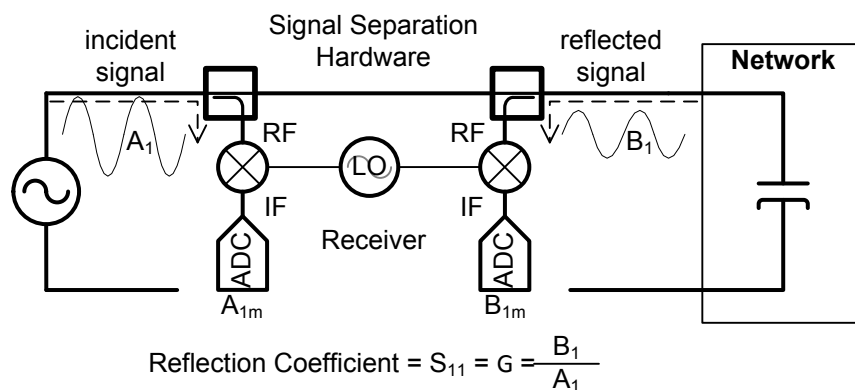


Figure 1.2. Measuring Impedance at Higher Frequencies

Modern network analyzers cover a wide range of frequencies from the low hertz (Hz) to tens of terahertz (THz).

Applications for Network Analyzers

The primary use of a VNA is to determine the S-parameters of a myriad of passive components, including cables, filters, switches, diplexers, duplexers, triplexers, couplers, bridges, transformers, power splitters, combiners, circulators, isolators, attenuators, antennas, and many more. In addition, VNAs can also characterize active devices such as transistors and amplifiers using S-parameters, as long as they are operating in their linear mode of operation. High-frequency devices can have one or two or more ports.

RF and Microwave

S-parameters provide complete insight into the linear behavior of RF and microwave components and are independent of the characteristics of the VNA itself. Through the VNA calibration, impairments of the instrument are completely removed from the measurement. Thus, S-parameters are an extremely accurate representation of the linear behavior of the component under test, describing how it behaves and how it interacts with other devices when cascaded. For this reason, you can use S-parameters in a wide range of applications, ranging from a simple metric to characterize circuit performance to a basic behavior model in electronic design automation (EDA) tools. In the latter case, S-parameters enable the simulation of complete RF circuit chains.

S-parameters are very useful because they are an integral part of different theories, such as filter, transmission line, and amplifier design theories. Also, S-parameters are essential to performing proper noise parameter measurements. While engineers are increasingly facing nonlinear problems, they mainly design for linear modes of operation, as the arsenal of engineering tools is still based on linear theories.

Other Uses

Over time, VNAs have even found their way into applications besides the pure S-parameter measurement capability. Modern VNAs are being used for the characterization of the dielectric properties of materials (both liquids and

solids) and even the human brain. Because the VNA can measure the reflection factor as function of frequency, the results it produces can be transformed into the time domain. As a result, one common application of modern VNAs is cable fault detection, which takes advantage of time-domain reflectometry (TDR).

Note that in addition to becoming the tool of choice for RF and microwave circuits, the VNA has also become widely used in signal integrity applications. In fact, the relatively high dynamic range of a VNA, in comparison to the traditional oscilloscope, makes it a useful diagnostic tool.

Additionally, electromagnetic imaging is one of the growing fields where VNA technology is fast becoming an essential tool. Similar to the detection of a cable fault, it is possible to reconstruct images from reflected waves by performing a form of scanning. VNAs have penetrated into medicine to help visualize, for example, breast tumors. Also in the security realm, microwaves are used for body scans, and some applications are taking advantage of a VNA as a basic instrument.

The growing ranges of applications for which a VNA can be used have driven their design into new form factors beyond the traditional box instrument. For example, because of the need to conduct VNA measurements in the field, VNAs are also available in portable formats. In addition, because applications such as electromagnetic imaging require significant software, the VNA is increasingly connected to the PC. In fact, in the case of VNAs in the PXI form factor, the network analyzer becomes a part of the PC itself.

History of Network Analysis

In the early 1950s, Rohde & Schwarz achieved a breakthrough in network analysis with a product called the Zg diagraph [1]. Until then, network analysis had been limited to manual measurements of standing waves in a slotted line. The Zg covered a frequency range from 30 MHz to 300 MHz, though the frequency range was later extended to 2.4 GHz. This instrument became the first complex network analyzer on the market that could directly measure the phase of S-parameters and display the S-parameters on a Smith Chart. The Zg was used to measure radios, TV antennas, and some telecommunication cables.

Between the 1950s and early 1960s, three other companies introduced a form of network analyzer based on the homodyne principle. In 1965, Wiltron (later acquired by Anritsu) introduced the Wiltron 310 Impedance Meter. During that time, other companies such as Hazeltine Corporation and Rantec Corporation [2] created some of the world's first network analyzers.

Improvements in the VNA Design

In 1966, Hewlett Packard introduced the 8405A vector voltmeter, which used step recovery diode (SRD) technology. Using SRD technology, you could measure both amplitude and phase difference between two signals up to 1 GHz. The SRD enabled sampling techniques to down convert high-frequency signals into manageable low frequencies. Note that the SRD technology in the HP 8405A was leveraged from the sampling oscilloscope product HP 185/7 that was brought to the market around the same time [3].

By combining the 8405A with couplers or bridges, you could measure the amplitude and phase difference of incident, transmitted, or reflected waves of a DUT at one frequency. Next, you would use a printed-paper Smith

chart to plot an impedance chart as function of frequencies. It is important to note that at that time, sophisticated calibration techniques were not available to remove the vast majority of measurement uncertainties. Instead, you could only compensate for the difference in length between the various measurement paths.

In 1967, Hewlett Packard introduced the 8410A network analyzer, which contained the ability to directly plot a Smith chart on the instrument's screen. One year later, HP introduced the world's first automatic network analyzer, the HP8540A. This instrument was a 2-cabinet system that included the HP 8410 network analyzer and a HP 2001A "computer". The computer contained a maximum of 32 K words of memory, a 5 MB disk drive, a digital cassette tape, and a paper tape reader or puncher. The first version had a teletype as human interface, though a later version, the HP 8542A, had a "real" CRT display and improved software [3]. In many ways, the HP 8542A revolutionized RF and microwave component and system design. For the first time, it automated the measurement process by allowing users to perform multi-band frequency sweeps by measuring and plotting impedance and transfer characteristics on a Smith Chart.

Meanwhile, scattering parameters, or S-parameters, were quickly becoming the standard format to describe passive and active devices in their linear mode of operations at high frequencies. At first, the network analyzer could only measure the transmission and reflection wave. Thus, you were forced to reverse the DUT to measure these quantities at the opposite device port. Soon, more sophisticated test sets became available that could measure the full S-parameter set of a two-port device. Thanks to the automation of the measurements, it became possible to develop and deploy more sophisticated calibration techniques to eliminate the systematic errors of the measurement system.

In contrast to VNAs, scalar network analyzers measure standing wave ratios, reflection and transmission coefficients only as scalar quantities rather than as vector quantities. The first realizations of a scalar network analyzer date from 1954 and were developed in parallel with vector network analyzers. At the time, scalar network analysis was more adequate for production test than it was for design work. As a result, scalar network analyzers were sold in much larger quantities than VNAs.

Automated VNAs

While automated vector network analyzers were available since the late 1960s, it was not until 1985 that a fully automated, relatively compact vector network analyzer first became available. The first of this category was the 8510, which was introduced by Hewlett Packard [4]. The HP 8510 network analyzer took major advantage of the microprocessor technology that had become available. Using microprocessor technology, the instrument could perform frequency sweeps and labor-intensive calibration techniques relatively quickly. For the first time, a vector network analyzer was capable of giving almost real-time feedback while tuning a DUT, representing the measurements in both the time domain and the frequency domain. Soon after the 8510 introduction, Wiltron introduced the 360 Network Analyzer [5], which benefited from similar technological advancements.

In the 1980s, engineers conducted much research on the use a set of scalar detectors or power measurements to derive the amplitude and phase information of the incident and reflected waves. One example of this is a six-port reflectometer with three scalar detectors at each port to determine the coefficients-parameters of the DUT. Using only power detectors, the setup was highly simplified compared to a VNA. However, this approach never took off commercially.

The Modern Era

In the 1990s, growth in high-frequency systems such as cellular telephones and cable television created increasing demand for instrumentation that was capable of characterizing RF and microwave circuits. As the demand for VNAs increased, engineers began to enjoy a broader range of price points and performance capabilities for VNAs. As a result, the market saw various low-end VNAs being commercialized and deployed into test and manufacturing environments while high-end VNAs continued to target research and development. Meanwhile, the scalar network analyzer had been replaced by the VNA in many applications due to VNA benefits such as a reduced cost and technical advantages.

One key innovation in network analysis technology in the early 2000s was the evolution of the VNA towards measuring devices in their non-linear range of operation. As RF and microwave components have been pushed to their limits, an increasing number of engineers are required to characterize the nonlinear behavior of the devices under test. The need for measurements beyond S-parameters, which only describe linear behavior, has imposed a new requirement on the VNAs to measure additional device characteristics and requires the presence of the combined sources and receivers. As a result, a wide range of vendors such as Agilent Technologies, Rohde & Schwarz, and Anritsu began to develop non-linear network analysis technology.

Given the challenges of modern network analysis, the complexity of network analysis technology has elevated the role of software in VNA systems. As a result, PC-based form factors are an increasingly popular choice for VNAs. In 2011, National Instruments introduced the first PXI-based vector network analyzer, the NI PXIe-5630, which was followed one year later by the NI PXIe-5632 with full S-parameter measurement capability. A NI PXIe-5632 VNA is shown in a PXI system in **Figure 1.3**.

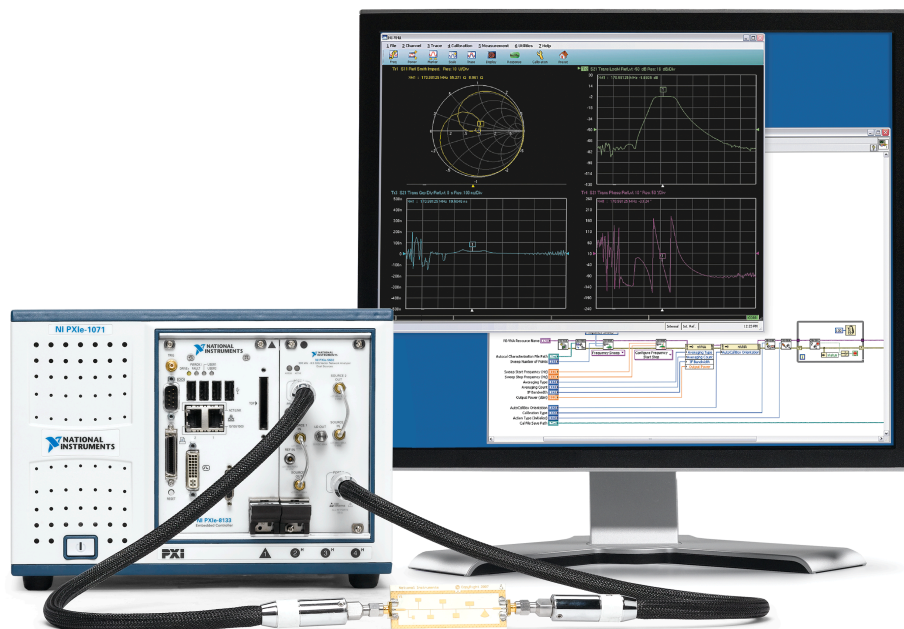


Figure 1.3. PXI VNA System

2. Understanding Wave Quantity Measurements

S-parameters describe the complete linear behavior of a high frequency device or circuit in linear mode of operation. The main function of a vector network analyzer is to provide S-parameters as a function of the frequency. Although modern VNAs return S-parameters with a push of the button (after proper calibration), an S-parameter is not a quantity that can be measured directly. Instead, the S-parameters are the result of solving a set of equations after measuring ratios of wave quantities while exciting the device at its various ports.

