Large-signal characterization and modeling of nonlinear devices using scattering parameters

John B. Call

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of
Master of Science
in
Electrical Engineering

William A. Davis, Chair
Warren L. Stutzman
Dennis G. Sweeney

September 13, 2002
Blacksburg, Virginia

Keywords: black box model, S-parameters, microwave measurement, load-pull

Large-signal characterization and modeling of nonlinear devices using scattering parameters

John B. Call

(ABSTRACT)

Characterization and modeling of devices at high drive levels often requires specialized equipment and measurement techniques. Many large-signal devices will never have traditional nonlinear models because model development is expensive and time-consuming. Due to the complexity of the device or the size of the application market, nonlinear modeling efforts may not be cost effective. Scattering parameters, widely used for small-signal passive and active device characterization, have received only cursory consideration for large-signal nonlinear device characterization due to technical and theoretical issues. We review the theory of S-parameters, active device characterization, and previous efforts to use S-parameters with large-signal nonlinear devices.

A robust, calibrated vector-measurement system is used to obtain device scattering parameters as a function of drive level. The unique measurement system architecture allows meaningful scattering parameter measurements of large-signal nonlinear devices, overcoming limitations reported by previous researchers.

A three-port S-parameter device model, with a nonlinear reflection coefficient terminating the third port, can be extracted from scattering parameters measured as a function of drive level. This three-port model provides excellent agreement with device measurements across a wide range of drive conditions. The model is used to simulate load-pull data for various drive levels which are compared to measured data.

To Eva, the light of my life

ACKNOWLEDGMENTS

I am indebted to my advisor, Dr. William A. Davis. His ability to challenge students in the classroom and laboratory has been an inspiration in pursuing this thesis. Dr. Davis has been patient in overseeing my work and instructing me in the details of S-parameter theory and practice. Thank you for the many hours of support and wise counsel you have contributed towards my education.

Thanks to Mike Bergin, Bob Keasler, and Walt Wright of Motorola's Semiconductor Product Sector Wireless Infrastructure Systems Division for providing funding, devices, and fixtures for measurement. Without their support this work would not have been possible.

A work of this scope would not be possible without the sacrifices of my family. There were many hours spent "scattering and plotting" and I never had a satisfactory answer to put to rest their questions on the absence of "gathering parameters." A heartfelt thanks is extended to my wonderful wife. Her dedication to a lifetime of learning has been proven time and again.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Nonlinear devices can be characterized as a function of input drive level using S-parameters. This characterization can be used for further analysis and design. A calibration and measuring procedure that makes these measurements practical was used to measure S-parameters for several commercial devices. A three-port model for nonlinear devices is extracted and used to simulate a typical nonlinear measurement. Simulation results using the three-port model are compared with measurement results.

1.1 Nonlinear Networks and Device Characterization

In the late 1960's and early 1970's there was a lot of activity related to describing the newly developed nonlinear devices with extensions of existing network theory. The idea that a nonlinear network could be linearized in special cases where the nonlinearity was small or the power range was restricted was convenient because it allowed engineers who had been trained to deal with linear networks using linear analytical tools (such as the Laplace transform) to make a first order approximation to the analysis of a nonlinear network. During the same time period there were some very enlightened approaches to applying network analysis to nonlinear networks. Attempts were made to measure and use S-parameters on large-signal nonlinear devices but researchers ran into measurement issues—working with high power levels and devices that were nonlinear (Leighton, 1973). One approach would restrict the measurement of the nonlinear device to a linear region of operation so that linear theory and analysis could be applied to the measurement results.

The efforts to describe nonlinear devices with network parameters were in large part set aside as the advent of relatively inexpensive computing power in the late 1970's facilitated use of computer-aided design (CAD) programs that could simulate nonlinear systems using nonlinear device models in the time or frequency domain. Designers use models to optimize the device source and load terminations for best performance. Using computer design tools a designer can try several ideas before implementing a circuit. A designer can use the computer to test and experiment with new ideas, gaining crucial insight into circuit operation that will be useful when debugging the prototype circuit. The

model itself may provide valuable insight into the device operation and may be useful for improving future device designs.

Device models can be technology specific or they can be generic black box models. Models for a specific device technology can be based on the actual physical properties of the device or they can use lumped circuit elements to describe device behavior. Physics-based models rely on knowledge of device structure and material composition along with sets of differential equations to describe device behavior. Lumped circuit models describe model behavior using capacitors, inductors, and resistors to capture physical behavior. Most large-signal RF device models developed for harmonic balance simulators use lumped-element device models with nonlinear capacitors and current sources. Physics-based and lumped-element circuit models can take years for researchers to develop and if the device structure or layout significantly changes, the model may need to be altered. Black box models treat the device as a function block where the internal details are not as important as the ability to predict input and output characteristics. The device characteristics using a black box model are modeled using the fewest number of parameters. Because of their flexibility, black box models can be applied to a wide variety of devices and are particularly attractive for modeling complex circuits and systems. In terms of computational effort, physics-based models require the most computational effort, black box models usually require the least computational time, with lumped-element models falling somewhere between the two.

Lumped-element models contain many circuit elements with values extracted from various device measurements. Large-signal lumped-element models and some black box models rely on pulsed measurement of dc I-V curves. Small-signal S-parameter measurements are often used in the extraction process. Most device models have a long history of use for small-signal devices. For high power RF devices, models are often scaled to achieve the correct output power based on model extraction methods from a smaller device. The limitations of the extraction method may also limit the correlation between the expected performance and the actual performance. For example, a model extracted using pulsed measurements may not predict device performance when operating temperatures are significantly above room temperature.

As a result of nonlinear models being complex and costly to generate, a device manufacturer may provide models for only a small handful of popular devices. The cost of developing and testing models can be prohibitive for commercial RF/microwave devices that are used in low volume applications. Unfortunately, there are many high-power active devices that will never have a large-signal model. At present, designers working on time-to-market critical applications have not been able to benefit from models because the model may be released up to a year after the first prototype devices are sampled and their designs are well underway.

1.2 Scattering Parameters for Nonlinear Devices, Measurement and Simulation

A three-port representation of a nonlinear device as a function of drive level can be obtained with a few relatively simple measurements. This characterization is used to analyze and simulate device performance for a range of large-signal drive levels. Using the three-port model, designers could "roll their own" nonlinear device model and make use of the results in their circuit simulations.

We lay the foundations of S-parameter theory for single and multiple port networks. Fundamental theory of network representation is presented to further understand the bounds and limitations on network parameters. S-parameters are not limited to small-signal linear measurement and analysis. The operating regions of a nonlinear device are explored and large-signal S-parameters that could be used to describe the device in a large-signal linearized mode of operation are presented. Large-signal S-parameters were first presented in the early 1970's (Leighton, 1973), but with new calibration and measurement techniques we are able to apply the theory to high power transistors without some of the constraints faced by earlier researchers.

Commercially available microwave measurement systems capable of making S-parameter measurements on nonlinear devices under high drive level conditions have not been available. Theory and implementation of a nonlinear measurement system are discussed. A system has been developed (Davis, 1993) that makes possible scattering parameter measurements on nonlinear devices under high drive level conditions. For this thesis, the system was used to obtain S-parameter measurements on several commercially available power transistors that previously had been considered unmeasurable because of their high gain and capability for large output levels. The system was also used to make high drive level scattering parameter measurements from which a three-port representation was extracted for two transistors.

The three-port representation of nonlinear devices was developed as a compact and convenient method for characterizing the behavior of an active device under high drive level conditions (Davis, 1999). Using the three-port model, a device can be characterized using S-parameter measurements as a function of drive level. Advantages of representing a device as a three-port model are 1) any two-port nonlinear device can be represented, the model is not technology specific; 2) the model coefficients are themselves S-parameters which simplifies the simulation overhead and suggests that the three-port would be ideal for characterizing devices using new physical structures or exotic materials and RFIC or MMIC devices which are not adequately handled by existing models; 3) most functional RF design labs have the measurement equipment required to extract a three-port model. For this thesis, new methods of extracting three-port model parameters were used to obtain an excellent fit of the model to measured power sweep data.

The three-port model is well suited for characterizing devices from the circuit designer's standpoint, i.e. what type of circuit needs to be wrapped around a device in order to realize the desired performance? The three-port model provides a convenient

method that could be used by individual designers to obtain their own models for any nonlinear device they may be using, especially for high-power RF/microwave devices. Though the effort required to obtain a three-port model is minimal, the information and insight available from the model make it possible for high-power amplifier designers to adopt tools and techniques that have previously been available only to small-signal circuit designers.

The S-parameter nature of the three-port model lends itself well to simplifying simulation efforts and provides the capability of obtaining analytical expressions for important device performance parameters. Three-port model simulations are compared with measurements for devices running in two different modes of operation. Results indicate that there are still some technical measurement challenges that need to be addressed, but the three-port model does replicate power sweep measurements and promises to be a useful tool for circuit designers using nonlinear devices.

CHAPTER 2

SCATTERING PARAMETERS

This chapter provides a review of S-parameter definitions and a behavioral description of networks using S-parameters. Beginning with fundamental definitions of S-parameters, the discussion touches on n-port networks and presents the theory of nonlinear device representation. In chapter 4 a three-port network will be used to model nonlinear devices. Existing techniques for applying S-parameter concepts to active devices are summarized to lay the foundation for later chapters.

2.1 Representation of Networks with Scattering Parameters

A nonlinear model may describe the behavior of a complex system in terms of voltage and current as seen from outside the system. Most models restrict the system to a narrow region of operation. Network parameters such as S-, Y- or Z-parameters are often useful for describing system performance. When S-parameter theory is discussed it is usually narrowly restricted to linear, small-signal systems. Limiting the application of S-parameters exclusively to small-signal linear networks may be necessary for practitioners lacking a solid fundamental understanding of basic S-parameter theory. Introductory engineering texts motivate the derivation of S-parameters using linear networks. Most measurement systems are not designed to measure nonlinear networks. When S-parameters are measured, they are often measured in a small-signal regime of device operation—for measurement convenience. Alas, the shortcomings in theoretical instruction and measurement technology has fostered a climate in which it is believed nonlinear behavior cannot be described using S-parameters. Engineers building nonlinear models will routinely measure S-parameters of a device (as a function of voltage), convert the S-parameters to Y-parameters and extract an expression for a nonlinear capacitor as a function of voltage, yet these same engineers will adamantly claim that it is heresy for S-parameters to be used directly to describe nonlinear behavior. The theoretical underpinnings for a true nonlinear network parameter are in their infancy and the possibilities of extending S-parameters to directly describe nonlinear behavior presents

some formidable theoretical challenges. Essential in understanding the limitations and possibilities of describing network behavior (linear and nonlinear) using S-parameters is a theoretical and intuitive feel for the origins of scattering parameters.

In presenting the theoretical foundation for S-parameters it is useful to lay a sufficiently broad foundation such that the usual S-parameter theory can be enlarged to encompass the description of nonlinear devices (and not just the description of nonlinear devices restricted to a linear region of operation). The intent of the approach is to prepare the reader to accept the possibility of using S-parameters to describe nonlinear systems. Without rigorous derivation, the basis of S-parameters will be presented to give the reader a fundamental understanding of the depth of device behavioral information encapsulated by S-parameters.

Consider the equivalent circuit of a linear generator as illustrated in Fig. 2.1. The impedance Z_G is the Thevenin equivalent generator impedance and may be complex. The open circuit voltage is given by E_O. The magnitude of current I_i into load Z_L

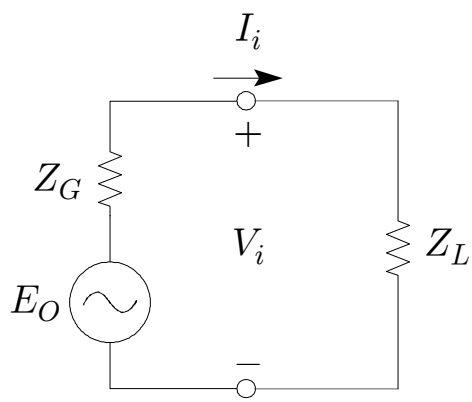


Figure 2.1. Equivalent circuit of a linear generator. The Thevenin equivalent generator is given by E_O and Z_G, where Z_G may be complex. For this discussion the load Z_L is the device being characterized.

is given by

$$|I_i| = \left| \frac{E_O}{Z_L + Z_G} \right|. \tag{2.1}$$

The power delivered to the load is

$$\begin{aligned} P_L &= \Re\left(Z_L |I_i|^2\right) \\ &= \Re\left(Z_L \left| \frac{E_O}{Z_L + Z_G} \right|^2 \right) = \frac{R_L |E_O|^2}{(R_L + R_G)^2 + (X_L + X_G)^2} \end{aligned} \tag{2.2a}$$

where R and X are respectively the resistive and reactive elements of Z. It is instructive to express the denominator of (2.2a) in an alternate format:

$$P_L = \frac{|E_O|^2}{4R_G + (R_L - R_G)^2/R_L + (X_L + X_G)^2/R_L}. \tag{2.2b}$$

When $R_G > 0$, maximum power is delivered to a linear load Z_L when

$$R_L = R_G,\ X_L = -X_G$$

with the maximum power, labeled P_a, given as

$$P_a = \frac{|E_O|^2}{4R_G}, \tag{2.3}$$

which is power delivered by the source to a conjugately matched load, sometimes referred to as available power.

The power delivered by the generator and dissipated by the load can be conveniently described in terms of incident and reflected power waves. The incident and reflected power waves a_i and b_i are defined (Kurokawa, 1965) as

$$a_i = \frac{V_i + Z_{oi} I_i}{2\sqrt{\Re(Z_{oi})}}, \quad b_i = \frac{V_i - Z_{oi}^* I_i}{2\sqrt{\Re(Z_{oi})}}, \tag{2.4}$$

where Z_{oi} is the port reference impedance and may not have any correlation to the generator impedance. The real part of the reference impedance is the normalization impedance. Typically the port reference impedance is chosen to be real and the terms reference impedance and normalization impedance are used interchangeably. A commonly used value of $Z_{oi} = Z_0 = 50\,\Omega$ is used in most commercial S-parameter measurement systems with both ports being normalized to the same impedance. It is common for sources to have $Z_G = 50\,\Omega$ in which case the normalization impedance coincidentally equals the port terminating impedance, but it is not required that the normalization impedance be the same as the port terminating impedance[1]. For the purpose at hand the port reference impedance will be positive real and denoted Z_{oi}. The definition of a_i and b_i is chosen such that the magnitude squared, i.e. $|a_i|^2$, has units of power ($\mathrm{V}^2/\Omega = \mathrm{Watts} = \mathrm{J/s}$). It is instructive to expand (2.4) into an explicit function of current I_i

$$a_i = \frac{Z_L I_i + Z_{oi} I_i}{2\sqrt{Z_{oi}}}, \quad b_i = \frac{Z_L I_i - Z_{oi} I_i}{2\sqrt{Z_{oi}}}. \tag{2.5}$$

From (2.4) and (2.5) it is plain to see why the real part of Z_{oi} is termed the normalization impedance, as a_i and b_i consist of a voltage vector normalized by the impedance $\Re(Z_{oi})$.

[1]When measuring active devices there are many cases where it is desirable to have a reference impedance $Z_{oi} \ll 50\,\Omega$, for example $Z_{oi} = 10\,\Omega$. In such cases the physical constraints of the measurement setup may mean that $Z_{oi} \neq Z_G(\omega)$ where the actual port termination impedance is a complex function of frequency.

The vectors that constitute the power wave expression of (2.4) are shown graphically in Fig. 2.2. The expression for a_i consists of the unnormalized vector $E_O = V_i + Z_{oi}I_i$ in Fig. 2.2 which is the sum of voltages across the load Z_L and normalizing impedance (in the case of Fig. 2.1 the generator impedance $Z_G = Z_{oi}$) and is proportional to the square root of the power out of the source (since all power delivered by the source will be absorbed by the combination of Z_{oi} and Z_L). The a_i term is the incident power wave on the load Z_L, the magnitude squared being the incident power. The expression for b_i consists of the $V_i - Z_{oi}^* I_i$ vector in Fig. 2.2 which is the voltage delivered to the load minus what *could* have been delivered *if* the source were optimally matched ($Z_{oi} = Z_L^*$). It may also be thought of as an equivalent source voltage, where the load is modeled by

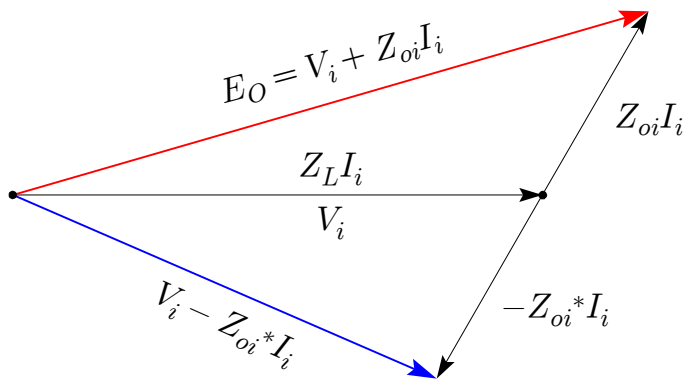


Figure 2.2. Graphical representation of power wave transformation. This representation is approximate for real normalizing impedances. A complex normalizing impedance would further stretch and rotate the transformation component vectors.

Z_{oi}^* and a dependent source. The vector $V_i - Z_{oi}^* I_i$ is a figure of merit that when squared is proportional to the power that could not be delivered to the load. If $Z_L = Z_{oi}^*$ then $V_i - Z_{oi}^* I_i$ is zero, corresponding to the conjugate match case of maximum power transfer and the figure of merit indicates that all power was transferred to the load. When $Z_L \neq Z_{oi}^*$, $V_i - Z_{oi}^* I_i$ would indicate a residual portion of power was *reflected* by the load. Fig. 2.2 suggests that a_i and b_i contain sufficient information about the port voltage and current to characterize V_i, I_i, and Z_L.

The port normalization impedance determines the properties of the transform. For a linear n-port, changing the reference terminations will not alter the linear relationship between v and i. For a nonlinear network, device characteristics that define the relationship between v and i are often a function of terminating impedance and a change in port reference impedance may not properly account for changes in device behavior.

The transformation of V and I into power waves as given by (2.4) has some advantages in describing the behavior of a network. Using power waves there is a simple relation between a_i and b_i, whereas when using V and I directly to compute power transferred and reflected the math can become quite cumbersome. Measurements with microwave test equipment yield quantities that lend themselves to a power wave description of the device under test. Consider the expression for power delivered to the

load Z_L. From circuit theory, power delivered to the load Z_L is given by

$$P_L = \Re(V_i I_i^*)$$

which in terms of power waves can be equated, through substitution of (2.4), to

$$\begin{aligned} |a_i|^2 - |b_i|^2 &= \frac{(V_i + Z_{oi} I_i)(V_i^* + Z_{oi}^* I_i^*) - (V_i - Z_{oi}^* I_i)(V_i^* - Z_{oi} I_i^*)}{4\Re(Z_{oi})} \\ &= \frac{(Z_{oi} + Z_{oi}^*)(V_i^* I_i + V_i I_i^*)}{4\Re(Z_{oi})} \\ &= \Re(V_i I_i^*) \ . \end{aligned} \tag{2.6a}$$

Using the above, the power delivered to the load Z_L can be expressed using power waves as

$$P_L = \Re(V_i I_i^*) = \left(|a_i|^2 - |b_i|^2\right). \tag{2.6b}$$

Note that for $\Re(Z_{oi}) > 0$ this expression is power delivered to the load. The power wave represented by $|b_i|^2$ is the backward-traveling power emanating from the load. If the load does not contain a power source this would be the power that was not absorbed by the load, i.e. the reflected power. The term $-|b_i|^2$ is always negative (even if the load contained a power source) so that the quantity $|a_i|^2$ can be considered the power incident on the load. When $\Re(Z_{oi}) > 0$ the quantity $|a_i|^2$ is the maximum power that can be delivered by the generator.

This section has presented the fundamental theory of scattering parameters. The presentation has followed that of most texts where the device being characterized, Z_L in the case of Fig. 2.1, is a linear passive network. In the next section the theory will be extended to n-port networks.

2.2 Representation of n-port Networks

A majority of active networks are fundamentally two-port networks and it would be useful to extend the theory of section 2.1 to two-port devices, as well as n-port systems. Power waves can be applied to n-port networks (Ha, 1981). Consider the n-port network, N, represented in Fig. 2.3. At each port is a Thevenin's equivalent voltage generator E_i and impedance Z_{oi}. In general the generator impedance may be a function of frequency. As in the previous section Z_{oi} is not constrained to be the generator impedance, it could be set arbitrarily to any reference impedance value, though for the development of this section it may be useful to consider the reference impedance and generator impedance to be the same, i.e. $Z_{oi} = Z_{Gi} = 50\,\Omega$. At each port the incident power wave a_i, reflected power wave b_i, port voltage V_i, and port current I_i are represented by the $n \times 1$ vectors $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{V}$, and $\boldsymbol{I}$. The incident and reflected power wave vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are defined as

$$\boldsymbol{a} = \boldsymbol{R}(\boldsymbol{V} + \boldsymbol{Z}\boldsymbol{I}) \tag{2.7a}$$

$$\boldsymbol{b} = \boldsymbol{R}(\boldsymbol{V} - \boldsymbol{Z}^*\boldsymbol{I}) \tag{2.7b}$$

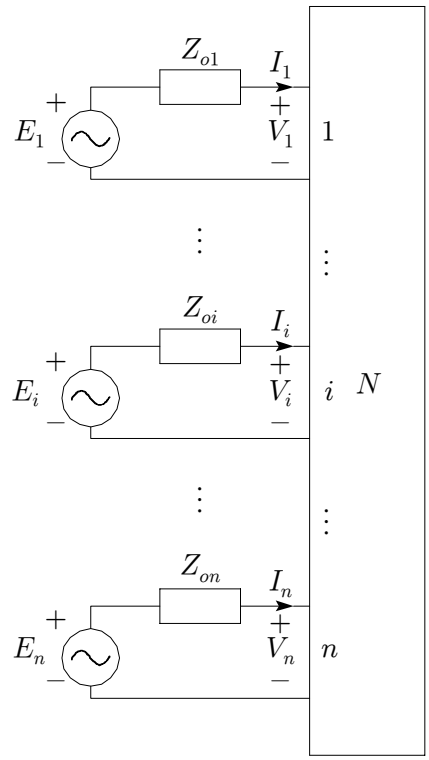


Figure 2.3. The n-port network N for extending power waves to multi-port networks.

where $\boldsymbol{R}$ is the diagonal matrix consisting of the square root of the normalization impedance

$$\boldsymbol{R} = \begin{bmatrix} \frac{1}{2\sqrt{\Re(Z_{o1})}} & & & & 0 \\ & \ddots & & & \\ & & \frac{1}{2\sqrt{\Re(Z_{oi})}} & & \\ & & & \ddots & \\ 0 & & & & \frac{1}{2\sqrt{\Re(Z_{on})}} \end{bmatrix} \tag{2.8}$$

and $\boldsymbol{Z}$ is the diagonal matrix of the reference impedance Z_{oi}

$$\boldsymbol{Z} = \begin{bmatrix} Z_{o1} & & & & 0 \\ & \ddots & & & \\ & & Z_{oi} & & \\ & & & \ddots & \\ 0 & & & & Z_{on} \end{bmatrix}. \tag{2.9}$$

The network N is characterized by the $n \times n$ scattering matrix $\boldsymbol{S}$

$$\boldsymbol{b} = \boldsymbol{S}\boldsymbol{a} \tag{2.10}$$

with the S-parameters normalized by the reference $Z_1, Z_2, \ldots, Z_n$. The incident and reflected power waves given by (2.7) defines a mapping function from the $\boldsymbol{V}$, $\boldsymbol{I}$ plane into $\boldsymbol{a}$ and $\boldsymbol{b}$ device behavioral space. The scattering matrix (2.10) is a transformation or mapping that relates the incident and reflected power waves at the port terminals and could be alternatively expressed in terms of $\boldsymbol{V}$ and $\boldsymbol{I}$ directly. To derive definitions for

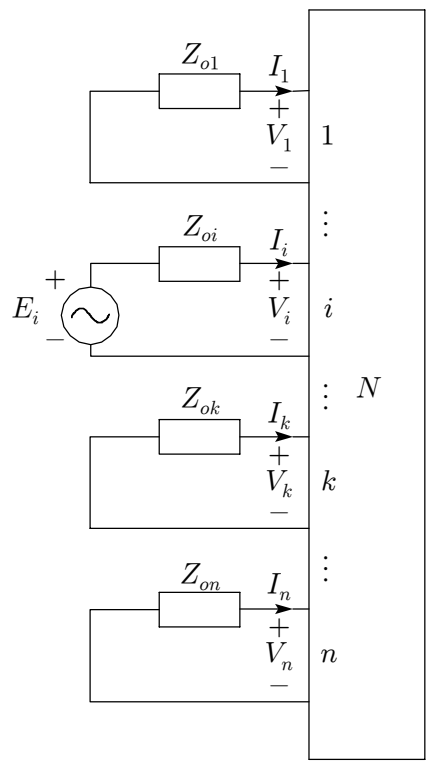


Figure 2.4. Terminating network N to derive multi-port S-parameters.

the scattering parameters, we may terminate all ports in the corresponding reference impedance Z_{ok} and allow only port i to be driven by generator voltage E_i as shown in Fig. 2.4. The voltage at port k is given by

$$V_k = -Z_{ok} I_k, \quad k \neq i. \tag{2.11}$$

Substituting port voltage V_k into (2.7a) results in $a_k = 0$ (for $k \neq i$). As in the one-port

case, the generator voltage at port i can be expressed as $E_i = V_i + Z_{oi}I_i$. Substituting the generator voltage E_i into (2.10) results in the relationship for the reverse power wave at port k due to an input signal on port i

$$b_k = s_{ki}a_i = s_{ki}\frac{E_i}{2\sqrt{\Re(Z_{oi})}}, \quad i = 1, 2, \ldots n. \tag{2.12}$$

The power wave b_k can also be expressed using (2.11) and the kth equation of (2.7b) as

$$b_k = -I_k\sqrt{\Re(Z_{ok})}, \quad k \neq i. \tag{2.13}$$

The scattering parameter s_{ki} is obtained from (2.11) and (2.13)

$$s_{ki} = \frac{b_k}{a_i} = \frac{-2I_k\sqrt{\Re(Z_{oi})\,\Re(Z_{ok})}}{E_i}\ . \tag{2.14}$$

Multiplying s_{ki} by its conjugate results in

$$|s_{ki}|^2 = \frac{|I_k|^2\Re(Z_{ok})}{|E_i|^2/4\Re(Z_{oi})} = \frac{|b_k|^2}{|a_i|^2}\ . \tag{2.15}$$

As in the one-port case (2.3), the maximum power available at port i is $P_a|_i = |E_i|^2/4\Re(Z_{oi})$. The quantity $|I_k|^2\Re(Z_{ok})$ is the power dissipated in the termination Z_{ok}. The ratio $|s_{ki}|^2$ is the transducer power gain from port i to port k when the ports are terminated with the reference impedance[2]. Expressing the current I_k in (2.14) in terms of (2.11) yields an expression for s_{ki} in terms of the generator and port voltages

$$s_{ki} = \frac{b_k}{a_i} = \frac{V_k}{E_i}\frac{2\Re(Z_{ok})}{Z_{ok}}\sqrt{\frac{\Re(Z_{oi})}{\Re(Z_{ok})}}\ . \tag{2.16}$$

To understand the physical meaning of (2.16) consider the case where the reference impedances are purely real and $Z_{oi} = Z_{ok} > 0$. Under such ideal conditions

$$s_{ki} = \frac{2V_k}{E_i} \tag{2.17}$$

and the scattering parameter s_{ki} directly relates the voltage at port k with the generator voltage at port i in the sense of a voltage gain. To determine s_{ii} of the multi-port network N in Fig. 2.4, substitute the ith equations in (2.7a) and (2.7b) into (2.11) to give

[2]The effects of the termination impedance (as well as reference impedance) are included in the expression for transducer gain. For $|s_{ki}|^2$ to be consistent with transducer gain as developed in fundamental circuit analysis, the reference impedance and the termination impedance should be the same.

$$s_{ii} = \frac{b_i}{a_i} = \frac{V_i - Z_{oi}^* I_i}{V_i + Z_{oi} I_i} \tag{2.18}$$

Let $Z_i = V_i / I_i$ be the impedance seen looking into port i when all other ports are terminated. Substituting Z_i into the expression for s_{ii} gives

$$s_{ii} = \frac{Z_i - Z_{oi}^*}{Z_i + Z_{oi}} \tag{2.19}$$

The ratio s_{ii} is the power wave reflection coefficient for the condition that all other ports are terminated with their reference impedances. When Z_{oi} is positive real, s_{ii} is also the voltage reflection coefficient. The square of the magnitude of (2.18) $|s_{ii}|^2$, is the power reflection coefficient.

The elements of S in (2.10) relate the incident and reflected power waves of the multi-port network N. Specifically the elements of the scattering matrix are

$$s_{ii} = \left. \frac{b_i}{a_i} \right|_{a_k=0,\, k \neq i} \tag{2.20a}$$

$$s_{ki} = \left. \frac{b_k}{a_i} \right|_{a_k=0,\, k \neq i} \tag{2.20b}$$

The scattering parameters are defined for the condition $a_k = 0$, $k \neq i$, such that there is no contribution to the parameter from power waves incident on other ports. This condition is typically met by disabling sources on all other ports and terminating the ports in their respective reference impedances, forcing $a_k = 0$. From a measurement perspective it is relatively straightforward to obtain the S-parameters using a technique similar to that shown in Fig. 2.4, though this requires that the multi-port network N be linear since the technique requires superposition to hold true. It would be equally valid to derive the S-parameters with signal sources simultaneously at all ports as shown in Fig. 2.3, provided that the effects of all generators $k = 1, 2, \ldots N$, $k \neq i$ be removed from the evaluation of the S-parameters related to port i as given by (2.20). It should be noted that the derivation of power waves and network descriptions using S-parameters does not require that the magnitudes of V or I be small. The elements of the S matrix are functions of the incident and reflected power waves at the port terminals. For a nonlinear network to be described by scattering parameters the elements of the S matrix would become functions of power waves.

2.3 Generalization of Network Representations

The previous sections have shown how S-parameters can be used to describe the behavior of a network by transforming the terminal v_k and i_k relationships to an alternate device

behavioral space, in the case of power waves the space is spanned by $\boldsymbol{a}$ and $\boldsymbol{b}$. A set of network parameters such as S-, Y-, or Z- parameters represents a transform from one device behavioral space to another behavioral space representation. Before employing S-parameters for the task of describing active devices, we first consider in more fundamental detail the basis of network representations and the conditions under which they can be used. How many ways can a network be characterized? Do the transformations that make network parameters possible have limitations and bounds? And most importantly for the purposes at hand, how does this theory apply to nonlinear networks? Most texts present network parameters in the context of linear networks with no mention of nonlinear networks. First-time students may conclude that the theory of network parameters (specifically S-parameters) cannot be applied outside the realm of linear networks. This section presents some fundamentals of network representations. The underpinnings of network theory are broad and solid enough to support both linear and nonlinear networks. Indeed, linear network theory is a well explored subset of circuit theory. The insight gained by reviewing the fundamentals of network representation will motivate an understanding of previous attempts to use network parameters for nonlinear devices and prepare the reader for the possibilities of using nonlinear modeling techniques to expand the useful application of S-parameter design techniques. For reasons of clarity and conciseness, many aspects of nonlinear circuit theory that are interesting and useful in a broader context will not be covered.

The terminal characteristics of an n-port device can be specified in terms of network variables voltage $v_k(t)$, current $i_k(t)$, flux $\phi_k(t)$[3], and charge $q_k(t)$ for each port k. The relationship between flux-voltage and charge-current is given by

$$\phi_k(t) \equiv \phi_k(t_0) + \int_{t_0}^{t} v_k(\tau) d\tau, \quad k = 1, 2, \ldots, n \tag{2.21}$$

$$q_k(t) \equiv q_k(t_0) + \int_{t_0}^{t} i_k(\tau) d\tau, \quad k = 1, 2, \ldots, n \tag{2.22}$$

where network variables ϕ_k and q_k are related by a time dependence to v_k and i_k. The n-port is uniquely described by any one of four $n \times 1$ pairs of independent port vectors of network variables consisting of one of the following possible combinations

$$\{(\boldsymbol{v}, \boldsymbol{i}), (\boldsymbol{\phi}, \boldsymbol{i}), (\boldsymbol{v}, \boldsymbol{q}), (\boldsymbol{\phi}, \boldsymbol{q})\}. \tag{2.23}$$

The waveform vectors are paired such that they span a vector space $\mathbb{R}^n$ defining device behavior. It may not always be desirable, or practical, to describe device performance using one of the waveform vector pairs in (2.23). For convenience, we follow the development of Chua (1978) and define an independent port vector pair $(\underset{\sim}{\varepsilon}, \underset{\sim}{\eta})$ which is a mapping of the n-port network variables to a space with desirable properties. Let the transformation T represent a mapping of a pair of independent port waveform vectors to

[3]The flux $\phi_k(t)$ can be considered to be a magnetic charge.

the vector pair $(\underset{\sim}{\varepsilon}, \underset{\sim}{\eta})$, expressed as

$$T : \{(\boldsymbol{v}, \boldsymbol{i}), (\boldsymbol{\phi}, \boldsymbol{i}), (\boldsymbol{v}, \boldsymbol{q}), (\boldsymbol{\phi}, \boldsymbol{q})\} \mapsto (\underset{\sim}{\varepsilon}, \underset{\sim}{\eta}). \tag{2.24}$$

Transformation T is chosen such that the device behavioral space spanned by $(\underset{\sim}{\varepsilon}, \underset{\sim}{\eta})$ has desirable properties, typically simplifying network analysis.

Measurements of an n-port over the time interval $[t_0, \infty)$ results in a pair of dynamically independent port vector waveforms $(\underset{\sim}{\varepsilon}(t), \underset{\sim}{\eta}(t))$ which are defined as an admissible signal pair. By making multiple measurements with respect to t_0, admissible signal pairs $(\underset{\sim}{\varepsilon}(\,\cdot\,), \underset{\sim}{\eta}(\,\cdot\,))$ can be collected to characterize the n-port (Chua, 1978).

If an n-port can be characterized by an algebraic[4] relation between v_k and i_k it is said to be an n-port resistor. Examples of memoryless nonlinear resistors are

one-port: pn junction diodes, zener diodes
two-port: transistors
three-port: op-amp, analog multiplier

as shown in Fig. 2.5.

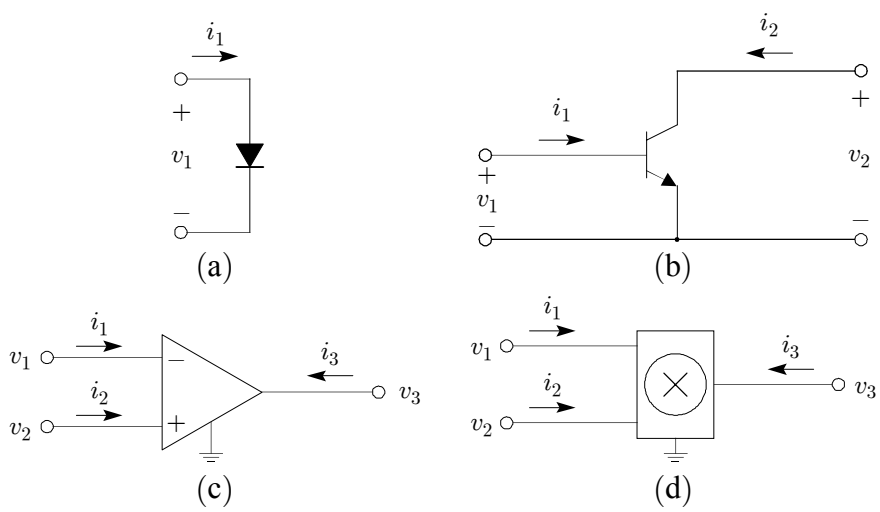


Figure 2.5. Examples of nonlinear resistors where an algebraic relation characterizes the relationship between v and i at each port: a) diode, b) npn transistor, c) operational amplifier, and d) analog multiplier.

We seek a general way for representing an n-port in terms of a relationship between $\underset{\sim}{\varepsilon}$ and $\underset{\sim}{\eta}$. Consider an n-port resistor N as shown in Fig. 2.5 with port voltage vector $\boldsymbol{v}$ and port

[4]The term algebraic is used loosely and is intended to imply a functional relationship between v and i which could include transcendental functions, but does not involve time derivatives or integrals which would imply the network has memory.

current vector $\boldsymbol{i}$.[5] The $n \times 1$ vectors $(\underset{\sim}{\varepsilon}, \underset{\sim}{\eta})$ are related to $\boldsymbol{v}$ and $\boldsymbol{i}$ by a coordinate transformation matrix $\underset{\sim}{\Omega}$ as

$$\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{i} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{a} & \underset{\sim}{b} \\ \underset{\sim}{c} & \underset{\sim}{d} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\eta} \end{bmatrix} = \underset{\sim}{\Omega} \begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\eta} \end{bmatrix} \tag{2.25}$$

$$\begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\eta} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{\alpha} & \underset{\sim}{\beta} \\ \underset{\sim}{\gamma} & \underset{\sim}{\delta} \end{bmatrix} \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{i} \end{bmatrix} = \underset{\sim}{\Omega}^{-1} \begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{i} \end{bmatrix} \tag{2.26}$$

The coordinate transformation matrix $\underset{\sim}{\Omega}$ is a $2n \times 2n$ nonsingular constant matrix. The condition that $\underset{\sim}{\Omega}$ be nonsingular means that information about the n-port is not lost in the transformation and there is a clearly defined inverse transformation. Vectors $\underset{\sim}{\varepsilon}$ and $\underset{\sim}{\eta}$ are defined as the generalized port coordinates of the network. As an example, a coordinate transform matrix $\underset{\sim}{\Omega}$ exists for the definitions of a_i and b_i as in (2.4) for a one-port network. The definition of incident and reflected power waves given by (2.4) can be expressed in the form of (2.26) where the inverse of the coordinate transformation matrix for a one-port with real reference impedance Z_0 is given by

$$\underset{\sim}{\Omega}^{-1} = \begin{bmatrix} \frac{1}{2\sqrt{Z_0}} & \frac{\sqrt{Z_0}}{2} \\ \frac{1}{2\sqrt{Z_0}} & \frac{-\sqrt{Z_0}}{2} \end{bmatrix}. \tag{2.27}$$

The port vector waveforms become the incident and reflected power waves

$$\begin{bmatrix} \underset{\sim}{\varepsilon} \\ \underset{\sim}{\eta} \end{bmatrix} = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}. \tag{2.28}$$

If $\underset{\sim}{\Omega}^{-1}$ is nonsingular (the determinant of $\underset{\sim}{\Omega}^{-1}$ is non-zero) then a valid coordinate transform matrix $\underset{\sim}{\Omega}$ exists. The determinant of $\underset{\sim}{\Omega}^{-1}$ in (2.27) is computed to be

$$\det \underset{\sim}{\Omega}^{-1} = -\frac{1}{2} \tag{2.29}$$

indicating that a valid coordinate transform matrix exists for the one-port transformation of the form (2.4) and is given by

$$\underset{\sim}{\Omega} = \begin{bmatrix} \sqrt{Z_0} & \sqrt{Z_0} \\ \frac{1}{\sqrt{Z_0}} & \frac{-1}{\sqrt{Z_0}} \end{bmatrix}. \tag{2.30}$$

The inverse coordinate transform matrix (2.26) for a two-port network with real port reference impedances Z_{o1}, and Z_{o2} using the power wave expression of (2.4) would be expressed as

[5]Analysis would be similar for n-port devices using network port vectors $(\boldsymbol{\phi}, \boldsymbol{i})$, $(\boldsymbol{v}, \boldsymbol{q})$, $(\boldsymbol{\phi}, \boldsymbol{q})$. For the purposes at hand it is sufficient to consider a nonlinear resistor.

$$\begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2\sqrt{Z_{o1}}} & 0 & \frac{\sqrt{Z_{o1}}}{2} & 0 \\ 0 & \frac{1}{2\sqrt{Z_{o2}}} & 0 & \frac{\sqrt{Z_{o2}}}{2} \\ \frac{1}{2\sqrt{Z_{o1}}} & 0 & \frac{-\sqrt{Z_{o1}}}{2} & 0 \\ 0 & \frac{1}{2\sqrt{Z_{o2}}} & 0 & \frac{-\sqrt{Z_{o2}}}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} . \tag{2.31a}$$

This is a coordinate transformation matrix transforming the terminal $(\boldsymbol{v}, \boldsymbol{i})$ relationship to power waves, i.e.

$$\underset{\sim}{\Omega}^{-1} : (\boldsymbol{v}, \boldsymbol{i}) \mapsto (\boldsymbol{a}, \boldsymbol{b}). \tag{2.31b}$$

S-parameters are short hand notation representing transformation (2.31), but the S-parameters themselves are *not* the transformation. S-parameters define the relationship between $\boldsymbol{a}$ and $\boldsymbol{b}$ as defined in (2.10). The only restriction on the transformation (2.26) and (2.31b) is that an inverse transformation exist. Note that while $\underset{\sim}{\Omega}$ is a linear transform there is no constraint limiting $(\boldsymbol{v}, \boldsymbol{i})$ to be *only* voltages and currents of linear n-ports. It is common to consider the waveform vectors in (2.23) to be functions of time, though they could also be frequency domain amplitudes and phases or, for some nonlinear networks, functions of drive level and termination impedance, for example $i_k = f(v_k, Z_{ok})$. If the terminal $(\boldsymbol{v}, \boldsymbol{i})$ characteristics of the n-port can be described mathematically, (2.26) can be used to obtain a useful network representation.

If an n-port with transformation $\underset{\sim}{\Omega}$ can be described by a function

$$\underset{\sim}{\varepsilon} = \boldsymbol{f}(\underset{\sim}{\eta}), \quad \underset{\sim}{\eta} \in \mathbb{R}^n \tag{2.32}$$

for all $\underset{\sim}{\eta}$ within the device behavioral space, then $\underset{\sim}{\Omega}$ and $\boldsymbol{f}(\,\cdot\,)$ globally characterize the n-port. The global representation (2.32) has general application to many nonlinear networks and for added flexibility can be transformed into a parametric representation of the n-port using (2.25).

If $\boldsymbol{f}(\,\cdot\,)$ is continuous, the n-port can be linearized about an operating point Q by differentiating $\boldsymbol{f}(\,\cdot\,)$ about the operating point specified by $(\underset{\sim}{\varepsilon}_Q, \underset{\sim}{\eta}_Q)$ which results in the $n \times n$ Jacobian matrix $\boldsymbol{J}$, defined by

$$\boldsymbol{J}(\underset{\sim}{\eta}_Q) = \left. \frac{\partial \boldsymbol{f}(\underset{\sim}{\eta})}{\partial \underset{\sim}{\eta}} \right|_{\underset{\sim}{\eta} = \underset{\sim}{\eta}_Q} . \tag{2.33}$$

The linearized representation of the n-port about the operating point Q is expressed as a function of $\boldsymbol{J}(\underset{\sim}{\eta}_Q)$ and defines a linear n-port resistor

$$\hat{\underset{\sim}{\varepsilon}} = \boldsymbol{J}(\underset{\sim}{\eta}_Q)\, \hat{\underset{\sim}{\eta}} \quad , \hat{\underset{\sim}{\eta}} \in \mathbb{R}^n \tag{2.34}$$

with $(\hat{\xi}, \hat{\eta})$ being distinct, dynamically independent linearized port vectors within the device behavioral space $\mathbb{R}^n$, and their magnitudes need *not* be small (Chua, 1978). The global representation (2.32) and linearized representation (2.34) are widely used to simplify behavioral descriptions of nonlinear n-port networks.

To illustrate the advantages of network transformations and the global representation of (2.32), let us consider an ideal pn junction diode with a terminal v-i relationship

$$i(t) = I_S \exp\left(\frac{qv(t)}{kT}\right) \tag{2.35}$$

where I_S is the junction scaling constant, q is the electron charge, k is Boltzman's

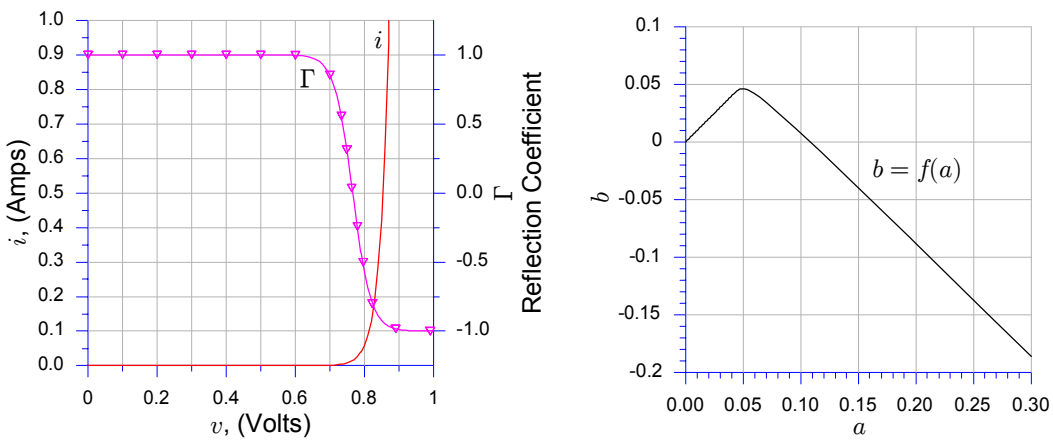


Figure 2.6 Transformation of diode terminal relationships from $(v(t), i(t))$ to $(a(t), b(t))$ domain. The current i approaches ∞ as v increases. The function $b = f(a)$ is well behaved, i.e. bounded linearly. Here a snapshot of the (a, b) relationship could be modeled by a simple linear piece-wise characteristic that describes the diode behavior at the terminals.

constant, and T is the junction temperature. The diode current i as a function of voltage v is shown in Fig. 2.6. As the voltage increases beyond 0.8 Volts the current approaches an infinite value. For computerized network analysis tools to model this ideal diode, the behavior of the unbounded exponential is problematic as it tends to be unstable. Using the power wave transform of (2.4), the corresponding incident and reflected power waves are computed from which the reflection coefficient Γ is computed using (2.18) and is plotted in Fig. 2.6. When the diode is not conducting $\Gamma = +1$ indicating a high impedance (open) and when the diode turns on and current approaches infinity $\Gamma = -1$ indicating a short. In terms of reflection coefficient Γ the behavior of the diode is bounded between $[-1, +1]$. Consider the advantage of representing the terminal characteristics of the diode using power waves by expressing the reflected power wave as a function of the

incident power wave using the global representation form of (2.32)

$$b = f(a) \tag{2.36}$$

which is plotted in Fig. 2.6. Using the global representation and power waves, the behavior of the diode could be expressed using two linear, piece-wise functions, greatly simplifying the behavioral description when compared with using an exponential function[6].

The coordinate transform matrix $\underset{\sim}{\Omega}$ can be constructed for familiar network parameters such as Z-, Y-, $ABCD$-, and S-parameters. Observe that these network parameters represent nothing more than a transformation on network port vectors $(\boldsymbol{v}, \boldsymbol{i})$ to a device behavioral space that is convenient from a network analysis viewpoint. While it may be common in practice to consider Z- and Y-parameters as not having restrictions on the domain of definition, allowing port waveforms to have arbitrary magnitudes, many would be quick to point out that S-parameters are considered to be only a small-signal parameter. Such a constraint on S-parameters is not found in the theoretical foundation but is more a practical constraint on the measurement and application of S-parameters to nonlinear devices operating in a small-signal regime.

2.4 S-Parameters and Nonlinear Networks

Consider the terminal characteristics of a memoryless nonlinear two-port device and how S-parameters could be used to characterize the device behavior[7]. The quiescent drain current I_{DQ} is set for class A operation such that operation is centered within the $I_{DS} - V_{DS}$ plane of Fig. 2.7. An input signal v_G is applied at the gate as shown in Fig. 2.7. The forward transconductance of the device is nonlinear as shown by the I_{DS}–V_{GS} curve. The I_{DS}–V_{GS} curve is a slice taken out of the I_{DS}–V_{DS} plane at a given drain source voltage and is often termed the device characteristic curve. The output current i_D is a function of the input voltage v_G as it is applied across the device characteristic curve. Associated with i_D is the corresponding output voltage v_D. The relationship between i_D and v_D when the device is driving a fixed resistive load R_L is indicated on the I_{DS}–V_{DS} plane and is known as the load-line. The load-line has a slope

$$slope = -\frac{1}{R_L} \tag{2.37}$$

where R_L is considered to be the parallel combination of the device shunt output

[6]For inputs consisting of multiple sinusoids, the power wave transformation would need to be computed for each frequency as it would for a time domain analysis (i.e. harmonic balance). There would still be a computational advantage using power waves because of the elegance of a piece-wise function compared with a Taylor series expansion of an exponential.

