A Generalized Theory and New Calibration Procedures for Network Analyzer Self-Calibration

Hermann-Josef Eul, *Member, IEEE*, and Burkhard Schiek, *Member, IEEE*

***Abstract*—A general theory for performing network analyzer calibration is presented. New calibration procedures are derived which allow for partly unknown standards. The most general procedure derived is called TAN and allows for five unknown parameters in the three calibration standards. The values of the unknown parameters are determined during the calibration procedure via eigenvalue conditions. The good performance of all the procedures is shown by measured results.**

I. Introduction

The accuracy of network analyzers is enhanced by calibrating the setup at its measurement ports. Usually this is performed by applying the well-known 12-term procedure [1], [2], employing the standards thru, match, short and open.[1] While the 12-term procedure depends only on fully known standards, there are two other methods in use allowing for partly unknown standards. They were introduced in [3] as TSD (thru, short, delay) and in [4] as TRL (thru, reflect, line) and they can be implemented in a double reflectometer as shown in Fig. 1 as well as in other configurations [5].

Besides the above-mentioned advantages TRL and TSD have significant drawbacks because the electrical length of the line must differ from multiples of the half-wavelength. Therefore they show a lower band limit and, depending on the frequency, periodically repeating bands of unreliable calibration.

In [6] the results of [4] and [3] were derived in a different mathematical manner. In [7] a family of new self-calibration procedures was proposed allowing for a higher number of unknown parameters in the standards. Here we will present the theoretical background of these concepts, along with certain generalizations and experimental results. The criteria for the kind and the number of these unknown parameters are deduced from eigenvalue solutions.

II. Theory

A. Basic Model

As shown in Fig. 1, the main element of each reflectometer is a four-port, which normally contains the directional couplers or bridges to separate the waves propagating toward and emerging from the device under test (DUT). We will regard it as an arbitrary four-port described by its scattering parameters, e.g.

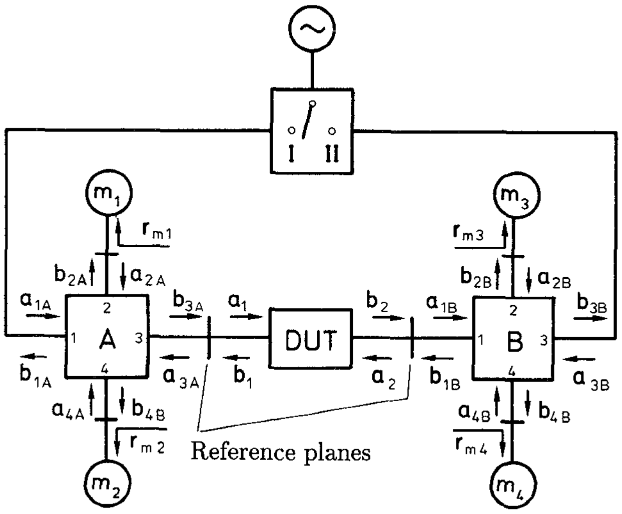

Fig. 1. Block diagram of a double reflectometer.

$$\begin{pmatrix} b_{1A} \\ b_{2A} \\ b_{3A} \\ b_{4A} \end{pmatrix} = \begin{pmatrix} S_{11A} & S_{12A} & S_{13A} & S_{14A} \\ S_{21A} & S_{22A} & S_{23A} & S_{24A} \\ S_{31A} & S_{32A} & S_{33A} & S_{34A} \\ S_{41A} & S_{42A} & S_{43A} & S_{44A} \end{pmatrix} \begin{pmatrix} a_{1A} \\ a_{2A} \\ a_{3A} \\ a_{4A} \end{pmatrix} \quad (1)$$

for the left part of the system. As boundary conditions we use

$$a_{2A} = r_{m1} b_{2A} \quad \text{and} \quad a_{4A} = r_{m2} b_{4A} \quad (2)$$

where the r_{mi} are the reflection coefficients of the measurement channels connected to the four-ports. To consider mismatch losses, mixer conversion, IF amplifier transmission, etc., we introduce complex conversion factors η_i, relating the RF signals to the readings m_i of the A/D converters sampling the complex IF signal at the very end of the measurement channels:

$$m_1 = \eta_1 b_{2A} \quad \text{and} \quad m_2 = \eta_2 b_{4A}. \quad (3)$$

Using (2) and (3) we eliminate a_{1A}, a_{2A}, a_{4A}, b_{1A}, b_{2A}, and b_{4A}, obtaining

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} a_{3A} \\ b_{3A} \end{pmatrix} \quad (4)$$

where the elements A_{ij} are functions of the S_{ijA} and the

Manuscript received March 8, 1990; revised November 13, 1990. This work was supported by the Deutsche Forschungsgemeinschaft (DFG).

The authors are with the Institut für Hoch- und Höchstfrequenztechnik der Ruhr-Universität Bochum, Universitätsstrasse 150, D-4630 Bochum 1, Germany.

IEEE Log Number 9042350.

[1] Acronyms are in use as well, for example SOLT (short, open load, through) and OSLT. However these two will not be used here to avoid a conflict. They use the letter "L" for the load, which has already been claimed by self-calibration procedures for the line standard [4].

0018-9480/91/0400-0724$01.00 ©1991 IEEE

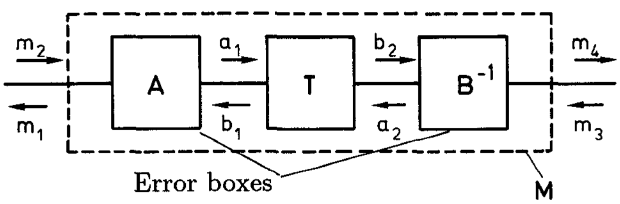

Fig. 2. To the interpretation of (9).

η_i, r_{mi} only. Equation (4) fully describes the mapping of the waves of interest, a_3 and b_3, onto the measurements available, m_1 and m_2. In a similar manner we derive the mapping for the right reflectometer,

$$\begin{pmatrix} m_3 \\ m_4 \end{pmatrix} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} b_{1B} \\ a_{1B} \end{pmatrix}. \tag{5}$$

The two reflectometers are linked twice. One linkage is provided by the three-port (Fig. 1) controlling the energy splitting. In practical setups it is realized as an RF switch but any other power-splitting behavior is allowed as long as independent readings are ensured. The behavior needs neither to be known nor to be reproducible because it does not affect the A_{ij} and B_{ij}.