Measuring Wave Quantities

To properly understand S-parameters, you must first understand the theory and process of measuring wave quantities. Consider, for example, a load connected to a sine wave generator through a piece of cable, as illustrated in **Figure 2.1**. The sine wave generator has a certain source impedance, and the electrical power travels across the wire from the source to the load as the combination of an electric and magnetic field. The propagation speed (v_p) depends on the dielectric properties of the medium and is a fraction of the speed of light (c). The wavelength, λ , is defined as a function of the propagation speed (v_p or c) and the sine wave generator frequency (f_0) in **Equation 2.1**.

$$\lambda = \frac{v_p}{f_0}$$

Equation 2.1. Wavelength as a function of frequency and propagation speed

When the frequency (f_0) is low, the wavelength is large, and the length of the cable is negligible compared to the size of the wavelength. As a result, the measured voltage and current are independent of the location on the cable. This situation is illustrated in **Figure 2.1**, and the circuit is referred to as being a lumped element circuit or lumped circuit.

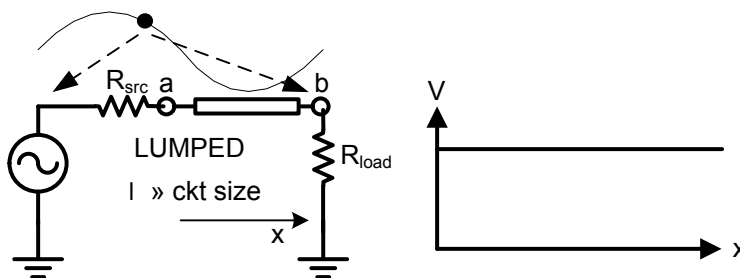


Figure 2.1. A Lumped Circuit

When the frequency (f_0) of the source increases, however, the wavelength is reduced. Thus, as frequency increases, the wavelength eventually becomes similar in size or even smaller than the length of the cable.

In a scenario where the wavelength of the signal is similar or smaller in size to the length of the cable, the measured voltage and current will depend on the position, as shown in **Figure 2.2**. Thus, measuring the voltage with a voltage

probe is invalid because the result will be dependent on the probe's position. In this scenario, the circuit must be treated as a distributed element circuit rather than as a lumped circuit.

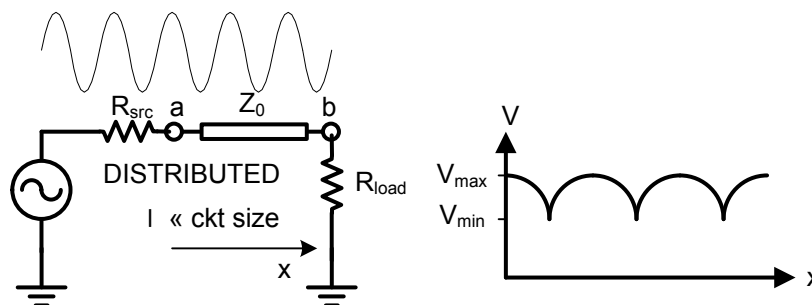


Figure 2.2. A Distributed Circuit

Analysis of a distributed circuit is more complex and involves the use of transmission line theory. Using transmission line theory, you can think of the electrical power traveling along the line starting from the source. While traveling on the line, a certain voltage (E-field) and current (H-field) relation is imposed by the electrical properties of the line. In fact, the cable itself will behave such that it is characterized by an inherent impedance that does not change as long as the properties of the line or cable do not change. This impedance is called the characteristic impedance (Z_0).

As the electrical power hits the termination (R_{load}), the voltage to current relationship is now imposed by the impedance of the load. Under the condition where the load impedance is equal to the characteristic impedance, the power is fully absorbed. To keep the reasoning simple, real impedances are assumed here. If the load impedance is different from the characteristic impedance, the ratio of voltage and current will change at the point where the transmission medium occurs. As a result, the load will not absorb all the power, resulting in a portion of the power traveling back towards the source. Given this characteristic, you can refer to the system as having both an incident and reflected wave.

The ability to separate incident and reflected waves is a critical element of network analysis not only because it identifies how much power is absorbed by the load but because it reveals information about its impedance. As a result, one of the most important functions of a VNA is to separate these incident and reflected waves by means of either couplers or bridges in order to measure each of these waves independently. Thus, while measurements from simple volt and current meters are able to determine impedances and transfer functions of circuits at lower frequencies, measuring similar characteristics at higher frequencies requires the measurement of incident and reflected waves. In fact, the VNA is able to measure the amplitude and phase differences between individual waves, using one of the waves as a reference. Modern VNAs are even able to measure these quantities in an absolute way instead of in a relative way. **Figure 1.2** clearly shows the signal separation hardware, which separates incident and reflected waves using directional couplers.

Introduction to S-parameters

Determination of scattering parameters, or S-parameters, is the basic measurement capability of vector network analyzers. The S-parameters describe the magnitude and phase relationship between incident and reflected waves and are numbered according to where a wave originates from and where it propagating to. The term *scattering parameters* comes from the “Scattering Matrix” described in a 1965 IEEE article entitled, “Power Waves and the Scattering Matrix.” Today, S-parameter results, such as S_{21} for example, are displayed on the VNA display screen with the simple push of the button. Moreover, the instrument is able to apply appropriate calibration techniques to remove a wide range of inherent impairments of the instrument itself.

In order to better understand the S-parameters in an intuitive way, it is useful to first consider an example from the field of optics—the behavior of light through varying transmission mediums. Suppose a light beam is shining on a lens. In this example, some light will pass through the lens, undergoing the influence of the lens. However, some light will also be reflected back by the lens towards the direction of the light source, as illustrated in **Figure 2.3**.

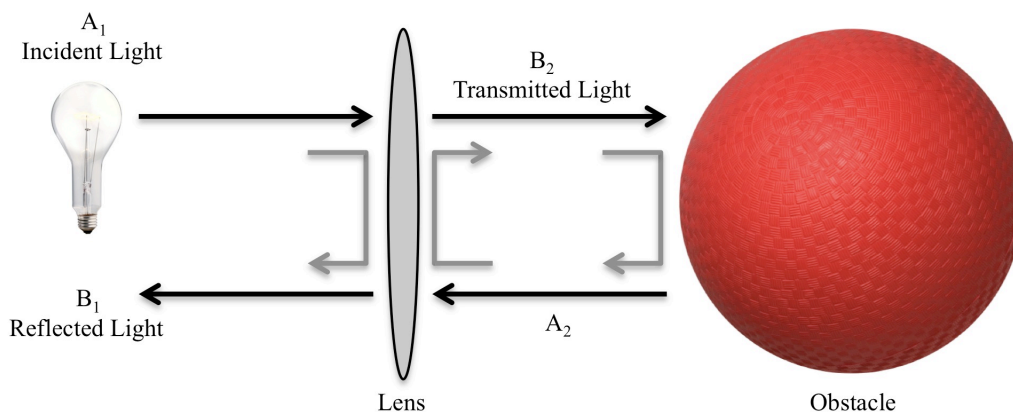


Figure 2.3. Optical Equivalent of S-parameter Behavior

In addition, it is also useful to investigate what happens once the light passes through the lens. If there is no obstacle behind the lens, the light will continue to travel forever. If an obstacle exists, however, part of the transmitted beam is reflected back towards the lens. This light will be partially reflected back towards the obstacle and will be partially transmitted through the lens. Thus, the reflection from the obstacle is added to the initial reflection from the light source. As a result, the reflected wave before the lens consists of two parts: the original beam reflected by the lens and a part of the reflection from the obstacle. Similarly, the transmitted wave consists of the original beam transmitted through the lens and partially the reflected beam from the obstacle, also shown in **Figure 2.3**.

The lens and light analogy translates easily to the understanding of the way waves travel through an electrical circuit. As illustrated in **Figure 2.3**, both the B_1 and the B_2 wave can be described as a combination of both the A_1 wave and the A_2 wave. These relationships can be formally represented as S-parameter equations for a two-port device as shown in the following equations.

$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

Equation 2.2. 2-port S-parameters Equations

Figure 2.4 shows a two-port device under test (DUT) along with the corresponding four S-parameters and the four waves.

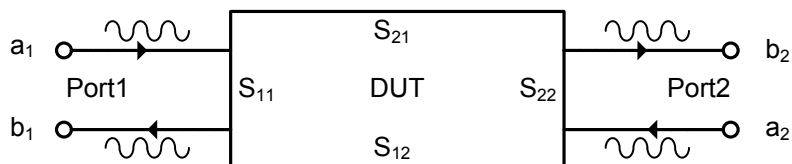


Figure 2.4. Two-port S-parameter Model

It is important to realize from **Equation 2.2**, that a wave leaving the DUT (b_1 or b_2) is a linear combination of the waves entering the DUT (a_1 or a_2). As a result, terminating the DUT in the characteristic impedance (Z_0) allows the extraction of the individual S-parameters. For example, when the characteristic impedance (Z_0) is equal to 50 Ohms, and if a 50 Ohm termination is present at port 2, a_2 is reduced to zero, resulting in equations for S_{11} and S_{21} . This principle can be applied in the reverse direction as well. By setting a_1 to zero, you can obtain equations for S_{22} and S_{12} , as shown in **Equation 2.3**.

$$S_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0} \quad S_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0}$$

$$S_{21} = \left. \frac{b_2}{a_1} \right|_{a_2=0} \quad S_{22} = \left. \frac{b_2}{a_2} \right|_{a_1=0}$$

Equation 2.3. Formal Definitions of S-parameters

Note that a common notation is to use the subscripts m and n for the general S-parameter (S_{mn}), where m is the receiver port and n is the source port. For example, S_{12} is the transmission coefficient for a wave sourced at port 2 and received at port 1. Keep in mind that these ratios are complex-valued, and the S-parameter will carry both a magnitude and phase component as a function of frequency.