[7]Notation in this section is for a MOSFET, but the nonlinear device could be any physically realizable two port.

resistance R_o and the actual load resistance R_L'[8]. For small-signal analysis the output resistance is typically large and $R_L \simeq R_L'$. Adjusting the load resistance R_L will result in different trajectories of the drain current and voltage in the I_{DS}–V_{DS} plane. If the magnitude of the input signal v_G is small, the effects of the nonlinear transconductance will be "linearized" and the output will have a linear relationship to the input as given by (2.34). Linear small-signal analysis performed on nonlinear devices requires the output current and voltage be small so that the relationship between v_G and i_D is linear. As the input signal amplitude increases, the nonlinear transconductance results in a nonlinear distortion of i_D. The device has a quasi-linear region of operation about the bias point, denoted by a box in the I_{DS}-V_{GS} plane and I_{DS}-V_{DS} plane of Fig. 2.7. Within the quasi-linear region of operation, output signals have a linear relationship to the input signals.

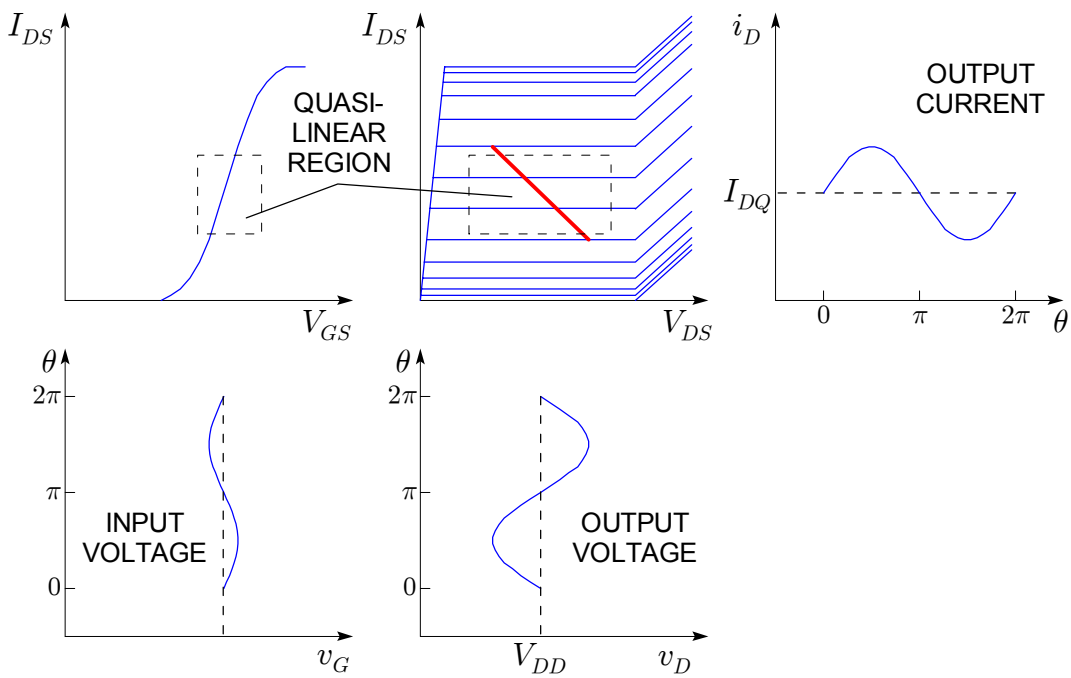


Figure 2.7. Nonlinear device operating in class A linear region.

Further increases in the magnitude of the input signal v_G results in the output current i_D being affected by the nonlinear transconductance. In the I_{DS}–V_{DS} plane the load-line would extend beyond the bounds of the quasi-linear region and sweep through the nonlinear region where the V_{GS} curves are compressed. If the input signal is further increased the load-line would extend to the limits of the device characteristic in the I_{DS}–V_{DS} plane as shown in Fig. 2.8. The output current i_D in Fig. 2.8 exhibits the effects of the nonlinear device characteristic while the output voltage v_D is even further distorted by the physical limits of the nonlinear device at the left and right edges of the I_{DS}–V_{DS} characteristic. On the high current portion of the output cycle for $0 \leq \theta < \pi$ the voltage

[8]A physically realizable device has parasitic capacitance and inductance which forces a complex load line trajectory.

cannot follow the current because the finite device "on" resistance limits the minimum value of v_D resulting in additional output distortion. This physical limit on device output voltage appears as a bend in the load-line. One might ask how the load-line could bend given that R'_L is fixed. As v_D and i_D diverge the apparent device output resistance R_o changes and this in turn alters the value of R_L. During the high voltage portion of the output cycle $\pi \leq \theta < 2\pi$ the device enters the break down region, further distorting v_D. Large-signal inputs distort the output signal, forcing the load-line trajectories against the limits of the I_{DS}–V_{DS} device characteristic. Distortion at the physical limits of the I_{DS}–V_{DS} characteristic is not readily predicted from the I_{DS}–V_{GS} curve as changes in the device output resistance begin to dominate the output characteristics.

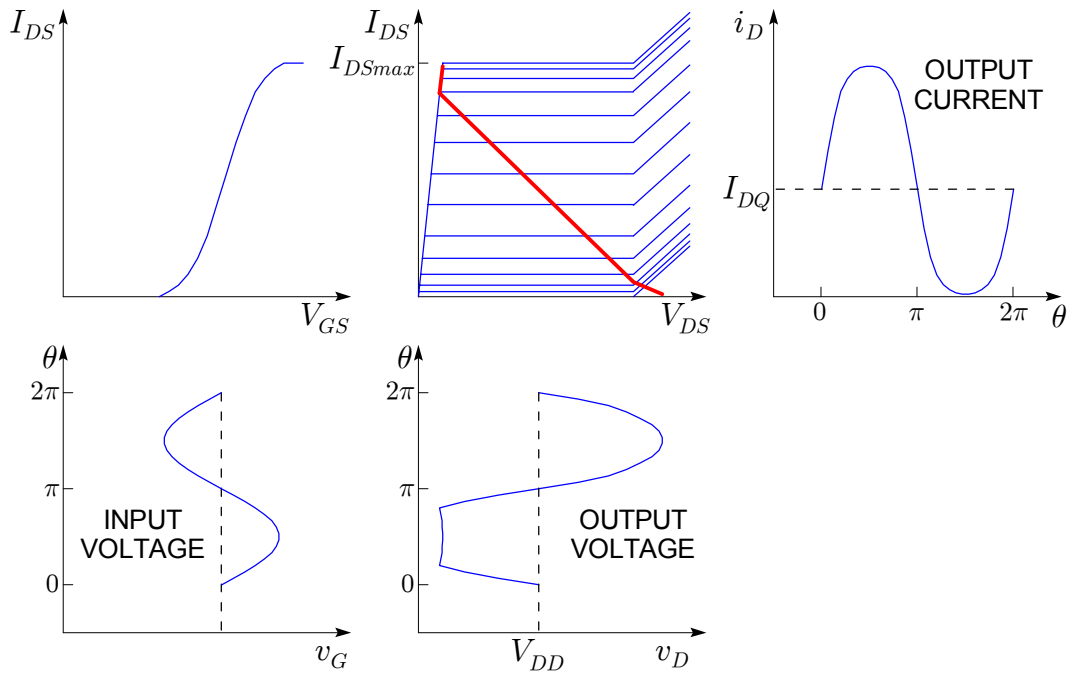


Figure 2.8. Large-signal nonlinear device biased in class A.

A device biased in class A operation always dissipates power, whether amplifying a signal or idling in a quiescent state. A more efficient way to bias a device would be class B, where the device is biased at cutoff, so that a small input signal will turn the device "on". A device biased in class B conducts for only half a cycle and therefore dissipates less power than a class A device. When two class B devices are paired in a push-pull configuration, or the output matching circuit is a resonant tuned circuit, class B operation can yield a good compromise between linearity and efficiency. Class AB is a class B bias shifted slightly more into the active region (more toward a class A operation) and corrects for the cutoff distortion that may be seen in class B operation. Fig. 2.9 illustrates the case where the nonlinear device is biased in a class AB mode of operation. Class AB mode of operation is more efficient than class A operation because the device is on for only half the conduction cycle and maximum output current occurs when there is minimal output

voltage resulting in less power dissipation within the device. The class AB load resistance R_L is smaller than the load used for Fig. 2.7 and Fig. 2.8 resulting in a load-line with a steeper slope. A steeper slope with a bias near cutoff produces more efficient circuit.

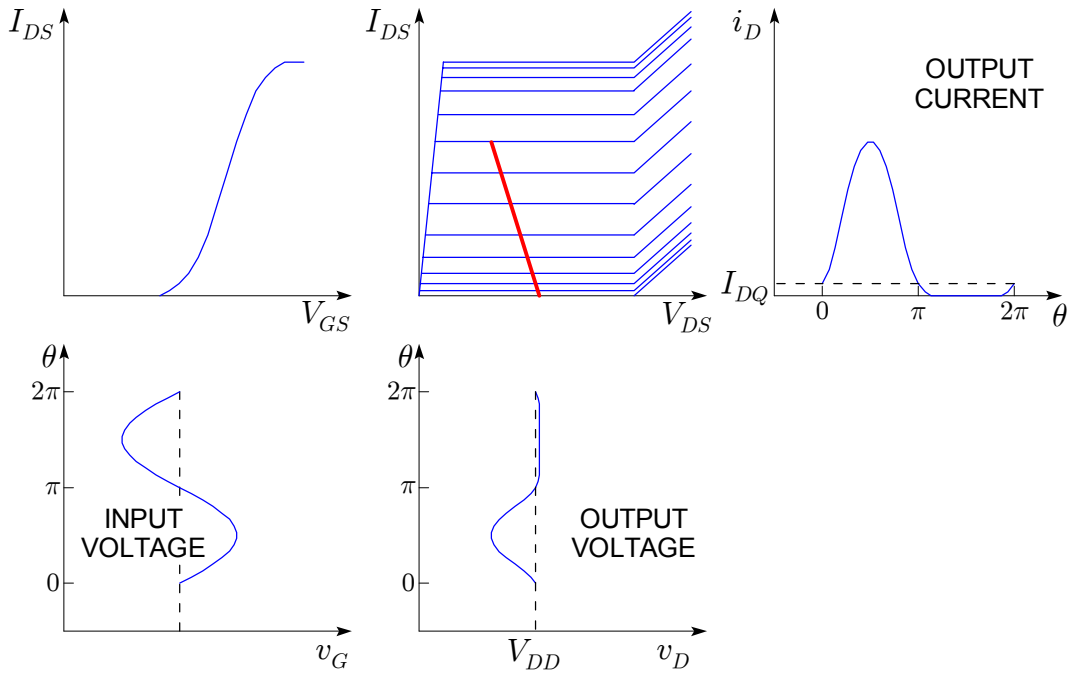


Figure 2.9. Nonlinear device biased in class AB mode of operation.

The bias is selected such that the device conducts for $\theta \approx 180°$. The cycle is completed by either a resonant circuit or a second device. There is some nonlinear distortion associated with the device shutting off for the second half of the cycle, though this is minimal relative to the distortion of the input as it sweeps across the nonlinear transconductance in the I_{DS}–V_{GS} plane. If the amplitude of the input signal were further increased, the physical limits of the device would limit the total amount of current and the load-line would hit the ceiling of the I_{DS}–V_{DS} characteristic curve at I_{DSmax}, the maximum drain source current for a given V_{DS}. Comparing the load-lines of Figs. 2.7, 2.8, and 2.9 it can be seen that a steeper load-line (small R_L) corresponds to a higher output current. The maximum current capability of the device is reached more quickly for small R_L and large input signals. Large R_L corresponds to a load-line which is "more horizontal", or less sloped as shown in Fig. 2.7. When the load-line slope is small, large-signal inputs are more likely to force the output beyond the saturation or break down limits.

The power delivered to the load is

$$P_L = \left(\frac{V_{DS}}{\sqrt{2}}\right)^2 \Big/ R_L \ . \tag{2.38}$$

To increase power delivered to the load, either the dc drain source voltage V_{DD} could be increased, or the load resistance R_L could be decreased. The motivation to squeeze as much linear power out of a device as possible forces the designer to select the highest possible drain to source voltage that will not exceed device breakdown and the smallest possible R_L that will not impact linearity by forcing the load-line trajectory to hit the minimum drain voltage condition (or the maximum drain current condition). Observe that for linearity it may be desirable to have a large R_L for a load-line that does not go into a second region of breakdown.

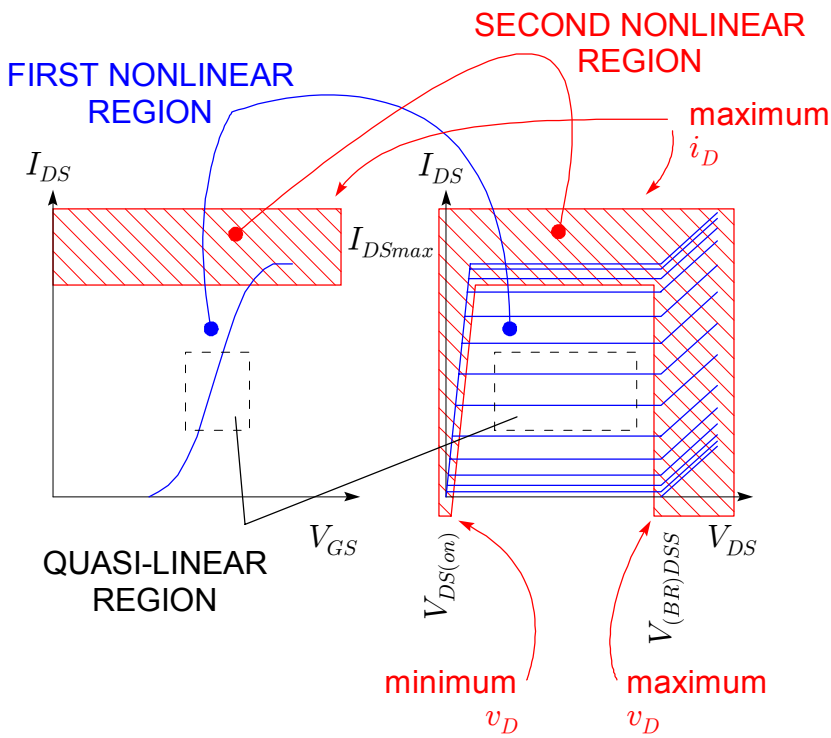


Figure 2.10. Regions of dominant nonlinearities for nonlinear two-port device.

To summarize the nonlinear device operation, it has been shown that output voltage and current can be mapped in the I_{DS}-V_{DS} plane along a load-line trajectory. For small-signal operation the nonlinear device can be approximated as a linear device. For large-signal operation, choosing the bias point appropriately can provide a quasi-linear mode of operation. As the input drive traverses the nonlinear transconductance under large-signal input conditions, the device is physically limited at the boundaries of the $I_{DS} - V_{DS}$ plane which results in further nonlinear output distortion. Fig. 2.10 shows the nonlinearities associated with large-signal operation. Overall the device of Fig. 2.10 is nonlinear, but there is a quasi-linear region of operation where the performance is approximately linear. By properly choosing the bias point and bounding the input signal, device operation can be restricted to the quasi-linear region. Outside this region the device exhibits a nonlinear behavior—for a FET the device characteristic curve would be approximately a square law. As the device physical limits are approached a second nonlinear region begins to dominate.

The second nonlinear region is associated with minimum v_D, maximum v_D, and maximum i_D and occurs as the load-line trajectory departs from a straight line as shown in Fig. 2.8.

Small-signal measurements restrict device operation to a single quiescent bias point in the I_{DS}-V_{DS} plane—usually biased class A and centered in the quasi-linear region of Fig. 2.7 and Fig. 2.10. Traditional large-signal S-parameters (as will be discussed in section 2.5) were intended to describe device operation within the quasi-linear region of operation as shown in Fig. 2.10, (though in practice they are often forced into use well beyond a quasi-linear region). The three-port device model that will be described in Chapter 4 is intended to describe device behavior within the nonlinear region of device operation (unshaded region of Fig. 2.10). As input drive increases and the physical limits of the device are reached, the second region of nonlinearity begins to dominate (shaded region of Fig. 2.10)[9].

Given an estimate of R_L and the device output power P_L, a rough check can be performed to verify if the device is operating within the first nonlinear region, or if it has entered the second nonlinear region. The second nonlinearity associated with minimum v_D, as shown in Fig. 2.10, is the most likely physical limit to be reached first. Using (2.38) the peak magnitude of the drain source voltage can be estimated to determine if a minimum v_{DS} condition is occurring. First determine V_{pk}, the amplitude of the sine wave across the load during one cycle, which will then determine the voltage waveform on the drain

$$\begin{aligned} V_{pk} &= \sqrt{2P_L R_L} \\ v_{DS} &= V_{DD} - V_{pk}\sin\omega_o t \end{aligned} \tag{2.39}$$

with ω_o being the fundamental frequency of the input signal, a good approximation for narrowband networks with high Q. For class A operation if $V_{pk} \geq V_{DD}$ the trajectory along the load-line extends into the second nonlinear region. Class B and AB bias halves the allowable V_{pk} because the magnitude of the fundamental output signal will be between the supply rails for operation within the first nonlinear region. From Fig. 2.9 it can be seen that for a class AB bias, $2V_{pk} \geq V_{DD}$ would mean that minimum v_{DS} is being reached.

Maximum allowable v_{DS} is usually stated on the device data sheet as $V_{(BR)DSS}$, the drain source voltage for which device breakdown begins to occur. If the drain dc bias voltage is chosen too high, the second nonlinear region may be entered for large-signal inputs. If the output circuit is resonant, this breakdown may be a source of distortion, even if the device is in cutoff. A sufficient condition to check if the drain source breakdown voltage is exceeded is

$$V_{pk} + V_{DD} \geq V_{(BR)DSS}$$

[9]The three port model is robust enough that device operation can be modeled into the second nonlinear region, but as the secondary nonlinearities begin to dominate the accuracy of the model begins to suffer. A second nonlinearity could be modeled by adding an additional port to the three-port model creating a four-port model.

or, expressed in terms of the peak output signal from (2.39)

$$V_{pk} \geq V_{(BR)DSS} - V_{DD} \quad \text{for Class A bias.} \tag{2.40}$$

The maximum v_{DS} condition for class B bias is primarily a function of peak output signal when biased near cutoff

$$\begin{aligned} 2V_{pk} &\geq V_{(BR)DSS} \quad \text{or} \\ V_{pk} &\geq V_{(BR)DSS}/2 \quad \text{for Class B bias.} \end{aligned} \tag{2.41}$$

It is more likely that the minimum v_{DS} condition (V_{sat}) will be reached before the maximum v_{DS} condition is exceeded for most class B bias choices because of the high currents and small load impedances.

With a value for I_{DSmax}, a check can be performed to determine if the value of i_D is entering the second nonlinear region of operation. From P_L and R_L, the peak current becomes

$$I_{pk} = \sqrt{\frac{2P_L}{R_L}} \tag{2.42}$$

The conditions on I_{pk} to determine if device operation has reached a physical limit would be

$$\begin{aligned} 2I_{pk} &\geq I_{DSmax} \text{ for Class A} \\ I_{pk} &\geq I_{DSmax} \text{ for Class B.} \end{aligned}$$

These checks on the operating limits of a device with high-input drive level are typically given as limits on the device data sheet and designers usually select the device and/or circuit parameters to avoid pushing these limits. For a device operating with large-signal levels but within the bounds of the first nonlinear region of Fig. 2.10, it is feasible to extend S-parameter design techniques from the small-signal linear realm of design to the large-signal region of operation.

The power dissipation of the device may also impact the ability to characterize and predict device performance. Large-signal drive levels and high bias currents will result in device self-heating which may significantly affect the device performance.

2.5 Large-Signal S-Parameters

Early attempts to extend small-signal S-parameter measurements to large-signal levels has given rise to what are referred to in the literature as "large-signal S-parameters". Researchers experimented with the characterization of devices and circuits operating in class C by extending small-signal S-parameter analytical design tools. Three different methods of measuring large-signal S-parameters are reported in the literature. A group of researchers at Sandia Laboratories used a high-power signal generator to measure

S-parameters of a BJT with a conventional test setup using a reference impedance of $4\,\Omega$ and $50\,\Omega$ (Leighton, 1973), (Chaffin, 1973). Care was taken to ensure that the output signal of the device was sinusoidal and that there was agreement between the $4\,\Omega$ and 50 Ω reference impedance S-parameters when renormalized and compared in a common reference impedance. If the S-parameters were in agreement when in a common normalization impedance, then assumptions about linearity could be maintained. A drawback to the measurement approach was that s_{12} and s_{22} could only be measured by switching the signal source from port 1 to port 2, in effect changing the device operation. The reported results showed only slight variation in s_{11} and s_{22} with input power, while s_{21} exhibited the largest change with input power. The large-signal S-parameters were used to design class C BJT amplifiers at 500 MHz and 1000 MHz, (Webb, 1973). The final results were satisfactory, recommending the large-signal approach as a useful design tool.

Harmonics were terminated during measurement of the large-signal S-parameters with the same harmonic terminations used in the final design. Measurements were made with and without filtering of harmonic components. For the reported measurements, the harmonics on the transistor output were very low and filtering did not make an appreciable difference in the results. The high-frequency performance of the transistor die and package could have been a factor in limiting the amount of high-frequency distortion seen during the measurements.

An improvement on measurement of class C devices was later demonstrated using two signal sources simultaneously driving ports 1 and 2 (Mazumder, 1978). In section 2.2, S-parameters were derived by switching off sources at all ports except one. By limiting the derivation to one source, the equations were greatly simplified. From a measurement standpoint, the cost and complexity of the measurement system is minimized when using a single source. The use of multiple signal sources enables nonlinear devices (for which superposition does not hold) to be characterized without making any linearization assumptions. The measured values of s_{12} and s_{22} were closer to the actual device characteristics under typical operation by driving the device with two signal sources. In the work of Mazumder, amplifier design using the large-signal S-parameters measured with two signal sources gave improved design results at 2 GHz compared with measurements made with one signal source. The measurements were made with the same harmonic terminations as used in the final amplifier design. Care was taken to ensure that the output signals were approximately sinusoidal.

An elegant approach to large-signal measurement of class C devices was reported by Müller in a German technical journal (Müller, 1968). Müller's measurement setup used a single driving source, but allowed for measurement of s_{12} and s_{22} while the device was being driven from port 1, providing the advantage of nonlinear measurement capability without the added equipment cost and complexity. One drawback to the method was that it required an additional measurement step and calibration of the system would have benefited from the use of a computer, a luxury that was not widely available at the time.

As conceived, large-signal S-parameters were an attempt to extend small-signal S-parameter measurement and design techniques to high-power devices. Early

practitioners were familiar with fundamental S-parameter theory and limitations. Great care was exercised in the measurement and use of the data to achieve valid design results. By forcing a linear approximation on the nonlinear device characteristics, an acceptable engineering trade-off, S-parameter network theory could be applied to large-signal circuit design. The alternatives were no S-parameters, or S-parameters restricted to a quasi-linear region of operation. Successful application of S-parameters to a design was dependent on understanding the design goals and configuring device measurement to closely approximate the final design objectives given limitations of the measurement equipment, responsibilities that cannot be over emphasized.

To an extent, traditional large-signal S-parameters *are* dependent on the source and load terminations under the assumption that the system is nonlinear and power-dependent. In the cases reported in the literature great care was taken to ensure that the device was operating in a "linear" region, i.e. harmonics on the output were minimal and S-parameters did not change appreciably with input level and source/load terminations. In terms of Fig. 2.10, efforts were made to maintain device operation within the quasi-linear region. In the years since initial work was done on large-signal S-parameters it is difficult to assess how successfully the approach has been used in practice. It is apparent from the literature that the care used in the original large-signal S-parameter work has not always been exercised by those making large-signal S-parameter measurements. Vendelin, et al., state that large-signal S-parameters are an "obvious" approach to characterizing nonlinear devices (Vendelin, 1990). The authors commented on how "easily" device S-parameters can be obtained by simply measuring at elevated power levels. Following brief comments on selection of power level and load terminations the authors were quick to point out the "negative aspects of using large-signal S-parameters." Rather than using S-parameter theory to guide the selection of input power level and load terminations as was done by Leighton (Leighton, 1973), Vendelin used a simplistic definition of S-parameters to argue that the load and power dependence of a nonlinear device seriously limit the usefulness of large-signal S-parameters. While Vendelin, et al. are critical of large-signal S-parameters, if one were to follow their cursory comments, any large-signal S-parameters obtained would be almost certainly unusable outside the narrow region they define. Another author has written that traditional large-signal S-parameter techniques are "a brutal attempt to force nonlinear circuits to obey linear circuit theory" (Maas, 1997). Large-signal S-parameters as originally conceived were more an attempt to force linear measurement systems to work with nonlinear devices. Proponents of nonlinear lumped-element device models tend to dismiss the idea of using large-signal S-parameters. Despite the widespread availability of nonlinear device models and tools, S-parameters may still be the front line approach to nonlinear device characterization and design challenges.

Concerns about linearity restrictions on nonlinear devices aside, large-signal S-parameters do present a formidable barrier to a design engineer who may not have the measurement and network theory background (not to mention time and resources) required to successfully obtain a useful set of measurement data. Experiments with large-signal S-parameters required expertise with high-power measurements, vector network analyzer calibration techniques (which were still being developed), and network theory to

obtain useful parameters from a collection of time-consuming measurements that demanded a great deal of attention-to-detail. Many years later, the challenge of obtaining large-signal S-parameters has not become easier.

2.6 S-Parameters Applied to Class B Power Amplifiers

Applying S-parameters to class B power amplifier design would equip the power amplifier designer with a convenient and powerful tool. The truest behavioral description of the device requires that it be characterized while operating under large-signal conditions. However, if the class B operating point provides an approximate linear response at the fundamental over a wide range of drive levels for varying loads, why not use S-parameters to describe the terminal behavioral characteristics? A first-cut design could be achieved by using input and output tuning stubs (i.e. source- and load-pull) to measure the optimally loaded device running at the rated output power. Load-pull measurement data cannot be used to predict device performance if drive and source/load conditions are varied from those of the original measurements. It would be preferable if the class B amplifier could be designed for high output power operation using S-parameters measured at a lower power level to ease the requirements for equipment capable of handling a potentially unstable device at high power levels. If the measurements require the device to be operating at or near the rated output power then there is little advantage to measuring and using S-parameters in the design process.

Useful S-parameters for a device operating in class B can be obtained by measuring at an average current lower than that encountered under actual operating conditions provided that the S-parameters are not critically dependent on the termination impedances and drive level. The effects of self heating can be minimized by measuring the device at a temperature close to the intended design operating temperature. The nonlinear operation of the device may not be adequately determined from the S-parameters alone, but for most applications where linear performance is desired, the S-parameters may yield acceptable results.

A suitable class B bias point must be selected such that the amplifier is linear at the fundamental input frequency before measuring the device S-parameters. To illustrate this point, consider the I_{DS} vs. V_{GS} curve in Fig. 2.9 for a MOSFET. The transfer characteristic curve models the MOSFET as a square law device with $I_{DS} = k(V_{GS} - V_T)^2$ until the onset of saturation. For class B operation the device is biased as close to the turn-on threshold as possible so that conduction angle $\phi \approx 180°$.

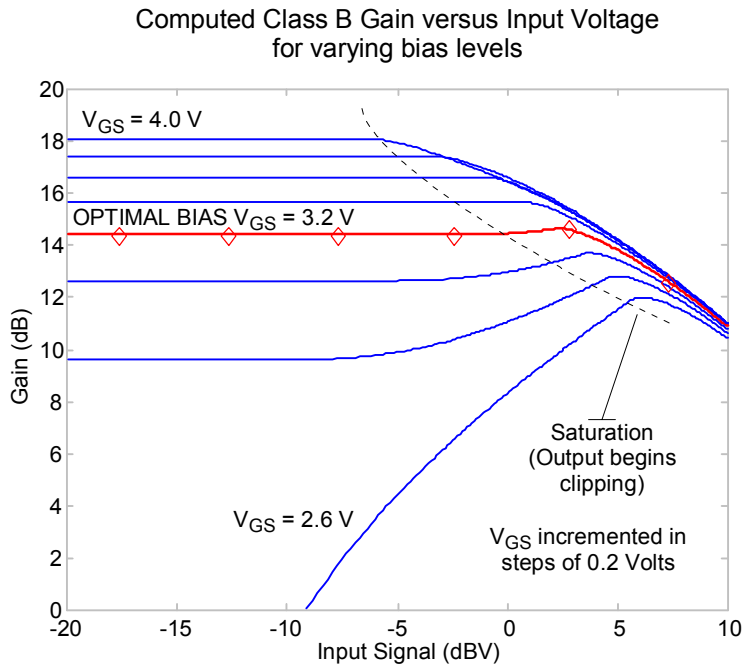


Figure 2.11. Effect of dc bias on class B linearity. Gain curves are computed at the fundamental frequency for varying dc bias voltages. Harmonics are not considered. Optimal bias results in a flat gain.

Simulating the transfer characteristic of Fig. 2.9 and adjusting the bias results in a plot of gain vs. input signal level as shown in Fig. 2.11. Fig. 2.11 shows the ratio of the output voltage to the incident voltage ($|s_{21}|$) as a function of the gate source bias voltage and input signal. The input signal voltage is swept from a small-signal level to the level where the device output saturates and begins clipping the output voltage waveform. Note that at each bias level the gain varies as a function of input power. As seen in Fig. 2.11, for a swept input level there is a dc bias point for which the gain is optimally flat[10]. Below the optimum point there is gain expansion and above the optimum bias point the gain begins compressing early. When optimally biased, the device will exhibit the efficiency of a class B amplifier while maintaining the linearity of a class A amplifier at the fundamental input frequency over a wide range of input levels, a bias condition often referred to as class AB.

With the device biased to achieve a maximally linear gain profile, the S-parameters s_{11} and s_{21} can be measured directly. However, the parameters related to port 2 must be measured with some care. Previous researchers noted that measurement of s_{12} by driving port 2 with the device biased near cutoff resulted in a parameter of questionable value. Class A bias could be used to make the measurement of s_{12} (Leighton, 1973). Measurement of s_{22} by driving port 2 while the device was biased near cutoff did not

[10]The term optimally flat is used here to indicate flat gain for a wide range of input drive level. Note that for all gain curves there is a slight gain expansion just prior to compression.

account for any variations in bias and the researchers concluded that s_{22} was not a strong function of bias. The preferred approach would be to measure s_{12} and s_{22} while the device is biased appropriately in class B and being driven from port 1 (*not* port 2 which is an artificial condition and would not be encountered in practical use). Measuring parameters associated with port 2 while the device is being driven from port 1 is often referred to as "hot s_{22} measurement."[11]

To determine if changes in frequency and/or load terminations impact device linearity for an optimal class B bias, conditions which would invalidate underlying linearity assumptions, measurements were made with input signals swept from the small- to the large-signal region of operation. The transmission, τ, for an MRF166W as a function of bias and swept input power was measured. The MRF166W is a typical push-pull device with two 20 Watt transistor die in a single package as shown in Fig. 2.12. For the experiments, one transistor was measured while the gate, drain, and source of the other

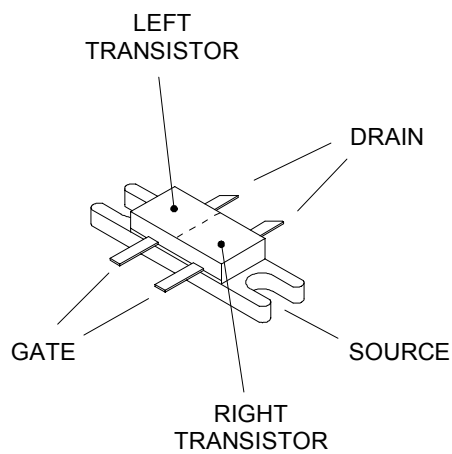
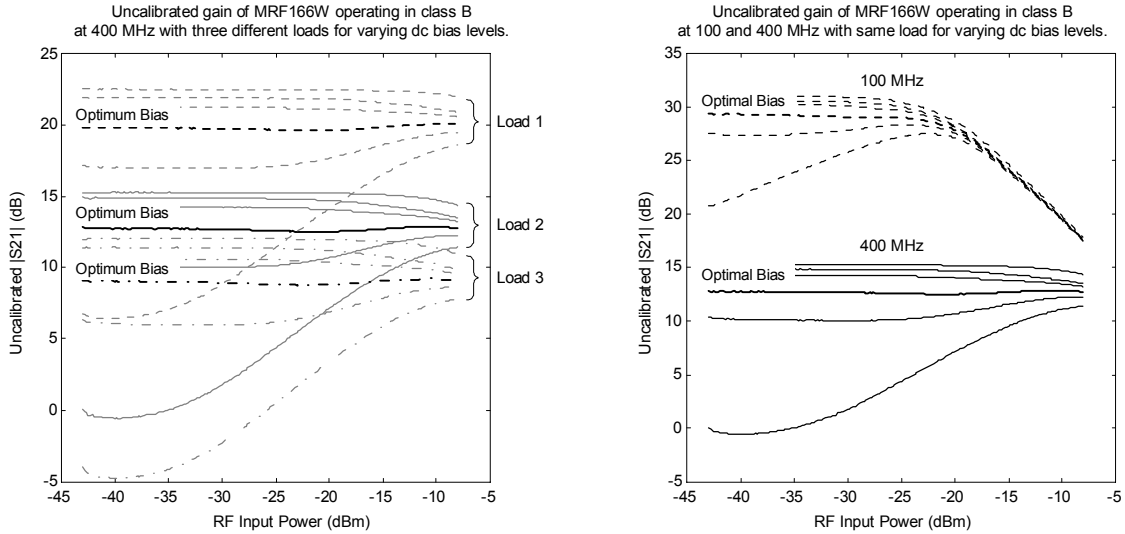


Figure 2.12. MRF166W push-pull transistor rated at 40 Watts output power at 500 MHz.

were grounded. Three power sweeps were run at 400 MHz for three different load terminations with the results shown in Fig. 2.13a. Two power sweeps were run at 100 MHz and 400 MHz for a single load termination as shown in Fig. 2.13b. The measured curves in Fig. 2.13 are similar to the theoretical curves in Fig. 2.11. Load and frequency variations can be seen to have little effect on the linearity of the gain until compression for an optimal linear bias.

With an optimal linear bias, the MRF166W was measured at 400 MHz with a high input drive level and two different loads. Load 1 is approximately $Z_1 = 50\,\Omega$ and load 2 is approximately $Z_2 = 12.5\,\Omega$. Load 2 was chosen to maximize the available output power before the onset of clipping on the drain output voltage. The transmission τ, plotted as a function of input power in Fig. 2.14, remains almost constant across a wide range of input power indicating that the device is operating in a linear region. At higher

[11]A measurement technique to properly measure devices operating in class B is discussed in detail in chapter 3.



(a) (b)

Figure 2.13. Effects of load and frequency variation for optimal linear class B bias. Frequency is held constant in (a) while load terminations are changed. The optimal bias is the same for the three loads. Changing frequency for a constant load termination in (b) results in the optimal bias point being constant with frequency.

output power levels, the magnitude and phase of transmission τ are no longer constant as the transistor gain compresses[12].

For an optimally linear bias, the drive level at port 1 could be chosen across a wide range of input levels during S-parameter measurement provided gain compression was not occurring. For a fixed load, the higher device gain at 100 MHz results in the curves of Fig. 2.13b compressing at a lower input power than for the 400 MHz case. The average drain current in Fig. 2.13 varies as a function of input drive level from 200 - 800 mA. Under optimal loading conditions at its rated output power, the MRF166W would have a higher average current.

The harmonic output for both loads are given in Fig. 2.15. The harmonics are well below the carrier. Load 2 had higher harmonics at 1200 MHz whereas load 1 only had a harmonic at 800 MHz with any additional harmonics below the noise floor of the spectrum analyzer.

Based on the peak RF drain voltage and power dissipated in the loads it would be reasonable to consider gain compression due to thermal heating as the cause for changes in τ for the high input drive levels shown in Fig. 2.14. For high-power devices, the thermal properties of the device will heavily influence the measurement results, most often

[12]Data in Fig. 2.14 is consistent with the linearity requirement that device performance not change for a variation in load termination. The variation in magnitude and phase of τ is similar for both loads. The offset in magnitude and phase of τ for both loads is a result of different normalization impedances $Z_0 = Z_1$ for load 1 and $Z_0 = Z_2$ for load 2.

negatively. Measured performance will more closely match operational performance if the temperature of measurement is close to the temperature of operation, often $T > 25\,\mathrm{C}$. One way to elevate device temperature during measurement is to increase the average current through the device so that it is close to the current seen by the device during operation (Leighton, 1973).

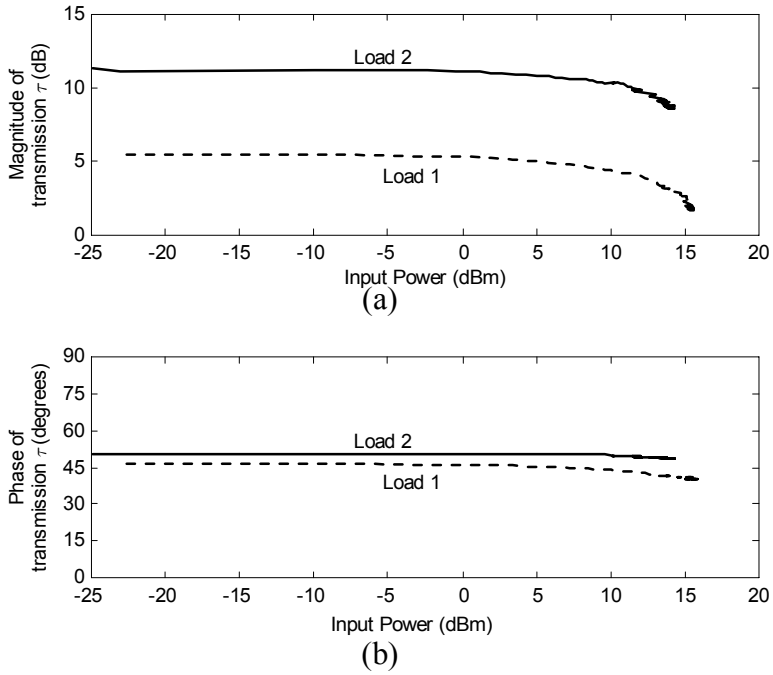


Figure 2.14. Linearity of class B transmission τ at high input power levels.

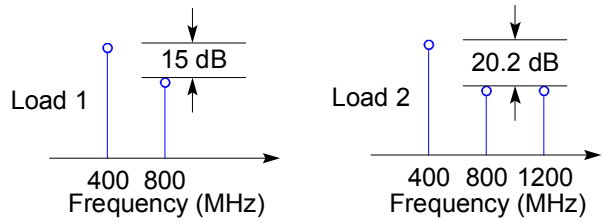


Figure 2.15. MRF166W Class B output spectrum for loads 1 and 2.

For the class B measurement, simply biasing the device with a quiescent current equivalent to the average operating drain current may exaggerate thermal effects. Under large-signal drive conditions a significant portion of the average current is dissipated in the load, not in the device. Increased device self-heating due to an artificially high bias current could degrade performance, skew the S-parameter measurements, and shorten the lifetime of reliable operation. Careful selection of the dc bias current could replicate thermal conditions encountered during normal operation and account for thermal heating effects which will impact the S-parameters. A properly mounted device with an adequate heat

sink may not exhibit appreciable thermal effects on the S-parameters until driven into nonlinear compression.

S-parameters were measured for an optimally linear class B bias and for a high bias current equivalent to the average class B operating current. The S-parameters as a

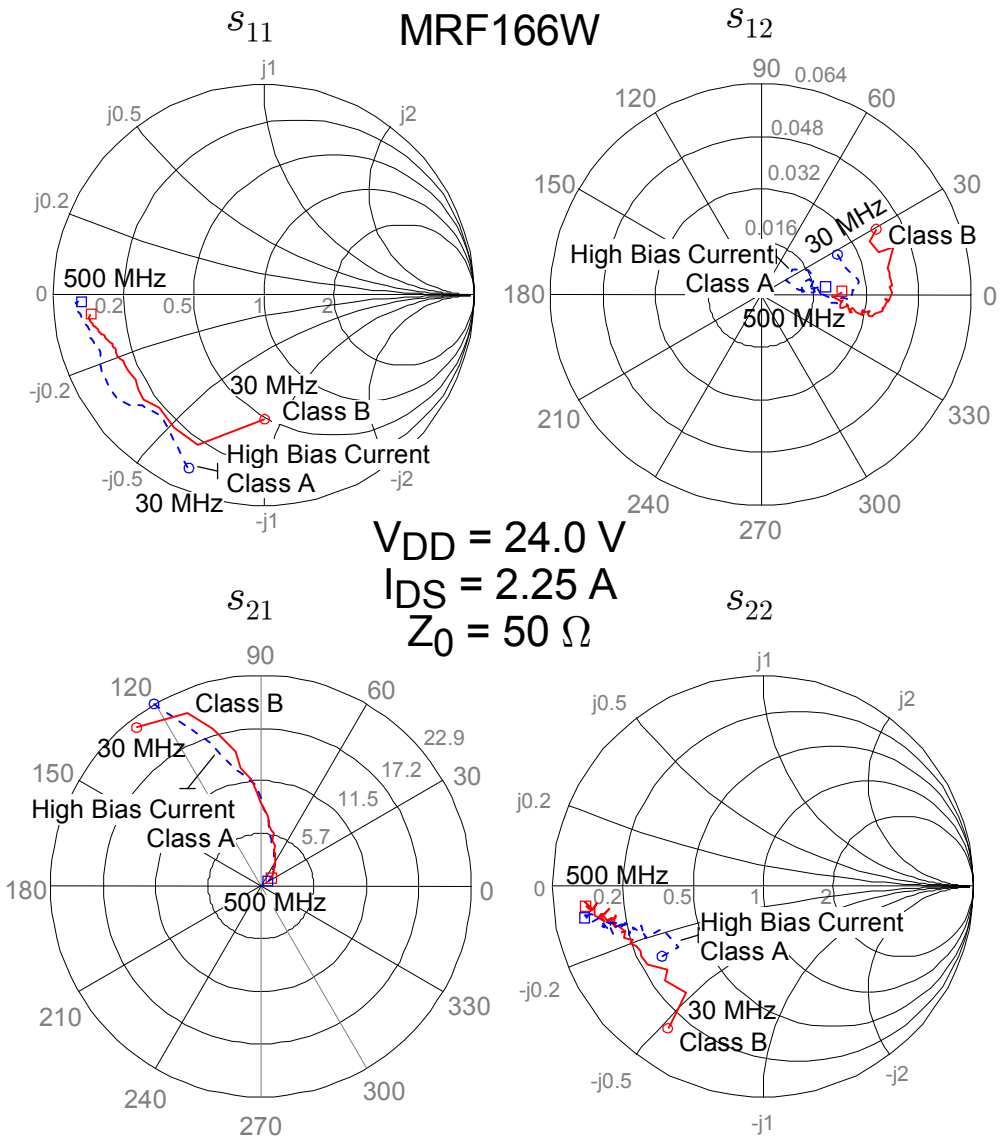


Figure 2.16. Comparison of S-parameters for linearized class B and high bias current class A ($I_{DS} = 2.25$ A).

function of frequency are plotted in Fig. 2.16 for both a linear class B bias and a high-current class A bias. As the frequency increases, the two S-parameter measurements begin to converge, suggesting that at the higher frequencies, designers may be able to use the high-bias current class A S-parameters as a substitute for optimally linear class B S-parameters. The differences for s_{11} and s_{22} at the lower frequencies are significant.

The device parasitic input and output capacitances will change substantially from a class B bias to a high current class A bias. Bias dependent parasitics may affect gate and drain impedances more at lower frequencies than at higher frequencies where structural parasitics dominate, suggesting that s_{11} and s_{22} show differences in the two measurements. Across a broad frequency range it is expected that the optimally linear class B parameters would give the designer a more accurate set of design parameters.

From the class B measurement results we see it is possible to measure the S-parameters of a device operating in class B and use the results to design high power amplifiers. There are slight differences in the device characteristics when biased in class B compared with a high-bias current class A measurement approach.

This chapter has reviewed the fundamentals of S-parameters and their application to network representations. Existing techniques for characterizing devices with S-parameters have been summarized and measurement results for class B and high-bias current class A measurements have been presented. The emphasis in this presentation has been on linearized devices. The linear restrictions are lifted in the following chapters.

CHAPTER 3

MICROWAVE MEASUREMENTS

Limitations of commercial microwave measurement systems restrict active device measurements to a region of small-signal linear operation. The need to measure nonlinear devices as a function of frequency and power has led to the development of the Reflect-Thru-Line (RTL) measurement system (Davis, 1993). The RTL measurement architecture provides a safe, effective means, of obtaining two-port scattering data on potentially-unstable nonlinear devices that are otherwise difficult (or dangerous) to measure using conventional microwave measurement systems. A theoretical overview of RTL calibration and measurement is discussed along with some practical aspects of device measurement. The potential and practical uses for the RTL measurement system is illustrated through the presentation of measured device data. The chapter concludes with the derivation of equations to normalize S-parameters taken using the RTL system to arbitrary impedance levels, a useful and practical aspect of S-parameter measurement that finds wide application in many areas of microwave measurement.

3.1 Microwave Measurement System

Two-port scattering parameter measurements require sampling the reflection and transmission of a test signal at ports 1 and 2. Fig. 3.1a illustrates a two-port measurement system where both ports are driven simultaneously by two sources. The impedance at the reference plane looking toward the generators is determined by Γ_S. For practical reasons, most measurement systems do not use dual sources. Adding an additional signal generator or splitter impacts the total cost and complexity of the measurement system. A common solution is to switch a single source between ports 1 and 2 as shown in Figs. 3.1b and 3.1c. With the aid of directional couplers and relays, the measurement system switches the generator and load to the second port as shown in Fig. 3.1c. The port not being driven is ideally terminated with a load $\Gamma_L = Z_0$ so that switching between port 1 and port 2 will not alter the impedances seen by the device. With the source and load reflections zero, the measurement of the transmission and reflection at each port directly gives the scattering parameters. Port 1 and 2 measurements require $a_2 = 0$ and $a_1 = 0$

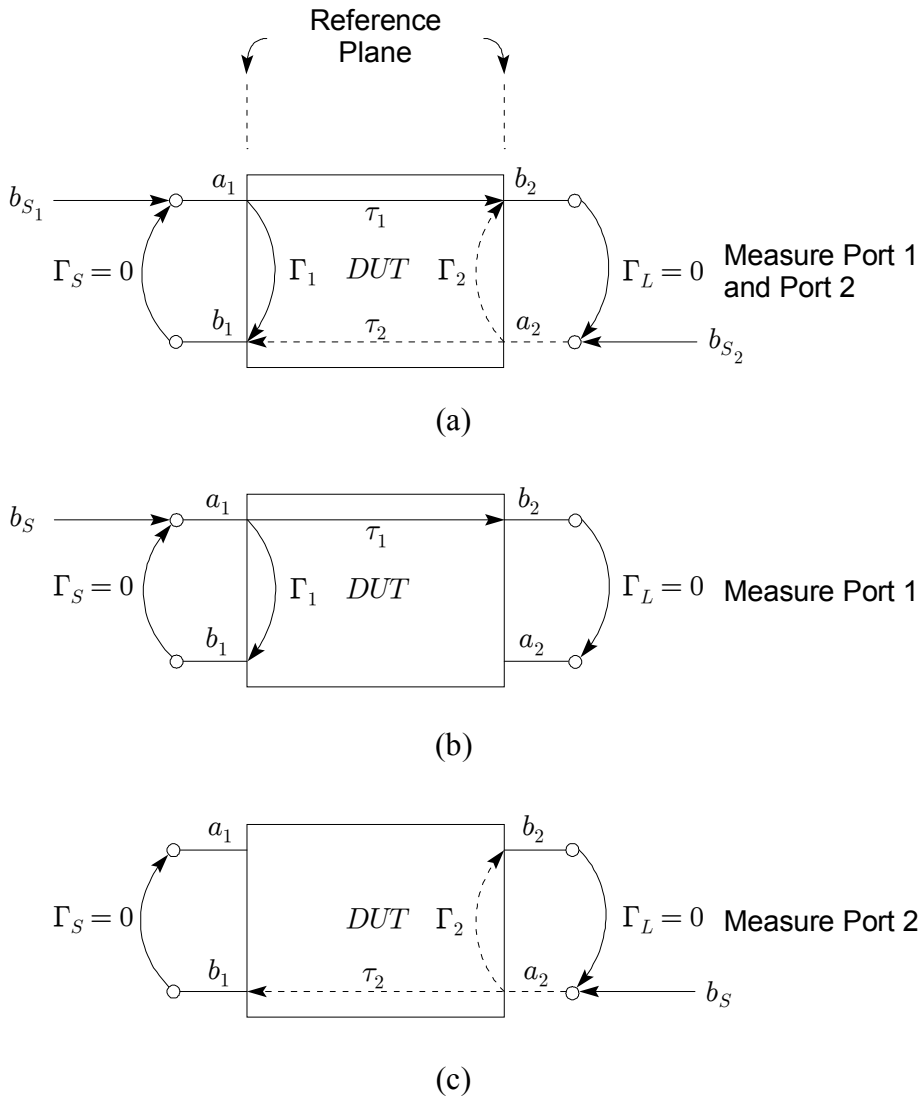


Figure 3.1. Diagram of an ideal two-port scattering parameter measurements.

respectively. A measurement system that switches a single source between two ports is intended primarily for measurement of linear systems. Superposition must hold for the DUT in order to obtain valid linear measurements. An accurate characterization of the device under test may be difficult to achieve if a nonlinear two-port device is measured in a system where superposition is implicit. Two-port measurements of a nonlinear device requires a system similar to that shown in Fig. 3.1a where both ports are driven and measured simultaneously. One solution for driving the two-ports without using a second signal generator has been reported (Takayama, 1976) whereby the signal generator is followed by a splitter and phase shift network simultaneously driving ports 1 and 2 by a single signal generator.

