Short Papers

Frequencies of a Truncated Circular Waveguide—Method of Internal Matching

C. Y. Wang

***Abstract*—The frequencies of a truncated circular waveguide is determined using the method of internal matching. Both TE and TM modes are considered. As the aspect ratio increases, the modes change continuously from those of a long rectangle to those of a circular waveguide. The TM modes also describe vibrating membranes.**

***Index Terms*—Cutoff frequency, normal mode, waveguide.**

I. Introduction

The physics of electromagnetic waves propagating in a hollow waveguide is mainly governed by the Helmholtz equation [1], [2]. The E modes or TM modes satisfy zero boundary conditions and the H modes or TE modes require zero normal derivatives on the boundary. The former problem is analogous to the vibrations of membranes, while the latter problem has no equivalent in mechanical vibrations. The methods used to solve the Helmholtz equation for various cross-sectional geometries were reviewed by Ng [3].

This paper concerns the truncated circular waveguide whose cross section is shown in Fig. 1. This geometry is sometimes used to change the orientation of the wave axis. Pyle and Angley [4], using a finite strip method, Sinnott [5], using finite differences, and Levy [6] using approximations, published the cutoff wavelengths (TE) for aspect ratio a from 0.7 to 1. Their results, however, differ considerably.

The purpose of this paper is several fold. We extend the published results to the full range $0 \leq a \leq 1$, while considering the higher TE and TM modes. In the process, the accuracy of either [4] or [5] is ascertained. We also introduce a new internal matching method, which is more efficient and especially suited for this geometry.

II. Formulation

The governing equation in cylindrical coordinates (r, θ) is

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} + k^2 u = 0 \tag{1}$$

where all lengths have been normalized by the circle radius R, u is the scalar wave function, and k is the normalized frequency

$$k = 2\pi R(\text{frequency})\ (\text{inductivity} \cdot \text{capacitivity})^{1/2}. \tag{2}$$

For TM modes, u is to be zero on the boundary. Due to the geometry, existing methods are tedious, if not impossible, to apply. The internal matching method is as follows. Complete the circle by the dashed lines in Fig. 1(b). Consider the lens-shaped region $AP\widehat{B}\widehat{P'}A$ where P' is the reflection of P about the line $\overline{AB}$. To ensure u is zero on $\overline{AB}$, we let the fictitious segment $A\widehat{B}\widehat{P'}A$ be antisymmetrical about $\overline{AB}$, i.e.,

$$u(P') = -u(P). \tag{3}$$

Manuscript received May 19, 1998.

The author is with the Department of Mathematics, Michigan State University, East Lansing, MI 48824 USA.

Publisher Item Identifier S 0018-9480(00)08721-4.

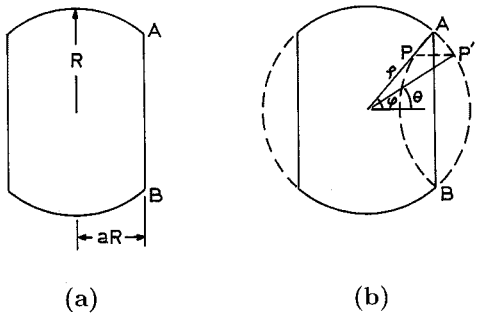


Fig. 1. (a) Truncated circular waveguide. (b) Coordinate system showing extended domain.

Notice $u(P')$ is not zero since the field is analytically extended into the fictitious region.

If point P' is at $(1, \theta)$, trigonometry yields P at (ρ, ϕ) where

$$\rho \equiv \sqrt{1 + 4a^2 - 4a\cos\theta} \tag{4}$$

$$\phi \equiv \tan^{-1}\left(\frac{\sin\theta}{2a - \cos\theta}\right). \tag{5}$$

Now the domain is the full circular region with zero boundary conditions on the solid arcs and (3) on the dashed arcs.

The lowest TM mode is symmetric about $\theta = 0$ and $\theta = \pi/2$. We name this as the first $\mathrm{TM}_{\mathrm{SS}}$ mode. The general solution satisfying the symmetry conditions (1) and bounded at $r = 0$ is

$$u(r, \theta) = \sum_{n=0}^{\infty} C_n (2n)! \cos(2n\theta) J_{2n}(kr) \tag{6}$$

where C_n are coefficients to be determined, J_{2n} are Bessel functions, which decay to zero as $r \to \infty$, and the factorial is to assure C_n to be of reasonable magnitudes for large n. Let $\beta = \cos^{-1} a$. The boundary conditions are

$$u(1, \theta) = 0, \qquad \beta < \theta \leq \frac{\pi}{2} \tag{7}$$

$$u(1, \theta) = -u(\rho, \phi), \qquad 0 \leq \theta < \beta. \tag{8}$$

We use point match on N points along the quarter circle. Let

$$\theta_j = (j - 0.5)\pi/2N, \qquad j = 1, \ldots, N. \tag{9}$$

Truncate (6) to N terms and apply (7) and (8)

$$\sum_{0}^{N-1} C_n (2n)! \left\{ \cos(2n\theta_j) J_{2n}(k) + \begin{bmatrix} 0 & \beta < \theta_j \\ \cos\left(2n\phi(\theta_j)\right) J_{2n}\left(k\rho(\theta_j)\right) & \beta > \theta_j \end{bmatrix} \right\} = 0. \tag{10}$$

For nontrivial C_n, the determinant of coefficients is set to zero. The eigenvalue k is obtained by a root search algorithm. Spurious roots, if any, are rejected by considering the eigenfunction u.

TABLE I
WAVEGUIDE CUTOFF FREQUENCY

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
k	1.571	1.573	1.583	1.595	1.613	1.640	1.672	1.716	1.764	1.807	1.841

TABLE II
CUTOFF FREQUENCY FOR VIBRATING MEMBRANE

a	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
k	∞	15.73	8.010	5.470	4.235	3.528	3.081	2.785	2.588	2.462	2.405

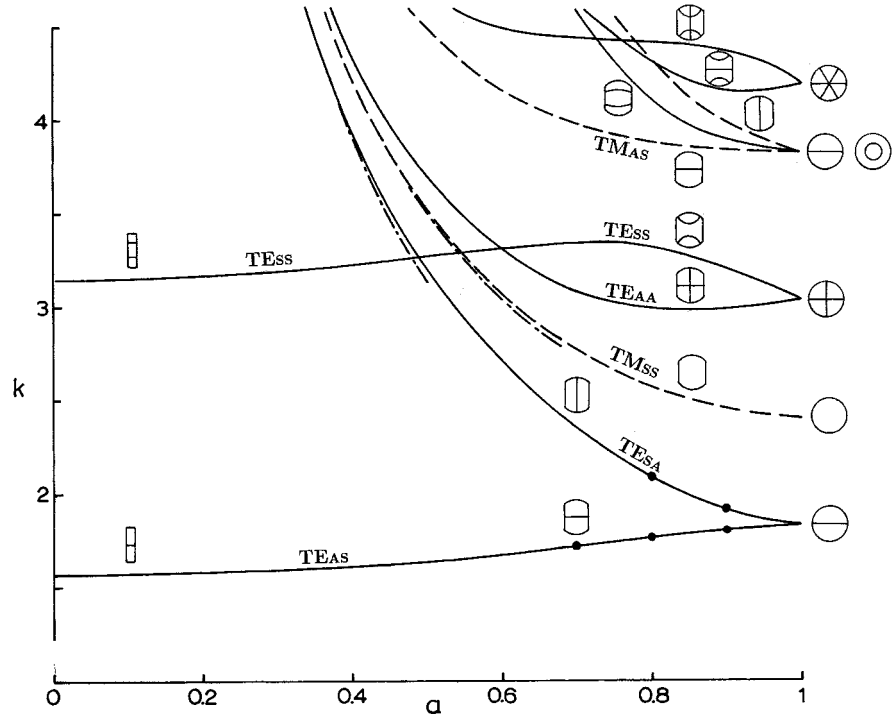


Fig. 2. Natural frequencies of the truncated circular waveguide. Circles denote the results by Sinnott [5]. ——: TE modes, – – – – –: TM modes, —·—·—: approximations of (19) or (20).

The next-order TM mode is antisymmetrical about $\theta = 0$ and symmetrical about $\theta = \pi/2$. We call this the $\mathrm{TM}_{\mathrm{AS}}$ mode. Equation (6) is replaced by

$$u(r,\theta) = \sum_{n=1}^{\infty} C_n (2n-1)! \sin\Big[(2n-1)\theta\Big] J_{2n-1}(kr). \tag{11}$$

Point match is then applied to (7) and (8) as before. For the $\mathrm{TM}_{\mathrm{SA}}$ mode, we assume

$$u(r,\theta) = \sum_{n=1}^{\infty} C_n (2n-1)! \cos\Big[(2n-1)\theta\Big] J_{2n-1}(kr). \tag{12}$$

For the $\mathrm{TM}_{\mathrm{AA}}$ mode

$$u(r,\theta) = \sum C_n (2n)! \sin(2n\theta) J_{2n}(kr). \tag{13}$$

In the case of TE modes, the normal derivatives of u is zero on the boundary. Using internal matching, we set

$$u(P') = u(P) \tag{14}$$

on the dashed arcs. The boundary conditions are thus

$$\frac{\partial u}{\partial r}(1,\theta) = 0, \qquad \beta < \theta \le \frac{\pi}{2} \tag{15}$$

$$u(1,\theta) = u(\rho,\phi), \qquad 0 \le \theta < \beta \tag{16}$$

The lowest TE mode is antisymmetrical about $\theta = 0$ and symmetric about $\theta = \pi/2$. We use (11) for the $\mathrm{TE}_{\mathrm{AS}}$ mode. The derivative is

$$\frac{\partial u}{\partial r}(r,\theta) = \sum_{n=1}^{\infty} C_n (2n-1)! \sin\Big[(2n-1)\theta\Big] \times \left[k J_{2n-2}(kr) - \frac{(2n-1)}{r} J_{2n-1}(kr)\right]. \tag{17}$$

$$\sum_{1}^{N} C_n(2n-1)! \begin{Bmatrix} \sin\Big[(2n-1)\theta_j\Big]\Big[kJ_{2n-2}(k)-(2n-1)J_{2n-1}(k)\Big], & \beta<\theta_j \\ \sin\Big[(2n-1)\theta_j\Big]J_{2n-1}(k)-\sin\Big[(2n-1)\phi(\theta_j)\Big]J_{2n-1}\Big(k\rho(\theta_j)\Big), & \beta>\theta_j \end{Bmatrix} = 0 \tag{18}$$

Equations (15) and (16) then yield (18), shown at the top of this page.

Again, we have $\mathrm{TE}_{\mathrm{SA}}$, $\mathrm{TE}_{\mathrm{AA}}$, and $\mathrm{TE}_{\mathrm{SS}}$ modes in the form of (12), (13), and (6), respectively, but with the boundary conditions of (18).

The convergence and accuracy of our solution can be ascertained by increasing N the number of terms retained. A typical example of $\mathrm{TE}_{\mathrm{AS}}$ mode, i.e., $a = 0.7$, gives $k = 1.715$ 1.716, and 1.716 for $N = 10$, 20, and 30, respectively. In general, $N = 20$ is adequate in obtaining four-digit accuracy. For low values of a, such as 0.1, $N = 30$ is sufficient.