The second linkage of the reflectometers is provided by a device connected to the measurement ports. It shall be denoted in transmission parameters,

$$\begin{pmatrix} b_1 \\ a_1 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \tag{6}$$

and leads due to the identities $a_1 = b_{3A}$, $b_1 = a_{3A}$, $a_2 = b_{1B}$, and $b_2 = a_{1B}$ to the vector equation

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}^{-1} \begin{pmatrix} m_3 \\ m_4 \end{pmatrix} \tag{7}$$

which must be satisfied by the measurements. If the three-port is turned to its second state it provides a further set of readings m_i' which must satisfy (7) as well. The two vector equations are combined to a matrix equation, yielding

$$\begin{pmatrix} m_1 & m_1' \\ m_2 & m_2' \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \cdot \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}^{-1} \begin{pmatrix} m_3 & m_3' \\ m_4 & m_4' \end{pmatrix} \tag{8}$$

and finally

$$\boldsymbol{M} = \boldsymbol{A}\boldsymbol{T}\boldsymbol{B}^{-1} \tag{9}$$

with the measurement matrix

$$\boldsymbol{M} = \begin{pmatrix} m_1 & m_1' \\ m_2 & m_2' \end{pmatrix} \begin{pmatrix} m_3 & m_3' \\ m_4 & m_4' \end{pmatrix}^{-1}. \tag{10}$$

With a knowledge of $\boldsymbol{A}$ and $\boldsymbol{B}$, which has to be provided by a calibration procedure, one is able to evaluate the parameters $\boldsymbol{N}_x$ of a DUT from its measurement matrix $\boldsymbol{M}_x$ via

$$\boldsymbol{N}_x = \boldsymbol{A}^{-1}\boldsymbol{M}_x\boldsymbol{B}. \tag{11}$$

Equation (9) can be interpreted as a cascade of two-ports (Fig. 2). Thus the elimination process from (1) to (4) is called the four-port to two-port reduction and provides justification for the heuristic approach via error-boxes (e.g. [4], [3], [1], [11], [6]).

It should be stressed that the cascade of two-ports is only an interpretation, because

1) m_2 and m_3 are not waves propagating towards the four-port;
2) the error two-ports do not exhibit the properties of real two-ports. For example, even if the four-ports were only assembled with reciprocal elements, reciprocity could not be applied for $\boldsymbol{A}$ and $\boldsymbol{B}$ as their determinants generally do not equal unity, i.e., $\det \boldsymbol{A} \neq \det \boldsymbol{B} \neq 1$.

It is also possible to treat cascades of real two-port networks, e.g. de-embedding problems, in the same way as is described here for network analyzer calibration. In the de-embedding case additional assumptions can sometimes be made such as reciprocity or symmetry, leading to special solutions [9]. The investigations presented here are mainly focused on network analyzer calibration, and no assumptions of this kind will be made. However, the results can be applied to de-embedding procedures as well.

B. Calibration

As will be shown in the following, the setup can be calibrated by measuring three different standards

$$\boldsymbol{M}_1 = \boldsymbol{A}\boldsymbol{N}_1\boldsymbol{B}^{-1} \tag{12a}$$

$$\boldsymbol{M}_2 = \boldsymbol{A}\boldsymbol{N}_2\boldsymbol{B}^{-1} \tag{12b}$$

$$\boldsymbol{M}_3 = \boldsymbol{A}\boldsymbol{N}_3\boldsymbol{B}^{-1}. \tag{12c}$$

Here the transmission matrices $\boldsymbol{N}_1$, $\boldsymbol{N}_2$, and $\boldsymbol{N}_3$, of the standards are, for the moment, supposed to be known. Combining (12a) and (12b) we eliminate $\boldsymbol{B}$, yielding

$$\boldsymbol{P} = \boldsymbol{A}^{-1}\boldsymbol{Q}\boldsymbol{A} \quad \text{with} \quad \boldsymbol{P} = \boldsymbol{N}_2\boldsymbol{N}_1^{-1} \quad \text{and} \quad \boldsymbol{Q} = \boldsymbol{M}_2\boldsymbol{M}_1^{-1}. \tag{13}$$

Further straightforward algebraic treatment leads to a linear equation for the calibration constants

$$\hat{\boldsymbol{C}}\underline{A} = \underline{0} \tag{14}$$

with

$$\hat{\boldsymbol{C}} = \left(\begin{array}{c|c} Q_{11}\boldsymbol{P}^{t^{-1}} - \boldsymbol{E} & Q_{12}\boldsymbol{P}^{t^{-1}} \\ \hline Q_{21}\boldsymbol{P}^{t^{-1}} & Q_{22}\boldsymbol{P}^{t^{-1}} - \boldsymbol{E} \end{array}\right),$$

$$\underline{A} = \begin{pmatrix} A_{11} \\ A_{12} \\ A_{21} \\ A_{22} \end{pmatrix}, \qquad \boldsymbol{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{15}$$

where $\boldsymbol{P}^{t^{-1}}$ is the inverse matrix of the transposed $\boldsymbol{P}$.

Because (14) is a homogeneous equation there must be at least a one-dimensional ambiguity in the solution. This can be accepted, as will be shown later. Actually the ambiguity is a two-dimensional one, which can be shown as follows. Due to (13), $\boldsymbol{P}$ and $\boldsymbol{Q}$ are similar matrices. Similar matrices have several invariances, i.e., equal traces and determinants or, what is equivalent, equal eigenvalues λ_1, λ_2 [8]:

$$\operatorname{eig}(P) = \operatorname{eig}(Q) = \lambda_{1,2}, \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{16}$$

where eig( ) is the eigenvalue operator. Λ denotes the matrix of eigenvalues which is related to P and Q via the transformations

$$\Lambda = X^{-1}PX \quad \text{and} \quad \Lambda = Y^{-1}QY. \tag{17}$$

The columns of the transforming matrices X and Y are the eigenvectors of P and Q, respectively. As P and Q are assumed to be known, the eigenvectors can be evaluated. Because eigenvectors are determined except for an arbitrary factor, we denote

$$X = (\alpha_1\vec{x}_1, \alpha_2\vec{x}_2) \quad \text{and} \quad Y = (\alpha_3\vec{y}_1, \alpha_4\vec{y}_2). \tag{18}$$

Eliminating Λ in (17) gives

$$P = XY^{-1}QYX^{-1}. \tag{19}$$

Introducing (18) and comparing with (13), it becomes obvious that A remains with an ambiguity of second order. Thus $\hat{C}$ only contains two independent equations, rank $\hat{C} \leq 2$.