Finally, the equations in **Equation 2.2** can be converted into matrix notation and generalized for any number (n) of ports as illustrated in **Equation 2.4**.

$$\mathbf{b} = \mathbf{S} \cdot \mathbf{a}$$
$$\mathbf{S} = \begin{bmatrix} S_{11} & \dots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \dots & S_{nn} \end{bmatrix}$$

Equation 2.4. S-parameter Equations in Matrix Notation

In theory, determining the S-parameters using a VNA is a straightforward process. In practice, however, it is impossible to terminate the port opposite to the driving port exactly in the characteristic impedance (Z_0), resulting in the presence of a reflected wave at all times. Given this challenge, a solution is to first connect the source to port 1 and terminate port 2, which allows the VNA to measure ratios b_1/a_1 , b_2/a_1 and a_2/a_1 as a function of frequency. Next, for a two-port device, the VNA switches the source to the output and terminates the input port of the DUT. With this configuration, the VNA can now measure ratios b_1/a_2 , b_2/a_2 and a_1/a_2 . Using basic linear algebra, it is possible to calculate the S-parameters as a function of frequency.

In practice, a VNA will generally measure the incident and reflected waves through a series of couplers or bridges, referred to as directional devices. The directional device is able to separate the incident from reflected waves. Using this measurement practice, the resulting measurements will be quantities that are subject to imperfections of the VNA, such as the coupling factor and directivity of the directional devices. Thus, more complicated calculations are needed in order to determine calibrated S-parameters starting from a set of raw quantities.

3. Network Analyzer Architectures

Historically, scalar network analyzers were a common tool used to characterize a network using only the magnitude of the signal. However, as network analysis technology matured, the development of digital components, such as analog-to-digital converters (ADCs), greatly simplified the design of vector network analyzers. As a result, most modern VNAs are capable of measuring both scalar (magnitude) and vector (magnitude and phase) information about a signal.

Scalar Network Analyzers

Scalar network analyzers typically capture a broadband signal and convert it to DC or low frequency AC in order to measure the power of the signal. Examples of the hardware used to accomplish this include diodes and thermoelectric devices. The main advantage of scalar network analyzers is that the hardware required for downconversion and power detection is relatively simple and inexpensive. In addition, because the detectors are broadband devices, it is unnecessary to re-tune the receiver to measure power at a different frequency. Thus, performing a frequency sweep is as simple as re-tuning the frequency of the microwave source and measuring the power at each frequency step.

Due to their relatively simple architecture, scalar network analyzers are capable of relatively fast frequency sweeps. Note that some scalar network analyzers simplify the hardware even further by removing the reference sensor. In these designs, however, a slightly more complicated measurement sequence is required. A simplified block diagram of a scalar network analyzer is shown in **Figure 3.1**.

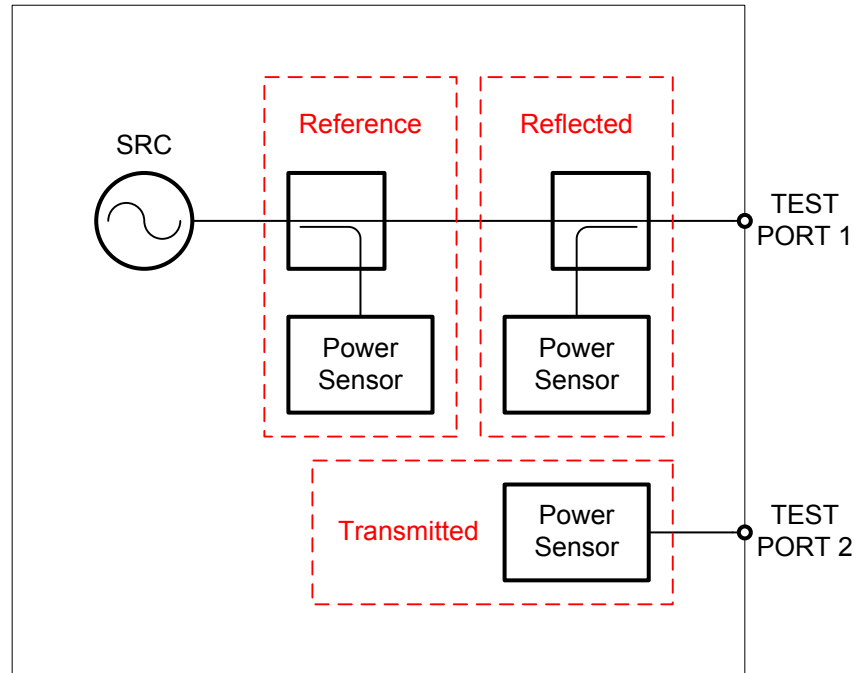


Figure 3.1. Simplified Scalar Network Analyzer Block Diagram

Scalar analyzers make transmission or insertion loss measurements by using a signal source that sweeps repetitively over the frequency range of interest. If the reference detector (as seen **Figure 3.1**) is not available, you can calculate the transmission coefficient as the power ratio of the transmitted signal with and without the device under test (DUT) in the signal path. In this case, two sweeps are required to characterize the device. If a reference detector is present, you can calculate the transmission coefficient as the ratio of the transmitted power to the incident power as measured by the reference detector. For reflection measurements, a directional device, either coupler (as seen in **Figure 3.1**) or bridge, is necessary to detect the signal reflected by the DUT. A ratio of the reflected signal power over the incident signal power at the device is the reflection coefficient.

While scalar network analyzers have the benefit of simplicity of design, they are also prone to inherent challenges. For example, broadband detectors are susceptible to spurious tones and broadband noise. In addition, because the calibration is scalar in nature, it is not as accurate as full vector calibration. Due to their lack of selectivity, scalar network analyzers tend to have limited dynamic range compared to vector network analyzers.

Vector Network Analyzers

Vector network analyzers generally use full heterodyne receivers to measure both the phase and magnitude of signals and are often significantly more complex than scalar network analyzers. Measurements made with vector network analyzers are often more accurate, and the narrowband nature of the receivers provides better rejection of

broadband noise and spurious tones, allowing for improved dynamic range. Furthermore, calibration can use more complex error models, which provide greater accuracy.

Due to the complexity of the heterodyne receiver architectures of vector network analyzers, these instruments generally perform frequency sweeps more slowly than broadband scalar network analyzers. In addition, the added complexity often makes them more expensive.

The fundamental principle of a vector network analyzer is to measure the amplitude and phase of both incident and reflected waves at the various ports of the DUT. The general design of a VNA is to stimulate an RF network at a given port with a stepped or swept continuous wave (CW) signal and to measure the travelling waves, not only at the stimulus port but at all the ports of the network terminated with specific load impedances, typically 50 Ohms or 75 Ohms. A typical but simplified VNA architecture is illustrated in **Figure 3.2**.

As observed in **Figure 3.2**, the block diagram of a VNA can be decomposed into different elementary blocks. First, note that one (or more) signal source with controllable frequency is used to drive the DUT. In addition, the test set includes two test ports. Each test port contains some signal separation hardware to split out the incident and reflected travelling waves. Optionally, the test set can contain switches to route the signal source to the different test ports and terminates other ports with specific load impedances. Finally, one or more receivers (marked ref 1, ref 2, test 1, and test 2 in **Figure 3.2**) capture both the phase and magnitude of the RF signals before processing them.

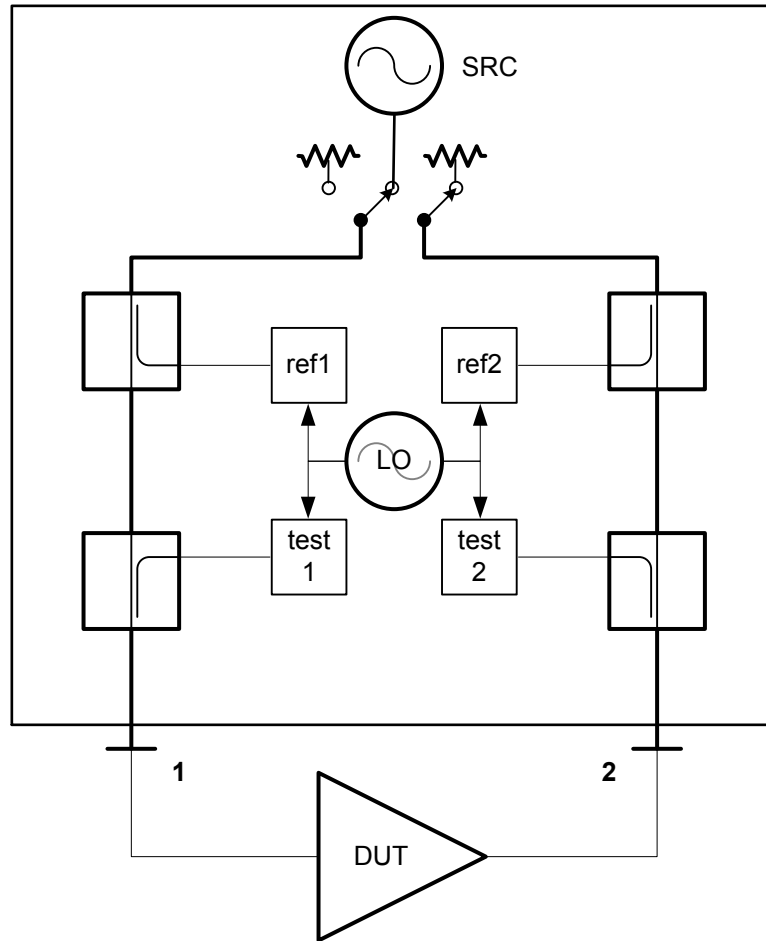


Figure 3.2. Typical VNA Block Diagram

VNA Source

The source of a VNA provides the stimulus that is used to characterize the response of a DUT. The source must be tunable for output frequency and power, and its performance is dictated largely by the types of measurements the VNA will make. The output signal is typically a sine tone for single-frequency S-parameter measurements. However, some advanced applications require multi-tone or modulated signals to provide the most complete characterization of a device's behavior.

In general, the most significant requirement of a VNA signal source is the speed and accuracy of its tuning capabilities. For most measurements, the source will rapidly sweep across frequencies to collect the frequency response of a device. However, for measurements such as gain compression, the source will sweep across output power at a fixed CW frequency.

When measuring the linear behavior of the device, the requirement for spectral purity of the output signal is more relaxed than general-purpose signal generators. Because the frequency of the source is known, the VNA's receivers can selectively tune to the appropriate frequency. Thus, spurious tones and other undesirable frequency components can often be present without affecting the accuracy of a measurement. One caveat is that phase noise of the source signal contributes to the noise measured at the output. As a result, phase noise cannot be neglected entirely.

The spectral content of the source's output becomes much more critical when making nonlinear measurements. Multi-tone measurements, such as intermodulation distortion, are negatively impacted by undesirable signals, as are frequency-translation measurements.

Signal-separation Hardware

The signal-separation hardware is a key element of vector network analyzers. Each test port of the VNA makes use of signal-separation hardware, such as a directional coupler or bridge, to separate incident from reflected waves travelling the test port. Each wave is then captured with a dedicated receiver, allowing the instrument to measure the phase and magnitude of both waves independently.

As illustrated in **Figure 3.3**, a generic directional device separates the incident wave from the reflected wave, while a small portion of the incident wave travelling along the through path is diverted into the coupled path. Note that forward coupling is shown in **Figure 3.3**.