III. Results and Discussion

Table I shows the cutoff frequency for the waveguide ($\mathrm{TE}_{\mathrm{AS}}$ mode). Our values for $a = 0.7, 0.8$, and 0.9 agree with the graph published by Sinnott [5] using finite differences. The lowest TM mode is important for the cutoff frequency of vibrating membranes and is tabulated in Table II ($\mathrm{TM}_{\mathrm{SS}}$ mode).

Fig. 2 shows the frequencies of higher modes. For waveguides, both TE and TM modes are possible, while for vibrating membranes, only TM modes exist.

When $a \to 0$, the cross section resembles a long rectangle. In this case, only the $\mathrm{TE}_{\mathrm{AS}}$ or $\mathrm{TE}_{\mathrm{SS}}$ modes give finite frequency. The exact values for $\mathrm{TE}_{\mathrm{AS}}$ modes are $k = \pi/2, 3\pi/2, 5\pi/2, \ldots$, and the $\mathrm{TE}_{\mathrm{SS}}$ modes are $\pi, 2\pi, 3\pi, \ldots$. Considering an infinite strip of width $2a$, an asymptotic formula for the lowest $\mathrm{TE}_{\mathrm{SA}}$ mode can be derived

$$k = \frac{\pi}{2a}, \quad a \to 0. \tag{19}$$

Similarly, considering a rectangle of $2 \times 2a$, the lowest $\mathrm{TM}_{\mathrm{SS}}$ mode is approximated by

$$k = \frac{\pi}{2}\sqrt{1+\frac{1}{a^2}}, \quad a \to 0. \tag{20}$$

When $a = 1$, the cross section is a circle. The frequencies are the roots of either the Bessel function J_n or its derivative J_n' [2]. In ascending order they are: 1.841 (TE), 2.405 (TM), 3.054 (TE), 3.832 (TE and TM), 4.201 (TE), 5.136 (TM), etc.

Note for small a, the frequencies are widely separated, which is the reason why narrow rectangular (or near narrow rectangular) waveguides are used. For $a \approx 1$ (near circle), numerous different frequencies can be excited. In order to change the orientation of the electromagnetic waves, the waveguide cross section can be made to change gradually from a narrow rectangle to a circle then to a narrow rectangle of a different orientation. Fig. 2 shows the transition of frequency and mode shapes as geometry is changed, and would be useful in the design of waveguides.

Our method of internal matching is accurate and efficient. In comparison, finite differences would have to contend with curved boundaries and a much larger system of equations. Of course, finite differences are more versatile in terms of treating complex geometries.

References

[1] R. F. Harrington, *Time Harmonic Electromagnetic Fields*. New York: McGraw-Hill, 1961, ch. 8.

[2] J. A. Edminister, *Schaum's Outline of Electromagnetics*. New York: McGraw-Hill, 1979, ch. 16.

[3] F. L. Ng, "Tabulation of methods for the numerical solution of the hollow waveguide problem," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 322–329, Mar. 1974.

[4] J. R. Pyle and R. J. Angley, "Cutoff wavelengths of waveguides with unusual cross sections," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-12, pp. 556–557, Sept. 1964.

[5] D. H. Sinnott, "The application of finite difference techniques to electromagnetic problems," *Elect. Eng. Trans.*, vol. EE-6, pp. 6–11, Mar. 1970.

[6] R. Levy, "The relationship between dual mode cavity cross-coupling and waveguide polarizers," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 2614–2620, Nov. 1995.

A Novel Scheme for Reception Using an Active Microstrip Antenna

B. N. Biswas, A. Bhattacharya, P. Lahiri, and D. Mondal

***Abstract*—An unique aspect of the injection-locked active microstrip antenna incorporating a Gunn diode, i.e., the variation of the device current within the lock band, is exploited for the demodulation of an FM microwave signal. A simple experimental arrangement has been presented to demonstrate the demodulation performance of the receiver even in a multichannel environment. Using two such separate antennas as transmitter and receiver, a two-way short-range communication system has been demonstrated that may be suited for certain commercial and military applications.**

***Index Terms*—Active microstrip antenna, current valley.**

I. Introduction

Active microstrip antennas are potentially attractive, offering an effective alternative to conventional phase and adaptive arrays for a few commercial applications [1], short-range microwave communication, local area networks (LANs), microwave identification systems, etc. [2]. Obviously they offer savings in size, weight, cost, and improved efficiency. Recently self-mixing Gunn mounted active antennas have been reported [2], [3] at L-band for a half-duplex communication system and at X-band for transceiver and spatial power-combining applications. A Gunn diode integrated active antenna has the unique advantage of performing the combined functions of transmitter, local oscillator, and a very low conversion loss mixer [3]. However, the processing of the IF signal requires additional electronics, particularly for the detection of an FM signal.

In this paper, we propose a scheme for the reception of a microwave FM signal at X-band for short-range communication and LANs, using

Manuscript received May 7, 1999. This work was supported by the Ministry of Information Technology, Government of India, and by the University Grants Commission.

The authors are with the Radionics Laboratory, Physics Department, Burdwan University, Burdwan 713 104, India.

Publisher Item Identifier S 0018-9480(00)08719-6.

the variation of the device current in the lock band of an injection synchronized active microstrip antenna. This approach eliminates the use of intermediate stages for the recovery of message or data signals, in contrast to earlier schemes [2], [3].

II. Demodulation Performance of Active Antenna

A. Active-Antenna Configuration

Fig. 1 shows configuration of the active patch receiver (APRx). A rectangular patch has been designed by etching on Taconic Substrate TLY-5, with a relative dielectric constant of 2.2 and thickness 0.7870 mm. The patch antenna is supported from the other side (ground plane) by a thick aluminum plate to form a good heat sink and a mechanical support for experimental mounting. The dimensions (length $=10.1975$ mm and width $=15.06943$ mm) of the antenna are fixed [4] depending on the operating frequency (f_o), which, is in our case, is 9.5 GHz. An M/A Com, Burlington, MA (MA 49 104-111) 25-mW Gunn diode is screw-mounted at a point $l = 4.326$ mm (measured from the top of the patch) on the patch for the highest possible radiated power from the antenna. DC bias is applied to the diode through an RF choke realized by means of etching a hi–lo-type two-section transformer.

B. Device Current Well in Injection-Locked Gunn Oscillator

The variation of the device current over the lock-band of an injection-locked Gunn oscillator (ILGO) resembles [5] the data in Fig. 2. This happens basically due to the dynamic (i–v) characteristic of the Gunn diodes, viz.

$$i = -\alpha v + \beta v^2 + \gamma v^3. \tag{1}$$

As a result, the average current through the diode is obtained as

$$I_{\rm av} = I_0 - \beta \langle v \rangle^2 \tag{2}$$

where $\langle v \rangle^2$ denotes the mean square output voltage of the oscillator. The frequency dependence of (2) can be explained as follows. When the open-loop frequency error (Ω) is zero (i.e., the error between the free-running and injected signal frequencies), the synchronizing signal power directly adds up with that of the oscillator as the phase difference between the two signals is zero. As Ω increases, the phase difference between the two signals increases and the in-phase component of the synchronizing signal reduces. Hence, the output voltage, i.e., the locked oscillator output, decreases with Ω. Therefore, the device current increases on both sides of the center frequency, as shown in Fig. 2. The frequency response characteristic of an ILGO is known to be given by [6]

$$\left(1 - \langle v \rangle^2\right)^2 \langle v \rangle^2 + \left(\frac{\Omega}{\omega_o/2q}\right)^2 \langle v \rangle^2 = E \tag{3}$$

where E is directly related to the strength of the synchronizing signal. The shape and depth of the valley depend on the circuit parameters. Carefully looking into this current well/valley phenomenon, the following observations can be made. The frequency variation of the incoming signal is first converted into amplitude variation of the locked oscillator, which is manifested as a device current variation. Thus, in brief

Current Well/Valley Phenomenon

$$\Rightarrow \text{FM-AM Conversion} + \text{Square Law Detection}. \tag{4}$$

The sensitivities of the active antenna receiver is

$$S = \frac{\delta I}{\delta \Omega} = \frac{\delta I}{\delta \langle v \rangle^2} \frac{\delta \langle v \rangle^2}{\delta \Omega}. \tag{5}$$

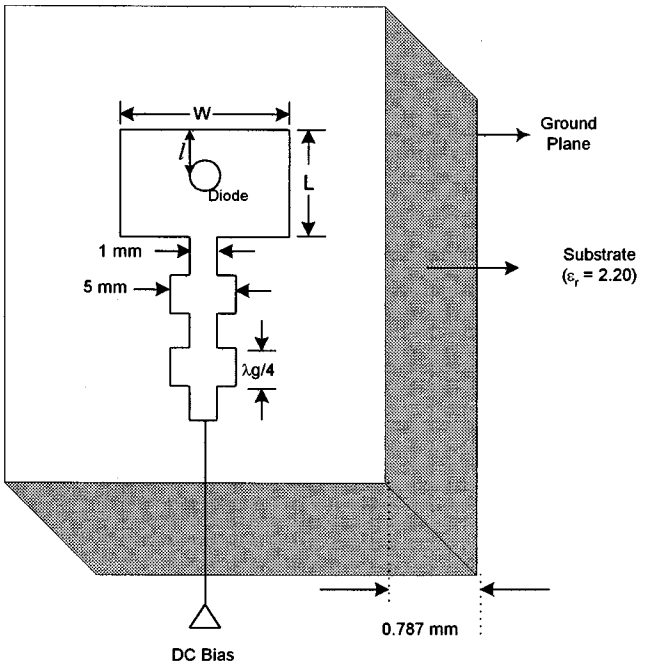


Fig. 1. APRx configuration showing the location of the Gunn diode on the patch and the diode biasing arrangement.

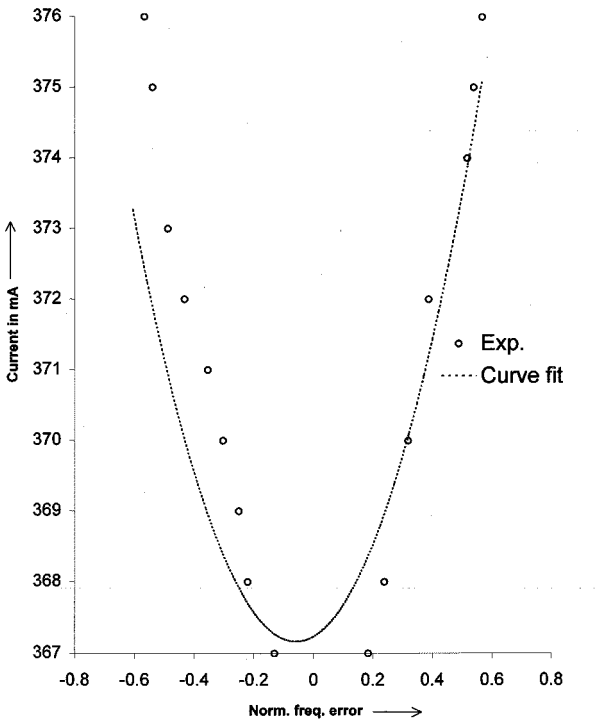


Fig. 2. Current valley characteristic as observed for the injection synchronized active patch oscillator. Variation of the diode bias current with normalized frequency detuning $((f_{\rm inj} - f_0)/\Omega)$ is plotted.

