A Scattering Variable Approach to the Volterra Analysis of Nonlinear Systems

DONALD D. WEINER, MEMBER, IEEE, AND GERALD H. NADITCH, MEMBER, IEEE

Abstract—A new mathematical model is developed which extends Volterra series analysis of nonlinear systems with memory to high-frequency systems, including those containing linear distributed component devices. A generalized set of nonlinear scattering parameters is defined which can be used to describe power transfer and distortion in nonlinear multiports, and which reduce to the classical scattering parameters for linear networks.

The methodology is based on Volterra functional series, and is most useful for the small-signal case where the response can be approximated by a finite number of terms of the series. Nonlinear scattering kernels, derived by extending the Volterra analysis, are simply related to previously developed nonlinear voltage and current Volterra kernels. For sinusoidal inputs nonlinear scattering parameters are defined which are shown to be particularly helpful when power relationships are studied. The principal applications are for microwave networks terminated in real-valued finite reference impedances. To evaluate the average power dissipated in a load at some intermodulation frequency, the concept of nonlinear transducer gain is defined and shown to be proportional to the squared magnitude of a nonlinear scattering parameter. Examples are presented illustrating the analysis procedure for a tunnel diode reflection amplifier and for a linear lossless transmission line terminated by a nonlinear network.

I. Introduction

NONLINEAR considerations often arise in the analysis and design of microwave systems. For example, given two out-of-band interfering signals at the input to a microwave transistor amplifier, it may be desirable to predict the magnitude of an in-band intermodulation component generated by the nonlinearities of the transistor. On the other hand, for a linear transmission line terminated by a square-law diode, the magnitude and phase of the second harmonic in the wave reflected from the load may be of interest. Alternatively, it may be important to evaluate the nonlinear distortion in the output of a tunnel diode amplifier employing a circulator.

In recent years the Volterra functional series [1]–[4] has emerged as a promising tool for the analysis of such problems. The Volterra approach assumes that the response of a nonlinear system, having input $x(t)$ and output $y(t)$, can be expressed as

$$y(t) = \sum_{n=1}^{\infty} y_n(t) \tag{1}$$

Manuscript received July 28, 1975; revised December 16, 1975. This work was supported in part by the U.S. Air Force Systems Command, Rome Air Development Center, Griffiss Air Force Base, NY, as part of the RADC Post-Doctoral Program under Contract F30602-72-C-0360.

D. D. Weiner is with the Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY 13210.

G. H. Naditch is with the Technical Staff, the MITRE Corporation, Bedford, MA.

where $y_n(t)$ is the nth-order portion of the response and is given by

$$y_n(t) = \int h_n(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} x(t-\tau_p)\, d\tau_p. \tag{2}$$

The symbol $\int$ denotes an n-fold integration from $-\infty$ to $+\infty$ while $\prod_{p=1}^{n}$ denotes an n-fold product. $y_n(t)$ is of nth order in the sense that multiplication of the input $x(t)$ by the constant A results in multiplication of $y_n(t)$ by A^n. Systems having Volterra functional representations are referred to as "Volterra systems" [5].

In practice, the Volterra approach is most useful when the response $y(t)$ can be adequately approximated by a finite number of terms. This is the situation commonly encountered in communications circuits such as amplifiers and small-signal mixers where the nonlinear distortions are usually 20 dB or more below the input signals. In this paper it is assumed that all responses are adequately represented by the finite sum

$$y(t) = \sum_{n=1}^{N} \int h_n(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} x(t-\tau_p)\, d\tau_p. \tag{3}$$

The nth-order Volterra kernel $h_n(\tau_1,\cdots,\tau_n)$ is referred to in this paper as the nth-order impulse response. In actuality the impulse response may not be identically zero above order N. However, the finite sum of (3) implies that higher order terms contribute negligibly to the output.

The nth-order nonlinear transfer function is defined to be the n-dimensional Fourier transform of $h_n(\tau_1,\cdots,\tau_n)$. This results in the Fourier transform pair

$$H_n(f_1,\cdots,f_n) = \int h_n(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} e^{-j2\pi f_p \tau_p}\, d\tau_p \tag{4}$$

$$h_n(\tau_1,\cdots,\tau_n) = \int H_n(f_1,\cdots,f_n) \prod_{p=1}^{n} e^{j2\pi f_p \tau_p}\, df_p. \tag{5}$$

Substitution of (5) into (3) yields the input–output relation

$$y(t) = \sum_{n=1}^{N} \int H_n(f_1,\cdots,f_n) \prod_{p=1}^{n} X(f_p) e^{j2\pi f_p t}\, df_p \tag{6}$$

where $X(f)$ is the Fourier transform of the input $x(t)$. Equation (3) suggests the block diagram representation shown in Fig. 1. Equations (3) and (6) indicate that the nonlinear system is completely characterized by either the nonlinear impulse responses or the nonlinear transfer functions. Once these are known, it is possible to determine

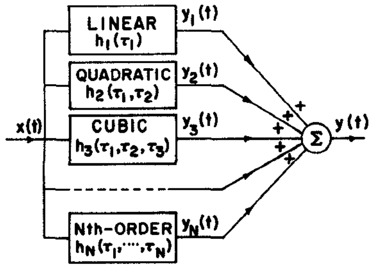

Fig. 1. Characterization of a nonlinear system with memory in terms of its N impulse responses, $h_n(\tau_1, \cdots, \tau_n)$; $n = 1, 2, \cdots, N$.

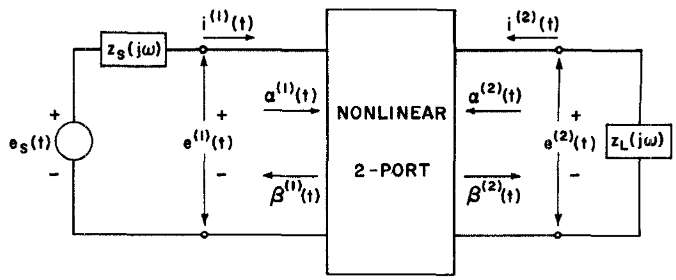

Fig. 2. Loaded nonlinear 2-port with notation and conventions used for port voltages, currents, and scattering variables.

the system response for arbitrary inputs expressed either in the time or frequency domain.

The special case for which the input is a sum of M sinusoidal terms is of particular interest. Assume

$$x(t) = \sum_{m=1}^{M} |E_m| \cos (2\pi f_m t + \theta_m). \tag{7}$$

Define the complex voltage

$$E_m = |E_m| e^{j\theta_m}. \tag{8}$$

Let

$$f_{-m} = -f_m \qquad E_0 = 0 \qquad E_{-m} = E_m^* \tag{9}$$

where the asterisk denotes complex conjugate. It follows that the input can be written as

$$x(t) = \frac{1}{2} \sum_{m=-M}^{M} E_m e^{j2\pi f_m t}. \tag{10}$$

Substitution of (10) into (3) and use of (4) results in the response

$$\begin{aligned} y(t) = {} & \frac{1}{2} \sum_{m=-M}^{M} E_m H_1(f_m) e^{j2\pi f_m t} \\ & + \frac{1}{2^2} \sum_{m_1=-M}^{M} \sum_{m_2=-M}^{M} E_{m_1} E_{m_2} H_2(f_{m_1}, f_{m_2}) e^{j2\pi (f_{m_1} + f_{m_2})t} \\ & + \cdots + \frac{1}{2^N} \sum_{m_1=-M}^{M} \cdots \sum_{m_N=-M}^{M} E_{m_1} \cdots E_{m_N} \\ & \cdot H_N(f_{m_1}, \cdots, f_{m_N}) e^{j2\pi (f_{m_1} + \cdots + f_{m_N})t}. \end{aligned} \tag{11}$$

As expected, when a sum of n tones is applied to a nonlinear system of highest significant order N, additional frequencies are generated consisting of all possible combinations of the tones taken from one up to N at a time. Equation (11) clearly demonstrates that $H_n(f_1, \cdots, f_n)$ is the nonlinear transfer function associated with the sinusoidal output at frequency $(f_1 + \cdots + f_n)$.

In previous applications of the Volterra approach inputs and outputs have generally been characterized in terms of voltages and currents. In microwave problems scattering variables are the natural choice. The purpose of this paper is to develop nonlinear scattering parameters, analogous to the commonly used linear scattering parameters, which can be used to describe power transfer and distortion in nonlinear multiports.

II. Nonlinear Network Analysis Using Scattering Variables

Consider the loaded nonlinear 2-port shown in Fig. 2. In this paper subscripts are used to denote the order of a nonlinear term or parameter. As a result, superscripts in parentheses are used to denote port numbers. Because the only excitation in the system is at port 1, the response at any point in the network depends directly on $e_s(t)$. We first develop expressions for the port voltages and currents.