In order to proceed with the calibration, (12a) and (12c) are combined in the same way to yield

$$U = A^{-1}VA \quad \text{with} \quad U = N_3N_1^{-1} \quad \text{and} \quad V = M_3M_1^{-1} \tag{20}$$

resulting in

$$\hat{\hat{C}}\underline{A} = \underline{0} \quad \text{with} \quad \hat{\hat{C}} = \left(\begin{array}{c|c} V_{11}U^{t^{-1}} - E & V_{12}U^{t^{-1}} \\ \hline V_{21}U^{t^{-1}} & V_{22}U^{t^{-1}} - E \end{array}\right) \tag{21}$$

with a rank of 2 as well. Finally we extract the independent equations in (14) and (21), e.g. by using a Gaussian elimination procedure, and combine them to build C:

$$C\underline{A} = \underline{0} \tag{22}$$

which allows for the determination of A except for a common factor α, the one-dimensional ambiguity. Let $\tilde{A}$ be an arbitrary solution of the infinity of possible solutions, which are all proportional to the original A, $A = \alpha\tilde{A}$. Substituting B in (11)

$$N_x = \frac{1}{\alpha}\tilde{A}^{-1}M_xM_1^{-1}\alpha\tilde{A}N_1 = \tilde{A}^{-1}M_xM_1^{-1}\tilde{A}N_1 \tag{23}$$

proves that $\tilde{A}$ is sufficient for the system error removal. It should be noted that this degree of freedom can be used to enforce the determinant of $\tilde{A}$ to unity. Even under this assumption in general $\det\tilde{B} \neq 1$ remains.

C. Exploiting Redundancies (Self-Calibration)

While there remain seven unknowns to be determined in $\tilde{A}$ and $\tilde{B}$, the measurement of the three two-port standards yields 12 equations. This redundancy is the reason that it is not necessary to know all the parameters of all the standards. Ultimately it is possible to evaluate five different quantities. In the following it will be shown how to exploit this redundancy.

For one of the standards, e.g. N_1, all parameters are supposed to be known. The second standard is assumed to be partly unknown. This is considered by denoting its transmission parameters as known functions of the unknowns:

TABLE I

N_1	Fully known	N_1
N_2	Maximum of two free parameters	$N_2(x_1, x_2)$
N_3	Maximum of three free parameters	$N_3(y_1, y_2, y_3)$

$$N_2(\underline{x}) = \begin{pmatrix} N_{2,11}(\underline{x}) & N_{2,12}(\underline{x}) \\ N_{2,21}(\underline{x}) & N_{2,22}(\underline{x}) \end{pmatrix}$$

$$\text{with} \quad \underline{x} = (x_1, \cdots, x_n)^t, \quad n \in \mathbb{N}. \tag{24}$$

Therefore P is also a function of $\underline{x}$. Recalling the similarity of Q and $P(\underline{x})$, the most effective way is to exploit the trace and determinant invariances, i.e.,

$$\operatorname{trace}(P(\underline{x})) = \operatorname{trace}(Q) \quad \text{and} \quad \det(P(\underline{x})) = \det(Q) \tag{25}$$

to develop two equations which must be satisfied by the parameters of the standards and the measurements

$$P_{11}(\underline{x}) + P_{22}(\underline{x}) = Q_{11} + Q_{22} \tag{26}$$

$$P_{11}(\underline{x})P_{22}(\underline{x}) - P_{12}(\underline{x})P_{21}(\underline{x}) = Q_{11}Q_{22} - Q_{12}Q_{21}. \tag{27}$$

These two equations allow for the determination of two unknown parameters $\underline{x} = (x_1, x_2)^t$ in N_2.

The matrix N_3 is treated in the same manner:

$$N_3(\underline{y}) = \begin{pmatrix} N_{3,11}(\underline{y}) & N_{3,12}(\underline{y}) \\ N_{3,21}(\underline{y}) & N_{3,22}(\underline{y}) \end{pmatrix}$$

$$\text{with} \quad \underline{y} = (y_1, \cdots, y_n)^t, \quad n \in \mathbb{N}. \tag{28}$$

The similarity of the matrices U and V provides two equivalent equations as well:

$$U_{11}(\underline{y}) + U_{22}(\underline{y}) = V_{11} + V_{22} \tag{29a}$$

$$U_{11}(\underline{y})U_{22}(\underline{y}) - U_{12}(\underline{y})U_{21}(\underline{y}) = V_{11}V_{22} - V_{12}V_{21}. \tag{29b}$$

Two additional equations are derived by combining N_3 and N_2, which now is known from the previous step:

$$R_{11}(\underline{y}) + R_{22}(\underline{y}) = W_{11} + W_{22} \tag{30a}$$

$$\left(R_{11}(\underline{y})R_{22}(\underline{y}) - R_{12}(\underline{y})R_{21}(\underline{y}) = W_{11}W_{22} - W_{12}W_{21}\right) \tag{30b}$$

with abbreviations similar to (13):

$$R = N_3N_2^{-1} \quad \text{and} \quad W = M_3M_2^{-1}. \tag{31}$$

Equation (30b) is set in parentheses because it provides the same information as (29b). Thus three equations remain, limiting the degrees of freedom in $N_3(\underline{y})$ to three, $\underline{y} = (y_1, y_2, y_3)^t$. Therefore the standards have to meet the requirements given in Table I.

This maximum of five unknowns in the standards and the number of seven effective calibration standards $\tilde{A}_{ij}, \tilde{B}_{ij}$ fully exploit the twelve equations in (12). Based on the results condensed in Table I, calibration procedures with practical usefulness will be derived.

