COMSAT
Technical Review

COMSAT TECHNICAL REVIEW

Volume 15 Number 2A, Fall 1985

A 4-GHz QPSK modulator for high bit rates

K. KARIMULLAH AND R. G. EGRI

(Manuscript received December 6, 1984)

Abstract

A 4-GHz microwave integrated circuit (MIC), quadrature phase-shift keying (QPSK) modulator has been developed for 120-Mbit/s data rates. The measured phase and amplitude balance are within ±1° and ±0.15 dB, respectively, over a 200-MHz bandwidth. The measured rise time of a typical phase-state transition is less than 1.5 ns. These features guarantee good performance at bit rates even higher than 120 Mbit/s.

The QPSK modulation is achieved by driving two balanced binary phase-shift keying (BPSK) modulators in parallel and using a 90°, 3-dB hybrid to combine their outputs. The BPSK modulators have been designed using a different approach, and each exhibits an excellent amplitude balance, ±0.1 dB, and a phase balance of ±0.5° over the 200-MHz bandwidth. The QPSK modulator is laid out on a 1.8 × 2.3-in. alumina substrate and is directly compatible with standard emitter-coupled logic (ECL) gate drivers. Its size, power consumption, and weight make it suitable for use in a spacecraft regenerative modem.

Introduction

Future communications satellites will employ high-capacity data transmission links using digital modulation. These links will employ satellite-switched, time-division multiple-access (SS/TDMA) techniques to help meet the ever-increasing demand for transponder use. In SS/TDMA systems, transponder interconnectivity is made possible by switching traffic between transponders

on request from the ground station. Although this switching can be done at RF [1], advances in high-speed digital technology make baseband switching a viable alternative. On-board demodulation is needed to accomplish baseband switching, and regeneration is needed to gain link performance advantages over the conventional frequency translation satellite system.

The channel performance of QPSK transmission and the simplicity of QPSK modulation/demodulation circuits make a QPSK regenerative modem a likely candidate for use in future communications satellite payloads. A typical modem will consist of an input receive filter, a QPSK demodulator, a QPSK modulator, and a transmit filter. Figure 1 shows a simplified configuration for the modem.
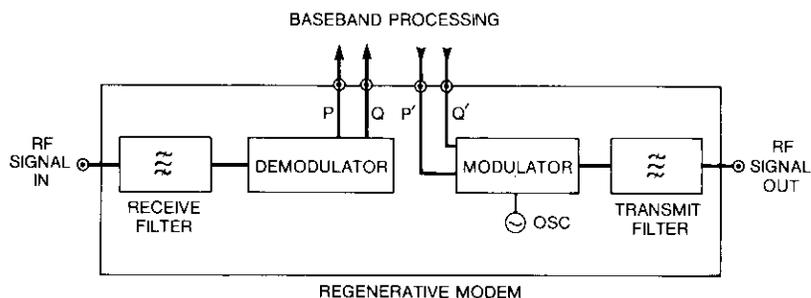

Figure 1.   *Simplified Block Diagram of a Regenerative Modem*

This paper describes the design, implementation, and testing of a 4-GHz MIC QPSK modulator for high-bit-rate applications. The QPSK modulation is obtained by combining the outputs of two parallel BPSK modulators through a 90° hybrid. The nominal output phases of the QPSK modulator are given in Table 1, where A and B are the data inputs to the BPSK modulators.

TABLE 1. PHASE OF THE QPSK MODULATOR AS A FUNCTION OF BPSK DATA INPUT

| INPUT | | RELATIVE |
A	B	PHASE
0	0	45°
0	1	135°
1	0	−45°
1	1	−135°

The following sections describe the design considerations leading to this approach, followed by details of the design. The performance of the QPSK modulator is then described by presenting experimental data on the static and dynamic behavior of phase-state transitions.

General design considerations

All designs for QPSK modulators at microwave frequencies have either a serial or a parallel configuration [2]-[5]. The serial configuration consists of a 0°/90° binary phase shifter cascaded to a 0°/180° binary phase shifter. This circuit gives good amplitude balance but may have undesirable switching transients in the output [2].

A more commonly used design is the parallel configuration shown in Figure 2. Here, the CW signal is divided by an in-phase, 3-dB hybrid and applied to identical BPSK modulators on both arms. The output of these modulators is combined by means of a phase quadrature hybrid such as a Lange coupler to produce the desired QPSK signal.
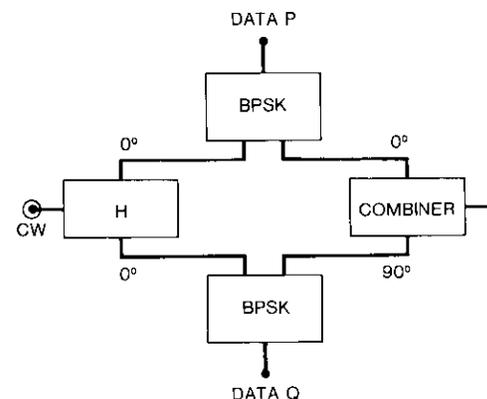
Figure 2.   *Parallel Configuration for Design of a QPSK Modulator*

The BPSK modulators which determine the performance of the QPSK modulators have been realized at C-band by using various techniques [6], [7]. There are two categories of BPSK modulators: the transmission type and the reflection type. A reflection modulator consists of either a circulator with one reflecting element or a hybrid with a matched pair of reflecting elements. Switching diodes are most frequently used to achieve strong reflections in forward and reverse bias modes. Transmission modulators connect the input

port to the output port via two switchable paths whose phase differential is 180°. Dual-gate FETs are well-suited as the switching elements for transmission modulators [8].

The choice between an FET modulator and a diode modulator is based on such factors as size, weight, and power consumption. In most applications, high RF output power-handling capability can be traded off against amplitude and phase balance. For such applications, diode modulators appear to be a better choice when thermal stability of the circuit is a critical requirement.

A parallel configuration (Figure 2) was adopted as the baseline design. The approach given in Reference 5 was used for the C-band implementation, with a slight variation. First, balanced BPSK modulators were designed using beam-lead Schottky diodes as reflecting elements at the arms of a Lange coupler. Then, a BPSK modulator with excellent amplitude balance (<0.1 dB variation) and a phase imbalance of ±0.5° over a 200-MHz bandwidth was developed.

BPSK modulator design

Reflection modulator configuration

Reflection modulators have been constructed by using microwave circulators [2] or an MIC 3-dB branch line or interdigital coupler. In both of these designs, the carrier-to-output port isolation depends on the isolation of the uncoupled ports. The isolation of microwave isolators is usually between 20 and 23 dB. The measured isolation of the Lange coupler, designed using the SUPER-COMPACT℗ computer program, was 26 dB or more over a 300-MHz bandwidth centered at 3.95 GHz. Since better carrier suppression in the output signal can be achieved by using a Lange coupler, that configuration was selected.

The underlying principle of the reflection modulator is to create reflections of equal amplitude and 180° phase difference, by the two logic levels. These reflections are created by driving the diodes that are connected to the arms of the coupler into two distinct states. Theoretically, a strong forward bias or reverse bias can create a true short circuit or a true open circuit, respectively, thereby achieving the required reflection coefficients. However, the finite values of the forward bias resistance (dynamic resistance) and the package capacitance could create an amplitude imbalance, and could also offset the phase in the two states. Additionally, the finite resistance affects the insertion loss (modulation loss) of the modulator. Because these effects worsen at higher frequencies, the choice of suitable diodes is a critical factor in the design of reflection modulators at microwave frequencies.

Schottky diode vs PIN diode

A beam-lead Schottky diode with the excellent RF characteristics needed to achieve strong reflections in the forward and reverse biased states has been reported [5], but is not commercially available. Most commercial Schottky diodes have a dynamic resistance of between 15 and 20 Ω. This gives reflection coefficients of nearly −2.7 to −3.7 dB, respectively. On the other hand, a beam-lead PIN diode will have a resistance as low as 5 Ω, giving a reflection coefficient of −0.9 dB. The package capacitance of both diodes is 0.12 to 0.15 pF. Comparison of the reflection coefficients alone would indicate that PIN diode modulators are a better choice, since the higher reflection coefficient results in greater reflected power at the output, and hence a lower modulation loss. However, from an overall design standpoint, PIN diode driver circuits are more complex than the Schottky diode driver circuit. A PIN diode has a turnoff delay caused by diffusion capacitance. To reduce the turnoff time, as would be required for high-bit-rate operation, the driver circuit applies a large reverse voltage (6 to 8 V) across the diode, and will therefore consume more power.

In Figure 3, a typical circuit that would drive a PIN diode [2] is compared with the circuit that drives the Schottky diodes of the BPSK modulator. Both circuits assume ECL data inputs. Clearly, if simplicity, low weight, and reduced power consumption are essential, the Schottky diode is a better choice.

Characterization of Schottky diodes

Beam-lead Schottky diodes (HP5082-2716), which are readily available as batch-matched diodes, were used in this modulator. These diodes (see Figure 4) have about 18 Ω of forward RF resistance and 0.15 pF of capacitance for moderate current (5 mA) and low reverse voltages (1 V), respectively. Typical characteristics are listed in Table 2. Other beam-lead Schottky diodes with similar RF characteristics would do equally well.

Initially, a test fixture (Figure 5) was constructed to compare the performance characteristics of a diode to the values provided by the manufacturer. The diode mounting configuration was made identical to the one used in the modulator. Three diodes were measured in the test fixture. The return loss of the connector to microstrip was first verified by through line, and was measured to be greater than 27 dB.

Next, circuit models of the diode and the test fixture were used in the SUPER-COMPACT℗ program to determine the reflection coefficients of the diode for forward and reverse bias conditions. A forward bias current of 5 mA and a reverse voltage of 0.5 V were used. The results were repeatable,

(a) Driver Used for the BPSK
Schottky Diode Modulator Design

(b) Typical PIN-Diode Modulator Driver
in High-Speed Applications [2]
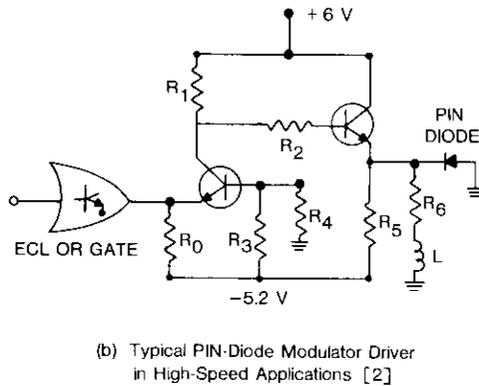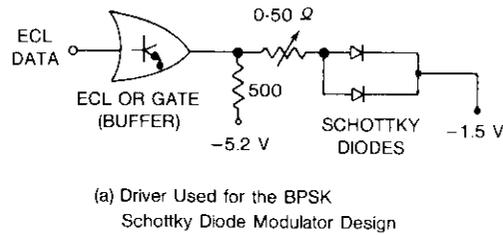
Figure 3.   *Comparison of Driver Circuits*
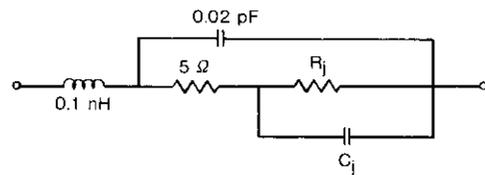
Figure 4.   *Schottky Diode Model*

TABLE 2. SCHOTTKY DIODE CHARACTERISTICS

|	MEASURED		SPECIFIED	
BIAS	R_j	C_j (pF)	R_j	C_j (pF)
Forward	13 Ω	—	15 Ω	—
Reverse	2.4 K	0.13	—	0.15
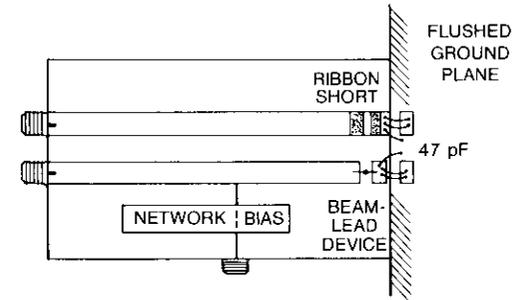
Figure 5.   *Test Fixture for Diode Characteristics Measurement*

but slight variations were observed between the measured and specified values. These values are also presented in Table 2.