Following the Volterra approach, the nth-order nonlinear impulse response relating the voltage at port k ($k = 1,2$) to the source voltage $e_s(t)$ is denoted by $h_n^{(ks)}(\tau_1, \cdots, \tau_n)$. Assuming terms above Nth order are negligible, the voltage response at port k is given by

$$e^{(k)}(t) = \sum_{n=1}^{N} e_n^{(k)}(t), \qquad k = 1,2 \tag{12}$$

where the nth-order portion of the response is expressed as

$$e_n^{(k)}(t) = \int h_n^{(ks)}(\tau_1, \cdots, \tau_n) \prod_{p=1}^{n} e_s(t - \tau_p)\, d\tau_p. \tag{13}$$

Similarly, the nth-order nonlinear impulse response relating the current into port k to the source voltage $e_s(t)$ is denoted by $y_n^{(ks)}(\tau_1, \cdots, \tau_n)$. It follows that

$$i^{(k)}(t) = \sum_{n=1}^{N} i_n^{(k)}(t), \qquad k = 1,2 \tag{14}$$

where

$$i_n^{(k)}(t) = \int y_n^{(ks)}(\tau_1, \cdots, \tau_n) \prod_{p=1}^{n} e_s(t - \tau_p)\, d\tau_p. \tag{15}$$

Given a specific network for the nonlinear 2-port, it is a straightforward matter to determine the impulse responses $h_n^{(ks)}(\tau_1, \cdots, \tau_n)$ [3]. These are assumed to be known during the remainder of this discussion. The nonlinear transfer functions corresponding to $h_n^{(ks)}(\tau_1, \cdots, \tau_n)$ and $y_n^{(ks)} \cdot (\tau_1, \cdots, \tau_n)$ are denoted by $H_n^{(ks)}(f_1, \cdots, f_n)$ and $Y_n^{(ks)} \cdot (f_1, \cdots, f_n)$, respectively. Since $y_n^{(ks)}(\tau_1, \cdots, \tau_n)$ relates a current to a voltage, $Y_n^{(ks)}(f_1, \cdots, f_n)$ is referred to as an nth-order nonlinear admittance function.

The nonlinear admittance functions $Y_n^{(ks)}(f_1, \cdots, f_n)$ are readily expressed in terms of the known nonlinear transfer functions $H_n^{(ks)}(f_1, \cdots, f_n)$ by utilizing Kirchhoff's voltage law at each port in conjunction with the harmonic input method [2]. Let $z_s(\tau)$ be the inverse Fourier transform of the source impedance $Z_s(j\omega)$. Application of Kirchhoff's voltage law at port 1 results in

$$e_s(t) = \int_{-\infty}^{\infty} z_s(\tau) i^{(1)}(t-\tau)\, d\tau + e^{(1)}(t)$$
$$= \sum_{n=1}^{N} \int_{-\infty}^{\infty} z_s(\tau) i_n^{(1)}(t-\tau)\, d\tau + \sum_{n=1}^{N} e_n^{(1)}(t). \quad (16)$$

Assume the excitation consists of the sum of M unit amplitude complex exponentials involving the noncommensurable positive frequencies $f_1,\cdots,f_M$. Thus

$$e_s(t) = \sum_{m=1}^{M} e^{j2\pi f_m t}. \quad (17)$$

This results in

$$e_n^{(1)}(t) = \sum_{m_1=1}^{M} \cdots \sum_{m_n=1}^{M} H_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} e^{j2\pi f_{m_p} t} \quad (18)$$

and

$$i_n^{(1)}(t) = \sum_{m_1=1}^{M} \cdots \sum_{m_n=1}^{M} Y_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} e^{j2\pi f_{m_p} t}. \quad (19)$$

Substitution of (17)–(19) into (16) and recognition that

$$\int_{-\infty}^{\infty} z_s(\tau) e^{-j2\pi(f_{m_1}+\cdots+f_{m_n})\tau}\, d\tau = Z_s(f_{m_1}+\cdots+f_{m_n}) \quad (20)$$

yields

$$\sum_{m=1}^{M} e^{j2\pi f_m t} = \sum_{n=1}^{N} \sum_{m_1=1}^{M} \cdots \sum_{m_n=1}^{M} \cdot [Z_s(f_{m_1}+\cdots+f_{m_n}) Y_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) + H_n^{(1s)}(f_{m_1},\cdots,f_{m_n})] \prod_{p=1}^{n} e^{j2\pi f_{m_p} t}. \quad (21)$$

Using the linear independence of the exponentials it is possible to equate terms involving identical frequencies. For example, equating terms involving $e^{j2\pi f_m t}$, it is necessary to focus attention only on those terms for which $n = 1$. The linear transfer functions are then related by

$$Y_1^{(1s)}(f_m) = \frac{1 - H_1^{(1s)}(f_m)}{Z_s(f_m)}. \quad (22)$$

Similarly, for $n > 1$, equating terms involving $\prod_{p=1}^{n} e^{j2\pi f_{m_p} t}$ results in

$$Y_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) = -\frac{H_n^{(1s)}(f_{m_1},\cdots,f_{m_n})}{Z_s(f_{m_1}+\cdots+f_{m_n})}, \quad n > 1. \quad (23)$$

In an analogous manner, application of Kirchhoff's voltage law at port 2 in conjunction with the harmonic input method gives

$$Y_n^{(2s)}(f_{m_1},\cdots,f_{m_n}) = -\frac{H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})}{Z_L(f_{m_1}+\cdots+f_{m_n})}, \quad n \geq 1 \quad (24)$$

where $Z_L(j\omega)$ is the load impedance. Because of the form of $Y_n^{(1s)}(f_{m_1},\cdots,f_{m_n})$ and $Y_n^{(2s)}(f_{m_1},\cdots,f_{m_n})$ for $n > 1$, it is apparent that the voltage source $e_s(t)$ can be considered a short circuit as far as the higher order responses are concerned. This is not surprising in view of the fact that the frequency content of the higher order terms differs from that of $e_s(t)$. Since $e_s(t)$ is an independent voltage source with frequencies specified in (17), the higher order currents in $i^{(1)}$ do not create any voltage drops across the terminals of the source.

Having found the port voltages and currents, it is now possible to determine the scattering variables as a function of the source voltage. By definition, the scattering variables at port k ($k = 1,2$) are related to the corresponding port voltage and current according to the relations [6]

$$\sqrt{r_k}\, \alpha^{(k)}(t) = \tfrac{1}{2}[e^{(k)}(t) + r_k i^{(k)}(t)]$$
$$\sqrt{r_k}\, \beta^{(k)}(t) = \tfrac{1}{2}[e^{(k)}(t) - r_k i^{(k)}(t)], \qquad k = 1,2. \quad (25)$$

$\alpha^{(k)}(t)$ and $\beta^{(k)}(t)$ are the incident and reflected scattering variables, respectively, at port k while r_k is the reference impedance associated with port k. Although complex reference impedances are possible [6], [7], r_k in this paper is restricted to be a positive real resistance in order to simplify the discussion. It is not necessary that port k be terminated in r_k in order to use r_k as the reference impedance. Equation (25) can be inverted to yield

$$e^{(k)}(t) = \sqrt{r_k}\, [\alpha^{(k)}(t) + \beta^{(k)}(t)]$$
$$i^{(k)}(t) = \frac{1}{\sqrt{r_k}} [\alpha^{(k)}(t) - \beta^{(k)}(t)], \qquad k = 1,2. \quad (26)$$

Observe that the instantaneous power into port k can be expressed as

$$p^{(k)}(t) = e^{(k)}(t) i^{(k)}(t) = [\alpha^{(k)}(t)]^2 - [\beta^{(k)}(t)]^2. \quad (27)$$

As a result, we have the interpretation that $[\alpha^{(k)}(t)]^2$ is the instantaneous power in the wave incident on port k while $[\beta^{(k)}(t)]^2$ is the instantaneous power in the wave reflected from port k.

Using the Volterra approach, the nth-order nonlinear impulse response relating the incident scattering variable at port k ($k = 1,2$) to the source voltage $e_s(t)$ is denoted by $q_n^{(ks)}(\tau_1,\cdots,\tau_n)$. Similarly, the nth-order nonlinear impulse response relating the reflected scattering variable at port k ($k = 1,2$) to the source voltage $e_s(t)$ is denoted by $r_n^{(ks)}(\tau_1,\cdots,\tau_n)$. Assuming terms above Nth order to be negligible, the scattering variables at port k are given by

$$\alpha^{(k)}(t) = \sum_{n=1}^{N} \alpha_n^{(k)}(t)$$
$$\beta^{(k)}(t) = \sum_{n=1}^{N} \beta_n^{(k)}(t), \qquad k = 1,2 \quad (28)$$

where the nth-order portions of the response are expressed as

$$\alpha_n^{(k)}(t) = \int q_n^{(ks)}(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} e_s(t-\tau_p)\, d\tau_p$$

$$\beta_n^{(k)}(t) = \int r_n^{(ks)}(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} e_s(t-\tau_p)\, d\tau_p. \quad (29)$$

The nonlinear transfer functions corresponding to $q_n^{(ks)}\cdot(\tau_1,\cdots,\tau_n)$ and $r_n^{(ks)}(\tau_1,\cdots,\tau_n)$ are denoted by $Q_n^{(ks)}\cdot(f_1,\cdots,f_n)$ and $R_n^{(ks)}(f_1,\cdots,f_n)$, respectively. $Q_n^{(ks)}(f_1,\cdots,f_n)$ and $R_n^{(ks)}(f_1,\cdots,f_n)$ are referred to as the nth-order nonlinear incidence and reflection functions, respectively.