D. The Procedures TAN and TLN

The first two-port must be fully known. The simplest two-port meeting this condition is the thru connection of the two measurement ports:

$$N_1=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}. \tag{32}$$

All procedures which use this standard shall have a "T" as the first letter in their acronym. Without loss of generality they are the basis for the following explanations. However, "L" procedures are possible and will be explained later.

For the second standard, a transmission line of known characteristic impedance is one realization:

$$N_2=\frac{1}{e^{-\gamma l}(1-\rho^2)}\begin{pmatrix}e^{-2\gamma l}-\rho^2 & \rho(1-e^{-2\gamma l})\\ -\rho(1-e^{-2\gamma l}) & 1-\rho^2 e^{-2\gamma l}\end{pmatrix}$$

$$\text{with}\quad \rho=\frac{Z_L-Z_0}{Z_L+Z_0} \tag{33}$$

where Z_L denotes the actual characteristic impedance of the line and Z_0 is the reference impedance. The quantity γ is the complex propagation constant of the transmission line and l its mechanical length.

Substituting (32) and (33) into (26), we find

$$e^{-\gamma l}+e^{+\gamma l}=Q_{11}+Q_{22} \tag{34}$$

which is not a function of ρ, whether known or unknown. Equation (27) provides no further information because reciprocity is already introduced by (33). Thus it cannot be applied to evaluate the characteristic impedance of the transmission line. But its transmission coefficient may be left unknown since it can be determined by (34) whether the impedance is known or not. However, the solution of (34) is ambiguous as $-\gamma l$ and $+\gamma l$ are interchangeable. To find the true value, the passivity criterion requiring

$$|e^{-\tilde{\gamma} l}|\leqslant 1 \tag{35}$$

could be used. In practice this criterion is not reliable because of the low line loss; even small measurement errors such as noise or quantization errors are able to influence the small deviation from unity in a way that leads to the wrong decision. Thus the estimation of the electrical length of the line is recommended:

$$\operatorname{sign}(\arg(e^{\pm\tilde{\gamma} l}))=\operatorname{sign}(\varphi_{\text{estimate}}). \tag{36}$$

This criterion is very reliable apart from the regions of uncertainty around 0° and 180°, which are of minor interest as they must be avoided anyway. This is due to the fact that the second standard has to be different from the first and the line has to satisfy the conditions

$$e^{-\gamma l}\neq \pm 1 \qquad \text{resp.}\quad l\neq n\frac{\lambda}{2}\Rightarrow l\approx(2n+1)\frac{\lambda}{4},\qquad n\in\mathbb{N}. \tag{37}$$

In practice it is useful to define the line to be matched, $\rho=0$, thus implying that its characteristic impedance becomes the reference for the calibration.

Although transmission lines can be produced with high accuracy they show the drawback of limited bandwidth. At lower frequencies the line tends to become too long and at higher frequencies it exhibits, depending on the frequency, periodically repeating ranges of unreliable calibration.

If a transmission line is used as second calibration standard it will be denoted by an "L" in the procedure's acronym.

While maintaining the property of being matched, we will show that realizations other than transmission lines are possible and circumvent the problems shown above. As the standard is allowed to be nonreciprocal we start with the transmission matrix

$$N_2=\begin{pmatrix}N_{2,11} & 0\\ 0 & N_{2,22}\end{pmatrix}. \tag{38}$$

The substitution of (38) into (26) yields the two equations

$$N_{2,11}+N_{2,22}=Q_{11}+Q_{22}\quad\text{and}\quad N_{2,11}N_{2,22}=\det Q. \tag{39}$$

Due to the symmetry in (39) it is obvious that it holds for interchanged $N_{2,11}$ and $N_{2,22}$ as well. However these parameters may be left unknown and will be calculated via (39). For passivity of N_2 the evaluated elements $\tilde{N}_{2,11}$ and $\tilde{N}_{2,22}$ must satisfy

$$|\tilde{N}_{2,11}|\leqslant 1\quad\text{and}\quad |\tilde{N}_{2,22}|\geqslant 1 \tag{40}$$

a very reliable criterion for the selection, if N_2 is designed to have significant losses. Therefore the need for prior information concerning N_2 is avoided and the argument of the transmission is allowed to have any value.

We will denote this choice of the second calibration standard with an "A" for attenuation and summarize the advantages:

1) no estimation of the transmission required,
2) attenuation and phase shift may be arbitrary and unknown,
3) no lower band limit,
4) no periodically repeating ranges of unreliable calibration.

It should be noted that in either case, "L" or "A", reciprocity is not required, but normally is given. Thus the procedures provide a measure of the calibration performance by building a first figure of merit using the values evaluated by the self-calibration procedure, such as

$$F_1=1-\det\tilde{N}_2. \tag{41}$$

For a good calibration F_1 should be as small as possible.

In the last calibration step standard N_3 is measured. We substitute (32) and (38) in (29) and (30), yielding

$$N_{3,11}+N_{3,22}=V_{11}+V_{22} \tag{42}$$

$$N_{2,22}N_{3,11}+N_{2,11}N_{3,22}=N_{2,11}N_{2,22}(W_{11}+W_{22}) \tag{43}$$

$$N_{3,11}N_{3,22}-N_{3,12}N_{3,21}=\det V \tag{44}$$

where N_2 is known from the step before. Equations (42)–(44) must be satisfied by $N_{3,11}$, $N_{3,22}$, and $N_{3,12}N_{3,21}$. Thus these parameters may be left unknown and evaluated via (42)–(44). The product can be separated if either of the factors is known or if N_3 is symmetrically reflecting. In the latter case, which is of practical usefulness, one has to remove the sign ambiguity of the square root by an estimation of the sign of the reflection. This is the only piece of information needed.

As the first two standards are well matched, for proper calibration, N_3 should be highly reflecting. The standard may be nonreciprocal as well, but in the reciprocal case the procedure provides a further figure of merit,

$$F_2 = 1 - \det \tilde{N}_3 \tag{45}$$

which also should be as small as possible.

In the following this realization of N_3 will be referred to by an "N" and can be combined with both the line (TLN) and the attenuation (TAN). Here TLN is the special case of vanishing attenuation in the more general TAN procedure. The order of the letters in the acronym is chosen according to the increasing number of unknown parameters in the standards. Unfortunately, in the following this leads to a conflict with the well-known TRL procedure.