Dependence of $\langle v \rangle^2$ on Ω is given by (3), which can be rewritten as

$$p = \sqrt{\left[\frac{E}{a^2} - (1 - a^2)^2\right]} = f(a^2) \quad \text{(say)} \tag{6}$$

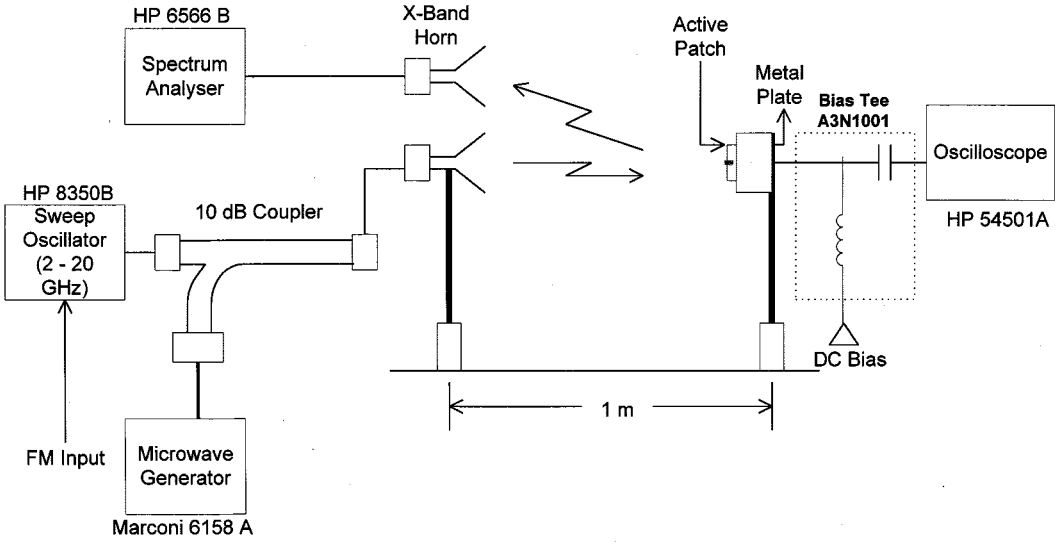


Fig. 3. Experimental setup for observing demodulation by the APRx.

where

$$p = \frac{\Omega}{\omega_o/2q} \quad \text{and} \quad \langle v \rangle^2 = a^2.$$

The Taylor expansion of (6) along with (5) gives the sensitivity of APRx as

$$S = -\beta \left(\frac{\omega_o}{q}\right) \frac{p_o}{\left[\frac{E}{a_o^4} - 2\left(1 - a_o^2\right)\right]} \tag{7}$$

a_o^2 corresponds to the power at a normalized frequency error p_o. Using the above expression, one can also find the distortion components at the output of the APRx [5].

C. Demodulation Performance

To audit the detection performance of the APRx, an experimental arrangement, as shown in Fig. 3, is made. The patch is biased by using a broad-band bias T from Anritsu, Tokyo, Japan. The antenna oscillates at a frequency of 9.866 GHz when the diode bias is 14.2 V. The spectral purity of the signal is quite good. In the steady state of operation, the temperature variation of frequency is about 400 KHz around its center frequency and has a close to carrier FM noise of -99.7 dBc/Hz measured at 200-kHz offset. A voltage tuning range of 170 MHz is obtained (power variation <1.5 dB), when the bias voltage is varied from 10 to 14.6 V.

The active antenna is irradiated from a standard gain (18-dB specified) X-band horn antenna from a line-of-sight distance of 1.0 m and connected to a microwave frequency generator. A 1.0-MHz sinusoidal tone signal obtained from a Philips 807/DRF RF oscillator is used to frequency modulate the microwave carrier. The generator has a frequency sensitivity of 6 MHz/V and peak-to-peak swing of the tone signal is 40.78 mV. This gives an FM index $(\Delta f/f_m)$ of 0.122. The reception capability of APRx was also tested in a multichannel (two-channel, in our case) environment. The interfering signal (frequency separation $= 40$ MHz) is generated from a Marconi source and power combined in order to transmit both the carriers simultaneously from the horn antenna. The oscillator frequency is slightly bias tuned to capture the desired carrier. Injection-locking phenomenon is observed on a Spectrum Analyzer by using another similar type horn antenna (cf. Fig. 3). The lock range is approximately 6.0 MHz at an approximate locking gain of 20 dB (transmitted carrier power $= 10.0$ dBm). The demodulated output is obtained at the output port of the bias T and its amplitude depends upon the demodulation efficiency of the injection-locked oscillator, which is adjusted by varying the diode bias. The bias T has been used for experimental demonstration. However, the RF choke would have some low frequency rolloff. In actual circuit, it should be replaced by a low dc resistance in order to have a uniform response over the demodulation band.

The demodulated signal (in the absence of any interfering signal) is compared to the original test tone, and an online fast Fourier transform (FFT) analysis reveals that distortion is only about 1%. However, slight disadvantage with the present scheme is the reduction in the locking range, which can be calculated [6] by using the relation

$$\Omega_e = \Omega_l - 8\Delta f. \tag{8}$$

In our case, $\Omega_l = 6.0$ MHz and $\Delta f = 0.122$ MHz, thus, the effective locking range reduces to 5.024 MHz.

In the presence of the interfering carrier, the information signal becomes heavily corrupted by beat frequency noise, as shown in Fig. 4(a). The beat frequency noise may easily be removed by using even a passive low-pass filter. An oscillogram of the filtered signal is shown in Fig. 4(b).

III. Application Proposal

The APRx as proposed here will be very useful for short-range communication, as it is based on the principle of injection locking. Such proposition will be extremely suitable for indoor or premises communication and LANs. For duplex communication, it would be suitable to use two separate active patches as the transmitter (Tx) and receiver (Rx). A varactor tuned active patch, as demonstrated in [3], and fitted with a high power diode may be used as the Tx. Data outputs from computers or multiplexed voice signals for wireless private branch exchange (PBX) application are to be applied to the varactor terminal for frequency modulation of the microwave carrier. Frequency modulation is also possible by modulating the bias current of the Gunn diode. Bias modulating an APTx by test-tone signals have frequencies of 1.0 and 2.0 MHz, respectively, and demodulation performance of the APRx is observed to be quite satisfactory. This demonstrates the possibility of the proposed scheme. However, the distance between the Tx and Rx

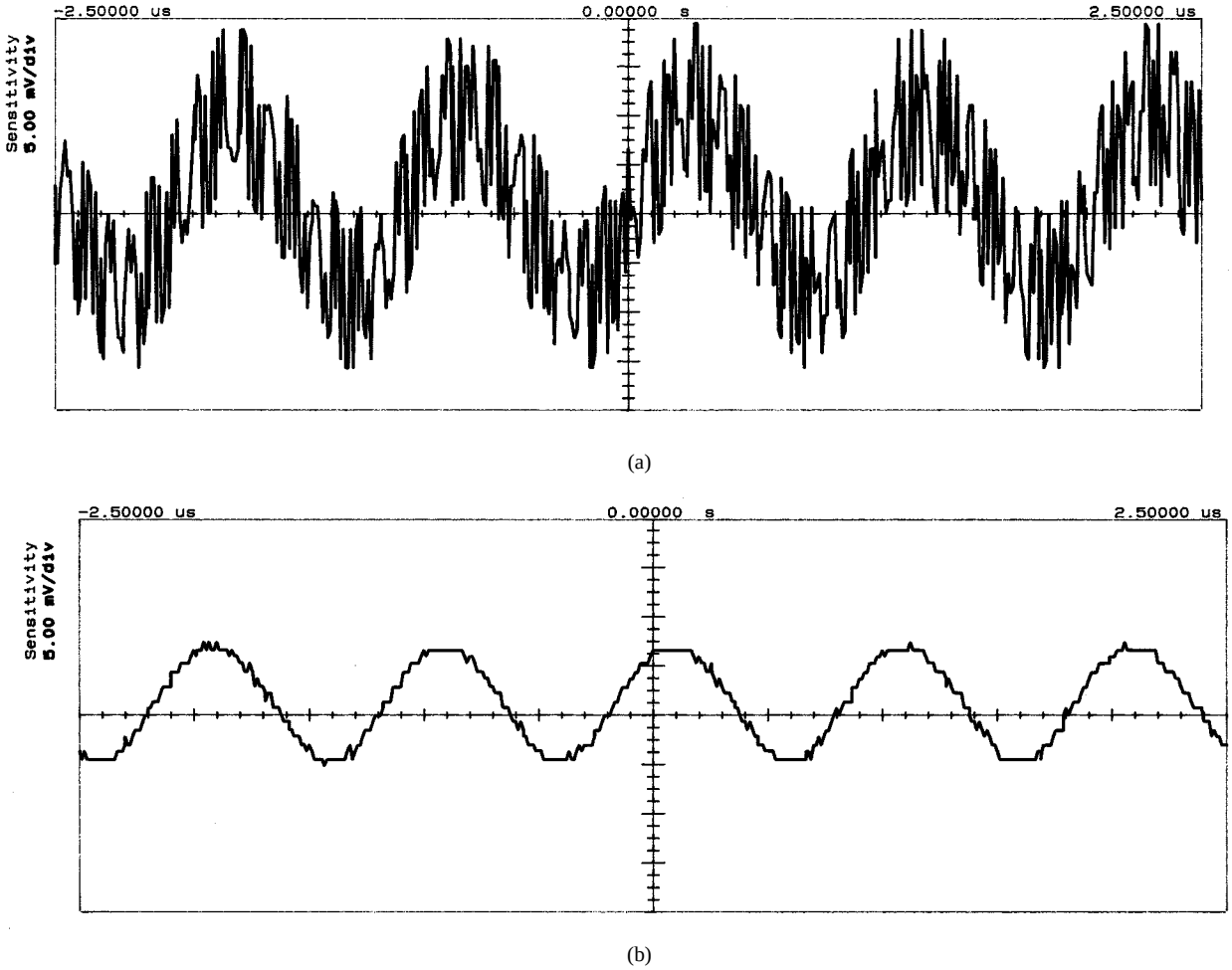


Fig. 4. Demodulation of the test-tone signal in an interfering environment. (a) Signal heavily corrupted by the beat noise. (b) Filtered signal completely removed off beat noise.

has been reduced to 0.6 m in order to keep the locking range the same as earlier. This restriction is imposed due to the low radiated power (25-mW diode, $\mathrm{EIRP} < 125$ mW) by the Tx antenna. Use of a high power diode (preferably 100 mW), accurate broadside alignment of the Tx antenna, and low-power diode (10 mW) for the APRx can increase this range. For the transmission of voice or data signals, this will not impose much restriction on available locking bandwidth. For two-way communication, two separate frequency ranges may be used. An adjacent channel may be a signal transmitted from another active patch transmitter carrying data from a separate hub. Oscillator lock range may be increased by means of varactor control of frequency and, in that case, close-circuit video transmission may also be possible.

IV. Conclusion

A Gunn diode mounted APRx has been investigated for the reception of FM microwave signals. Current well/valley phenomenon has been successfully utilized to demodulate the baseband information. The carrier is locked on to the patch oscillator by means of bias tuning (tuning sensitivity $\sim$38 MHz/V) the APRx. With the present system configuration, it is not difficult to realize a faithful demodulation bandwidth, which can successfully accommodate quite a large number of voice or data channels. The uniqueness of the receiver lies in its simplicity, as it avoids the use of IF stages. It also works well in a multichannel environment due to noise filtering property of an ILGO. Due to its simple circuit configuration and similarity in Tx and Rx architecture, an active patch antenna as demonstrated, is well suited for commercial and military application as a two-way microwave communication system.