The harmonic input method is now used to obtain expressions for the incidence and reflection functions in terms of the known nonlinear transfer functions $H_n^{(ks)}(f_1,\cdots,f_n)$. Assuming $e_s(t)$ to consist of the M unit amplitude complex exponentials given by (17), the nth-order portions of the incident and reflected scattering variables become

$$\alpha_n^{(k)}(t) = \sum_{m_1=1}^{M}\cdots\sum_{m_n=1}^{M} Q_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} e^{j2\pi f_{m_p} t}$$
$$\beta_n^{(k)}(t) = \sum_{m_1=1}^{M}\cdots\sum_{m_n=1}^{M} R_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} e^{j2\pi f_{m_p} t}. \quad (30)$$

Substitution of (12), (14), and (28) into (25) results in

$$\sqrt{r_k}\sum_{n=1}^{N}\alpha_n^{(k)}(t) = \frac{1}{2}\left[\sum_{n=1}^{N} e_n^{(k)}(t) + r_k\sum_{n=1}^{N} i_n^{(k)}(t)\right]$$
$$\sqrt{r_k}\sum_{n=1}^{N}\beta_n^{(k)}(t) = \frac{1}{2}\left[\sum_{n=1}^{N} e_n^{(k)}(t) - r_k\sum_{n=1}^{N} i_n^{(k)}(t)\right], \quad k = 1,2. \quad (31)$$

Making use of (18), (19), and (30) [with the superscript 1 replaced by the superscript k in (18) and (19)] and equating terms involving $\prod_{p=1}^{n} e^{j2\pi f_{m_p} t}$, it follows, for $n \geq 1$, that

$$\sqrt{r_k}\, Q_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) = \tfrac{1}{2}[H_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) + r_k Y_n^{(ks)}(f_{m_1},\cdots,f_{m_n})]$$
$$\sqrt{r_k}\, R_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) = \tfrac{1}{2}[H_n^{(ks)}(f_{m_1},\cdots,f_{m_n}) - r_k Y_n^{(ks)}(f_{m_1},\cdots,f_{m_n})]. \quad (32)$$

To simplify further it is necessary to specify the superscript k. Focusing attention on port 1, $k = 1$. Referring to (22) and (23), the nonlinear incidence and reflection functions at port 1 become

$$\sqrt{r_1}\, Q_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) = \begin{cases} \tfrac{1}{2}H_1^{(1s)}(f_m)\left[1 - \dfrac{r_1}{Z_s(f_m)}\right] + \dfrac{r_1}{2Z_s(f_m)}, & n = 1 \\ \tfrac{1}{2}H_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) \cdot \left[1 - \dfrac{r_1}{Z_s(f_{m_1}+\cdots+f_{m_n})}\right], & n > 1 \end{cases}$$

$$\sqrt{r_1}\, R_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) = \begin{cases} \tfrac{1}{2}H_1^{(1s)}(f_m)\left[1 + \dfrac{r_1}{Z_s(f_m)}\right] - \dfrac{r_1}{2Z_s(f_m)}, & n = 1 \\ \tfrac{1}{2}H_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) \cdot \left[1 + \dfrac{r_1}{Z_s(f_{m_1}+\cdots+f_{m_n})}\right], & n > 1. \end{cases} \quad (33)$$

A particularly interesting case arises when the source impedance $Z_s(j\omega)$ is identical to the port 1 reference impedance r_1. Then

$$Z_s(f_m) = Z_s(f_{m_1}+\cdots+f_{m_n}) = r_1. \quad (34)$$

The expressions for the incidence and reflection functions at port 1 simplify to

$$Q_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) = \begin{cases} \dfrac{1}{2\sqrt{r_1}}, & n = 1 \\ 0, & n > 1 \end{cases}$$

$$R_n^{(1s)}(f_{m_1},\cdots,f_{m_n}) = \begin{cases} \dfrac{1}{2\sqrt{r_1}}[2H_1^{(1s)}(f_m) - 1], & n = 1 \\ \dfrac{H_n^{(1s)}(f_{m_1},\cdots,f_{m_n})}{\sqrt{r_1}}, & n > 1. \end{cases} \quad (35)$$

Taking the inverse transform of $Q_n^{(1s)}(f_{m_1},\cdots,f_{m_n})$, it follows that

$$q_n^{(1s)}(\tau_1,\cdots,\tau_m) = \begin{cases} \dfrac{1}{2\sqrt{r_1}}\delta(\tau_1), & n = 1 \\ 0, & n > 1. \end{cases} \quad (36)$$

With reference to (28) and (29), it is seen that the incident scattering variable is comprised solely of the first-order portion of the response. Specifically, $\alpha_n^{(1)}(t) = 0$ for $n > 1$ and

$$\alpha^{(1)}(t) = \alpha_1^{(1)}(t) = \int_{-\infty}^{\infty} q_1^{(1s)}(\tau_1)e_s(t-\tau_1)\, d\tau_1 = \frac{e_s(t)}{2\sqrt{r_1}}. \quad (37)$$

Hence, when port 1 is terminated in its reference impedance, the incident scattering variable is linearly dependent on the excitation even though the 2-port is nonlinear. In addition, $\alpha^{(1)}(t)$ is independent of the 2-port and is determined entirely by the source voltage and reference impedance. On the other hand, the nonlinear 2-port does reflect components of order greater than unity. Except for $n = 1$, the nonlinear reflection functions are simply the conventional nonlinear transfer functions normalized by the square root of the reference impedance.

Attention is now focused on port 2 for which $k = 2$. Substitution of (24) into (32) results, for $n \geq 1$, in

$$\sqrt{r_2}\, Q_n^{(2s)}(f_{m_1},\cdots,f_{m_n}) = \tfrac{1}{2}H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})\left[1 - \frac{r_2}{Z_L(f_{m_1}+\cdots+f_{m_n})}\right]$$
$$\sqrt{r_2}\, R_n^{(2s)}(f_{m_1},\cdots,f_{m_n}) = \tfrac{1}{2}H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})\left[1 + \frac{r_2}{Z_L(f_{m_1}+\cdots+f_{m_n})}\right]. \quad (38)$$

As before, particularly simple results are obtained when port 2 is terminated in its reference impedance r_2. Then

$$Z_L(f_{m_1} + \cdots + f_{m_n}) = r_2. \quad (39)$$

For this special case the nonlinear incidence and reflection functions at port 2 reduce to

$$Q_n^{(2s)}(f_{m_1},\cdots,f_{m_n}) = 0$$
$$R_n^{(2s)}(f_{m_1},\cdots,f_{m_n}) = \frac{H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})}{\sqrt{r_2}}, \qquad n \geq 1. \quad (40)$$

The zero nonlinear incidence functions imply that $q_n^{(2s)} \cdot (\tau_1,\cdots,\tau_n) = 0$ for all n. In addition, from either (29) or (30) it follows that $\alpha^{(2)}(t) = 0$. Hence the response at port 2 consists entirely of the reflected wave $\beta^{(2)}(t)$. This wave represents the signal impressed upon the load due to the excitation at port 1. Because the load impedance is identical to the reference impedance, there are no reflections from the load. As a result, no waves are incident upon port 2. For all order n observe that the nonlinear reflection functions at port 2 are simply the conventional nonlinear transfer functions normalized by the square root of the reference impedance.

III. Nonlinear Scattering Parameters

Just as 2-port parameters, such as Z parameters, Y parameters, $ABCD$ parameters, etc., have been used to represent linear 2-ports, parameters can also be defined to characterize nonlinear 2-ports [8]. However, the characterization is much more complicated. To completely characterize a nonlinear 2-port, all possible interactions between the independent port variables must be accounted for. This is illustrated as follows in terms of the well-known Z parameters [8].

Consider the nonlinear 2-port shown in Fig. 3. For simplicity assume all voltages and currents are in the sinusoidal steady state. Denote the complex voltage and current associated with sinusoidal waveforms at frequency f by $E(f)$ and $I(f)$, respectively. Using Z parameters, the linear portion of the 2-port is characterized by

$$E_1^{(1)}(f) = Z_1^{(1,1)}(f)I_1^{(1)}(f) + Z_1^{(1,2)}(f)I_1^{(2)}(f)$$
$$E_1^{(2)}(f) = Z_1^{(2,1)}(f)I_1^{(1)}(f) + Z_1^{(2,2)}(f)I_1^{(2)}(f). \quad (41)$$

In (41) the subscript 1 indicates that all variables and parameters are of first order. The superscript k ($k = 1,2$) on the variables indicates the port with which they are associated. Finally, the double superscript j,k ($j,k = 1,2$) on the 2-port parameters indicates that the corresponding dependent variable is at port j while the corresponding independent variable is at port k. Observe that, as usual, four parameters are required to completely characterize the linear portion of the nonlinear 2-port.