E. The Procedures TAR, TAS, TLR, and TLS

The following procedures use two-ports without transmission for the third calibration standards, actually one-ports. This standard can be realized either by using two reflections, which are sufficiently equal, or by measuring one reflection subsequently at each measurement port. The latter is assumed for the mathematical derivation. Based on (4) and (5) we define

$$\frac{m_1}{m_2} = \frac{A_{11}r_R + A_{12}}{A_{21}r_R + A_{22}} \stackrel{\text{def}}{=} \Gamma_A \tag{46a}$$

$$\frac{m_3'}{m_4'} = \frac{B_{11} + B_{12}r_R}{B_{21} + B_{22}r_R} \stackrel{\text{def}}{=} \Gamma_B \tag{46b}$$

with

$$r_R = \frac{a_{3A}}{b_{3A}} \quad \text{and} \quad r_R = \frac{a_{1B}'}{b_{1B}'}. \tag{47}$$

Taking (12a) the B_{ij} are eliminated in (46b). Then using the abbreviations

$$c_1 = M_{1,22} + M_{1,21}\Gamma_B \qquad c_2 = M_{1,12} + M_{1,11}\Gamma_B \tag{48}$$

equations (46) can be reorganized to fit into the matrix notation (14) and assembled with the two independent rows of (14) to provide an equivalent to (22):

$$C(r_R)A = 0 \quad \text{with}$$

$$C(r_R) = \left(\begin{array}{cccc} \multicolumn{4}{c}{\text{the two independent}} \\ \multicolumn{4}{c}{\text{rows of } \hat{C} \text{ in (15)}} \\ \hline r_R & 1 & -\Gamma_A r_R & -\Gamma_A \\ c_1 & c_1 r_R & -c_2 & -c_2 r_R \end{array} \right). \tag{49}$$

Now we have to decide whether we use a TxS or a TxR procedure.

First we treat the more general TxR case. Apart from the trivial solution $A_{ij} = 0$ this homogeneous equation is only satisfied if the determinant vanishes. Thus,

$$\det C(r_R) = 0 \Rightarrow r_R = \left[\frac{(c_1C_{13} + c_2C_{11})(C_{24} + C_{22}\Gamma_A)}{(c_1C_{24} + c_2C_{22})(C_{13} + C_{11}\Gamma_A)} \right]^{1/2} \tag{50}$$

must be fulfilled by the reflection r_R of the third calibration standard. In (50) the C_{ij} are elements of C and the c_i are given by (48). If (50) is used to determine the unknown reflection, the sign ambiguity of the square root must be removed. This small piece of prior information has to be provided by the user. Thus the preferred realization is a short or an open circuit since they allow for an easy and reliable sign estimation. After the calculation of $\tilde{r}_R$, C in (49) is determined and the calibration proceeds with the evaluation of $\tilde{A}$ as described above.

In order to prove the capability of self-calibration, a well-defined calibration standard is applied and a third figure of merit,

$$F_3 = 1 - \frac{\tilde{r}_R}{r_R} \tag{51}$$

is introduced, where r_R is the value from the data sheet and $\tilde{r}_R$ the value evaluated via (50). The performance will not degrade if reflections of minor quality are used. Procedures using an unknown reflection are indicated by an "R" in the scheme, actually TAR and TLR. The latter (TLR) has previously been mentioned in the literature as TRL[2] in [4] and appears as a special case of this general concept.

Second we will treat the TxS procedures. They assume the third standard as to be known, simply but not necessarily to be a short. Due to the knowledge of $r_R = -1$ (49) is overdetermined because equations (46) are not independent of each other. Thus measurement requirements can be reduced by omitting one of them; i.e., the short has *only* to be measured at *one* measurement port and step (50) is skipped. Procedures using the short "S" are TAS and TLS. Here TLS is quite similar to the TSD procedure [3], but it overcomes the necessity of connecting the short at both ports, as is required in [3] and [6].

F. The Procedures TMN, TMR, and TMS

First we will treat the TMR and TMS procedures, which employ two standards without transmission; i.e., the second and the third one become one-ports. As the second standard only has its degrees of freedom in transmission, its reflection has to be known. Without loss of generality we assume one reflection r_M measured at each measurement port in turn. The third standard is the same as in the other TxR procedures already explained. With the measurements

$$\frac{m_{11}}{m_{21}} = \frac{A_{11}r_M + A_{12}}{A_{21}r_M + A_{22}} \stackrel{\text{def}}{=} \Gamma_{A1}$$

$$\frac{m_{31}'}{m_{41}'} = \frac{B_{11} + B_{12}r_M}{B_{21} + B_{22}r_M} \stackrel{\text{def}}{=} \Gamma_{B1} \tag{52}$$

$$\frac{m_{12}}{m_{22}} = \frac{A_{11}r_R + A_{12}}{A_{21}r_R + A_{22}} \stackrel{\text{def}}{=} \Gamma_{A2}$$

$$\frac{m_{32}'}{m_{42}'} = \frac{B_{11} + B_{12}r_R}{B_{21} + B_{22}r_R} \stackrel{\text{def}}{=} \Gamma_{B2} \tag{53}$$

and the abbreviations due to (12a):

$$c_1 = M_{1,22} + M_{1,21}\Gamma_{B1} \qquad c_2 = M_{1,12} + M_{1,11}\Gamma_{B1}$$

$$c_3 = M_{1,22} + M_{1,21}\Gamma_{B2} \qquad c_4 = M_{1,12} + M_{1,11}\Gamma_{B2} \tag{54}$$

[2]It is not the intention to alter the common, well-established terminology for this special case. But is it is proposed to maintain the scheme of notation for the whole family of procedures for reasons that are a result of the logic of the derivation explained above.

we obtain a homogeneous system of equations:

$$C(r_R)\underline{A} = 0 \quad \text{with}$$

$$C(r_R) = \begin{pmatrix} r_M & 1 & -\Gamma_{A1} r_M & -\Gamma_{A1} \\ r_R & 1 & -\Gamma_{A2} r_R & -\Gamma_{A2} \\ c_1 & c_1 r_M & -c_2 & -c_2 r_M \\ c_3 & c_3 r_R & -c_4 & -c_4 r_R \end{pmatrix}. \tag{55}$$

In practice the known reflection is realized as a matched load $r_M = 0$, leading to an "M" in the acronym. Depending on the frequency and on the accuracy desired, a fixed load or a sliding load with a circle fitting procedure can be applied.