The difference between the measured and specified values is a result of the mounting technique used. A standard mounting method is used for the specifications, while in the test fixture the mounting configuration conforms to the geometry of the BPSK modulator.

The average reflection coefficients obtained by measuring the three diodes are shown in Figure 6. In the reverse bias condition, the magnitude of the reflection coefficient is −1 dB, but in the forward bias condition it is a function of the bias current and ranges from −4 to −6 dB. Furthermore, these amplitudes remained within ±0.1 dB over the 200-MHz bandwidth around the 3.950-GHz center frequency. The fact that at any given frequency the phase deviation of the forward bias reflection did not exceed 3° as the amplitude varied from −4 to −6 dB allowed total freedom in adjusting the amplitude balance of the BPSK modulator.

Design approach

Once the diodes were properly characterized, a circuit (shown in Figure 7a) was devised to make the reflection coefficients in the forward and reverse bias equal in magnitude and 180° out of phase. The requirement for equal amplitudes was achieved by using a 200-Ω resistor in parallel with the diode. Under reverse bias conditions, the magnitude of the reflection coefficient is constant at nearly −4.5 dB. By adjusting the forward bias current, the magnitude of the forward bias reflection coefficient can also be brought equal to this value by using the variable resistor shown in Figure 3a. Amplitude balancing of the reflection coefficients in the forward and reverse bias states could have been achieved by using stubs (or possibly spiral inductors).

IMPEDANCE OR ADMITTANCE COORDINATES
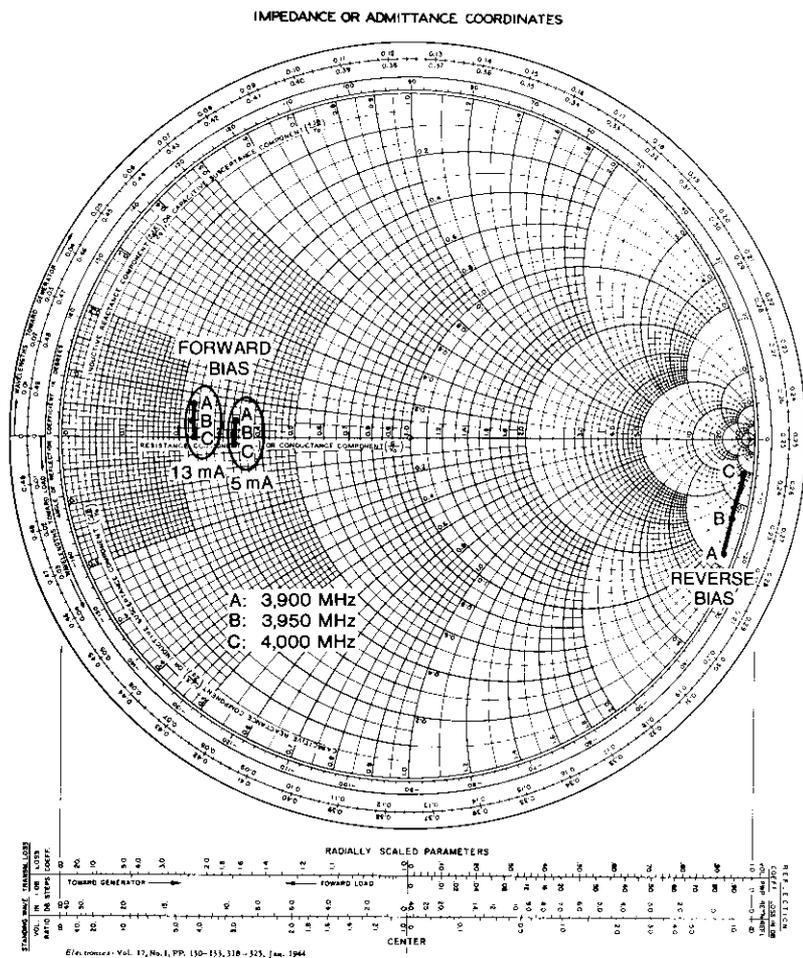
Figure 6.   *Average RF Characteristics of Diodes on the Smith Chart*

However, since the reactance of these components is frequency-dependent, their use would degrade the bandwidth of the modulator.

The 180° phase difference between the forward and reverse bias states was easily achieved by using a single stub tuner. Once the location and length of this stub were determined using the experimental values for the diode

(a)   Basic Circuit
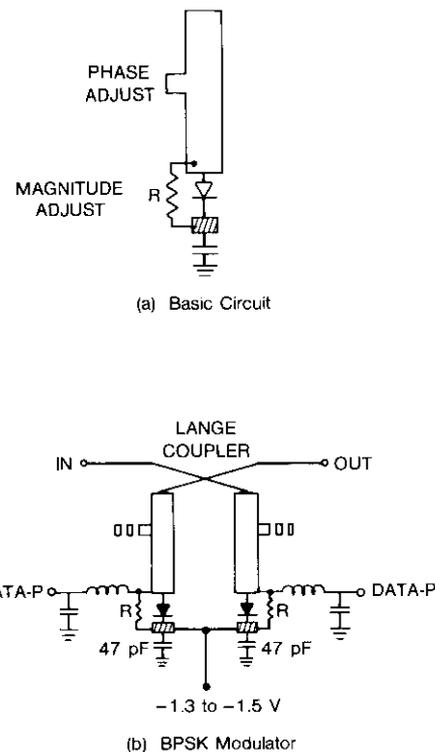
(b)   BPSK Modulator

Figure 7.   *Method of Achieving Amplitude and Phase Balance in a BPSK Modulator Design*

characterization, the BPSK configuration of Figure 7b was simulated on SUPER-COMPACT™, and these lengths and location were fine-tuned.

The circuit was laid out on an alumina substrate and tested on the network analyzers. After tuning, a phase imbalance of ±0.5° and amplitude imbalance of ±0.1 dB were measured over the 200-MHz bandwidth. This BPSK circuit was used as the basis for fabricating the QPSK modulator.

QPSK modulator

The QPSK modulator was designed using parallel configuration on alumina substrate. This microstrip circuit is shown in Figure 8. Since the modulator is ECL-compatible, the diodes have an RF ground (47-pF chip capacitor) and a −1.5-V DC bias at their cathodes.

ECL DATA

RF IN

RF OUT
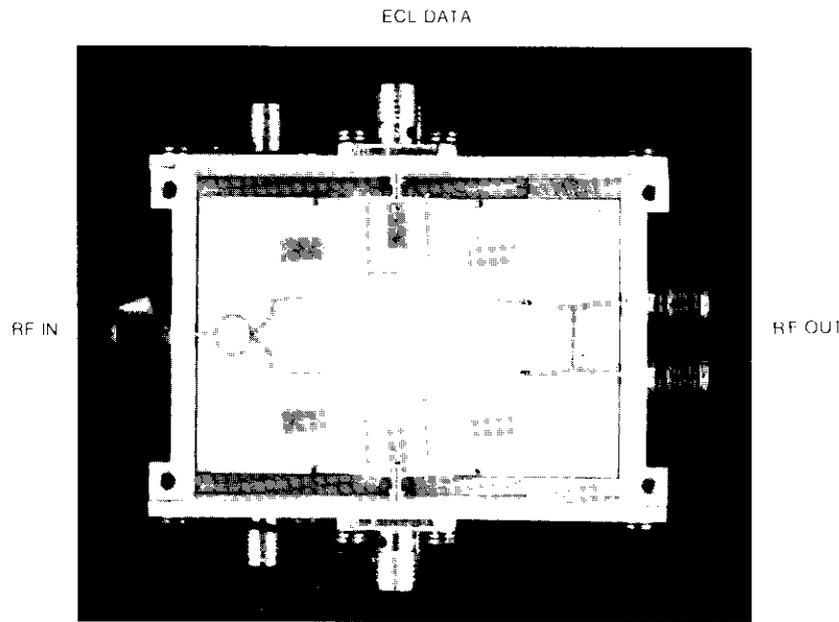
Figure 8.   *QPSK Modulator on 2.3 × 1.8-in. Alumina Substrate*

The QPSK modulator was tested in two modes: a static mode (100-Hz random data rate) to measure the amplitude and phase balance on the network analyzer, and a dynamic mode (60-MHz 1010 sequence) to determine the rise and fall times of the phase state transitions. The static characteristics of the modulator are shown in Figure 9. The test setup and measured dynamic characteristics are depicted in Figure 10. The latter were obtained by splitting the carrier into two paths, one going to the modulator and the other to the local oscillator port of a phase detector. The output of the modulator was applied to the RF port of the phase detector, and the resulting IF signal was observed on a fast oscilloscope. Figure 10a is a schematic of this measurement setup.

Rise time (<1.5 ns), which can be obtained from the pulse shown in Figure 10b and 10c, is a function of the modulator frequency response, the ECL data waveform, and the phase detector used. The phase detector IF port has a 500-MHz bandwidth, and the ECL driver has a 1.2-ns rise time. Considering these limitations, the contribution from the modulator itself is quite small.

(a) f_0 = 4 05 GHz

(b) f_0 = 3 85 GHz

(c) f_0 = 3 95 GHz
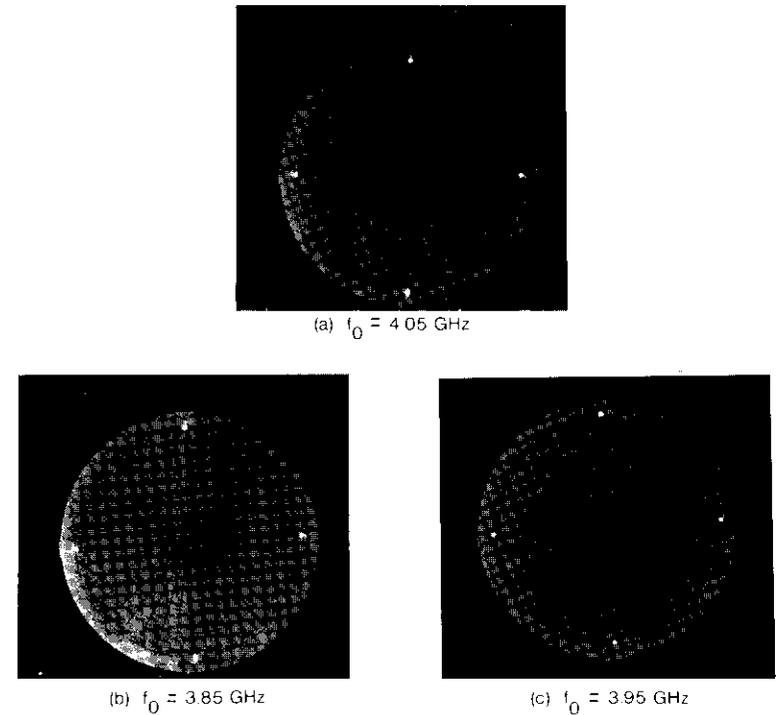
Figure 9.   *Network Analyzer Display for Very Low Rate Random Data*

The performance characteristics of the QPSK modulator are summarized in Table 3.