To investigate characterization of the second-order portion of the 2-port, assume sinusoidal excitations at frequencies f_1 and f_2. In general, the second-order response at each port consists of sums of sinusoids whose frequencies are the second harmonics and the sums and differences of

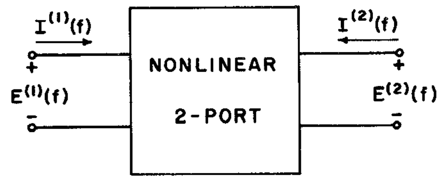

Fig. 3. Nonlinear 2-port with notation and conventions used for port voltages and currents.

the input frequencies. Let the response at the sum frequency $(f_1 + f_2)$ be of interest. To completely characterize the second-order portion of the nonlinear 2-port in terms of the Z parameters, it is necessary to account for each of the possible mechanisms by which components at $(f_1 + f_2)$ may be generated in the port voltages. Accordingly, at port k, ($k = 1,2$),

$$\begin{aligned} E_2^{(k)}(f_1 + f_2) = {} & Z_1^{(k,1)}(f_1 + f_2)I_2^{(1)}(f_1 + f_2) \\ & + Z_1^{(k,2)}(f_1 + f_2)I_2^{(2)}(f_1 + f_2) \\ & + Z_2^{(k;1,1)}(f_1,f_2)I_1^{(1)}(f_1)I_1^{(1)}(f_2) \\ & + Z_2^{(k;1,2)}(f_1,f_2)I_1^{(1)}(f_1)I_1^{(2)}(f_2) \\ & + Z_2^{(k;2,1)}(f_1,f_2)I_1^{(2)}(f_1)I_1^{(1)}(f_2) \\ & + Z_2^{(k;2,2)}(f_1,f_2)I_1^{(2)}(f_1)I_1^{(2)}(f_2). \quad (42) \end{aligned}$$

In (42) the subscripts and superscripts have the same interpretation as in (41). However, now the triple superscript $j;k,l$ ($j,k,l = 1,2$) on the 2-port parameters indicates that the corresponding dependent variable is at port j while the corresponding independent variable at f_1 is at port k and the corresponding independent variable at f_2 is at port l. For example, $Z_2^{(1;2,1)}(f_1,f_2)I_1^{(2)}(f_1)I_1^{(1)}(f_2)$ is that portion of the second-order voltage response at port 1 and frequency $(f_1 + f_2)$ due to the mix of the linear portion of the current at port 2 and frequency f_1 with the linear portion of the current at port 1 and frequency f_2. Note that eight second-order parameters are needed in addition to the four linear parameters in order to completely characterize the second-order behavior of the 2-port. In general, 2^{n+1} nth-order parameters are needed in the characterization of the nth-order portion of a nonlinear 2-port. Because this number grows large even for relatively small n, representation of the behavior of nonlinear 2-ports is a difficult and cumbersome task.

In many practical situations scattering variables can be used to considerably simplify the characterization of nonlinear 2-ports. As before, assume the network is in the sinusoidal steady state. Let the complex incident and reflected scattering variables be denoted by $a(f)$ and $b(f)$, respectively. Hence

$$a_n^{(2)}(f) = |a_n^{(2)}(f)|e^{j\theta_n^{(2)}(f)} \quad (43)$$

implies that

$$\alpha_n^{(2)}(t) = |a_n^{(2)}(f)| \cos [2\pi ft + \theta_n^{(2)}(f)]. \quad (44)$$

In terms of scattering parameters, the linear portion of the nonlinear 2-port is characterized by

$$b_1^{(1)}(f) = S_1^{(1,1)}(f)a_1^{(1)}(f) + S_1^{(1,2)}(f)a_1^{(2)}(f)$$
$$b_1^{(2)}(f) = S_1^{(2,1)}(f)a_1^{(1)}(f) + S_1^{(2,2)}(f)a_1^{(2)}(f). \quad (45)$$

The parameters in (45) are the conventional scattering parameters encountered in linear circuit theory.

Following the discussion leading to (42), if the 2-port is excited by sinusoids at f_1 and f_2, the second-order behavior at port k ($k = 1,2$) for the sum frequency ($f_1 + f_2$) is given by

$$\begin{aligned} b_2^{(k)}(f_1 + f_2) = {} & S_1^{(k,1)}(f_1 + f_2)a_2^{(1)}(f_1 + f_2) \\ & + S_1^{(k,2)}(f_1 + f_2)a_2^{(2)}(f_1 + f_2) \\ & + S_2^{(k;1,1)}(f_1,f_2)a_1^{(1)}(f_1)a_1^{(1)}(f_2) \\ & + S_2^{(k;1,2)}(f_1,f_2)a_1^{(1)}(f_1)a_1^{(2)}(f_2) \\ & + S_2^{(k;2,1)}(f_1,f_2)a_1^{(2)}(f_1)a_1^{(1)}(f_2) \\ & + S_2^{(k;2,2)}(f_1,f_2)a_1^{(2)}(f_1)a_1^{(2)}(f_2). \quad (46) \end{aligned}$$

In many microwave applications the source and load impedances are pure resistances. This is the situation, for example, when a port is loaded by a transmission line terminated in its characteristic impedance. The reference impedances r_1 and r_2 can then be chosen equal to Z_S and Z_L, respectively. This, in turn, results in $\alpha_n^{(1)}(t) = 0$ for $n > 1$ and $\alpha_n^{(2)}(t) = 0$ for all n. Consequently,

$$a_2^{(1)}(f_1 + f_2) = a_2^{(2)}(f_1 + f_2) = a_1^{(2)}(f_2) = a_1^{(2)}(f_1) = 0. \quad (47)$$

With (47) substituted into (46), the second-order behavior at the sum frequency ($f_1 + f_2$) reduces to

$$b_2^{(1)}(f_1 + f_2) = S_2^{(1;1,1)}(f_1,f_2)a_1^{(1)}(f_1)a_1^{(1)}(f_2)$$
$$b_2^{(2)}(f_1 + f_2) = S_2^{(2;1,1)}(f_1,f_2)a_1^{(1)}(f_1)a_1^{(1)}(f_2). \quad (48)$$

It follows that only two nonlinear scattering parameters are needed to completely characterize the second-order portion of the nonlinear 2-port when the source and load impedances equal the reference impedances at ports 1 and 2, respectively. Similarly, the nth-order portion of the 2-port is characterized by

$$b_n^{(1)}(f_1 + \cdots + f_n) = S_n^{(1;\underbrace{1,\cdots,1}_{n})}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(1)}(f_p)$$
$$b_n^{(2)}(f_1 + \cdots + f_n) = S_n^{(2;\underbrace{1,\cdots,1}_{n})}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(1)}(f_p). \quad (49)$$

Hence, under matched conditions, the nth-order behavior of a nonlinear 2-port can be completely characterized using only two parameters instead of the 2^{n+1} parameters mentioned earlier.

The previous discussion was based upon sinusoidal inputs in order to simplify the presentation. Nonlinear scattering functions are now derived in terms of arbitrary inputs. In particular, it is shown how the nonlinear scattering functions are related to the nonlinear incidence and reflection functions defined previously. Throughout this development matched conditions are assumed such that

$$Z_S(j\omega) = r_1 \qquad Z_L(j\omega) = r_2. \quad (50)$$

Since $\alpha_n^{(1)}(t) = 0$ for $n > 1$,

$$\alpha^{(1)}(t) = \alpha_1^{(1)}(t) = \int_{-\infty}^{\infty} q_1^{(1s)}(\tau_1)e_s(t - \tau_1)\, d\tau_1 = \frac{e_s(t)}{2\sqrt{r_1}}. \quad (51)$$

Converting to the frequency domain, let $a^{(1)}(f)$, $a_1^{(1)}(f)$, $Q_1^{(1s)}(f)$, and $E_S(f)$ be the Fourier transforms of $\alpha^{(1)}(t)$, $\alpha_1^{(1)}(t)$, $q_1^{(1s)}(t)$, and $e_s(t)$, respectively. Then

$$a^{(1)}(f) = a_1^{(1)}(f) = Q_1^{(1s)}(f)E_S(f) = \frac{E_S(f)}{2\sqrt{r_1}}. \quad (52)$$

Because $\alpha_n^{(2)}(t) = 0$ for all n in addition to $\alpha_n^{(1)}(t) = 0$ for $n > 1$, the reflected waves at both ports depend only on the linear portion of the incident wave at port 1. Let the nth-order impulse response relating the reflected waves at port k to $\alpha_1^{(1)}(t)$ be denoted by $\sigma_n^{(k1)}(\tau_1,\cdots,\tau_n)$. It follows that

$$\beta^{(k)}(t) = \sum_{n=1}^{N} \beta_n^{(k)}(t) \quad (53)$$

where

$$\beta_n^{(k)}(t) = \int \sigma_n^{(k1)}(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} \alpha_1^{(1)}(t - \tau_p)\, d\tau_p. \quad (54)$$

The nonlinear transfer function corresponding to $\sigma_n^{(k1)} \cdot (\tau_1,\cdots,\tau_n)$ is denoted by $S_n^{(k1)}(f_1,\cdots,f_n)$. In analogy with (6), it can be shown that $\beta_n^{(k)}(t)$ can also be expressed as

$$\beta_n^{(k)}(t) = \int S_n^{(k1)}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(1)}(f_p)e^{j2\pi f_p t}\, df_p. \quad (55)$$

Since $S_n^{(k1)}(f_1,\cdots,f_n)$ relates a reflected wave to an incident wave, it is referred to as an nth-order nonlinear scattering function.