The further steps are as described for TAR with

$$\det C(r_R) = 0 \Rightarrow r_R = \left[\frac{(c_1 c_4 - c_2 c_3)(\Gamma_{A1} - \Gamma_{A2})}{(c_2 - c_1 \Gamma_{A2})(c_4 - c_3 \Gamma_{A1})} \right]^{1/2} \tag{56}$$

as the counterpart to (50).

As with the TAR procedure, in the TMR procedure there exists a version which replaces the unknown reflection standard with a known reflection. This procedure, TMS, has the advantage that the short (or any other known reflection) has to measured at *only one* measurement port.

Finally we treat the TMN procedure. From a mathematical point of view, this method can be regarded as a TAR procedure with switched meanings of the second and the third standard and rearranged degrees of freedom. To apply the TMN procedure, use the TAR algorithm but connect the N-standard instead of the A-standard and assume, for the moment, that it might be well matched. Thus the calibration is related to an unknown reference impedance introduced by the standard N. Then employ the M-standard instead of the R-standard and solve for the reflection via (50). Having this value, one is able to evaluate the above-mentioned reference impedance. With this information the parameters of N are renormalized to the desired reference impedance. Now proceed with TAR as usual.

Similar to the TAx procedures the TMx show no inherent band limit. However, they can be viewed as a special case of infinite attenuation in TAx.

G. *Deriving Lxx Procedures from Txx*

Another convenient realization of the first calibration standard is a known, well-matched transmission line:

$$N_1 = \begin{pmatrix} e^{-\gamma_1 l_1} & 0 \\ 0 & e^{+\gamma_1 l_1} \end{pmatrix}. \tag{57}$$

For the special case of TRL [4] this idea has been described in detail in [10].

For completeness in the following a brief description will be given of how to adopt this idea to the whole family of calibration procedures discussed here.

The calibration steps are performed as in Txx without any changes, yielding the matrices $\tilde{\tilde{A}}$ and $\tilde{\tilde{B}}$. Obviously the reference planes will be located in the center of l_1. Thus the desired matrices $\tilde{A}$ and $\tilde{B}$ are simply found through

$$\begin{aligned} \tilde{A} &= \tilde{\tilde{A}} L^{-1} \\ \tilde{B} &= \tilde{\tilde{B}} L \end{aligned} \quad \text{with } L = \begin{pmatrix} e^{-\frac{1}{2}\gamma_1 l_1} & 0 \\ 0 & e^{+\frac{1}{2}\gamma_1 l_1} \end{pmatrix}. \tag{58}$$

To perform this transformation either we have to know $\gamma_1 l_1$ directly or we must know r_R to allow the evaluation of $\gamma_1 l_1$ via

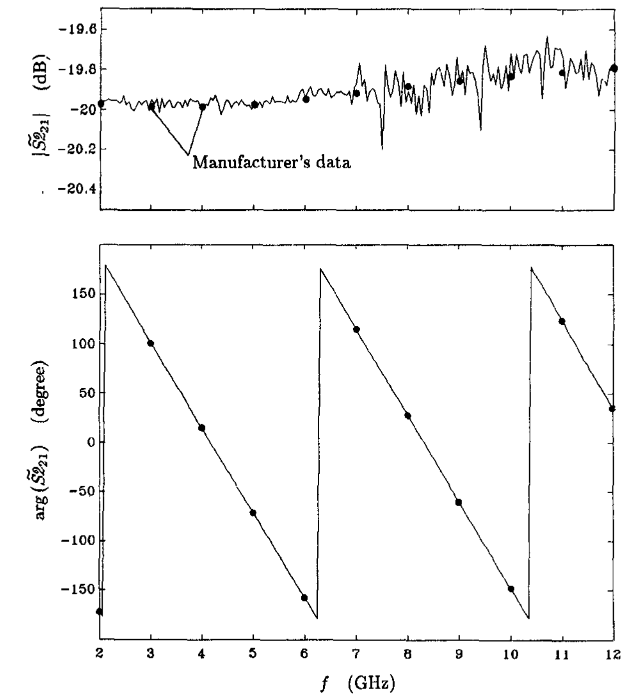

Fig. 3. The transmission of the second standard evaluated via self-calibration (solid line) in comparison with the manufacturer's data (black dots).

$$\gamma_1 l_1 = \ln\left(\frac{\tilde{r}_R}{r_R}\right) \tag{59}$$

where $\tilde{r}_R$ is the value from the TxR procedure. For TxN, $\tilde{S}_{3,11}$ takes the position of $\tilde{r}_R$. In the special case where N_2 is a transmission line with the same propagation constant as N_1, but different length (LLx), a knowledge of $\arg(r_R)$ is sufficient [10].

III. Experimental Results

The following measurements have been carried out on a Hewlett Packard HP8510 with APC-7 standards over a frequency range of 2 to 12 GHz. Since, to the authors' knowledge [12], the equipment has no capability for the full variety of the new calibration procedures, the raw data for the m_i (the unratioed values a_i and b_i can be taken for that purpose) have been read out and processed on an external personal computer.

The figure of merit F_1 is the first indicator for the calibration reliability we examine. It is already available after the second calibration step of each xAx and xLx procedure. In this TAx example it remains under 0.035 over the full bandwidth, promising a good calibration. In order to check the data for F_1 against other measurements, the transmission $\tilde{S}_{2,21}$ calculated via self-calibration is plotted in Fig. 3 as well as the manufacturer's data, the black dots. The figure shows very good agreement.