TABLE 3. PERFORMANCE CHARACTERISTICS OF THE 4-GHz QPSK MODULATOR

Center Frequency	3,950 MHz
Bandwidth	200 MHz
Modulation Loss	8 dB
Maximum RF Input	−10 dBm
Amplitude Imbalance	<±0.15 dB
Phase Imbalance	<±1°
Input Return Loss	≥26 dB
Output Return Loss	≥20 dB
Phase-State Transition	<1.5 ns
Carrier Suppression	≥30 dB

(b) 0° - 180° Transition

(c) 0° - 90° Cyclic Steps
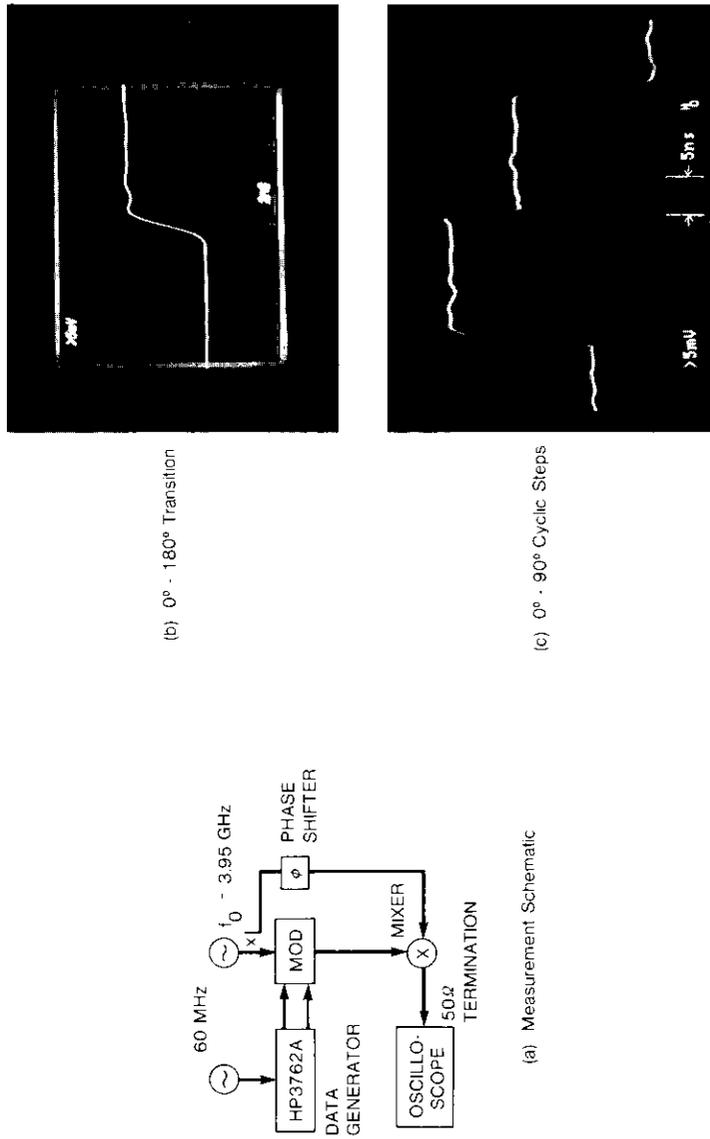
Figure 10. *Coherently Detected Waveforms for 0°–180° Transitions and 90°-Step Cyclic Transitions Using a 60-MHz Clock*

(a) Measurement Schematic

Summary

Modulators with 4-GHz MIC BPSK and QPSK have been fabricated using a new approach for obtaining a good amplitude and phase balance over a 200-MHz bandwidth. The QPSK modulator was designed to operate at a data rate of 120 Mbit/s. Since a transition in the output phase can be accomplished in less than 1.5 ns, this device can easily be used for higher bit rates.

The new approach is based on using Schottky diodes in a high-current (5 to 10 mA) mode in a reflection modulator configuration. The magnitude of the forward bias reflection coefficient and its reverse bias reflection coefficient are equalized by placing a fixed resistor in parallel with the diode. This method ensures a wide operating bandwidth (at the expense of some increase in insertion loss).

The QPSK modulator has an amplitude balance of ± 0.15 dB and a phase balance of $\pm 1°$ over the 200-MHz bandwidth centered around 3.95 GHz. The device is laid out on a 2.3×1.8-in. alumina substrate, and is directly drivable from standard ECL gates such as the MC1600 series by Motorola. The device draws about 7 mA of current from a standard ECL output. The size, power consumption, and weight make it suitable for spacecraft applications.

References

[1] F. Assal, R. Gupta, J. Apple, and A. Lopatin, "Satellite Switching Center for SS-TDMA Communications," *COMSAT Technical Review*, Vol. 12, No. 1, Spring 1982, pp. 29–68.

[2] G. Ohm and M. Alberty, "11-GHz MIC QPSK Modulator for Regenerative Satellite Repeater," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-30, No. 11, November 1982, pp. 1921–1926.

[3] Y. Chang et al., "High Data Rate Solid-State Millimeter Wave Transmitter Module," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-23, No. 6, June 1975, pp. 470–477.

[4] C. L. Cuccia and E. W. Matthews, "PSK and QPSK Modulators for Gigabit Data Rates," IEEE MTT International Microwave Symposium, San Diego, California, June 1977, *Digest*, pp. 208–211.

[5] H. Ogawa et al., "Integrated Balanced BPSK and QPSK Modulators for the K_a-Band," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-30, No. 3, March 1982, pp. 227–234.

[6] T. Matsumoto and M. Aikawa, "A High-Speed Balance Modulator Composed of Slot Lines," *Electronics and Communications in Japan*, Vol. J60-B, May 1977, pp. 350–357.

[7] R. B. Mouw, "A Broadband Hybrid Junction and Application to the Star Modulator," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-16, No. 11, November 1968, pp. 911–918.

[8] Y. Ayasli, "Microwave Switching With GaAs FETs," *Microwave Journal*, Vol. 25, No. 11, November 1982, pp. 61–74.

Khalid Karimullah received a B.E. from Karachi Engineering University, Pakistan, in 1973, and an M.S. and Ph.D. from Michigan State University in 1976 and 1979, respectively. From 1979 to 1980, he was an Assistant Professor in the Department of Electrical Engineering at Michigan State University. He joined COMSAT Laboratories in 1980 and is currently a Member of the Technical Staff in the Satellite Transponders Group in the Microwave Technology Division. He has been involved in the design of a differential QPSK demodulator and a coherent QPSK modem at microwave frequencies. His activities have also included the design of clock synchronization circuits and pulse-shaping channel filters, and communications channel optimization studies.

Robert G. Egri received a Dipl. Eng. degree from the Budapest Technical University in 1974, and a Ph.D. from the Johns Hopkins University in 1978. From 1974 to 1976, he worked in the Research Institute for Telecommunications in Budapest as a microwave circuit designer. Since 1978, he has been a Member of the Technical Staff in the Transponders Department at COMSAT Laboratories. His responsibilities have included the optimization and design of contiguous-band multiplexers, carrier and clock synchronization circuits, and pulse-shaping filters. He is a part-time faculty member at The George Washington University, where he teaches coding and communications theory.

A contour generation algorithm applied to antenna patterns

K. D. FULLETT AND V. E. RIGINOS

(Manuscript received June 26, 1985)

Abstract

Algorithms for the generation of contour plots as applied to antenna pattern data obtained during in-orbit tests of communications satellites are described. These contours are used for operational planning and to verify spacecraft in-orbit compliance with specifications. These algorithms have been developed to enable contour generation from data which are nonuniform and sparsely spaced along the azimuth and elevation planes in spacecraft antenna coordinates, conditions under which conventional contour generation algorithms are not entirely applicable. These simple algorithms generate contours which can be further enhanced with a graphics editor, are general in nature, and can be applied to other types of data.

Introduction

During the past few years COMSAT Laboratories has developed techniques for generating and displaying contour plots of satellite antenna patterns from measurements performed on satellites in orbit. These in-orbit measurements, performed after the launch of a new communications satellite, verify *in situ* compliance with specifications.

Several cuts from each spacecraft antenna beam are obtained by performing repeated antenna gain measurements while changing the azimuth and elevation angles [1],[2]. From these antenna cut data, a two-dimensional contour map is generated which represents the three-dimensional antenna gain surface.

The spacing of in-orbit measured data is very sparse compared to typical antenna pattern data measured on the ground. In addition, the data spacing is not homogeneous in either dimension because of the time and operational constraints imposed by the fuel needed to achieve different elevation angles and the necessity of transmitting a separate command for each azimuth angle. Because of this data sparseness and nonuniformity, known contour generation algorithms are not entirely applicable. As part of a spacecraft antenna pattern analysis package, new algorithms have been developed at COMSAT Laboratories to overcome this problem. While these algorithms were developed for use with antenna pattern data, they are also applicable to other data types. This paper presents a brief survey of existing contour generation algorithms as well as a discussion of the new algorithms.

Existing Algorithms

All existing contour generation algorithms [3],[4] can be divided into two general classes: algorithms that generate contours from data points located on a uniform grid, and those using unevenly spaced data points (i.e., not on a grid). The first class of algorithms applies to data originating from either a deterministic function or from antenna measurements performed on a slant range. This paper will be limited to a discussion of the second class of algorithms, which includes in-orbit test data.

Algorithms for nonuniformly distributed data [4],[5] utilize two steps: interpolation between data points to locate the position (in the two-dimensional plane) of a given contour value, and connection of the contour. Numerical interpolation between data points in one dimension is a well documented [6] and straightforward procedure. However, the problem of interpolation in two dimensions is more difficult. Several algorithms have been reported in the literature. Some of these algorithms use orthogonal polynomials [7], splines [8], trigonometric functions [9], and finite differences [6]. There is no universal algorithm because the *best* algorithm depends on the nature of the data. The algorithm described here uses the finite difference method.

Interpolation between nonuniform distributed points reveals the location of the contour points. Next, these points are connected by line segments, an operation referred to as *contouring*. There are two basic contouring methods: grid scan [10] and contour following [11]. Grid scan is simple to implement, although it does not easily distinguish between different contour values as required for labeling, and it does not always prevent different contours from crossing. Contour following avoids these problems but unfortunately exhibits instabilities unless the data are considerably more dense than in the in-orbit test case. The method implemented here is a modified contour-following technique.

The algorithms

Contour plots from antenna pattern data are generated in two steps: first, all points that belong on the contour are found; then the points on the contour are connected and the contour curves drawn. The first step requires information in the following form:

$$\text{Value, } D(*), G(*), \text{Gnpnt}$$

where *Value* = value of the contour to be found

$D(n,m,p)$ = a three-dimensional array containing the in-orbit test data. The array $D(*)$ typically has the form shown in Figure 1, where the integer indices n, m, and p are defined as:

n = index of a cut. Its range is $[1:n']$ where n' is the total number of cuts.

m = index of points along a cut n. Its range is $[1:m']$ where m' is the maximum allowed number of points in a cut.

p = the coordinate index defined as:
  1 = x coordinate of the cut
  2 = y coordinate of the cut
 3 = gain parameter associated with the x,y point.
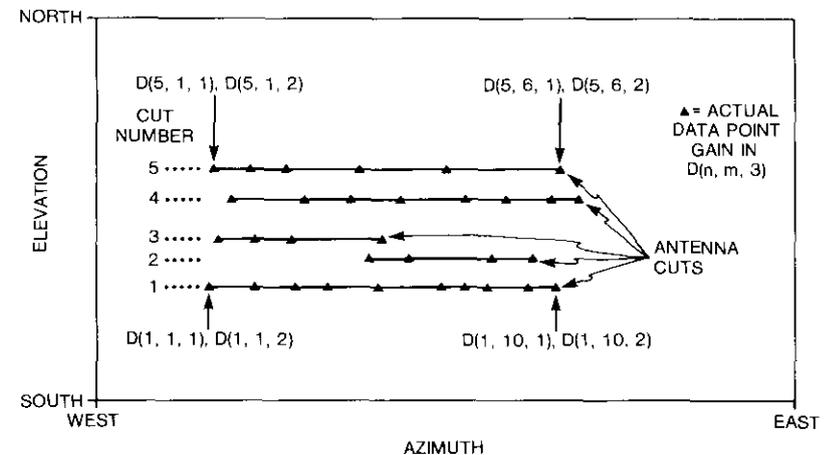
Figure 1.   *Example of the Input Array D(*)*

$G(m,n)$ = a two-dimensional array containing the points that lie on the contour.

m = index of contour points having value *Value*. Its range is $[1:m']$ where m' is the maximum number of points expected to be found. This number is chosen based on the memory limitations of the system.

n = coordinate index defined as:

1 = x coordinate of the found point

2 = y coordinate of the found point

3 = magnitude of the gradient at this point

4 = angle of the gradient at this point

5 = type of point flag, either extrapolated or interpolated

$Gnpnt$ = number of points found in $G(*)$ described above.

The cut array $D(*)$ is scanned four times to find all the locations where data points have a gain equal to the variable *Value*. The data points at these locations are the sought-after contour points. The scan process finds both interpolated and extrapolated points with gains equal to *Value*. The interpolated points are to be connected to form the contour curves. The extrapolated points are used as an aid to this connection process. When any point is found, either by interpolation or extrapolation, its gradient is also calculated by using neighboring points. In detail, the four scans used in the search process are as follows:

a. A scan along a cut in the x direction to look for crossings of the contour curve between the points of the cut.

b. A scan at either end of the cut using the first two points or the last two points to determine if a point can be extrapolated.

c. A look between adjacent cuts to determine if the contour curve crosses between them.

d. A search above and below the cuts in the y direction using the first two and last two cut pairs to find extrapolated points.

Each of these four scans finds points in different areas of the in-orbit testing measured cuts. The four scans are shown in Figure 2.