Another alternative to dual signal generators would be to use a single generator and at least two different loads, Γ_{L_1} and Γ_{L_2} as illustrated in Fig. 3.2. The load reflection a_2 is used to probe port 2 and serves the same purpose as a second RF signal generator as in the system of Fig 3.1a. Four sampling receivers measure values for a_1, b_1, a_2, and b_2 from which the measured reflection and transmission coefficients Γ_1, Γ_2, τ_1, and τ_2 are obtained for the two independent loads Γ_{L_1} and Γ_{L_2}. The two loads Γ_{L_1} and Γ_{L_2} used in the measurement process are required to be independent, but do not need to be known precisely, since a_2 and b_2 are measured for each load. Though the measurement system requires only a single RF source, a two-port nonlinear device can be meaningfully measured. The RF source is not subjected to high incident power levels when high-power RF devices are measured (unlike the system of the type shown in Fig. 3.1a where the port 2 signal generator is directly in the path of high power output). Extra attenuation may be added between the DUT and the measurement path of Fig. 3.1b to avoid overdriving the b_2 receiver during high-power measurements. When the RF source is switched in Fig. 3.1c, additional protective attenuation is needed directly in the signal path and hinders port 2 measurements.

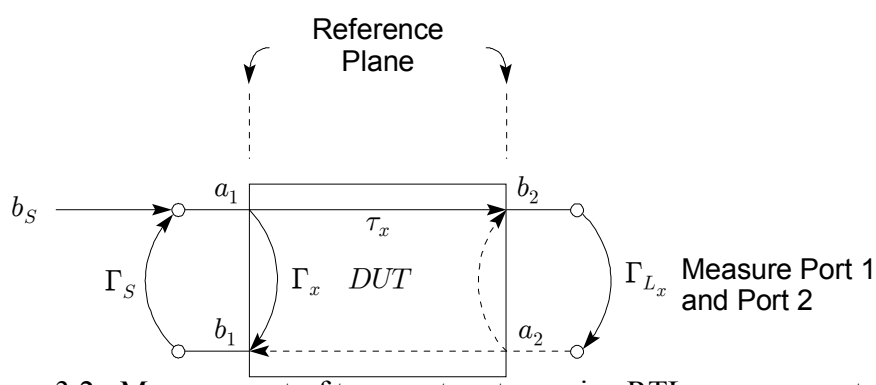


Figure 3.2. Measurement of two-port system using RTL measurement setup.

The Reflect-Thru-Line (RTL) measurement system provides an environment where the drive signal can be safely decoupled from the measurement path without extra attenuation in the direct signal path. The ability to measure high power nonlinear two-port devices makes this system well suited for power amplifier measurement. Freedom to select independent loads for the measurements provides an added degree of flexibility. Loads can be selected to enhance stability, minimize saturation, or satisfy similar constraints that might be encountered in a fixed termination impedance system.

Fig 3.3 shows the RTL measurement system signal flow diagram and Fig. 3.4 shows the complete system using directional couplers. A four port sampling receiver is used to measure the signals a_{1m}, b_{1m}, a_{2m}, and b_{2m}. A prototype system was constructed using a modified S-parameter test set. Coaxial relays provided necessary switching to measure a_{1m}, b_{1m}, a_{2m}, and b_{2m} with a dual channel magnitude-phase measurement system (a vector voltmeter and a network analyzer have both been used). For measurement of high power input and output signals, attenuation can be placed directly in the measurement path without affecting the drive signal path to the DUT. The additional attenuation would

be included in the calibration as e_{40} and error correction terms e_{51}, e_{62}, and e_{73} as shown in Figs. 3.3 and 3.4. The measurement system uses the Reflect-Thru-Line (RTL) (Davis, 1993) calibration and measurement technique to correct for systematic measurement errors due to losses inherent in the measurement system. The RTL calibration technique uses readily constructed standards to achieve a system accuracy comparable to TRL calibration. The RTL calibration and measurement procedure are ideal for a measurement system where a single source is used without switching the source from port 1 to port 2. Section 3.2 discusses the RTL calibration method and a full derivation of the procedure is included in Appendix A.

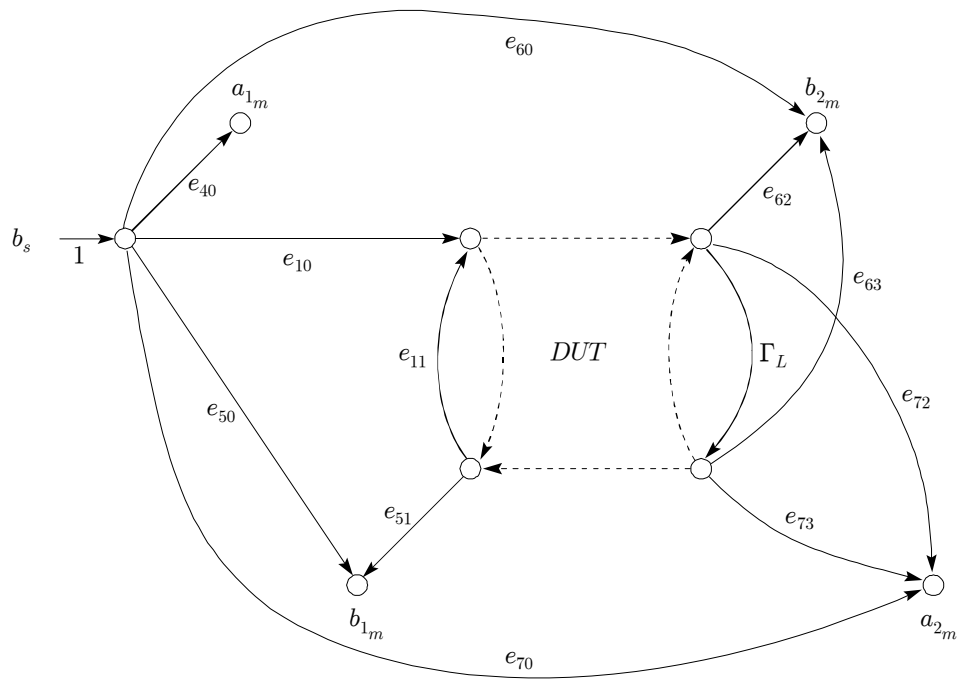


Figure 3.3. Signal flow diagram of RTL measurement system.

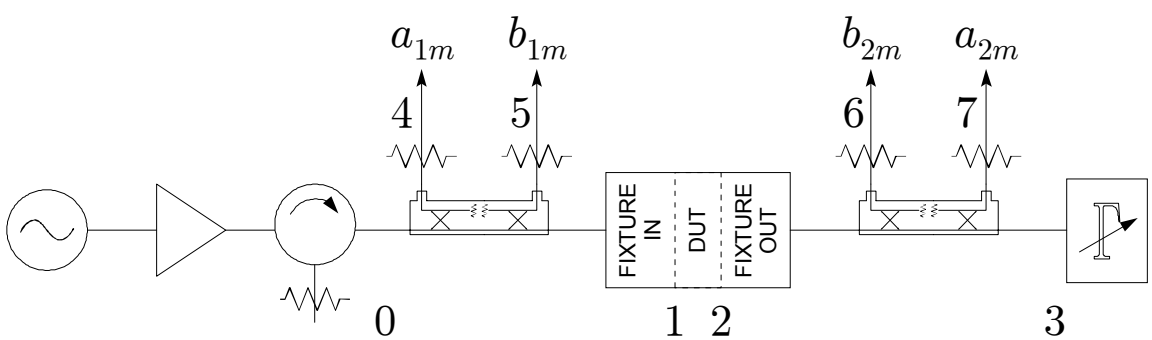


Figure 3.4. One possible realization of the RTL measurement system.

The ability to vary the load Γ_L, seen by the DUT at the measurement reference plane, simplifies load-pull measurements or device parameter extraction methods that require

multiple loads. The actual value of Γ_L is known at any given time because of the ability to measure both b_{2m} and a_{2m}. The RTL measurement system works equally well for linear and nonlinear devices under small- *and* large-signal drive conditions.

3.2 Correction of Systematic Measurement Errors

The scattering parameter measurements are taken at ports away from the desired reference plane. As seen from Fig. 3.3, it is not possible to directly obtain the desired signals at the DUT reference plane. Losses in directional couplers, cables, connectors, and switches are embedded in the measurements and are modeled as error terms e_{ij} in the signal flow graph of Fig. 3.3. The error terms are considered time invariant systematic measurement errors. The effects of the error terms e_{ij} must be removed from the measured values to obtain scattering parameter values at the reference plane. Measurement system calibration is the process of determining the values of the error terms e_{ij}. The derivation of the error terms for a realizable measurement system is given in Appendix A and is summarized below.

If the measurement instrument were ideal there would be no need to consider the error terms e_{ij}. It is difficult to make perfect instruments and perfect calibration standards. A perfect microwave measurement system without sources of error is an unreachable goal. As the ability to make accurate microwave measurements has improved, the task of constructing a perfect system has become even more difficult. The popular open-short-load-thru (OSLT) method of network analyzer calibration relies on precise, highly repeatable standards—open, short, load, and thru. Network analyzer calibration kits include documentation indicating the accurately measured values of the individual standards. The measured calibration standards are stored directly in the network analyzer memory or as a model for use with a computerized calibration routine.

As measurement system theory evolved, researchers began investigating alternative calibration methods that would require neither a perfect instrument or perfect calibration standards. One early method proposed was the "thru-short-delay" (TSD) (Franzen, 1975) which required known precision standards. A thru, precision short, and unknown length of precision non-reflecting delay line are easier to construct and characterize than an open. TSD requires the values associated with the standards to be known in advance.

In the late 1970's researchers at the National Bureau of Standards (now the National Institute of Standards and Technology, NIST) introduced the "thru-reflect-line" (TRL) two port calibration method (Engen, 1979). TRL was an improvement to the TSD calibration method in which several of the terms in the TSD error model were set to zero with no significant effects. TRL does not require *a priori* characterization of the calibration standards. The measured values of the calibration standards are obtained as an artifact of the calibration process. The thru standard is achieved by directly connecting port 1 to port 2. A short of known reflection is not required. In its place, a reflection standard with a high reflection coefficient such as an open or a short, may be used. The standard needs to have a reflection coefficient $\Gamma_L \gg 0$ (with phase known to within $\pm 90°$). The line standard is used as the phase reference and can be any arbitrary length

(other than approximately $n\lambda/2$) of unknown value. The line may be lossy. The reference impedance of the measurement system is determined by the line impedance. The ability of being able to use imperfect standards to obtain accurate calibration of the measurement system has had a major impact on advancing the state of the microwave art.

3.2.1 RTL Calibration Procedure

The "reflect-thru-line" (RTL) two port calibration method has an underlying theoretical form similar to TRL calibration. RTL differs from TRL calibration in one significant respect. RTL allows a full two-port calibration of a one-way measurement system of the type shown in Fig. 3.2, making possible measurement of nonlinear two-port devices with a single RF source. The RTL calibration procedure is a three step process, 1) reflection calibration to characterize measurement port 1; 2) transmission calibration to characterize measurement port 2 and the load characterization; and 3) system power calibration for accurate measurement of power levels at both measurement ports. The three main steps in the calibration process will be briefly discussed, followed by a practical discussion on creating calibration standards, fixturing, and procedures for reliable and accurate measurement.

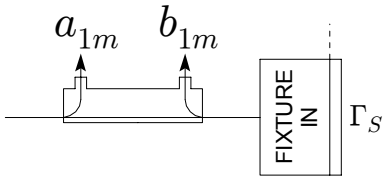
3.2.2 Reflect Calibration

The reflect standard Γ_S is placed at the reference plane and measured (b_{1m}/a_{1m}) as in Fig. 3.5a. The thru standard is placed between ports 1 and 2 and reflection measurements (b_{1m}/a_{1m}) are made for independent terminations Γ_i, $i = 1, 2, 3$ as in Fig. 3.5b. The line standard is placed between ports 1 and 2 and the terminations Γ_i are again measured (b_{1m}/a_{1m}) as in Fig. 3.5c. These measurements are sufficient to determine the unknown error coefficients e_{10}, e_{11}, e_{50}, and e_{51}. The remaining unknown term e_{40} will be determined during system power calibration, but for measurements where absolute power is not required e_{40} could be set to unity. A byproduct of the calibration measurements is that the unknown line standard and independent terminations are now known. The reflection calibration could be done by using OSL calibration but it is not used because of the difficulty in fabricating and characterizing the standards, though for a purely coaxial measurement system this may not be an issue.

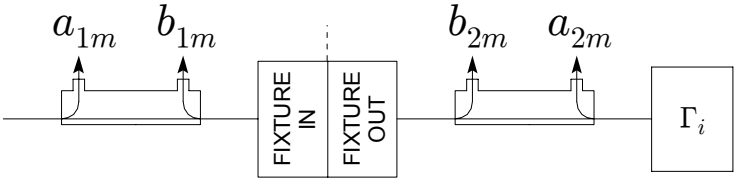
3.2.3 Transmission Calibration

With the thru standard connecting ports 1 and 2, as in Fig. 3.5b, transmission τ (b_{2m}/a_{1m}), and source ratio r (a_{2m}/b_{2m}) measurements are made for independent terminations Γ_i, $i = 1, 2, 3$. The transmission and source ratio measurements of Γ_i are repeated with the line standard as in Fig. 3.5c. Once the terminations are measured for both the thru and reflect standard the unknown error coefficients e_{62}, e_{63}, e_{72}, and e_{73} can be computed. The isolation error terms e_{60} and e_{70} can be measured directly with no

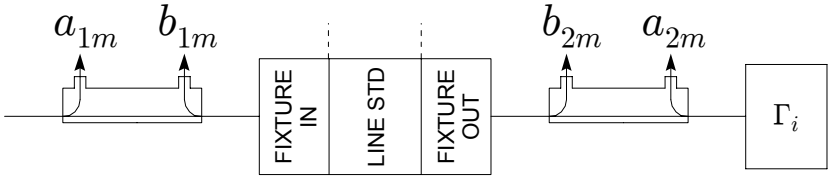
connection between ports 1 and 2 as shown in Fig. 3.5d.



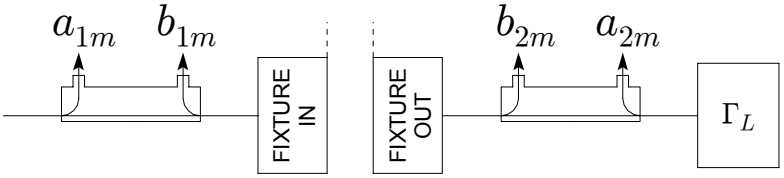
(a) Measure Reflect Standard



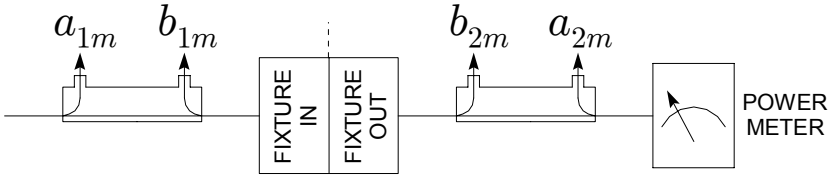
(b) Measure Thru Standard with three independent loads



(c) Measure Line Standard with three independent loads



(d) Measure isolation between measurement ports



(e) Power calibration for absolute power measurements

Figure 3.5. Measurement procedure for RTL calibration.

3.2.4 Power Calibration

For swept power measurements and applications that require an accurate power level measurement a power calibration can be performed to obtain a value for e_{40}. With a thru inserted at the DUT measurement plane a second reflect calibration is performed to obtain the measurement system loss and reflection at the power meter plane. The reflect calibration provides an accurate measurement of the reflection of the power meter sensor.

With these values known a power measurement can be made as shown in Fig. 3.5e, and a value for e_{40} can be obtained.

3.2.5 RTL Measurement Summary

RTL calibration requires a number of measurements to determine the error correction coefficients, but the mechanical make-and-break connections can be streamlined to simplify the measurement process. As presented in Fig. 3.5a-c the measurement sequence can be reduced to a handful of steps. The reflect standard Γ_S is placed at the reference plane and measured. Using the thru standard ports 1 and 2 are connected while reflection Γ (b_{1m}/a_{1m}), transmission τ (b_{2m}/a_{1m}), and source ratio r (a_{2m}/b_{2m}) are measured with terminations Γ_i on the output. Reflection, transmission, and source ratio measurements of Γ_i are repeated for the line standard.

3.3 Implementation of RTL Calibration

3.3.1 Realization of RTL Calibration Standards

The standards used in RTL calibration can be readily constructed using microstrip circuit techniques. The reflection standard must be known and determines the phase reference. In most measurement applications a "perfect" short can be readily manufactured such that $\Gamma_S = -1$. Any known, highly reflective termination, could be used as a reflect standard. An advantage of using a short is that it does not require expert craftsmanship to obtain a reflection coefficient of -1. Note that the reflect standard is required only for port 1 error correction and is measured only once during the calibration process. Fig. 3.6 illustrates one possible realization of a reflection standard. The reflect standard can be mechanically sandwiched between the break-apart fixture halves to ensure a reliable and consistent connection during measurement as shown in the cutaway view of Fig. 3.6. A practical matter to keep in mind when constructing the reflect standard is to ensure a consistent RF signal path at the interface between the transmission line as well as the ground plane. Careful construction techniques will ensure a consistent and reliable short for frequencies well above 2 GHz. Typically it is assumed that $\Gamma_S = -1$. For higher microwave frequencies where fixture topology or construction techniques may result in substantial parasitic effects the reflect standard could be measured prior to its use in the RTL calibration using standard TRL methods. Fig 3.7a is a photograph of a reflect standard constructed as shown in Fig. 3.6. Fig. 3.7b shows the reflect standard inserted at the measurement reference plane. The reflect standard in Fig. 3.7a is constructed from double sided copper 1.5 mm FR4 circuit board material. The spring tab is securely fastened to the microstrip transmission line using a short segment of transparent cellophane tape.

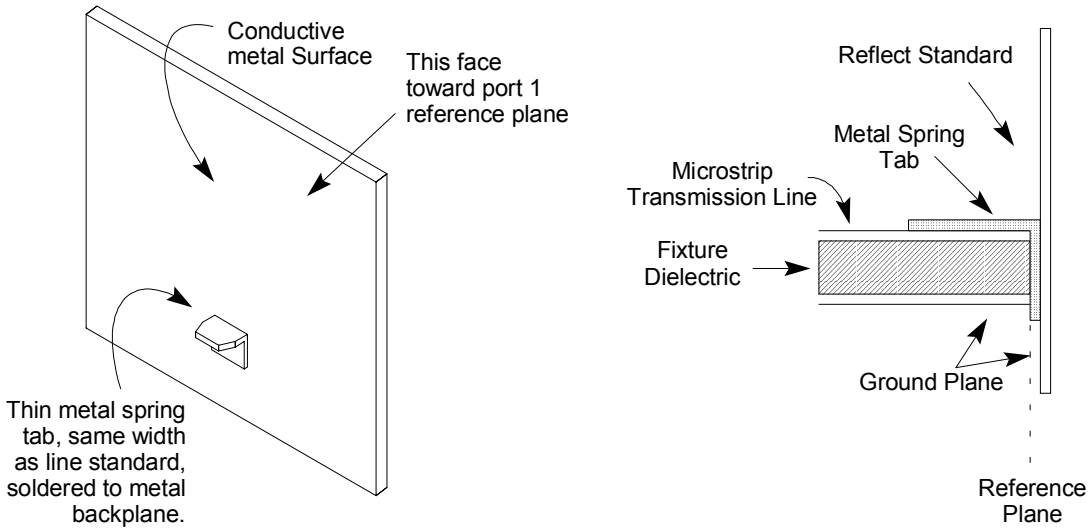
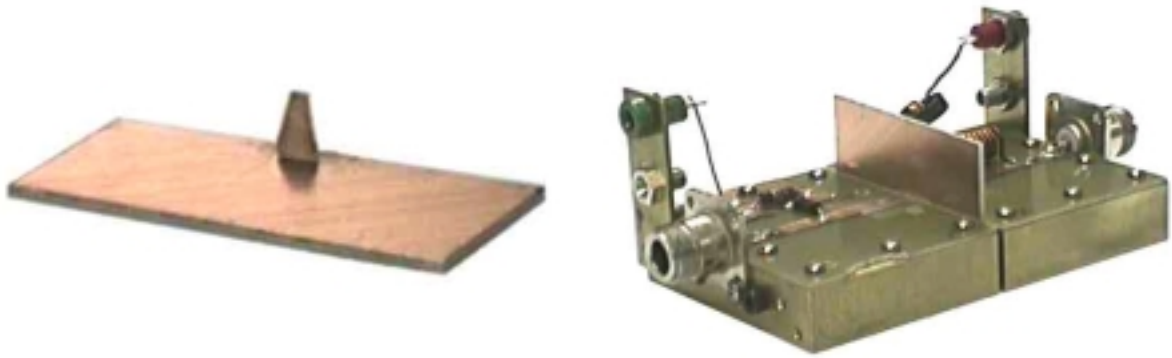


Figure 3.6. Reflect standard and cutaway showing fixture interface.



(a) (b)

Figure 3.7. Reflect standard (tab facing up) and measurement of standard with break-apart fixture. [Fixturing provided by Motorola, Inc.]

The thru standard is assumed to be an ideal thru, i.e. perfect transmission and zero reflection ($s_{11} = s_{22} = 0$ and $s_{21} = s_{12} = 1$ for the thru standard). The thru is realized by direct connection of port 1 to port 2. Because the thru is a "known", deliberate care must be exercised to ensure a clean connection between ports 1 and 2 at the reference plane. A contrast between RTL and TRL calibration is that for the most common form of TRL the thru standard determines the phase reference whereas for RTL the phase reference is determined by the reflect standard. This would suggest that when using break-apart fixtures RTL calibration would be more tolerant of parasitic effects associated with the thru measurement. Fig. 3.8 shows a top and cutaway view of the fixture thru connection. Fig. 3.9 is a photograph of the break-apart fixture with the thru standard in place ready for

measurement. To ensure a reliable connection between the fixture halves, a short strip of copper foil cut the same width as the transmission lines at the reference plane is used to bridge the discontinuity between fixture halves. The foil strip can be secured in place using cellophane tape. Care must be taken during fixture construction to ensure that when butted together the fixture halves have a close fit at the reference plane. To maintain a continuous ground connection at the reference plane a wide strip of copper foil was inserted between the circuit board and the metal base on the underside of the fixture. The foil above and below the circuit board spans the discontinuity guaranteeing a continuous signal path with minimal effect on the calibration measurement. The foil strip across the

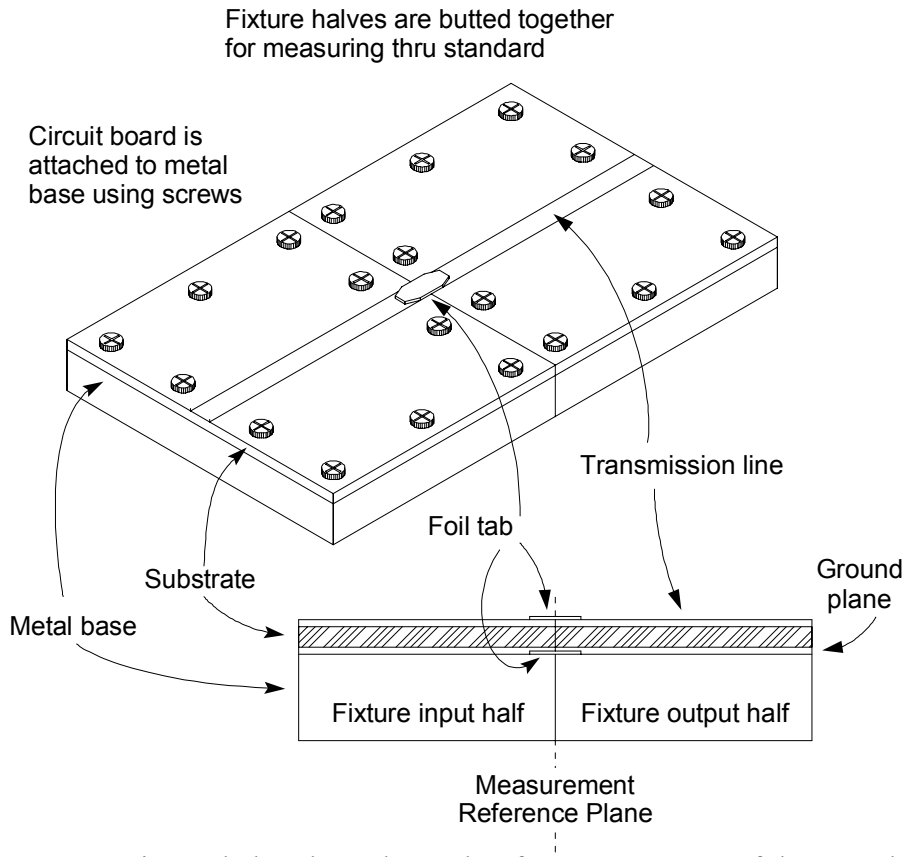


Figure 3.8. Fixture halves butted together for measurement of thru standard.

ground plane is securely held in place by the downward pressure of the screws holding the circuit board to the metal base. Alternatively the circuit board could be soldered to the metal base and the interface between the fixture halves could be precisely machined to provide a good ground and signal path connection at the reference plane.

The length of the line standard is not known *a priori* and may be of any length (other than $n\lambda/2$). For multiple half-wavelengths, the line appears similar to the thru and provides little useful information. The line standard sets the reference impedance at the measurement plane. As in TRL calibration it is permissible to have a lossy line standard. For practical reasons it is best to limit the electrical length of the line to within 20° to 160°

to minimize the effects of computational noise in the calibration solution. Modern computers usually have enough digits of precision that these length limits are rather conservative and it is possible to exceed the limits with 5° to 175° being near a maximum threshold, particularly with line loss. Inspection of the calibration data will clearly indicate the frequencies for which the line is not suited for calibration as the data will tend to spike towards positive or negative infinity at or near the critical $n\lambda/2$ data points. To calibrate across a broad range of frequencies, multiple line standards may be measured with the data interleaved in the calibration solution.

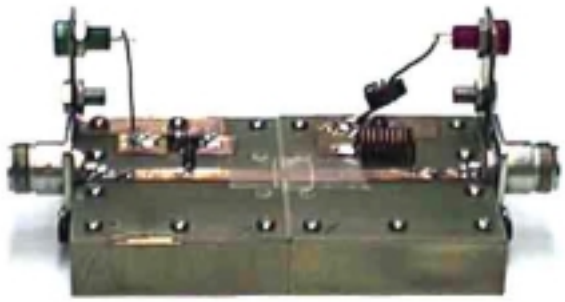


Figure 3.9. Break-apart fixture setup for measuring the thru standard.

Fig. 3.10 illustrates the line standard construction techniques used for the break-apart fixture. Thin metal tabs are attached to the transmission line at the reference planes to provide signal continuity at the junction between the line standard and the reference plane. If metal foil is used for the tabs, cellophane tape can be used to hold the foil next to the transmission line during the measurement. A stiffer metal could be used that has some spring tension. The tabs could be slightly bowed to provide the necessary contact. Mechanical assemblies constructed of Teflon or Nylon could also be used to apply consistent pressure without significantly affecting the electromagnetic field structure at the reference planes. To provide a solid ground plane connection between the fixture and the line standard a thin strip of metal is soldered to the ground plane at each end of the line standard. The extension tab slides beneath the fixture circuit board and is held securely in place by the downward pressure of the screws holding the circuit board to the metal base as was done for the thru standard. The impedance of the line standard determines the reference impedance at the measurement ports.

Overall measurements require accurate knowledge of the reference impedance. The line impedance may be computed if the dielectric constant ϵ_r of the substrate and dimensions are known. For some substrate material, such as fiberglass epoxy resin material, ϵ_r may vary substantially from manufacturer to manufacturer and batch to batch. The substrate material shown in Fig. 3.11a is FR4 fiberglass epoxy resin material. The impedance of the line standard was measured directly using a time domain reflectometer (TDR). A TDR measurement was also used to obtain an estimate of ϵ_r from which the line impedance could be determined, verifying the line impedance measurement. For line

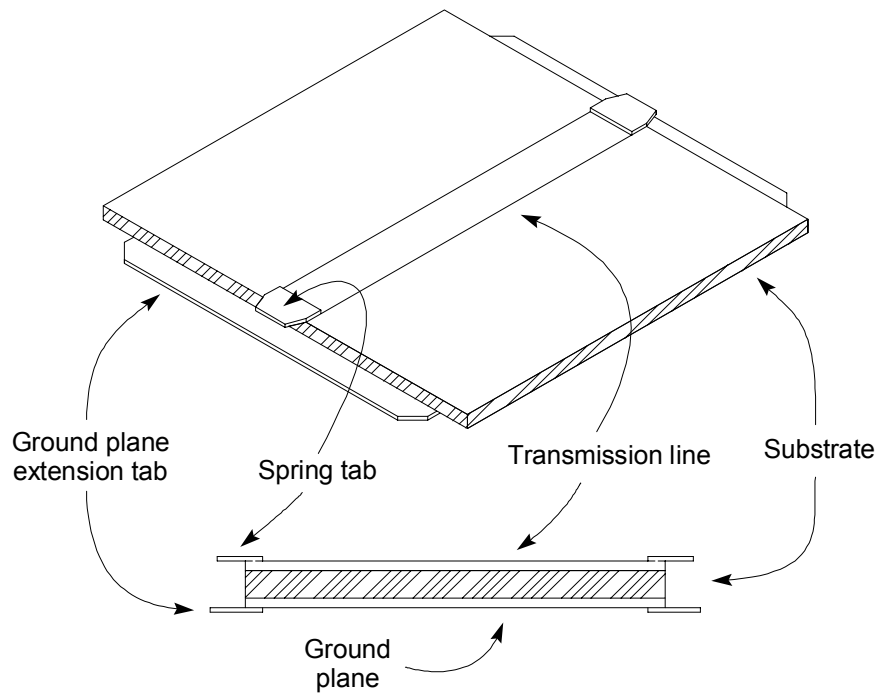


Figure 3.10. Detail of line standard. The line standard sets the measurement reference impedance. The width of the line standards is typically chosen to be the same width as the device package lead width.

impedances with $Z_0 < 20\,\Omega$ it is best to use a high quality TDR for improved accuracy. Inexpensive TDR meters are best suited for measurement of 50 Ω or 75 Ω transmission line and their absolute resolution decreases as the magnitude of the reflection increases. A laboratory grade TDR is capable of good measurement accuracy for line impedances down to approximately 10 Ω. Another measurement option that has shown good measurement accuracy is the TDR option on the HP8510 series network analyzer, though it uses a frequency to time domain transformation that is not as straightforward as the dedicated TDR meter. Another option for measuring the line impedance is to solder a small chip resistor with a resistance value close to the expected line impedance value on one end of the line and a small SMA connector on the other end of the line and run a frequency sweep across a broad range of frequencies looking for a null in the reflected signal. With the null frequency, length of the line, and resistance value, an accurate estimate of the actual line impedance can be determined. For frequencies below 1 GHz this frequency sweep method gives excellent results even if parasitic effects are neglected.

A complete break-apart fixture and set of calibration standards used for device measurements are shown in Fig. 3.12. A fixture was constructed for each package lead

(a)

(b)

Figure 3.11. Line standard and measurement of standard with break-apart fixture.

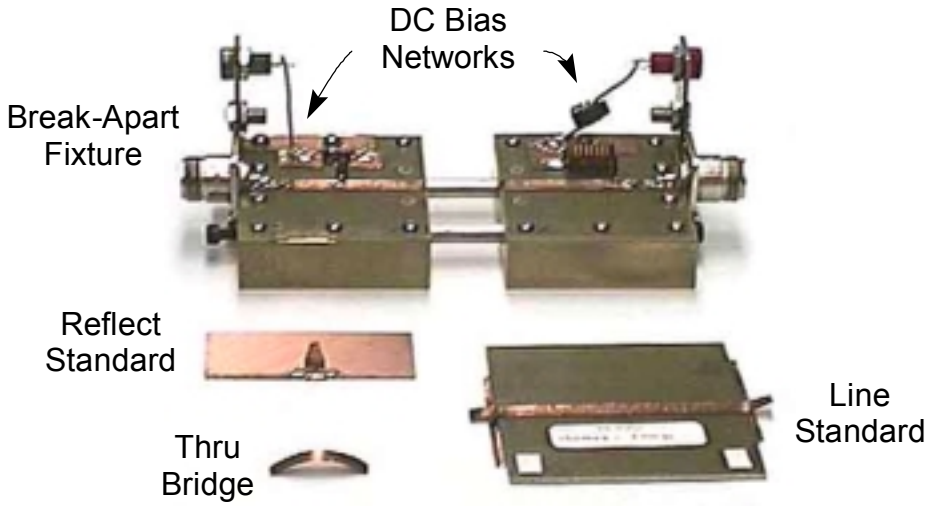


Figure 3.12. Break-apart fixture and calibration standards for RTL calibration and measurement.

width style along with multiple line standards to span the frequency range of 10 MHz to 1 GHz. The fixture includes onboard dc bias networks for the gate and drain supply as well as dc blocking capacitors on the signal lines. Onboard bias and decoupling networks provide flexibility for maintaining device stability across a broad range of frequencies and should be characteristic of typical bias networks to be used in the field. Bias networks can be easily designed to handle large operating voltages and currents, whereas output bias tees (as is more common in small-signal device measurement) require expensive, specially constructed high-power bias tees. Since the fixture will be included in the error correction there is no need to build a "perfect" fixture across the frequency range of measurement, as long as reasonable operation is available. Fixture effects will be calibrated out of the measurement.

After calibration of the system a metal insert the same width as the DUT is clamped between the fixture halves. The DUT is mounted to the insert using bolts through the device heat sink flange or using a mechanical clamping arrangement. The method of mounting is selected to provide a good thermal connection between the device and the insert and a good RF connection between the package leads and the transmission lines on the fixture. As with the thru and line standard measurements, care must be taken to ensure a good ground connection between the insert and the fixture halves. Careful machining of the fixture halves and insert will provide for a good ground connection. One caution about using break-apart fixtures with high power devices is that often the thermal heat sinking capacity of the insert is limited. It is difficult to get a good thermal connection and a good ground connection simultaneously with a break-apart fixture. The mechanical interface between the insert and fixture halves should be cleaned with a fine metal polish periodically to avoid oxidation and dirt from interfering with the quality of the connection.

Mechanical assembly of the break-apart fixture requires some practice to obtain consistent and accurate connection of the calibration standards. It is recommended that a TDR be used to check the fixture and calibration standard connections before and after a measurement. Parasitics due to imperfect connections can be minimized by using the TDR to observe the connection. The TDR will detect poor quality connections or mechanical problems that might be difficult or impossible to detect by looking at the measured data alone. The TDR can give immediate feedback if a connection changes before or during a measurement. For observing the mechanical connection of the fixture it is useful to have a true TDR rather than using a frequency domain technique such as the HP8510. A true TDR injects a dc component that can locate problems that might not be seen using a frequency domain method. A high resolution TDR allows problems to be pinpointed to an actual location on the fixture or in the standard, a feature that is most useful when corrective action must be taken.

3.3.2 Independent Terminations $\Gamma_1, \Gamma_2, \Gamma_3$

The RTL calibration procedure requires three independent terminations be presented at the measurement reference plane during the calibration measurement procedure. The

terminations Γ_i, $i = 1, 2, 3$ are shown collectively in Fig. 3.5. The Γ_i are connected to the measurement system output and it is convenient to use coaxial terminations such as open, short, and load standards. Depending upon the reference impedance for the line standard and phase shift between the Γ_i load plane and the reference plane, the resultant loads may not necessarily be an "open", "short", or "load", however the requirement that the loads be independent should still be satisfied. The load terminations should also be repeatable for the complete set of calibration measurements. It is acceptable to use stub tuners to obtain arbitrary loads, but the value of the load termination must remain consistent for all the calibration measurements. For the case where tuners are used to obtain the terminations this might require separate tuners, locked into preset positions, for each termination, or a precision tuner where the load positions can be accurately and repeatably set during the course of calibration.

The load terminations Γ_i afford a considerable degree of flexibility in the calibration process. Load terminations can be carefully selected for special measurement applications to ensure a good calibration. For example, measuring high power RF devices with low input and output impedances a measurement system may have non-ideal transformation networks on ports 1 and 2 that transform 50 Ω to 5 Ω over the frequency range of interest. Depending on the transformation network, a short, 50 Ω load, and an open standard may not be the optimal set of independent loads. A better choice of load terminations might be a short and two offset shorts such that $\Gamma_1 = 1\angle 180°$ (short), $\Gamma_2 = 1\angle 60°$, and $\Gamma_3 = 1\angle -60°$. An advantage to using offset shorts is that they can be readily realized using a coaxial short termination and two additional lengths of transmission line. The calibration process will compute the actual load termination values as seen at the reference plane. A check of the values of measured load terminations Γ_i will provide an indication of the degree to which they are independent. Care should be taken when calibrating across a broad frequency band to ensure that the Γ_i are independent for all frequencies of interest.

3.3.3 Measuring Port to Port Isolation

The RTL calibration procedure is a robust calibration method capable of accounting for imperfect isolation between measurement ports 1 and 2. It may be possible for the reference signal b_s from the generator to leak into the output measurement path. In a well designed measurement system, the leakage between the signal generator and the output receiver would be minimal. For a network analyzer such as the HP8510 the isolation term is dominated by cross talk in the dual receiver architecture. For some measurement setups a leakage path could exist across the measurement fixturing, which should be minimized as much a possible. System calibration accounts for any leakage in the receiver or fixturing by measuring the isolation between the reference port a_{1m} and the output ports b_{2m} and a_{2m}. The isolation is measured with the DUT removed from the fixture and no connection path inserted in the fixture.

3.3.4 Drive Calibration

For large signal measurements, it is required that the measurement system be capable of measuring the absolute power levels at the measurement reference plane. Power calibration determines the absolute power wave measured at the input a_{1m}. A power meter is used to provide a reliable, absolute power level measurement. The power calibration process first characterizes the ratio of power reflected off the power sensor head to the amount of power delivered to the power sensor. Most power meters are set to a reference impedance of $50\ \Omega$. A $50\ \Omega$ power sensor will provide a non zero reflection coefficient at the measurement reference planes in a non-$50\ \Omega$ measurement environment (as will typically be the case when measuring high power RF devices). A power measurement is made, and with the known reflection of the sensor head and frequency response of the power meter an accurate value for the input power at the reference plane can be determined. Once the system is power calibrated it is possible to know the incident and reflected power wave levels (and hence the incident and reflected power levels) at the measurement reference planes. The capability of knowing input and output power for a DUT is useful in characterizing the performance of a device across a range of input drive levels. A power calibration would need to be performed for each measurement frequency. The ability to measure across a wide range of input power levels to accomplish both small and large signal device measurements requires that some care must be taken when selecting measurement equipment. Equipment should be isolated using attenuators and couplers to ensure that safe operating ranges are not exceeded. For measurements across a wide dynamic range a check of the system linearity is recommended to ensure consistent and reliable measurements across power level. System linearity will be discussed in more detail in the following section.

This section has presented an overview of the RTL calibration procedure and methods that can be used to obtain reliable and accurate results. A detailed description and derivation of the RTL calibration and measurement procedure is presented in Appendix A.

3.4 Measurement System Linearity

When making S-parameter measurements it is desirable to have a linear measurement system. A linear measurement system could be defined as a measurement system that is capable of accurate and repeatable measurements across frequency and power levels of interest. When making small-signal measurements it is usually not critical to know the absolute input power level at the measurement reference plane. Most modern vector network analyzers provide the capability of specifying the output power, but the output power level is typically referenced internally to the signal source and often does not account for losses in the test setup or measurement system. If, for example, a measurement setup is to have a power level of 0 dBm, most likely the actual power level at the reference plane will be -6 dBm due to losses through the S-parameter test set. As the network analyzer sweeps across a broad range of frequencies the power level could

vary due to the frequency response of the test setup. Input power variations of ± 1 dB would not be unusual across 500 MHz of bandwidth. This variation does not affect the accuracy of the small-signal S-parameter measurements because of the error correction. For active devices measured in the small-signal region of operation these variations would not significantly alter device performance since they are being held to a linear region of operation and performance should not be a function of the input drive level. Basic power calibration was addressed in the previous section, but nonlinearities of the measurement system are still of concern.

A device subject to a large signal input level could exhibit different performance for a change in input drive level. For this reason, when measuring device performance as a function of input drive level it is desirable to be able to quantify the input (and output) power level linearity. The receiver architecture used to measure amplitude and phase of the input signals should give reliable results across the range of signals that are to be measured. Many modern network analyzers have internal nonlinear checking and correction. However, linearity checks are still a good practice. For the RTL measurement system described, the linearity of the vector voltmeter used to make measurements was measured across the range of power levels used for measurement of the DUT. The system was connected for a simple thru measurement similar to that shown in Fig. 3.13 with a non-reflective load Γ_L on the output. The measured values of a_{1m} and b_{1m} were then

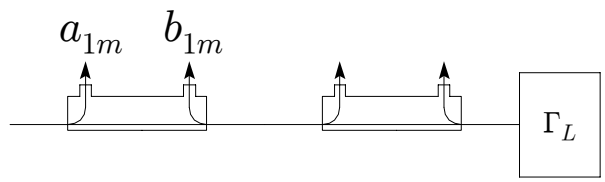


Figure 3.13. System setup to measure linearity as a function of power.

examined to determine if the system was linear across the range of input drive. The system would not be linear if the receiver were saturated by either of the input signals or the system losses were changing with drive level (excessive heating in the signal path attenuators). A typical system setup as shown in Fig. 3.4 used a signal generator with a voltage controlled output level. Fig. 3.14 shows the measured values at 450 MHz for channels A and B as a function of the signal generator drive voltage. As the signal generator control voltage is decreased it can be seen from the plot of measured voltage amplitude that the signal generator gain control is not a strict linear function of the control voltage. Since the system measures the input power level for each measurement this nonlinearity is not a concern since it will not limit the accuracy of the measurements. The ratio of the magnitudes gives an average value for the load reflection of $\Gamma_L = -30.80$ dB. An ideal system would give the same result for Γ_L regardless of drive level. The measured values show a discontinuity at 9.54 V, 7.57 V, and 7.18 V respectively. At 9.54 V and 7.57 V the receiver internally switched ranges and there is a 0.1 dB difference between the internal receiver calibration for these ranges. From 7.57 V to 7.18 V it is not clear what is causing the deviation in magnitude. Note that as the

power levels increase above 7.18 V the measured values tend to settle around a consistent −30.80 dB level. The observed discontinuities show that the system is not truly linear, though the discrepancy between the different ranges is small enough to be neglected. The nonlinearity associated with internally switched modes of the receiver is potentially the most likely form of measurement system nonlinearity. Another observation worth noting is the deviation of the magnitude measurement about the average. Just above 7.57 V the measurement noise increases considerably. This increase is due most likely to quantization error and the number of significant bits being used by the measurement system. As the input power level increases, fewer bits are used in the analog to digital converters to

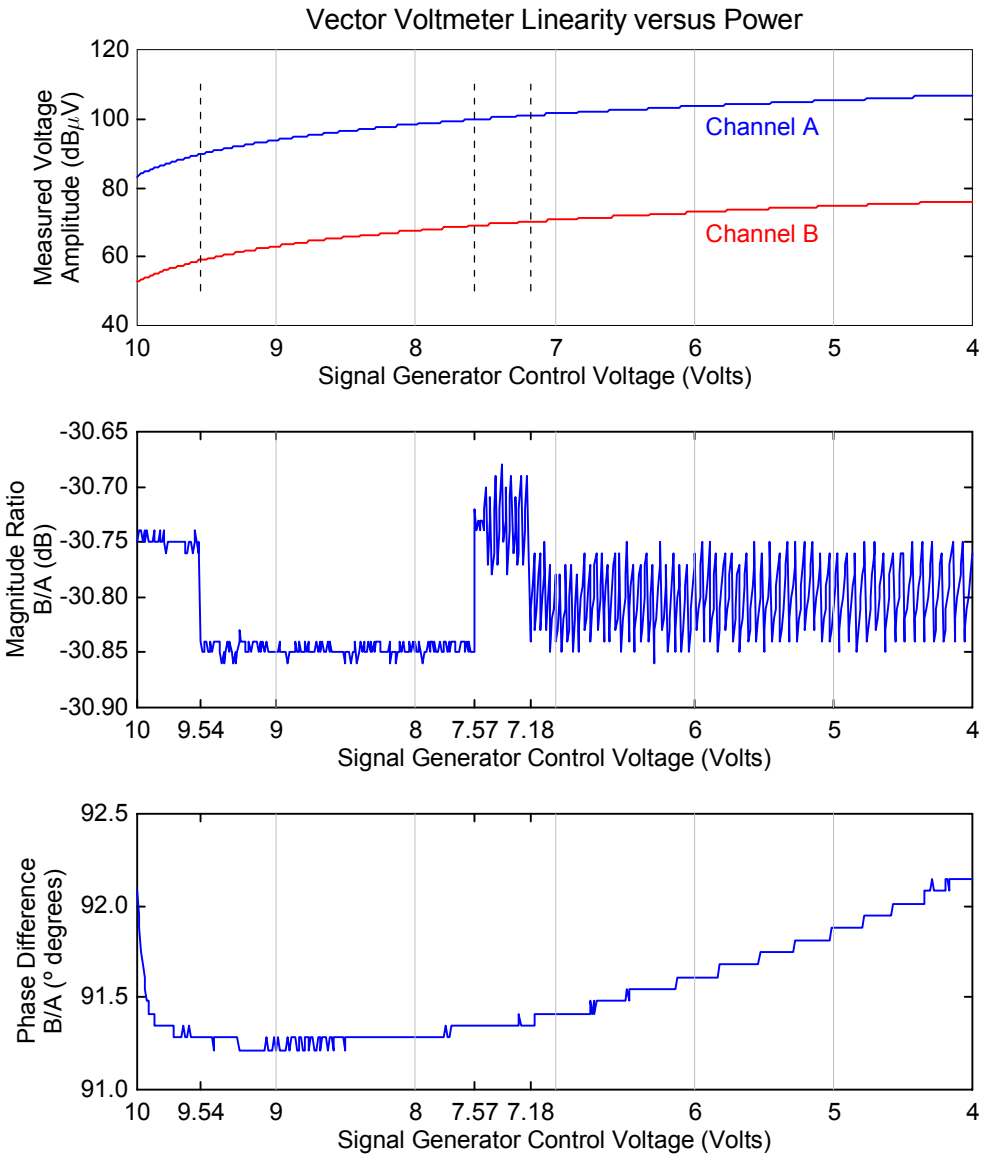


Figure 3.14. Measuring system linearity as a function of power level.

obtain a higher dynamic range for input power levels. The result is that quantization noise increases for larger signal levels. The phase difference as a function of control voltage shows less than $\pm 0.5°$ of total phase error across the power range of interest. These measurements show that while the measurements do change as power level is increased the changes are well within a reasonable error limit for the measurement system. The discontinuities in magnitude are on the same order as the measurement system noise for higher input power levels, i.e. the measurement system nonlinearities are on the same order as the measurement system noise.