References

[1] R. A. York and T. Itoh, "Injection- and phase-locking technique for beam control," *IEEE Trans. Microwave Theory Tech.*, vol. 46, pp. 1920–1929, Nov. 1998.

[2] C. M. Montiel, L. Fan, and K. Chang, "A novel active antenna with self-mixing and wide-band varactor-tuning capabilities for communication and vehicle identification applications," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2421–2430, Dec. 1996.

[3] ——, "An X-band self-mixing oscillator antenna for transceiver and spatial power combining," *IEEE Trans. Microwave Theory Tech.*, vol. 46, pp. 1546–1551, Oct. 1998.

[4] B. N. Biswas, A. Bhattacharya, D. Mondal, P. Lahiri, A. Bose, and S. Pal, "Rectangular active microstrip patch antenna revisited," *IETE J. Res.*, vol. 45, no. 2, pp. 135–145, Mar.–Apr. 1999.

[5] B. N. Biswas, S. Chatterjee, D. Mondal, and P. Pal, "Current valley in Gunn diode oscillators," *J. IETE*, vol. 34, pp. 446–449, June 1988.

[6] B. N. Biswas, *Phase Lock Theories and Applications*. Oxford, U.K.: Oxford Univ. Press, 1988.

High Combining-Efficiency X-Band Spatial Power-Combined Array Using a Multilayered Packaging Architecture

Sean M. Duffy and Mark A. Gouker

***Abstract*—The design of a high combining-efficiency spatial power-combined array is described in this paper. A multilayered stacked stripline architecture enables a compact stable design. An array incorporating antenna active impedance and proper amplifier matching is measured with a combining efficiency of 87%, radiating 6.8 W of an available 7.8 W into the ideal uniformly illuminated array directivity at 10.1 GHz.**

***Index Terms*—Active arrays, antenna array mutual coupling, MMIC amplifiers, multilayered microwave circuitry, power combiners, packaging.**

I. Introduction

The promise of spatial power combining [1] using solid-state devices has been illustrated with designs for high power [2]–[4] and at millimeter-wave frequencies reported in [5]–[8]. However, the success of any spatial power-combined array rests on the ability to effectively combine the output power from a number of solid-state devices. Therefore, the output circuitry and antenna are a critical aspect to the overall design. An effective measure of the output of a spatial power-combined array is the combining efficiency [1], [9].

Previous spatial power-combining efforts have primarily focused on maximizing output power without as much consideration to maximizing combining efficiency [2]–[8]. The best combining efficiency results in X-band have been 71% [2], [4]. In the work described in this paper, high combining efficiency is achieved by presenting the optimum impedance to the amplifier found using a load–pull technique and by considering active antenna impedance through inclusion of array mutual coupling with edge effects. In addition, a compact multilayered circuit architecture is used to minimize line lengths and associated line loss. An example 4×4 array design is presented which has a combining efficiency of 87% at 10.1 GHz. The design techniques and package architecture described have wider application than spatial power combining, including phased arrays and other active arrays.

II. Overview of Design

The spatial power combining in this paper combines the outputs from an array of monolithic-microwave integrated-circuit (MMIC) amplifiers using a circuit-fed tile-approach architecture [4], [7]. The input signal is distributed by a corporate-fed circuit structure, and the output signals are radiated to space via cavity-backed microstrip antennas. The improvement described in this paper is the multilayered stripline packaging architecture shown in Fig. 1(a) and (b). The MMIC amplifiers are placed on the bottom stripline layer where they can be attached to the heat sink for proper thermal management and the antennas are placed on the top stripline layer. The architecture features the isolation of circuit elements using ground-plane-to-ground-plane shorting vias, the placement of amplifiers in individual closed cavities, and the compactness inherent in the use of multilayered circuitry [10]. In addition, the multilayered packaging architecture improves the stability of the active array, an issue that becomes increasingly important when high-gain amplifiers are employed [10]. The various circuit components including the cavity-backed patch, RF stripline layer-to-stripline layer via, and buried cavity provide an isolated efficient circuit design.

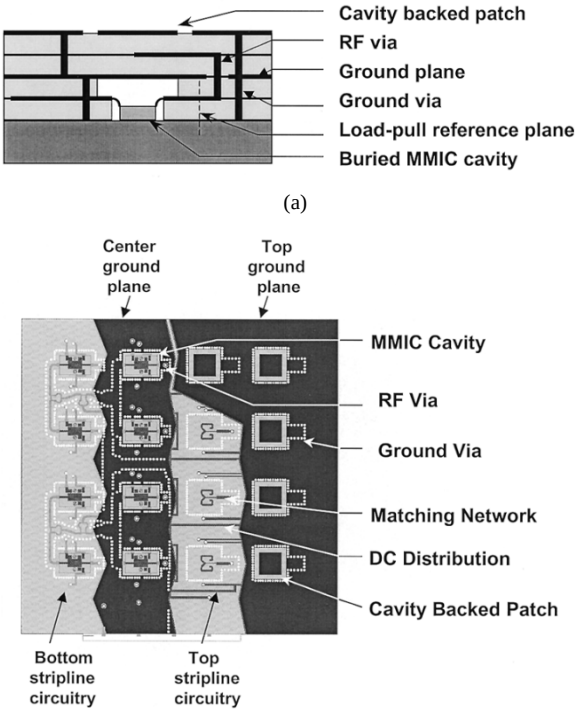


Fig. 1. Spatial power-combined array package consisting of two stacked stripline layers. (a) Side view. (b) Cut-away view.

The operating characteristics of the 16 MMICs (TGA8031-SCC, Texas Instruments Incorporated, Dallas, TX) used in the array design are large-signal gain of 30.9-dB dc–RF efficiency of 29.7%, output power of 490 mW with a $\pm 10^{\circ}$ phase variation. The first aspect to the improved combining efficiency is load–pull measurements performed on the MMIC amplifiers in the cavity with a reference plane in the stripline section just outside the cavity, as shown in Fig. 1(a). This choice of reference plane bundles the effect of two sets of wire bonds, substrate, and ground-plane discontinuities, and the cavity-to-stripline transition. The load impedance for maximum output power at 10 GHz is approximately $90 - j\ 10\ \Omega$ for the MMIC amplifiers in the array.

The second aspect to the improved combining efficiency is incorporating the antenna active impedance for a small array with edge effects. To accomplish this, careful measured and calculated input impedances of the output circuitry at the load–pull reference plane are found. This impedance includes the antenna impedance, quarter-wave matching section, transmission line, and RF via. The antenna active impedance is calculated using Momentum,[1] while the rest of the components use circuit models [10]. Mutual coupling creates variations in the antenna impedance across the small array. These variations have two potential impacts on the array performance. First, they must be accounted for optimum amplifier matching. Second, left uncorrected, the variations cause amplitude and phase errors in the radiated signals, leading to reduced antenna gain. By knowing the active impedance variations across

Manuscript received May 18, 1999. This work was supported by the Department of the Army and by the Defense Advanced Research Projects Agency under Contract F19628-95-C-0002.

The authors are with the Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Lexington, MA 02420-9108 USA.

Publisher Item Identifier S 0018-9480(00)08729-9.

[1]Agilent Momentum, Agilent Technology, Santa Rosa, CA.

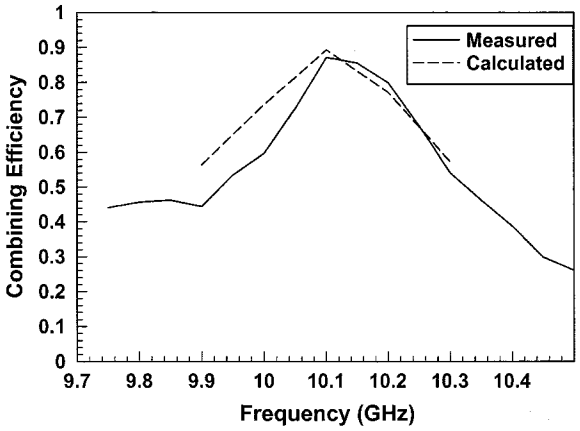


Fig. 2. Measured and calculated combining efficiency for the array.

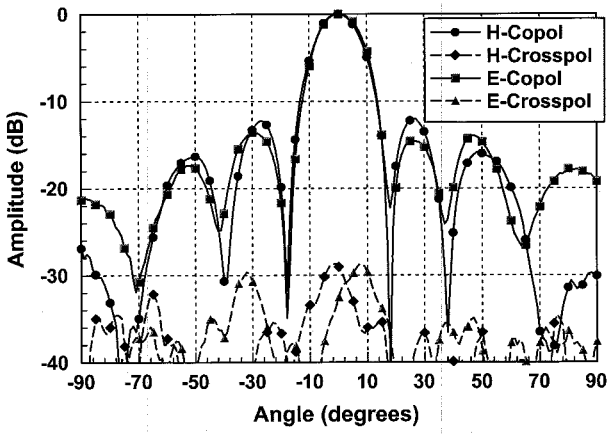


Fig. 3. Radiation patterns for the array at 10.1 GHz.

the array, the matching network can be adjusted to present $90 - j\ 10\ \Omega$ at the reference plane for all elements in the array. In the array, two matching states are used for the dominant E-plane coupling: one for the inner elements and one for the edge elements.

III. Measured Results

The figures-of-merit for spatial power-combined arrays have been discussed in [1]. The effective isotropic radiated power (EIRP) is measured using the Friis transmission formula for the array in the far field using a standard gain horn with power meter. The peak EIRP is measured to be 29.0 dBW. The directivity of a uniformly illuminated array with $0.8\lambda_o$ spacing is calculated[2] to be 20.65 dBi at 10.1 GHz. The peak effective radiated power is 6.8 W with an available power across this band of 7.8 W. The combining efficiency is the ratio of the effective radiated power to the available power of the MMICs [1] leading to a peak value of 87%, as shown in Fig. 2. The system gain is defined as the effective radiated power divided by the input power [9] and is 28.3 dB for an input power to the array of 10 mW. The far-field patterns, H- and E-planes with co- and cross-polarization patterns are shown in Fig. 3. The patterns demonstrate the nearly uniform illumination of the array with symmetric sidelobes and well-defined nulls.

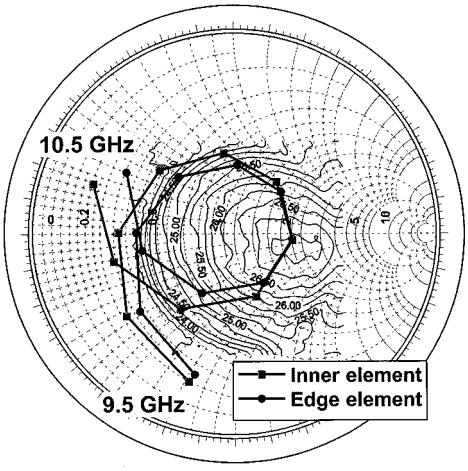


Fig. 4. Measured load–pull contours of typical MMIC (contour lines in dBm) and calculated input impedance of antennas (with mutual coupling compensation), matching section, RF via and transmission line at load–pull reference plane.