Scanning the antenna cuts

The first scan is performed along each cut stored in $D(*)$. That is, $D(n,m,p)$ is scanned for the values of $p = 1$ and $p = 2$ (x and y coordinates) where $p = 3$ (the gain) equals *Value*, while n is held constant and m is varied. Points are treated in pairs, starting with the pair consisting of points one and
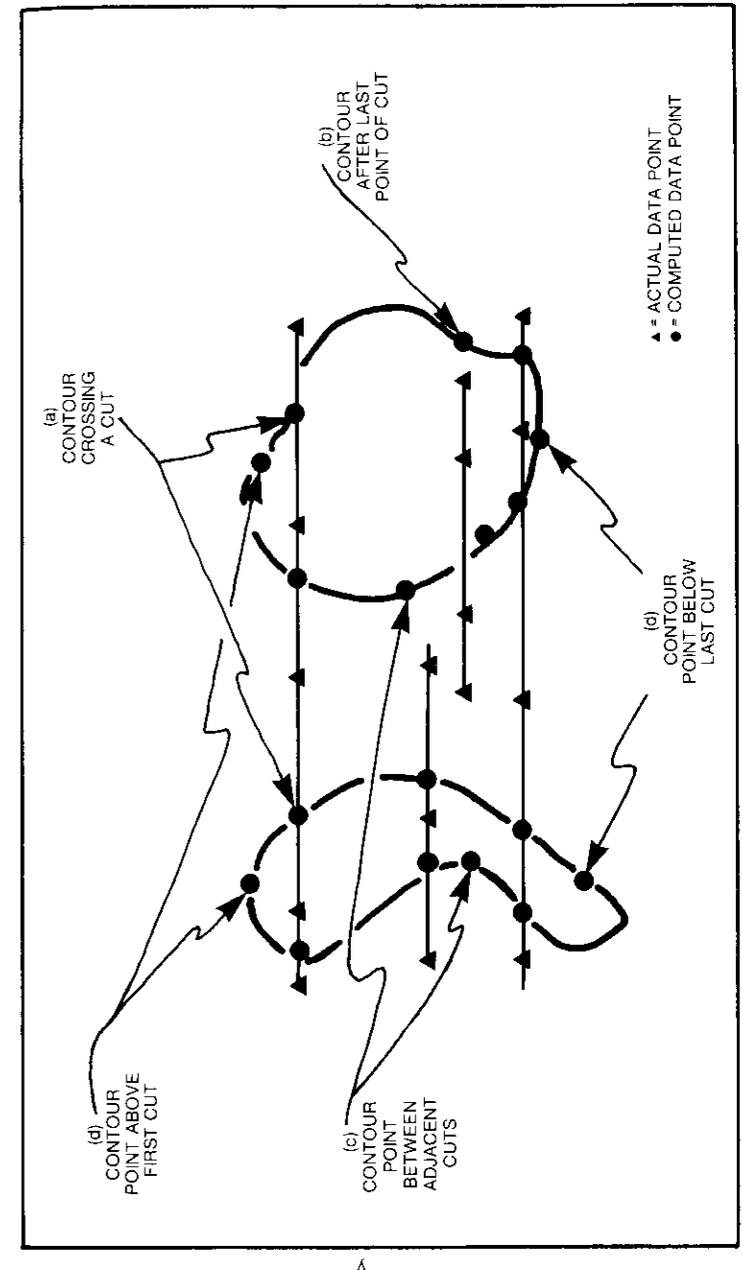
Figure 2.   *Points Found Using the Four Types of Scans*

two and continuing through the pair consisting of points $N - 1$ and N, where N is the number of valid points in the cut. Each pair of points defines a line. Since checking is done along the cut, the y coordinate will be constant and the problem is reduced to two dimensions. The values of interest are as follows:

X_1 = x coordinate of the first point
G_1 = gain coordinate of the first point
X_2 = x coordinate of the second point
G_2 = gain coordinate of the second point.

As stated before, this point pair defines a line where the equation of the line is

$$G = mX + b \tag{1}$$

where m = slope given by $(G_2 - G_1)/(X_2 - X_1)$
b = the intercept given by $G_1 - mX_1$.

The information desired is the X value at which G is equal to *Value*, that is, the location in the field of points described by $D(*)$ where the gain is equal to *Value*. Therefore, this equation must be rearranged as follows:

$$X = mG + b \tag{2}$$

where m = slope $(X_2 - X_1)/(G_2 - G_1)$
b = intercept $X_1 - mG_1$.

The value of X where G is equal to the desired contour value can always be found by using equation (2). However, this X, defined as X_0, may lie in any of three regions:

a. below the pair of points $(X_0 < X_1)$,
b. above the pair of points $(X_0 > X_2)$, or
c. between the pair of points $(X_1 \le X_0 \le X_2)$ inclusive.

The first two cases above indicate that the point was extrapolated; it does not lie between the two data points. The third case describes an interpolated point, the only points of interest during this search mode. If point X_0 must be found by extrapolation (cases a and b above), it is simply ignored because an extrapolated point about the point pair is meaningless except for the first two and last two points in the cut, which are handled separately.

Once an interpolated point is found (case c), it is necessary to calculate the gradient at this point (X_0, Y_0) by separately finding the gradient in the x

and y directions and combining the results. This procedure requires a second point with the same value of Y_0 along the x-axis, and a third point with the same value of X_0 along the y-axis, as shown in Figure 3.
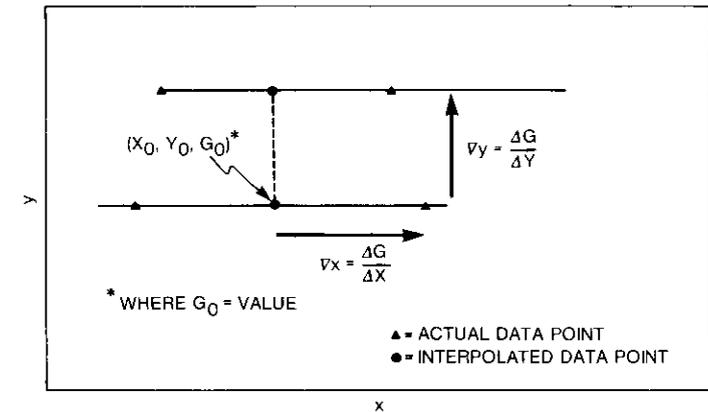
Figure 3.   *Gradient Computation*

In general, the vectors used to generate the gradient are broken into their X and Y components. The gradient along the x-axis is found by looking at the next available data point in the positive x direction. The gradient along the y-axis is found by looking in the positive y direction.

An example of the search process for an interpolated point within the raster of points follows. First, the cuts of constant elevation (constant Y) are scanned in turn. Within each cut, each pair of adjacent points is checked to see if the contour passes between them. The point with the lower X value is called *point one* and has coordinates X_1, Y_1, and G_1. The point in the pair with the higher X value is called *point two* and has coordinates X_2, Y_2, and G_2. The point at which the contour crosses between point one and two is called *point zero* and has coordinates of X_0, Y_0, and G_0. This relationship is illustrated in Figure 4, where

a. Y_1, Y_2 and Y_0 are all equal,
b. G_0 is the searched-for contour value, and
c. $X_1 \le X_0 \le X_2$.

The *point zero* is found by performing a linear interpolation using as inputs *point one*, *point two*, and the desired G_0, which must be equal to *Value*. The Y values are not used in the linear interpolation routine because they are all
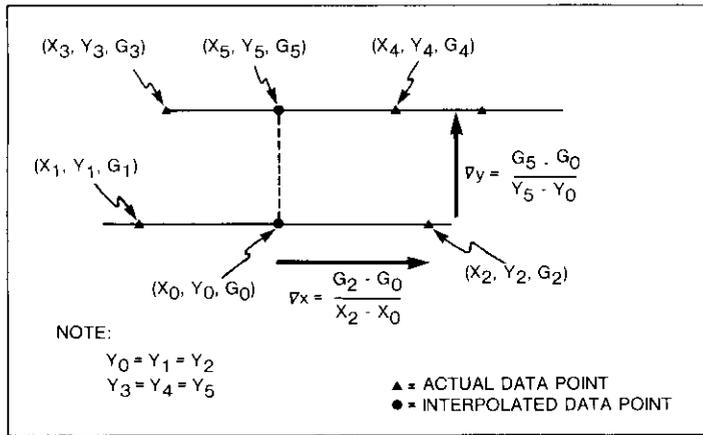
Figure 4.   *Gradient of Interpolated Points*

equal. Once the point has been found, its gradient is calculated. First the gradient in the x direction is found:

$$\nabla x = \frac{G_2 - G_0}{X_2 - X_0} \tag{3}$$

as shown in Figure 4. The gradient is represented by the arrow parallel to the x-axis. The tail of the arrow indicates the second parameter in equation (3) and the tip of the first parameter in equation (3). Thus, a positive gradient indicates a positive slope (uphill) when moving from the tail to the tip of the arrow and a negative gradient indicates a negative slope (downhill). Next, the gradient parallel to the y-axis is found. To do this, a point is needed that lies in an adjacent cut with an X value equal to X_0. First, two points are found in the adjacent cut. These will be called *point three* and *point four* with coordinates X_3, Y_3, G_3 and X_4, Y_4, G_4, respectively. These points are chosen such that X_0 lies between X_3 and X_4. Then, through linear interpolation, *point five* is found such that X_5 is equal to X_0, as shown in Figure 4. Then the gradient in the y direction is given by

$$\nabla y = \frac{G_5 - G_0}{Y_5 - Y_0} \tag{4}$$

Again, the tip of the arrow in the figure points to the first parameter in the

equation and the tail the second. Combining ∇x and ∇y to give the two-dimensional gradient in polar coordinates yields

$$\text{Magnitude Grad} = \sqrt{\nabla x^2 + \nabla y^2} \tag{5}$$

$$\text{Angle Grad} = \text{ATAN} \frac{\nabla y}{\nabla x} \tag{6}$$

It is important to note that equation (6) must provide the proper quadrant for the angle. That is, angles returned will be in the range of $0°$ through $360°$. This information is required later for use in connecting the contour points found.

If X_5 is greater than or less than all the points in the adjacent cut, then *point three* and *point four* cannot be found such that *point five* lies between them. When this occurs, it is necessary to extrapolate a point off the end of the adjacent cut. For simplicity, *point three* and *point four* are still found. If X_5 is less than all X values on the adjacent cut, then X_3 and X_4 will be found so that X_3 will be equal to X_5. Likewise, if X_5 is greater than all values in the adjacent cut, then X_3 and X_4 are found such that X_5 is equal to X_4. The algorithm used for finding the gradient can then continue as before.

Since there are a finite number of cuts, the antenna cut with the greatest Y value, by definition, does not have a cut with a greater Y value to be used as the adjacent cut. In this case, the adjacent cut is the cut with the next lower Y value. The top cut, or the cut with the greatest Y value, is easily found. Remember that the cuts are sorted by increasing Y. If the total number of cuts is N, then the last cut is located in $D(N,*,*)$.

EXTRAPOLATING BEYOND THE ENDS OF THE ANTENNA CUTS

The next step is to search for points that lie on the contour off the negative and positive extremes of each antenna cut. For each cut the first two and the last two points form point pairs that are used to extrapolate points in the contour. The first point pair for the cut in Figure 5 is given by

$$X_1 = D(\text{cut},1,1)$$
$$G_1 = D(\text{cut},1,3)$$
$$X_2 = D(\text{cut},2,1)$$
$$G_2 = D(\text{cut},2,3).$$

These values are passed to a linear interpolation routine that looks for the contour value on the line defined by these two points. The resulting value is called *point zero*, with coordinates X_0, G_0 and Y_0, where all Y values are
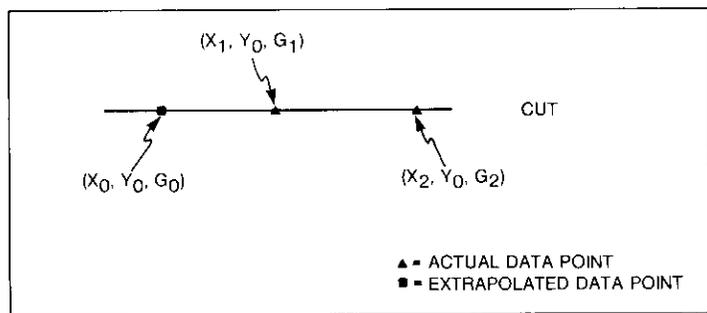
Figure 5.   *Extrapolated Points Off the End of a Cut*

equal because they are all along the same cut. Coordinate X_0 is valid only if $X_0 < X_1$. If X_0 is greater than or equal to X_1, then the point lies within data that already exist and is not valid. The process of finding points beyond the negative x end of an antenna cut is illustrated in Figure 5.