If the measurement system were not linear across input drive level due to limitations of the receiver it would still be possible to obtain reliable measurements if multiple calibration files were used. Regions of discontinuity could each have their own set of error coefficients and these coefficients would be used according to the drive levels being measured. Such an approach complicates the system calibration and would require additional software overhead, but would be an effective way to extend the measurement capability of a non-ideal receiver.

A check was performed to quantify the phase linearity of the measurement system across frequency. Fig. 3.15 shows the system setup for measuring phase linearity across frequency. A short segment of coaxial transmission line is inserted between the directional couplers and the input signal a_{1m} and b_{2m} are measured as a function of frequency. Without using error correction the signal magnitudes will change across frequency. The magnitude changes will be compensated for using error correction. The receiver should properly track the phase between b_{2m} and a_{1m}. The measured phase is plotted in Fig. 3.16 along with the ideal linear response. Across the entire measured range there are slight variations in the measured phase, though these appear to be negligible. The phase linearity is close to being ideal. There are no noticeable phase discontinuities as the receiver switches from one frequency band to another across the range of measured frequencies.

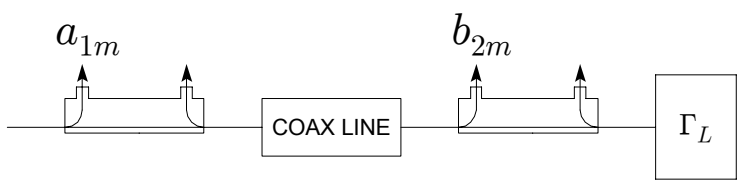


Figure 3.15. System setup to measure phase linearity across frequency.

Error correction would be able to compensate for non ideal phase response across frequency. Poor phase linearity could, however, be an indication that the measuring system might be prone to exhibit phase problems at varying power levels. Checking the system linearity can aid in diagnosing or preventing systemic problems that could impact overall measurement accuracy. The nonlinear measurement environment can subject the measurement system to a wide range of signal frequencies and power levels. To minimize the reactions between the harsh measurement environment and the measurement equipment it is desirable to operate comfortably within the linear operating range of the signal receiver.

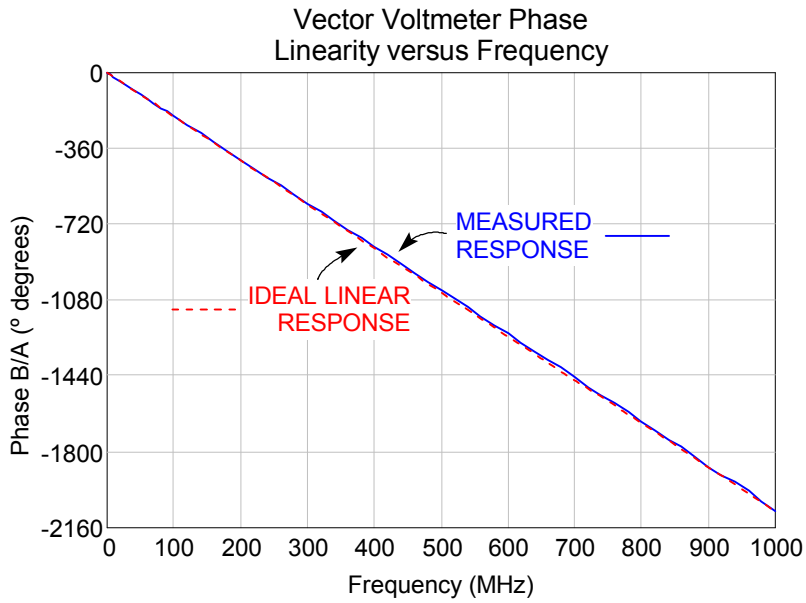


Figure 3.16. Measuring system linearity as a function of frequency.

3.5 Small-Signal Measurements

With a calibrated measurement system, the input reflection Γ, transmission τ, and load Γ_L can be measured at the reference plane as a function of input power and/or frequency. Two-port active device measurement has traditionally been constrained to small-signal measurements. For small-signal measurements the input power is fixed at a low level to ensure that the device is in a "linear" region of operation. Measurements are made for constant input power across a range of frequencies. For the RTL measurement system the device can be measured with two independent loads Γ_{L1} and Γ_{L2}. The reflection Γ, transmission τ, and load values Γ_{L1} and Γ_{L2} are measured. Using the measured values a system of four equations can be written

$$\begin{aligned} s_{21} + s_{22}\tau_x\Gamma_{Lx} &= \tau_x \\ s_{11} + s_{12}\tau_x\Gamma_{Lx} &= \Gamma_x \end{aligned} \tag{3.1a}$$

where $x = 1, 2$ for both loads. Solving the system of four equations for the four unknown S-parameters results in the following set of equations.

$$
\begin{aligned}
s_{11} &= \frac{\tau_1\Gamma_{L1}\Gamma_2 - \tau_2\Gamma_{L2}\Gamma_1}{\tau_1\Gamma_{L1} - \tau_2\Gamma_{L2}} \\
s_{21} &= \frac{\tau_1\tau_2(\Gamma_{L1} - \Gamma_{L2})}{\tau_1\Gamma_{L1} - \tau_2\Gamma_{L2}} \\
s_{12} &= \frac{\Gamma_1 - \Gamma_2}{\tau_1\Gamma_{L1} - \tau_2\Gamma_{L2}} \\
s_{22} &= \frac{\tau_1 - \tau_2}{\tau_1\Gamma_{L1} - \tau_2\Gamma_{L2}} \ .
\end{aligned}
\tag{3.1b}
$$

For a linear device, the change in load will not affect the two-port S-parameters as the device performance is not a function of the load.

Measurement of the S-parameters using this method allows excellent isolation between the measurement receivers and the active device. Often times a device that is potentially unstable could begin to oscillate as the input signal is swept in frequency. The RTL measurement system isolates the sensitive measurement equipment from any large signal bursts. The independent loads can be chosen such that the device is stable during the measurement process.

Motorola provided eighteen different TMOS device types that were measured at two or three different bias points at frequencies from 10 MHz up to 1 GHz. Three to five individual devices were measured from each device type. Motorola had requested the measurements be made using a high dc bias current in a class A mode of operation with a small-signal input level. Appendix B lists the devices and bias conditions for which S-parameter data were measured and published in the product data sheet.

A graph of the S-parameters for the MRF275L is shown in Fig. 3.17. The MRF275L, which is rated by the manufacturer as a 100 Watt device, was measured with a bias setting of $V_{DD} = 28\,\text{V}$, $I_{DQ} = 4.5$ A. The bias current was selected artificially high so that the effects of device self-heating could be captured in the measurements. Under such high bias currents, if a device were to begin oscillation it could potentially emit a destructive burst of power capable of exceeding the power limitations of most laboratory equipment. Even though the bias supplies were set to current limit, the quiescent currents were large enough that damage could easily be done to the test equipment without triggering the drain supply dc current limit. If the transistor were to go into oscillation, it could potentially deliver a maximum of 126 Watts at either the input or the output of the fixture. If during oscillation the transistor were to deliver even $1/100$ of the total input power to the measurement system, the power levels would be large enough to damage most laboratory test equipment. The attenuation in the measurement path needs to take the worst case power delivery into account. The RTL system proved to be a safe and reliable method for measuring these devices. Load terminations Γ_i, $i = 1, 2$, were a coaxial 50 Ω power termination and a coaxial short which provided independent Γ_{Li} at the measurement plane. These terminations were selected because they were readily available, could safely handle the power levels involved, and provided stable device operation in the measurement fixture. Some of the devices were unstable in the 10 to 100 MHz region

when an open termination was used. It is a good idea to test the device for oscillation in the fixture prior to connecting the fixture to the test system. The device can be tested for varying power levels, frequencies, and loading conditions to access the likelihood of oscillation. This is still not a guarantee that the device will not oscillate under test. Swept signal sources inject a broad spectrum of frequencies into the test system at the beginning and end of the sweep cycle. The step function at the end of the sweep cycle could stimulate a potentially unstable device to begin oscillation.

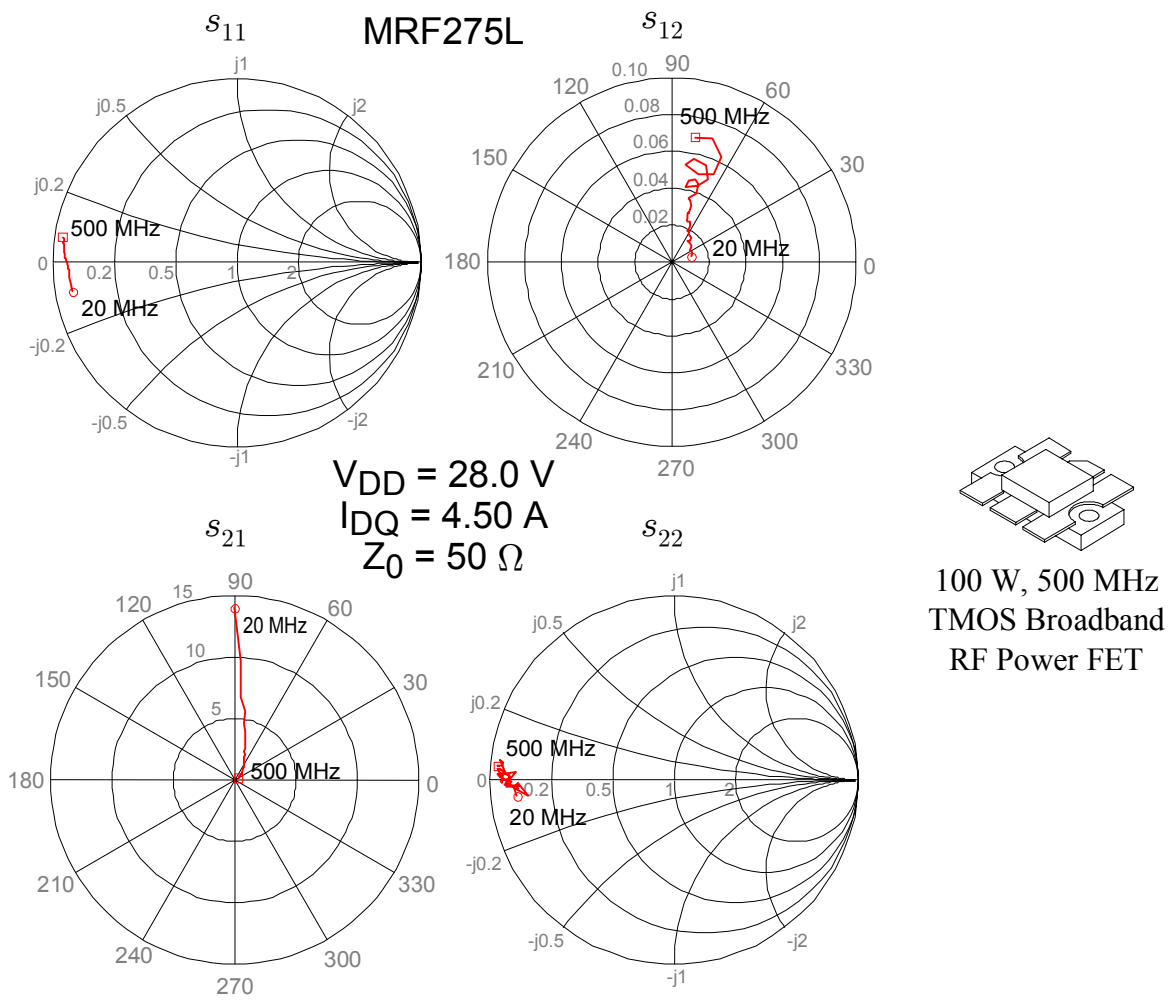


Figure 3.17. MRF275L S-parameters for high bias current class A operation.

Parameters associated with the second port, i.e. s_{12}, and s_{22}, may be subject to slight inaccuracies due to the dynamic range of the measurement receivers and numerical noise in the computation of the S-parameters using (3.1b). From s_{11} in Fig. 3.17 the MRF275L has a resonance on the input at 170 MHz where the input reflection is purely real. Using the data sheet value for $C_{iss} = 135\,\text{pF}$ as a rough estimate of device input capacitance and computing the inductance necessary to resonate out the input capacitance we arrive at a

parasitic inductance value $L \approx 6.5\,\text{nH}$ on the input. This inductance is consistent with the values that could be expected for package parasitics and gate bond wires. The value for s_{12} in Fig. 3.17 has some curlicues in the data which may be due to package parasitics or device parasitic resonances in the reverse path. The s_{21} value for the MRF275L decreases with frequency. The S-parameters can be used to compute $|h_{fe}(\omega)|$ to determine an f_T, the gain-bandwidth frequency. The measured data ends at 500 MHz, so an extrapolation of the data is necessary. Assuming that there is not an additional corner frequency between 500 MHz and 5000 MHz the extrapolated $f_T \approx 3000$ MHz. The initial corner frequency, f_β is lower than 20 MHz and is within the measured frequency range. The gain of the device rolls off at 6 dB/octave across the entire range of measurement from 20 MHz to 500 MHz. The measured s_{22} is noisy, but the data is sufficient to be useful for design. Observing the impedance at which s_{22} crosses the real axis in Fig. 3.17 we note that the conjugate of s_{22} would give a load resistance of $R_L = 2.9\,\Omega$. From (2.38) we can compute an estimate of the power that can be delivered to a load R_L as $P_L = 134\,\text{Watts}$, which is consistent with the values given in the manufacturer's data sheet for operation at approximately $f = 225\,\text{MHz}$. Using the data sheet values for $C_{oss} = 140\,\text{pF}$ as an approximation to the output capacitance and the frequency for which s_{22} is purely real, (taken at 230 MHz, the data is a bit noisy here so this is an approximate frequency point), the parasitic inductance is $L \approx 3.4\,\text{nH}$, which would be a practical value for the output parasitic inductance. S-parameters provide a wealth of device information, both from a device characterization and circuit design perspective. The S-parameters in Fig. 3.17 could be used directly to design an amplifier with the MRF275L for the given bias conditions.

The unique one-way measurement capability of the RTL measurement system allows meaningful S-parameter measurements on devices operating in class B or class AB mode of operation. For small-signal measurement of devices, the class B bias should be adjusted for a linear mode of operation as discussed in section 2.6 (often times this would be considered a class AB bias condition). The RTL measurement system is capable of measuring s_{12} and s_{22} while port 1 is being driven so that the S-parameters are a snapshot of true class B device performance. A push-pull transistor may be intended for class AB mode of operation and often contains two transistors in a single package. Push-pull configuration requires that the two devices have near identical performance. By manufacturing a device with two transistors in a single package, the manufacturer can ensure that the two die are picked off the wafer from adjacent positions and will have near identical characteristics. S-parameters for a push-pull device can be obtained by measuring one of the two transistors in the package. The other device in the package can be safely terminated and need not be biased on during measurement. The S-parameters for the individual transistor can be used in a simulation to design the matching networks for the push-pull device.

The S-parameters for the MRF275G are shown in Fig. 3.18. The MRF275G uses the same silicon die as the MRF275L in Fig. 3.17 with two devices placed side by side in the same package. One of the two transistors in the package was measured, the other was left unbiased with the drain and gate grounded. On the input side, s_{11} has a resonance at

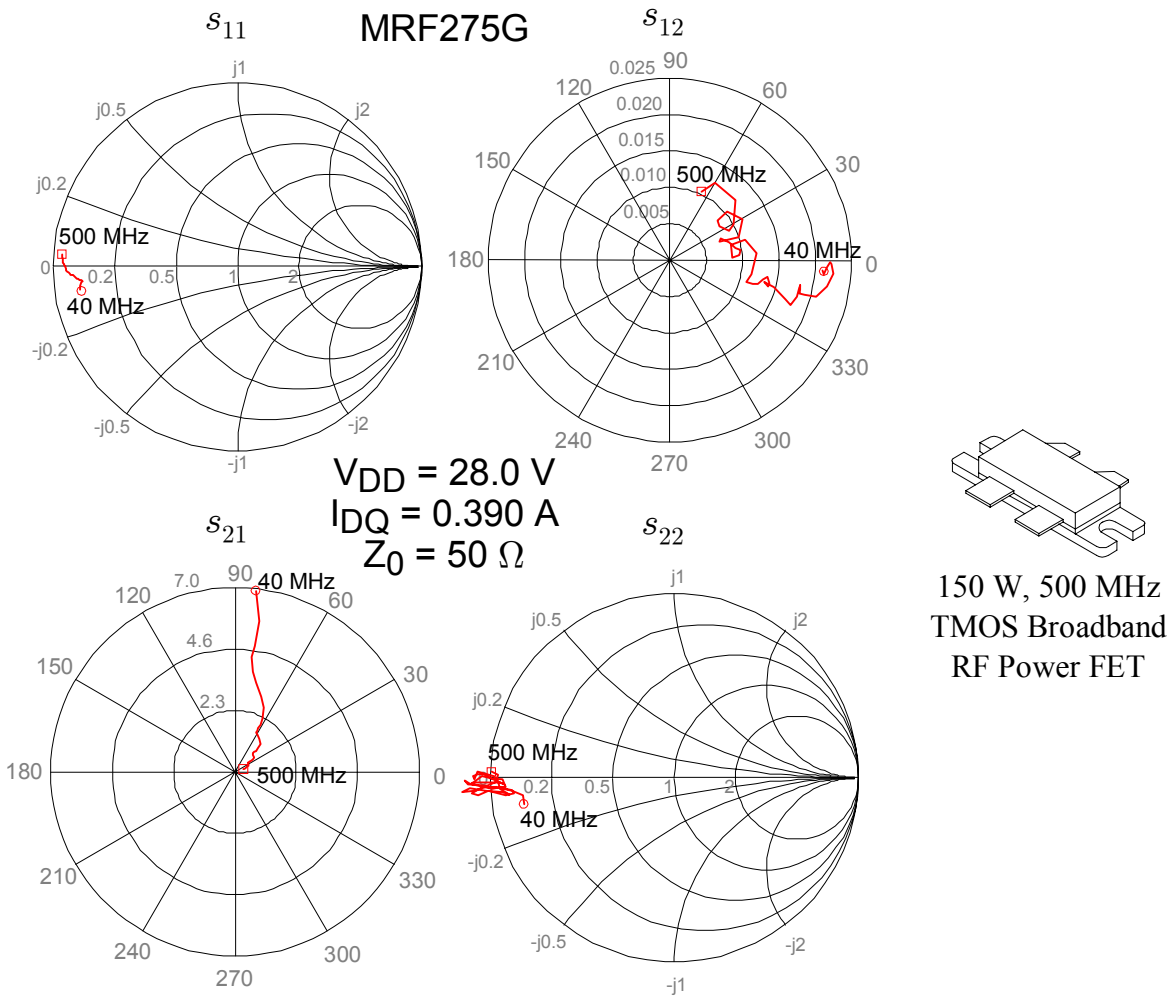


Figure 3.18. MRF275G S-parameters for class AB operation (one side of push-pull device).

330 MHz. Using $C_{iss} = 135$ pF from the device data sheet as an approximation to the input capacitance, the parasitic inductance is computed to be $L = 1.7$ nH, an indication that this device has lower input inductance than the MRF275L, possibly due to shorter bond wires on the gate. The device has a computed maximum stable gain of $G_{\text{MSG}} = 18$ dB at 250 MHz with $|s_{21}|$ rolling off at 6 dB/octave across the entire measurement range. The values for s_{12} are an order of magnitude smaller than the values for the MRF275L. It would be reasonable to conclude that the smaller s_{12} should help stabilize the device and oscillation is not the cause of $|s_{22}| > 1$ for several frequencies. The large s_{22} value is due to numerical noise in solving (3.1b) for s_{22} as the values of $|s_{22}|$ plotted in Fig. 3.18 are on the order of 1.05 and are not significantly greater than 1. This device exhibits a higher input and output reflection coefficient than the MRF275L due in large part to the class B mode of operation.

Researchers have looked at using four-port S-parameters to describe a push-pull device (Chandra, 1998). The detail of a four-port description is of little benefit for most high-power transistors and often a representative two-port description is extracted from the four-port measurements. The main benefit of a four-port description is to capture the cross coupling between the two transistors in the single package. At lower frequencies this coupling is of minimal concern. The S-parameters for seven commercially available VHF and UHF push-pull devices have been published. Appendix B lists the RF power transistors measured using the RTL system for which data has been published.

The RTL measurement system has been used to make two-port small-signal S-parameter measurements on high-power RF transistors. The unique measurement architecture provides a safe, yet effective means, of obtaining two-port scattering data on potentially-unstable devices that are otherwise difficult (or dangerous) to measure using conventional microwave measurement systems.

3.6 Large-Signal Measurements

To quantify the nonlinear behavior of a DUT it may be desirable to drive the device with a large-signal input, and vary the load termination Γ_L. The measurement technique where the device load impedance is varied as the performance at the DUT is measured is known as load-pull[13] (Cusack, 1974; Mazumder, 1978; Poulin, 1980; Sechi, 1980; Pierpont, 1986; Raab, 2002). The RTL measurement system is well suited for measurement of nonlinear devices because it provides a means for varying the DUT drive level and load termination.

A traditional method of determining the nonlinear performance of a DUT is to vary the load terminations and drive level, measuring the DUT performance at each test condition. The resulting DUT performance parameters are mapped on a Smith chart as a series of contours. These source- and load-pull measurements are useful for determining the DUT operating characteristics and design parameters for a given drive level and termination. The disadvantage to load-pull is that it requires a large number of load terminations to provide sufficient detail in the Γ_L plane. The RTL measurement system provides Γ and τ as functions of input power drive level a_1 for any arbitrary load Γ_L. The ability to measure Γ_L *in situ* is an improvement over existing load-pull methods. Traditional methods either involve highly-repeatable mechanical tuners that are characterized prior to making measurements or tuning the test setup for best performance and then breaking the system apart and measuring Γ_L at the DUT plane with the DUT removed. Both methods require a separate measurement setup to quantify Γ_L which significantly adds to the time and effort required to properly characterize the load terminations prior to plotting the contours. The RTL measurement system allows real time *in situ* measurement of Γ_L and

[13]When the measurement technique was first practiced, varying the load impedance literally required pulling on stubs of a trombone tuner, since then the terminology "pulling" the load impedance has been in common use. The source impedance can also be varied while device operation is being measured and is termed source-pull.

provides immediate updates of the load state and the device performance. The system can use inexpensive stub tuners instead of expensive highly-repeatable tuners. Load-pull measurements were made with the RTL system using a $50\,\Omega$ power termination, sliding short in a shunt configuration, and a trombone line extender. This load termination setup gave good coverage of the Smith chart and provided a simple means of systematically changing Γ_L. The measurement system was fully calibrated. Measurements were made while adjusting the input power to maintain a constant P_{in} and varying the load termination for each point. Power incident on the device is given by

$$P_{\text{incident}} = |a_1|^2 \ . \tag{3.2}$$

The input level a_1 is the forward traveling input rms power wave at the measurement reference plane. Power output by the device is

$$P_{\text{out}} = |a_1 \tau|^2 \tag{3.3}$$

at the drive frequency. The transmission τ includes the effects of load mismatch. From the measured output power and the measured load reflection, the power delivered to the load is computed

$$P_L = |a_1 \tau|^2 \left(1 - |\Gamma_L|^2\right) \ . \tag{3.4}$$

The incident power and the measured reflection provide the power input to the DUT

$$P_{\text{in}} = |a_1|^2 \left(1 - |\Gamma|^2\right) \ . \tag{3.5}$$

From these measurements the transducer and power gain can be computed

$$\begin{aligned} G_T &= P_L / P_{\text{incident}} \\ G_P &= P_L / P_{\text{in}} \ . \end{aligned} \tag{3.6}$$

These measurements are sufficient to provide load-pull contours as shown in Fig. 3.19.

The G_P contours were computed from the data points using (3.6) and a contouring algorithm (Falconer, 1971). Gain contours $G_P \geq 11$ dB close to the edge of the Smith chart are not closed because the device is unstable for low impedance load terminations. The unstable region is roughly indicated by a dotted curve in Fig. 3.19. Device self-heating effects, evident for several data points as variations in the contours, provide a truer indication of potential performance under actual operating conditions. Coupled with the known input Γ, the load-pull data provides a good starting point for design and allows optimal source and load terminations to be empirically determined. As with large-signal S-parameters, load-pull source and load terminations are power-dependent. The ability to make measurements for multiple loads, rather than a single fixed termination, does provide a more complete picture of the DUT performance.

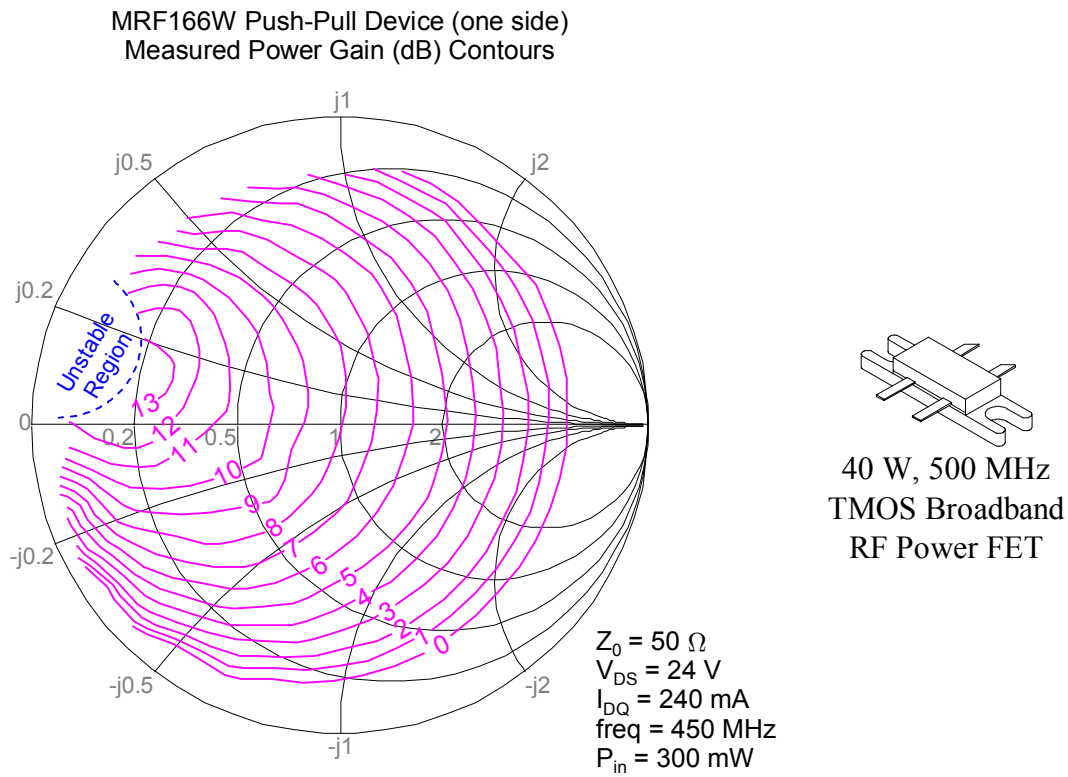


Figure 3.19. Load-pull power gain contours measured using the RTL measurement system. RTL measurement provides real-time *in situ* measurements of Γ_L and DUT performance without requiring precision tuners.

3.7 Normalization of Measurements

RF power device measurements are often made in a non-$50\,\Omega$ reference impedance environment. The reference impedance is determined by the impedance of the line standard, and is typically designed to be the width of the transistor leads. The measurement reference impedance Z_{o1} is assumed to be purely real for the frequencies of interest. Small-signal S-parameters are usually given for a $50\,\Omega$ environment. It is often necessary to renormalize the measured data to an impedance of $Z_{o2} = 50\,\Omega$. Either reference impedance Z_{o1} or Z_{o2} could be complex, in which case a more general derivation of the normalization procedure would need to be used. For a complete derivation of the normalization procedure for complex impedances the reader is referred to (Ha, 1990).

Before deriving the normalization procedure, consider first the connection of two transmission lines with real impedance Z_{o1} and Z_{o2} as illustrated in Fig. 3.20a. To transform the values measured in reference impedance Z_{o1} to the normalization impedance

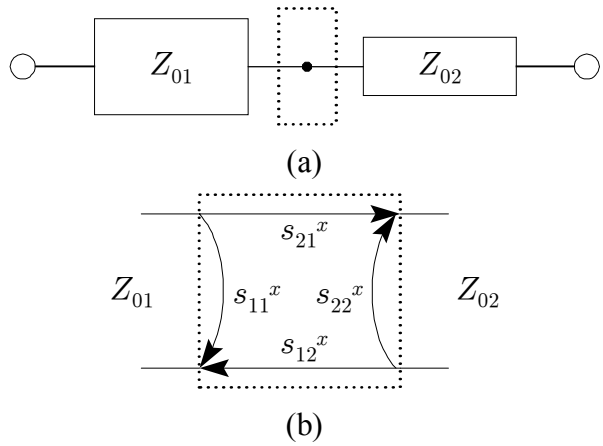


Figure 3.20. Connection between reference impedance Z_{o1} and normalization impedance Z_{o2}.

Z_{o2} consider the transition at the connection point between the two impedances as indicated by the dotted outline in Fig. 3.20. The connection point is shown schematically as in Fig. 3.20b as a two-port device. This two-port will be used to describe the transformation of the measurements from a reference impedance of Z_{o1} to that of reference impedance Z_{o2}. This transformation represents an equivalent lossless, reciprocal, two-port network.

The S-parameters of the network in Fig. 3.20b are denoted s_{ij}^x, where x indicates the network transformation. The reflection coefficient s_{11}^x, seen looking from the Z_{o1} impedance at port 1 into port 2 terminated by Z_{o2} is given in terms of the terminating impedances as

$$s_{11}^x = \frac{Z_{o2} - Z_{o1}}{Z_{o2} + Z_{o1}} \ . \tag{3.7}$$

Similarly the reflection coefficient s_{22}^x seen looking from the Z_{o2} impedance at port 2 into port 1 terminated by Z_{o1} is

$$s_{22}^x = \frac{Z_{o1} - Z_{o2}}{Z_{o1} + Z_{o2}} = -s_{11}^x \ . \tag{3.8}$$

The relationship between s_{22}^x and s_{11}^x as given above holds only if the terminating impedances are purely real. Note also that for the case of purely real termination impedances the reflection coefficients are real. The forward transmission through the network, given by s_{21}^x, can be determined by considering the voltage incident at port 1 and the voltage transmitted to port 2

$$
\begin{aligned}
s_{21}^x = \frac{b_2}{a_1} &= \frac{V_{2t}/\sqrt{Z_{o2}}}{V_{1i}/\sqrt{Z_{o1}}} = \frac{A_v V_{1t}}{V_{1i}}\sqrt{\frac{Z_{o1}}{Z_{o2}}} \\
&= A_v(1+s_{11}^x)\sqrt{\frac{Z_{o1}}{Z_{o2}}} \\
&= (1+s_{11}^x)\sqrt{\frac{Z_{o1}}{Z_{o2}}} = s_{12}^x
\end{aligned}
\tag{3.9}
$$

where V_{1i} = voltage incident at port 1
V_{1t} = total voltage at port 1
V_{2t} = total voltage at port 2
A_v = voltage gain of the network = 1.

The voltages are normalized by the termination impedances Z_{o1} and Z_{o2} to make the expression valid in a generalized power wave sense. With the transition between the reference impedance and normalization impedance specified by the two-port network of Fig. 3.20 the DUT can be embedded between the two impedances as shown in Fig. 3.21.

The measured data provides information on Γ, τ, and Γ_L as a function of a_1. The renormalized values are denoted by $'$ so that after renormalizing to Z_{o2} the data becomes a_1', Γ', τ', and Γ_L'.

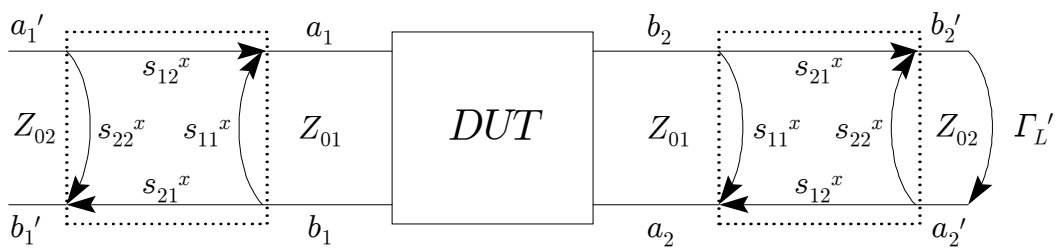


Figure 3.21. Transforming DUT S-parameters from reference impedance Z_{o1} to normalization impedance Z_{02}.

The reflection Γ' is given in terms of the known values as

$$
\Gamma' = \frac{b_1'}{a_1'} = s_{22}^x + \frac{s_{12}^x s_{21}^x \Gamma}{1 - s_{11}^x \Gamma} = \frac{s_{22}^x + \Gamma}{1 + s_{22}^x \Gamma} \ . \tag{3.10}
$$

The adjusted drive level a_1' is found by first writing an expression for a_1 in terms of a_1'

$$
a_1 = a_1' \frac{s_{12}^x}{1 - s_{11}^x \Gamma} \ . \tag{3.11}
$$

The corresponding expression for a_1' can be obtained by algebraic manipulation of (3.11), giving the final expression

$$a_1' = a_1 \frac{1 + s_{22}^x \Gamma}{s_{21}^x} \ . \tag{3.12}$$

The normalized load reflection Γ_L' can be obtained by first writing an expression for the load reflection Γ_L as

$$\Gamma_L = \frac{a_2}{b_2} = s_{11}^x + \frac{s_{12}^x s_{21}^x \Gamma_L'}{1 - s_{22}^x \Gamma_L'} \ . \tag{3.13}$$

Solving for Γ_L' gives the following for the normalized load reflection

$$\Gamma_L' = \frac{s_{22}^x + \Gamma_L}{1 + s_{22}^x \Gamma_L} \ . \tag{3.14}$$

The normalized transmission is defined as

$$\begin{aligned} \tau' &= \left. \frac{b_2'}{a_1'} \right|_{a_2' = b_2' \Gamma_L'} \\ &= \frac{b_2'}{a_1'} \frac{a_1 \tau}{b_2} = \frac{b_2'}{b_2} \tau \frac{s_{21}^x}{1 + s_{22}^x \Gamma} = \tau \frac{{s_{21}^x}^2}{(1 + s_{22}^x \Gamma)(1 - s_{22}^x \Gamma_L')} \ . \end{aligned} \tag{3.15}$$

Substituting (3.14) into the above expression and simplifying gives the expression for the normalized transmission

$$\tau' = \tau \frac{1 + s_{22}^x \Gamma_L}{1 + s_{22}^x \Gamma} \ . \tag{3.16}$$

Using the expressions for the normalized a_1', Γ', τ', and $\Gamma_L'$the data measured in a Z_{o1} reference impedance can be normalized to a Z_{o2} impedance system. To verify that the normalization is being properly implemented measurements can be made for a thru or a transmission line. For the thru the normalized S-parameters are

$$\begin{aligned} s'_{11_{Thru}} &= s_{22}^x + \frac{s_{11}^x s_{12}^x s_{21}^x}{1 - {s_{11}^x}^2} &&= 0 \\ s'_{21_{Thru}} &= \frac{s_{12}^x s_{21}^x}{1 - {s_{11}^x}^2} &&= 1 \\ s'_{12_{Thru}} &= s'_{21_{Thru}} \\ s'_{22_{Thru}} &= s'_{11_{Thru}} \end{aligned} \tag{3.17}$$

which is the expected result.

The RTL measurement system and robust calibration procedure have been described. Using the RTL measurement system, techniques were presented for obtaining reliable and accurate vector large-signal measurements on active devices, neglecting harmonics. As discussed in section 2.6, neglecting these harmonic content is a reasonable assumption in

the measurement system. The major drawbacks in neglecting harmonics appear in harmonic rejection for transmit applications and the estimate of efficiency. Example small- and large-signal measurements were presented. The line standard, typically chosen to correspond to the device lead width, sets the measurement system reference impedance. Equations for renormalizing the measurements to a standard $50\,\Omega$ reference impedance to provide typical S-parameter data were derived. The RTL measurement system makes available a wealth of nonlinear device operating information that cannot be readily obtained using existing microwave measurement systems.

CHAPTER 4

SCATTERING PARAMETER LARGE-SIGNAL DEVICE CHARACTERIZATION

When a nonlinear two-port device is dominated by a signal nonlinearity, which is often the case, design information may be obtained from a simple three-port model. The three-port model is a black box representation of the nonlinear device which enables the use of S-parameter design concepts in nonlinear circuit design. A method of extracting model parameters from reflection and transmission measurements versus drive level is presented. The three-port model is used to simulate load-pull measurements on a high-power RF transistor and the results are compared directly with measurements.

4.1 Three-Port Representation of Nonlinear Devices

A scattering parameter model consisting of a three-port linear system with one port terminated by a nonlinear reflection coefficient has been developed (Davis, 1999). The resulting three-port model, represented as a signal flow diagram in Fig. 4.1, describes the behavior of a nonlinear two-port device where performance is dominated by a single nonlinear characteristic. Port 3 is terminated by the nonlinearity Γ_n, which is a monotonic function of the input drive level and is constrained to be memoryless and lossless. The constraint that Γ_n be memoryless means that it will be purely real-valued. For small-signal (linear) operation $\Gamma_n = 0$ and the model reduces to a simple two-port linear network. As the input drive level increases, device performance is affected by Γ_n which could represent a nonlinear I vs. V, Q vs. V, or similar nonlinear relationship occurring within the device. Using the three-port model, a set of equivalent two-port parameters are defined

$$s'_{ij} = s_{ij} + \frac{s_{i3}s_{3j}\Gamma_n}{1 - s_{33}\Gamma_n}, \quad i, j \in [1, 2] \ . \tag{4.1}$$

For large-signal operation the equivalent S-parameters could be considered nonlinear

S-parameters in that they capture the nonlinear characteristics of the device. Under small-signal conditions the equivalent parameters are the small-signal S-parameters of the device. The linear model parameters s_{ij} would be expected to change across frequency, while Γ_n should be independent of frequency.

There is much that can be learned about nonlinear device behavior by undergoing a thorough study of the model coefficients for a given device type as well as expanding the model definition to include frequency effects. A designer could use the model to determine the appropriate input and output terminations, device stability, and output power as a function of input power. The model development has yet to incorporate

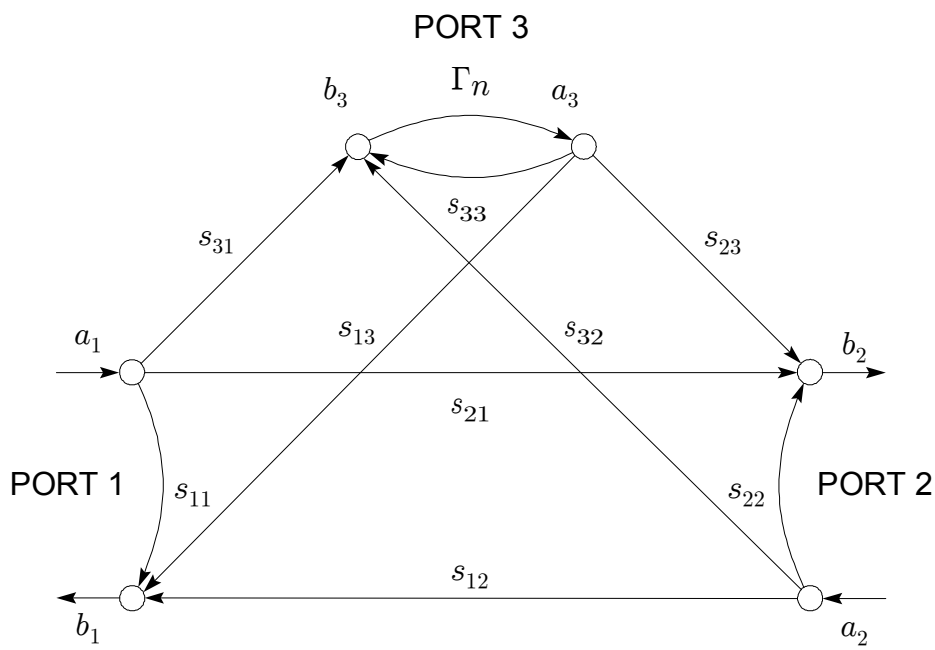


Figure 4.1. Signal flow diagram of three-port nonlinear model.

harmonic related effects. The three-port model also offers an intuitive view of the device nonlinearity that may be useful in evaluating and comparing the linear performance of several devices from various manufacturers.

Model parameters s_{11}, s_{12}, s_{21}, and s_{22} describe the linear operation of the device. The S-parameters associated with port 3 in conjunction with Γ_n describe the nonlinear behavior of the device. Just as analysis of s_{11}, s_{12}, s_{21}, and s_{22} can add insight and understanding to the linear device characteristics, the S-parameters associated with port 3 can bring insight and understanding to the large-signal, nonlinear characteristics of the device. As a result of not being able to directly measure Γ_n at port 3, the five model parameters associated with the third port cannot be determined explicitly, but they can be found as ratios. Figs. 4.2, 4.3, and 4.4 illustrate the three sets of parameter ratios that complete the large signal model : s_{13}/s_{23}, s_{32}/s_{31}, and $s_{23}s_{31}$, in addition to s_{33} and Γ_n.

The signal flow paths associated with s_{13}/s_{23} are indicated in Fig. 4.2 and can be thought of as the nonlinear port output coefficient. A signal a_3 applied at port 3 can

appear at either port 2 or port 1. The ratio s_{13}/s_{23} provides an indicator of how the nonlinearity at port 3 affects output ports 1 and 2 respectively. Consider the ideal case as

$$|s_{23}| > |s_{13}| \ ,$$

such that the net signal flow in the device being modeled is from port 1 to port 2 and a majority of the signal affected by Γ_n at a_3 would appear on port 2, typical of field effect devices.

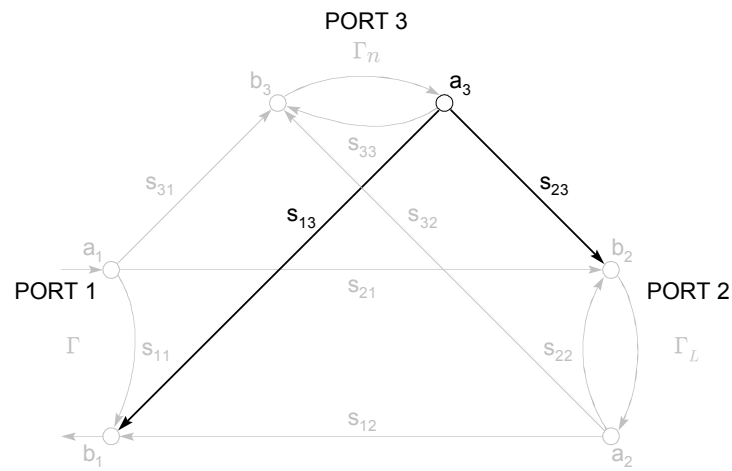


Figure 4.2. Diagram of three-port model with large signal output paths s_{13} and s_{23}. These model parameters indicate how the nonlinearity at port 3 affects ports 1 and 2. The ratio s_{13}/s_{23} is the nonlinear port output coefficient.

An alternative case could be

$$|s_{13}| > |s_{23}| \ ,$$

in which case port 1 would be affected more by the nonlinearity than port 2. This case could suggest potential instability. It is possible that when a device is oscillating the measured input power to port 1 is negative—an indication that the effects of the nonlinearity are being channeled back to the input port. It is hypothesized that the typical value for the nonlinear port output coefficient is given as

$$\left| \frac{s_{13}}{s_{23}} \right| < 1 \ ,$$

indicating the nonlinearity is dominant at the output. This inequality has been found to be appropriate for all the bipolar and field-effect devices studied.

Fig. 4.3 highlights the signal flow paths associated with the nonlinear port directivity coefficient s_{32}/s_{31}. Ideally the signal at port 1 would dominate the nonlinear behavior, suggesting that

$$|s_{31}| > |s_{32}|\,,$$

for a typical device. Even if this relation holds, a highly reflective load Γ_L could still affect the nonlinear performance. Analysis of the nonlinear device behavior could be facilitated

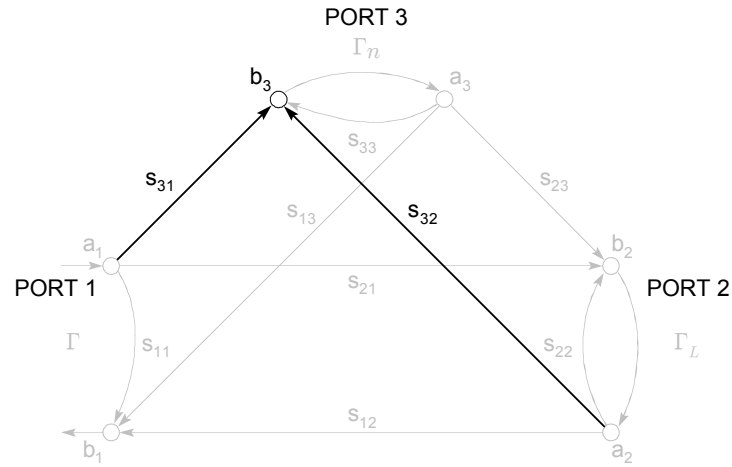


Figure 4.3. Diagram of three-port model with nonlinear port input paths s_{31} and s_{32}. These model parameters indicate how drive levels on ports 1 and 2 affect the value of Γ_n. The ratio of s_{32}/s_{31} is the nonlinear port directivity coefficient. Ideally port 1 input would dominate the value of Γ_n.

by considering the nonlinear port directivity coefficient across a broad band of frequencies, suggesting future possibilities for model extension. For example, an amplifier with a two-tone input signal f_1 and f_2 can exhibit improved linearity by selecting a highly reflective load Γ_L at the harmonics $2f_1$, $2f_2$ and the difference frequency $|f_1 - f_2|$ (Staudinger, 1997c; Sevic, 1998; Staudinger 1997b). The load at the fundamental frequency could also be selected to minimize an unnecessary drive level being applied to the nonlinear port input which would further affect the overall linearity of the output signal.

Fig. 4.4 illustrates the signal flow paths through port 3, in effect connecting the input port to the output port through the nonlinearity. The coefficient $s_{23}s_{31}$ could be termed the large signal forward path coefficient. This product could be a strong indicator of the potential linearity (or nonlinearity) of a particular device. For example, if this coefficient had the value

$$|s_{23}s_{31}| \approx 0\,,$$

it would be expected that the device output would be very linear with respect to changes on the input drive level. Even with a small value for $s_{23}s_{31}$ the input and output reflection could change with drive level due to the effects of the other two large signal .

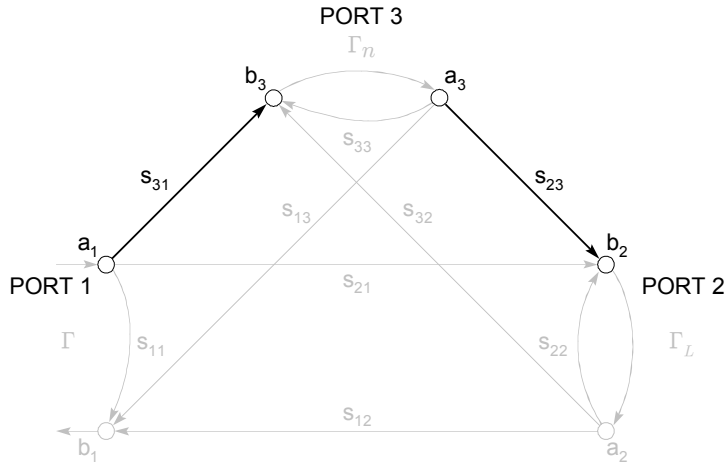


Figure 4.4. Diagram of the three-port model with large signal forward path coefficient $s_{23}s_{31}$. The magnitude of this coefficient is an indication of the effect Γ_n will have on the device output at port 2.

model coefficients associated with port 3. A study of $s_{23}s_{31}$ for several devices could provide an indicator of potential linear (or nonlinear) performance. Such an indicator would be a useful figure of merit to allow designers to choose a suitable device from among several possible candidates.

Model coefficient s_{33} cannot be measured and cannot be separated easily from the other coefficients. The value of s_{33} is chosen to simplify the extraction process, and is typically set to $s_{33} = -j$, with the sign determining whether Γ_n is monotonically increasing or decreasing.