IV. Tabulation of Losses

The loss is tabulated for the array design. The loss can be categorized into three main sources: amplifier load mismatch loss, circuit loss, and directivity loss. Amplifier load mismatch loss is the reduction in output power due to not presenting the ideal load impedance to the amplifier. This loss for a power amplifier is much more significant than typical small-signal mismatch loss [11] and requires careful consideration. Impedance transformers were designed to match to the "typical" or average load impedance, while the MMICs have some variation about that impedance. In this paper, every MMIC has been measured, and the deviation from average has been tabulated. A 2% reduction in combining efficiency is attributed to this loss factor. The second loss factor, i.e., circuit loss, is the conductor and dielectric loss of the output lines, RF via, and radiation efficiency of the antenna. These losses can be calculated and/or measured usually to a fairly high degree of accuracy. Circuit loss leads to a reduction in 7% to the combining efficiency. The final loss factor, i.e., directivity loss, is the reduced antenna gain due to amplitude and phase errors at the array elements. The errors may be caused by variation in the amplifier responses either through process variations or due to antenna mutual coupling. The predicted directivity loss due to phase errors based on individual amplifier phase variations [12] is 2%. A measurement of the directivity loss was made with a near-field measurement of the radiation from the array. Calculation of the phase and amplitude variations in the plane of the array shows a 4% decrease to the combining efficiency due to these nonuniformities.

A system-level loss analysis of the array also agrees with the measured performance. The input feed line and dividers cause approximately 2-dB loss, found from individual component measurements. Based on the average amplifier gain, input power, and above output losses, the predicted output power is 6.9 W with a system gain of 28.4 dB and combining efficiency of 89%. The close comparison of these loss calculations with the above-measured results provides a degree of consistency with the analysis and measurements.

A final measure of consistency is found from the calculation of combining efficiency versus frequency. The amount of MMIC power supplied versus frequency can be found by overlaying the calculated impedance looking toward the antenna from the reference plane (see Fig. 1) on the load–pull measurement (see Fig. 4). In this figure,

[2]PCAAD, ver. 3, Antenna Design Associates, Leverett, MA.

contour lines show the MMIC output power as a function of load impedance. This approximation is valid since the output power and contour lines are approximately constant over the bandwidth of the antenna elements. By repeating the overlay process for all 16 elements and accounting for the circuit and directivity loss, the effective radiated power is calculated and converted to combining efficiency, as shown in Fig. 2.

V. Conclusion

This paper described a spatial power-combined array using a multilayered packaging architecture. Efforts to maximize combining efficiency were undertaken due its direct impact on overall system performance. Measured results agree well with predicted values, demonstrating the concepts and implementation.

Acknowledgment

The authors would like to thank J. Delisle, Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, and L. Kushner, Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, for useful comments and suggestions, and S. Robertson, Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, for assembly of the arrays. The authors are also grateful to D. Snider, Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, and R. Bauer, Massachusetts Institute of Technology (MIT) Lincoln Laboratory, Cambridge, for their continued support and encouragement.

References

[1] R. A. York, "Quasioptical power combining," in *Active and Quasi-Optical Arrays for Solid-State Power Combining*, R. A. York and Z. B. Popović, Eds. New York: Wiley, 1997, pp. 1–48.

[2] A. Alexanian and R. A. York, "Broadband spatially combined amplifier array using tapered slot transitions in waveguide," *IEEE Microwave Guided Wave Lett.*, vol. 7, pp. 42–44, Feb. 1997.

[3] N. Cheng, T. Dao, M. G. Case, D. B. Rensch, and R. A. York, "A 60-watt X-band spatially combined solid-state amplifier," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1999, pp. 539–542.

[4] M. A. Gouker, J. T. Delisle, and S. M. Duffy, "A 16-element subarray for hybrid-circuit tile-approach spatial power combining," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 2093–2098, Nov. 1996.

[5] E. A. Sovero, J. B. Hacker, J. A. Higgins, D. S. Deakin, and A. L. Sailer, "A Ka-band monolithic quasioptic amplifier," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1998, pp. 1453–1456.

[6] J. Hubert, L. Mirth, S. Ortiz, and A. Mortazawi, "A 4 watt Ka-band quasioptical amplifier," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1999, pp. 551–554.

[7] J. T. Delisle, M. A. Gouker, and S. M. Duffy, "45-GHz MMIC power combining using a circuit-fed, spatially combined array," *IEEE Microwave Guided Wave Lett.*, vol. 7, pp. 15–17, Jan. 1997.

[8] J. J. Sowers, D. J. Pritchard, A. E. White, W. Kong, O. S. A. Tang, D. R. Tanner, and K. Jablinsky, "A 36W, V-band, solid state source," in *IEEE MTT-S Int. Microwave Symp. Dig.*, 1999, pp. 235–238.

[9] M. A. Gouker, "Toward standard figures-of-merit for spatial and quasi-optical power-combined arrays," *IEEE Trans. Microwave Theory Tech.*, vol. 43, pp. 1614–1617, July 1995.

[10] S. M. Duffy, "Design and analysis of microwave/millimeter wave active arrays using a multilayered packaging architecture," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Massachusetts, Amherst, MA, 1999.

[11] S. C. Cripps, "A method for the prediction of load-pull power contours in GaAs MESFETs," in *Proc. IEEE MTT-S Int. Microwave Symp.*, 1983, pp. 221–223.

[12] M. I. Skolnik, "Nonuniform arrays," in *Antenna Theory*, R. E. Collin and F. J. Zucker, Eds. New York: McGraw-Hill, 1969, p. 233.

Design of Waveguide Narrow-Wall Short-Slot Couplers

Louis W. Hendrick and Ralph Levy

***Abstract*—It is shown that the broad-band coupling properties of narrow-wall short-slot directional couplers and hybrids are due to the effect of the evanescent TE_{30} mode in the coupling region. This causes a change to the effective phase length of the even-mode circuit, especially when a central tuning element is used. Using the results of a previous paper on the formation of an equivalent circuit from the generalized scattering matrix, an equivalent circuit for the even mode of the coupler is formed, clearly depicting the effect of the TE_{30} mode.**

***Index Terms*—Directional couplers, hybrid junctions, waveguide couplers.**

I. Introduction

The motivation for this paper arose from the requirement for a large number of narrow-wall WR51 directional couplers for an antenna feed network operating in the 20-GHz band with coupling values in the range of 3–8 dB. Existing design theories in the literature do not appear to have the required accuracy. Initially, a finite-element field analysis program (HFSS) was used in an attempt to optimize the dimensions of the four-port circuit, but this proved to be too slow and failed to converge in a reasonable time. It was necessary to develop means to design the couplers reasonably quickly with only minor degrees of optimization.

A major first step is to perform the HFSS analysis for the two-port odd and even modes of the symmetrical structure shown in Fig. 1, the circuit theory of symmetrical four-ports being described in [1]. The device becomes a good directional coupler when the individual modes are simultaneously matched. The coupling to port 3 is given by

$$S_{31}\ [\mathrm{dB}] = 20 \log_{10} \left[\sin(\theta/2) \right] \tag{1}$$

where θ is the S_{21} phase difference between the two modes. Hence, the condition for this to be a 3-dB coupler or hybrid is that θ is close to 90° [2].

The odd mode has an electric wall at the symmetry plane and is simple to match. It is also simple to analyze using circuit methods. However, the even mode has a magnetic wall in the slot region and is a relatively complex circuit.

Some work on the junction between the terminating waveguides and the even-mode region has been reported in the literature [3]–[5]. Reference [3] applies only to common walls of zero thickness, while the results are presented in a form not readily applicable to an engineering problem. The attempt to use the results in [2] appears to be incomplete because it fails to take into account the major effect of interactions between the discontinuities, which turns out to be of greater importance than any variation or compensating effect of the phase of the reflected waves at the ends of the coupling region. The same argument applies to [4, Sec. 7.7], which (as intended) gives the equivalent circuit of the discontinuity without any information on higher order mode interactions. Reference [5] has been found to give useful results for 3-dB hybrids

Manuscript received June 16, 1999.

L. W. Hendrick is with Hughes Aircraft Company, El Segundo, CA 90009 USA.

R. Levy is with R. Levy Associates, La Jolla, CA 92037 USA.

Publisher Item Identifier S 0018-9480(00)08717-2.

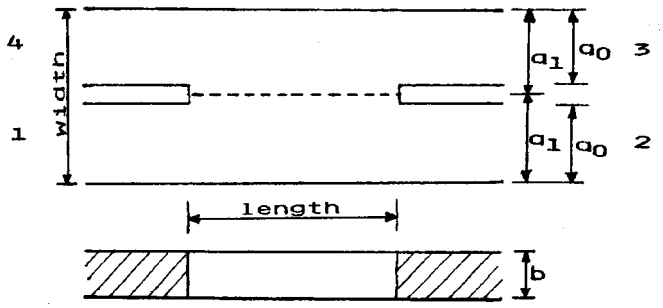


Fig. 1. Waveguide narrow-wall short-slot coupler.

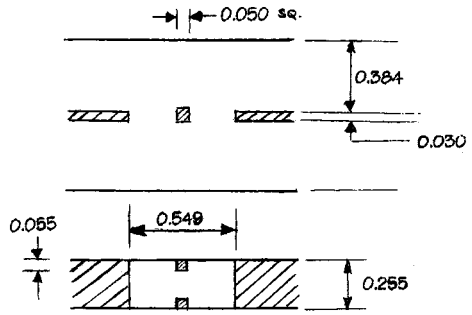


Fig. 2. Dimensions of WR51 short-slot hybrid.

with slight deviations in coupling and band centering, but the error increases for coupling values looser than 3 dB. This is due to more significant interactions when the coupling slot becomes shorter, as discussed in Sections III and IV.

A description of a somewhat different version of the short-slot coupler where the coupling slot is less than the full height of the waveguide is given in [6]. In this case, the TE_{30} mode is propagating in the coupling region, which is not the case for the "normal" short-slot coupler considered here. The frequency dependence of the coupling of this version of the short-slot coupler [6] is not particularly flat, varying almost linearly over the band. Other interesting contributions [7], [8] describe multislot narrow wall couplers having good coupling characteristics. Further description and other references may be found in [9].

Evidently, the design of a short-slot coupler is a situation where modern field theory software may be applied to analyze the even- and odd-mode circuits, followed by derivation of equivalent circuits in order to facilitate realization of a given coupling value with balanced coupling over the designated frequency band.

Returning to consideration of the coupler of Fig. 1, the width of the coupling region is usually considerably less than twice the width of the terminating waveguides, and transitions are used to match into the odd-mode circuit. Thus, for the Ku-band hybrid to be discussed here, the slot width is 0.798 in and transitions are required to match from the 0.51-in-wide input waveguides. Inevitably their presence complicates the analysis and the basic theoretical work concentrated on solving the problem without transitions, terminating the coupler in nonstandard narrow-width waveguides. The transitions may be incorporated as a later refinement.

II. Hybrid Design for the 18.25–19.75-GHz Band

Initial attempts to analyze the even-mode circuit were based on the formation of transfer matrices of the component parts of the circuit, including the transition region, uniform coupling regions, and central capacitive matching element (necessary to match the even-mode circuit). The product of the transfer matrices should then equal the overall transfer matrix determined from the overall scattering matrix of the even mode. However, it was found that the transfer matrix formed from the component parts of the circuit did not equate to the overall matrix formed by direct HFSS analysis. Evidently there has to be some evanescent-mode interaction taking place.

The dimensions of a typical hybrid are shown in Fig. 2, and the scattering transfer matrices (STMs) of the component parts of the even- and odd-mode circuits, i.e., the end discontinuities, connecting waveguides, and central capacitive matching element, were formed and multiplied together in the appropriate order to form overall STMs. The same operation was performed on the circuits without the capacitive matching element to give extra information, which proved to be of some use.