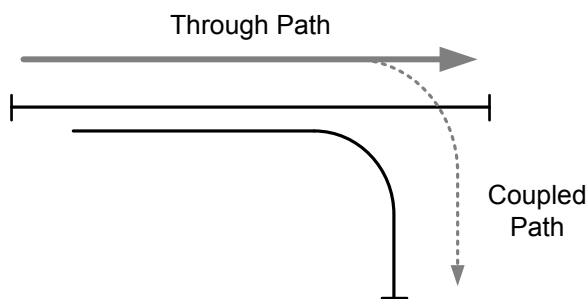


Figure 3.3. Directional Device

Key characteristics of a directional device are its insertion loss, coupling factor, isolation, and directivity. Insertion loss is a metric of the loss of signal power when it is traveling through the directional device (through path). A high quality directional coupler can present insertion loss as low as 1 dB up to microwave frequencies. You can calculate insertion loss using **Equation 3.1**.

$$\text{Insertion Loss (dB)} = -10 \log_{10} \left(\frac{\text{output power}}{\text{input power}} \right)$$

Equation 3.1. Calculating Insertion Loss from Input and Output Power

With any coupling device, only a portion of the incident signal going through the directional device is diverted to the coupled path. The coupling factor is a metric that describes the ratio of the coupled power to the total forward signal

power. In typical VNAs, the coupling factor varies anywhere from 10 dB to 20 dB, and the coupling factor is calculated using **Equation 3.2**.

$$\text{Coupling Factor (dB)} = -10 \log_{10} \left(\frac{\text{forward coupled power}}{\text{forward power}} \right)$$

Equation 3.2. Calculating Coupling Factor

Ideally, a radio frequency signal travelling in the reverse direction of a directional device should not appear in the coupled path. In practice, however, some power traveling in the reverse direction is diverted in the coupled path due to the finite isolation between each port of the directional device. Isolation is one of the most critical characteristics of a coupling device, and performance in the range of 30 dB to 40 dB is required for accurate network analysis measurements. Isolation is expressed in **Equation 3.3**.

$$\text{Isolation (dB)} = -10 \log_{10} \left(\frac{\text{reverse coupled power}}{\text{reverse power}} \right)$$

Equation 3.3. Calculation of Isolation

Directivity is a metric that describes the overall performance of the ability of a directional device to separate out forward and reverse signals traveling through the directional device. This metric combines elements of isolation, coupling factor, and insertion loss and is expressed in **Equation 3.4**.

$$\text{Directivity (dB)} = \text{Isolation (dB)} - \text{Coupling Factor (dB)} - \text{Insertion Loss (dB)}$$

Equation 3.4. Directivity As a Function of Isolation, Coupling Factor, and Insertion Loss

The directivity of a directional coupler has a significant impact on the measurement accuracy of a network analyzer. A large value for directivity implies a high performance VNA. In essence, directivity accounts for the amount of measurement error that is introduced by a wave traveling in the reverse direction of the coupling device.

Challenges with Poor Directivity

In scenarios such as a return loss measurement, the directivity error can cause artifacts in the measurement result. For example, consider what happens when the directional device diverting the travelling wave reflected back from the DUT has poor directivity. When both the reflected wave and the directivity error vector (introduced by the incident wave) are in-phase, the electromagnetic waves combine at the coupled port level, resulting in a measured return loss that is worse than the actual DUT performance. Alternatively, the directivity error vector can cancel out the captured reflected wave, resulting in return loss measurement that is actually better than the DUT performance. In addition, the phase difference between the coupled waves and directivity error vector is frequency dependent. Thus, poor directivity manifests itself with typical ripple patterns in many measurements of DUT response. **Figure 3.4a** and **Figure 3.4b** illustrate the impact of directivity on a measured response both where the directivity error is combined and where the directivity error cancels out.

The finite directivity of a VNA also has a direct impact on the precision of the instrument. Though the raw directivity can be compensated for with calibration (see **Vector Network Analyzer Calibration**), the finite directivity of the couplers will still impact the stability of the calibration. In fact, as the estimated directivity error

vector is subtracted from the raw DUT measurement during correction, a small drift in the actual directivity error vector may significantly influence the final result.

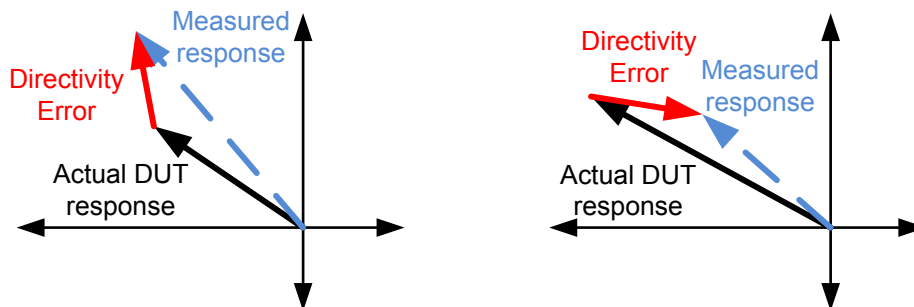


Figure 3.4a and Figure 3.4b. Impact of Directivity on Measured Response:

(a) Directivity Error Adds Up, (b) Directivity Error Cancels Out

The concept of directivity has implications to connectivity of the DUT. Because a signal separation device is required to separate a relatively low-power reflected wave from a relatively high-power incident wave, the directivity of the coupling device is one of the largest sources of measurement uncertainty. This uncertainty is further exacerbated when significant loss is added between the output of the directional device and the DUT through the use of poor quality test cables or attenuators. Because loss between the coupler and the DUT will further attenuate the reflected wave, the raw directivity of the directional device will be degraded by twice the attenuation value.

Note that both directional couplers and the directional bridges are common signal separation devices used in vector network analyzers. The choice between one or the other is frequently determined by the frequency range of the instrument. By design, directional bridges have higher insertion loss than directional couplers, but the former can operate at low frequency (down to kilohertz) while maintaining compact physical dimensions. At higher frequencies, directional couplers tend to have better performance than directional bridges, both in terms of insertion loss and directivity, and are generally preferred in microwave VNAs.

Receivers

Measurements of the phase and magnitude of a traveling wave in a VNA are determined by a receiver, which translates analog signals into the digital domain. **Figure 3.5** illustrates a generic block diagram of a single-stage downconversion heterodyne receiver.

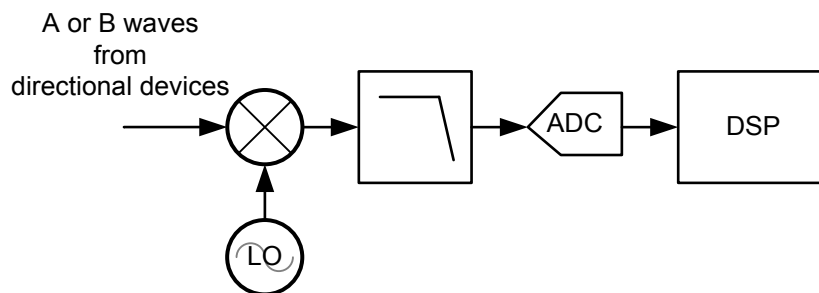


Figure 3.5. Generic VNA Receiver Block Diagram

Digitization and processing of the captured incident and reflected waves first require downconversion of the RF signal to an intermediate frequency (IF). As illustrated **Figure 3.5**, a local oscillator (LO) is used to drive the receiver's downconverter.

Although the receiver would achieve the best phase noise and spur performance using only the fundamental frequency of an LO, this design is often impractical. In fact, a broadband LO source able to cover the entire bandwidth of a microwave VNA is often challenging and costly in practice. As a result, it is more common to use harmonic-conversion process to limit the required frequency and power range of the LO source.

Sampling Receivers

The downconversion process from RF to IF can occur using either high-frequency mixers or sampling downconverters, or samplers. Historically, VNA receivers were mainly based on sampling downconverters due to simplicity and cost. In fact, such a downconverter requires only a relatively low frequency LO drive. The LO frequency limitation is then compensated by using a highly nonlinear circuit, such as shockline or step recovery diode, to generate a very short pulse with very high harmonic content. The sampling diode, acting as a switch, is then driven by this low frequency periodic pulse. As a result, the LO is effectively converted into a multi-tone stimulus, as shown in the following figure, which allows for simultaneous downconversion of multiple frequency bands at once.

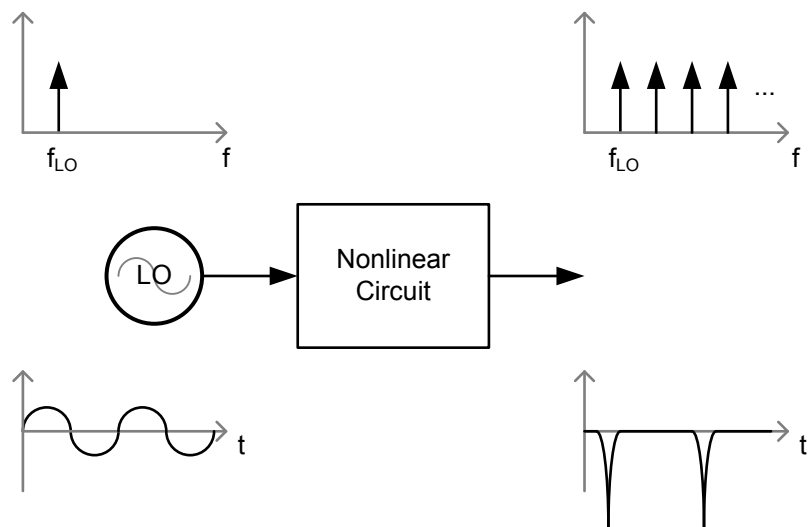


Figure 3.6. LO Source Converted into a Multi-tone Stimulus

Using sampler technology, the resulting downconversion process is then able to cover RF frequencies much higher than the LO drive frequency. Also, as the order of nonlinearity introduced by the pulse generator is very high, the harmonically related tones can have almost flat power over the full downconverter bandwidth. This results in about the same conversion efficiency for each LO harmonic. Unfortunately, the process of simultaneously downconverting multiple bands negatively impacts the dynamic range of the receiver. In this case, broadband noise and potential spurious signals are also downconverted into the same IF bandwidth.

Modern VNA Receiver Design

Today, almost all VNAs use mixers to perform the RF to IF downconversion since broadband LO sources have become less expensive. While several high-frequency mixer topologies exist, the most commonly used design is the doubly balanced diode-ring mixer [6]. These mixers use ring diodes, driven by the LO source, to perform the RF to IF downconversion.

Because broadband LO sources that cover the full bandwidth of microwave VNAs are quite expensive, the LO signal is typically clipped to look more closely like a square wave. Clipping results in generation of odd harmonically related tones. The captured RF signal is then mixed with the appropriate harmonic to produce the desired IF signal. The unused harmonics still contribute to the IF output, but in contrast with samplers, the power of the generated LO harmonics decays rapidly with the harmonic order N (proportional to $1/N$ for an ideal square wave). Because downconversion at higher harmonic order is less efficient, the IF bandwidth is less impacted by unwanted broadband noise and spurious signals. As a result, harmonic mixers typically present better dynamic range than sampling downconverters.