4.2 Three-Port Parameter Extraction

The nine model parameters and Γ_n are extracted from measurements of Γ and τ, using two independent load terminations Γ_{L1} and Γ_{L2}, while sweeping the input power level from small-signal operation into device gain compression. An expression for the reflection Γ and transmission τ can be written in terms of a single nonlinear reflection coefficient Γ_n using the signal flow graph of the three-port model shown in Fig. 4.1,

$$\Gamma = \frac{s_{11} - \Delta s\Gamma_L + (|S|\Gamma_L - s_{11}s_{33} + s_{13}s_{31})\Gamma_n}{(1 - s_{22}\Gamma_L) - [s_{33}(1 - s_{22}\Gamma_L) + s_{23}s_{32}\Gamma_L]\Gamma_n}$$

$$\tau = \frac{s_{21} - (s_{21}s_{33} - s_{23}s_{31})\Gamma_n}{(1 - s_{22}\Gamma_L) - [s_{33}(1 - s_{22}\Gamma_L) + s_{32}s_{23}\Gamma_L]\Gamma_n} \tag{4.2}$$

with Δs given as the determinant of the 2×2 small-signal S-parameter matrix

$$\Delta s = s_{11}s_{22} - s_{12}s_{21} \tag{4.3}$$

and $|S|$ is the determinant of the 3×3 S-parameter matrix of the three-port network

$$\begin{aligned} \det S = |S| = s_{11}(s_{22}s_{33} - s_{23}s_{32}) - s_{12}(s_{21}s_{33} - s_{23}s_{31}) \\ + s_{13}(s_{21}s_{32} - s_{22}s_{31}). \end{aligned} \tag{4.4}$$

The expressions for Γ and τ given in (4.2) are bilinear transformations of the variable Γ_n and can be expressed in the form

$$w = \frac{az + b}{cz + d} \tag{4.5}$$

where a, b, c, and d are complex constants and z is a complex variable. An important theorem of bilinear transformations is that any circle on the complex plane is transformed into a circle by a bilinear transformation. In the limit, the circle is a straight line created by increasing the radius of the circle and moving the center away from the origin (Kurokawa, 1969).

An illustration of the circular trajectory for the measured transmission τ as indicated by (4.2) is shown in Fig. 4.5 for an enhancement-mode MOSFET. Measured Γ has characteristics of a bilinear transform as indicated by (4.2) with a trajectory similar to the one in Fig. 4.5. Note that the denominator for both Γ and τ of (4.2) are the same. It can be shown that Γ is a bilinear transform of τ. The first measured data point in Fig. 4.5 is indicated by a small circle. The small-signal region of operation is a small cluster of data

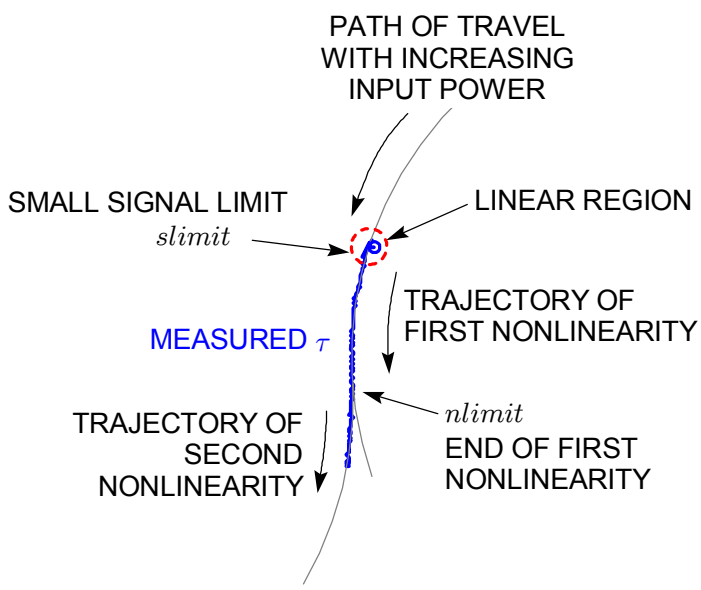


Figure 4.5. Illustration of τ as a function of large input drive level. The circular trajectory of τ is a bilinear transform of Γ_n within the region dominated by the first nonlinearity as given by (4.2).

points about a fixed point. There will be some variation of the data due to random noise within the measurement setup. The last data point within the small-signal region is denoted by $slimit(\ell)$, where ℓ denotes the load reflection for which data was measured, i.e. $\ell = 1, 2$ corresponds to Γ_{L1} and Γ_{L2} respectively. As the input drive increases, the trajectory of τ traces out a circular arc indicating nonlinear operation. The radius of the circle and the rate of travel along the circumference of the circle are device and load dependent. The bilinear transformation of (4.2) indicates the straight line variation of a memoryless nonlinearity as defined by a real function would transform to a circle in the complex plane. As the device varies from this ideal (single nonlinearity) model, the trajectory changes direction and begins tracking along the perimeter of another circular path corresponding to a second nonlinearity. The last data point in the region dominated by the first nonlinearity is labeled $nlimit(\ell)$. Data points beyond $nlimit(\ell)$ are within the second region of nonlinearity. The three-port model will describe device operation up to the onset of the second nonlinearity. This variation typically occurs far into compression. The second nonlinearity is not often encountered in normal device usage as it tends to indicate the onset of device break down modes that could severely decrease the reliable lifetime of operation.

The RTL measurement system is used to obtain the necessary measurement data because the parameter extraction measurements are of necessity nonlinear and typically require high drive levels. The device should be biased and operating at a temperature similar to its intended design application environment. The input drive levels should be similar for both load termination measurements. While it is not absolutely necessary for the device to be running at the intended output power, the input power for the measurements should include the intended range of input powers for the final design to ensure that the model will be able to predict performance at the required output power. Model parameter extraction can be performed with equipment commonly available in an RF design lab and requires only two power sweep measurements for each frequency of interest. The details of extracting the three-port model parameters are discussed in the sections below. The reader may skip to section 4.3 for a discussion of how the model may be used to simulate device performance.

4.2.1 Extraction of the Small-Signal Model Coefficients

The small-signal operation of the device will result in a cluster of measured reflection and transmission data about a fixed point for each load. When the measured data leaves the cluster and begins traversing a curved path as shown in Fig. 4.5, the device leaves the small-signal region of operation and enters a large-signal regime of operation. Fig. 4.6 is an enlarged detail of the cluster of measured τ within the small-signal region.

For extracting the model coefficients from the measured data it is necessary to define a value $slimit(\ell)$ that indexes the last small-signal data point in the measured data for each load $\ell = 1, 2$. If many low-level measurements are included in the data set then there may be some deviation of the small-signal data about a fixed point. The value of $slimit(\ell)$ should be chosen such that the device is just leaving the linear operating region.

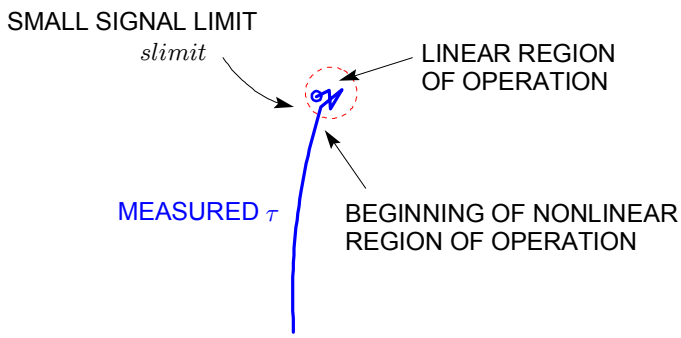


Figure 4.6. Enlarged view of τ showing detail of the small-signal region. While the device is operating in a linear region, the measured data will cluster about a fixed point. As drive level increases, τ traces out a circular path and departs the linear region of operation.

The small-signal reflection and transmission constants $\Gamma_{\ell ss}$ and $\tau_{\ell ss}$ are found for each load, $\ell = 1, 2$ using a weighted average up to and including the data point indexed by $slimit(\ell)$

$$\Gamma_{\ell ss} = \frac{\sum_{k=1}^{slimit(\ell)} \Gamma_{\ell k} |a_{1\ell k}|}{\sum_{k=1}^{slimit(\ell)} |a_{1\ell k}|} \ , \quad \tau_{\ell ss} = \frac{\sum_{k=1}^{slimit(\ell)} \tau_{\ell k} |a_{1\ell k}|}{\sum_{k=1}^{slimit(\ell)} |a_{1\ell k}|} \tag{4.6}$$

where the measured values of $\Gamma_{\ell k}$ and $\tau_{\ell k}$ have been restricted in the summation to the small-signal region. The data is weighted by the magnitude of the input power to smooth out any effects of noise on the small-signal data.

Using Γ_{ss} and τ_{ss}, the small-signal portion of the linear model is found using the equations:

$$\begin{aligned} \mathcal{D} &= \tau_{1ss}\Gamma_{L1} - \tau_{2ss}\Gamma_{L2} \\ s_{11} &= (\tau_{1ss}\Gamma_{2ss}\Gamma_{L1} - \tau_{2ss}\Gamma_{1ss}\Gamma_{L2})/\mathcal{D} \\ s_{21} &= \tau_{1ss}\tau_{2ss}(\Gamma_{L1} - \Gamma_{L2})/\mathcal{D} \\ s_{12} &= (\Gamma_{1ss} - \Gamma_{2ss})/\mathcal{D} \\ s_{22} &= (\tau_{1ss} - \tau_{2ss})/\mathcal{D} \ . \end{aligned} \tag{4.7}$$

Note that the above equations are identical in form to those of (3.1).

4.2.2 Extraction of the Large-Signal Model Coefficients

Several fundamental coefficients may be defined in association with port 3 rather than using the S-parameters directly. These coefficients form ratios and products of the

S-parameters that cannot be explicitly separated from each other since measurements cannot be directly made at port 3. To illustrate this point, we first define the nonlinear portion of the model and related parameters as

$$\mathfrak{z} = s_{23}s_{31}\gamma_n = s_{23}s_{31}\frac{\Gamma_n}{1 - s_{33}\Gamma_n} \tag{4.8}$$

with γ_n a complex nonlinear form related to Γ_n, the original real nonlinearity.

The measured transmission τ can be written in terms of (4.8) and the remaining model parameters as

$$\tau = \frac{(s_{21} + \mathfrak{z})/(1 - s_{22}\Gamma_L)}{1 - \left(\frac{s_{32}}{s_{31}}\right)\mathfrak{z}\frac{\Gamma_L}{1-s_{22}\Gamma_L}}$$

or equivalently

$$\tau = \frac{s_{21}}{1 - s_{22}\Gamma_L} + \mathfrak{z}\frac{\left[1 + \left(\frac{s_{32}}{s_{31}}\right)\frac{s_{21}\Gamma_L}{1-s_{22}\Gamma_L}\right]/(1 - s_{22}\Gamma_L)}{1 - \left(\frac{s_{32}}{s_{31}}\right)\mathfrak{z}\frac{\Gamma_L}{1-s_{22}\Gamma_L}} \ . \tag{4.9}$$

Similarly for Γ we obtain

$$\Gamma = \frac{s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1-s_{22}\Gamma_L} + \mathfrak{z}\left[\left(\frac{s_{13}}{s_{23}}\right) - \frac{\Gamma_L}{1-s_{22}\Gamma_L}\left(s_{11}\frac{s_{32}}{s_{31}} - s_{12} - s_{21}\left(\frac{s_{13}}{s_{23}}\right)\left(\frac{s_{32}}{s_{31}}\right)\right)\right]}{1 - \left(\frac{s_{32}}{s_{31}}\right)\mathfrak{z}\frac{\Gamma_L}{1-s_{22}\Gamma_L}}$$

or equivalently

$$\Gamma = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} + \mathfrak{z}\frac{\left(\frac{s_{13}}{s_{23}} + \frac{s_{12}\Gamma_L}{1-s_{22}\Gamma_L}\right)\left(1 + \frac{s_{32}}{s_{31}}\frac{s_{21}\Gamma_L}{1-s_{22}\Gamma_L}\right)}{1 - \left(\frac{s_{32}}{s_{31}}\right)\mathfrak{z}\frac{\Gamma_L}{1-s_{22}\Gamma_L}} \ . \tag{4.10}$$

This data, comprised of Γ and τ, form the set of measured data for each load versus drive level. The expressions for τ and Γ are similar to the bilinear form given earlier in (4.2) with a slight change in form due to the introduction of complex function $\mathfrak{z}$ in place of Γ_n. For simplicity, the small-signal data may be substituted in (4.9) and (4.10) as

$$\tau_{ss} = \frac{s_{21}}{1 - s_{22}\Gamma_L}$$

and

$$\Gamma_{ss} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} = s_{11} + s_{12}\tau_{ss}\Gamma_L \ , \tag{4.11}$$

where τ_{ss} and Γ_{ss} are the measured small-signal transmission and reflection coefficients. From this data, we now consider the process of determining the coefficients.

4.2.3 Extraction of the Nonlinear Port Output Constant s_{13}/s_{23}

The derivation of the nonlinear port output constant s_{13}/s_{23} is motivated by an attempt to use the measured device data to directly obtain parameters for the three-port nonlinear model. We may take the small-signal terms to the left of eqs. (4.9) and (4.10) and divide the transmission term by the reflection term to obtain

$$\frac{\Gamma - \Gamma_{ss}}{\tau - \tau_{ss}} = \left(\frac{s_{13}}{s_{23}} + \frac{s_{12}\Gamma_L}{1 - s_{22}\Gamma_L}\right)(1 - s_{22}\Gamma_L)\,,$$

which simply gives

$$\frac{s_{13}}{s_{23}} = \frac{\Gamma - \Gamma_{ss} - s_{12}\Gamma_L(\tau - \tau_{ss})}{(\tau - \tau_{ss})(1 - s_{22}\Gamma_L)} = \frac{\Gamma - \tau s_{12}\Gamma_L - s_{11}}{\tau(1 - s_{22}\Gamma_L) - s_{21}}\,. \tag{4.12}$$

The ratio s_{13}/s_{23} is an indication of the relative large signal flow from port 3 to ports 1 and 2 and can be directly obtained from the measured data. To obtain the value of the ratio s_{13}/s_{23} it is desirable to use as much of the measured data as possible for both loads Γ_{L1} and Γ_{L2}. To accomplish this a least squares solution is obtained for the first nonlinear constant by writing an error function ρ which is summed across both loads for all measured data points that are within the first (or dominant) nonlinear region. The error function ρ is defined in terms of (4.12)

$$\rho = \sum_{\ell=1}^{2}\sum_{k=1}^{nlimit(\ell)}\left|\Gamma_{\ell k} - \tau_{\ell k}\, s_{12}\Gamma_{L\ell} - s_{11} - \frac{s_{13}}{s_{23}}[\tau_{\ell k}(1 - s_{22}\Gamma_{L\ell}) - s_{21}]\right|^2\,. \tag{4.13}$$

The value of s_{13}/s_{23} will minimize the error function and can be found by differentiating (4.13) with respect to the real and imaginary parts of s_{13}/s_{23}. For notational convenience, the values x and y are defined to be

$$\begin{aligned} x &= \tau_{\ell k}(1 - s_{22}\Gamma_{L\ell}) - s_{21} \\ y &= \Gamma_{\ell k} - \tau_{\ell k}\, s_{12}\Gamma_{L\ell} - s_{11}\,. \end{aligned} \tag{4.14a}$$

Differentiation gives

$$\frac{\partial\rho}{\partial\,\frac{s_{13}}{s_{23}}^{\Re}} = \sum_{\ell}\sum_{k} -x^*y - xy^* + 2\frac{s_{13}}{s_{23}}^{\Re}|x|^2 = 0 \tag{4.14b}$$

$$\frac{\partial\rho}{\partial\,\frac{s_{13}}{s_{23}}^{\Im}} = \sum_{\ell}\sum_{k} jx^*y - jxy^* + 2\frac{s_{13}}{s_{23}}^{\Im}|x|^2 = 0\,. \tag{4.14c}$$

The superscripts $^{\Re}$ and $^{\Im}$ denote the real and imaginary components respectively. Multiplying (4.14c) by j and adding to (4.14b), we obtain

$$\frac{s_{13}}{s_{23}} = \frac{\sum_{\ell}\sum_{k} x^* y}{\sum_{\ell}\sum_{k} |x|^2} \tag{4.15}$$

The error function ρ as given in (4.13) might be useful for estimating the $nlimit(\ell)$ values. An iterative method could be used to compute s_{13}/s_{23} and then sum (4.13) up to an estimated $nlimit(\ell)$, when the value of ρ deviates from the minimum value by an appreciable amount, the corresponding upper limit on the summation of k could be taken as the value of $nlimit(\ell)$.

4.2.4 Extraction of the Nonlinear Port Directivity Coefficient s_{32}/s_{31}

The nonlinear port directivity coefficient s_{32}/s_{31} is directly related to $s_{23}s_{31}\gamma_n$ or equivalently $\mathfrak{z}$ in eq. (4.8). This relationship may be rewritten as

$$s_{23}s_{31}\gamma_n = \frac{(\tau - \tau_{ss})(1 - s_{22}\Gamma_L)}{1 + \frac{s_{32}}{s_{31}}\tau\Gamma_L} \tag{4.16}$$

The right hand side of the above is known except for the coefficient s_{32}/s_{31}. To use data from both loads in determining the unknown coefficient, consider relating (4.16) using measured data for load 1 with measured data for load 2 which would give the form

$$\frac{(\tau_{1k} - \tau_{1ss})(1 - s_{22}\Gamma_{L1})}{1 + \frac{s_{32}}{s_{31}}\tau_{1k}\Gamma_{L1}} = \frac{(\tau_{2k} - \tau_{2ss})(1 - s_{22}\Gamma_{L2})}{1 + \frac{s_{32}}{s_{31}}\tau_{2k}\Gamma_{L2}} \tag{4.17}$$

Transmission $\tau_{\ell k}$ is the measured transmission with load $\Gamma_{L\ell}$. Index k indicates the measured transmission for both loads corresponds to the same input drive level, i.e. for a given a_{1k} there is a corresponding τ_{1k}, τ_{2k}. The details of extracting τ_{1k} and τ_{2k} will be discussed below.

For the sake of clarity, the derivation will continue with the assumption that τ_{1k} and τ_{2k} are known. From (4.17), the coefficient s_{32}/s_{31} is given by

$$\begin{aligned}\frac{s_{32}}{s_{31}}&[\tau_{2k}\Gamma_{L2}(\tau_{1k} - \tau_{1ss})(1 - s_{22}\Gamma_{L1}) - \tau_{1k}\Gamma_{L1}(\tau_{2k} - \tau_{2ss})(1 - s_{22}\Gamma_{L2})] \\ &+ (\tau_{1k} - \tau_{1ss})(1 - s_{22}\Gamma_{L1}) - (\tau_{2k} - \tau_{2ss})(1 - s_{22}\Gamma_{L2}) = 0\end{aligned} \tag{4.18a}$$

which may be simplified along with the substitution of the small-signal terms as

$$\frac{s_{32}}{s_{31}} = \frac{\tau_{2k}(1 - s_{22}\Gamma_{L2}) - \tau_{1k}(1 - s_{22}\Gamma_{L1})}{\tau_{2k}(\tau_{1k} - s_{21})\Gamma_{L2} - \tau_{1k}(\tau_{2k} - s_{21})\Gamma_{L1}} \ . \tag{4.18b}$$

Use (4.18b) to define an error function that will be optimized to obtain a value for the s_{32}/s_{31} coefficient. Define new constants $\mathcal{T}_{1k}$ and $\mathcal{T}_{2k}$ which will be used for shorthand notation in writing (4.18b)

$$
\begin{aligned}
\mathcal{T}_{1k} &= \tau_{1k}(1 - s_{22}\Gamma_{L1}) \\
\mathcal{T}_{2k} &= \tau_{2k}(1 - s_{22}\Gamma_{L2}) \\
\mathcal{T}_{\ell k} &= \tau_{\ell k}(1 - s_{22}\Gamma_{L\ell}) \ .
\end{aligned}
\tag{4.19}
$$

The error function will take into account all measurement data for both loads within the nonlinear region of the device operation. Rewriting (4.18b) gives the error function

$$
\rho = \sum_{k=slimit(\ell)+1}^{a3limit} \left| \frac{s_{32}}{s_{31}} [\tau_{2k}(\tau_{1k} - s_{21})\Gamma_{L2} - \tau_{1k}(\tau_{2k} - s_{21})\Gamma_{L1}] - [\mathcal{T}_{2k} - \mathcal{T}_{1k}] \right|^2 \tag{4.20}
$$

The error function spans the measured data associated with the first (dominant) nonlinearity. The summation begins at $slimit(\ell) + 1$ which is the first nonlinear data point outside of the small-signal region and ends at $a3limit$. The data index $a3limit$ is chosen to be

$$
a3limit = \min(nlimit(1), nlimit(2))
$$

since measured τ data from both loads is used in (4.20) and one load will have fewer measured data points within the first region of nonlinearity. It is possible that measured data will extend beyond $a3limit$ for one of the loads without leaving the first region of nonlinearity, while for the other load the data would be well within a region dominated by second nonlinearity. Choosing $a3limit$ to be the minimum of $nlimit(\ell)$ ensures that the model parameters are extracted from measured data within the first region of nonlinearity.

The desired value of s_{32}/s_{31} will minimize the error ρ and can be obtained by differentiating (4.20) with respect to the real and imaginary parts of s_{32}/s_{31} to obtain

$$
\frac{s_{32}}{s_{31}} = \frac{\sum_k [\tau_{2k}(\tau_{1k} - s_{21})\Gamma_{L2} - \tau_{1k}(\tau_{2k} - s_{21})\Gamma_{L1}]^* (\mathcal{T}_{2k} - \mathcal{T}_{1k})}{\sum_k |\tau_{2k}(\tau_{1k} - s_{21})\Gamma_{L2} - \tau_{1k}(\tau_{2k} - s_{21})\Gamma_{L1}|^2} \tag{4.21}
$$

where * denotes the complex conjugate.

The solution for s_{32}/s_{31} requires paired values for the transmission τ_{1k} and τ_{2k} corresponding to the measured values for Γ_{L1} and Γ_{L2} where the transmission values are related to each other by (4.16)

$$
\begin{aligned}
\tau_{1k} &\Rightarrow s_{23}s_{31}\gamma_n \\
\tau_{2k} &\Rightarrow s_{23}s_{31}\gamma_n
\end{aligned}
$$

for a given drive level. For a drive level a_{1k} it is unlikely that the signal at port 3 will be the same value for both Γ_{L1} and Γ_{L2}. Consequently, the resultant τ_{1k} and τ_{2k} will not correlate to the same value of $s_{23}s_{31}\gamma_n$. Further, $s_{23}s_{31}\gamma_n$ is unknown. The $s_{23}s_{31}\gamma_n$ term is not directly available, but may be related to $s_{23}a_3$ which has a one-to-one correspondence to $s_{23}s_{31}\gamma_n$. The nonlinear reflection coefficient γ_n can be expressed in terms of $\tau_{\ell k}$ using (4.16) and (4.17)

$$\gamma_n = \frac{(\tau_{\ell k} - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell})}{(s_{31} + s_{32}\tau_\ell \Gamma_{L\ell})s_{23}} \ . \tag{4.22}$$

The output of port 3 can be written in terms of γ_n by considering the signal flow path through port 3 and substituting in (4.8)

$$\begin{aligned} a_3 = b_3\Gamma_n &= a_1 \frac{s_{31} + s_{32}\tau\Gamma_L}{1 - s_{33}\Gamma_n}\Gamma_n \\ &= a_1(s_{31} + s_{32}\tau\Gamma_L)\gamma_n \ . \end{aligned} \tag{4.23}$$

Substituting (4.23) into (4.22), we obtain a relation between γ_n and $s_{23}a_3$

$$1 = \frac{a_1(\tau_\ell - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell})}{s_{23}a_3} \ . \tag{4.24}$$

Simplifying the above yields an expression that is a function of $(\tau_\ell - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell})$

$$s_{23}a_3 = a_1(\tau_\ell - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell}) \ . \tag{4.25}$$

The values on the right hand side of (4.25) are measured, therefore the $s_{23}a_3$ values are known. As an approximation, the $s_{23}a_3$ values can be used to estimate a matching set of $\tau_{\ell k}$ to initiate the process of solving for s_{32}/s_{31}. The $s_{23}a_3$ product can be written in terms of τ for input drive level a_{1k}

$$|s_{23}a_3|_k = |a_{1k}(\tau_{\ell k} - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell})| \ . \tag{4.26}$$

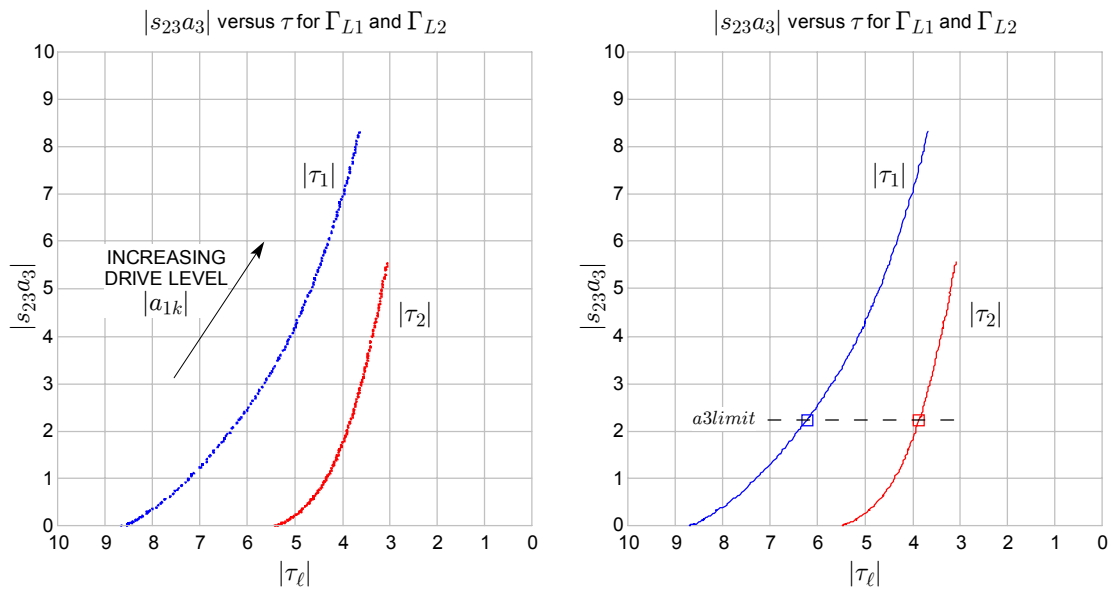
A table is generated by first computing the $|s_{23}a_3|_k$ values for all measured $\tau_{\ell k}$ and $\Gamma_{L\ell}$. The table of $|s_{23}a_3|_k$ is sorted into ascending order and a record of the $\tau_{\ell k}$ and $\Gamma_{L\ell}$ used to compute the $|s_{23}a_3|_k$ value is also recorded. The resulting data structure is shown in Table 4.1 below. The computed values of $|s_{23}a_3|_k$ will have one measured transmission value $\tau_{\ell k}$ corresponding to either load $\ell = 1, 2$. There will be an unknown transmission

$\lvert s_{23}a_3\rvert_k$	Γ_{L1}	Γ_{L2}
$\vdots$	τ_{1k1}	$\vdots$
Ascending order $\downarrow$	$\vdots$	$\vdots$
	$\vdots$	τ_{2kx}
	$\vdots$	$\vdots$
	τ_{1ky}	$\vdots$
$\vdots$	$\vdots$	$\vdots$

Table 4.1. Ordering of $|s_{23}a_3|_k$ for use in estimating unknown values of $\tau_{\nu k}$. For each $|s_{23}a_3|_k$ value, at least one measured $\tau_{\ell k}$ is known.

value $\tau_{\nu k}$ for the other load $\nu \neq \ell$ (i.e. the subscript ν will be used to denote the load with the unknown $\tau_{\ell k}$ value).

Table 4.1 is represented graphically in Fig. 4.7 using measured data for the MRF160 at $f = 450\,\text{MHz}$. The loads used in Fig. 4.7 are $\Gamma_{L1} = 0.681\angle 125.2°$ and $\Gamma_{L2} = 0.941\angle -176.5°$. Note that the two curves are different in overall shape. The $|s_{23}a_3|$ values for τ_{2k} are lower than the values for τ_{1k} for approximately the same input drive level a_1. The load corresponding to τ_{2k} is considerably more reflective than the first



(a) (b)

Figure 4.7. Graph showing $|s_{23}a_3|$ as a function of $\tau_{\ell k}$ for (a) measured values only and (b) measured and interpolated values. Interpolated values are necessary for extracting the directivity coefficient s_{32}/s_{31}. Data in (b) above $a3limit$ is within a second nonlinear region of operation.

load and from Fig. 4.3 it could be reasoned that the phase alignment across s_{32} creates destructive interference at the third port such that $|s_{23}a_3|$ is smaller for an equivalent input power. The gaps between $|s_{23}a_3|$ data points in Fig. 4.7a indicates an unknown transmission value $\tau_{\nu k}$ that will be estimated. The unknown values $\tau_{\nu k}$ in the table are estimated by using linear interpolation to obtain a complete set of values. Fig. 4.7b shows a plot of the $\tau_{\ell k}$ data after linear interpolation is used between the measured $\tau_{\ell k}$ data points. The ratio s_{32}/s_{31} is computed using data below the $a3limit$ in Fig. 4.7b. Beyond the $a3limit$ a second nonlinearity becomes dominant indicating the device is operating in a highly nonlinear region.

Once the table of $\tau_{\ell k}$ values is complete, (4.21) can be used to compute the nonlinear port directivity coefficient s_{32}/s_{31}. With the coefficient s_{32}/s_{31} known, the measured values of $\tau_{\ell k}$ can be used with (4.16) to compute $\mathfrak{z}_k$

$$\mathfrak{z}_k = s_{23}s_{31}\gamma_n = \frac{(\tau_{\ell k} - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell})}{1 + \frac{s_{32}}{s_{31}}\tau_{\ell k}\Gamma_{L\ell}} \tag{4.16}$$

where the sequencing of $\mathfrak{z}_k$ will be based on the ascending order of $|s_{23}a_3|_k$ as computed in (4.26). The ordering of $\tau_{\ell k}$ used to compute $\mathfrak{z}_k$ is taken from the sorted $|s_{23}a_3|_k$ in Table 4.1, using only the measured transmission values in the computation. A plot of $\mathfrak{z}$ for the MRF160 data set is shown in Fig. 4.8. The values of $\mathfrak{z}$ computed using load 1 and load 2 measured data are shown separately in Fig. 4.8. Ideally the plot of $\mathfrak{z}$ should be single valued, giving the same value for measured data from both loads. The behavior of $\mathfrak{z}$

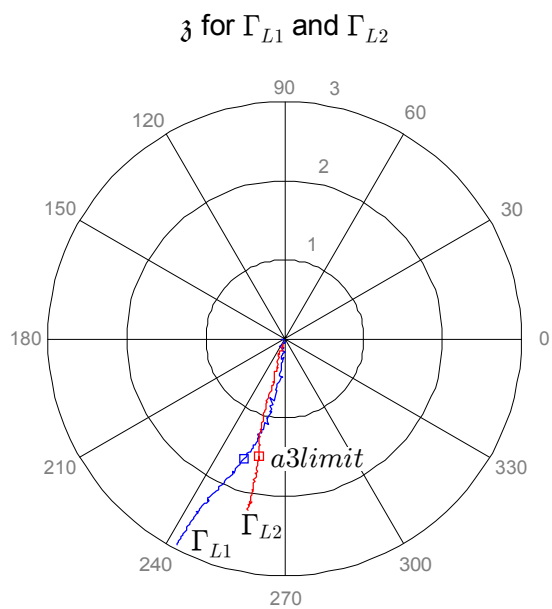


Figure 4.8. Computed $\mathfrak{z}$ for measured data from both loads Γ_{L1} and Γ_{L2}.

in Fig. 4.8 would suggest that the ratio s_{32}/s_{31} is not optimal. A better estimate of s_{32}/s_{31} should bring the computed $\mathfrak{z}$ for both loads closer together.

One method of refining the value of s_{32}/s_{31} would be to use the computed $\mathfrak{z}$ for load 1 and load 2 to obtain a refined value for $\mathfrak{z}$, label this value $\mathfrak{z}'$. One method to obtain this refined value $\mathfrak{z}'$ would be to obtain an expression for the transformation from $\tau_{\ell k}$ to $\mathfrak{z}_{\ell k}$. These two expressions could then be used to interpolate a new $\mathfrak{z}'$. From (4.16) write an error function that can be used for least squares optimization of the s_{32}/s_{31} ratio using the computed $\mathfrak{z}$

$$\rho = \sum_{k=1}^{a3limit} \left| (\tau_{\ell k} - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell}) - \mathfrak{z}_k \left(1 + \frac{s_{32}}{s_{31}}\tau_{\ell k}\Gamma_{L\ell}\right) \right|^2 . \tag{4.27}$$

Differentiating (4.27) with respect to the real and imaginary part of s_{32}/s_{31} , we obtain a

form with the nonlinear port directivity coefficient given in terms of $\mathfrak{z}_k$

$$\frac{s_{32}}{s_{31}} = \frac{\sum_k (\mathfrak{z}_k \tau_{\ell k} \Gamma_{L\ell})^* [(\tau_{\ell k} - \tau_{\ell ss})(1 - s_{22}\Gamma_{L\ell}) - \mathfrak{z}_k]}{\sum_k |\mathfrak{z}_k \tau_{\ell k} \Gamma_{L\ell}|^2} \ . \tag{4.28}$$

This form of solution for s_{32}/s_{31} is slightly different from (4.21) in that measured data from both loads *and* the initial computed s_{32}/s_{31} are used.

4.2.5 Extraction of the Large-Signal Forward Path Coefficient $s_{23}s_{31}$ and s_{33}

We seek a way to obtain the forward path coefficient $s_{23}s_{31}$ in terms of what has been measured or already computed. Modify (4.8) by multiplying through by $1/\left(1 - s_{33}^{\Re}\Gamma_n\right)$ to isolate the imaginary component of s_{33} in the denominator

$$\begin{aligned} \mathfrak{z} &= s_{23}s_{31}\frac{\alpha_n}{1 - js_{33}^{\Im}\alpha_n} \\ \alpha_n &= \frac{\Gamma_n}{1 - s_{33}^{\Re}\Gamma_n} \ . \end{aligned} \tag{4.29}$$

The value for α_n is purely real since by definition Γ_n is the reflection coefficient for a memoryless nonlinearity and is therefore purely real. The value for $\mathfrak{z}$ is complex due to the values of $\tau_{\ell k}$ and $\Gamma_{L\ell}$ used in (4.16) requiring either one or both of $s_{23}s_{31}$ and s_{33} to be complex values as shown in (4.29) above. Solving for α_n in terms of $\mathfrak{z}$

$$\alpha_n = \frac{\mathfrak{z}}{s_{23}s_{31} + js_{33}^{\Im}\mathfrak{z}} \ . \tag{4.30}$$

The requirement that α_n be real leads to the following

$$\frac{\mathfrak{z}}{s_{23}s_{31} + js_{33}^{\Im}\mathfrak{z}} - \frac{\mathfrak{z}^*}{s_{23}^*s_{31}^* - js_{33}^{\Im}\mathfrak{z}^*} = 0 \ . \tag{4.31}$$

It is required to determine the values of $s_{23}s_{31}$ and $s_{33}^{\Im}$ that will make (4.31) a true statement. Multiply through by the denominators in (4.31) and define an error function

$$\rho = \sum_{k=1}^{a3limit} \left| S^{\Re}\mathfrak{z}_k^{\Im} - S^{\Im}\mathfrak{z}_k^{\Re} - s_{33}^{\Im}|\mathfrak{z}_k|^2 \right|^2 \ , \tag{4.32}$$

where $S = s_{23}s_{31}$. The values of $s_{23}s_{31}$ and $s_{33}^{\Im}$ should be selected such that they minimize the error. Differentiate (4.32) with respect to the real and imaginary parts of $s_{23}s_{31}$ and divide by $s_{33}^{\Im}$ to obtain two equations for the unknown real and imaginary parts of $s_{23}s_{31}$

$$
\begin{aligned}
\frac{\partial \rho}{\partial S^{\Re}} &= \frac{S^{\Re}}{s_{33}^{\Im}} \sum_k \mathfrak{z}^{\Im 2} - \frac{S^{\Im}}{s_{33}^{\Im}} \sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} = \sum_k \mathfrak{z}^{\Im} |\mathfrak{z}|^2 \\
\frac{\partial \rho}{\partial S^{\Im}} &= \frac{S^{\Re}}{s_{33}^{\Im}} \sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} - \frac{S^{\Im}}{s_{33}^{\Im}} \sum_k \mathfrak{z}^{\Re 2} = \sum_k \mathfrak{z}^{\Re} |\mathfrak{z}|^2 \ .
\end{aligned}
\tag{4.33}
$$

The above equations form a system of linear equations in the form

$$
By = c
$$

where

$$
B = \begin{bmatrix} \sum_k \mathfrak{z}^{\Im 2} & -\sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} \\ \sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} & -\sum_k \mathfrak{z}^{\Re 2} \end{bmatrix} \quad y = \begin{bmatrix} \Re(s_{23} s_{31}) / s_{33}^{\Im} \\ \Im(s_{23} s_{31}) / s_{33}^{\Im} \end{bmatrix} \quad c = \begin{bmatrix} \sum_k \mathfrak{z}^{\Im} |\mathfrak{z}|^2 \\ \sum_k \mathfrak{z}^{\Re} |\mathfrak{z}|^2 \end{bmatrix}. \tag{4.34}
$$

The solution of this system gives the unknown value for $s_{23} s_{31} / s_{33}^{\Im}$, where $s_{33}^{\Im} \neq 0$. The eq$s$. in (4.34) can be used to obtain

$$
\frac{s_{23} s_{31}}{s_{33^{\Im}}} = \frac{\sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} \sum_k \mathfrak{z}^{\Re} |\mathfrak{z}|^2 - \sum_k \mathfrak{z}^{\Re 2} \sum_k \mathfrak{z}^{\Im} |\mathfrak{z}|^2 + \jmath \sum_k \mathfrak{z}^{\Im 2} \sum_k \mathfrak{z}^{\Re} |\mathfrak{z}|^2 - \jmath \sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} \sum_k \mathfrak{z}^{\Im} |\mathfrak{z}|^2}{\left(\sum_k \mathfrak{z}^{\Re} \mathfrak{z}^{\Im} \right)^2 - \sum_k \mathfrak{z}^{\Re 2} \sum_k \mathfrak{z}^{\Im 2}}. \tag{4.35}
$$

All that remains is to determine the proper value for $s_{33}^{\Im}$. For simplicity, the value of $s_{33}^{\Im}$ is chosen to be

$$
s_{33}^{\Im} = -1. \tag{4.36a}
$$

The sign of $s_{33}^{\Im}$ is chosen such that Γ_n is monotonically increasing. The value of $s_{33}^{\Re}$ is selected to be

$$
s_{33}^{\Re} = 0 \tag{4.36b}
$$

without a loss of generality.

4.2.6 Extraction of Γ_n and Remaining Port 3 Model Parameters

With the value of $s_{23} s_{31}$ determined, the remaining three-port model parameters can be computed. The following values are set

$$
\begin{aligned}
s_{33} &= j s_{33}^{\Im} \\
s_{31} &= 1
\end{aligned}
\tag{4.37}
$$

such that s_{33} is either purely imaginary or zero, and s_{31} is purely real. If the model were to be expanded to include frequency then s_{31} and s_{33} could be chosen to be complex such that Γ_n is frequency independent. Using the coefficient ratios of the three-port model, the remaining signal flow graph terms are found

$$
\begin{aligned}
s_{23} &= \frac{s_{23}}{s_{31}} s_{31} \\
s_{13} &= \frac{s_{13}}{s_{23}} s_{23} \\
s_{32} &= \frac{s_{32}}{s_{31}} s_{31} \ .
\end{aligned}
\tag{4.38}
$$

This completes the computation of all the linear model terms.

With the model terms known, the nonlinear reflection coefficient can be computed from the measured data versus $|s_{23}a_3|$ by using (4.8) to obtain

$$
\Gamma_n = \frac{\mathfrak{z}}{s_{23}s_{31} + s_{33}\mathfrak{z}} \ . \tag{4.39}
$$

The value of $\mathfrak{z}$ was computed from measured data using (4.16). Typical data is given in Fig. 4.9.

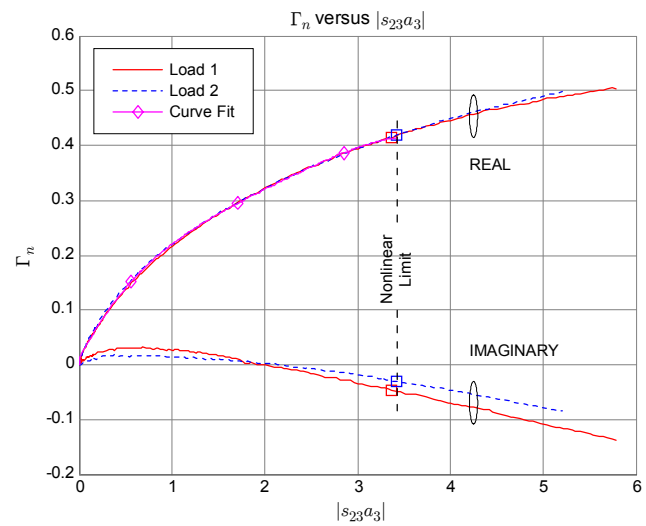


Figure 4.9. Plot of Γ_n as a function of $|s_{23}a_3|$ for MRF160. Only $\Re(\Gamma_n)$ is used for the three-port network. Ideally, $\Im(\Gamma_n) = 0$, the residual $\Im(\Gamma_n)$ indicates how well the extracted three-port network fits the measured data. The curve fit to Γ_n is used in the load-pull simulation to provide intermediate data points between discrete measured values.

The three-port model can be used in simulations to determine device performance characteristics as well as obtain useful information that can be used to design the surrounding impedance transformation and bias circuitry.

4.3 Load-Pull Simulation

4.3.1 Simulation Theory and Implementation

The three-port nonlinear model can be used in simulations to predict device performance about the bias operating point for which the measured data was taken. Fig. 4.10 illustrates the signal flow diagram for the three-port model terminated with a load Γ_L which will be used for simulating load-pull measurements (Cusack, 1974; Mazumder, 1978; Poulin, 1980; Sechi, 1980; Pierpont, 1986; Raab, 2002). As discussed in section 3.6, load-pull is a measurement technique where the device load impedance is varied while measuring the performance of the DUT. Load-pull measurements could be simulated by varying Γ_L,

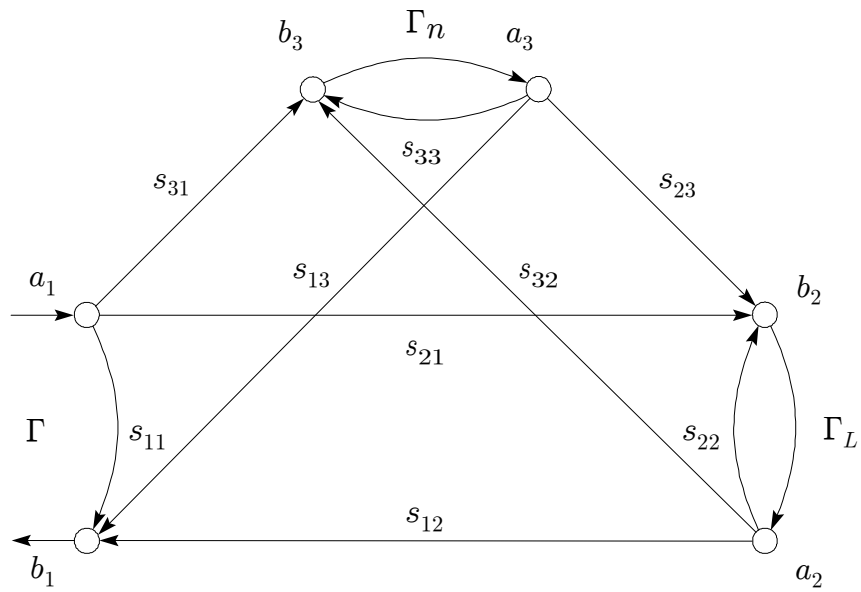


Figure 4.10. Signal flow diagram for load-pull simulation.

tabulating the device performance for a given parameter such as gain, output power, or input reflection, and then using the simulated data to draw contours for optimum Γ_L on a Smith chart.

If a traditional nonlinear model were used to compute load-pull contours, the load termination Γ_L would be varied and a simulation run for each load state. Using the three-port nonlinear model, such a brute force simulation method could also be used to obtain the data for the load-pull contours; however, a more elegant approach has been used that takes advantage of the concise nature of the three-port model. In this section we will

derive equations for two circles in the Γ_L plane for a given Γ_n. The first circle will describe the load states Γ_L for which a given power can be delivered to the load. The second circle will give the load states Γ_L for which a given input power is delivered to the nonlinear device terminated with Γ_L. If both circles exist and intersect, the intersection points will be on the load-pull contour for power gain G_P which describes a given power delivered to the load for a given input power delivered to the device. By computing the intersection points for the circles as the input power is swept, a closed contour can be obtained for G_P. A single contour can be generated by computing the intersection points for a subset of input power levels, a smaller set of computations than in the brute force approach where a single input power and a large set of Γ_L are used to compute device performance. The derivations that follow are for power gain G_P. A similar approach could be used for other performance parameters of interest. Power gain was chosen for this demonstration because of the availability of measured data to verify the simulation results.

The power delivered to the load Γ_L was expressed in chapter 3 as

$$\begin{aligned} P_L &= |b_2|^2\left(1-|\Gamma_L|^2\right) \\ &= |\tau\, a_1|^2\left(1-|\Gamma_L|^2\right) . \end{aligned} \tag{3.4}$$

The power input into a device represented by Fig. 4.10 was given by

$$P_{\text{in}} = |a_1|^2\left(1-|\Gamma|^2\right). \tag{3.5}$$

For the load-pull simulation, P_{in} will be held constant and the loads Γ_L for a given G_P will be plotted as contours in the Γ_L plane (Cusack, 1974; Pierpont, 1986). The power gain of the device is given by (3.6) as the ratio of (3.4) and (3.5) and can be expressed as

$$G_P = \frac{P_L}{P_{\text{in}}} = \frac{|\tau\, a_1|^2\left(1-|\Gamma_L|^2\right)}{|a_1|^2\left(1-|\Gamma|^2\right)} = \frac{|\tau|^2\left(1-|\Gamma_L|^2\right)}{1-|\Gamma|^2} . \tag{4.40}$$

The transmission and reflection coefficients τ and Γ can be written in terms of the three-port S-parameters as

$$\tau = \frac{s'_{21}}{1-s'_{22}\Gamma_L} \tag{4.41}$$

$$\Gamma = \frac{s'_{11}-\Delta s'\Gamma_L}{1-s'_{22}\Gamma_L}. \tag{4.42}$$

The $'$ after the S-parameter is a shorthand notation for the large signal S-parameters of the equivalent two port device given by (4.1) as

$$s'_{ij} = s_{ij} + \frac{s_{i3}s_{3j}\Gamma_n}{1-s_{33}\Gamma_n} \qquad i,j \in [1,2] \tag{4.1}$$

where s_{ij} and Γ_n are the parameters of the three-port model. In (4.42) the quantity $\Delta s'$ is given by

$$\Delta s' = s'_{11}s'_{22} - s'_{12}s'_{21} \ . \tag{4.43}$$

An expression for power gain is obtained by substituting (4.41) and (4.42) into eq. (4.40) to obtain

$$G_P = \frac{P_L}{P_{\text{in}}} = \frac{|s'_{21}|^2\left(1 - |\Gamma_L|^2\right)}{|1 - s'_{22}\Gamma_L|^2 - |s'_{11} - \Delta s'\Gamma_L|^2}. \tag{4.44}$$

The right hand side of (4.44) defines a circle in Γ_L as a function of Γ_n. Expanding (4.44) shows the circle nature of the power gain circles in the Γ_L load plane, resulting in an equation for a circle

$$\left(\Gamma_L^{\Re} - \frac{c_0^{\Re}}{d_1}\right)^2 + \left(\Gamma_L^{\Im} + \frac{c_0^{\Im}}{d_1}\right)^2 = \frac{|c_0|^2}{d_1^2} + \frac{1}{d_1} - \frac{d_0}{d_1}\left(1 - |s'_{11}|^2\right) \tag{4.45}$$

where the real constants d_0 and d_1 are given by

$$\begin{aligned} d_0 &= \frac{1 - \Gamma_L\Gamma_L^*}{(1 - s'_{22}\Gamma_L)(1 - s'_{22}\Gamma_L)^* - (s'_{11} - \Delta s'\Gamma_L)(s'_{11} - \Delta s'\Gamma_L)^*} \\ &= \frac{G_P}{|s'_{21}|^2} \end{aligned} \tag{4.46a}$$

$$d_1 = d_0\left(|s'_{22}|^2 - |\Delta s'|^2\right) + 1 \tag{4.46b}$$

and complex constant c_0 is given by

$$c_0 = d_0(s'_{22} - s'^*_{11}\Delta s') \ . \tag{4.46c}$$

The center and radius of the circle are given by

$$\begin{aligned} C_1 &= \frac{c_0^*}{d_1} \\ R_1^2 &= \frac{|c_0|^2}{d_1^2} + \frac{1}{d_1} - \frac{d_0}{d_1}\left(1 - |s'_{11}|^2\right) \ . \end{aligned} \tag{4.47}$$

Given a power gain G_P, the load terminations Γ_L for which the device will have the specified power gain will fall on the circle with center C_1 and radius R_1. Since the equivalent S-parameters are input power dependent, the Γ_L circle will vary according to the input drive level due to the variation in Γ_n. The Γ_L circle will exist if $R_1^2 > 0$. For small-signal conditions, (4.47) would give the constant power gain circles for the device.