Extrapolation beyond the upper end of the antenna cut is performed in a similar fashion as that for the lower end. If the number of points in the cut is given by M, then the point pairs at the positive end of the cut used to define a line are given by

$$X_1 = D(\text{cut}, M\text{-}1, 1)$$
$$G_1 = D(\text{cut}, M\text{-}1, 3)$$
$$X_2 = D(\text{cut}, M, 1)$$
$$G_2 = D(\text{cut}, M, 3).$$

The criterion for a valid point is that the resulting X_0 must be greater than X_2. When an extrapolated point is found, the gradient is calculated for that point using the technique explained earlier for interpolated points within a cut.

LOOKING BETWEEN CUTS FOR POINTS

The search for points that lie on the contour continues with the third step. In this step, a search is performed for points that lie on the contour and are between adjacent antenna cuts. This search repeats the steps used to look between points along a cut, but interpolation is performed along the y-axis instead of along the x-axis. The algorithm is slightly different due to variations in the input data format.

In summary, the algorithm considers each point on each antenna cut and looks to an adjacent cut to find a point that can be used for interpolation along the y-axis. The interpolation is performed, and if a valid point is found, the gradient is calculated. The search begins by choosing an actual data point on an antenna cut (see Figure 6) called *point zero*, which has the following coordinates:

$$X_0 = D(\text{cut}, \text{point}, 1)$$
$$Y_0 = D(\text{cut}, \text{point}, 2)$$
$$G_0 = D(\text{cut}, \text{point}, 3)$$

where cut is within the range of valid cuts in $D(*)$ and point is within the range of valid points in the antenna cut.
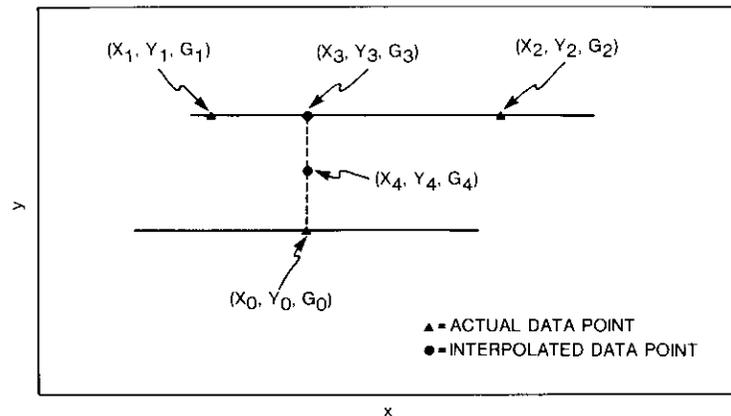
Figure 6.   *Interpolated Points Between Two Cuts*

Next, two points are found called *point one* and *point two*, with coordinates X_1, Y_1, G_1 and X_2, Y_2, G_2, respectively. *Point one* and *point two* are found such that $X_1 \leq X_0 \leq X_2$. *Point one* and *point two* will be found even if extrapolation is required beyond the end of the antenna cut.

Next, *point three* is found between *point one* and *point two* by interpolating on X and G such that $X_3 = X_0$. The value of Y is ignored during this interpolation because $Y_1 = Y_2$. Figure 6 shows the following relationships:

$$Y_1 = Y_2$$
$$X_3 = X_0$$
$$X_1 \leq X_0$$
$$X_2 \geq X_0.$$

Point four is found by interpolating between *point zero* and *point three*. This interpolation is performed on the variables Y and G since X is constant. The interpolation searches for the contour *Value* that occurs on the line defined by *point zero* and *point three*. This point, which has coordinates X_4, Y_4, G_4, is valid only if Y_4 lies between Y_3 and Y_0. If Y_4 lies outside of the range specified by Y_0 and Y_3, it is ignored at this time. The relationships between the points in addition to those specified above are

$$X_4 = X_3 = X_0$$
$$G_4 = \text{the contour } Value$$
$$Y_4 \text{ lies between } Y_0 \text{ and } Y_3.$$

If contour *point four* is a valid point, then the gradient must be calculated (see Figure 7). Again, the gradient is calculated by finding the X and Y vectors separately and combining them. The x direction vector at *point four*
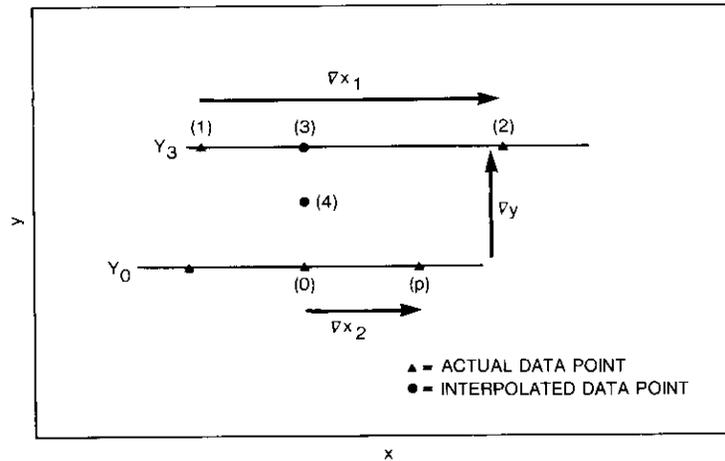

Figure 7. *Gradient Calculations Looking Between Two Cuts*

is found by combining two x direction gradient vectors, one along the antenna cut with value Y_0 and a second along the antenna cut with value Y_3. The first X gradient, called ∇x_1, is found along the Y_0 antenna cut. The two points used to define the gradient vector are *point zero* and an adjacent point. The adjacent point normally used is that of the next higher X value. If a point of

higher X value does not exist in the cut, then X_0 is the last point in the cut and the next lower X value is used. Let *point zero* be the mth point in the Y_0 antenna cut. Then the coordinates of these points are given by

$$X_0 = D(\text{cut}, m, 1)$$
$$Y_0 = D(\text{cut}, m, 2)$$
$$G_0 = D(\text{cut}, m, 3).$$

If the number of points in the antenna cut Y_0 is M and m is not equal to M, then the adjacent point, called p, is *point m + 1*. If M is equal to m, then the adjacent point is $M - 1$. The first contribution to the X gradient along the Y_0 antenna cut is given by

$$\nabla x_1 = \frac{G_p - G_0}{X_p - X_0} . \tag{7}$$

The second X gradient is found along the Y_3 axes and is always given by

$$\nabla x_2 = \frac{G_2 - G_1}{X_2 - X_1} . \tag{8}$$

These two gradients are combined, giving the final X gradient for *point four*:

$$\nabla x = \frac{\nabla x_1 + \nabla x_2}{2} . \tag{9}$$

The Y gradient at *point four* is always given by

$$\nabla y = \frac{G_3 - G_0}{Y_3 - Y_0} . \tag{10}$$

Equations (9) and (10) are then used along with equations (5) and (6) to give the gradient at *point four* in polar coordinates.

This search is performed between all the given antenna cuts. For the first cut, that is, the cut with the minimum Y or elevation value, the search is performed for increasing Y only. For the last cut, that is, the cut with the largest Y value, the search is performed looking down to the cut with the next lowest Y value. For all other cuts, the search is performed looking at the adjacent cut in both the forward (increasing) and backward (decreasing) y directions, as shown in Figure 8.
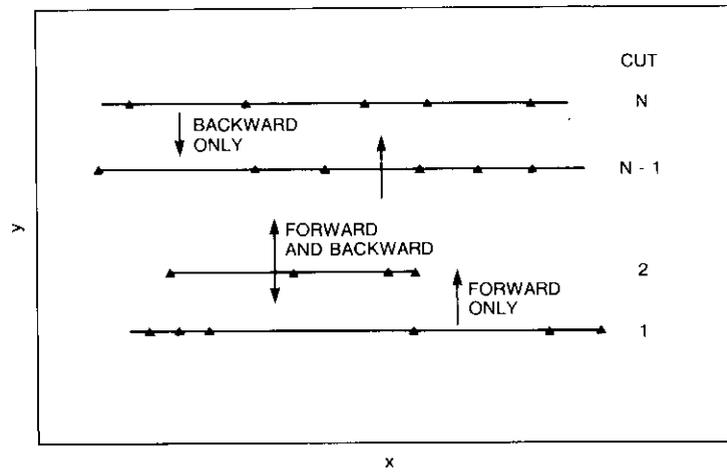
Figure 8.   *Direction of Searches Along Antenna Cuts*

SEARCHING ABOVE AND BELOW THE CUTS FOR EXTRAPOLATED POINTS

The final step in searching for points that lie on the contour is to look above and below the set of cuts (step 4 of Figure 2) for extrapolated points with the value of *Value*. This search repeats the steps performed while looking beyond the ends of the cuts (along the *x* direction) in $D(*)$ using the first and last point pairs. It also uses the first and the last cut pairs to look beyond the data in the *Y* direction and to determine if any points in that region lie on the contour (Figure 9).

The first cut pair processed is cut one and cut two, which are the cuts with the lowest *Y* values. A point called *point zero* with coordinates X_0, Y_0, G_0 is found in cut one. Next, a pair of points called *point one* and *point two* are found in cut two such that $X_1 \leq X_0$ and $X_0 \leq X_2$. This is the same procedure used previously for finding points in an adjacent cut. As explained before, it is always possible to satisfy this condition although it may be necessary for either *point one* or *point two* to be extrapolated.

Given *point one* and *point two*, another point can be found along cut two. This point, called *point three*, lies between *point one* and *point two* such that $X_0 = X_3$. Finally, as before, a linear interpolation routine is used to find a point called *point four* with coordinates X_4, Y_4, G_4 such that G_4 is equal to the desired *Value*. Figure 9 illustrates the following conditions:

a. X_0, X_3, and X_4 are all equal,
b. Y_1, Y_2, and Y_3 are all equal,
c. $X_1 \leq X_3 \leq X_2$

and that *point four* will lie in one of the following three regions:

a. $Y_4 < Y_0$
b. $Y_4 > Y_3$
c. $Y_0 \leq Y_4 \leq Y_3$.

For the example shown in Figure 9, region c corresponds to an interpolated point and is of no interest. Region b corresponds to a point that is either above all other cuts in the *y* direction or within the range of the other cuts in the *y* direction. In either case, it does not provide any additional information. Therefore, only points that lie within region a are of interest, and these are points that lie outside the current set of cuts.
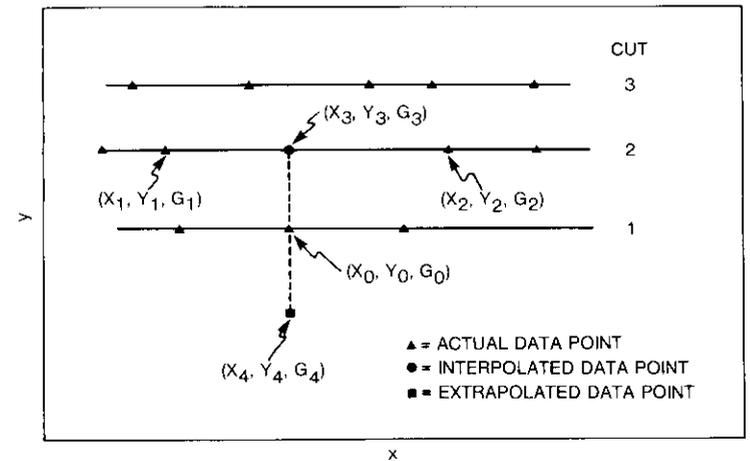
Figure 9.   *Extrapolating Points Below Cuts*

If *point four* is in the proper region, then the gradient must be found. The points used for the gradient calculation (the same technique depicted in Figure 7) are shown in Figure 10. *Point one* and *point two* are used to calculate ∇x_2. *Point zero* and an adjacent point are used to calculate ∇x_1. *Point zero* and *point three* are used to calculate ∇y. Finally, these are combined using equations (9), (10), (5), and (6) to give the gradient in polar coordinates.