A simple expression for the nonlinear scattering function in terms of the complex scattering variables is obtained by introducing the multidimensional time function

$$\beta_n^{(k)}(t_1,\cdots,t_n) = \int \sigma_n^{(k1)}(\tau_1,\cdots,\tau_n) \prod_{p=1}^{n} \alpha_1^{(1)}(t_p - \tau_p)\, d\tau_p. \quad (56)$$

With reference to (54), observe that $\beta_n^{(k)}(t_1,\cdots,t_n) = \beta_n^{(k)}(t)$ when $t_1 = \cdots = t_n = t$. Define the n-dimensional Fourier transform of $\beta_n^{(k)}(t_1,\cdots,t_n)$ to be $b_n^{(k)}(f_1,\cdots,f_n)$. It follows that

$$b_n^{(k)}(f_1,\cdots,f_n) = S_n^{(k1)}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(1)}(f_p). \quad (57)$$

Hence the nonlinear scattering function is given by

$$S_n^{(k1)}(f_1,\cdots,f_n) = \frac{b_n^{(k)}(f_1,\cdots,f_n)}{\prod_{p=1}^{n} a_1^{(1)}(f_p)}. \tag{58}$$

For the preceding expression to be valid, recall that the source and load impedances must equal the reference impedances at ports 1 and 2, respectively.

An alternate expression for $S_n^{(k1)}(f_1,\cdots,f_n)$ in terms of the nonlinear incidence and reflection functions results by observing that $\beta_n^{(k)}(t)$ can also be written as

$$\beta_n^{(k)}(t) = \int_{\underline{f}} R_n^{(ks)}(f_1,\cdots,f_n) \prod_{p=1}^{n} E_s(f_p) e^{j2\pi f_p t}\, df_p. \tag{59}$$

This results by transforming (29) such that the input is in the frequency domain while the output is in the time domain, as is the case with (6) and (55). By comparing (55) and (59), it follows that

$$S_n^{(k1)}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(1)}(f_p) = R_n^{(ks)}(f_1,\cdots,f_n) \prod_{p=1}^{n} E_s(f_p). \tag{60}$$

From (52), $a_1^{(1)}(f) = Q_1^{(1s)}(f)E_s(f)$. Substitution into (60) results in

$$S_n^{(k1)}(f_1,\cdots,f_n) = \frac{R_n^{(ks)}(f_1,\cdots,f_n)}{\prod_{p=1}^{n} Q_1^{(1s)}(f_p)}. \tag{61}$$

As with (58), this equation is valid only when impedance matches exist at both ports. Equation (61) demonstrates that the nonlinear scattering, incidence, and reflection functions are related in a straightforward manner. In fact, substitution of (35) and (40) into (61) yields

$$S_n^{(11)}(f_1,\cdots,f_n) = \begin{cases} 2H_1^{(1s)}(f_1) - 1, & n = 1 \\ (2)^n (r_1)^{(n-1)/2} H_n^{(1s)}(f_1,\cdots,f_n), & n > 1 \end{cases}$$

$$S_n^{(21)}(f_1,\cdots,f_n) = (2)^n (r_1)^{n/2} \frac{1}{\sqrt{r_2}} H_n^{(2s)}(f_1,\cdots,f_n), \qquad n \geq 1. \tag{62}$$

Hence knowledge of the conventional nonlinear transfer functions $H_n^{(ks)}(f_1,\cdots,f_n)$ is adequate to determine the nonlinear scattering functions.

The remaining task is to relate the nonlinear scattering functions $S_n^{(k1)}(f_1,\cdots,f_n)$ to the nonlinear scattering parameters $S_n^{(k;\underbrace{1,\cdots,1}_{n})}(f_1,\cdots,f_n)$ introduced in (49) for sinusoidal excitation. Assume

$$\alpha_1^{(1)}(t) = \sum_{m=-M}^{M} \frac{1}{2} a_1^{(1)}(f_m) e^{j2\pi f_m t} \tag{63}$$

where

$$f_{-m} = -f_m \qquad a_1^{(1)}(f_0) = 0 \qquad a_1^{(1)}(f_{-m}) = (a_1^{(1)}(f_m))^*. \tag{64}$$

The incident scattering variable at port 1, therefore, consists of the sum of m sinusoidal tones at frequencies $f_1,\cdots,f_m$. Substitution of (63) into (56) yields

$$\beta_n^{(k)}(t) = \frac{1}{(2)^n} \sum_{m_1=-M}^{M} \cdots \sum_{m_n=-M}^{M} S_n^{(k1)}(f_{m_1},\cdots,f_{m_n}) \cdot \prod_{p=1}^{n} a_1^{(1)}(f_{m_p}) e^{j2\pi(f_{m_1}+\cdots+f_{m_n})t}. \tag{65}$$

The nth-order portion of the response is seen to consist of an n-fold summation. By definition, $a_1^{(1)}(f_0) = 0$. Hence, ignoring the zero index, each summation extends over $2M$ indices. Thus the n-fold summation contains $(2M)^n$ individual terms. Since the frequency associated with each term is $(f_{m_1} + \cdots + f_{m_n})$, the output frequencies are those that can be generated by adding together all possible combinations of the input frequencies $-f_m,\cdots,-f_1, f_1,\cdots,f_m$ taken n at a time.

The output frequencies can also be expressed in the form $(q_{-M}f_{-M} + \cdots + q_{-1}f_{-1} + q_1 f_1 + \cdots + q_M f_M)$ where q_m is a nonnegative integer that denotes the number of times the index $m = -M,\cdots,-1,1,\cdots,M$ occurs in the various frequency combinations. Since exactly n frequencies are involved in each frequency mix, the q_m obey the constraint

$$q_{-M} + \cdots + q_{-1} + q_1 + \cdots + q_M = n. \tag{66}$$

The output frequencies are then those that can be generated by all possible choices of the q_m such that (66) is satisfied.

It can be shown that the nonlinear transfer functions are symmetrical in their arguments [3]. As a result, all of the terms in (65) involving the same frequency mix are identical. The number of different ways in which the n indices $\underline{m} = (m_1,\cdots,m_n)$ can be partitioned such that $-M$ appears q_{-M} times, $\cdots$, -1 appears q_{-1} times, 1 appears q_1 times, $\cdots$, and M appears q_M times is given by the multinomial coefficient, denoted here by

$$\binom{n}{\underline{q}} = \frac{n!}{(q_{-M}!)\cdots(q_{-1}!)(q_1!)\cdots(q_M!)} \tag{67}$$

where $\underline{q} = (q_{-M},\cdots,q_{-1}, q_1,\cdots,q_M)$. Combining identical terms, (65) can be written as

$$\beta_n^{(k)}(t) = \frac{1}{(2)^n} \underline{\sum_{m_1=-M}^{M} \cdots \sum_{m_n=-M}^{M}} \binom{n}{\underline{q}} S_n^{(k1)}(f_{m_1},\cdots,f_{m_n}) \cdot \prod_{p=1}^{n} a_1^{(1)}(f_{m_p}) e^{j2\pi(f_{m_1}+\cdots+f_{m_n})t} \tag{68}$$

where the underline under the multiple summation sign serves as a reminder that only distinct terms are to be included in the summation. From (68) the complex scattering variable associated with the sinusoidal response at $(f_{m_1} + \cdots + f_{m_n})$ is

$$b_n^{(k)}(f_{m_1} + \cdots + f_{m_n}) = \frac{1}{(2)^{n-1}} \binom{n}{\underline{q}} S_n^{(k1)}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} a_1^{(1)}(f_{m_p}) \tag{69}$$

where $b_n^{(k)}(f_{m_1} + \cdots + f_{m_n})$ is analogous to the complex amplitude $a_1^{(1)}(f_m)$ in (63). However, from (49),

$$b_n^{(k)}(f_{m_1} + \cdots + f_{m_n}) = S_n^{(k;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n}) \prod_{p=1}^{n} a_1^{(1)}(f_{m_p}).$$

It follows that the nonlinear scattering functions in (55) and the nonlinear scattering parameters in (49) are related by

$$S_n^{(k;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n}) = \frac{1}{(2)^{n-1}} \binom{n}{q} S_n^{(k1)}(f_{m_1},\cdots,f_{m_n}). \quad (70)$$

Of the two quantities, $S_n^{(k1)}(f_{m_1},\cdots,f_{m_n})$ is the more basic because it is valid for arbitrary inputs. However,

$$S_n^{(k;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n})$$

is the quantity of interest when dealing with sinusoidal excitations.