The downconverted IF signal is further converted and detected in the IF processing chain. Modern network analyzers use a high-speed analog-to-digital converter (ADC) to perform direct sampling of the IF signal. Also, an

anti-aliasing filter is placed just before the ADC input for proper IF image rejection. Note that signal conditioning hardware, generally an amplifier, can be used to optimize the signal-to-noise ratio in the ADC.

Finally, the digitized IF signal is processed using various digital signal processing (DSP) techniques to improve the ADCs as well as digitally downconvert, decimate, and filter the IF signal to retrieve the magnitude and phase (or real and imaginary part) of the measured signal. The results are then sent to the main computer processing unit (CPU) to apply error correction and calculate the DUT S-parameters or other derived quantities before displaying them.

Common VNA Architectures

Today, modern network analyzers rely on a combination of receivers and coupling devices to perform S-parameter measurements. In general, the theory of operation for both VNA architectures are similar, though each design will have a varying degree of complexity. The following sections discuss two VNA architectures: the transmission/reflection (T/R) test set and the full S-parameter test set.

T/R Test Set

One of the simplest network analyzer test sets is the transmission/reflection (T/R) test set, which uses one source fixed to port one to measure forward S-parameters, as shown in **Figure 3.7**. Because the source is fixed, the reverse S-parameters may not be measured without reversing the orientation of the DUT using either external switching or physical disconnection and reconnection, both of which can be difficult to calibrate for. However, the reduced complexity of the T/R test set makes it a less expensive option than a full two-port test set when measuring reciprocal devices or when only the forward measurements are necessary.

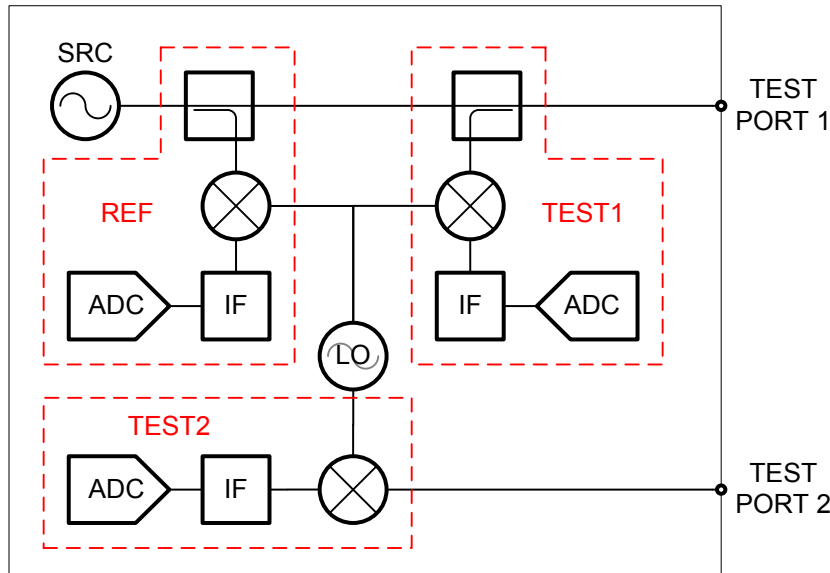


Figure 3.7. Transmission/Reflection Network Analyzer Block Diagram

As shown in **Figure 3.7**, the T/R test set contains three receivers for measuring the components of the forward S-parameters. The reference receiver (REF) measures the incident power from the source, a_{1F} . The test receivers (TEST1 and TEST2) measure the forward reflection and transmission parameters, b_{1F} and b_{2F} , respectively. Using these measurements, S_{11} and S_{21} are calculated.

In the T/R VNA architecture, the test receiver on port 2 is usually directly connected to the test port, in contrast to the coupled receivers on port 1. As a result, T/R test sets can often have higher dynamic range of S_{21} measurements, because the receiver is not subject to the coupling factor in the port 2 receiver path. However, this design may also lead to problems with load matching on port 2. Because a load mismatch cannot be fully corrected during user calibration, the resulting measurements will be affected by an uncorrected return loss at the reference plane of test port 2.

Full S-parameter Test Set

In contrast to the T/R test set, full S-parameter VNA architectures introduces an RF stimulus at each test port of the VNA. In its simplest version, shown in **Figure 3.8**, a full two-port VNA contains three receivers, an RF source and a transfer switch to route the RF stimulus to either test port 1 or 2. With the switch in the port-one switch position, the VNA measures forward reflection and transmission parameters a_{1F} , b_{1F} , and b_{2F} , and in the port-two switch position, it measures the reverse reflection and transmission parameters a_{2R} , b_{2R} , and b_{1R} without the need to physically disconnect and reverse the DUT.

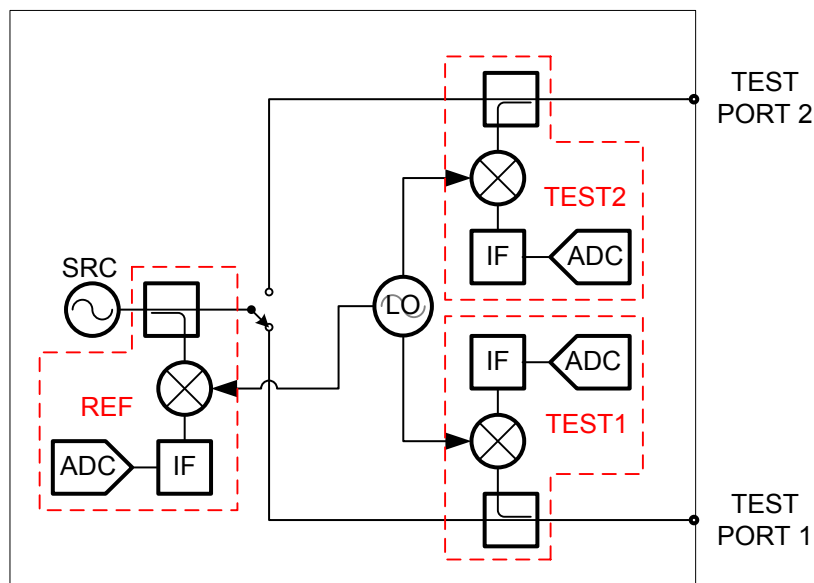


Figure 3.8. Simple VNA Using One Source and Three Receivers

As observed in **Figure 3.8**, the port 1 reference receiver (REF1) measures the forward and reverse reference parameters a_{1F} and a_{2F} . The port 2 test receiver (TEST2) measures the forward transmission parameter, b_{2F} , and the reverse reflection parameter, b_{2R} . The port 1 test receiver (TEST1) measures the forward reflection parameter, b_{1F} , and the reverse transmission parameter, b_{1R} .

In the simplest VNA architecture, the transfer switch is part of the network analyzer measurement path. As a result, VNA calibration accounts for the switch variation. However, slight differences in the two switch positions can lead to measurement uncertainty. In addition, switch contacts may wear over time, requiring more frequent VNA calibration. To resolve this, in alternative VNA architectures, the switch has been moved directly at the source output. Here the forward and reverse reference parameters are measured using two reference receivers, REF1 and REF2 respectively, as shown in **Figure 3.9**.

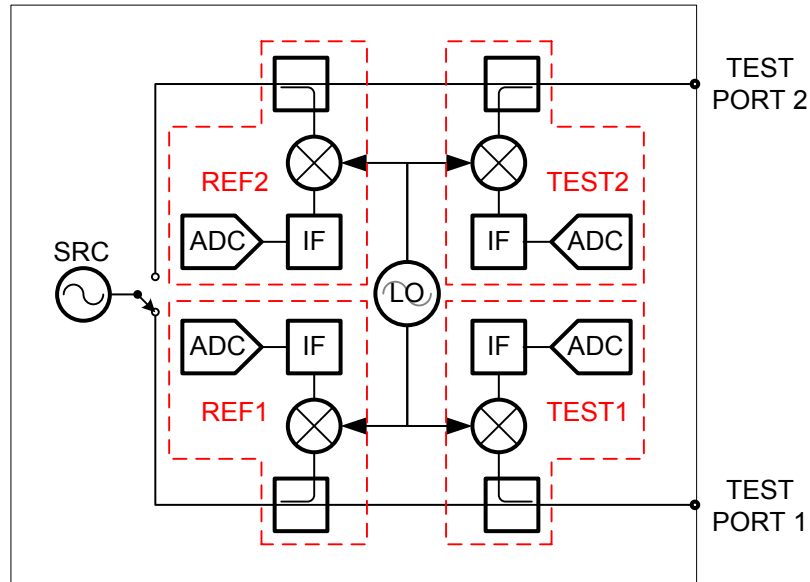


Figure 3.9. VNA Using One Source and Four Receivers

An alternative solution is to eliminate the switch entirely and simply use one RF source per port. This approach, as illustrated in **Figure 3.10**, provides the best performance but at the highest cost. Note that this architecture allows the VNA to perform advanced measurements beyond S-parameters, such as hot S22 measurements.

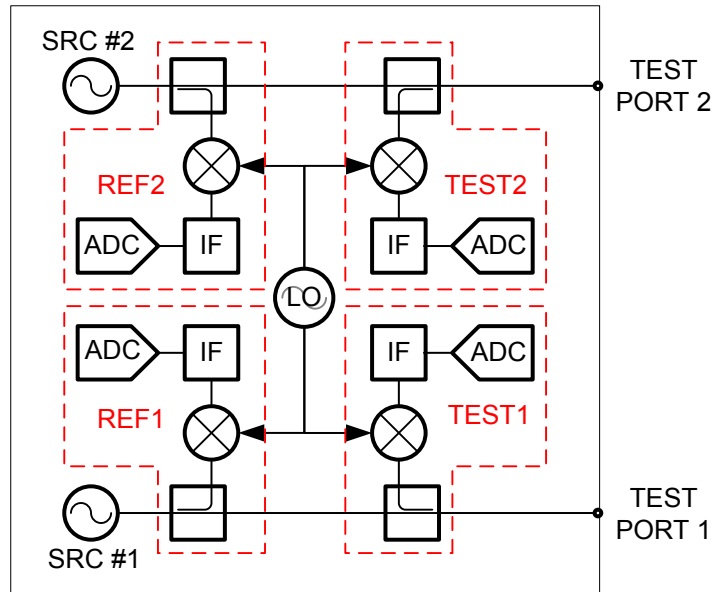


Figure 3.10. VNA Using Two Sources and Four Receivers

As previously mentioned, higher-performance VNA architectures often come with additional cost and complexity. Therefore, consider the application, performance, required accuracy, and cost among other factors, in order to select the proper VNA architecture.

4. Vector Network Analyzer Calibration

In order to understand VNA calibration, first consider the purpose of a vector network analyzer (VNA). The VNA is designed to *accurately* characterize the linear behavior of a device by evaluating the phase and magnitude of incident and reflected waves. By measuring the phase and magnitude of these waves, you can determine a wide range of device characteristics including impedance, return loss, insertion loss, and even group delay. Thus, the accuracy with which a VNA can determine the linear behavior of a device under test (DUT) is fundamentally determined by the accuracy with which the instrument can measure the phase and magnitude relationship of the incident or reflected waves.

Why Calibrate?