The above example presents only the first search performed for extrapolated points. The search is repeated by moving along cut two and finding points that lie on cut one. Similarly, if the number of cuts is N, the procedure is
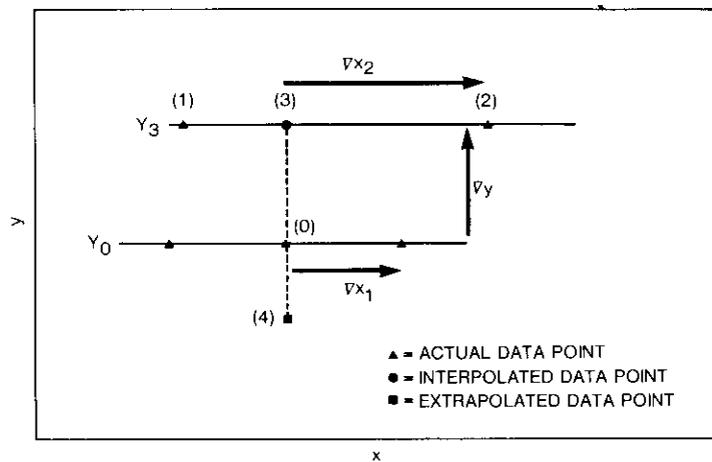
Figure 10.   *Gradient of Extrapolated Point Located Below Cuts*

repeated for the cut pair N and $N - 1$ and the cut pair $N - 1$ and N.

This section fully described the first step in drawing a contour; that is, finding the location of the points which lie on the contour. The results of this process are returned in an array that contains the x and y coordinates of the point, the gradient in polar coordinates, and a flag that describes the point type (either interpolated or extrapolated). If a duplicate point is found in this search process, it is simply not included.

Data derived from this process are used to perform the next step, in which the points are connected to generate the contour lines.

The contour connection process

Once the locations of all points that lie on the specified contour have been found, they are connected to generate contour curves, a process that can be very subjective. For example, given the set of points in Figure 11, most persons would connect them as in Figure 12. A smaller number of persons would connect the points as shown in Figure 13. Figure 14 shows additional ways these four points could be connected, although most people would be unlikely to connect the points in these ways. Based upon the available information, no one method of connecting the points is any more unrealistic than the others, although the human mind favors certain connections. Clearly, when data are sparse, uncertainty in the connectivity of points is introduced. Very complicated contours with many peaks and saddle points aggravate the problem. The human mind, unlike the computer, solves the problems with
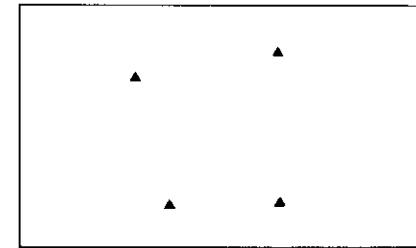
Figure 11.   *Points on a Contour*
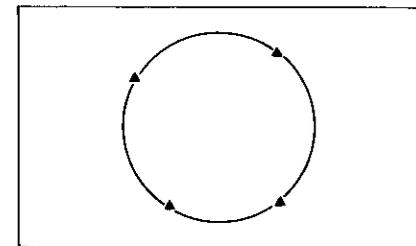
Figure 12.   *Points on a Contour Connected One Way*
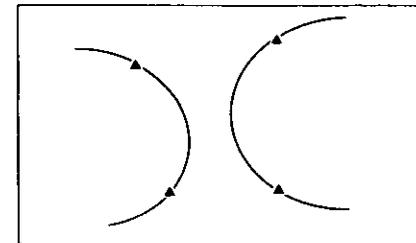
Figure 13.   *Points on a Contour Connected a Different Way*

the insight of prior knowledge. Given a sparse amount of data, programming a computer to be as competent as a human being in interpreting data for connection has been found to be difficult, if not impossible.

A fool-proof algorithm would make no mistakes under any circumstances. When efforts are made to improve the connection algorithm so that fewer and fewer mistakes are made, soon a point is reached in which the length of the program code, development time, and processing time increase disproportionately for a very small improvement in performance.
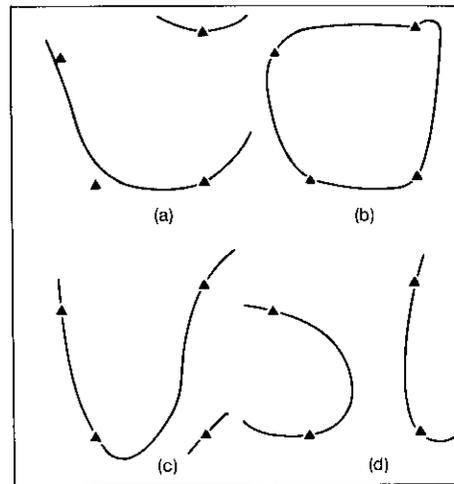
Figure 14.   *Additional Ways to Connect the Four Points*

Since connectivity criteria are subjective and since no algorithm can be totally error free, the operator must have some means for altering the connections between points generated or correcting mistakes, such as some form of connection editor. Hence, a decision was made to invest the available program development time in a powerful graphics editor as opposed to spending a large amount of time on developing heuristics for a connectivity algorithm. The graphics editor allows the user to alter the contour curves generated by the algorithm, making the connectivity algorithm compact and simple.

THE CONNECTION ALGORITHM

As previously discussed, coordinate pairs (contour points) are found that occur at the contour value. When found, the gradients at those points are calculated. The connection algorithm uses as inputs an array $G(*)$ in the following format:

$$G(m,n) = \text{contour intersection array}$$

where m = point index range $[1:m']$, where m' is the maximum number of points (memory dependent)

n = coordinate index in the range $[1:5]$ where
  1 = x coordinate of point
  2 = y coordinate
  3 = magnitude of the gradient
  4 = angle of the gradient
  5 = a flag set to (-1) for interpolated or to (-2) for extrapolated contour points.

The algorithm begins with the first contour point in the list and compares it to every other contour point as it looks for the best connection point. The criterion for "best" is based upon the distance between the pair of points being considered and the direction of their gradients. More specifically, there are three tests (explained later in detail) performed before a contour point is even considered for connection:

a. The gradients must lie on the same side of their displacement vector.
b. The distance between the two points must not exceed two units.
c. If a prior point has been connected, the difference between the direction vectors for the first and second point pairs must be less than 60°.

Once a point has passed these three tests, a mathematical value called the *criterion* is calculated and maintained for the point. These criteria are used to evaluate all the points which have passed the above three tests in order to decide which will actually be used.

The first test is shown graphically in Figure 15. In this figure, four sets of point pairs are shown, each connected by the displacement vector $\vec{D}$ and each shown with a gradient vector. The top two cases in the figure show gradient vectors that lie on the same side of $\vec{D}$ and are therefore acceptable. The bottom two pairs of points each have gradient vectors that do not lie on the same side of the corresponding distance vector $\vec{D}$ and are therefore unacceptable. These criteria, along with the others, have been chosen through a combination of mathematical reasoning and empirical testing. Because the actual spacecraft in-orbit test measurements are both sparse and noisy, the gradient calculations are approximations. These approximations to the gradient are used to assist in connection decisions and cannot be their sole determinant.

In the second test listed above, absolute distance between two points is used to quickly eliminate very distant points from further consideration. This increases the speed of the connection algorithm and assists in the reduction of connection errors. The acceptable distance between points depends on the
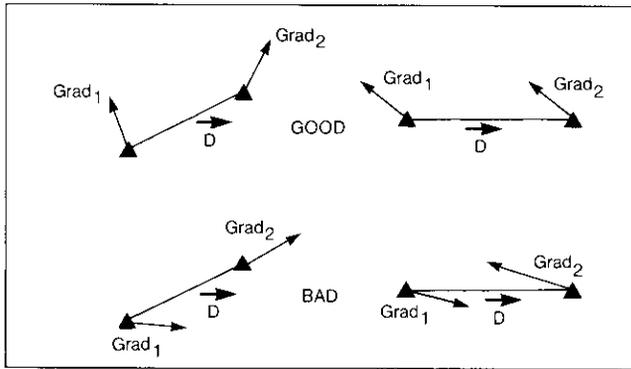
Figure 15.   *First Test for Connecting Two Points*

shape of the antenna beam since the in-orbit measurement points are discrete samples of the beam shape. Beams with many saddle points, peaks, and steep slopes require a greater density of measurement points than smoother beams of the same width. An acceptable absolute distance between points is derived from a prior knowledge of the beam shape and the density of the measurement points. It can either be actively calculated based on an examination of the in-orbit test data or it can be a fixed number which is based on expected measurement point density and beam shape. For example, consider a global beam of a communications satellite. The maximum possible range for both coordinates is approximately 17° in the spacecraft coordinate system. More typically, the range will be 12° in the x direction and 6° in the y direction. If the measurement points are taken in 0.5° increments along the x-axis, then 2° is a reasonable absolute distance for this test; thus, any points greater than this spacing are unacceptable for connection.

The third test (shown in Figure 16) checks the circulation of the connections that are being made. It requires a set of three points, two of which are already connected and a third being considered for connection. Points one and two in the figure had been previously connected and define displacement vector $\vec{D}$. The line defined by $\vec{D}$ is extended in the direction of the third point and a 60° sector is defined 30° above and below the extended displacement vector. Since point three in the figure lies within the 60° sector, it is acceptable. Alternately, the point labeled four in the figure is unacceptable because it lies outside the 60° sector. If a preceding point such as the figure's point one is unavailable, this test is not performed.

Once a point has passed all three of these tests, the connectivity criterion is calculated. The criterion is given by

$$\text{Criterion} = \cos(\Delta\,\text{Grad}) + \frac{2}{|\vec{D}|}$$

where $\Delta\,\text{Grad}$ = smaller angle between Grad_1 and Grad_2
Grad_1 = gradient at the first point
Grad_2 = gradient at the second point
$|\vec{D}|$ = magnitude of the displacement vector:

$$|\vec{D}| = \sqrt{\Delta x^2 + \Delta y^2}\ \ .$$
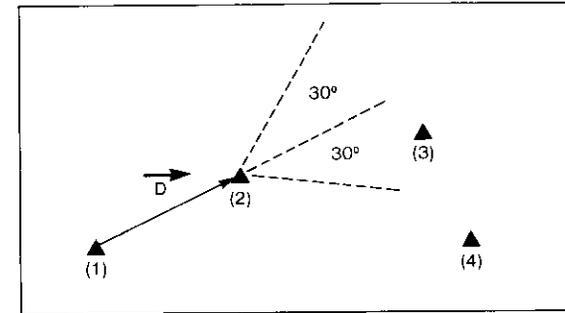
Figure 16.   *Second Test for Connecting Two Points*

Figure 17 shows the physical meaning of some of these values. The criterion formula was optimized through experimentation. Alternate criteria may work better for contour data that are different from the type for which this system was designed.
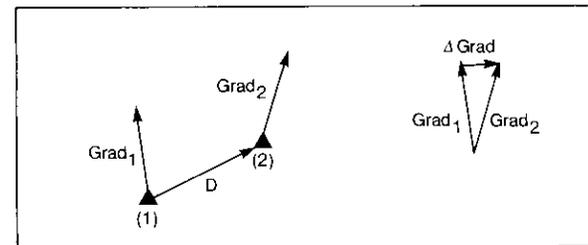

Figure 17.   *Third Test for Connecting Two Points*

Each contour point is compared with every other point to arrive at the best pair of points to be connected. Once a point is connected, it will no longer be considered for future connection. If there are four points, numbered one through four, to be connected, *point one* is compared with every other point. If all the points pass the three tests, then the criteria between *point one* and the other three are calculated and *point one* will be connected to the point with the maximum criterion. For example, if *point three* has the highest criterion, then *point one* will be connected to *point three* and *point three* will be compared in the same fashion with *point two* and *point four* for the best fit. This process continues until all the points have been considered.

One step in this process remains to be explained. Earlier a distinction was made between extrapolated and interpolated points. Extrapolated points are used to steer the connectivity algorithm near the edges of the contour. The fact that a point may be extrapolated is not considered until the actual connection is about to be made. Thus, as a result of the previous tests, it is possible for an extrapolated point to be the best match. In this case, the point is simply discarded and the current curve being generated is terminated. The next available free point is picked up and the search begins as it does during the start of the connection routine. That is, extrapolated points are not displayed but are merely used to determine the proper circulation.