The results for the coupling of the hybrid with the matching element included are given in Fig. 3, which compares the coupling derived from

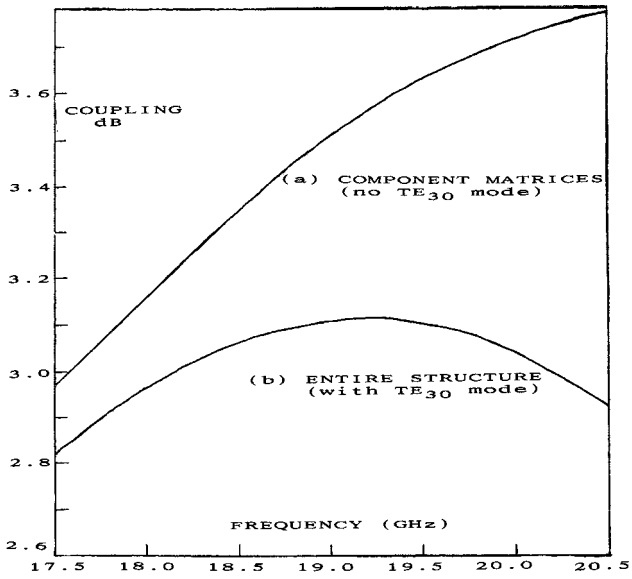


Fig. 3. Analyses of Ku-band hybrid: (a) by multiplication of two-port scattering transfer matrices and (b) by analysis of the entire even- and odd-mode circuits.

the component STMs with the HFSS analysis of the entire even- and odd-mode circuits, the latter, of course, representing the correct performance. The deviation of the 31 couplings is seen to be quite large, and increases with increasing frequency. Inspection of the even- and odd-mode circuits showed that this was due almost entirely to the lack of agreement between the even-mode circuits and, in particular, to a phase difference between the values of S_{21}, as shown in Fig. 4 (plot labeled "With Cap."). This phase difference displays a monotonic increase with increasing frequency, varying from just over 2° at 17.5 GHz to 10.7° at 20.5 GHz. This is the reason for the worsening deviation with increasing frequency of the 31 couplings, shown in Fig. 3, since the hybrid coupling is dependent mainly on the phase difference between the even and odd modes, as given by (1).

Calculations were also carried out without the central matching element, and the deviations between the transfer phases of the even modes formed from the component matrices and from the entire structure are also plotted in Fig. 4 ("No Cap."). It is seen that the deviations are now much less. This is a significant result, demonstrating the more severe interactions between the ends of the coupler and matching element.

III. Interaction Hypothesis

It soon became evident that the significant interaction effects seen only in the even mode must be due to the TE_{30} mode below cutoff at 22.185 GHz, which is not too far above the band of operation. At

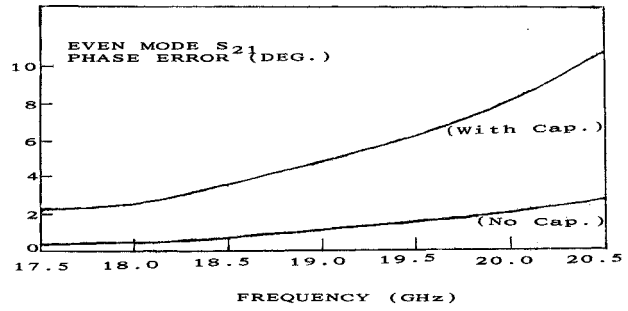


Fig. 4. Even-mode phase error due to TE_{30}-mode interaction for the Ku-band hybrid with and without the central matching element.

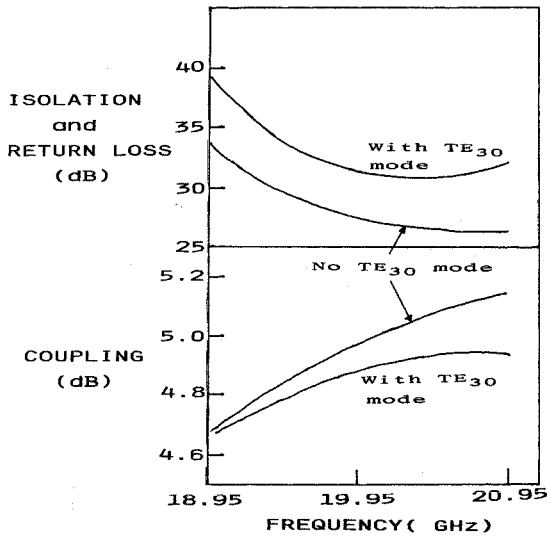


Fig. 5. Coupling, return loss, and isolation of the 4.77-dB coupler with and without the TE_{30}-mode interaction.

the upper frequency of 20.5 GHz, the attenuation per unit length of the TE_{30} mode, as given by

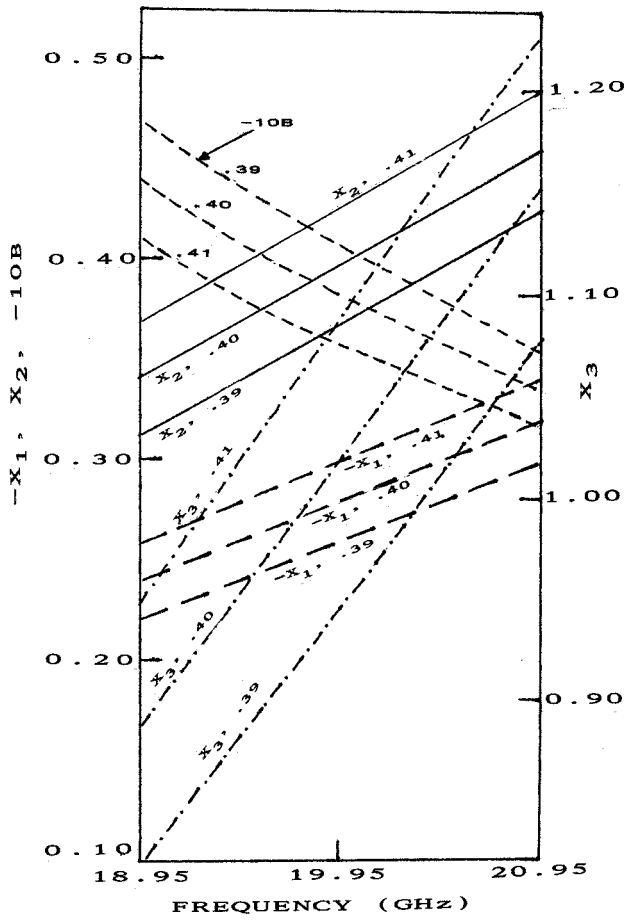
$$\mathrm{Atten.} = 54.6/\lambda_g \text{ dB} \tag{2}$$

where the guide wavelength of the mode below cutoff, $j\lambda_g$, is 39.2 dB/in. Thus, within the slot of length 0.549 in, the attenuation is 21.5 dB, equivalent to a voltage reduction of 0.084. This is quite significant, but at the center of the coupling region is much larger since, here, the attenuation is only 10.75 dB, corresponding to a voltage reduction of 0.29.

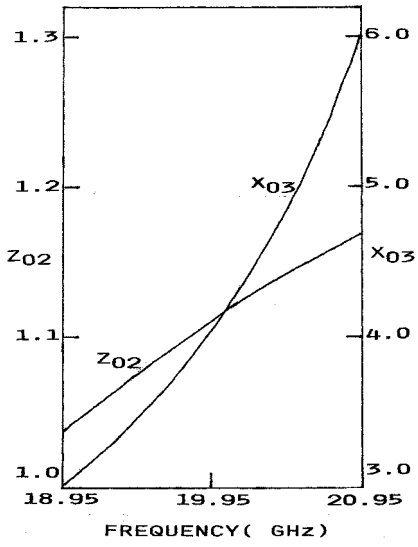
In the case of the odd-mode circuit, the next higher mode is TE_{40} with a cutoff at 29.58 GHz. At 20.5 GHz, the attenuation here is 98.64 dB/in, which is comparatively insignificant. There is no large central discontinuity in the odd mode to be of any concern, and the voltage interaction between the slot ends is only 0.002.

IV. Design of a 4.77-dB Coupler with No Central Matching Element

It has been found possible to design couplers with couplings in the approximate range of 4.5–7 dB without a central matching element. In this range, it is possible to choose a slot width to match out the discontinuities at the ends of the even-mode circuit. The form of the nominal 4.77-dB coupler is shown in Fig. 1, with dimensions of $a_0 = 0.400$ in, $a_1 = 0.395$ in, and $b = 0.05$ in with a slot length of 0.495 in. The theoretical coupling and isolation are shown in Fig. 5, derived from a complete analysis of the even- and odd-mode circuits, with and without the



(a)



(b)

Fig. 6. (a) Equivalent circuit normalized element values X_1, X_2, X_3, and B for the coupler for the 4.77-dB coupler. ———: X_1, ——: X_2, –·–·–: X_3, – – – –: $-10 \times B$. (b) Normalized impedance values Z_{02} and jX_{03}.

TE_{30}-mode interaction. In the exact case the analysis indicated an isolation >30 dB over the operating frequency band of 18.95–20.95 GHz. The same coupler was treated as an example in [10], which presents the theory for the analysis of the even-mode circuit to include the TE_{30} mode, followed by the generation of an equivalent circuit. Since couplers of coupling looser than 3 dB have slots of shorter length, the interaction effects of the TE_{30} mode are quite pronounced, even when there is no central matching element.

The equivalent circuit of the coupler takes the form shown in [10, Fig. 5]. Values of the elements in this equivalent circuit for several values of the half-slot width a_1 are shown in Fig. 6(a). It is seen that the series circuit elements X_1, X_2, and X_3 have linear frequency dependence over the 18.95–20.95-GHz band, while the shunt element B has smooth monotonic variation with frequency. The characteristic impedances Z_{02} and X_{03} shown in Fig. 6(b) are also monotonic, with X_{03} rapidly increasing at higher frequencies as the cutoff of the TE_{30} mode is approached. It is obvious from [10, Fig. 5] that the higher order mode has more effect at higher frequencies since then the attenuation of the parallel path is less.

The coupling obtained from multiplication of the individual scattering transfer matrices neglecting the TE_{30} mode, as shown in Fig. 5, demonstrates the same increasing deviation with increasing frequency (compared to the precise analysis) as the hybrid described in Section III. Calculation of the S_{21} phase error, as also described in Section III, gives a monotonic increase from 0.134° to 1.697° over the band, which agrees with the observed coupling differences. The coupling and isolation of the coupler derived from the equivalent circuit of [10, Fig. 5] is virtually identical to that obtained by direct analysis given in Fig. 5 here, confirming the validity of the equivalent circuit.

V. Conclusion

The flat coupling obtainable in the classic short-slot coupler is due to the effect of the evanescent TE_{30} mode in the coupling region, which has a major effect on the performance. Actually, it appears to be difficult or even impossible to obtain flat coupling unless this mode is significant, requiring a defined width of the coupling region so that the mode cutoff frequency is correctly located. The frequency variation of the phase "error" caused by the TE_{30} mode has an opposite slope to that of the dominant-mode phase, and the resulting phase compensation results in flatter coupling.

An equivalent circuit for the even mode, based on the formation of equivalent circuits from the generalized scattering matrix, has been derived, and all of the circuit elements have monotonic variation with frequency. The effect of the TE_{30} mode is clearly indicated as a parallel path connecting the input and output ports of the equivalent circuit.