$S_n^{(2;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n})$ has a simple interpretation in terms of the forward nonlinear transducer gain of the 2-port. Let the source voltage be given by

$$e_s^{(1)}(t) = \frac{1}{2} \sum_{m=-M}^{M} E_m e^{j2\pi f_m t} \quad (71)$$

where

$$f_{-m} = -f_m \qquad E_0 = 0 \qquad E_{-m} = E_m^*. \quad (72)$$

The complex voltage at port 2 associated with the sinusoidal response at frequency $f_{m_1} + \cdots + f_{m_n} = \sum_{p=-M,\, p\neq 0}^{M} q_p f_p$ is

$$E_n^{(2)}(f_{m_1} + \cdots + f_{m_n}) = \frac{1}{2^{n-1}} \binom{n}{q} \prod_{\substack{p=-M\\p\neq 0}}^{M} (E_p)^{q_p} H_n^{(2s)}(f_{m_1},\cdots,f_{m_n}). \quad (73)$$

Assuming the load impedance to equal the reference impedance at port 2, the average power dissipated in r_2 is

$$P_n^{(L)}(f_{m_1} + \cdots + f_{m_n}) = \tfrac{1}{2}|E_n^{(2)}(f_{m_1} + \cdots + f_{m_n})|^2 \frac{1}{r_2}. \quad (74)$$

Assuming the source impedance to equal the reference impedance at port 1, the available power from the source at frequency f_m is

$$P_{\text{ava}}^{(s)}(f_m) = \frac{1}{8} \frac{|E_m|^2}{r_1}. \quad (75)$$

The forward nonlinear transducer gain is defined to be

$$g_n^{(T)}(f_{m_1} + \cdots + f_{m_n}) = \frac{P_n^{(L)}(f_{m_1} + \cdots + f_{m_n})}{\prod_{\substack{p=-M\\p\neq 0}}^{M} [P_{\text{ava}}^{(s)}(f_p)]^{q_p}}. \quad (76)$$

This is a natural definition for the transducer gain if the gain is to be independent of the complex voltages of the input tones. Substitution of (74) and (75) into (76) and use of (66) results in

$$g_n^{(T)}(f_{m_1} + \cdots + f_{m_n}) = (2)^{n+1} \binom{n}{q}^2 \frac{(r_1)^n}{r_2} |H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})|^2. \quad (77)$$

Combining (70) and (62), it follows that

$$|S_n^{(2;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n})|^2 = (2)^2 \binom{n}{q}^2 \frac{(r_1)^n}{r_2} |H_n^{(2s)}(f_{m_1},\cdots,f_{m_n})|^2 = \frac{1}{(2)^{n-1}} g_n^{(T)}(f_{m_1} + \cdots + f_{m_n}). \quad (78)$$

Thus the forward nonlinear transducer gain is simply related to the squared magnitude of the nonlinear scattering parameter $S_n^{(2;\overbrace{1,\cdots,1}^{n})}(f_{m_1},\cdots,f_{m_n})$. Hence knowledge of this scattering parameter is sufficient to determine the average power dissipated in the load at some intermodulation frequency assuming the 2-port is matched at both ports. For $n = 1$, (78) reduces to

$$g_1^{(T)}(f_m) = |S_1^{(2,1)}(f_m)|^2. \quad (79)$$

This is the usual interpretation of the linear scattering parameter in terms of the linear forward transducer gain. Thus, once again, a familiar concept from linear scattering theory is seen to be nothing more than a special case of the more general nonlinear problem.

IV. Nonlinear Applications of Scattering Variables

Scattering variables are a natural choice for the analysis of microwave systems involving nonlinearities. To demonstrate their use, a tunnel diode amplifier employing a circulator (Section IV-A) and a nonlinearly loaded transmission line (Section IV-B) are discussed as follows.

A. Tunnel Diode Amplifier

A tunnel diode has the typical i–v characteristic and nonlinear equivalent circuit shown in Fig. 4. In amplifier and/or oscillator applications the diode is usually operated in the negative conductance region of the i–v characteristic. The tunnel diode junction is modeled by a nonlinear conductance in parallel with a nonlinear capacitance. The currents through these nonlinear elements are assumed to be adequately characterized by the power series representations [3]

$$i_c(t) = \sum_{n=1}^{N} \frac{d}{dt} [C_n e^n(t)]$$

$$i_G(t) = \sum_{n=1}^{N} G_n e^n(t). \quad (80)$$

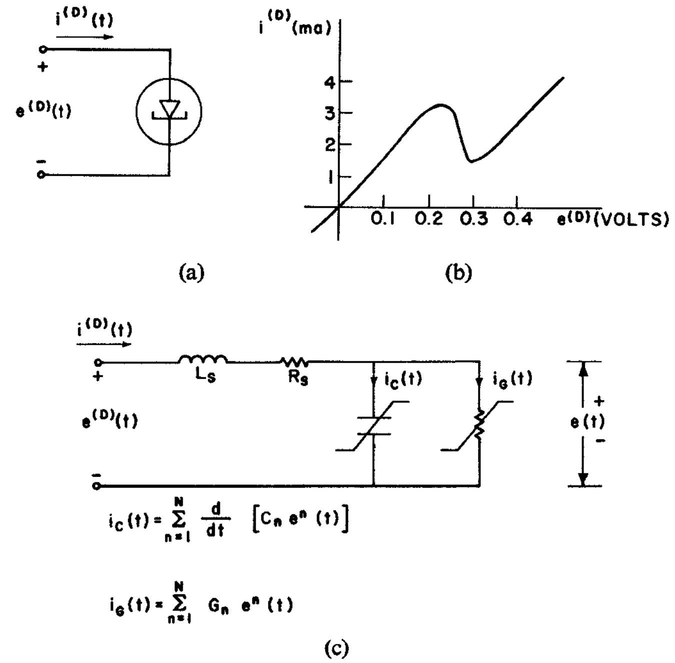

Fig. 4. Tunnel diode representation. (a) Circuit symbol. (b) i–v characteristic. (c) Equivalent circuit.

The bulk resistance of the materials and the lead inductance are characterized by the linear elements R_S and L_S, respectively. In this example the tunnel diode is imbedded in a reflection amplifier employing a circulator as shown in Fig. 5. For the ideal circulator, assuming that the source impedance, load impedance, and all reference impedances are chosen equal to the real characteristic impedance of the circulator R, the linear scattering parameters are identically zero except for $S_1^{(1,3)}$, $S_1^{(2,1)}$, and $S_1^{(3,2)}$ which equal unity. Hence the scattering variables into and out of the circulator ports are constrained by

$$\beta^{(1)}(t) = \alpha^{(3)}(t) \qquad \beta^{(2)}(t) = \alpha^{(1)}(t) \qquad \beta^{(3)}(t) = \alpha^{(2)}(t). \tag{81}$$

We see that the circulator has a cyclic power transmission capability in that the wave reflected from port 1 is the wave incident on port 3, the wave reflected from port 2 is the wave incident on port 1, and the wave reflected from port 3 is the wave incident on port 2. Because the load impedance at port 3 and the source impedance at port 1 equal the reference impedance,

$$\alpha_n^{(3)}(t) = 0, \qquad n \geq 1$$
$$\alpha_n^{(1)}(t) = 0, \qquad n > 1. \tag{82}$$

The problem is to predict second-order intermodulation power levels in the load.

Note that all of the power from the source is transferred to the tunnel diode while all of the power reflected from the tunnel diode is transferred to the load. Consequently, it is only necessary to evaluate the intermodulation power in $\beta^{(D)}(t)$ as shown in Fig. 6. With reference to Figs. 5 and 6,

$$e^{(D)}(t) = e^{(2)}(t) \qquad i^{(D)}(t) = -i^{(2)}(t) \qquad \alpha^{(D)}(t) = \beta^{(2)}(t)$$

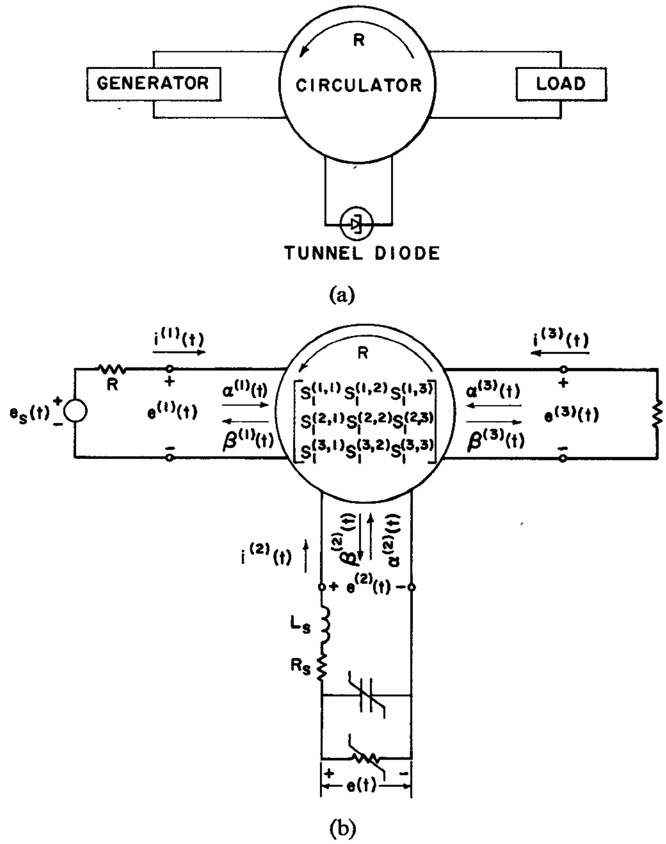

Fig. 5. Tunnel diode reflection amplifier with equivalent circuit.