The intersection of the G_P and input power circles will give the load terminations for which the device will meet the specified power gain and input drive conditions for a fixed

Γ_n. The drive level incident on port 1 of the device can be expressed using (3.5)

$$|a_1|^2 = \frac{P_{\text{in}}}{1 - |\Gamma|^2} \tag{4.48}$$

The drive level out of the nonlinearity on the third port can be related to the incident drive level with the expression

$$|s_{23}a_3|^2 = \left|\frac{s_{23}\Gamma_n}{1 - s_{33}\Gamma_n}\right|^2 |a_1|^2 |s_{31} + s_{32}\Gamma_L\tau|^2 \tag{4.49}$$

Substituting (4.48), (4.41), and (4.42) into (4.49) gives the drive level incident on Γ_n as a function of drivel level incident on port 1

$$\frac{|s_{23}a_3|^2}{\left|\frac{s_{23}s_{31}\Gamma_n}{1-s_{33}\Gamma_n}\right|^2 P_{\text{in}}} = \frac{\left|1 - s'_{22}\Gamma_L + \frac{s_{32}}{s_{31}}\Gamma_L s'_{21}\right|^2}{|1 - s'_{22}\Gamma_L|^2 - |s'_{11} - \Delta s'\,\Gamma_L|^2} \tag{4.50}$$

This form also gives circles in Γ_L as a function of input drive level. For convenience we use the complex nonlinear form γ_n which is related to the original real nonlinearity as shown in (4.8)

$$\gamma_n = \frac{\Gamma_n}{1 - s_{33}\Gamma_n} \tag{4.51}$$

The left hand side of (4.50) is a function of γ_n and can be denoted as

$$d_2 = \frac{|s_{23}a_3|^2}{|s_{23}s_{31}\gamma_n|^2 P_{\text{in}}} \tag{4.52}$$

which is purely real. The bilinear nature of (4.50) can be clearly seen by rewriting the equation in the form

$$\Gamma_L\Gamma_L^* d_3 - \Gamma_L c_2 - \Gamma_L^* c_2^* + d_2\left(1 - |s'_{11}|^2\right) - 1 = 0 \tag{4.53}$$

with constants c_2 and d_3 given by

$$\begin{aligned} c_1 &= s'_{22} - \frac{s_{32}}{s_{31}}s'_{21} \\ c_2 &= d_2\left(s'_{22} - s'^*_{11}\Delta s'\right) - c_1 \\ d_3 &= d_2\left(|s'_{22}|^2 - |\Delta s'|^2\right) - |c_1|^2 \end{aligned} \tag{4.54}$$

Constant c_2 is complex, and constant d_3 is purely real. For small-signal conditions c_2 and d_3 are undefined. By completing the square, an expression for constant input power circles in the Γ_L plane is given by

$$\left(\Gamma_L^{\Re} - \frac{c_2^{\Re}}{d_3}\right)^2 + \left(\Gamma_L^{\Im} + \frac{c_2^{\Im}}{d_3}\right)^2 = \frac{|c_2|^2}{d_3^2} + \frac{1}{d_3} - \frac{d_2}{d_3}\left(1 - |s'_{11}|^2\right). \tag{4.55}$$

The center and radius of the circle are given by

$$\begin{aligned} C_2 &= \frac{c_2^*}{d_3} \\ R_2^2 &= \frac{|c_2|^2}{d_3^2} + \frac{1}{d_3} - \frac{d_2}{d_3}\left(1 - |s'_{11}|^2\right) \end{aligned} \tag{4.56}$$

Eq. (4.55) is a constant input power circle in the Γ_L load plane. For a given incident power level $|a_1|^2$ there is a set of load terminations Γ_L for which P_{in} will be constant. The Γ_L terminations corresponding to constant P_{in} are on the circle with center C_2 and radius R_2. The circle will exist if $R_2^2 > 0$ and d_2 is defined. The equivalent S-parameters are input power dependent and the Γ_L circle described by (4.55) will vary with changes in input drive level.

For small-signal conditions the derivation of (4.55) does not hold because $\Gamma_n = 0$ and the three-port model reduces to a simple two-port network consisting of the small-signal S-parameters as given by (4.1). The derivation of input power circles as a function of Γ_L as previously given does not apply when the device is operating linearly and port three has no effect on device performance. For the case when $\Gamma_n = 0$ we derive an alternative equation to (4.55). Begin with (4.48), the drive level incident on port 1, and substitute in for Γ the equivalent small-signal reflection Γ_{ss} defined as

$$\Gamma_{ss} = s_{11} + \frac{s_{12}s_{21}\Gamma_L}{1 - s_{22}\Gamma_L} \tag{4.57}$$

which yields a small-signal equation for drive level incident on port 1

$$|1 - s_{22}\Gamma_L|^2 - |s_{11} - \Delta s\Gamma_L|^2 - \frac{P_{\text{in}}}{|a_1|^2} = 0 \tag{4.58}$$

where the quantity Δs is the small-signal determinant of the two-port S-parameter matrix, namely

$$\Delta s = s_{11}s_{22} - s_{12}s_{21} \tag{4.3}$$

which is the small-signal form of (4.43). Further manipulation of (4.58) gives an equation for an input power circle in the Γ_L plane similar to that derived earlier in (4.55)

$$\left(\Gamma_L^{\Re} - \frac{c_2^{\Re}}{d_3}\right)^2 + \left(\Gamma_L^{\Im} + \frac{c_2^{\Im}}{d_3}\right)^2 = \frac{|c_2|^2}{d_3^2} + \frac{1}{d_3} - \frac{d_2}{d_3}\left(1 - |s_{11}|^2\right) \tag{4.59}$$

with d_2 defined for the small-signal case as

$$d_2 = 1 + \frac{1 - P_{\text{in}}/|a_1|^2}{1 - |s_{11}|^2} \tag{4.60}$$

and small-signal constants c_2 and d_3 given by

$$\begin{aligned} c_2 &= s_{22} - s_{11}^* \Delta s \\ d_3 &= |s_{22}|^2 - |\Delta s|^2 \end{aligned} \tag{4.61}$$

As in (4.54), constant c_2 is complex and d_3 is purely real. Eq. (4.59) describes a circle with center and radius

$$\begin{aligned} C_2 &= \frac{c_2^*}{d_3} \\ R_2^2 &= \frac{|c_2|^2}{d_3^2} + \frac{1}{d_3} - \frac{d_2}{d_3}\left(1 - |s_{11}|^2\right) . \end{aligned} \tag{4.62}$$

For small-signal conditions the input power circle will exist if $R_2^2 > 0$. Because of the similarity between (4.55), (4.56), (4.59), and (4.62) and to promote clarity of notation, eqs. (4.55), (4.56) will be referred to when discussing the input power circles. If the

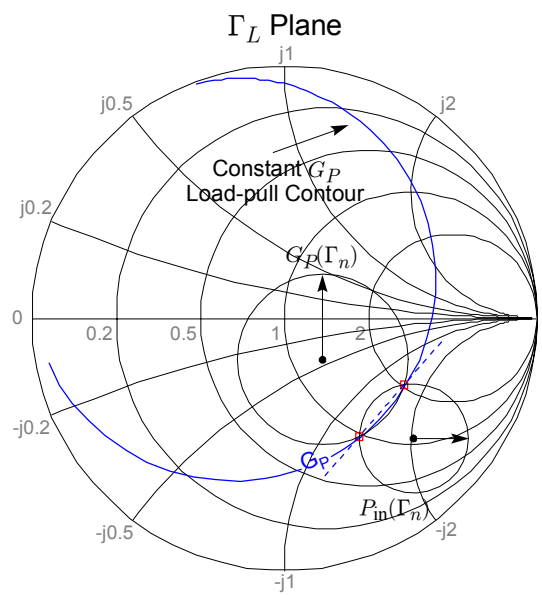


Figure 4.11. Simulation of constant power gain circles using the three-port model. In the Γ_L load plane circles for power gain G_P and input power P_{in} are defined as functions of Γ_n (which in turn is a function of drive level). If the circles exist and intersect the set of Γ_L defined form the constant power gain load-pull contours.

device is operating in a small-signal region where $\Gamma_n = 0$ then it will be understood that (4.59), and (4.62) should be used, along with the proper definitions of the constants c_2, d_2, and d_3.

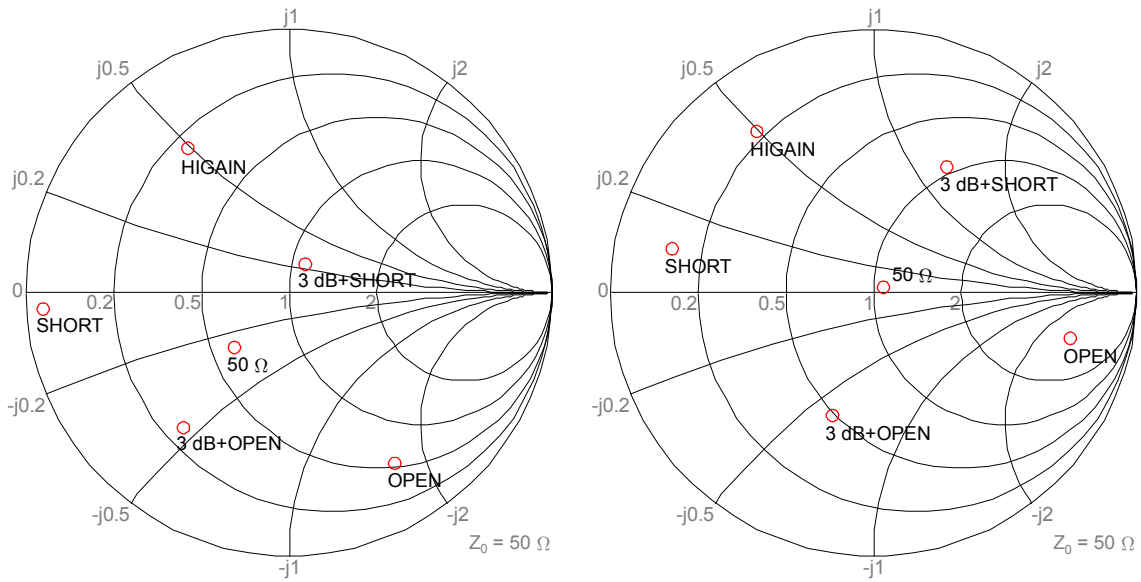
Eqs. (4.45) and (4.55) define circles in the Γ_L plane corresponding to the power gain G_P and power input into the nonlinear device. If both circles exist and intersect, i.e. $C_1 + R_1 e^{j\phi} = C_2 + R_2 e^{j\theta}$ for some ϕ and θ as shown in Fig. 4.11, the intersection points will be the Γ_L values for a given input power that will result in a given power delivered to the load. If the set of all possible Γ_L for a fixed input and output power are plotted on a Smith chart, the result will be a load-pull contour for G_P.

Simulation Implementation

A set of load-pull contours can be quickly generated for a given G_P and P_{in} by sweeping Γ_n across the range of measured input drive levels and computing C_1, R_1, and C_2, R_2 from (4.47) and (4.56) at each value of Γ_n. If both circles exist and intersect, the intersection points can be plotted directly on a Smith chart. To obtain a smooth contour, the individual points are connected using a "connect-the-dots" approach with a straight line between nearest neighbor points. The simulation contours are considerably easier to plot than the contours for arbitrary data points. The smaller the step size increment of Γ_n, the closer the intersection points will be on the Smith chart. The data points available for Γ_n may not be spaced sufficiently close to obtain an adequate distribution of intersection points for smooth contours. For the simulation, a polynomial curve fit using a fourth order polynomial is used to obtain a continuous function representing Γ_n as a function of input drive level as shown in Figs. 4.9 and 4.20. The LAPACK math library routine DGELSS is used to compute the polynomial coefficients for the curve fits by solving the over determined linear least squares problem using single value decomposition (Anderson, 1999). The curve fit allows for arbitrary step size within the simulation, as well as a compact representation of the nonlinear characteristics of the device. The curve fit also filters the effects of random noise on the measurement data. There are some drawbacks to using a polynomial curve fit, $\Gamma_n = 0$ must be forced, otherwise the polynomial would not cross zero. Forcing the curve fit through specific points may compromise the fit elsewhere along the data set. There are fewer computations when computing load-pull contours by stepping through Γ_n than would be the case if a fine mesh of Γ_L were used to compute G_P and P_{in} as in the brute force approach.

Simulation Results

Two TMOS devices were measured, the MRF160 and the MRF166W. These two devices were measured in separate test fixtures using an input power sweep with six different coaxial terminations. Additionally, load-pull measurements were made with these two devices. The MRF160 was load-pull measured for $P_{\text{in}} = 50\,\mu\text{W}$, 5 mW, and 50 mW. One side of the MRF166W was load-pull measured for $P_{\text{in}} = 50\,\mu\text{W}$, 50 mW, and 300 mW. The MRF160 was measured running small and large-signal class A operation. The MRF166W was measured running small and large-signal class B operation. The



(a) Measured Γ_L for MRF160

(b) Measured Γ_L for MRF166W

Figure 4.12. Load terminations, plotted at the device measurement plane, used for power sweep measurements of the MRF160 and MRF166W . Though the coaxial terminations are identical, each fixture has a different phase rotation from the Γ_{Li} plane and the MRF160 fixture has a slight impedance transformation. The resulting device plane Γ_L are labeled for reference and are listed in Table 4.2 in varying combinations for three-port model extraction. The HiGAIN terminations are optimized for each device.

power sweep measurements were at the same bias point as the load-pull measurements. The load terminations used for the MRF160 and MRF166W power sweeps, measured at the device plane, are shown in Fig. 4.12. The load terminations are labeled according to the actual termination used, not according to the location on the Smith chart at the measurement reference plane. As shown in Fig. 4.12 there were differences between the measurement fixtures used for the MRF160 and MRF166W such that even though similar coaxial terminations were used at the system output, Γ_L measured at the port 2 reference plane varied from fixture to fixture.

The HiGAIN termination was found by tuning Γ_L for the highest output power for each device. The 3 dB + OPEN and 3 dB + SHORT terminations were measured using a 3 dB attenuator before the open or short termination. The open termination was measured by not terminating the load port and leaving it open, for the measurement setup this would have been an open female N-connector port. The short termination was a N-connector short. The $50\,\Omega$ load was a high-power $50\,\Omega$ termination capable of handling the expected output power from the MRF166W. The two devices were measured in

separate break-apart fixtures. Each fixture was constructed to have a microstrip line the same width as the device lead at the measurement reference plane. The effect of the non-$50\,\Omega$ transmission line was a slight impedance transformation which is seen in Fig. 4.12 by the $50\,\Omega$ terminations being skewed slightly from the center of the $50\,\Omega$ Smith chart. The six Γ_L terminations will be paired to obtain up to fifteen possible combinations for obtaining a three-port representation of the measured device. The three-port representation requires two independent Γ_L terminations and no further restrictions have been placed on the values of those terminations. By mixing and matching among the six terminations it will be possible to see if some terminations are more desirable than others. The measured devices should give a good indication of the possibilities of the three-port nonlinear representation to predict device performance for class A and class B large signal devices.

4.3.2 Simulation of Small- and Large-Signal Class A Device

The possible Γ_L load combinations for obtaining a three-port representation of the MRF160 from the power sweep measurements are listed in Table 4.2. For each Γ_L sequence a three-port representation was obtained that best fit the measured power sweep

Γ_L sequence	Load 1	Γ_{L1}	Load 2	Γ_{L2}
1	HiGAIN	$0.6805\angle\ 125.20°$	SHORT	$0.9411\angle\ -176.53°$
2	50 Ω	$0.2952\angle\ -136.01°$	OPEN	$0.7553\angle\ -58.44°$
3	3dB+SHORT	$0.1266\angle\ 65.01°$	3dB+OPEN	$0.6498\angle\ -128.92°$
4	HiGAIN	$0.6805\angle\ 125.20°$	3dB+OPEN	$0.6498\angle\ -128.92°$
5	SHORT	$0.9411\angle\ -176.53°$	3dB+SHORT	$0.1266\angle\ 65.01°$
6	HiGAIN	$0.6805\angle\ 125.20°$	3dB+SHORT	$0.1266\angle\ 65.01°$
7	SHORT	$0.9411\angle\ -176.53°$	3dB+OPEN	$0.6498\angle\ -128.92°$
8	SHORT	$0.9411\angle\ -176.53°$	OPEN	$0.7553\angle\ -58.44°$
9	SHORT	$0.9411\angle\ -176.53°$	50 Ω	$0.2952\angle\ -136.01°$
10	HiGAIN	$0.6805\angle\ 125.20°$	OPEN	$0.7553\angle\ -58.44°$
11	HiGAIN	$0.6805\angle\ 125.20°$	50 Ω	$0.2952\angle\ -136.01°$
12	50 Ω	$0.2952\angle\ -136.01°$	3dB+SHORT	$0.1266\angle\ 65.01°$
13	50 Ω	$0.2952\angle\ -136.01°$	3dB+OPEN	$0.6498\angle\ -128.92°$
14	OPEN	$0.7553\angle\ -58.44°$	3dB+SHORT	$0.1266\angle\ 65.01°$
15	OPEN	$0.7553\angle\ -58.44°$	3dB+OPEN	$0.6498\angle\ -128.92°$

Table 4.2. Combinations of Γ_L for MRF160 power sweep measurements.

data. For some Γ_L there were one or two measured points in the power sweep data that were removed from the analysis because of apparent oscillation for a combination of a given load termination and input power level. Except for these occasional bad points the measured data was sufficient to obtain three-port representations for the MRF160. All combinations of Γ_L appeared to be capable of obtaining good agreement with the measured power sweep data. The small-signal limit and nonlinear limits were manually chosen to obtain the best fit of the three-port representation to the measured power sweep data.

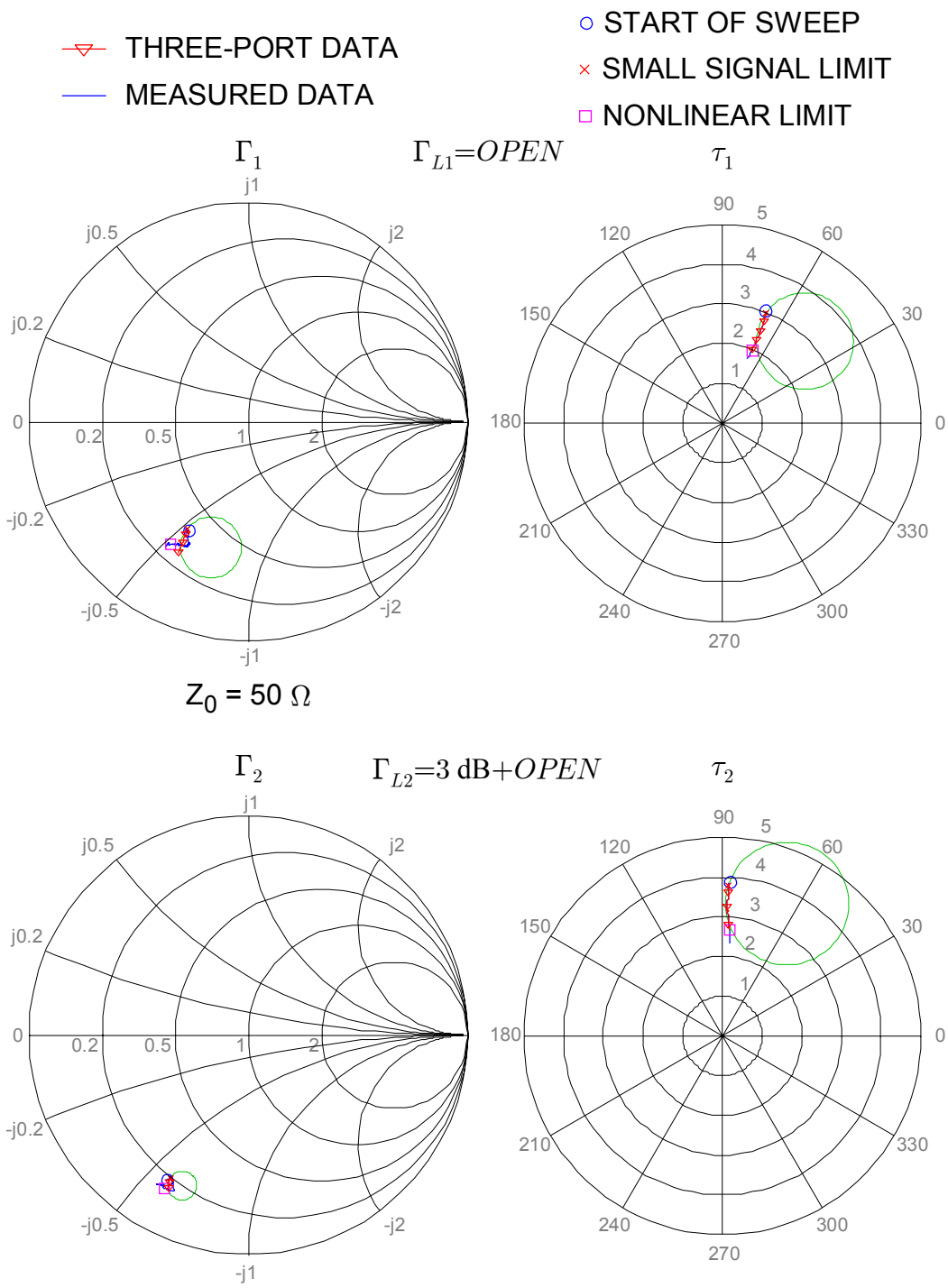


Figure 4.13. Measured and computed power sweep data for MRF160 using Γ_{L1} and Γ_{L2} that resulted in a best-fit match with measured load-pull data.

The fit of the three-port model to the measured data was determined by computing the Γ and τ for Γ_{L1} and Γ_{L2} and comparing the computed reflection and transmission with the measured data. Fig. 4.13 shows two power sweeps of Γ and τ, 1) computed using a three-port model, and 2) measured Γ and τ for $\Gamma_{L1} = \text{OPEN}$ and $\Gamma_{L2} = 3\,\text{dB} + \text{OPEN}$. The load terminations, small-signal limit, and nonlinear limit values used to extract the three-port model gave the best overall fit of the computed three-port data to the measured load-pull data out of all the possible combinations. The computed data for τ_1 and τ_2 is in excellent agreement with the measured power sweep data. The circles shown in Fig. 4.13 are computed based on the bilinear transform properties of the three-port parameters on Γ and τ as given by (4.2). For Γ_{L1} the computed power sweep data follows the expected circular path predicted by the three-port parameters. Though the circles for Γ do not closely track the measured data, the trend is predicted adequately. The MRF160 is a TMOS transistor and Γ changes very little with input power level when compared with a bipolar transistor. The radius of the Γ circles of Fig. 4.13 is $|\Gamma_x| < 0.15$, $x = 1, 2$. For a bipolar transistor Γ would be as circular and smooth as the τ circles, whereas for the TMOS device Γ covers only a small region in the complex plane and exhibits a sharp kink at a high drive level. Despite the discrepancy between measured Γ_1 and the computed data the modeled data is sufficiently close to simulate the device performance.

A load-pull simulation was run using the three-port network parameters computed from the measured data. The output results from the load-pull simulation were compared with measured load-pull data to check the simulation results. Based on this comparison, the selection of small-signal and nonlinear limits for some load combinations improved the load-pull simulation match with the measured data without noticeably effecting the agreement between computed and measured power sweep data. The nonlinear limit for Γ_{L2} is pushed further than required to characterize the first dominant nonlinearity. By including a small portion of data outside the first nonlinearity, the three-port model gave a slightly better fit to the measured load-pull data.

The area in the Γ plane covered by changes in the TMOS transistor input reflection is relatively small across a broad range of input drive level. For the best fit to load-pull measurements it was desirable to have the computed data approximate the overall path of Γ rather than being a close fit to a particular segment of Γ. Fig. 4.14 is a close-up of Γ_2 of Fig. 4.13 with an ideal trajectory for the nonlinear Γ_2 as input power increases. The three-port fit to measured data is a trajectory circle with radius $R = 0.18$ which is an approximation of the path of the measured Γ_2 across the input power range from the small-signal limit to the nonlinear limit. The expected radius based on the bilinear transform of Γ_2 is shown in Fig. 4.14 as radius $R = 0.07$, which is a good approximation of the $R = 0.18$ trajectory. The ideal trajectory has a smaller radius of $R = 0.05$ and is centered to the left of the computed circles to give a trajectory that lines up more closely with the measured Γ_2. The differences in the ideal trajectory and the computed trajectories would indicate that there is a relationship among the three-port network parameters that is not being optimally defined. The discrepancies in radius R affect a small region in the Γ plane and the three-port network does adequately predict nonlinear behavior of the device.

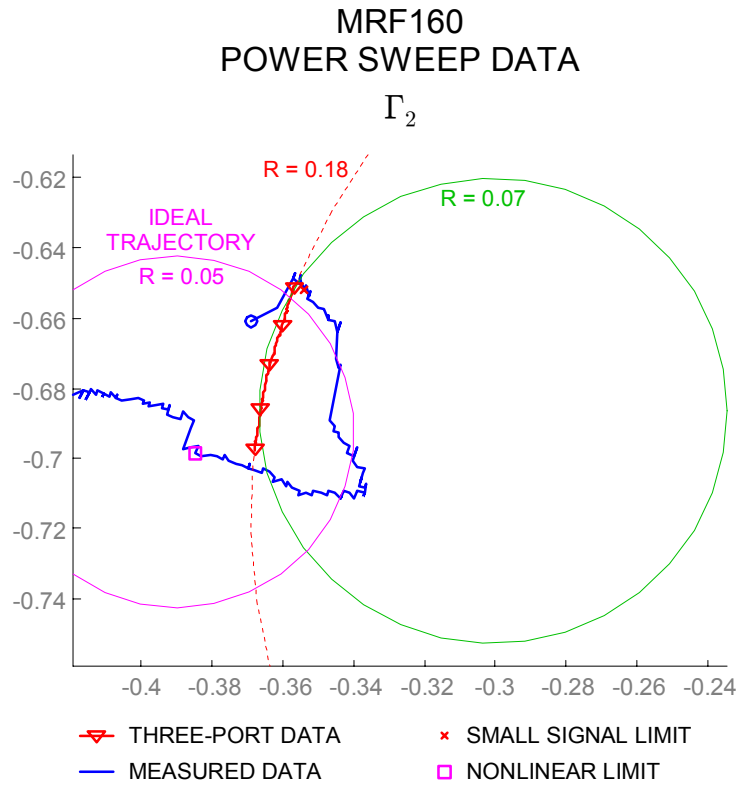


Figure 4.14. Enlarged view of Γ_2 showing detail of three-port computed and measured data. Circle trajectory is influenced by the choice of small signal and nonlinear limits. For best fit to load-pull data it was desirable to have the computed data approximate the overall path of Γ_2 rather than being a close fit to a particular segment.

The nonlinear reflection coefficient Γ_n is computed from the measured data and was plotted as a function $|s_{23}a_3|$ in Fig. 4.9 for the case of $\Gamma_{L1} = \text{OPEN}$ and $\Gamma_{L2} = 3\ \text{dB} + \text{OPEN}$. The Γ_n values associated with each load termination are shown separately in the graph. Ideally the two load terminations would result in identical values for Γ_n, which is almost the case for the real part of Γ_n in Fig. 4.9. The input drive level is closely related to $|s_{23}a_3|$ and this value is readily obtained from the measured data and the three-port network. The computed Γ_n will have a residual imaginary component due to minor discrepancies between the three-port network representation and the actual device. The three-port model was defined with Γ_n being a real valued function of input power. The imaginary residual of Γ_n provides a good indication of how well the nonlinear model fits the measured data. As the drive level increases to a point where the second nonlinearity is dominant, the imaginary component of Γ_n increases substantially in magnitude. The small-signal and nonlinear limits directly affect the overall characteristics of Γ_n. The choice of the nonlinear limit in Fig. 4.9 indicates that some of the second

nonlinearity behavior is being included in Γ_n to achieve a slightly better fit of the simulated load-pull contours to the measured data. The curve fit to the real part of Γ_n averages the effects of Γ_{L1} and Γ_{L2} and facilitates simulation using the model.

The three-port model for the MRF160 extracted using $\Gamma_{L1} = \text{OPEN}$ and $\Gamma_{L2} = 3\text{ dB} + \text{OPEN}$ is shown in Fig. 4.15. The value of $s_{33} = j1$ was selected to give a monotonically increasing Γ_n as shown in Fig. 4.9.

Load-pull simulations were run for three input power levels of $P_{\text{in}} = 50\,\mu\text{W}$, $5\,\text{mW}$, and $50\,\text{mW}$. The simulation results are compared with measured load-pull data in Figs. 4.16, 4.17, and 4.18. In Fig. 4.16 the simulated load-pull contours are in close agreement with the measured load-pull contours. In Fig. 4.17 there is some discrepancy between the measured and simulated contours. One possible cause for the discrepancy may be the operating temperature difference between the load-pull measurements and the power sweep measurements used for model extraction. Additional work needs to be done to determine the cause and resolve differences between measured and modeled data. The measured contours exhibit self-heating effects where the transistor has become hot and the gain is reduced. This reduction in gain is manifested by a reduction in the area of the gain contours as they effectively "shrink" in size. The measured load-pull data took a couple of hours to measure. The temperature difference between points measured early and later can be seen in the 10 and 15 dB contour. The general sequence for load-pull measurements was to start near the short side of the Smith chart and work around the

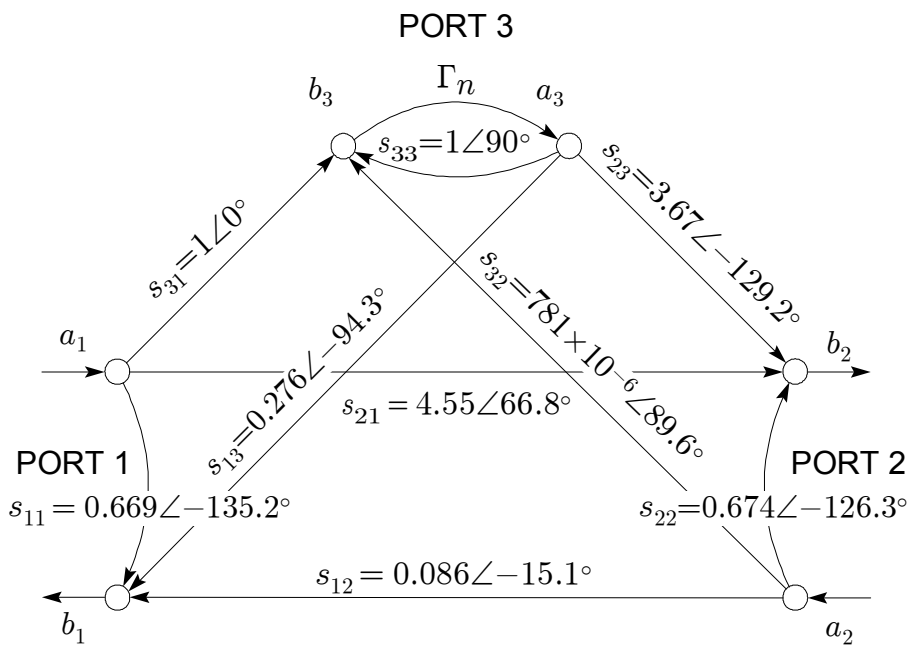


Figure 4.15. Three-port model parameters for the MRF160 at $f = 450\,\text{MHz}$. Device bias was $V_{DD} = 12.5\,\text{V}$, $I_{DQ} = 120\,\text{mA}$, ($V_{GS} = 3.70\,\text{V}$). The model was extracted using $\Gamma_{L1} = \text{OPEN}$ and $\Gamma_{L2} = 3\,\text{dB} + \text{OPEN}$. The Γ_n for this model is shown in Fig. 4.9.

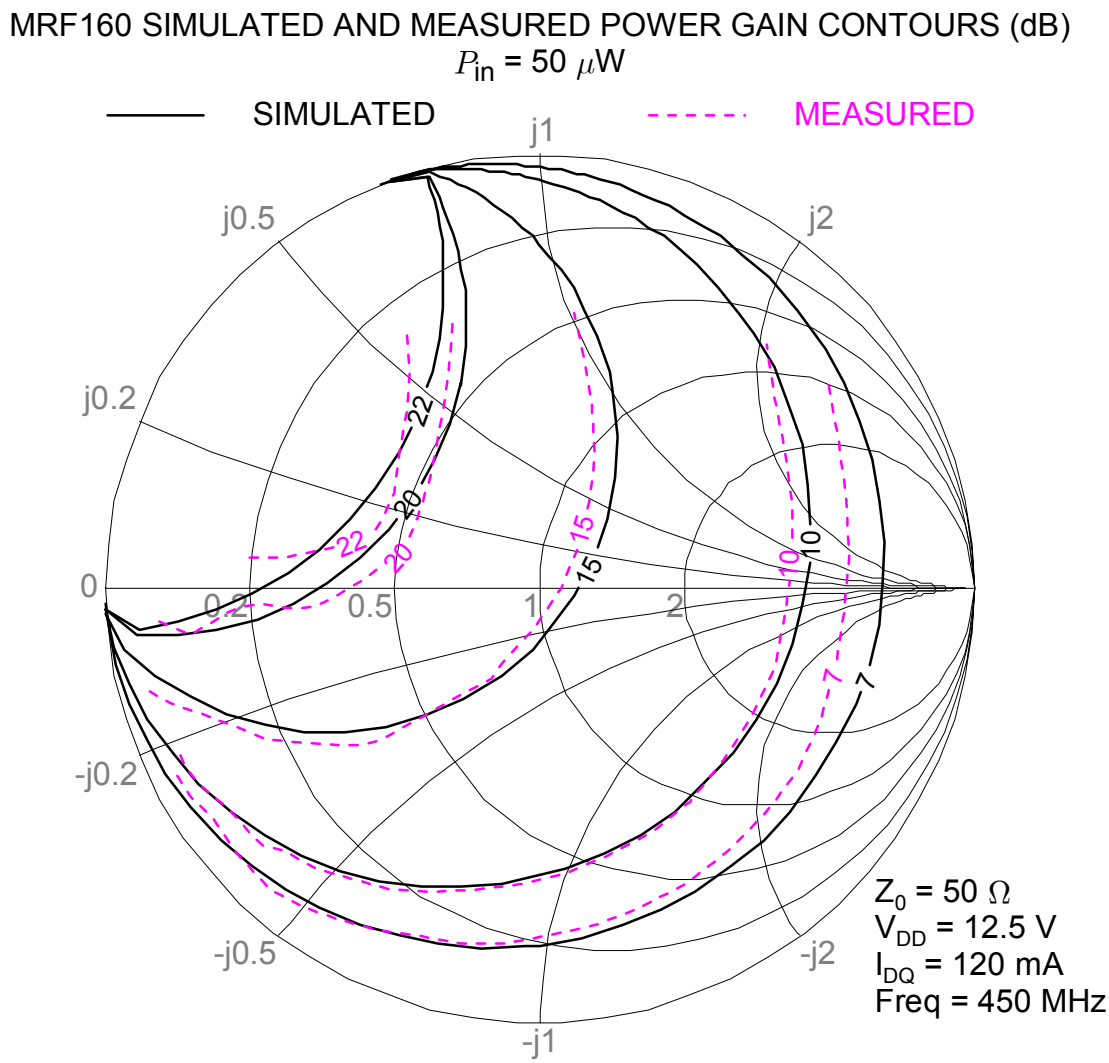


Figure 4.16. MRF160 measured and simulated power gain contours. The simulated contours compare well with the measured contours up to $G_P = 20$ dB. For high gains the measured contours are not smooth arcs as would be expected. The device is unstable for Γ_L terminations in the far upper left quadrant. Fluctuations in the 20 and 22 dB contours could a result of being close to a potentially unstable region.

chart in a clockwise manner. The gain contours are skewed with the bottom of the contours exhibiting more severe heating effects than the top of the contours. The power sweep measurements used for the simulation data were run in a matter of minutes for all six loads. In general, the load terminations used in extracting the model that were

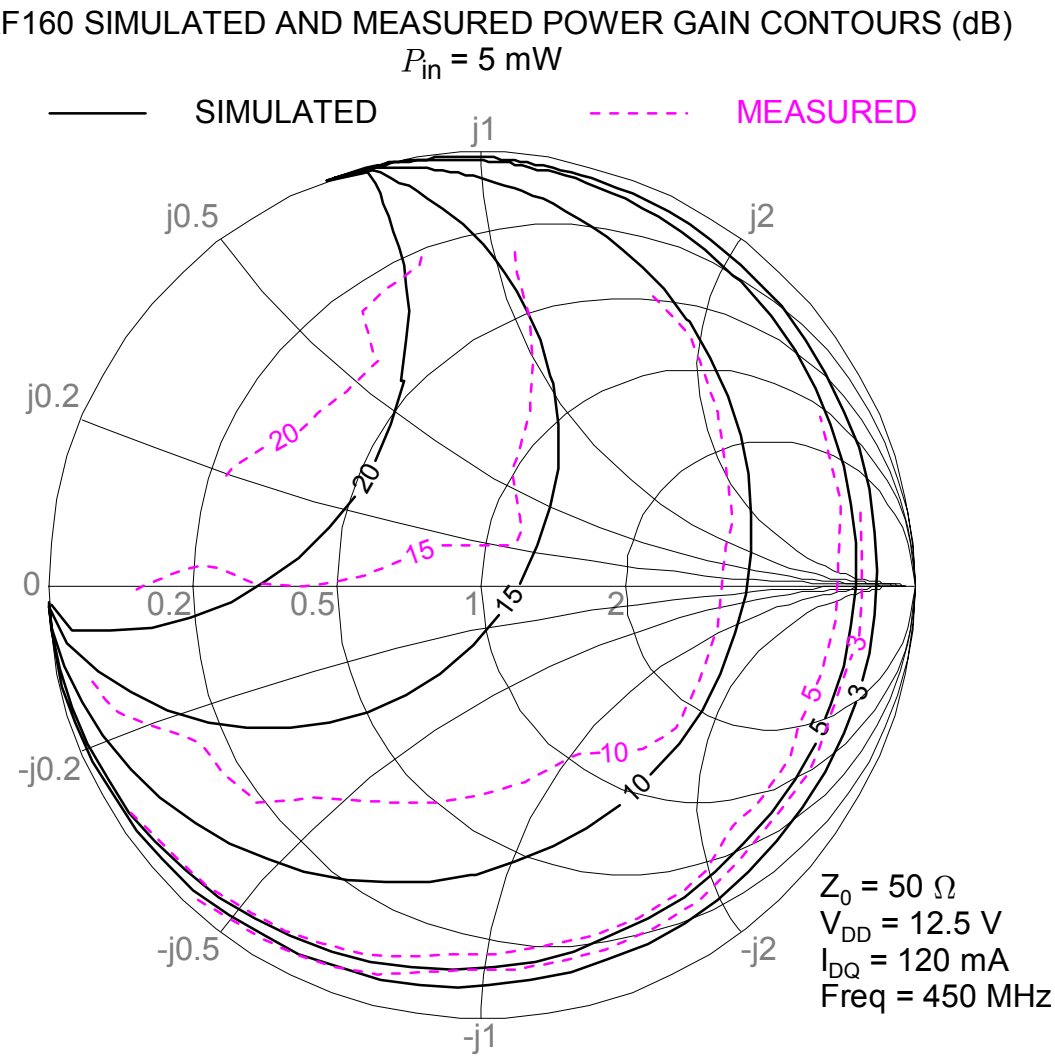


Figure 4.17. MRF160 measured and simulated power gain contours. The irregular measured contours show signs of gain compression due to elevated device temperature. Agreement between simulation and measured contours is closer when the measured contours are not distorted by heating effects.

measured last give a closer fit to the measured load-pull contours than do those model terminations that were measured first.

The effects of heating are most severe on the contours of Fig. 4.18 where it is clearly visible that the gain contours have shrunk significantly in size. The 8 and 11 dB contours show signs of heating. The 13 dB contour should suffer similar heating effects since it corresponds to the highest output power. The simulated load-pull contours are off considerably for $P_{in} = 50\,\text{mW}$. The MRF160 is a 4.0 Watt transistor and for a drive level

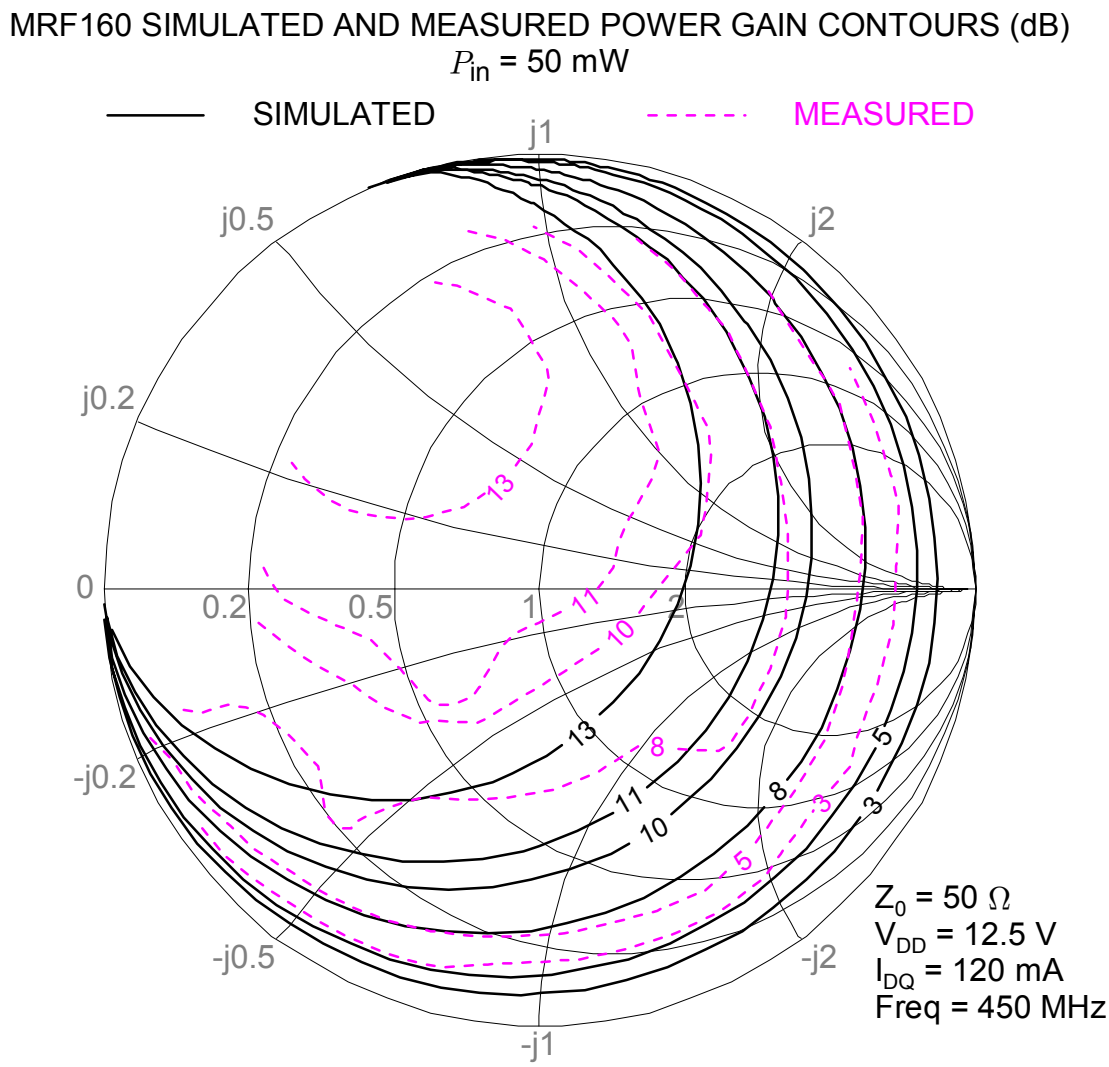


Figure 4.18. MRF160 measured and simulated power gain contours. The effects of heating on the measured contours is evident. The simulation predicts higher gain and the measured and simulated contours show little correlation. Measured gain is lower than what would reasonably be expected from the device data sheet.

of $P_{\text{in}} = 50\,\text{mW}$ the power gain should be at least 17 dB based on the published output power versus input power characteristics in the data sheet. All of the power gain contours in Figs. 4.16, 4.17, and 4.18 are in approximately the same location, even though the input power and Γ_n value increase. This would suggest that the power sweep data does not take into account the rise in temperature that would occur for prolonged operation at elevated input power levels. Efforts are usually made to remove temperature effects from the model parameter extraction process. It would be better, as evidenced by the plots, for

the power sweep measurements to be made under temperature conditions more closely approximating the load-pull measurements, or conversely the load-pull measurements made under temperature controlled conditions similar to the power sweep measurements.

Fig. 4.19 shows a plot of the computed small-signal G_P circles superimposed with the simulated contours for $P_{\text{in}} = 50\,\mu\text{W}$, $5\,\text{mW}$, $50\,\text{mW}$ and the measured contours for $P_{\text{in}} = 50\,\mu\text{W}$. The small-signal G_P contours are slightly larger than the $P_{\text{in}} = 50\,\mu\text{W}$ gain contours. As the input power increases the gain contours shrink in size, though the difference between the size of the simulation contours is not significant except at higher gain levels. When $G_P = 15$, 20 and 22 dB the simulated gain contours associated with

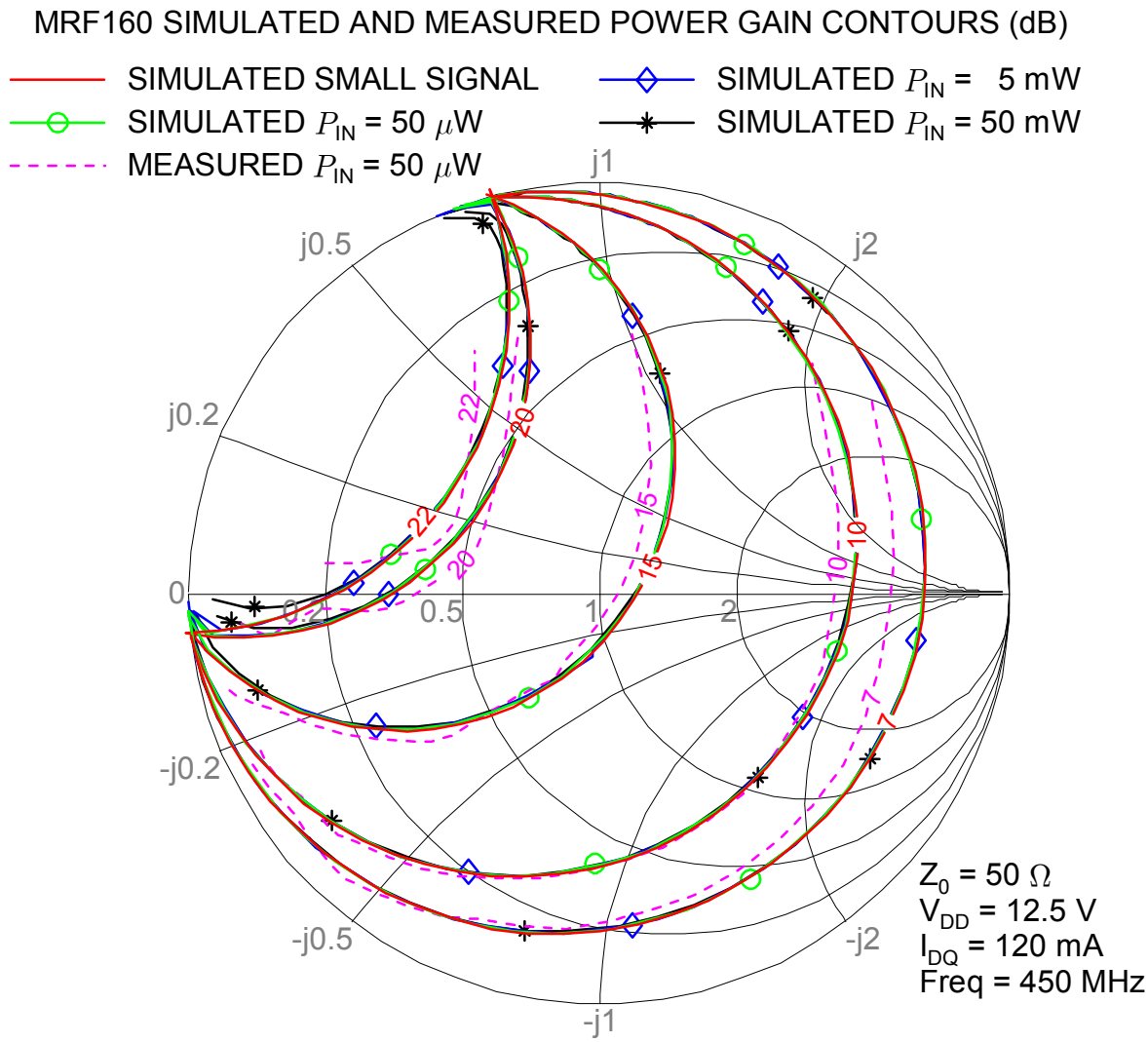


Figure 4.19. Small-signal G_P contours compared with MRF160 measured and simulated load-pull contours. At these drive levels gain contours are dominated by Γ_L, not Γ_n.