CONTOUR CLOSURE

A special case, contour closure, occurs if the point giving the best fit is the first point of the contour. Support of contour closure requires that the first point in the contour be treated differently from the remaining points, yet this first point must be considered along with each other point when determining the best fit. Assume there are five points numbered one through five and let the first two connections join *point one* to *point three* and *point three* to *point five*. The best match for *point five* must consider *point two*, *point three*, and *point one*. *Point one* must be considered to allow for contour closure.

Conclusion

This paper has presented an algorithm for generating antenna contour patterns from in-orbit test antenna gain measurements. Because of time and operational limitations, the in-orbit test data are sparse and do not fall on a uniform grid. The algorithm developed for this purpose generates acceptable contours all the time and correct contours most of the time. A human operator combining this algorithm with a graphics editor for post-processing after contour generation can create correct and unambiguous contours, such as those shown in Figure 18, very quickly and with a minimum of effort. In
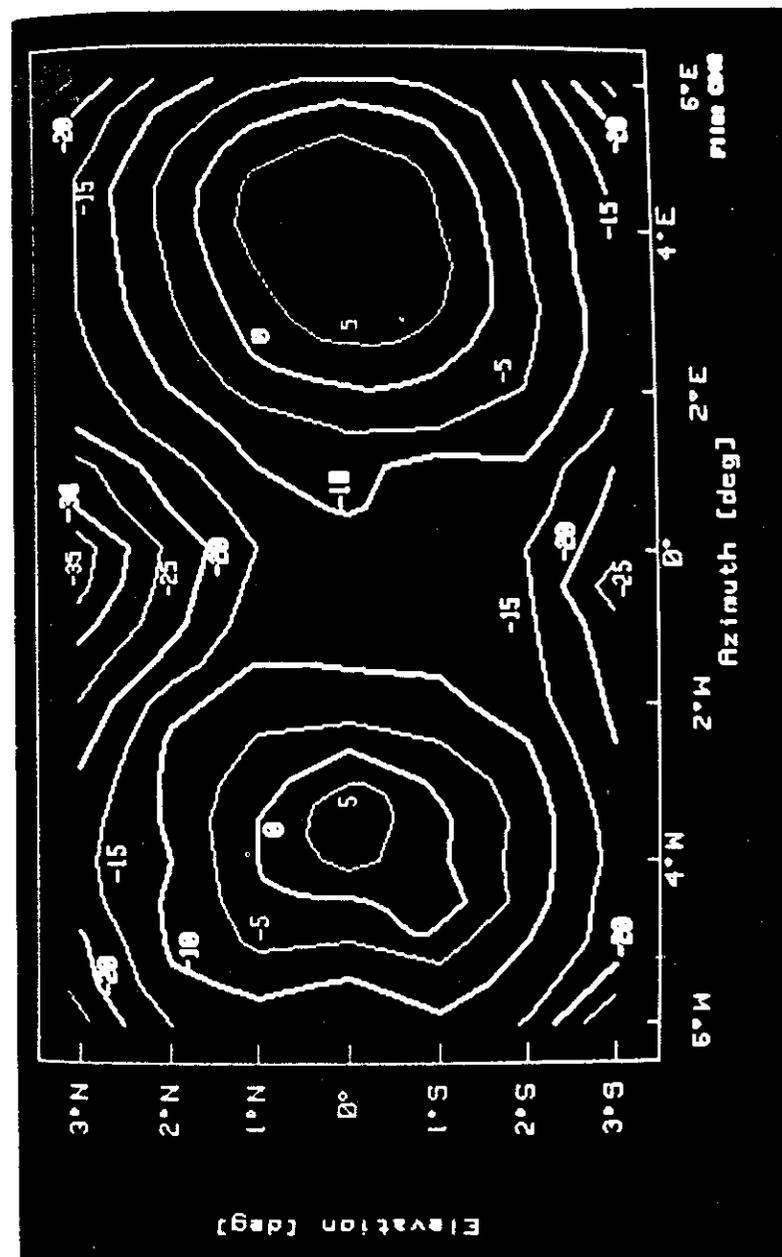
Figure 18.   *Contour of Antenna Pattern*
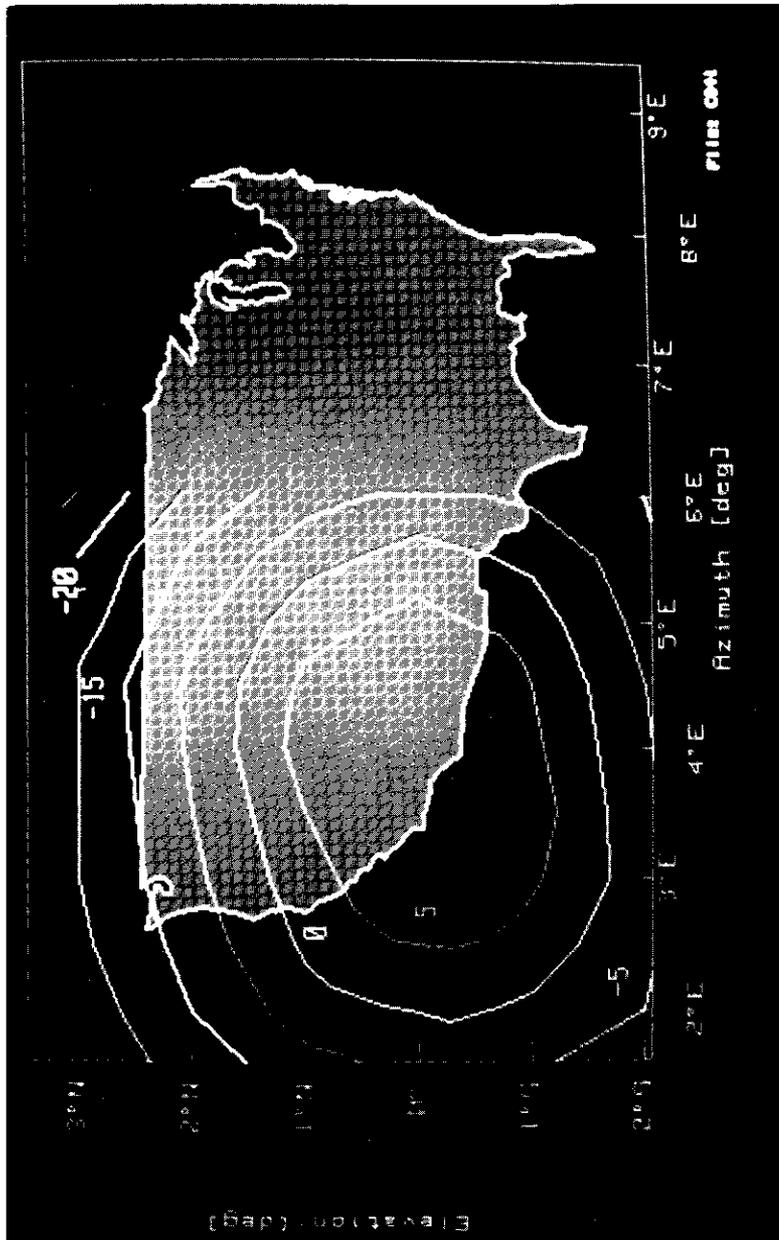
Figure 19.   *Mapped Spacecraft Antenna Beam*

addition, this contour generation system allows further processing of the antenna contour. For example, the spacecraft beam footprint can be displayed on a map of the earth showing coverage areas, as shown in Figure 19 (with simulated data), where the shape of United States is distorted due to the orbital location of the spacecraft. In these cases, different spacecraft subsatellite points and antenna pointings can be input to allow different orbital conditions to be analyzed using the same set of in-orbit test data.

Acknowledgments

References

[1] I. Dostis et al., "In-Orbit Testing of Communications Satellites," *COMSAT Technical Review*, Vol. 7, No. 1, Spring 1977, pp. 197–226.

[2] V. Riginos, C. Mahle, and I. Atohoun, "ATE Checks Health of Orbiting Spacecraft," *Microwaves*, Vol. 19, No. 8, August 1980, pp. 46–50.

[3] S. L. Simons, Jr., "Make Fast and Simple Contour Plots on a Microcomputer," *Byte*, Vol. 8, No. 11, November 1983, pp. 487–492.

[4] I. K. Crain, "Computer Interpolation and Contouring of Two-Dimensional Data: A Review," *Geoexploration*, Vol. 8, No. 2, 1970, pp. 71–86.

[5] H. Schultheis and R. Schultheis, "Algorithm 35: An Algorithm for Non-Smoothing Contour Representations of Two-Dimensional Arrays," *Computing*, Vol. 19, 1978, pp. 381–387.

[6] A. Ralston, *A First Course in Numerical Analysis*, New York: McGraw Hill, 1965, pp. 46–57.

[7] I. K. Crain and B. K. Bhattacharyya, "The Treatment of Two-Dimensional Non-Equispaced Data With a Digital Computer," *Geoexploration*, Vol. 5, 1967, pp. 173–194.

[8] J. H. Ahlberg, E. N. Nilson, and J. L. Walsh, *The Theory of Splines and Their Applications*, New York: Academic Press, 1967.

[9] D. F. Merian, and N. C. Cocke (editors), "Computer Applications in the Earth Sciences," *Colloquium on Trend Analysis—Kansas Geological Survey of Computer Contributions*, Vol. 12, 1967, p. 62.

[10] P. B. McIntyre, D. D. Pollard, and R. Smith, "Computer Programs for Automatic Contouring," *Kansas Geological Survey of Computer Contributions*, Vol. 28, 1968, p. 76.

[11] M. O. Dayhoff, "A Contour Map Program for X-Ray Crystallography," *Communications of the Association for Computing Machinery*, Vol. 7, 1963, pp. 620–622.

Kenneth D. Fullett received a B.S.E.E. and M.S.E.E. from the University of Illinois, Urbana-Champaign, in 1979 and 1981, respectively. He joined COMSAT Laboratories in 1981. As a Member of the Technical Staff in the Transponders Department of the Microwave Technology Division, he has participated in the in-orbit testing of various communications satellites and contributed to the development of several in-orbit automatic test equipment systems. In addition, he has participated in the development of a microwave CAD *system at COMPACT Software, and in the research and development of a high-efficiency L-band amplifier.*

Vasilis E. Riginos received B.E., M. Eng., and Ph.D. degrees in electrophysics from the Stevens Institute of Technology, Hoboken, New Jersey, in 1964, 1970, and 1973, respectively. He was NDEA Fellow at the Stevens Institute of Technology, where he worked on the transport properties of GaAs. He also helped develop the Physical Electronics Laboratory and taught physical electronics for several years. In 1974, he joined the Transponders Department of the Microwave Technology Division at COMSAT Laboratories as a Member of the Technical Staff, working on Gunn-effect device modeling, solid-state amplifier design, filtering techniques, and the design of automated microwave measurement systems. He has participated in the in-orbit testing of the INTELSAT IV, IV-A, V, and VI, COMSTAR, and SBS satellites, and was the primary engineer responsible for implementation of the INTELSAT and SBS automated in-orbit test system. Since 1981, he has been Manager of the Transponders Department responsible for directing research and development on state-of-the-art communications transponders and microwave circuits. He directed the development and implementation of the GTE GSTAR automatic in-orbit measuring system, as well as the INTELSAT Maritime Communications subsystem in-orbit test system software. Dr. Riginos is a member of Sigma Xi, AAAS, IEEE, the American Physical Society, and the New York Academy of Sciences.

A modeling system for simulation of GaAs FET performance

T. SMITH

Abstract

A new computer system for predicting field effect transistor (FET) performance is described which can be used extensively for design optimization, troubleshooting the fabrication process, and tailoring performance by varying the fabrication process parameters. The computer programs are based on the physics governing the operation of FETs and input parameters are chosen which relate closely to the fabrication process. The intrinsic physics of FET devices is covered as it relates to the modeling and computation strategies. The usefulness of the system is illustrated by modeling a GaAs power FET.

Introduction

The primary objective of the modeling described here is to guide the design and fabrication of FETs at COMSAT Laboratories. The computer programs are based on the physics of FET operation and organized so that performance predictions can be related to both process and device design parameters.

Since there is no one FET which is useful for all applications, the device designer must attempt to optimize the FET for the intended specific application. In general, this will involve both the design of the maskset and specification of material properties such as the carrier profile. Both tasks must take into account limitations due to existing fabrication processes and attainable

properties of materials, including metal and dielectric layers. The model described here is a valuable tool for the device designer.