Although it is possible to calibrate certain elements of the VNA at manufacturing time, such as source power and receiver accuracy, much of the calibration information required to perform an accurate measurement depends on the measurement setup. For example, consider the example of measuring the impedance of an open circuit. In this example, illustrated in **Figure 4.1**, the open circuit is connected to the instrument through a cable.

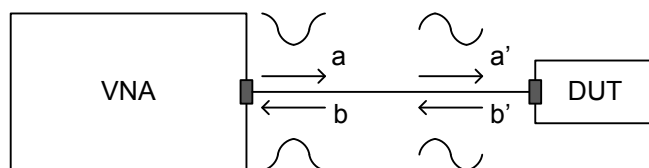


Figure 4.1. Effect of Test Port Cable on Impedance Measurement

It is well known that the open circuit has an infinite impedance and that the reflected wave will have the same magnitude as the incident wave and will be in-phase with the incident wave. However, as illustrated in **Figure 4-1**, the length of the cable between the VNA port and the open circuit will affect the phase as measured by the VNA. At certain frequencies, the open will even show up as a short when using only the factory calibration performed at the test ports.

In order to prove this phenomenon mathematically, assume the test port cable is lossless and introduces a delay (τ) that corresponds to the length of the cable. As a result, the incident wave (a'), at the end of the test port cable, will be a delayed version of the incident wave (a), at the test port, as shown in **Equation 4.1**.

$$\begin{cases} \mathbf{a}' = e^{-j2\pi f\tau} \cdot \mathbf{a} \\ \mathbf{b} = e^{-j2\pi f\tau} \cdot \mathbf{b}' \end{cases}$$

Equation 4.1. Impact of the Test Port Cable on the Incident and Reflected Waves

$$\Gamma_{test\ port} = \frac{\mathbf{b}}{\mathbf{a}} = e^{-j4\pi f\tau} \cdot \frac{\mathbf{b}'}{\mathbf{a}'} = e^{-j4\pi f\tau} \cdot \Gamma_{DUT}$$

Equation 4.2. Impact of the Test Port Cable on the Measured Mismatch

As observed in **Equation 4.2**, the open at the end of the test port cable will show up as a short at the test port for all frequencies, such that $4\pi f\tau = \pi + k \cdot 2\pi$ or $\tau = 1/(4f) + k/(2f)$ for $k \in \mathbb{Z}$. More specifically, an open will appear as a short when the cable corresponds to a quarter wavelength.

Although a factory calibration might be useful to verify the accuracy of certain elements of a VNA, such as the power of the stimulus and the frequency response of the receivers, other characteristics related to the measurement setup, such as cable length, are impossible to account for during a factory calibration. As a result, the user calibration is an important step that not only accounts for variables due to measurement configuration, but user calibration can also account for instrument variation over temperature and frequency.

Sources of Error in VNA Measurements

Although the location of the reference plane on VNA measurements is one factor that affects measurement results, VNAs are also subject to a range of inherent instrument impairments that can also be accounted for through calibration. In general, there are four main contributions to this imperfection, including:

- port match
- directivity
- frequency response
- isolation

The sources of error listed above are often referred to as *systematic* sources of error because they systematically affect the measurement at all times. The effects of systematic errors on a measurement result, shown in **Figure 4.2**, can largely be removed through calibration.

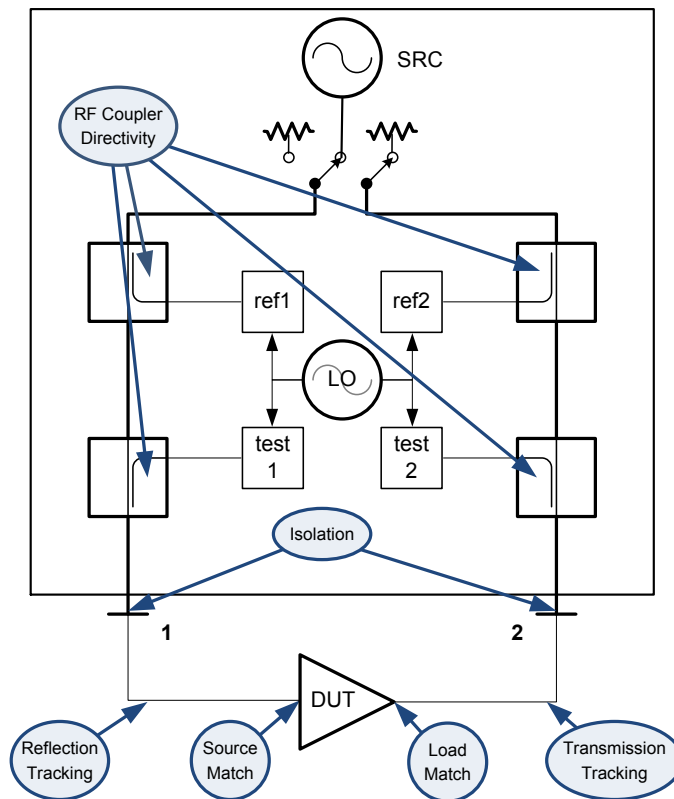


Figure 4.2. Major Contributions to Systematic Errors

Port Match

The port match of a VNA port has significant impact on the measurement accuracy of the instrument. Ideally, the test ports of a 50 ohm VNA would present itself as exactly 50 ohms. However, in practice, the precise impedance of the test port of a VNA can vary according to the quality of the hardware design. Moreover, even well designed hardware has a return loss typically around 20 dB. As a result, the effect of the impedance of a VNA test port cannot be neglected while performing accurate VNA measurements and is referred to as source match and load match.

Directivity

A VNA utilizes separation hardware such as directional couplers or bridges to independently measure the incident and reflected waves at both the input and output of the DUT. These measured quantities are referred to as the *raw* quantities. Although a perfect coupler would only sense the wave it is expected to sense, this is not absolutely true. For example, as observed in **Figure 4.3**, some portion of the reflected wave traveling in the reverse direction of the through path will still present itself in the coupled path.

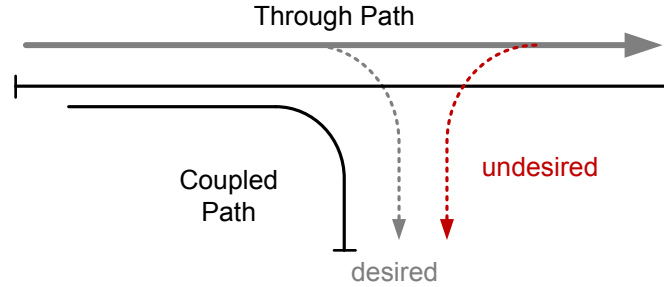


Figure 4.3. Imperfect Directional Coupler

Thus, the isolation of the coupled arm and the signals traveling in the reverse direction is finite. As a result, the coupled arm will also sense some of the wave traveling in the other direction. This phenomenon is referred to as *imperfect directivity*. Imperfect directivity is a systematic error and can be compensated for as part of the calibration process.

Frequency Response

The fundamental design of a VNA involves multiple receivers. Because the signal path corresponding to each receiver is not identical, each receiver can have a slightly different frequency response. The coupler can actually contribute to the frequency response variation between receivers, because each coupler has a slightly different coupling factor as a function of frequency. The phenomenon of each receiver having a slightly varying frequency response is often referred to as a *reflection and transmission tracking* within the context of a VNA calibration.

Isolation

Although a VNA would ideally have infinite isolation between ports, this is not always the case in practice. In fact, some of the signal measured at one port can leak into the receiver channels corresponding to another port. This phenomenon is sometimes referred to as *crosstalk* and can also be compensated for as part of the calibration process.

Types of VNA Calibration

Given the wide range of impairments that can prevent a VNA from performing an accurate network analysis measurement, calibration is required in order to measure each of these impairments and correct the measurement result. There are a wide range of methods to calibrate a VNA, depending on the number of ports, the frequency range, the available calibration standards, and even the port type of the DUT. The port type of the VNA can be coaxial, on wafer, in fixture, or wave guide. Another differentiation of VNA calibration types is the trade-off of accuracy versus speed. For a two-port, full S-parameter VNA, the following calibration types are most common:

- Frequency response calibration
- One-path, two-port calibration
- Full S-parameter calibration

The frequency response calibration (or transmission normalization) is the simplest of all calibration types and merely corrects for the frequency response of the instrument. In general, you should employ this calibration type only when a rough idea of the linear performance (S_{21} or S_{12}) of a DUT is sufficient. Transmission frequency response calibration does not compensate for directivity and mismatch errors. As a result, the quality of this basic calibration depends on the raw directivity and mismatch performance of the VNA.

A second calibration type, the one-path, two-port forward calibration, allows you to perform forward S-parameter measurements. This calibration type is often sufficient when measuring the input match (S_{11}) and the forward transmission (S_{21}) of the DUT. It is important to note that the forward transmission is only partly corrected using this calibration type, since port 2 of the VNA is not fully characterized as part of the calibration process. This calibration assumes that port 2 is perfectly matched and allows for reasonably accurate S_{11} and S_{21} measurements. S_{11} not only includes the input match of the DUT but also the (load) match at port 2 in combination with S_{12} of the DUT. In the case of a unilateral device, where the amplitude of $|S_{12}|$ is very small, the negative effect of a non-perfect match at port 2 will be minimal.

Finally, the full two-port calibration is the most common and complete calibration type of a two-port VNA. With this method, all four S-parameters are fully corrected using various calibration techniques. Note that correction for imperfect isolation is optional and can often be omitted for most commercially available VNAs. Generally, the isolation portion of the calibration process is only required when measuring high-isolation or high-dynamic range devices.

Modeling Two-Port Calibration

In order to understand exactly how the calibration process is used to correct for impairments in the instrument, it is first worthwhile to understand how to model a vector network analyzer and each source of error. For the purposes of this discussion, consider the full two-port calibration model, as it is the most common. Without diving into the mathematics, and not considering stochastic contributions such as measurement noise or connection repeatability, calibration models are based on linear relationships between the true wave quantities at the calibration planes and the raw measured wave quantities "inside" the VNA.

The classical representation of systematic errors uses an error adapter model, as shown in **Figure 4.4**, and is easily mapped onto the imperfections of the hardware as described above. For instance, $e_{10}e_{01}$ represents the frequency response of the reflected wave (b) path relative to the incident wave (a) path, as S-parameters are ratios.