$P_{\text{in}} = 5\,\text{mW}$ and $50\,\text{mW}$ begin to diverge from the small-signal simulations near the edge of the Smith chart. For small input drive levels the contours of Fig. 4.19 suggest reasonable agreement between measured and simulated device performance.

Fig. 4.20 shows a close-up of Γ_n from Fig. 4.9 over which the MRF160 load-pull simulations occurred. Values for Γ_n are small, indicating that the gain contours in Fig. 4.19 are not strong functions of Γ_n. Though the curve fit to Γ_n used in the simulation

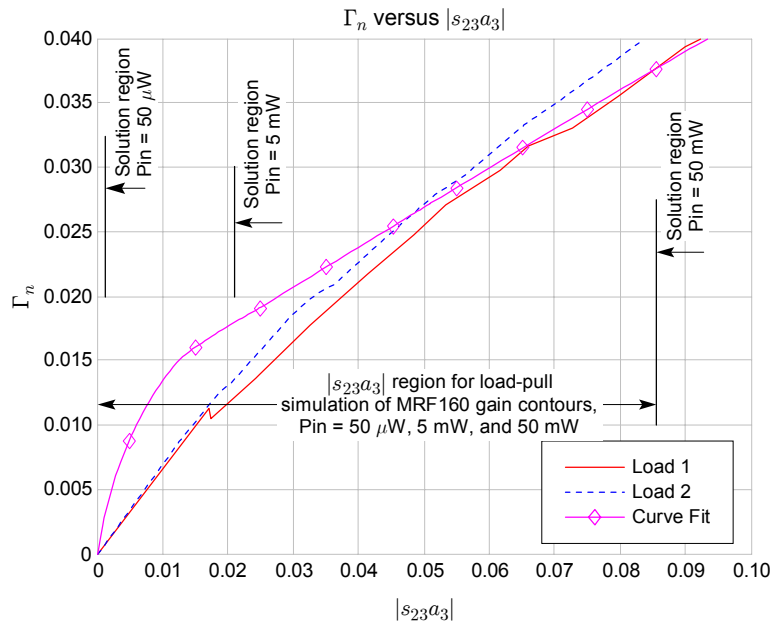


Figure 4.20. Close-up view of Γ_n from Fig. 4.9 showing the $|s_{23}a_3|$ region used in the load-pull simulations for $G_P = 15\,\text{dB}$. The value of Γ_n is small across this region, indicating that for the conditions under which power sweep measurements were taken, gain contours are not a strong function of Γ_n.

is in close agreement over the entire range of $|s_{23}a_3|$, there are local regions where the fit is not as good, especially where $\Gamma_n = 0$ is forced for $|s_{23}a_3| = 0$.

Experiments were run with a higher input power to observe the effect larger Γ_n values had on gain contours. Values for $|s_{23}a_3|$ used in the simulation ranged from 0.0014 to 1.83 which corresponds to Γ_n between approximately 0 and 0.3 as plotted in Fig. 4.9. Gain contours for $P_{\text{in}} = 500\,\text{mW}$ and $50\,\text{mW}$ are shown in Fig. 4.21. The gain contours do collapse as expected with the larger Γ_n values, but they are not dramatically different and do not collapse as quickly as observed in the load-pull measurements.

The fact that the device performance drops off dramatically as seen in Fig. 4.18 further suggests that temperature effects need to be monitored and controlled. The break-apart fixturing used for both the load-pull and power sweep measurements did not have adequate thermal sinking to remove the substantial amount of heat generated during hours of device operation. The poor thermal conductivity of the break-apart fixture has minimal

impact on the short duration power sweep measurements. Through out the power sweep measurements the device is operating close to room temperature. The power sweep data does contain a slight temperature differential between the first load termination and final load termination data, representing a temperature increase over the several minutes it took to measure the various loads. None of the power sweep measurements approached the high operating temperatures seen by the device during the extended operational time

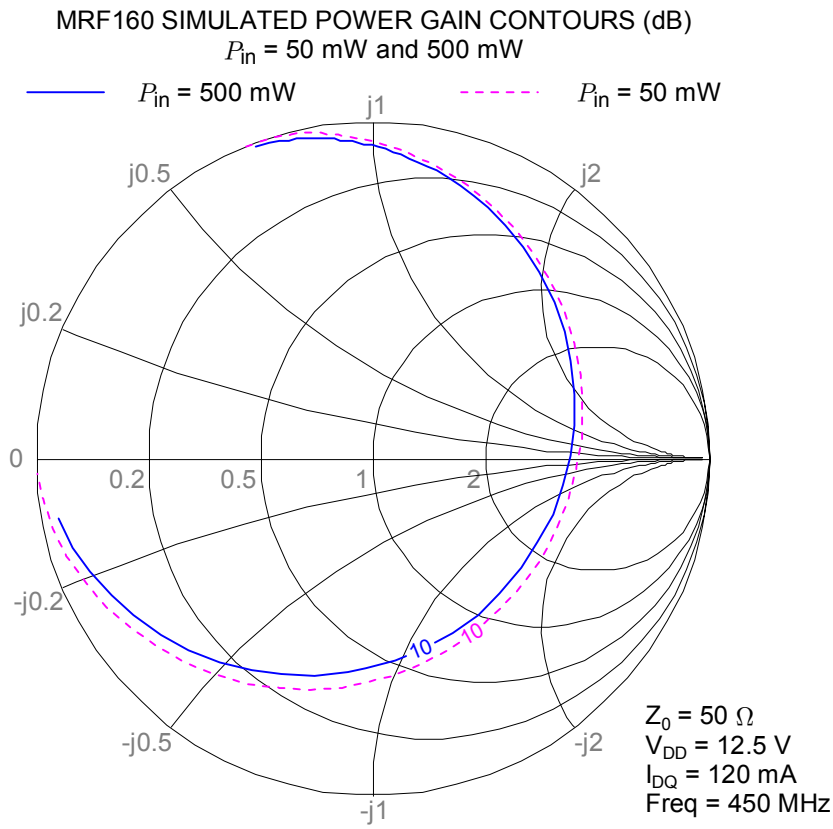


Figure 4.21. MRF160 simulated gain contours with $P_{\text{in}} = 50\,\text{mW}$ and $500\,\text{mW}$. Values for $|s_{23}a_3|$ used in the simulation ranged from 0.0014 to 1.83, allowing larger values of Γ_n to influence the gain contours.

period required to make load-pull measurements. The capability of measuring device temperature was not available when either the power sweep measurements or the load-pull measurements were made. A small fan was used to improve convection cooling and was used primarily during the load-pull measurements. The devices did get hot during the load-pull measurements. At the time the load-pull measurements were taken, it was not known how sensitive the device performance would be with fluctuations in temperature. The power levels used to measure the MRF160 are small enough that all the contours should have been near perfect circles as shown in Fig. 4.19. Had the load-pull measurements been run at constant temperatures the measured G_P contours should have

been in similar locations as predicted by the load-pull simulations because the transistor was operating well below its rated output power of 4.0 Watts.

When simulating the load-pull contours the starting value for Γ_n at which the input power and power gain circles exist does increase with input power as expected. The load-pull simulation also predicts that as input power is increased the area covered on the Smith chart by G_P circles will shrink in size, though with the small input powers used in the simulation this is not a dramatic effect. The measurements do show the trend for shrinking G_P contours as input power increases.

The MRF160 exhibited oscillation in the upper left-hand region of the Smith chart in Figs. 4.16 and 4.17. This region is not completely covered by measured contour data due to this tendency for the device to oscillate when presented with those load terminations. The three-port network could be used to compute the potential instability of the transistor by computing Γ for a given Γ_L. A check could be made for the Γ_L at which $|\Gamma| > 1$ indicating a potentially unstable device for the given input power level.

4.3.3 Simulation of Small- and Large-Signal Class B Device

The possible Γ_L load combinations for obtaining a three-port representation of the MRF166W from the power sweep measurements are listed in Table 4.3. The MRF166W was biased at $V_{DD} = 24.0\,\text{V}$, $I_{DQ} = 240\,\text{mA}$, ($V_{GS} = 2.90\,V$). All measurements are at $f = 450\,\text{MHz}$. This bias condition was chosen to be an "optimally linear" class B bias point. For some Γ_L there were one or two measured points in the power sweep data that were removed from the analysis because of apparent oscillation for a given input power level and load termination. For example, the power sweep data for the HiGAIN and SHORT terminations exhibited some oscillation as the device shifted from the small-signal region of operation into a larger signal regime. For higher power levels the data was

Γ_L sequence	Load 1	Γ_{L1}	Load 2	Γ_{L2}
1	HiGAIN	$0.7664\angle\ 126.08°$	SHORT	$0.7898\angle\ 167.44°$
2	50 Ω	$0.2952\angle\ -136.01°$	OPEN	$0.7610\angle\ -12.49°$
3	3dB+SHORT	$0.5542\angle\ 60.28°$	3dB+OPEN	$0.4898\angle\ -108.94°$
4	HiGAIN	$0.7664\angle\ 126.08°$	3dB+OPEN	$0.4898\angle\ -108.94°$
5	SHORT	$0.7898\angle\ 167.44°$	3dB+SHORT	$0.5542\angle\ 60.28°$
6	HiGAIN	$0.7664\angle\ 126.08°$	3dB+SHORT	$0.5542\angle\ 60.28°$
7	SHORT	$0.7898\angle\ 167.44°$	3dB+OPEN	$0.4898\angle\ -108.94°$
8	SHORT	$0.7898\angle\ 167.44°$	OPEN	$0.7610\angle\ -12.49°$
9	SHORT	$0.7898\angle\ 167.44°$	50 Ω	$0.2952\angle\ -136.01°$
10	HiGAIN	$0.7664\angle\ 126.08°$	OPEN	$0.7610\angle\ -12.49°$
11	HiGAIN	$0.7664\angle\ 126.08°$	50 Ω	$0.2952\angle\ -136.01°$
12	50 Ω	$0.2952\angle\ -136.01°$	3dB+SHORT	$0.5542\angle\ 60.28°$
13	50 Ω	$0.2952\angle\ -136.01°$	3dB+OPEN	$0.4898\angle\ -108.94°$
14	OPEN	$0.7610\angle\ -12.49°$	3dB+SHORT	$0.5542\angle\ 60.28°$
15	OPEN	$0.7610\angle\ -12.49°$	3dB+OPEN	$0.4898\angle\ -108.94°$

Table 4.3. Combinations of Γ_L for MRF166W power sweep measurements.

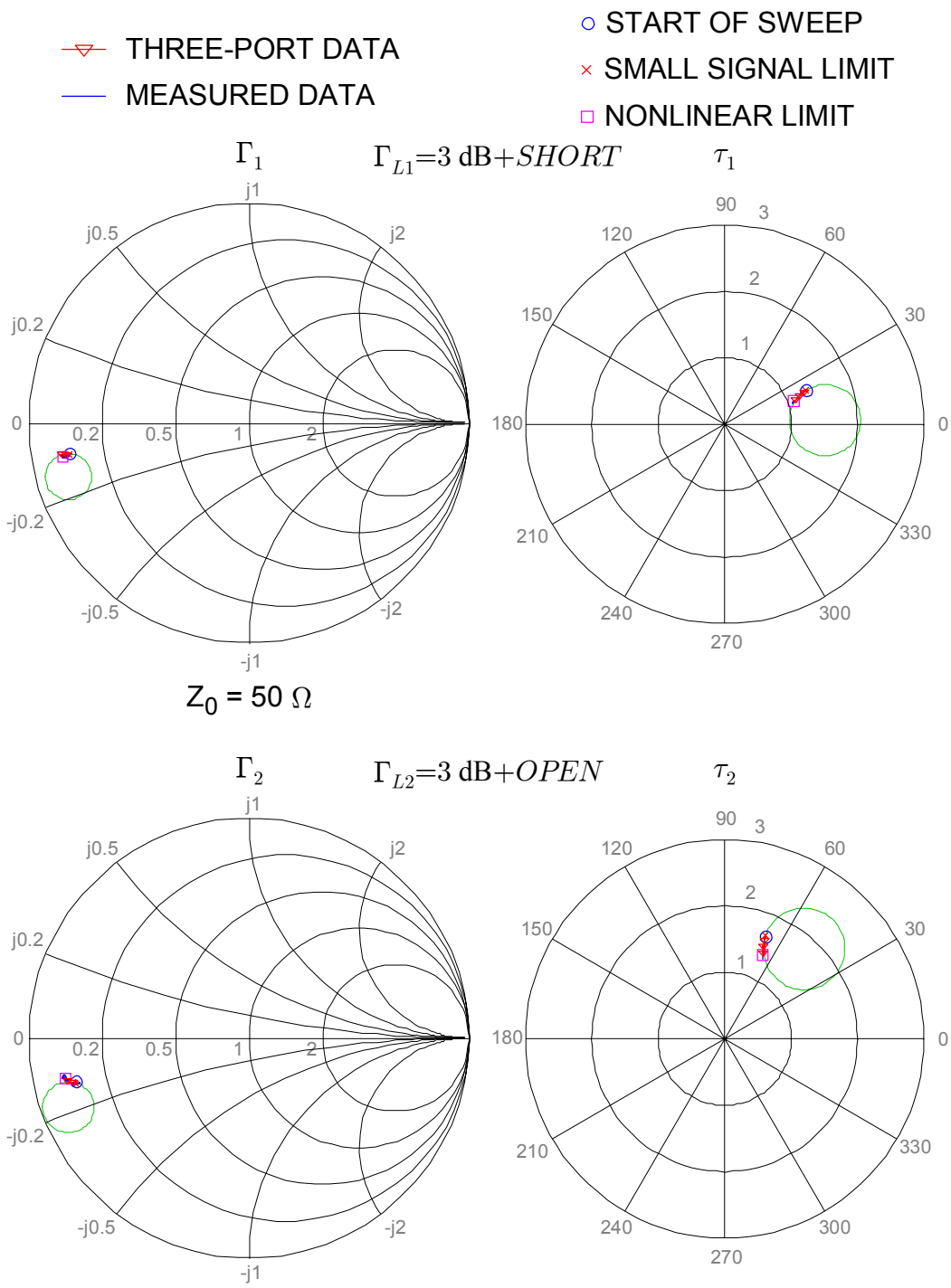


Figure 4.22. Measured and computed power sweep data for MRF166W using Γ_{L1} and Γ_{L2} that resulted in a best-fit match with measured load-pull data.

stable and there was no observable indication of oscillation. The measured data was sufficient to obtain three-port representations for the MRF166W. As with the MRF160 data, the small-signal limit and nonlinear limits were manually chosen to obtain the best fit of the three-port representation to the measured data. A best fit condition was determined by comparing the measured and modeled Γ and τ for Γ_{L1} and Γ_{L2}. A comparison plot of Γ and τ for the measured performance and modeled is shown in Fig. 4.22. The best fit values for the small-signal limit and nonlinear limit used for the model are marked in Fig. 4.22. The simulated Γ and τ were computed using a three-port model extracted with $\Gamma_{L1} = 3\ \text{dB} + SHORT$ and $\Gamma_{L2} = 3\ \text{dB} + OPEN$. Computed τ_1 and τ_2 is in close agreement with the measured power sweep data. Measured Γ_1 and Γ_2 is noisy, but the computed Γ_1 and Γ_2 track favorably with the measured data.

The nonlinear reflection coefficient Γ_n is computed from the measured MRF166W data and plotted as a function of $|s_{23}a_3|$ in Fig. 4.23 for the case $\Gamma_{L1} = 3\ \text{dB} + SHORT$ and $\Gamma_{L2} = 3\ \text{dB} + OPEN$. The Γ_n values associated with each load termination are shown for Γ_{L1} and Γ_{L2} separately in Fig. 4.23. For this data set the Γ_n approximates the ideal case where the values of Γ_n are approximately identical for both load reflections. The residual complex part of Γ_n is approximately zero until $|s_{23}a_3|$ increases beyond the nonlinear limit, at which time the complex portion of Γ_n increases, suggesting that the

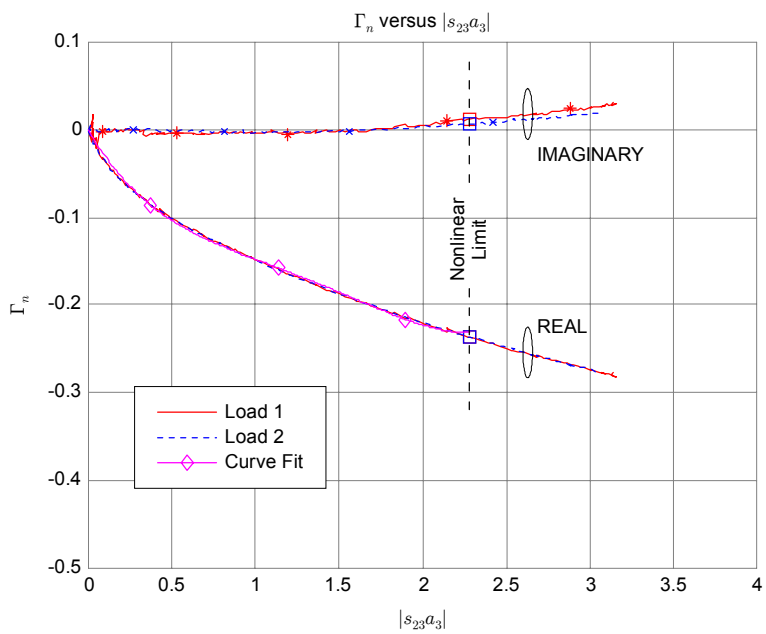


Figure 4.23. Plot of Γ_n real and imaginary residual for MRF166W. The $\Gamma_n^{\Re}$ used for the three-port network is in close agreement for both Γ_{L1} and Γ_{L2}. The residual $\Gamma_n^{\Im}$ indicates a good fit for the three-port network with the measured data.

nonlinear model should fit the measured data closely, which is confirmed by the plot in Fig. 4.22.

The three-port model for the MRF166W extracted using $\Gamma_{L1} = 3$ dB $+ SHORT$ and $\Gamma_{L2} = 3$ dB + OPEN is shown in Fig. 4.24. The value of $s_{33} = -j1$ gives a monotonically decreasing Γ_n as shown in Fig. 4.23.

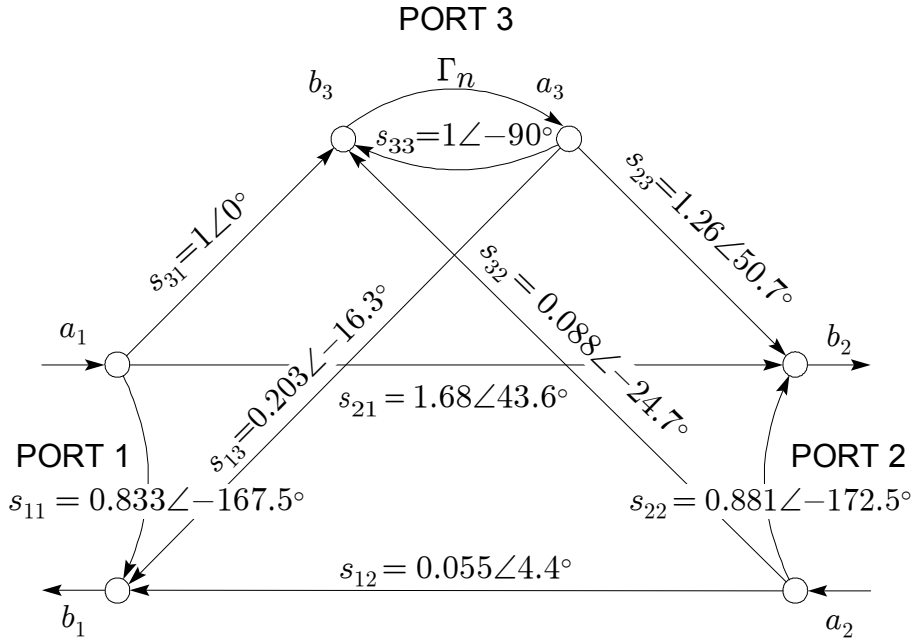


Figure 4.24. Three-port model parameters for the MRF166W at $f = 450$ MHz. Device bias was $V_{DD} = 24.0$ V, $I_{DQ} = 240$ mA, ($V_{GS} = 2.90$ V). The model was extracted using $\Gamma_{L1} = 3$ dB + SHORT and $\Gamma_{L2} = 3$ dB + OPEN. The Γ_n for this model is shown in Fig. 4.23.

Load-pull simulations were run for three input power levels of $P_{\text{in}} = 50\,\mu$W, 50 mW, and 300 mW. The input power levels were chosen to give approximately small-signal linear operation, linear operation outside of small-signal, and linear operation in a large-signal region of operation. The simulation results are compared with the measured load-pull data in Figs. 4.25, 4.26, and 4.27. The MRF166W is a 40 Watt TMOS transistor intended for wide band power amplifier output and driver stages. A single side of the MRF166W should be capable of providing 20 Watts of power at 500 MHz. The bias for both the power sweep measurements and the load-pull measurements was identical and chosen so that the device would be operating in a class B region.

In Fig. 4.25 the simulated load-pull contours optimistically over predict power gain for the 10 and 13 dB contours. The simulation predicts a larger region of 13 dB or greater power gain than the actual load-pull measurements. Heating effects are evident on the 10 and 13 dB load-pull contours. Interestingly, the 5 dB contour is well predicted by the simulation. In Fig. 4.26 the input power is increased to $P_{\text{in}} = 50$ mW and the simulation

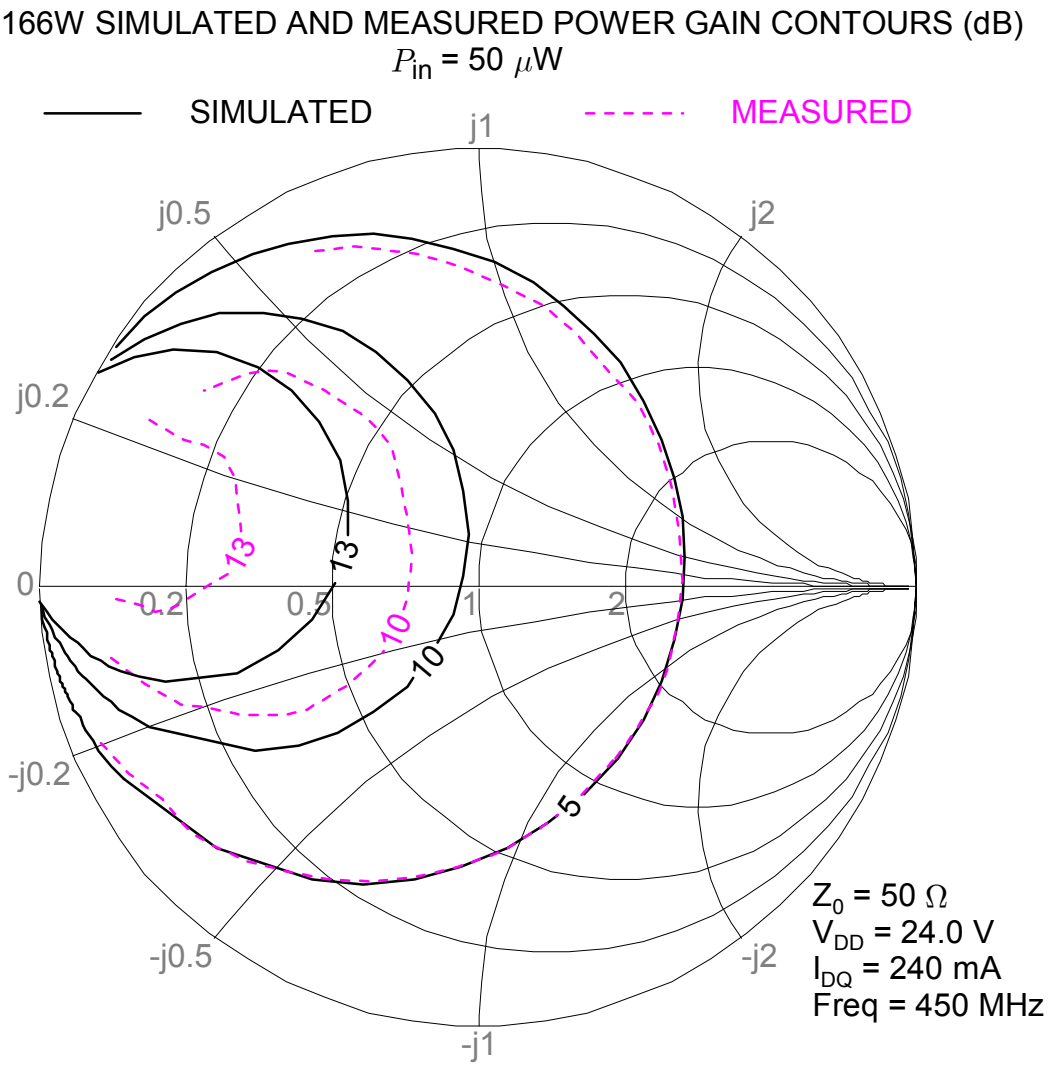


Figure 4.25. MRF166W measured and simulated power gain contours.

gain contours remain in basically the same location as that predicted by the $P_{in} = 50\ \mu$W simulation in Fig. 4.25. For a large device such as the MRF166W, both $P_{in} = 50\ \mu$W and $P_{in} = 50$ mW should be in a linear region. The measured load-pull contours for $P_{in} = 50$ mW are slightly smaller than for the $P_{in} = 50\ \mu$W. The heating effects on the 10 dB contour are more visible for the $P_{in} = 50$ mW case. While it is certain that as input power increases thermal heating should increase, the fact that gain is impacted (as seen from the measured load-pull data) by such a small input (and output) power level may suggest that the thermal sinking of the test fixture for the load-pull measurements was not adequate. In Fig. 4.27 the measured load-pull contours continue to "shrink" in size as $P_{in} = 300$ mW as can be seen by the almost closed contour for $G_P = 13$ dB. The simulation contours have not moved and are similar to the contours for the $P_{in} = 50\ \mu$W

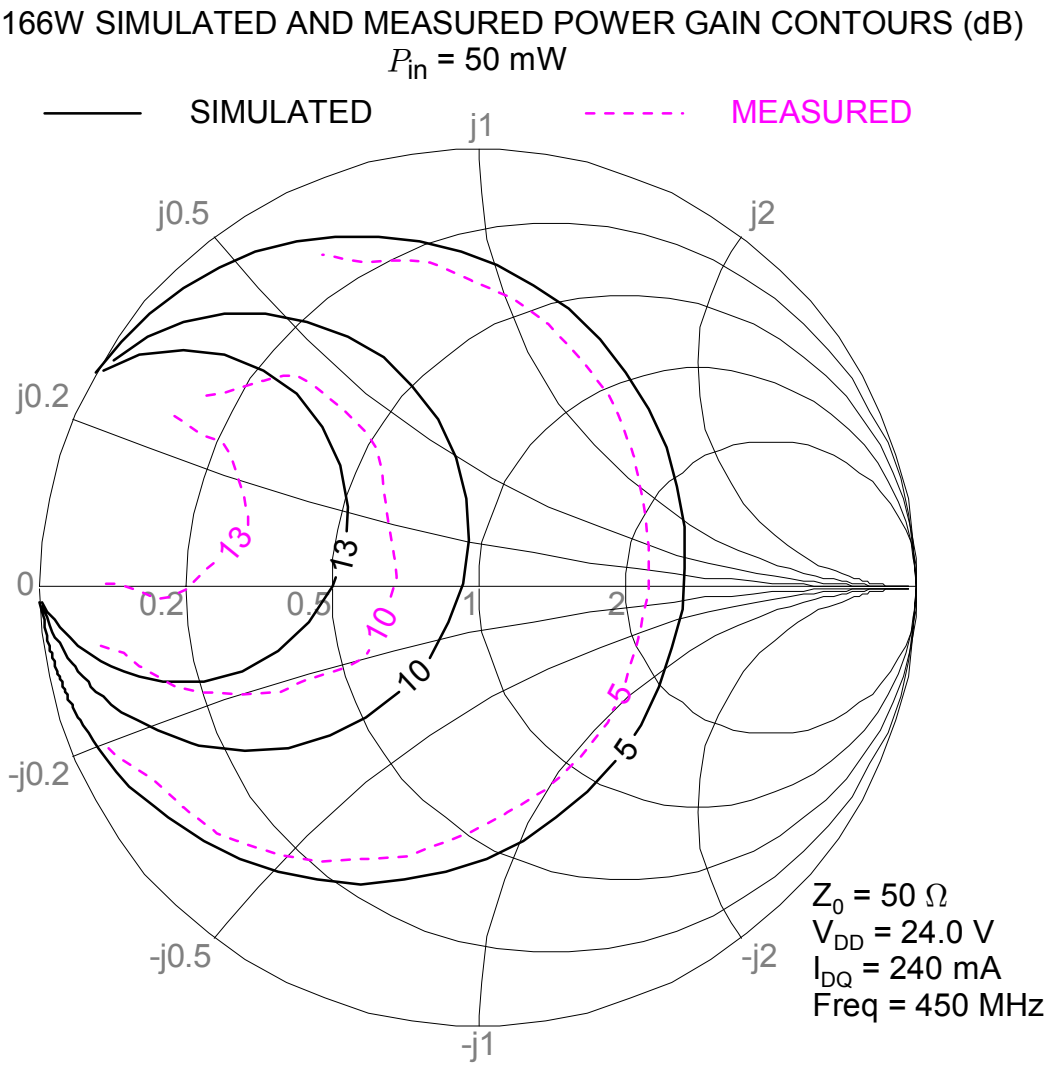


Figure 4.26. MRF166W measured and simulated power gain contours.

and $P_{\text{in}} = 50$ mW test cases. The device operating temperature was not regulated in the break-apart fixture during the two to three hour time period required to complete load-pull measurements. The power sweep data was taken in the same break-apart fixture over a period of several minutes (for all six loads) so that temperature effects, while not controlled, were minimal for the data used in the simulation. The isothermal nature of the power sweep data results in a steady set of contours for all three input power levels. Certainly, if the mechanical structure of the break-apart fixture were improved to allow the device to operate at a fixed temperature regardless of the input or output power, the measured load-pull data should be in closer agreement with the simulated contours based on power sweep measured data where the device ΔT is small and close to the lab temperature (approximately 25 C).

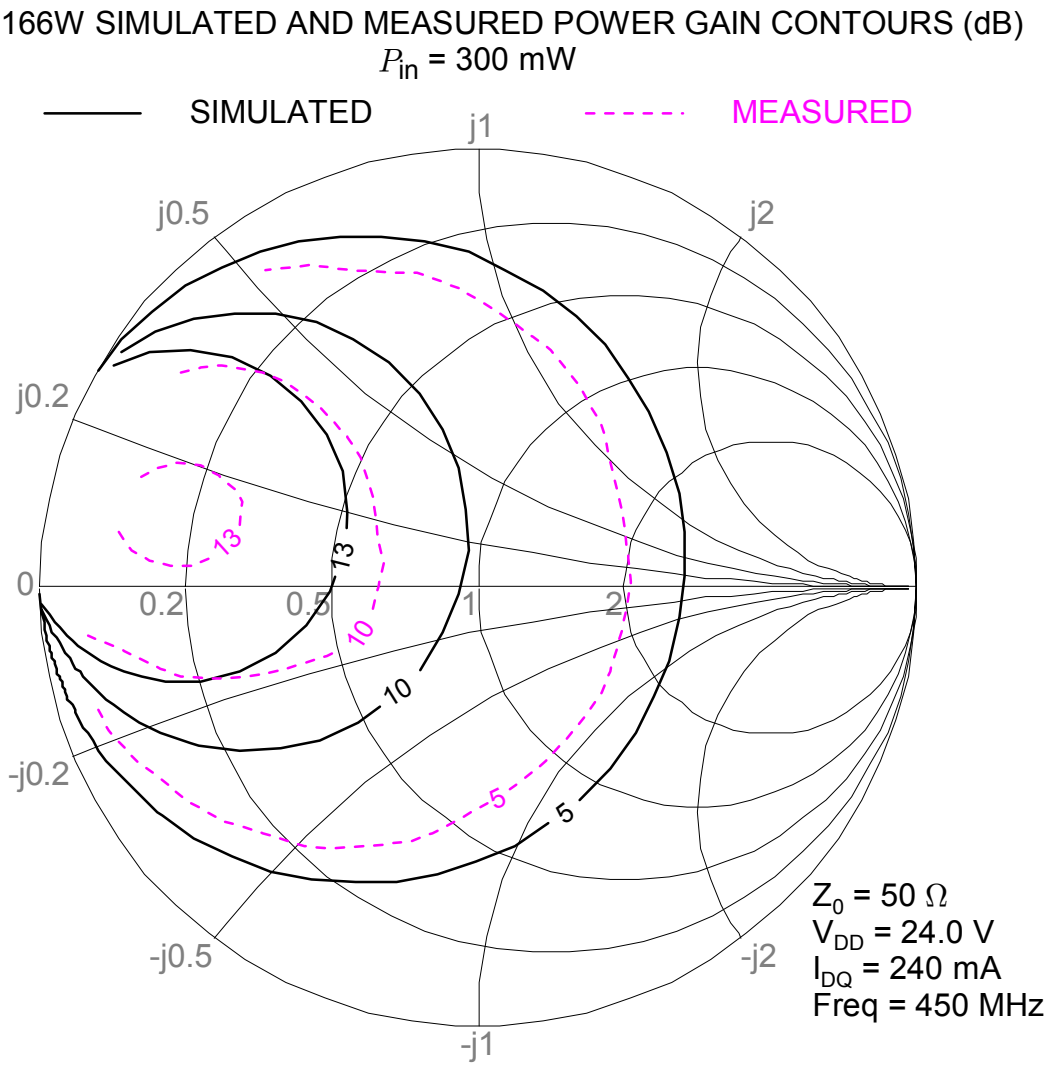


Figure 4.27. MRF166W measured and simulated power gain contours.

The measured power gain contours for three input power levels are plotted in Fig. 4.28 to highlight the changes in the MRF166W performance. The measured gain contours of Fig. 4.28 collapse into evenly spaced concentric contours as input power is stepped from $P_{in} = 50$ μW, 50 mW to 300 mW, an unexpected result. The $P_{in} = 300$ mW contour in Fig. 4.28 should be spaced further from the $P_{in} = 50$ mW than the spacing between the $P_{in} =$ 50 mW and 50 μW contours. Simulated gain contours in Figs. 4.25, 4.26, and 4.27 show no appreciable change in gain as power increases, a result that would suggest the nonlinear three-port model is isothermal. The changes in the gain contours for small drive levels suggests that the device is exhibiting signs of self-heating due to the poor thermal characteristics of the break-apart fixture.

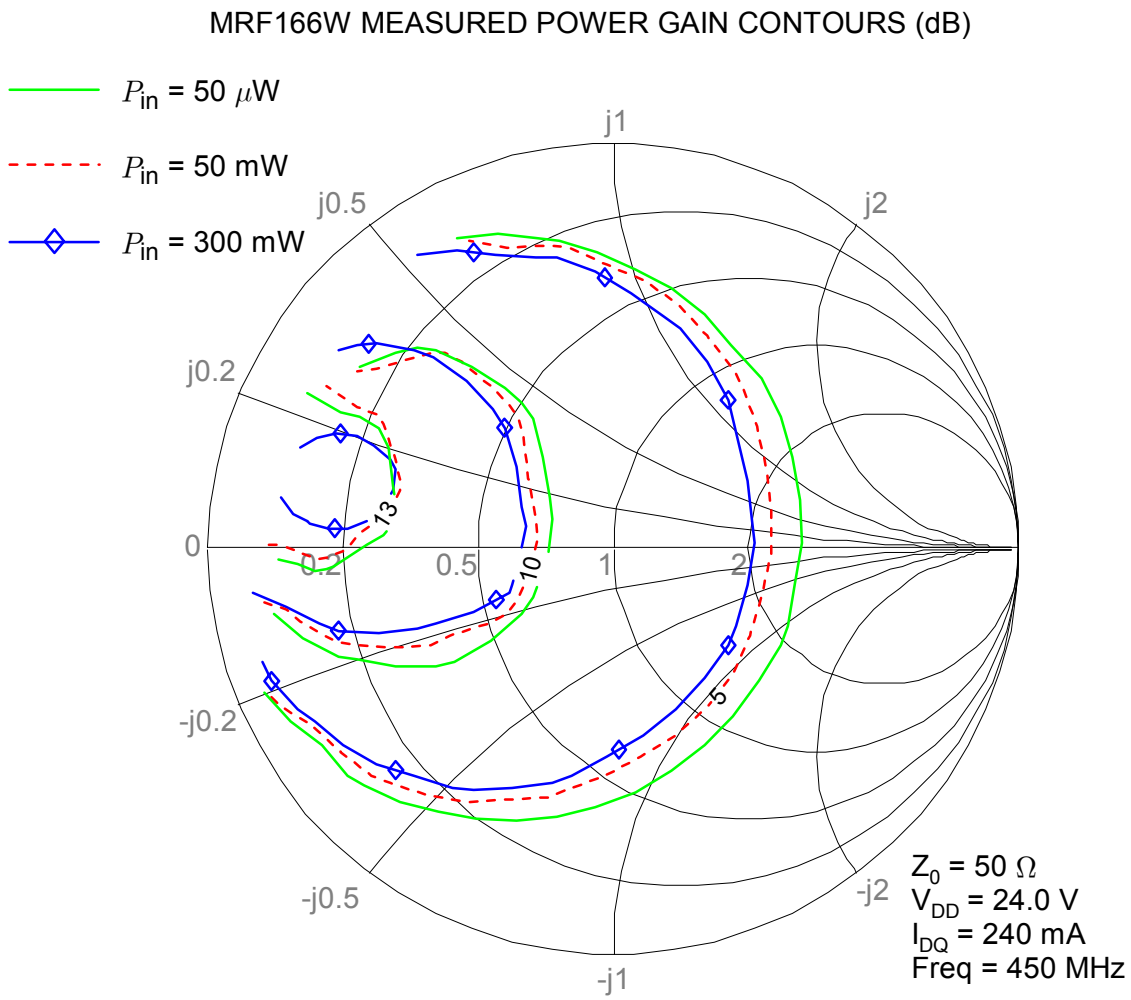


Figure 4.28. MRF166W measured power gain contours illustrating thermal degradation of performance associated with the break-apart fixture.

Based on the class A and class B simulated and measured data presented in this section, for the measured and simulated data to be meaningful to a designer (and for the measured and modeled data sets to be in closer agreement) measurements should be taken under controlled thermal conditions. Preferably the transistor should be run at (or slightly above) the expected operating temperature so that the thermal effects of the device are embedded in the model and in the measured load-pull data. The temperature of the device should be held constant near a steady-state operating temperature for the duration of the measurements. This type of temperature control will require the ability to measure the temperature of the device and a suitable control system to cool or heat the fixture base plate as necessary.

For many device models in popular use it is common to take pulsed measurements in an attempt to remove thermal effects from the measurements. For devices that do not generate a significant amount of thermal heat this approach may be satisfactory. For large-signal devices, such as high-power RF transistors, this approach would neglect a significant contributor to overall device performance. Recently work has been done on electrothermal models that attempt to model the electrical performance of the device and its interrelationship with the thermal performance. In theory, electrothermal models sound attractive, but they introduce an additional element of complexity to the model as challenging and difficult to measure and model as the nonlinear RF performance. From a design perspective, measuring the device at the operating temperature is less complicated and provides ready access to useful results in the shortest possible time frame.

This chapter has presented an approach to measurement based large-signal design using a nonlinear device model. The device model is a linear three-port network with a power dependent nonlinear reflection coefficient on port three. Simulated and measured data was presented for two high-power devices operating in class A and class B. The data suggests the ability of the three-port model to provide useful design information. The measurement of high-power devices presents some unique challenges. Measured data illustrates that care must be taken in holding the device temperature at a constant steady-state operating value for modeling and validation measurements. Thermal effects on device performance can be embedded in the three-port model by maintaining the device at a typical operating temperature during measurement. The three-port nonlinear model provides the designer with an affordable, effective tool for determining device performance and design information.

CHAPTER 5

CONCLUSIONS

Theory and practical considerations have been presented for a previously developed microwave calibration and measurement system capable of measuring scattering parameters for high-power nonlinear devices. Using the RTL measurement system, techniques were presented for obtaining reliable and accurate vector large-signal measurements on active devices. The RTL measurement system makes available a wealth of nonlinear device information that cannot be readily obtained using existing microwave measurement systems. For this thesis, the system was used to obtain S-parameter measurements on commercial power transistors, load-pull power gain contours, and high drive level scattering parameter measurements for use in three-port model extraction.

A device model consisting of a linear three-port network with a power-dependent nonlinear reflection coefficient on port three was applied to high-power transistors. The model parameters were extracted from drive level dependent scattering parameter measurements using an improved extraction method. Close agreement between power sweep measurements and three-port model sweep simulations indicates that the model can represent the nonlinear behavior of a device as a function of drive level and is capable of providing useful design information. A simulation of load-pull measurements was run using the three-port model for two high-power devices. The three-port model simplifies some nonlinear device simulations by encapsulating complex behavior in a straightforward S-parameter network. For the load-pull simulation the intersection of two circles, a drive level-dependent power gain circle and an input drive level circle, are points on the load-pull gain contour. Gain contours were compared with measured gain contours for three different power levels. At low drive levels the simulation and measured contours were in agreement. As the gain circles began to collapse the simulated and measured contours began to diverge, possibly due to device heating. Modeling and validation measurements should be taken with the device held at a constant steady-state operating temperature.

5.1 Areas for Further Research

Work presented in this thesis suggests many avenues for possible further research. Further experimentation on the model under thermal loading would be of interest to further expand the capabilities of the model. Would the three-port model be able to account for thermal effects if model extraction measurements were made at the operating temperature, or would individual model coefficients change with temperature? Thermal effects would only be a concern for modeling high-power devices where self-heating is an issue.

The load-pull simulations could be expanded to include stability contours as a function of drive level. Expressions for gain at constant output power could be derived (instead of constant input power as was done in chapter 4). Model measurements could be expanded to include dc bias current so that the model could be used to compute power-added efficiency or drain efficiency. The model could be adapted to characterize unique amplifier topologies, for example, model a FET where input drive and drain voltage are varied, a design challenge for next generation transmitter circuits.

The three-port model extraction process could be expanded to provide a nonlinear model across frequency. For broad frequency ranges this may take considerable development of the model. The model could be applied to broadband monolithic integrated circuit amplifiers and other nonlinear circuits that are not easily modeled using lumped circuit elements. The model could be used to design a broadband push-pull amplifier with the transistors operating in class AB operation.

Further research could be done on adapting network synthesis and analysis techniques for use with the three-port model. A practical application of the three-port model could be in the synthesis of "designer" impedance matching networks to wrap around high-power transistors to create a drop-in replacement for a competitor's device. The relative ease of extracting a three-port representation of a nonlinear device would lend itself to the task of characterizing multiple devices, comparing the performance as a function of frequency and drive level, and then synthesizing networks to minimize the differences, a task that is unthinkable using lumped-element nonlinear models.

An interesting research area would be to use the three-port model to compare devices, either from different manufacturers, or several different styles from the same manufacturer; for example, multiple prototypes for the same product. Because the model is not technology specific with only a small number of model coefficients it may be a concise tool for comparing and assigning a figure of merit to multiple devices, for instance selecting the most linear device, or the device that would make the best large-signal oscillator, or the best mixer. What would Γ_n look like for a BJT, LDMOS, or GaAs FET in a three-port model extracted across frequency? Are there parameters in the three-port model that capture unique intrinsic characteristics for a given technology platform?

In the future, researchers will return to fundamental network parameters to find ways to concisely express behavior of nonlinear devices. The three-port model is an important first step in the development of generalized nonlinear scattering parameter theory as it provides a way to express S-parameters as a function of drive level while removing the

dependence on a *specific* source and load termination. Expanding the three-port model to include frequency effects would be another step towards generalized nonlinear scattering parameter theory.

REFERENCES

Anderson, E. , Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., & Sorensen, D. (1999). LAPACK users' guide. Philadelphia: Society for Industrial and Applied Mathematics.

Chaffin, R. J. & Leighton, W. H. (1973). Large-signal S-parameter characterization of UHF power transistors. Digest of Technical Papers, 1973 IEEE-MTT International Microwave Symposium, (1973, June 5), pp. 155-7.

Chandra, K. S. & Chen, S. T. (1998). A new measurement technique and characterization tool for push-pull circuit design. 51st ARFTG Conference Digest. (1998, June 12), pp. 113-120.

Chua, L. O. (1978). Nonlinear circuit theory. In G. S. Moschytz & J. Neirynck (Ed.), Modern network theory—an introduction (pp. 65-172). St Saphorin, Switzerland: Georgi.

Cripps, S. C. (1999). RF power amplifiers for wireless communications. Boston: Artech House.

Cusack, J. M., Perlow, S. M., & Perlman, B. S. (1974). Automatic Load Contour Mapping for Microwave Power Transistors. IEEE Transactions on Microwave Theory and Techniques, MTT-22 (12), (1974, December), pp. 1146-52.

Davis, W. A., Keller, D. M., & Ditz, M. W. (1993, January). A new one-way, two-port network analyzer calibration: Reflect-Thru-Line (RTL). Paper presented at the National Radio Science Meeting, Boulder CO.

Davis, W. A. (1999). Three-port modeling for large-signal devices. 53rd ARFTG Conference Digest. (1999, June), pp. 15-19.

Engen, G. F. & Hoer, C. A. (1979). "Thru-Reflect-Line": An improved technique for calibrating the dual six-port automatic network analyzer. IEEE Transactions on Microwave Theory and Techniques, MTT-27 (12), (1979, December), pp. 987-93.

Falconer, K. J. (1971, September). A general purpose algorithm for contouring over scattered data points, National Physical Laboratory (Great Britain), Numerical Analysis & Computer Science, (6).

Franzen, N. R. & Speciale, R. A. (1975). A new procedure for system calibration and error removal in automated S-parameter measurements. Proceedings 5th European Microwave Conference. (1975, September 1-4), pp. 69-73.

Gonzalez, G. (1997). Microwave transistor amplifiers. Upper Saddle River, NJ: Prentice Hall.

Ha, T. T. (1981). Solid-state microwave amplifier design. New York: John Wiley & Sons.

Krauss, H. L., Bostian, C. W., & Raab, F. H. (1980). Solid state radio engineering. New York: John Wiley & Sons.

Kurokawa, K. (1969). An introduction to the theory of microwave circuits. New York: Academic Press.

Kurokawa, K. (1965). Power Waves and the Scattering Matrix. IEEE Transactions on Microwave Theory and Techniques, MTT-13 (3), (1965, March), pp. 194-202.

Leighton, W. H., Chaffin, R. J., & Webb, J. G. (1973). RF amplifier design with large-signal S-parameters. IEEE Transactions on Microwave Theory and Techniques, MTT-21 (12), (1973, December), pp. 809-14.

Maas, S. A. (1997). Nonlinear Microwave Circuits. New York: IEEE Press.

Mazumder, S. R. & Van Der Puije, P. D. (1978). "Two-signal" method of measuring the large-signal S-parameters of transistors. IEEE Transactions on Microwave Theory and Techniques, MTT-26 (6), (1978, June), pp. 417-20.

Motorola Semiconductor Technical Data. (1999) MRF166W The RF MOSFET Line Power Field Effect Transistor N-Channel Enhancement-Mode MOSFET. (3rd Revision) [Data sheet]. Phoenix, AZ.

Müller, O. (1968). Large-signal S-parameter measurements of class C operated transistors. Nachrichtentechnische Zeitschrift, (1968, October), pp. 644-7.

Pedro, J. C. & Carvalho, N. B. (1999). On the use of multitone techniques for assessing RF components' intermodulation distortion. IEEE Transactions on Microwave Theory and Techniques, MTT-47 (12), (1999, December), pp. 2393-402.