Future work could be carried out to express the elements of the equivalent circuit in terms of dimensions, probably requiring an application of mode-matching theory. Operations on the equivalent circuit, such as new reference plane locations, may be required.

References

[1] J. Reed and G. J. Wheeler, "A method of analysis of symmetrical four port networks," *IRE Trans. Microwave Theory Tech.*, vol. MTT-4, pp. 246–252, Oct. 1956.

[2] H. J. Riblet, "The short-slot hybrid junction," *Proc. IRE*, vol. 40, pp. 180–184, Feb. 1952.

[3] J. F. Carlson and A. E. Heins, "The reflection of an electromagnetic plane wave by an infinite set of plates, I," *Quart. Appl. Math.*, vol. 4, pp. 313–329, Jan. 1947.

[4] N. Marcuvitz, Ed., *Waveguide Handbook*. ser. MIT Radiation Lab.. New York: McGraw-Hill, 1951, vol. 10.

[5] N. Moriguchi, K. Rokushima, and S. Mori, "Properties of the short slot hybrid junction," *J. Inst. Elect. Commun. Eng.*, vol. 47, pp. 1–2, Feb. 1964.

[6] T. Tanaka, "Ridge-shaped narrow wall directional coupler using TE_{10}, TE_{20}, and TE_{30} modes," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-28, pp. 239–245, Mar. 1980.

[7] H. Schmeidel and F. Arndt, "Field theory design of rectangular waveguide multiple-slot narrow-wall couplers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-34, pp. 791–798, July 1986.

[8] K. Tomiyasu and S. B. Cohn, "The Transvar directional coupler," *Proc. IRE*, vol. 41, pp. 922–926, July 1953.

[9] R. Levy, "Directional couplers," in *Advances in Microwaves*. New York: Academic, 1966, vol. 1, pp. 115–209.

[10] ——, "Determination of simple equivalent circuits of interacting discontinuities in waveguides or transmission lines," *IEEE Trans. Microwave Theory Tech.*, vol. 48, pp. 1712–1716, Oct. 2000.

Nonlinear Gain Compression in Microwave Amplifiers Using Generalized Power-Series Analysis and Transformation of Input Statistics

Hector Gutierrez, Kevin Gard, and Michael B. Steer

***Abstract*—Two methods are presented for the estimation of gain compression generated by a digitally modulated carrier passed through a nonlinear RF circuit. The first one is based on developing an analytical expression for gain compression based on the transformation of input signal statistics. The second one is based on approximated expressions derived from generalized power-series analysis. The techniques are evaluated by comparing measured and predicted gain compression in a CDMA system.**

***Index Terms*—CDMA, gain compression, intermodulation distortion, microwave amplifiers, nonlinear power amplifiers.**

I. Introduction

Gain compression in nonlinear amplifiers with digitally modulated signals occurs at lower powers than that for one-tone signals. This is attributed to the nonunity peak-to-average ratio (PAR) of digitally modulated signals, with most of the compression being due to the peak signals. The PAR is a crude characterization of a digitally modulated signal, and a more accurate representation is to use the auto-correlation statistics of the signal or to treat the signal as being composed of a large number of uncorrelated tones. This paper develops expressions for gain compression with digitally modulated signals using both characterizations above. Gain compression is derived using a statistical approach previously used for arriving at the statistics of the signal at the output of a nonlinearity given the statistics of the input signal [1].

The alternative expression uses earlier research to evaluate the nonlinear response to a multitone signal [2], [3]. This expression has been

Manuscript received July 27, 1999. This work was supported in part by the Defense Advanced Research Projects Agency under Agreement DAA-01-96-K-3619.

H. Gutierrez is with the Department of Mechanical and Aerospace Engineering, Florida Institute of Technology, Melbourne, FL 32901 USA.

K. Gard is with Qualcomm Inc., San Diego, CA 92121 USA.

M. B. Steer is with the Center for Advanced Computing and Communication, North Carolina State University, Raleigh, NC 27695 USA.

Publisher Item Identifier S 0018-9480(00)08718-4.

modified to account for the tones being uncorrelated. Gain compression predicted using the two techniques is contrasted and compared to measurement.

II. Gain Compression of a Complex Gaussian Signal Passed Through a Nonlinear Device

A block diagram of the baseband equivalent quadrature modulator cascaded with a nonlinear element is shown in Fig. 1 [1], with the digital quadrature modulated signal $z(t)$ modeled as the quadrature addition of $x(t)$ and $y(t)$ as

$$\tilde{z}(t) = A(t)e^{j\theta(t)} = x(t) + jy(t). \tag{1}$$

The nonlinear device is represented by the complex power series

$$\tilde{G}(z) = \sum_{i=1}^{N} \tilde{a}_i \tilde{z}^i = \tilde{a}_1\tilde{z} + \tilde{a}_3\tilde{z}^3 + \tilde{a}_5\tilde{z}^5 + \ldots + \tilde{a}_N\tilde{z}^N \tag{2}$$

the coefficients of which are extracted from conveniently obtained AM–AM and AM–PM data, following conversion of an envelope characterization to a baseband equivalent characterization (see Section III). It is important to realize that (2) represents a general nonlinear transfer characteristic, not a gain expression. The procedure followed here is to calculate the auto-correlation function of the output of this system by applying an input signal, characterized by its complex auto-correlation function, to a model of the nonlinearity described by (2). This section derives a general gain expression for a modulated carrier passing through a memoryless bandpass nonlinearity, considering carrier effects. Following the development in [4], an amplitude- and phase-modulated carrier $w(t)$ with carrier frequency ω_c is modeled as

$$w(t) = A(t)\cos[\omega_c t + \theta(t)] = \frac{1}{2}\left[\tilde{z}(t)e^{j\omega_c t} + \tilde{z}^*(t)e^{-j\omega_c t}\right] \tag{3}$$

where $A(t)$ and $\theta(t)$ are the amplitude and phase modulation, and $z(t)$ is the quadrature signal (1). Using the binomial expansion, the nth power of $w(t)$ is

$$w^n(t) = \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\left[\tilde{z}(t)\right]^k\left[\tilde{z}^*(t)\right]^{n-k}e^{j\omega_c(2k-n)t}. \tag{4}$$

Consider now only the frequency terms centered at the carrier frequency (this is usually referred as the first zonal filter at the output of the nonlinearity). This implies $2k - n = \pm 1$ for odd n only. Equation (4) then becomes

$$w^n(t) = \frac{1}{2^{n-1}}\binom{n}{\frac{n+1}{2}}\left[\tilde{z}(t)\tilde{z}^*(t)\right]^{((n-1)/2)}\tilde{z}(t). \tag{5}$$

Combining (4) with the complex transfer characteristic (2) yields the bandpass nonlinear gain expression

$$\tilde{G}(z) = \tilde{z}(t)\sum_{n=1}^{N}\frac{\tilde{a}_n}{2^{n-1}}\binom{n}{\frac{n+1}{2}}\left[\tilde{z}(t)\tilde{z}^*(t)\right]^{((n-1)/2)}. \tag{6}$$

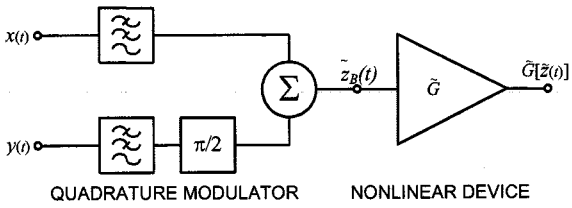


Fig. 1. Baseband equivalent quadrature modulated nonlinear amplifier.

The output auto correlation is then obtained by taking the moments of (6) as follows:

$$\begin{aligned}\tilde{R}_{gg}(\tau) &= E\left[\tilde{G}(\tilde{z}_1)\tilde{G}^*(\tilde{z}_2)\right] \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N}\frac{\tilde{a}_n\tilde{a}_m^*}{2^{n+m+1}}\binom{n}{\frac{n+1}{2}}\binom{m}{\frac{m+1}{2}} \\ &\quad\times E\left[\tilde{z}_1\tilde{z}_2^*(\tilde{z}_1\tilde{z}_1^*)^{((n-1)/2)}(\tilde{z}_2\tilde{z}_2^*)^{((m-1)/2)}\right]\end{aligned} \tag{7}$$

where N is the order of the polynomial model and n and m are integer indexes. The evaluation of (7) requires finding a closed-form expression for the expected value

$$E\left[\tilde{z}_1^{(n+1/2)}(\tilde{z}_1^*)^{((n-1)/2)}\tilde{z}_2^{(m-1/2)}(\tilde{z}_2^*)^{((m+1)/2)}\right]. \tag{8}$$

A general expression for this expectation can be induced by expanding (8) for several values of n and m, and collecting terms of equal order. This yields a closed-form expression for gain compression

$$\begin{aligned}\tilde{R}_{gg}(\tau) &= \tilde{R}_{zz}(\tau)\sum_{n=1}^{N}\sum_{m=1}^{N}\frac{n!m!}{2^{n+m-2}} \\ &\quad\times\binom{n}{\frac{n+1}{2}}\binom{m}{\frac{m+1}{2}}R_{zo}^{((n+m-2)/2)}\tilde{a}_n\tilde{a}_m^* \\ &= \tilde{R}_{zz}(\tau)\left[\sum_{n=1}^{N}\frac{n!}{2^{n-1}}\binom{n}{\frac{n+1}{2}}\tilde{a}_n R_{zo}^{((N-1)/2)}\right] \\ &\quad\times\left[\sum_{n=1}^{N}\frac{n!}{2^{n-1}}\binom{n}{\frac{n+1}{2}}\tilde{a}_n R_{zo}^{((N-1)/2)}\right]^*\end{aligned} \tag{9}$$

where $R_{zz}(\tau)$ and $R_{gg}(\tau)$ are the input and output auto-correlation functions, respectively, and $R_{zo} = R_{zz}(0)$ is a magnitude proportional to the average power of the input signal $z(t)$. This is a more general result than [1, eq. (9)] since none of the expected values of the cross products (8) are assumed to be zero.

III. Nonlinear Gain Compression Based on Generalized Power-Series Analysis

Previously, the authors have developed a formula for the phasors at the output of a nonlinear system described by a generalized power series (GPS) with an N-tone input [2], [3]. The GPS yields a deterministic expression for the output of a polynomial nonlinearity, considering a multitone input signal. When the input has a very large number of frequency components (such as in a CDMA signal), evaluation via power-series analysis might become computationally prohibitive, and

a statistical description of the input–output transformation such as the one outlined in Section II is necessary.

A phasor component of the output Y_q with radian frequency f_q can be expressed as a sum of intermodulation products U_q [2], [3]

$$Y_q = \sum_{n=0}^{\infty} \sum_{\substack{n_1,\ldots,n_N \\ |n_1|+\ldots+|n_N|=n \\ n_1 f_1+\ldots+n_N f_N=f_q}} U_q(n_1,\ldots,n_N) \tag{10}$$

where a set of n_k's define an intermodulation product, n is the intermodulation order, and

$$U_q(n_1,\ldots,n_N) = K(n_1,\ldots,n_N)\Big[1+T'(n_1,\ldots,n_N)\Big] \tag{11}$$

is given in terms of the intermodulation term K, which, for $n \neq 0$, is

$$K(n_1,\ldots,n_N) = \tilde{a}_n \frac{n!}{2^n} \prod_{k=1}^{N} \frac{(X_k^+)^{|n_k|}}{|n_k|!} \tag{12}$$

where X_k is the phasor of the kth input tone, and the saturation term T'

$$T'(n_1,\ldots,n_N) = \sum_{\alpha=1}^{\infty} \sum_{\substack{S_1,\ldots,S_N \\ S_1+\ldots+S_N=\alpha}} \left(\frac{(n+2\alpha)!}{n!2^{2\alpha}}\right) \frac{\tilde{a}_{n+2\alpha}}{\tilde{a}_n} \times \prod_{k=1}^{N} \frac{|X_k|^{2S_k}|n_k|!}{S_k!(|n_k|+S_k)!}. \tag{13}$$

In the above, $X_k^+ = X_k$ for $n_k \geq 0$, and $X_k^+ = X_k^*$, its complex conjugate, otherwise. The S_k's are integer sequences defined by the summation index α.