Next we test the self-calibration behavior in determining the unknown reflection r_R. Therefore a precision short is used and the figure of merit F_3 is evaluated, which is well below 0.02. It is instructive to plot the angle of $\tilde{r}_R / r_R$ (Fig. 4), which can be compared directly with the manufacturer's

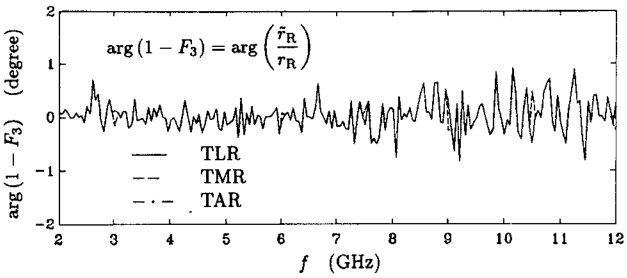

Fig. 4. The difference between the angle of $\tilde{r}_R$ evaluated via self-calibration and the ideal value r_R. Specified uncertainty, 0.5°.

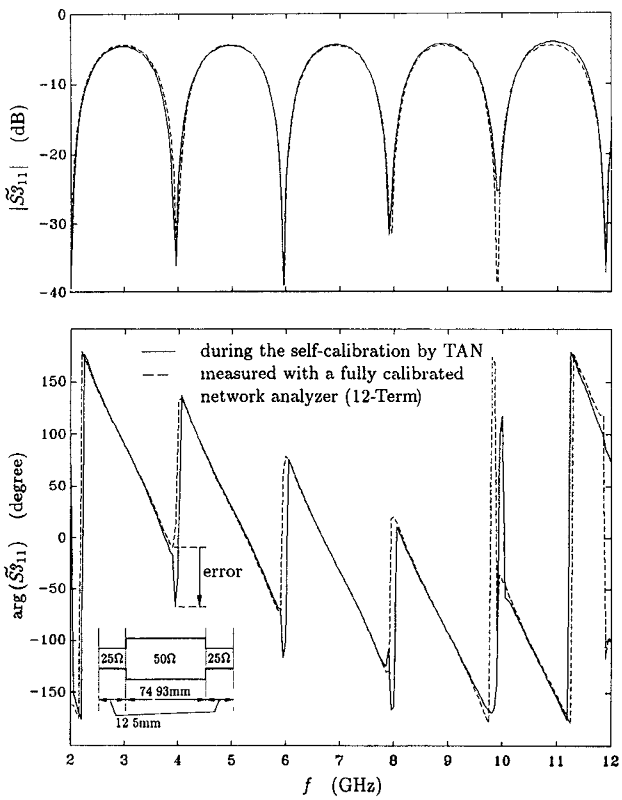

Fig. 5. The reflection of the third standard evaluated via self-calibration in comparison with a measurement with a fully calibrated network analyzer. The quality of the self-calibration result for the transmission quantities is like that of the second standard shown in Fig. 4.

specification of the uncertainty of $\pm 0.5°$. Fig. 4 demonstrates the good performance of the self-calibration. It should be stressed that the performance will not degrade if the quality of the short gets worse, because it is allowed to be unknown. However, it must be the same reflection at both ports. In anticipation of Fig. 7 it should be mentioned that the assumption of a precisely known short does not result in a better performance.

The TAN procedure is a good candidate for showing the errors to be expected if the third standard is not highly reflecting. An offset air-line line (Fig. 5) with 25 Ω characteristic impedance has been used to realize N_3. The transmissions evaluated via self-calibration are just as good as for the second standard N_2 and therefore are not plotted. The crucial step is the evaluation of the reflections. As can be seen from Fig. 5, the errors become quite high if the magnitude of the reflection becomes too small. Approximately

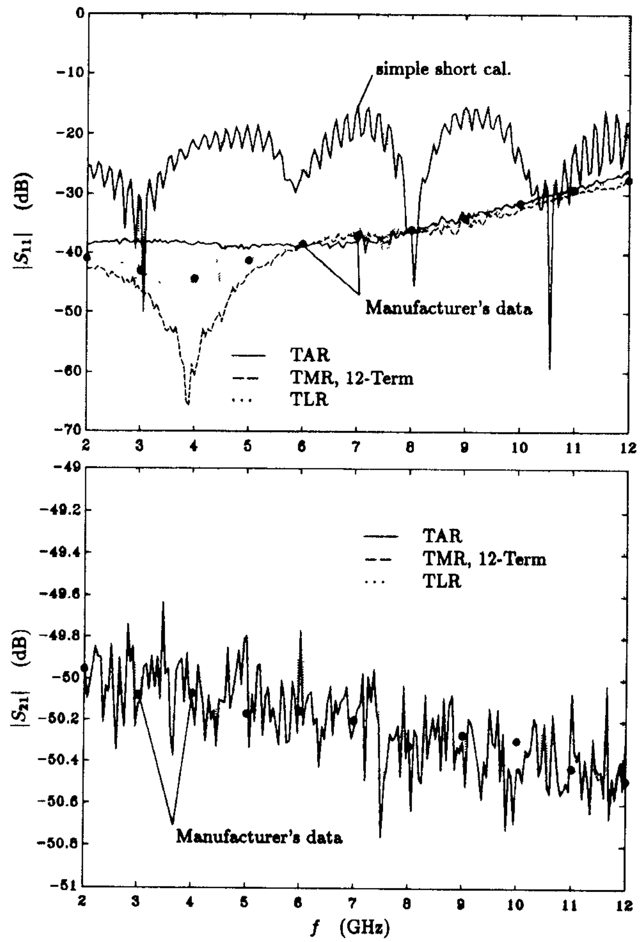

Fig. 6. Comparison of the error removal performance of different calibration procedures. Device under test: 50 dB attenuator.

these errors are to be expected in subsequent measurements. For this step the measure of reliability is $F_2 < 0.04$.

As one example to prove the measurement capability the magnitudes of the scattering parameters S_{11} and S_{21} of a 50 dB attenuator are plotted in Fig. 6. The performance of the arguments is equivalent. The difference between the 12-term and TMR procedures are, owing to the good standards, beyond the resolution of the drawing. The deviations are smaller than 0.1 dB for reflection and transmission as well over the full dynamic range. The dip in the curve at 4 GHz is due to the matched load used by these procedures. But, considering the absolute values of S_{11} and the specification $|r_M| \leq -34$ dB, there is no cause for criticism. To measure very small reflections one should prefer a sliding load or choose one of the other procedures. If the accuracy of the standards is not as good as in this example the performance of the 12-term procedure will degrade significantly faster than the TMR performance. This is because 12-term is based on three exactly known one-port standards, but TMR only on the knowledge of one of them.