Pierpont, M., Pollard, R. D., & Richardson, J. R. (1986). An automated measurement technique for measuring amplifier load-pull and verifying large-signal device models. IEEE MTT-S International Microwave Symposium Digest, (June 1986), pp. 625-8.

Poulin, D. (1980). Load-pull measurements help you meet your match. Microwaves, (1980, November), pp. 61, 63-5.

Raab, F. H., Asbeck, P., Cripps, S., Kenington, P. B., Popović, Z. B., Pothecary, N., Sevic, J. F., & Sokal, N. O. (2002). Power amplifiers and transmitters for RF and microwave. IEEE Transactions on Microwave Theory and Techniques, MTT-50 (3), (2002, March), pp. 814-26.

Sechi, F. N. (1980). Design procedure for high-efficiency linear microwave power amplifiers. IEEE Transactions on Microwave Theory and Techniques, MTT-28 (11), (1980, November), pp. 1157-63.

Sevic, J. F., Burger, K. L., & Steer, M. B. (1998). A novel envelope-termination load-pull method for ACPR optimization of RF/microwave power amplifiers. IEEE MTT-S International Microwave Symposium Digest, Vol. 2, (June 1998), pp. 723-26.

Staudinger, J. (1997a). Specifying power amplifier linearity via intermodulation distortion and channel spectral regrowth. Applied Microwave & Wireless, 9 (4), (1997, July-August), pp. 62, 64, 66, 68, 70.

Staudinger, J. (1997b). The importance of sub-harmonic frequency terminations in modeling spectral regrowth from CW AM-AM & AM-PM derived non-linearities. Proceedings 1997 Wireless Communications Conference, (August 1997), pp. 121-5.

Staudinger, J. & Norris, G. (1997c). The effect of harmonic load terminations on RF power amplifier linearity for sinusoidal and $\pi/4$ DQPSK Stimuli. IEEE MTT-Symposium on Technologies for Wireless Applications, (February 1997), pp. 23-8.

Takayama, Y. (1976). A new load-pull characterization method for microwave power transistors. IEEE MTT-S International Microwave Symposium Digest, (June 1976), pp. 218-20.

Vendelin, G. D., Pavio, A. M., & Rohde, U. L. (1990). Microwave circuit design using linear and nonlinear techniques. New York : John Wiley & Sons.

Webb, J. G. & Chaffin, R. J. (1973). Power transistor amplifier design using large-signal S-parameters. Digest of Technical Papers, 1973 IEEE-MTT International Microwave Symposium, (1973, June 5), pp. 239-41.

APPENDIX A

Measurement System Derivations

A.1 Derivation of Measurement System Error Model

A measurement system designed for making measurements using the "Reflect-Thru-Line" calibration procedure and measurement system as discussed in section 3.1 could be realized using an RF source and four-port directional or hybrid couplers. Typical S-parameter test sets will use directional or hybrid couplers depending on the intended

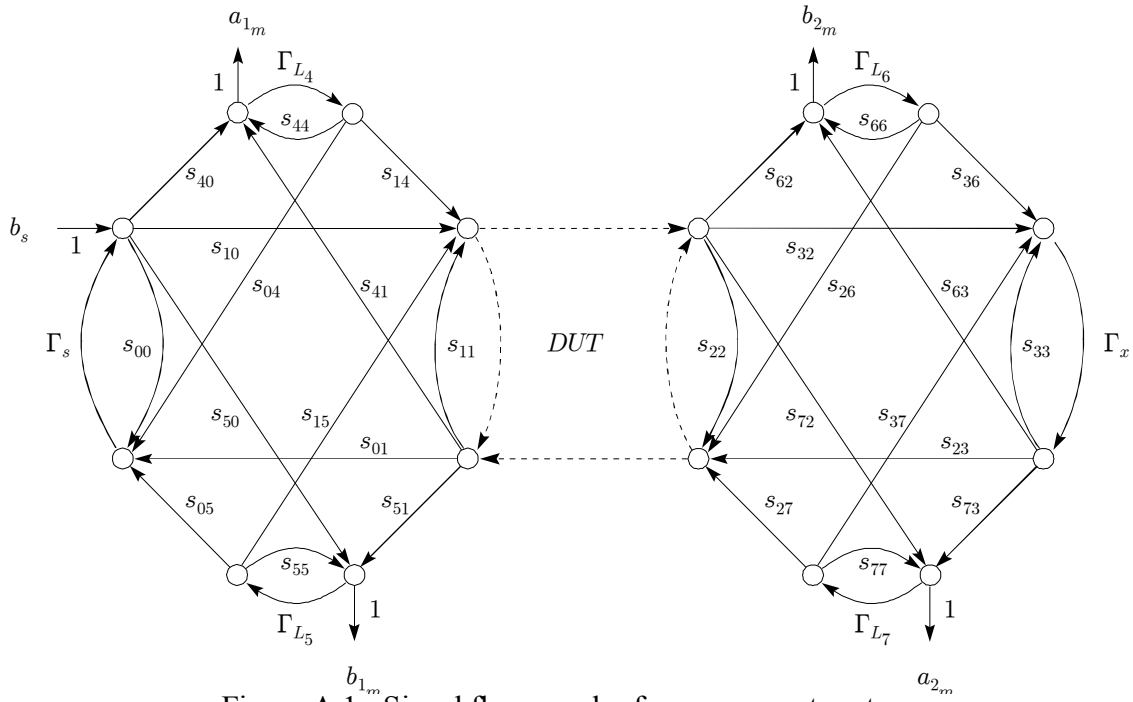


Figure A.1. Signal flow graph of measurement system.

frequency range of operation. The signal flow graph for the system is shown in Fig. A.1.

The directional couplers are shown schematically as four port networks. The device under test is placed at the measurement ports 1 and 2. The RF source b_s has a source reflection coefficient Γ_s. The load Γ_x is at port 3. Magnitude and phase measurements will be made at test ports 4, 5, 6, and 7. Isolation from ports 0 to ports 6 and 7 is not shown in Fig. A.1 for clarity. Isolation will be considered at a later stage.

The object of calibration is to make enough measurements of the system and calibration standards to determine the signals at ports 1 and 2 from the measurement data available at the test ports labeled a_{1_m}, a_{2_m}, b_{1_m}, b_{2_m}. Once the system is characterized the device under test (DUT) is placed between ports 1 and 2 for measurement. Load Γ_x can be selected to obtain the desired device loading. As shown in Fig. A.1 there are 24 unknown signal paths that would need to be determined before measurements at the test ports could be used to compute actual values at the measurement ports. Considering isolation terms would add an additional two unknowns for a total of 26 unknowns. One possible method of system calibration would be to devise 26 independent measurements that would completely determine the system. In practical measurement systems calibration standards and a system error model are used to reduce the number of necessary measurements. The following derivation for a system error model is from W. A. Davis (personal communication, October 1997).

The test port terminations Γ_{L_4}, Γ_{L_5}, Γ_{L_6}, and $\Gamma_{L_7}$are the reflection coefficients presented by the measurement instrument. The measurement instrument measures the complex voltage at each test port relative to a reference signal. The reference signal for this system will be taken at port 4 and is denoted by a_{1_m}. A well designed measurement instrument will have a small reflection coefficient due to insertion of the sensor probe, i.e. $\Gamma_{L_i} \simeq 0$, so that the measurement instrument will have minimal effect on the measurement setup and the signal present on the sensor probe will be optimally coupled into the measurement instrument. The test port reflection coefficients s_{44}, s_{55}, s_{66}, and s_{77} should be approximately zero and their effect on the measurement accuracy can be neglected because their contribution to measurement error will be proportional to Γ_{L_i}. The signal transmitted from the test ports back into the system along paths s_{i4}, s_{i5}, s_{i6}, and s_{i7} is also proportional to Γ_{L_i}. Consider the error signal that would be coupled from test port 4 on to measurement port 1 given by

$$b_1 = a_{1_m} \frac{s_{14}\Gamma_{L_4}}{1 - s_{44}\Gamma_{L_4}} \tag{A.1a}$$

The coupling from the test port back to the input or output port of a four port coupler will be minimal so that the effect on the system due to the terms s_{i4}, $i = 0, 1$ at test port 4 should be negligible when in series with Γ_{L_4}. Generalizing (A.1a) for the other test ports results in

$$b_i = a_{k_m} \frac{s_{ij}\Gamma_{L_j}}{1 - s_{jj}\Gamma_{L_j}} \tag{A.1b}$$
$$i = 0,\ 1,\ 2,\ 3$$
$$j = 4,\ 5,\ 6,\ 7$$
$$k = 1,\ 2$$

where i denotes the input or output port of the directional coupler, k is the input (reflection) or output (transmission) measurement referred to the DUT measurement ports, and j is the test port. The measurement system model that results from the above observations is illustrated in Fig. A.2. Typically the isolation between the through port

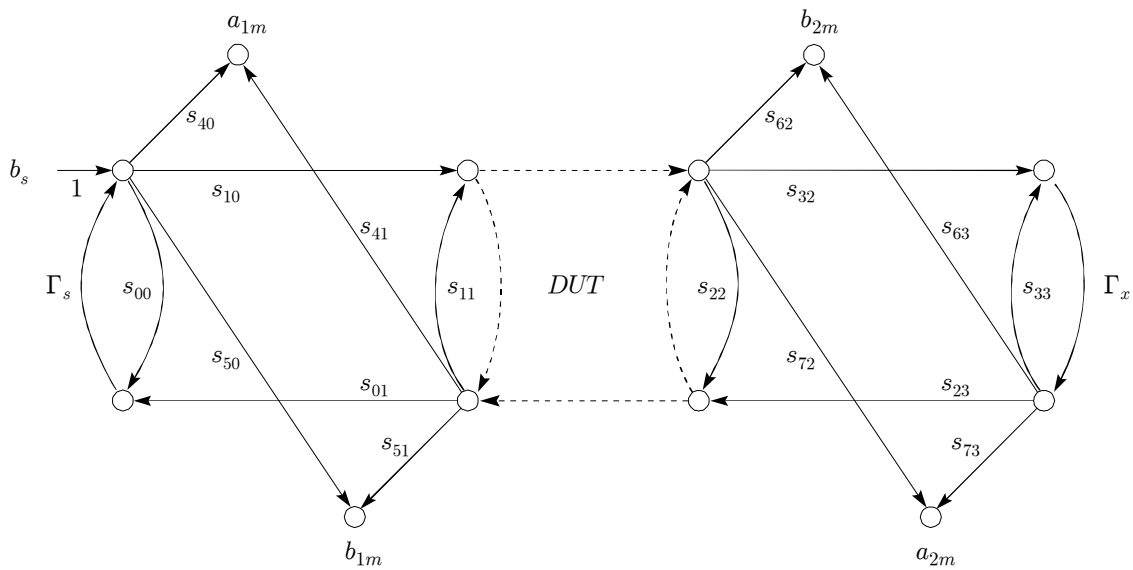


Figure A.2. Simplification of measurement system flow graph.

and coupled port on a directional coupler is quite high so that the cross terms s_{41}, s_{50}, s_{72}, and s_{63} are small values. The error introduced to a_{1_m} due to s_{41} would be maximized if a short or open were at port 1. Looking at the error on a_{1_m} due to s_{41} yields

$$a_{1_m err} = \pm\, b_s \frac{s_{10}s_{41}}{1 \mp s_{11}} \approx \pm\, b_s s_{10} s_{41} \tag{A.2}$$

It is assumed (though not necessary) that $s_{11} \simeq 0$ and (A.2) represents the worst case error. The value of $s_{41} \ll s_{10}$ and on the input port side of the measurement system the cross coupling due to s_{41} will be neglected. Cross coupling terms s_{50}, s_{72}, and s_{63} will be on the same order of magnitude as s_{41} but their contribution to error will not be neglected because the signals levels at ports 0, 2, and 3 are potentially much higher than at port 1.

Two additional error terms need to be added to the measurement system model to account for imperfect isolation between port 0 on the input side and ports 6 and 7 on the output side. These additional terms are shown in Fig. A.3 along with the further

simplification of the measurement system by removal of s_{41}. The source reflection Γ_s is also removed from Fig. A.3 because the effects on the measurement system should be minimal if the source is properly matched such that $\Gamma_s \simeq 0$. Any coupling back to the input through s_{00}, s_{01}, and Γ_s is folded into the input signal at port 0 and is included in the

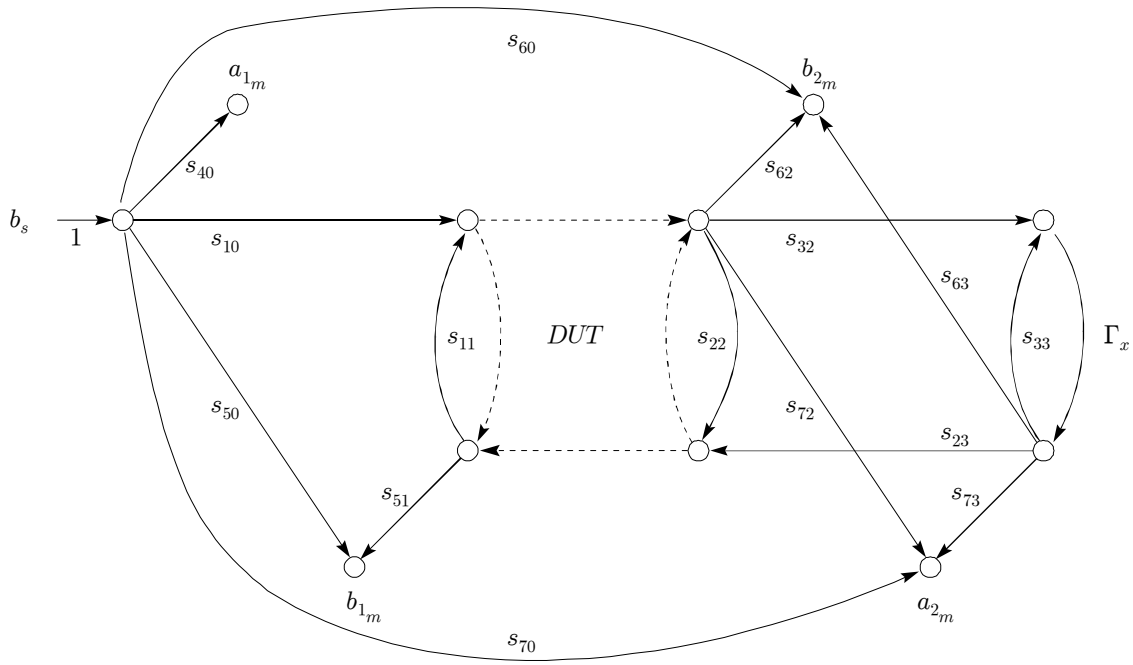


Figure A.3. Isolation terms added to measurement system model.

measurement of a_{1_m}. As shown in Fig. A.3 the measurement system model accounts for the significant sources of systematic measurement error. For measurement purposes it is not necessary to know the individual scattering parameters for the coupler on port 2. It is important to know the transmission from port 2 to ports 6 and 7. The device sees a load Γ_L at port 2 which consists of Γ_x in combination with the other terms in the signal flow graph. The load is given by

$$\Gamma_L = s_{22} + \frac{s_{23}s_{32}\Gamma_x}{1 - s_{33}\Gamma_x} \tag{A.3}$$

To obtain a more direct representation of the measurement system and reduce the number of measurements necessary to characterize the measurement system the four port network at port 2 can be further reduced to obtain a direct representation of the measurement setup.

At port 6 the measured signal is given by

$$b_{2_m} = b_s s_{60} + b_2\left(s_{62} + \frac{s_{32}s_{63}\Gamma_x}{1 - s_{33}\Gamma_x}\right) \tag{A.4a}$$

Substituting (A.3) into (A.4a)

$$b_{2_m} = b_s s_{60} + b_2 \left[\left(s_{62} - s_{22} \frac{s_{63}}{s_{23}} \right) + \frac{s_{63}}{s_{23}} \Gamma_L \right] \tag{A.4b}$$

The above equation defines three sources for the signal at port 6 and the signal flow graph of Fig. A.3 can be simplified to give

$$b_{2_m} = b_s e_{60} + b_2 (e_{62} + e_{63} \Gamma_L) \tag{A.4c}$$

where the scattering parameters of the system have been relabeled as error terms with

$$\begin{aligned} e_{60} &= s_{60} \\ e_{62} &= s_{62} - s_{22} \frac{s_{63}}{s_{23}} = s_{62} - s_{22} e_{63} \\ e_{63} &= \frac{s_{63}}{s_{23}} \end{aligned}$$

The measured value at port 7 is given by

$$a_{2_m} = b_s s_{70} + b_2 \left(s_{72} + \frac{s_{32} s_{73} \Gamma_x}{1 - s_{33} \Gamma_x} \right) \tag{A.5a}$$

Substituting (A.3) into (A.5a)

$$a_{2_m} = b_s s_{70} + b_2 \left[\left(s_{72} - s_{22} \frac{s_{73}}{s_{32}} \right) + \frac{s_{73}}{s_{32}} \Gamma_L \right] \tag{A.5b}$$

The above equation defines three sources for the signal at port 7 and the measured value can be simplified to give

$$a_{2_m} = b_s e_{70} + b_2 (e_{72} + e_{73} \Gamma_L) \tag{A.5c}$$

with the error terms defined by the system scattering parameters

$$\begin{aligned} e_{70} &= s_{70} \\ e_{72} &= s_{72} - s_{22} \frac{s_{73}}{s_{32}} \\ e_{73} &= \frac{s_{73}}{s_{32}} \end{aligned}$$

The simplified signal flow graph with error terms is shown in Fig. A.4. The scattering parameters of the system as shown in Fig. A.3 have been relabeled as error terms. The additional error terms shown in Fig. A.4 are

$$e_{10} = s_{10}$$
$$e_{11} = s_{11}$$
$$e_{40} = s_{40}$$
$$e_{50} = s_{50}$$
$$e_{51} = s_{51}$$

To calibrate the measurement system it is sufficient to determine values for the error terms as shown in Fig. A.4. The number of unknowns in the system has been reduced from

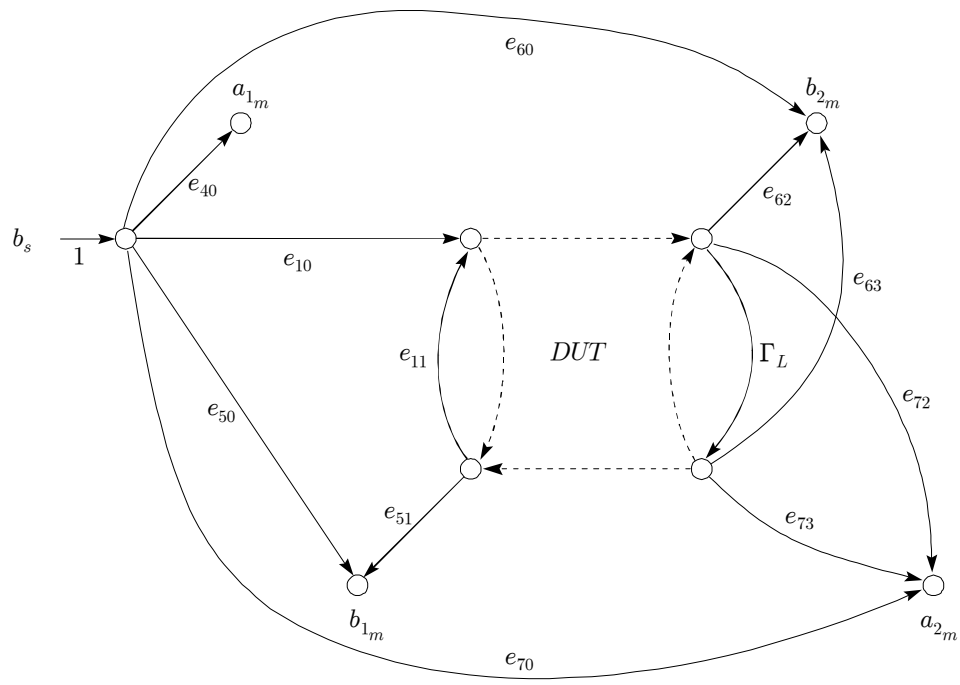


Figure A.4. Signal flow graph showing error terms for measurement system model.

twenty-six to twelve. By selecting a set of calibration standards and making appropriate measurements the unknown error terms can be determined.

A.2 Derivation of Reflect-Thru-Line Calibration

The measurement system is calibrated using a reflect, thru, and line standard along with three independent (unknown) terminations. The standards are placed between ports 1 and 2 with the independent terminations at port 3. The reflection standard has a known reflection coefficient, it is typically a short

$$\Gamma_S = -1 \tag{A.6a}$$

The reflection standard is the phase reference. The thru standard has a scattering matrix of

$$S_{thru} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{A.6b}$$

The line standard has a scattering matrix

$$S_{line} = \begin{bmatrix} 0 & \tau \\ \tau & 0 \end{bmatrix} \tag{A.6c}$$

where τ is not known a priori. The impedance of the line standard sets the reference impedance of the measurement system. The unknown line transmission τ is determined as a by-product of the calibration procedure. The three independent loads are assumed to be unknown and are labeled Γ_1, Γ_2, and Γ_3. Typically a short, open, and 50 Ω load are used. For fixturing that incorporates impedance transforming networks an open and $50\,\Omega$ load may not be independent at the measurement reference plane. The requirement that the independent terminations need not be known offers flexibility to use arbitrary networks or offset reflections as terminations for the calibration. The system calibration proceeds in three stages, 1) error terms associated with port 1 are determined; 2) error terms associated with port 2 and isolation are determined; and 3) the system is power calibrated.

Reflect Calibration (RL^2)

The reflection standard is placed at the reference plane and measured

$$\Gamma_R = \frac{b_{1m}}{a_{1m}} = \frac{1}{e_{40}}\left[e_{50} + \frac{e_{10}e_{51}\Gamma_S}{1 - e_{11}\Gamma_S}\right] \tag{A.7}$$

The thru standard is placed between ports 1 and 2 and reflection measurements are made for each load Γ_i, $i = 1, 2, 3$. The measured result is

$$\Gamma_{Ti} = \frac{b_{1m}}{a_{1m}} = \frac{1}{e_{40}}\left[e_{50} + \frac{e_{10}e_{51}\Gamma_i}{1 - e_{11}\Gamma_i}\right] \tag{A.8}$$

The line standard is placed between ports 1 and 2. The measured reflection for each of the three loads is similar to (A.8) but includes the unknown line transmission τ

$$\Gamma_{Li} = \frac{b_{1m}}{a_{1m}} = \frac{1}{e_{40}}\left[e_{50} + \frac{e_{10}e_{51}\tau^2\Gamma_i}{1 - e_{11}\tau^2\Gamma_i}\right] \tag{A.9}$$

Eq. (A.7) can be solved to obtain the shorthand expression Δ_e

$$\Delta_e = e_{11}e_{50} - e_{10}e_{51} = \frac{1}{\Gamma_S}[e_{40}(1 - e_{11}\Gamma_S)\Gamma_R - e_{50}] \tag{A.10}$$

Solving (A.8) in terms of the unknown load Γ_i

$$\Gamma_i = \frac{e_{40}\Gamma_{Ti} - e_{50}}{e_{11}e_{40}\Gamma_{Ti} - \Delta_e} \tag{A.11}$$

Similarly (A.9) can also be solved for the unknown load Γ_i. The two expressions for Γ_i yield an expression for the unknown transmission τ^2

$$\tau^2\big|_i = \frac{(e_{40}\Gamma_{Li} - e_{50})(e_{11}e_{40}\Gamma_{Ti} - \Delta_e)}{(e_{40}\Gamma_{Ti} - e_{50})(e_{11}e_{40}\Gamma_{Li} - \Delta_e)} \tag{A.12}$$

If independent load terminations are used with the thru and line standards, (A.12) can be used to eliminate the unknown line transmission τ by relating the measurements for the two loads to each other, i.e.

$$\tau^2\big|_i = \tau^2\big|_j \quad \text{where } i \neq j \tag{A.13a}$$

for independent loads i and j. The resulting expression is

$$\frac{(A - \Gamma_{Li})(\Gamma_{Ti} - B)}{(A - \Gamma_{Ti})(\Gamma_{Li} - B)} = \frac{(A - \Gamma_{Lj})(\Gamma_{Ti} - B)}{(A - \Gamma_{Tj})(\Gamma_{Lj} - B)} \tag{A.13b}$$

where

$$A = \frac{e_{50}}{e_{40}} \qquad B = \frac{\Delta_e}{e_{11}e_{40}} \tag{A.13c}$$

Expanding (A.13b) and collecting like terms gives the expression

$$\begin{aligned}&[(A+B)(\Gamma_{Lj}\Gamma_{Ti} - \Gamma_{Li}\Gamma_{Tj}) - AB(\Gamma_{Lj} - \Gamma_{Li} + \Gamma_{Ti} - \Gamma_{Tj})\\ &+ \Gamma_{Li}\Gamma_{Tj}(\Gamma_{Lj} + \Gamma_{Ti}) - \Gamma_{Lj}\Gamma_{Ti}(\Gamma_{Li} + \Gamma_{Tj})](A - B) = 0\end{aligned} \tag{A.14}$$

The trivial solution $A - B = 0$, which simplifies to

$$e_{10}e_{51} = 0$$

would not be consistent with a useful measurement system as can be seen from Fig. A.4 and will not be considered. By measuring with an additional independent load k and then selecting either loads j, k or i, k to complement loads i and j, the two unknowns A and B can be determined. The additional load results in two sets of eqs. (A.14) that are equal to each other and can be solved for the unknowns. The solution procedure will be briefly described. Choose loads i, j and j, k and substitute the measured values into (A.14). The resulting equations can be rewritten as

$$\begin{aligned}(A+B)c_1 - ABc_2 + c_3 = 0\\ (A+B)c_4 - ABc_5 + c_6 = 0\end{aligned} \tag{A.15a}$$

where

$$
\begin{aligned}
c_1 &= \Gamma_{Lj}\Gamma_{Ti} - \Gamma_{Li}\Gamma_{Tj} \\
c_4 &= \Gamma_{Lj}\Gamma_{Tk} - \Gamma_{Lk}\Gamma_{Tj} \\
c_2 &= \Gamma_{Lj} - \Gamma_{Li} + \Gamma_{Ti} - \Gamma_{Tj} \\
c_5 &= \Gamma_{Lj} - \Gamma_{Lk} + \Gamma_{Tk} - \Gamma_{Tj} \\
c_3 &= \Gamma_{Li}\Gamma_{Tj}(\Gamma_{Lj} + \Gamma_{Ti}) - \Gamma_{Lj}\Gamma_{Ti}(\Gamma_{Li} + \Gamma_{Tj}) \\
c_6 &= \Gamma_{Lk}\Gamma_{Tj}(\Gamma_{Lj} + \Gamma_{Tk}) - \Gamma_{Lj}\Gamma_{Tk}(\Gamma_{Lk} + \Gamma_{Tj})
\end{aligned}
\tag{A.15b}
$$

Solving for A and B gives the quadratic expressions

$$
\begin{aligned}
A^2(c_1c_5 - c_2c_4) - A(c_2c_6 - c_3c_5) + c_1c_6 - c_3c_4 = 0 \\
B^2(c_1c_5 - c_2c_4) - B(c_2c_6 - c_3c_5) + c_1c_6 - c_3c_4 = 0
\end{aligned}
\tag{A.16}
$$

It should be noted that this is the same quadratic form used in TRL calibration. Substituting in (A.15b) into the above quadratic equations and factoring out like terms gives

$$
\begin{aligned}
(\Gamma_{Tj} - \Gamma_{Lj})\left[a_2A^2 - a_1A + a_0\right] = 0 \\
(\Gamma_{Tj} - \Gamma_{Lj})\left[a_2B^2 - a_1B + a_0\right] = 0
\end{aligned}
\tag{A.16c}
$$

with

$$
\begin{aligned}
a_2 &= \Gamma_{Ti}(\Gamma_{Lj} - \Gamma_{Lk}) + \Gamma_{Tj}(\Gamma_{Lk} - \Gamma_{Li}) + \Gamma_{Tk}(\Gamma_{Li} - \Gamma_{Lj}) \\
a_1 &= \Gamma_{Ti}\Gamma_{Li}(\Gamma_{Lj} - \Gamma_{Tj} - \Gamma_{Lk} + \Gamma_{Tk}) \\
&\quad + \Gamma_{Tj}\Gamma_{Lj}(\Gamma_{Lk} - \Gamma_{Tk} - \Gamma_{Li} + \Gamma_{Ti}) \\
&\quad + \Gamma_{Tk}\Gamma_{Lk}(\Gamma_{Li} - \Gamma_{Ti} - \Gamma_{Lj} + \Gamma_{Tj}) \\
a_0 &= \Gamma_{Ti}\Gamma_{Tj}\Gamma_{Lk}(\Gamma_{Lj} - \Gamma_{Li}) \\
&\quad + \Gamma_{Ti}\Gamma_{Tk}\Gamma_{Lj}(\Gamma_{Li} - \Gamma_{Lk}) \\
&\quad + \Gamma_{Tj}\Gamma_{Tk}\Gamma_{Li}(\Gamma_{Lk} - \Gamma_{Lj})
\end{aligned}
\tag{A.16d}
$$

For a nontrivial solution it is required that $(\Gamma_{Tj} - \Gamma_{Lj}) \neq 0$ which places a constraint on the effective length of the line standard used in the calibration. From (A.6b), (A.6c)

$$
\tau = e^{\gamma\,\ell} \neq \pm 1 \tag{A.17}
$$

where γ is the complex propagation constant and ℓ is the length of the transmission line. The requirement that the line standard provide information not available from measurement of the thru standard forces the transmission line length ℓ to be chosen such that it is not a multiple of $\lambda/2$ for frequencies at which the calibration is being performed. The line standards used for thru-short-delay (TSD) and TRL are also subject to the same length requirement (Engen & Hoer, 1979). As in the Thru-Reflect-Line calibration process, the line determines the reference impedance of the measurement system. The quadratic equations in (A.16c) can now be solved to obtain values for A and B. Solving

for A

$$A = \frac{e_{50}}{e_{40}} = \frac{2a_0}{a_1 \pm \sqrt{a_1^2 - 4a_0a_2}} \tag{A.18}$$

In measurement systems e_{50} is typically the smallest root since the directional coupler is usually designed to have high isolation between port 0 and port 5.

It is possible that e_{11} equals zero so it is best to solve for the inverse of B

$$\frac{1}{B} = \frac{e_{11}e_{40}}{\Delta_e} = \frac{2a_2}{a_1 \pm \sqrt{a_1^2 - 4a_0a_2}} \tag{A.19}$$

with the sign chosen to give the smallest value of $1/B$. With the values of A and B it is now possible to solve for the unknowns. Solving for $e_{11}e_{40}$ gives

$$e_{11}e_{40} = \frac{1}{B}\Delta_e \tag{A.20}$$

the reflection standard measurement (A.7) can be solved for $e_{11}e_{40}$

$$e_{11}e_{40} = \frac{1}{\Gamma_R\Gamma_S}(e_{40}\Gamma_R - e_{50} - \Delta_e\Gamma_S) \tag{A.21}$$

Setting (A.20) equal to (A.21) and solving for Δ_e gives

$$\Delta_e = e_{11}e_{50} - e_{10}e_{51} = \frac{-e_{40}B(\Gamma_R - A)}{\Gamma_S(\Gamma_R - B)} \tag{A.22}$$

The value of e_{11} can now be determined using (A.22) and (A.20)

$$e_{11} = \frac{-(\Gamma_R - A)}{\Gamma_S(\Gamma_R - B)} \tag{A.23}$$

For most measurement systems, the unknown value e_{10} can be set equal to unity, i.e. $e_{10} \doteq 1$, without any loss of accuracy; any loss in the forward path will be accounted for in the other terms. There may be specialized applications where the loss in the forward path of port 1 would lead to values of $e_{10} \ll 1$ and it would not be desirable to compensate by scaling the other error correction terms. In applications where e_{10} cannot be approximated to unity, a precalibration measurement could be made to determine the proper value of e_{10}. An example of a special measurement setup that may require a precalibration measurement to obtain e_{10} would be a measurement system with a 6 dB attenuator in the signal path just before port 1 (but after the directional coupler). From (A.22) the value of e_{51} can be determined by collecting the known terms on the right hand side and substituting in (A.23). Further manipulation of the above equation gives an expression for the unknown e_{51}

$$\frac{e_{10}e_{51}}{e_{40}} = \frac{(B-A)(\Gamma_R - A)}{\Gamma_S(\Gamma_R - B)} \ . \tag{A.24}$$

The attenuation constant e_{40} is not known at this point and will be obtained from the power calibration measurements. The measurement system is now calibrated for reflection measurements at port 1.

Transmission Calibration

The reflection coefficients of the independent loads can be obtained from (A.11)

$$\Gamma_i = \frac{\Gamma_{Ti} - A}{e_{11}(\Gamma_{Ti} - B)} \tag{A.25}$$

The load reflection Γ_i is labeled Γ_L in Fig. A.4 and is the load seen at the reference plane of port 2. This load is the bilinear transform of the termination at port 3 through the cascaded output fixture and directional coupler networks as given by (A.3).

The isolation error terms e_{60}, and e_{70} can be measured directly with no connection between ports 1 and 2. The resulting measured values are

$$\begin{aligned} \frac{e_{60}}{e_{40}} &= \frac{b_{2m}}{a_{1m}} \\ \frac{e_{70}}{e_{40}} &= \frac{a_{2m}}{a_{1m}} \end{aligned} \tag{A.26}$$

The transmission measurements with the thru standard yields equations for the unknown e_{62} and e_{63}

$$\tau_m = \frac{b_{2m}}{a_{1m}} = \frac{1}{e_{40}}\left[e_{60} + e_{10}\frac{e_{62} + e_{63}\Gamma_L}{1 - e_{11}\Gamma_L}\right] \tag{A.27}$$

Transmission measurements are made for two independent loads $\Gamma_L \in \Gamma_i, \Gamma_j$. The load terminations Γ_L are known from (A.25). The two transmission measurements give τ_i and τ_j for each load (not to be confused with the unknown line transmission τ as used in eqs. (A.12) and (A.13a)). Eq. (A.27) can now be solved for the two unknowns e_{62} and e_{63}. The resulting equations are

$$\frac{e_{62}}{e_{40}} = \frac{\Gamma_j(\tau_i - e_{60}/e_{40})(1 - e_{11}\Gamma_i) - \Gamma_i(\tau_j - e_{60}/e_{40})(1 - e_{11}\Gamma_j)}{e_{10}(\Gamma_j - \Gamma_i)} \tag{A.28}$$

$$\frac{e_{63}}{e_{40}} = \frac{(\tau_j - e_{60}/e_{40})(1 - e_{11}\Gamma_j) - (\tau_i - e_{60}/e_{40})(1 - e_{11}\Gamma_i)}{e_{10}(\Gamma_j - \Gamma_i)} \tag{A.29}$$

The source ratio measurement r_m gives the ratio between the source at port 1 and the "source" at port 2 and can be written in terms of the error coefficients as

$$r_m = \frac{a_{2m}}{a_{1m}} = \frac{1}{e_{40}}\left[e_{70} + \frac{e_{10}(e_{72} + e_{73}\Gamma_L)}{1 - e_{11}\Gamma_L}\right] \tag{A.30}$$

which is similar to the expression for the measured transmission as given by (A.27). Two source ratio values r_i and r_j are measured using two independent loads Γ_i and Γ_j. Solving the two equations for the unknown e_{72} and e_{73} produces the equations

$$\frac{e_{72}}{e_{40}} = \frac{\Gamma_j(r_i - e_{70}/e_{40})(1 - e_{11}\Gamma_i) - \Gamma_i(r_j - e_{70}/e_{40})(1 - e_{11}\Gamma_j)}{e_{10}(\Gamma_j - \Gamma_i)} \tag{A.31}$$

$$\frac{e_{73}}{e_{40}} = \frac{(r_j - e_{70}/e_{40})(1 - e_{11}\Gamma_j) - (r_i - e_{70}/e_{40})(1 - e_{11}\Gamma_i)}{e_{10}(\Gamma_j - \Gamma_i)} \tag{A.32}$$

The measured load reflection can be obtained directly from measurements at ports 6 and 7.

$$\Gamma_{Lm} = \frac{a_{2m}}{b_{2m}} = \frac{r_m - e_{70}}{\tau_m - e_{60}} \tag{A.33}$$

Substituting (A.27) and (A.30) into the above equation and solving for Γ_L the load reflection at port 2 yields

$$\Gamma_L = \frac{e_{72} - e_{62}\Gamma_{Lm}}{-e_{73} + e_{63}\Gamma_{Lm}} \tag{A.34}$$

Even though the error terms in the above equation are known at this point only as a ratio with e_{40} the value of e_{40} is not required for the computation of the load reflection. Knowing the load presented to the DUT has some practical value in addition to being used for the power calibration. The RTL calibration and measurement procedure allows for accurate, real-time load measurement at the DUT reference plane.

The error terms for the system model are now known with the exception of e_{40}. For measurements that do not require knowledge of the input power level it would be sufficient to set the attenuation term to unity, i.e. $e_{40} = 1$.

Power Calibration

A power calibration provides calibrated power measurement at the port 1 reference plane by measuring a_{1m}, i.e. measure $e_{40}b_s$ to obtain the DUT drive level given by $e_{10}b_s$. The phase of e_{40} is neglected and the coefficient takes on a purely real value. Prior to making the power calibration measurement, a reflection calibration is performed with the thru standard placed between ports 1 and 2 of the measurement system of Fig. A.3 as shown in Fig. A.5. The reflection calibration is used to determine the power meter measurement path loss due to s_{32} and the power sensor reflection coefficient denoted Γ_P. The power meter reflection coefficient Γ_P takes into account the measurement port impedance being

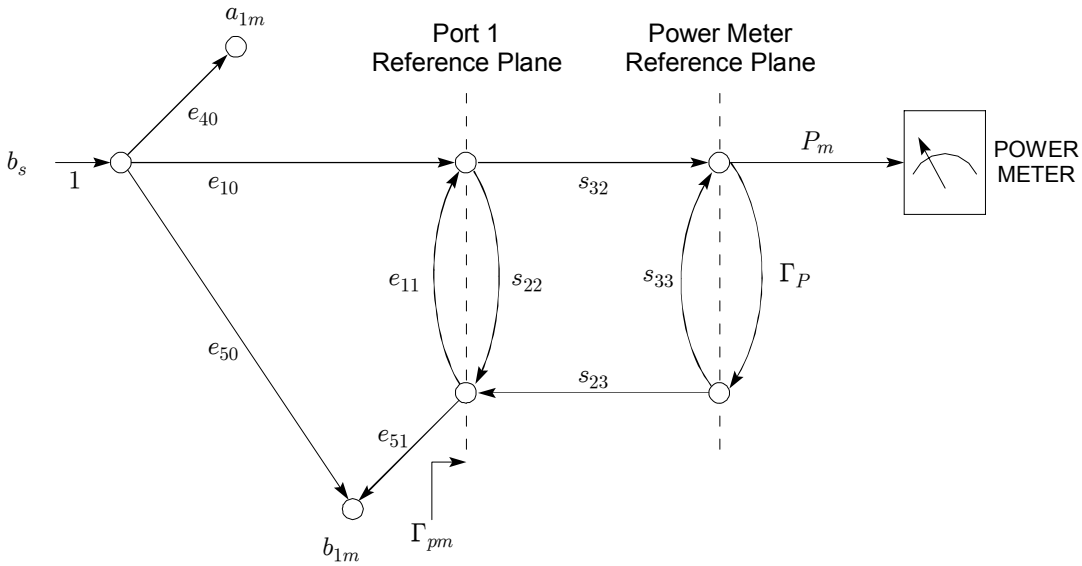


Figure A.5. Setting up measurement system for power calibration measurements. A reflection calibration is used to obtain the unknown s_{32} before measuring the power meter reflection Γ_P and the power P_m at the power meter reference plane. These measurements are sufficient to calibrate the power measurements at port 1 reference plane.

different from the power meter input impedance. The power meter may have an attenuator before the power sensor to allow for high-power drive levels. Any attenuation offset between the power sensor plane and the power meter should be properly accounted for when reading the power meter. The power measured by the power meter at the power meter reference plane is denoted P_m.

The reflection calibration may be performed using an open-short-load (OSL) method or the RL^2 method previously discussed. The reflection calibration is performed by placing the calibration standards at the power meter reference plane and measuring a_{1m} and b_{1m} to determine the unknown values for s_{22}, s_{33}, and the product $s_{23}s_{32}$. Of interest is s_{32}, and we note that the network between reference planes at port 1 and the power meter is reciprocal, hence $s_{23} = s_{32}$, and $s_{32} = \sqrt{s_{23}s_{32}}$ with the phase of s_{32} not required for the power calibration.

The value of Γ_P is measured by placing the power sensor at the power meter reference plane and measuring the reflection Γ_{pm} seen at the port 1 reference plane. From this measurement the power meter reflection coefficient is known

$$\Gamma_P = \frac{s_{33}(s_{22} + \Gamma_{pm}) - s_{23}s_{32}}{s_{22} - \Gamma_{pm}} \ . \tag{A.35}$$

Power incident on port 1 can be obtained from a measurement of a_{1m}

$$P_1 = |a_{1m}|^2 = |b_s e_{40}|^2 \ . \tag{A.36}$$

Power incident at the power meter reference plane is measured at the power meter

$$P_m = \left| b_s \frac{e_{10} s_{32}}{(1 - e_{11}\Gamma_{pm})(1 - s_{33}\Gamma_P)} \right|^2 \ . \tag{A.37}$$

The ratio of P_m and P_1 gives an expression for e_{10}/e_{40}

$$\left| \frac{e_{10}}{e_{40}} \right| = \left| \frac{(1 - e_{11}\Gamma_{pm})(1 - s_{33}\Gamma_P)}{s_{32}} \right| \sqrt{\frac{P_m}{P_1}} \tag{A.38}$$

which can be used with (A.24) to scale the individual error terms to complete the power calibration procedure.

The measurement system architecture and calibration procedure provide a high degree of accuracy in a robust measurement environment. Using RTL calibration provides the reference plane at the edge of the fixture where the transistor is to be mounted with all the fixture and dc bias parasitics embedded in the error correction coefficients.

The measurement system provides scattering parameters Γ, τ, and Γ_L versus drive level which are sufficient to describe *any* two-port network—linear or nonlinear. The system shown in Fig. A.4 also allows for vector measurement of harmonics on both the input and output of the DUT, a capability that may be useful for some applications.

APPENDIX B

Published S-Parameters

The following commercially available devices were measured using the RTL measurement system. The published S-parameter data is available in printed form (Motorola Semiconductor Technical Data, 1999).

Device	Mode of Operation	DC Bias		Frequency (MHz)	
		V_{DD} (V)	I_{DQ} (A)	Start	Stop
MRF140	Class A	24.0	5.0	10	500
	Class A	28.0	5.0	10	500
MRF141	Class A	24.0	5.0	10	500
	Class A	28.0	5.0	10	500
MRF141G	Class B	24.0		10	500
	Class B	28.0		10	500
MRF150	Class A	50.0	2.0	10	500
MRF151	Class A	50.0	2.0	10	500
MRF151G	Class B	50.0		10	500
MRF160	Class A	12.5	0.120	10	500
	Class A	28.0	0.250	10	500
MRF166C	Class A	12.5	1.25	30	1000
	Class A	28.0	1.25	30	1000
MRF166W	Class B	24.0		30	1000
	Class B	28.0		30	1000
MRF171A	Class A	12.5	0.5	10	500
	Class A	28.0	0.5	10	500
MRF173	Class A	12.5	4.0	10	500
	Class A	28.0	4.0	10	500
MRF173CQ	Class A	12.5	4.0	10	500
	Class A	28.0	4.0	10	500
MRF175GV	Class B	28.0		10	500
MRF175L	Class A	28.0	4.5	10	500
MRF176GV	Class B	50.0		10	500
MRF177	Class B	24.0		10	1000
	Class B	28.0		10	1000
MRF275G	Class B	12.5		10	500
	Class B	24.0		10	500
	Class B	28.0		10	500
MRF275L	Class A	12.5	4.5	10	500
	Class A	24.0	4.5	10	500
	Class A	28.0	4.5	10	500

VITA

John Brent Call

13 September 2002

EDUCATION
VIRGINIA POLYTECHNIC INSTITUTE & STATE UNIVERSITY
Candidate for M.S.E.E. degree
Completion date Fall 2002

UNIVERSITY OF UTAH
B.S.E.E. 1996

MILITARY INTELLIGENCE SCHOOL Goodfellow AFB, San Angelo, Texas
Graduated with Honors 1991

DEFENSE LANGUAGE INSTITUTE Monterey, California
Graduated as a Korean Linguist 1990

PROFESSIONAL EXPERIENCE
SILICON WIRELESS CORP. Raleigh/Durham, North Carolina
RF Design Engineer June 2001 - Present
• Characterize and evaluate proprietary silicon SLMOS RF power transistors ranging in output power from 5 to 50 Watts at 1 – 2 GHz. Designed and built circuits for device characterization and customer evaluation.

CREE, INC. Durham, North Carolina
RF Design Engineer January 2001 - May 2001
• Characterize and evaluate 30 Watt and 10 Watt silicon carbide MESFET microwave transistors from 1 – 3 GHz. Designed RF reliability test circuit. Designed calibration kit for S-parameter measurement on 30 Watt device. The commercial silicon carbide microwave group was disbanded and the RF engineers were reassigned to work on LDMOS products. Having joined the company to gain experience with commercial devices *other than* LDMOS I made the decision to pursue opportunities elsewhere.

MOTOROLA Phoenix, Arizona
RF Design Engineer August 1998 - December 2000
• Experience with commercial high-power RF discrete device design, single-ended and push-pull, 30 – 180 Watts from 470 MHz up to 2.1 GHz. Extensive use of load-pull measurement system to characterize and evaluate high-power RF/microwave next generation LDMOS die designs (up to 90 Watts). Designed and built test circuits for

device design and load-pull requirements. Applied advanced microwave calibration and measurement techniques to obtain test fixture S-parameters.

- Participated in an engineer exchange with the design center in Toulouse, France for four weeks in October 2000.

VIRGINIA POLYTECHNIC INSTITUTE & STATE UNIVERSITY
Research Assistant August 1997 - August 1998

- Measured S-parameters on several high-power TMOS devices. S-parameters were published in Motorola device data sheets. Developed novel load-pull measurement system.

MOTOROLA Phoenix, Arizona
Summer Intern May 1997 - August 1997

- RF Power Products Division, conducted thermal and ruggedness measurements to characterize an LDMOS device for an avionics application.

TARGET SYSTEMS, INC. Salt Lake City, Utah
Embedded Systems Programmer June 1995 - July 1996

- Developed 8051 software for a laser-based diameter measurement gage

JTP RADIATION Salt Lake City, Utah
Electronic Technician 1988-1989

- Constructed engineering prototypes of military electronic navigational equipment and accompanying test and measurement devices

TEACHING EXPERIENCE
VIRGINIA POLYTECHNIC INSTITUTE & STATE UNIVERSITY
Graduate Teaching Assistant August 1996 - May 1998

- Laboratory instructor for Radio Engineering and Electronics Lab

UNIVERSITY OF UTAH
Lab Teaching Assistant January 1995 - March 1996

- Laboratory instructor for Electronic Circuits and Engineering Electronics undergraduate lab courses for 2nd, 3rd, and 4th year students

MILITARY SERVICE
UNITED STATES ARMY
Military Intelligence Analyst/Korean Linguist 1989-1992

VOLUNTEER EXPERIENCE
Church of Jesus Christ of Latter-Day Saints New Zealand
Voluntary Missionary 1986-1988

- Supervised teams of up to 30 volunteers
- Worked with homeless youth

- Financial Clerk/Assistant Fleet Manager

PUBLICATIONS

- Call, J. B. & Davis, W. A. (2000). Large-signal scattering parameter measurements for RF power transistors. IEEE Radio and Wireless Conference (RAWCON). (2000, September), pp. 143-6.

ABSTRACTS

- Call, J. B. & Davis, W. A. (1999). Practical two-port scattering parameter measurement technique for large signal applications. United States National Commitee International Union of Radio Science Program and Abstracts. (1999, January), p. 161. Boulder CO.

PROFESSIONAL AND HONORARY SOCIETIES

- HKN Electrical Engineering Honor Society, 1997
- IEEE Member, 1993 - Present

AWARDS

- First place student paper contest, IEEE Northwest Region (University of Utah), 1996
- Army Commendation Medal, United States Army, 1992
- Army Good Conduct Medal, United States Army, 1992
- Gold medal in VICA Electronics competition for state of Idaho, 1985
- Gold medal in VICA Electronics competition for state of Idaho, 1984
- Eagle Scout, Boy Scouts of America, 1982