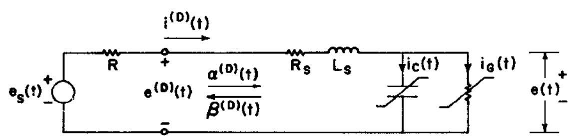

Fig. 6. Simplified nonlinear equivalent circuit.

$$\alpha^{(1)}(t) = \alpha^{(D)}(t) \qquad \beta^{(D)}(t) = \alpha^{(2)}(t) \qquad \beta^{(3)}(t) = \beta^{(D)}(t). \tag{83}$$

From (55), the nth-order portion of $\beta^{(D)}(t)$ is given by

$$\beta_n^{(D)}(t) = \int S_n^{(DD)}(f_1,\cdots,f_n) \prod_{p=1}^{n} a_1^{(D)}(f_p) e^{j2\pi f_p t}\, df_p \tag{84}$$

where $S_n^{(DD)}(f_1,\cdots,f_n)$ is the nth-order nonlinear scattering function relating the reflected wave from the tunnel diode to the incident wave on the diode. Making use of (62), this scattering function is related to the conventional nonlinear transfer function according to the relation

$$S_n^{(DD)}(f_1,\cdots,f_n) = \begin{cases} 2H_1^{(DS)}(f_1) - 1, & n = 1 \\ (2)^n (R)^{(n-1)/2} H_n^{(DS)}(f_1,\cdots,f_n), & n > 1 \end{cases} \tag{85}$$

where $H_n^{(DS)}(f_1,\cdots,f_n)$ is the nth-order nonlinear transfer function relating the diode voltage $e^{(D)}(t)$ to the source voltage $e^{(S)}(t)$.

Let $Z_S(f)$ denote the series impedance of R_S and L_S and let $Z_1(f)$ denote the parallel impedance of C_1 and G_1

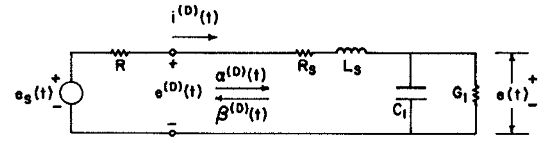

Fig. 7. Linearized equivalent circuit for the nonlinear circuit of Fig. 6.

where C_1 and G_1 are the coefficients of the linear terms in (80). The linearized equivalent circuit for the nonlinear circuit of Fig. 6 is shown in Fig. 7. It follows that

$$H_1^{(DS)}(f_1) = \frac{Z_S(f_1) + Z_1(f_1)}{R + Z_S(f_1) + Z_1(f_1)}. \tag{86}$$

Substitution of (86) into (85) results in

$$S_1^{(DD)}(f_1) = \frac{Z_D(f_1) - R}{Z_D(f_1) + R} \tag{87}$$

$$g_2^{(T)}(f_1 + f_2) = 2|S_2^{(D;D,D)}(f_1,f_2)|^2 = 32R^3 \left| \frac{[j2\pi(f_1 + f_2)C_2 + G_2]Z_1(f_1 + f_2)Z_1(f_1)Z_1(f_2)}{[Z_D(f_1 + f_2) + R][Z_D(f_1) + R][Z_D(f_2) + R]} \right|^2. \tag{92}$$

where $Z_D(f_1) = Z_S(f_1) + Z_1(f_1)$ is the impedance seen looking into the tunnel diode. $S_1^{(DD)}(f_1)$ is recognized to be a reflection coefficient. When the tunnel diode is operated in its negative conductance region, the magnitude of the reflection coefficient can be greater than unity yielding a reflected power which is greater than the incident power. Because of the circulator, all of this power is delivered to the load.

The second-order scattering function describes the second-order behavior which gives rise to second harmonics and second-order sum and difference frequencies. Applying the analysis technique of [3], the second-order nonlinear transfer function is found to be

$$H_2^{(DS)}(f_1,f_2) = -\frac{R[j2\pi(f_1 + f_2)C_2 + G_2]Z_1(f_1 + f_2)Z_1(f_1)Z_1(f_2)}{[Z_D(f_1 + f_2) + R][Z_D(f_1) + R][Z_D(f_2) + R]} \tag{88}$$

where C_2 and G_2 are the coefficients of the quadratic terms in (80). An interesting point is revealed by (88). Since $Z_1(f)$ appears in three of the factors in the numerator as well as in the three factors of the denominator whereas the second-order coefficients appear only once, it is more important to accurately characterize the linear portion of the nonlinearity than the second-order portion even though second-order effects are of interest. Substitution of (88) into (85) yields

$$S_2^{(DD)}(f_1,f_2) = -\frac{4R^{3/2}[j2\pi(f_1 + f_2)C_2 + G_2]Z_1(f_1 + f_2)Z_1(f_1)Z_1(f_2)}{[Z_D(f_1 + f_2) + R][Z_D(f_1) + R][Z_D(f_2) + R]}. \tag{89}$$

The nonlinear scattering parameter is related to the nonlinear scattering function, by (70), according to the relation

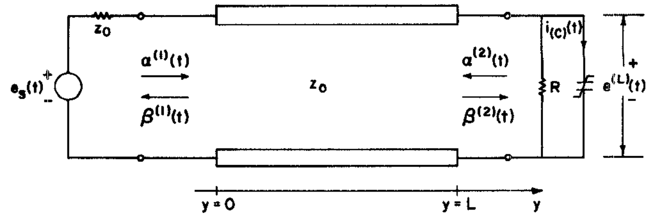

Fig. 8. Nonlinearly loaded transmission line.

$$S_n^{(D;D,D)}(f_1,f_2) = \frac{1}{2}\frac{2!}{1!\,1!} S_n^{(DD)}(f_1,f_2) = S_n^{(DD)}(f_1,f_2). \tag{90}$$

Hence, in analogy with (76) and (78), the average power dissipated in the load R by the second-order intermodulation component at $(f_1 + f_2)$ is

$$P_2^{(L)}(f_1 + f_2) = g_2^{(T)}(f_1 + f_2)P_{\text{ava}}^{(s)}(f_1)P_{\text{ava}}^{(s)}(f_2) \tag{91}$$

where

If desired, the power in other intermodulation components can be obtained in a similar manner.

B. Nonlinearly Loaded Transmission Line

Consider the nonlinearly loaded transmission line of length L shown in Fig. 8. The line is assumed to be lossless with real characteristic impedance $Z_0 = \sqrt{l/c}$ where l and c are the inductance and capacitance per unit length of the line, respectively. The linear scattering matrix of the line is given by

$$S = \begin{bmatrix} S_1^{(1,1)}(f) & S_1^{(1,2)}(f) \\ S_1^{(2,1)}(f) & S_1^{(2,2)}(f) \end{bmatrix} = e^{-j2\pi f\sqrt{lc}\,L} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{93}$$

Hence

$$\beta^{(1)}(t) = \alpha^{(2)}(t - \sqrt{lc}\,L)$$
$$\beta^{(2)}(t) = \alpha^{(1)}(t - \sqrt{lc}\,L) \tag{94}$$

and the transmission line acts as a pure delay. Let the reference impedance at both ports equal Z_0, the characteristic impedance of the line. Since the source impedance is also Z_0, $\alpha_n^{(1)}(t) = 0$ for $n > 1$ and the incident wave at port 1 is given by

$$\alpha^{(1)}(t) = \alpha_1^{(1)}(t) = \frac{e_s(t)}{2\sqrt{Z_0}}. \tag{95}$$

Note that $\beta_n^{(2)}(t) = \alpha_n^{(1)}(t - \sqrt{lc}\,L) = 0$ for $n > 1$ and that

$$\beta^{(2)}(t) = \beta_1^{(2)}(t) = \frac{e_s(t - \sqrt{lc}\,L)}{2\sqrt{Z_0}}. \qquad (96)$$

The nonlinear capacitance is characterized by the power series expansion

$$i_c(t) = \sum_{n=1}^{N} \frac{d}{dt}\,[C_n(e^{(L)}(t))^n]. \qquad (97)$$

The problem is to determine the average power in the second harmonic of the wave reflected from the load.

By virtue of the simple relationships between the scattering variables at ports 1 and 2, attention can be focused entirely on the nonlinear equivalent circuit shown in Fig. 9. With reference to Figs. 8 and 9

$$\alpha^{(L)}(t) = \beta^{(2)}(t) \qquad \beta^{(L)}(t) = \alpha^{(2)}(t). \qquad (98)$$

Since $\alpha_n^{(L)}(t) = \beta_n^{(2)}(t) = 0$ for $n > 1$, the load in Fig. 9 is driven by a source which is matched to the reference impedance. Since average power in the second harmonic of the wave reflected from the load is determined by the squared magnitude of $S_2^{(LL)}(f,f)$, the problem reduces to the evaluation of this parameter.