In contrast to the broad-band behavior of the new procedures presented, the loci for the TLR (= TRL) procedure show the expected ranges of unreliable calibration. The fast periodicity is due to the length of the line of approximately 100 mm.

In order to find an answer to the question whether it is advantageous if known standards instead of unknowns are

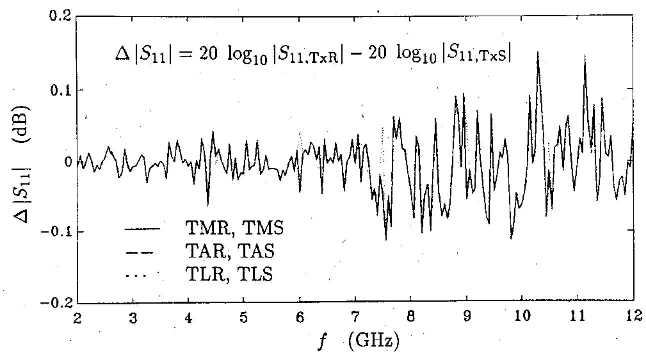

Fig. 7. The difference between corresponding TxR and TxS procedures

used, we compare the results of corresponding TxR and TxS procedures. Fig. 7 shows the differences in the magnitude of S_{11} for the same measurements as plotted in Fig. 6. The difference in the argument is below 1°. As can be taken from (46) in principle xxR and xxS show no difference in transmission measurements. Even for the high-quality shorts used here and the small absolute values shown in Fig. 6 the difference of less than 0.1 dB does not justify the application of TxS procedures. However, they can be highly important in many practical situations, e.g. if the two ports have connectors of different sex. Another remedy in this situation is to employ additional adapters [10].

IV. Conclusion

A general theory for network analyzer calibration has been presented. Based on this theory a family of calibration procedures has been presented, namely TAN, TLN, TMN, TAR, TLR, TMR, TAS, TLS, and TMS and their corresponding procedures LAN, LLN, LMN, LAR, LLR, LMR, LAS, LLS, and LMS. Compared with the well-known 12-term procedure, the new methods employ fewer standards, which in addition may be partly unknown. In contrast to the TRL and TSD procedures the methods xAx and xMx are in principle of unlimited bandwidth. The good performance of all the procedures has been shown experimentally. This wide spectrum of procedures using different calibration standards provides the opportunity to choose an optimal algorithm for any environment.

References

[1] Hewlett Packard, *Automating the HP 8410B Microwave Network Analyzer*, Application Note 221A, June 1980.
[2] Hewlet Packard *Specifying Calibration Standards for the HP 8510 Network Analyzer*, Product Note 8510-5a Feb. 1988.
[3] N. R. Franzen and R. A. Speciale, "A new procedure for System Calibration and error removal in automated S-parameter measurements," *Proc. 5th European Microwave Conf.* (Hamburg), 1975, pp. 69–73.
[4] G. F. Engen and C. A. Hoer, "Thru-reflect-line: An improved technique for calibrating the dual six port automatic network analyzer," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, Dec. 1979, pp. 987–993.
[5] H. J. Eul and B. Schiek, "Error-corrected two-state unidirectional network analyzer," *Electron. Lett.*, vol. 24, no. 19, pp. 1197–1198, Sept. 1988.
[6] R. A. Soares, P. Gouzien, P. Legaud, and G. Follot, "A unified mathematical approach to two-port calibration techniques and some applications," *IEEE Trans. Microwave Theory Tech.*, vol. 37, pp. 1660–1674, Nov. 1989.
[7] H. J. Eul and B. Schiek, "Thru-match-reflect: One result of a rigorous theory for de-embedding and network analyzer calibration," *Proc. 18th European Microwave Conf.* (Stockholm), 1988, pp. 909–914.
[8] R. Zurmühl, S. Frank, *Matrizen und ihre Andwendungen.* Berlin: Springer Verlag, 1984.
[9] Avantek, *Measurement and Modelling of GaAs FET Chips*, Application Note, Oct. 1983.
[10] C. A. Hoer and G. F. Engen, "Calibrating a dual six-port or four-port for measuring two-ports with any connector," *IEEE MTT-S Int. Microwave Symp. Dig.* 1986, pp. 665–668.
[11] Hewlett Packard, *Applying the HP8510B TRL Calibration for Non-coaxial Measurements*, Product Note 8510-8, Oct. 1987.
[12] D. Rytting, Private communication, May 1990.
[13] H. J. Eul and B. Schiek, "Experimental results of new self-calibration procedures for network analyzers," in *Proc. 20th European Microwave Conf.* (Budapest), 1990, pp. 1461–1466.

Hermann-Josef Eul (S'88–M'91) was born in Neustadt/Wied, Federal Republic of Germany, in 1959. He received the Dipl.Ing. (FH) degree from the Fachhochschule Koblenz (polytechnique) in 1984 and the Dipl.Ing. degree from the University of Bochum in 1987, both in electrical engineering. From 1987 to 1990 he was a research assistant at the University of Bochum, where he was mainly engaged in microwave measurement techniques. He received the Dr.Ing. degree in 1990 and is now with Siemens, Munich, working on mobile radio systems.

Burkhard Schiek (M'85) was born in Elbing, Germany, on October 14, 1938. He received the Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering, both from the Technische Universität Braunschweig, Germany, in 1964 and 1966, respectively.

From 1964 to 1969, he was an Assistant at the Institut für Hochfrequenztechnik, Technische Universität Braunschweig, where he worked on frequency multipliers, parametric amplifiers, and varactor phase shifters. From 1966 to 1969, he was involved in MIS interface physics and in the development of MIS varactors. From 1969 to 1978, he was with the Microwave Application Group of the Philips Forschungslaboratorium Hamburg GmbH, Hamburg, Germany, where he was mainly concerned with the stabilization of solid-state oscillators, oscillator noise, microwave integration, and microwave systems. Since 1978, he has been a Professor in the Department of Electrical Engineering, Ruhr-Universität Bochum, Germany, working on high-frequency measurement techniques and industrial applications of microwaves.