Considering only first-order intermodulation $(n = 1)$ enables gain compression to be described. The correlated component of the output at frequency f_q due to a multitone input signal (of uncorrelated components) of equal amplitude $|X_k| = |X_q|$, $k = 1,\ldots,N$ is

$$Y_{qc} = K_{qc}(1+T'_{qc}) \tag{14}$$

where $K_{qc} = \tilde{a}_1 X_q$ and

$$T'_{qc} = \sum_{\alpha=1}^{\alpha_m} \frac{(1+2\alpha)!}{2^{2\alpha}} \frac{\tilde{a}_{1+2\alpha}}{\tilde{a}_1} \sum_{\substack{S_1,\ldots,S_N \\ S_1+\ldots+S_N=\alpha}} \prod_{k=1}^{N} \frac{X_q^{2S_k}}{(1+S_k)!}. \tag{15}$$

Now, if the input is a single tone $(q = N = 1)$, (14) and (15) reduce to

$$Y_1 = \tilde{a}_1 X_1 + \sum_{\alpha=1}^{\alpha_m} b_{1+2\alpha}|X_1|^{2\alpha} X_1 \tag{16}$$

where the $\tilde{a}$ and b coefficients are related by

$$\tilde{a}_{1+2\alpha} = b_{1+2\alpha} \frac{2^{2\alpha}\alpha!(1+\alpha)!}{(1+2\alpha)!}; \quad \alpha = 1,\ldots,\alpha_m. \tag{17}$$

Thus, (16) is an envelope behavioral model with the effect of the carrier embedded in the model coefficients. However, an instantaneous model or baseband equivalent behavioral model, with coefficients $\tilde{a}_n$ obtained using (17) is required to determine the effects of multitone input signals, including gain compression. The standard nonlinear gain expression (16) relates a phasor at the input to the phasor at the output, and corresponds to the complex gain measured in AM–AM and AM–PM characterization.

The evaluation of (13) in CDMA systems can be computationally prohibitive due to the large number of frequency components required to accurately represent the modulated signal. An asymptotic expression for the compression term T' can be obtained by assuming that the tones of the input signal are uncorrelated and the number of tones N is sufficiently large. Thus, the asymptotic saturation term T'_{qc} is

$$T'_{qc} = \frac{1}{\tilde{a}_1} \sum_{\alpha=1}^{\alpha_m} \tilde{a}_{1+2\alpha} \frac{(1+2\alpha)!}{\alpha!2^{1+2\alpha}} |X_q|^{2\alpha} \tag{18}$$

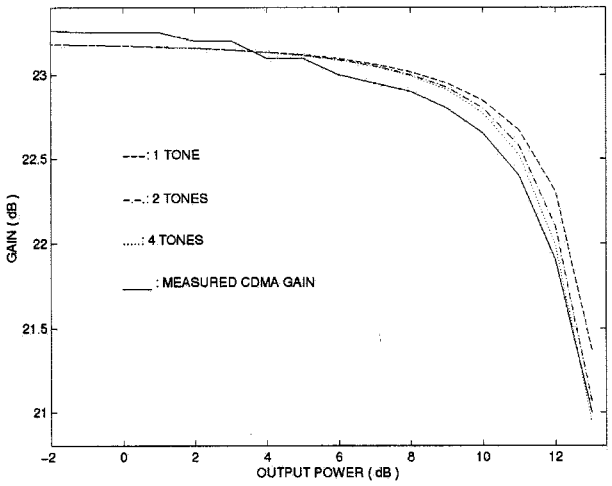


Fig. 2. Prediction of nonlinear compression based on GPS analysis of (13).

and is called the asymptotic gain-compression expression. This expression does not use the statistical properties of the phase since (18) is an approximation. On the other hand, the method outlined in Section II considers phase characteristics when computing the auto correlation of the output signal since phase is included in the signal model [see (3) and (4)].

IV. Results and Discussion

A. Behavioral Model

The AM–AM and AM–PM curves for a 900-MHz CDMA power amplifier with 23-dB gain were obtained using a single-frequency input power sweep and a vector network analyzer [1], [5]. A complex power series of order 13 was fitted to the AM–AM and AM–PM data from a 900-MHz CDMA power amplifier, using least-squares optimization. The resulting envelope behavioral model was then transformed into the baseband equivalent instantaneous behavioral model using (17). The power-series model was verified by comparing measured and predicted AM–AM and AM–PM data, as shown in [1].

B. Prediction of Gain Compression

The baseband equivalent behavioral model (1) was first used with the GPS model (13) to predict gain compression of the digitally modulated carrier, by simulating the expression for the saturation term (T') at different input power levels.

Due to the computational complexity involved in evaluating (13), only a few tones were used. Results are shown in Fig. 2, where the solid line represents measured CDMA gain, and the dashed and dotted lines correspond to compression factor [T' in (13)], calculated for several multifrequency inputs whose power match the measured available power, scaled by the linear gain coefficient $\tilde{a}_1$. The one-tone result corresponds identically to the measured AM–AM characteristic since it was measured using a single tone. The amplitudes of the phasors in (13) were derived from the input signal power, considering that tones were uncorrelated, so that the input power was divided equally among the tones. This assumption is not sufficient when predicting spectral regrowth [1]. Compression is expected to occur at lower output power levels in CDMA signals than in signals that only include a few tones. This effect is not properly predicted by this model, partly due to the least-squares process used in determining model coefficients, resulting in a small residual error in the first-order coefficient. The asymptotic expression that assumes a large number of input tones (18) was also

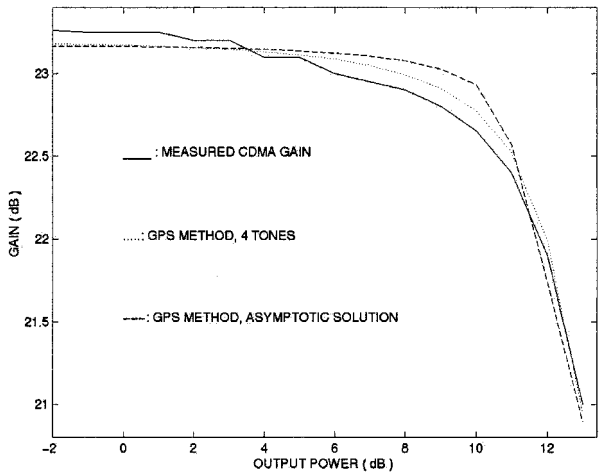


Fig. 3. Prediction of nonlinear compression based on GPS analysis. The asymptotic expression, i.e., (18), is compared to the exact solution of (13).

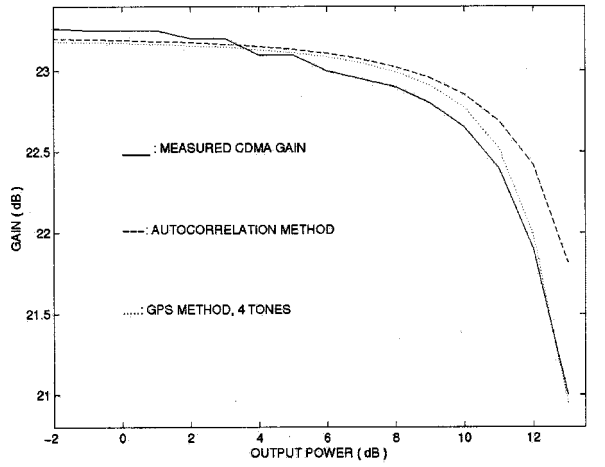


Fig. 4. Prediction of nonlinear compression. The exact solution of the GPS analysis of (13) is compared to the autorrelation transformation method of (9).

simulated and compared to the exact result for four tones, and to gain measurements. Results are shown in Fig. 3. The asymptotic gain compression (18), while computationally less demanding than the exact solution (13), yields less accurate estimates than the four-tone exact solution due to the approximations involved in (18).

The accuracy of the modeled gain-compression characteristic increases with the number of uncorrelated input tones considered. This provides important insight into why gain compression with digitally modulated signals occurs at lower power levels than with a single tone. With many uncorrelated input tones, gain compression is dominated by a saturation effect [T' in (13)]. Gain compression was also simulated using the auto-correlation method (9), and compared both to gain measurements and the four-tone simulation obtained from the exact GPS solution of (13). Results are shown in Fig. 4. The two methods track each other up to weak gain-compression levels. At high output power levels, the statistical method underestimates compression since the output auto-correlation expression [1] could only be expanded up to ninth order due to the computational complexity involved. Similarly, the GPS prediction tends to underestimate compression since higher order intermodulation components $(n > 1)$ in the output signal were neglected.

V. Conclusion

Two methods for the estimation of gain compression of digitally modulated signals passed through a nonlinear amplifier, modeled by a complex power series, have been presented in this paper. The first one uses statistical properties of the moments of the modulated signal to develop a simple expression for gain compression. The second, based on GPS analysis, provides two different expressions to predict gain compression, one based on the exact GPS solution (13), and another one based on assuming a very large number of tones [see (19)]. Both methods are qualitatively well suited to predict output power levels at which compression occurs, and the exact GPS solution can be used favorably even with relatively small numbers of tones. GPS analysis has also been used to derive an expression that relates an envelope behavioral model (for which parameters are readily available from sweep continuous wave (CW) tests) with an instantaneous power-series nonlinear transfer characteristic such as (2). These methods provide computational tools to the amplifier designer to predict gain compression given a polynomial model of the device that is readily available from AM–AM and AM–PM data.

References

[1] H. Gutierrez, K. Gard, and M. B. Steer, "Spectral regrowth in nonlinear amplifiers using transformation of signal statistics," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 1999, pp. 830–839.

[2] M. B. Steer, C. R. Chang, and G. W. Rhyne, "Computer-aided analysis of nonlinear microwave circuits using frequency domain nonlinear analysis techniques: The state of the art," *Int. J. Microwave Millimeter-Wave Computer-Aided Eng.*, vol. 1, no. 2, pp. 181–200, 1991.

[3] M. B. Steer and P. J. Khan, "An algebraic formula for the output of a system with large-signal, multifrequency excitation," *Proc. IEEE*, vol. 71, pp. 177–179, Jan. 1983.

[4] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*. New York: Plenum, 1992, sec. 2.11.7.2.

[5] V. Aparin, K. Gard, G. Klemens, and C. Persico, "GaAs RFICS for CDMA/AMPS dual-band wireless transmitters," in *IEEE MTT-S Int. Microwave Symp. Dig.*, June 1998, pp. 81–84.

[6] J. F. Sevic, M. B. Steer, and A. M. Pavio, "Nonlinear analysis methods for the simulation of digital wireless communication systems," *Int. J. Microwave Millimeter-Wave Computer-Aided Eng.*, vol. 6, no. 3, pp. 197–216.