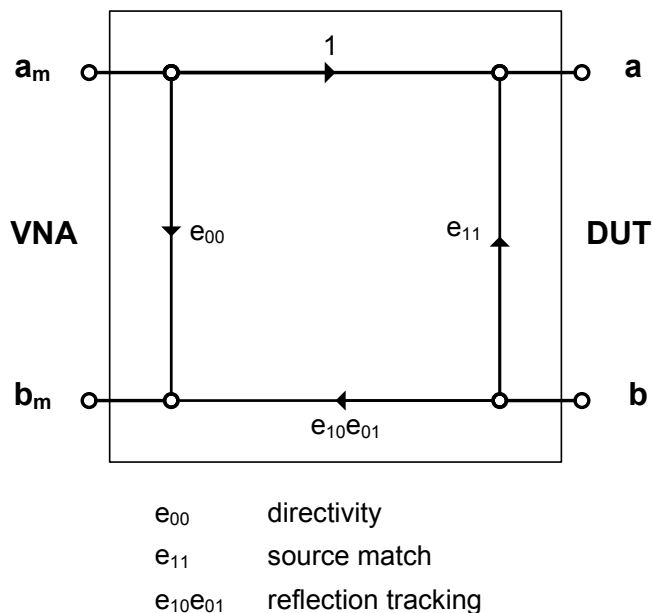


Figure 4.4. One-Port Error Adapter Model

At each (source) port, the calibration algorithms are designed to compensate for coupler directivity, reflection tracking, and source match. In addition to these errors, two-port calibration introduces three additional error contributions at the other (load) port, including load match, transmission tracking, and isolation. Applying both sets of errors to both the forward and reverse model results in 12 error contributions in total, hence the term *12-term error model*. In most scenarios, the contribution of the isolation can be neglected, which reduces the model to a 10-term error model. Finally, the VNA applies the error terms of the forward and reverse model onto the raw S-parameters to obtain the corrected S-parameters

In the case of a two-port VNA that has four receivers, you can mathematically describe the relationship between the raw measured waves and the corrected waves using matrix notation. The mapping of raw S-parameters onto calibrated S-parameters is referred to as a *wave formalism* approach. In order to understand this notation, it is important to review the block diagram of a two-port VNA, as illustrated in **Figure 4.5**.

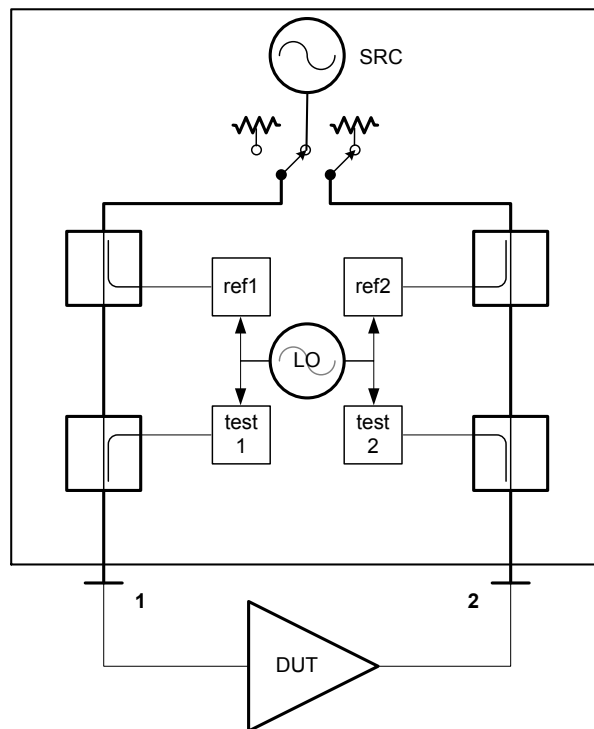


Figure 4.5. Typical VNA Block Diagram

As illustrated in **Figure 4.5**, the raw incident wave (a_{1m}) is measured at reference receiver 1 (ref 1) and the corresponding raw reflected wave (b_{1m}) at test receiver 1 (test 1). Meanwhile, the waves, a_1 and b_1 , are first measured at the reference plane at port 1 and then corrected using error coefficients. A similar scenario applies for port 2.

Assuming the crosstalk (due to imperfect isolation) between both ports can be neglected, the resulting matrix will contain eight non-zero elements within a 4x4 matrix. Because S-parameters only require ratios of wave quantities, one of the elements of the error correction matrix can be independently chosen and the value set to 1. The result is a *7-term error model*, as shown in **Equation 4.3**.

$$\begin{pmatrix} a_1 \\ b_1 \\ a_2 \\ b_2 \end{pmatrix} = K \begin{pmatrix} 1 & \beta_1 & 0 & 0 \\ \gamma_1 & \delta_1 & 0 & 0 \\ 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & \gamma_2 & \delta_2 \end{pmatrix} \begin{pmatrix} a_{1m} \\ b_{1m} \\ a_{2m} \\ b_{2m} \end{pmatrix}$$

Equation 4.3. Wave Formalism (7-Term Error Model)

Note that the ultimate error correction model is the *16-term error model* and is based on the determination of all 16 terms. Using a full (16-term) matrix, all possible systematic errors (resulting in contributions of any raw measured quantity to any corrected quantity) are taken into account.

Two-Port Calibration Methods

The general process of calibration involves connecting a set of known calibration standards to the appropriate reference plane. In general, think of a calibration standard as a device that has a precisely characterized and well-known impedance. The reference plane is the location at which you connect the DUT. Thus, with the calibration standard connected, the VNA measures the phase and magnitude relationship of incident and reflected waves with a well-known impedance attached to the reference plane. Using these results, the VNA calibration process compensates for internal impairments of the instrument and returns a more accurate measurement result at that reference plane.

In practice, there are several VNA calibration methods, each of which has unique benefits in various applications. The following two VNA calibration methods are discussed in the following sections:

- Short-Open-Load-Through (SOLT)
- Through-Reflect-Line (TRL)

Note that the number and type of standards, and the accuracy with which these standards need to be known, depend on the calibration methods. It is also important to recognize that because user calibration of a VNA takes characteristics from the measurement setup into account, you must take great care not to disturb the calibration plane or test setup after calibration. In addition, it is essential to use high-quality, phase-stable test port cables to minimize the impact of various bending of the cables during the measurement of the DUT compared to calibration.

SOLT Calibration

The most popular coaxial calibration method and probably the first manual calibration performed by a new VNA user is referred to as either SOLT (Short-Open-Load-Through) or TOSM (Through-Open-Short-Match). This calibration method involves the process of connecting well-known calibration standards to each port. Next, both ports are physically connected to each other. In a simple case, when preparing to measure an insertable coaxial device, which has mating connectors, both calibration ports can be directly connected to each other resulting in a zero-length through or flush through.

Dealing with non-insertable devices requires a slightly more complex variant of the SOLT calibration method. This situation is one in which the DUT has identical genders at both ports (for instance, two female SMA connectors) or different port types (for instance, 3.5 mm and 2.4 mm). The SOLR method assumes the use of a reciprocal through, hence the "R" instead of "T". In general, there are two primary techniques you can choose to deal with non-insertable devices: usage of phase-equal adapters or the application of the SOLR calibration technique. The SOLR technique is becoming more widely accepted, and hence much more common, as opposed to the traditional usage of phase equal adapters when dealing with non-insertable devices.

One of the simplest approaches to handling non-insertable devices is to use phase-equal precision adapters. Phase-equal adapters are precision matched adapters for various gender combinations, and each phase-equal adapter has identical dimensional and therefore identical electrical characteristics. Phase-equal adapters allow for a zero-length through during calibration, even when the test ports have identical genders, as shown in **Figure 4.6**.

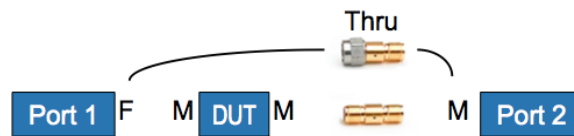


Figure 4.6. Usage of Phase-Equal Adapters to Support Non-Insertable Devices

As illustrated in **Figure 4.6**, a male-female adapter extends the test port cable at port 2. As a result, the reference plane at port 2 now corresponds to the male side of this adapter and thus the short, open, and load must be connected there as part of the calibration at port 2. Because the adapter has been included as part of the calibration process, a zero-length through connection is now possible between the original female connector of the test port cable at port 1 and the male connector of the test port cable at port 2.

After calibration, the male-female adapter is swapped by a phase-equal female-to-female adapter to appropriately connect to the male-to-male DUT. Given that both adapters have identical electrical characteristics, the calibration remains valid even though the genders of the adapters are different. Note that phase-equal adapters are typically provided as part of a calibration kit or as an option.

SOLR Adaptation to SOLT

A second and preferred solution to adaptor removal in SOLT calibration is the use of a slightly different calibration method, which is referred to as SOLR (Short-Open-Load-Reciprocal), also known as *Unknown Through*. This technique only requires a reciprocal ($S_{21} = S_{12}$) structure to be used during through calibration. In some situations, it is necessary to provide an estimate of the delay of the unknown through. When performing a SOLR calibration, the first step is to perform a one-port calibration at both ports using short, open, and load. Next, use a forward and reverse measurement to estimate both the S-parameters of the reciprocal structure (S_{11} , S_{22} and $S_{21}=S_{12}$). As a result of these measurements, you can determine the scaling factor which links the error coefficients at port 2 to those at port 1.

A final solution that allows calibrated S-parameter measurements of non-insertable devices is referred to as adaptor removal calibration. Adaptor removal requires two two-port calibrations. The first two-port calibration is performed using the adapter at port 1 and a second calibration is performed using the same adapter at port 2. Combining both calibrations yields a single full two-port calibration that allows you to accurately measure the non-insertable device.

Challenges with SOLT Calibration

The biggest drawback of SOL-based calibration techniques is the assumption that the load is ideal (generally 50 ohms). Fortunately enhancements do exist. One option is to measure the load after a TRL calibration (explained in the following section) and provide its S-parameters as part of the calibration kit. By providing the S-parameters for the load, it is no longer required to assume that the load is perfect. An alternative is to use an *ideal* load only at low frequencies (typically up to a couple of gigahertz) and to instead use a sliding load above that frequency. At higher frequencies, the combination of a precision air line and the *sliding load* allows you to create a *sliding mismatch* which is applied at different distances with respect to the reference plane. The corresponding mismatch results in a circle and the estimation of the center of that circle allows a better determination of the directivity.

A second challenge with the SOLT technique is the practical challenge of manufacturing the calibration standards. For example, in transmission mediums such as a wave guide, it is difficult or impossible to come up with an open. This difficulty can be somewhat circumvented by replacing the open with a second short, which has a different length and is referred to as offset short. Using a single additional offset short, you can accurately calibrate the VNA over a limited frequency range.

If you have a need for (significantly) more than two ports, it is worthwhile to remember that there is a QSOLT calibration technique, which only requires a short-open-load calibration at a single port and a known through connection between that port and every other port. The “Q” in QSOLT stands for “quick” and is a useful calibration tool given the complexity of the standard SOLT technique.

TRL Calibration

A second popular calibration technique is TRL (Through-Reflect-Line). TRL calibration is typically used for on-wafer or in-fixture measurements, where the use of precise calibration standards is impractical or impossible. In general, using the TRL calibration technique requires you to create your own in-circuit calibration standards that appropriately mate with the measurement fixture. Here, the *through* is essentially a zero-length microstrip that directly connects each end of the test fixture. The *reflect* standard is generally a crude (but repeatable) short or open circuit. Finally the *line* is similar to the through but has a non-zero length.

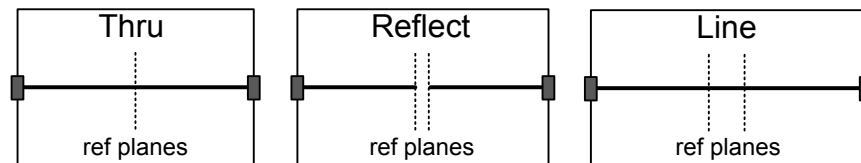


Figure 4.7. TRL Calibration Standards

The TRL calibration method first involves the connection of a reflect to each port. Note that the accuracy of this calibration technique requires that the reflect standard presented to each port be identical. In order to ensure that each port is presented with an identical reflect standard, one option is to use a single reflect standard and connect it to both ports. Next, measure a through and then a line. Here, the assumption is that the through and the line are perfectly matched and as a result, define the characteristic impedance of the system.