In the previous example the nonlinear scattering functions were determined by first obtaining the conventional nonlinear transfer functions. This approach is not necessary. In this example the nonlinear scattering functions are obtained directly.

Application of Kirchhoff's current law at the port in Fig. 9 yields

$$i^{(L)}(t) = \frac{1}{R}\,e^{(L)}(t) + i_c(t). \qquad (99)$$

By definition, the port voltage and current are related to the port scattering variables according to the equation

$$e^{(L)}(t) = \sqrt{Z_0}\,[\alpha^{(L)}(t) + \beta^{(L)}(t)]$$

$$i^{(L)}(t) = \frac{1}{\sqrt{Z_0}}\,[\alpha^{(L)}(t) - \beta^{(L)}(t)]. \qquad (100)$$

Substitution of (100) and (97) into (99) results in

$$\frac{1}{\sqrt{Z_0}}[\alpha^{(L)}(t) - \beta^{(L)}(t)] = \frac{\sqrt{Z_0}}{R}[\alpha^{(L)}(t) + \beta^{(L)}(t)] + \frac{d}{dt}\left\{\sum_{i=1}^{N} C_i(Z_0)^{i/2}[\alpha^{(L)}(t) + \beta^{(L)}(t)]^i\right\}. \qquad (101)$$

Because the source impedance equals the reference impedance,

$$\alpha^{(L)}(t) = \frac{e_g(t)}{2\sqrt{Z_0}}.$$

Hence the only unknown in (101) is $\beta^{(L)}(t)$. Rearranging (101) such that only terms linear in $\beta^{(L)}(t)$ are on the left-hand side, we obtain after some simplification

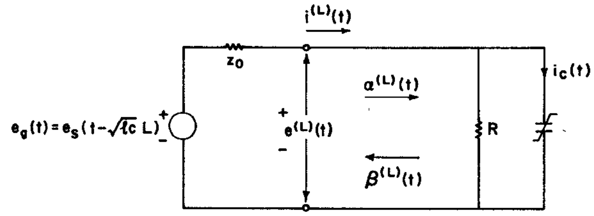

Fig. 9. Simplified equivalent circuit.

$$\left[1 + \frac{Z_0}{R} + pC_1Z_0\right]\beta^{(L)}(t) = \left[1 - \frac{Z_0}{R} - pC_1Z_0\right]\alpha^{(L)}(t) - p\sum_{i=2}^{N} C_i(Z_0)^{(i+1)/2}[\alpha^{(L)}(t) + \beta^{(L)}(t)]^i \qquad (102)$$

where p denotes the differential operator d/dt. Denote the nth-order nonlinear scattering function relating the wave reflected from the load to the wave incident on the load by $S_n^{(LL)}(f_1,\cdots,f_n)$. These scattering functions can be obtained directly by applying the harmonic input method to (102).

To determine the first-order scattering function, let

$$\alpha^{(L)}(t) = e^{j2\pi ft}. \qquad (103)$$

The reflected wave consists of the harmonics of f and is given by

$$\beta^{(L)}(t) = \sum_{n=1}^{N} S_n^{(LL)}(f,\cdots,f)e^{j2\pi nft}. \qquad (104)$$

Substitution of (104) into (102), equating terms involving $e^{j2\pi ft}$, and cancellation of the factor $e^{j2\pi ft}$ yields

$$\left[1 + \frac{Z_0}{R} + j2\pi fC_1Z_0\right] S_1^{(LL)}(f) = \left[1 - \frac{Z_0}{R} - j2\pi fC_1Z_0\right]. \qquad (105)$$

Noting that

$$Y_L(f) = \frac{1}{Z_L(f)} = \frac{1}{R} + j2\pi fC_1 \qquad (106)$$

the first-order scattering function becomes

$$S_1^{(LL)}(f) = \frac{Z_L(f) - Z_0}{Z_L(f) + Z_0}. \qquad (107)$$

As expected from linear scattering theory, $S_1^{(LL)}(f)$ is identical to the reflection coefficient of the 1-port.

The excitation given by (103) can also be used to determine the second-order scattering function $S_2^{(LL)}(f,f)$. Substitution of (104) into (102), equating terms involving $e^{j2\pi(2f)t}$, and cancellation of the factor $e^{j2\pi(2f)t}$ results in

$$\left\{1 + Z_0\left[\frac{1}{R} + j2\pi(2f)C_1\right]\right\} S_2^{(LL)}(f,f) = -j2\pi(2f)C_2Z_0^{3/2}[1 + 2S_1^{(LL)}(f) + (S_1^{(LL)}(f))^2]. \qquad (108)$$

Solving for $S_2^{(LL)}(f,f)$ yields

$$S_2^{(LL)}(f,f) = -j2\pi(2f)C_2Z_0^{3/2}Z_L(2f) \cdot \frac{1 + 2S_1^{(LL)}(f) + [S_1^{(LL)}(f)]^2}{Z_L(2f) + Z_0}. \quad (109)$$

Equation (109) clearly reveals how the reflection coefficient of the 1-port and the linearized impedance of the load enter into determination of the second-order response. Observe that the reflection coefficient and the linearized load impedance are functions of frequency.

In general, the nonlinear scattering functions $S_n^{(LL)} \cdot (f_1,\cdots,f_n)$ can be obtained from (102) by assuming an excitation of the form

$$\alpha^{(L)}(t) = \sum_{m=1}^{n} e^{j2\pi f_m t} \quad (110)$$

in conjunction with the harmonic input method.

V. Conclusion

Scattering variables are convenient to use when analyzing microwave systems. This paper has demonstrated that the conventional linear scattering parameter theory is a special case of a more general theory applicable to nonlinear systems. In addition, scattering variables can be used to simplify the characterization of a nonlinear multiport when the ports are matched to the reference impedance. The nonlinear scattering functions facilitate the calculation of power in nonlinear distortion products at microwave frequencies.

References

[1] N. Wiener, *Nonlinear Problems in Random Theory*. Cambridge, MA.: M.I.T. Press, 1959.
[2] E. Bedrosian and S. O. Rice, "The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and Gaussian inputs," *Proc. IEEE*, vol. 59, pp. 1688–1707, Dec. 1971.
[3] J. Bussgang, L. Ehrman, and J. Graham, "Analysis of nonlinear systems with multiple inputs," *Proc. IEEE*, vol. 62, pp. 1088–1119, Aug. 1974.
[4] S. Narayanan, "Application of Volterra series to intermodulation distortion analysis of transistor feedback amplifier," *IEEE Trans. Circuit Theory*, vol. CT-17, pp. 518–527, Nov. 1970.
[5] M. Brilliant, "Theory of the analysis of nonlinear systems," M.I.T. Research Lab. of Electronics, Cambridge, MA, Tech. Rep. 345, Mar. 3, 1958.
[6] E. Kuh and R. Rohrer, *Theory of Linear Active Networks*. San Francisco, CA.: Holden-Day, Inc., 1967.
[7] N. Balabanian and T. Bickart, *Electrical Network Theory*. New York: John Wiley and Sons, 1969.
[8] P. Penfield, "Exact cascading for nonlinear networks," Appendix A of RADC Technical Report 74-68, Rome Air Development Center, Rome, NY, Mar. 1974.

Experimental and Theoretical Studies on Electromagnetic Fields Induced Inside Finite Biological Bodies

BHAG SINGH GURU, STUDENT MEMBER, IEEE, AND KUN-MU CHEN, FELLOW, IEEE

***Abstract*—The total electric field inside some simulated biological bodies induced by an electromagnetic wave has been quantified by the recently developed tensor integral equation method and measured by an insulated probe. In general, the induced electric field inside a biological body was found to be quite complicated. An excellent agreement was obtained between theory and experiment.**

I. Introduction

IN THE STUDY of the interaction of electromagnetic radiation with biological bodies, the key physical quantity which determines the bioeffects on the body is the actual electromagnetic field induced inside the body by the incident electromagnetic wave. Since a biological body is usually a heterogeneous finite body with an irregular shape, the quantification of the internal electromagnetic fields becomes a difficult problem. For mathematical simplicity, commonly used models are the plane slab [1], [2], the sphere [3]–[5], the cylinder [6], and the spheroids [7], [8]. Although these simple models provide estimates of the internal electromagnetic fields, the results have limited applicability to the biological bodies with irregular shapes and illuminated by a microwave.

Recently, Livesay and Chen [9] have developed a theoretical method called the tensor integral equation method which can be used to quantify the internal electric field induced by an incident electromagnetic wave inside arbitrarily shaped biological bodies. This method has been utilized to quantify the induced electric field inside some simulated biological bodies illuminated by a microwave. The same induced electric field has been measured by a small insulated probe. In general, the induced electric field inside the body was found to be quite complicated even though the incident

Manuscript received October 23, 1975; revised January 19, 1976. This work was supported by the National Science Foundation under Grant ENG 74-12603.

The authors are with the Department of Electrical Engineering and Systems Science, Michigan State University, East Lansing, MI 48824.