In practice, it is often easier to fabricate a well-matched line than it is to fabricate a well-matched broadband load. Note that the characteristic impedance of the reflect does not need to be known exactly. When performing a TRL calibration, you only have to indicate whether it is approximately a short or an open for the frequency range of interest.

One of the fundamental limitations of the TRL approach is that it can only be used over a relatively small frequency range. The usable frequency range is defined by the length of the line according to **Equation 4.4**.

$$l \neq n \cdot \frac{\lambda}{2} \text{ with } \lambda = \frac{v_p}{f_0} \text{ and } v_p = \frac{c}{\sqrt{\epsilon_r}}$$

Equation 4.4. Impact of the Line Length on the Frequency Range of the TLR Calibration

In fact, at specific frequencies, the line looks too much like a through, such that it is impossible to extract meaningful error coefficients at these frequencies. As a rule of thumb, the frequency range is determined by a 20° to 160° region as shown in **Equation 4.5**.

$$f_{min} = \frac{v_p}{18l} \text{ and } f_{max} = 8 f_{min}$$

Equation 4.5. Calculating Frequency Range of the TLR Calibration

One solution to the frequency range limitations of the TRL technique is to use two or more lines, each of which have an appropriate length to cover different overlapping frequency bands. The corresponding calibration technique is referred to as *multi-line TRL*. In the situation where a zero-length through is impractical, a short line can be used instead, resulting in an LRL (Line-Reflect-Line) calibration. Note that it is important to accurately specify the propagation constant (delay and loss) of the shortest line in order to properly define the reference planes.

It makes sense to briefly comment on the usage of an automatic calibration module. The automatic calibration module contains different impedance standards that are carefully calibrated by the manufacturer, typically using a TRL-grade calibration. So, on top of the ease of use, an automatic calibration module also provides a high-quality calibration.

Tools for VNA Calibration

When calibrating a network analyzer, engineers can utilize a wide range of calibration and verification tools. For example, the load, short, or open used to calibrate a network analyzer is often part of a manual calibration kit. Verification kits are typically used to verify the performance of the VNA after a calibration and can be used to identify problems with the calibration procedure or even the calibration standard.

A calibration kit is the primary tool used to calibrate a VNA. Today, VNA manufacturers provide a wide range of calibration kits for various connector types and for both manual and automatic calibration. Calibration kits contain calibration elements, or calibration standards that have been carefully characterized based on their mechanical dimensions or based on measurements using a golden network analyzer, that was calibrated using traceable calibration standards.

A calibration kit also contains precise information about each calibration element behavior, and the VNA uses this information as part of the calibration routine. Calibration element data can be based on coefficients corresponding to polynomial models or based on a sufficiently dense set of S-parameter values as a function of frequency. Although connector care is considered to be standard good practice at RF and microwave frequencies, calibration elements should be treated extra carefully.

Manual Calibration Kits

Calibration kits can be either manual (mechanical) or automatic. When using a mechanical calibration kit, connect the correct calibration element to the appropriate port when requested to do so by the VNA software. Typically, you first connect each of the required calibration elements to each port one by one and then connect the VNA cables to realize a through connection between both ports. You should generally connect calibration standards through a direct connection if the genders of the calibration planes allow it. In scenarios where the gender of the calibration standards does not allow a direct connection, you can use a reciprocal ($S_{21} = S_{12}$) structure or a different technique based on phase-equal adapters or adapter removal.

Electronic Calibration Kits

In contrast to a manual calibration, electronic (or automatic) calibration modules are generally much easier to use. Calibrating a VNA with an automated device simply requires you to connect the ports of the calibration module to the ports corresponding to the calibration planes before pushing the “start calibration” button. In most cases, the calibration module is advanced enough to discover which ports are connected to which calibration planes. When using an electronic calibration module, the through connection can either be realized internally, inside the calibration module, or externally. When the through connection is realized externally, the connection is made in a manner similar to the manner used for a mechanical calibration kit.



Figure 4.8. Automatic VNA Calibration Kit Provided by National Instruments

Calibration Verification

Verification kits can be used not only to verify the accuracy of the calibration process, but also to verify that the VNA is still operating within its specifications. A verification kit typically contains precision air lines, mismatched air lines, and precision fixed attenuators. These elements are measured after carefully performing a calibration and compared to their known characteristic to determine if the VNA is still within specification.

It is also important to consider the frequency of the calibration interval. Assuming the current VNA settings (start and stop frequency, number of frequency points, IF bandwidth, source power, and receiver step attenuator values) can be used, the need for calibration depends on the time and the variation of environmental parameters, such as temperature, since the last calibration. Again, if you require extremely accurate measurements, it is worthwhile to

take the time and effort to perform a new calibration to eliminate all doubts about the validity of the error coefficients that are currently being used.

Calibration Considerations

When calibrating a VNA, it is important to use identical hardware and software settings during both the calibration of the VNA and the measurement of the DUT. For example, typical measurement settings options that should remain constant include frequency settings, IF bandwidth and averages, source power, and receiver step attenuator settings.

Frequency Settings

Basic frequency settings simply require a start and stop frequency and number of frequency points, although more complex frequency settings are possible. In fact, more advanced configurations allow for multiple frequency segments, each allowing for varying frequency resolution for selected frequency bands. Using this feature, you can zoom in to certain frequency ranges for which the behavior of the DUT is expected to change more rapidly with frequency than for other frequency ranges.

IF Bandwidth and Averaging

IF bandwidth and averaging can significantly affect both the measurement speed and accuracy of a VNA measurement. Decreasing the IF bandwidth will decrease the uncertainty on the measurements at the cost of increased measurement time and is an alternative to averaging. In general, calibration is most accurate when performed using a large number of averages. In fact, it is possible to use a large number of averages for the calibration routine and then reduce the number of averages used when measuring the DUT.

One final note about IF bandwidth is that various commercial VNAs indicate that the calibration is no longer valid when selecting a different IF bandwidth during a DUT measurement. When changing the IF bandwidth, be aware that the IF bandwidth setting used for the measurement must be the same or wider than the setting used for calibration, because the uncertainty on the calibrated measurement is defined by both the uncertainty of the raw measurement and the uncertainty of the error coefficients. The latter depends on the IF bandwidth that was selected during calibration.

Source Power

In principle, the source power should be chosen to be as high as possible to achieve the smallest uncertainty for a given IF bandwidth. However, remember to make sure that both the DUT and the receivers of the VNA are used in their linear mode of operation.

Receiver Step Attenuator Settings

Certain VNAs include receiver step attenuators in order to guarantee linear operation of the VNA receivers. In general, never change these attenuator settings in between calibration and DUT measurement.

Advanced Calibration Techniques

For some measurements, advanced calibration techniques are required due to the complexity of the test setup or the implementation of an advanced measurement. The following section discusses two advanced calibration techniques: de-embedding and power calibration.

De-embedding

Unfortunately, it is not always possible to perform a calibration and thereby establish a reference plane at the exact desired location. While the ideal reference plane is typically at the input and output of the DUT, it is often impossible to connect the appropriate calibration standards at this location. In these cases, use de-embedding techniques to move the calibration plane from one location to another, ideally as close as possible to the DUT.

Generally, de-embedding requires the knowledge of the S-parameters of whatever structure is between the reference plane and the DUT. For example, suppose you are testing a DUT in a test fixture. In this case, one of the most difficult tasks is to create an accurate model of the test fixture. In many cases, the S-parameters for the test fixture are provided by the manufacturer. In other instances, it is important to first extract the S-parameters using simulation tools in order to be able to de-embed the fixture. Today, most commercially available VNAs support simple corrections of fixture effects by means of port extensions.

Power Calibration

Today, vector network analyzer technology is moving beyond standard S-parameters by adding nonlinear measurement capabilities. Although the behavior of linear devices is, by definition, independent of the level of the input signal, the same is not true for the nonlinear behavior of the device. As a result, the precise power level applied to the DUT is an important element of making basic nonlinear measurements.

As already mentioned, a VNA is typically calibrated at manufacturing time at its test ports, with respect to both the source and the receivers. However, it is generally impossible to directly connect the DUT to the test port. Thus, in order to properly compensate for losses introduced by test port cables or other means to connect the DUT to the VNA, you must perform a *source calibration*. By comparing the power level applied by the source to the power level measured using a power meter at the location where the DUT will be connected, the VNA is able to build an internal correction table such that the desired source power is applied to the DUT.

Finally, in order to measure the compression characteristic of a nonlinear device, a *receiver calibration* can be performed, typically in combination with a source calibration, in much the same way as the source calibration. Here, the receiver is used to measure the absolute power levels of incident and reflected waves at the excitation frequency and its harmonics, instead of only the S-parameters that correspond to ratios of wave quantities.

A Final Word of Advice

Performing a calibration is only worth the effort when applying good measurement practices. Good measurement practices include the use of high-quality components, such as phase-stable test port cables, proper connector care, and the correct use of a torque wrench.

5. Glossary

A_1	incident power wave
A_2	reverse incident wave
B_1	reflected power wave
B_2	reverse reflected wave
CRT	Cathode Ray Tube
CW	Continuous Wave
DUT	Device Under Test
IF	Intermediate Frequency
RF	Radio Frequency
S	S-parameter Matrix
SNA	Scalar Network Analyzer
SOLT	Short-Open-Load-Through
SOLR	Short-Open-Load-Reciprocal
SRD	Step Recovery Diode
TDR	Time Domain Reflectometry
TOSM	Through-Open-Short-Match
TRL	Through-Reflect-Line
VNA	Vector Network Analyzer
Z_0	Characteristic Impedance

6. References

- [1] Rohde & Schwartz, "A Test and Measurement Retrospective — 75 Years of Rohde & Schwarz," *Microwave Journal*, 10 July 2008. [Online]. Available: <http://www.microwavejournal.com/articles/6586-a-test-and-measurement-retrospective-75-years-of-rohde-schwarz>. [Accessed 6 September 2013].
- [2] G. Simpson, "Vector Network Analysis and ARFTG: A Historical Perspective," in *50th ARFTG Conference Digest*, 1997.
- [3] M. Mislange, "HP Memory," [Online]. Available: <http://www.hpmemory.org/>. [Accessed 6 September 2013].
- [4] J. Love, "Electronic Test 1956-1987, a look back (Part 2)," *EDN Network*, 20 May 2013. [Online]. Available: <http://www.edn.com/design/design-and-prototyping/4414692/1/Electronic-test-1956-1987--a-look-back--Part-2->. [Accessed 6 September 2013].
- [5] N. Friedrich, "Microwave Legends," *Microwaves & RF*, 12 August 2011. [Online]. Available: <http://mwrf.com/community/microwave-legends>. [Accessed 6 September 2013].
- [6] S. A. Maas, *Microwave Mixers*, Artech House, 1993.