For a given maskset, considerable performance variation can be achieved by alteration of the process-dependent parameters. A modeling system such as the one described here makes it possible to use the same maskset but still tailor the process to meet the requirements of several applications. In addition, this modeling system provides the means to compensate for material properties which deviate from the nominal by adjusting process parameters.

The use of the model in troubleshooting is also very important. Sometimes the measured performance of an FET is not what had been expected. When this happens, measurements to determine the values of relevant parameters usually show several deviations from specifications. By varying these parameters in the model, determinations as to the cause of the problem can be made and corrective action taken.

Background

Several FET models are currently used at COMSAT Laboratories and tend to complement one another, since each has different strengths. This paper describes an addition to the set of models already in use.

The Naval Research Laboratory has published a model [1] mainly intended for small signal FETs, but which can accommodate design parameters appropriate for power FETs. This model does not attempt to compute a load impedance for maximum power, so the gain and power-added efficiency associated with maximum power are not calculated. However, some of the output of the model is helpful in estimating power FET performance parameters.

Ezzeddine [2] and Geller [3] have both developed useful FET models. B. Geller's model computes a parameter called *maximum efficient gain* and an associated load impedance based on work by Kotzebue [4]. Maximum efficient gain is basically a small signal parameter defined as the gain associated with maximum power added when the input signal amplitude is held fixed. This is easy to compute from the two-port parameters describing the FET and is a reasonable estimate of the actual gain for maximum power FET operation. The associated load impedance does not necessarily result in either the maximum possible output power or the maximum power because of constraints imposed by the physical limitations of current and voltage amplitude in the FET. The power FET model discussed in this paper can find the power limit for any load and thus can determine the load that allows the maximum power.

To understand modeling of FETs and their fabrication, it is necessary to review both the fabrication process and the basic physics of FET operation.

The fabrication process

In this section the fabrication process to demonstrate the choices of input parameters for the modeling programs is described.

First, a materials process produces a semiconductor wafer with appropriate properties which has a variation of carrier concentration vs depth below the top surface known as the *carrier profile*. Usually the carrier profile can be measured before any further fabrication processes are started. The profile data can be entered into one of the modeling system programs. Subsequent processes affect the top surface in two dimensional patterns defined by a set of masks called the *lithographic maskset*. Important geometric details provide further input to the modeling programs.

Process effects in the third (depth) dimension are often controllable through process parameters. For example, the gate recess etch is one of the most critical steps in the fabrication process. Unfortunately, it is not possible to directly measure the etch depth, and the etch rate is not sufficiently stable to rely on timing alone. However, it is possible to measure a current which is determined by the etch depth. The computer program called FETRECESS uses the carrier profile and other known or intended parameters to compute both this current and various performance parameters as a function of recess depth. This information can then be fed back to the fabrication process.

Geometric details of the two dimensional surface patterns of FETs are defined not only by the maskset, but also by the fabrication process itself. The alignment of each successive pattern with the preceding ones affects some of these details. So does pattern transfer by the photoresist method. Typically, a mask is a patterned, flat glass plate used to transfer the pattern to a removable layer of photoresist on the coated wafer. Subsequent process steps transfer the resist pattern to the required pattern in the actual device. In this pattern transfer process, the geometric details in the actual device may vary slightly from the corresponding details on the mask. This can affect results significantly when the possible amount of variation is comparable to the size of a critical dimension of the pattern. In actual practice, deliberate variation of the process can sometimes be used advantageously to adjust over a limited range certain dimensions such as FET gate length. Modeling can be used to predict the effects of intentional and unintentional variations in important geometric parameters.

FET device physics

The modeling system described here is based on the underlying physics of FET operation, with performance calculated through use of an equivalent

AC network representing the FET. The physics needed to define the network and to determine the values of the elements is discussed in this section.

The heart of the equivalent network is an AC current generator with amplitude proportional to that of an AC voltage across a capacitor, C_{gs}. The constant of proportionality is the transconductance, g_m. The AC power generated in this way is derived from the DC power supplies providing the bias voltages. Figure 1 shows an equivalent AC network adequate for discussion of the device physics needed to assign values to the elements of the network and to g_m. It should be noted that real FETs can have more complicated equivalent networks, as will be the case for the power FET example discussed later.
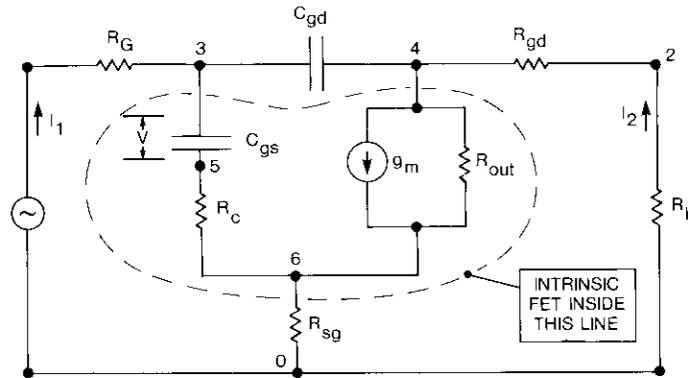

Figure 1. *Simplified* AC *Equivalent Network for an FET*

Intrinsic elements

Physically, the intrinsic circuit elements of the FET include the interface from the gate-metal to the semiconductor, the movable part of the depletion layer boundary, and the conducting layer immediately below it.

Figure 2 shows a cross section through an FET. The conducting layer will have a certain concentration of mobile charges or carriers which tend to stay in the layer because there is an equal concentration of opposite fixed charge due to the ionized impurity atoms that contribute the mobile carriers. A mobile charge which does wander into a region of lower fixed charge density creates a charge imbalance by setting up an electric field which attracts mobile charges so as to maintain neutrality.

The source and drain contacts are ohmic, meaning the potential difference between the metal and the semiconductor obeys Ohm's law (*i.e.*, the potential
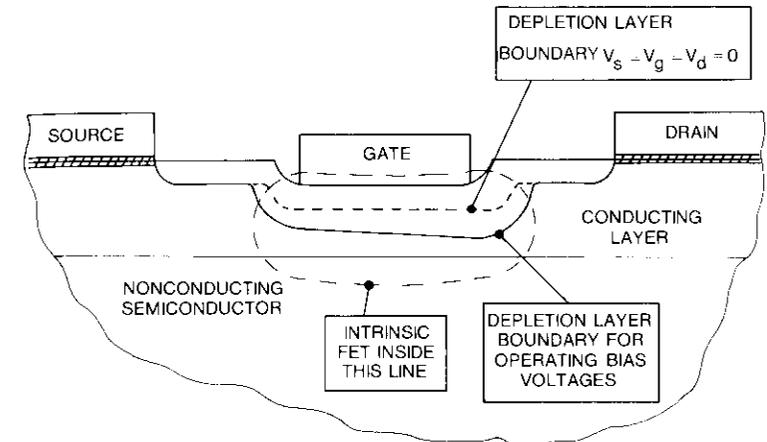
Figure 2. *Cross Section Through an FET*

difference is proportional to the current). When the source is grounded and a voltage is applied to the drain, carriers will flow through the semiconductor, passing under the gate.

The gate contact is not ohmic. Instead, a potential difference between the gate contact and the conducting region changes the depth of a depleted region where the electric field due to the metal-semiconductor junction drives out the mobile charges. This field vanishes at the depletion layer boundary because it has been balanced at this point by the field from the fixed charges of ionized dopant atoms in the depletion layer. As shown in Figure 2, a depletion layer is also associated with the bare surface and occurs in GaAs because of the field set up by ionized surface states.

The depletion layer boundaries can be regarded as walls confining the flow of carriers to the region between them and the lower boundary of the conducting layer. In this region the carriers behave like a nearly incompressible fluid. For low electric fields, carrier velocity is proportional to the field and hence the source drain-current will obey Ohm's law. However, at some higher field the magnitude of the velocity reaches a limiting value, v_s, which cannot be exceeded. Under the assumption of incompressibility, the source-drain current also reaches a limiting value determined by v_s and the cross-sectional area it flows through.

Actually, the depletion layer boundary is not abrupt and the carrier fluid is somewhat compressible. As a result, the source-drain current will increase slightly with increasing source-drain voltage beyond the onset of velocity saturation. The resistor labeled R_{out} in Figure 1 represents this finite

conductance. Furthermore, the negative differential mobility of the GaAs velocity vs field characteristic probably leads to the experimental finding that the AC value for R_{out} is lower than the DC value. For the above reasons, there is no adequate theory for computing R_{out} or any reactive elements which may be associated with it. These values are based on the measurements of similar FETS and reasonable assumptions about how R_{out} varies with device geometry.

The basic current generation mechanism in the FET is the movement of the depletion layer boundary under the gate in response to a change in the voltage applied to the gate. This alters the minimum height of the conducting layer and, therefore, the cross-sectional area through which the carriers flow at saturated velocity. Since the carrier velocity is constant, the current must change in proportion to the cross-sectional area. These factors cause the familiar drain current vs voltage characteristics shown in Figure 3.

Gain is proportional to the square of g_m/C_{gs}, where g_m is the intrinsic transconductance and C_{gs} is the gate-to-source capacitance. The physics related to these parameters deserves special attention. It has been shown how an AC voltage on the gate generates an AC component in the source-drain current, I_{ds}. The definition of g_m is

$$g_m = \frac{dI_{ds}}{dV} \quad . \tag{1}$$

I_{ds} is determined by the maximum depletion layer thickness, computed by

$$I_{ds} = e\,N\,W_g\,v_s\,(T - t) \tag{2}$$

where
e = the electronic charge
N = the number of ionized atoms per cm^{-3} in the depletion layer
W_g = the gate width
v_s = the carrier saturation velocity
T = the total active layer thickness
t = the depletion layer thickness (at the point of maximum thickness).

Therefore

$$g_m = -e\,N\,W_g\,v_s\,\frac{dt}{dV} \quad . \tag{3}$$
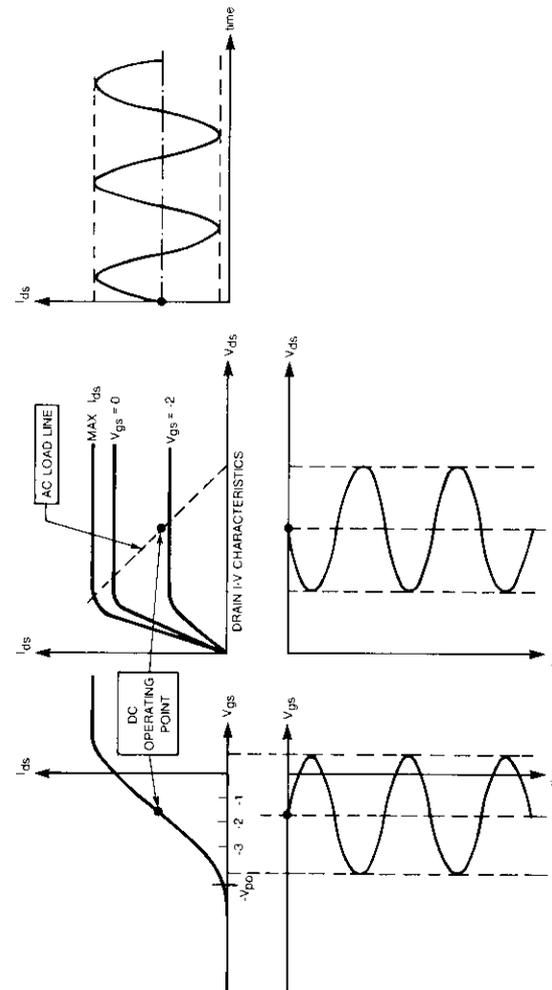
Figure 3. Current and Voltage Characteristics of an FET

By definition, the gate-source capacitance, C_{gs}, is given by

$$C_{gs} = \frac{dQ_f}{dV} \tag{4}$$

where Q_f is the total fixed charge in the depletion layer directly under the gate metal. Ignoring fringing capacity at the edges of the gate metal,

$$Q_f = e \, N \, l_g \, W_g \, t \tag{5}$$

where t = uniform depletion layer thickness
         l_g = gate length.

Differentiating (5) gives

$$C_{gs} = e \, N \, l_g \, W_g \frac{dt}{dV} \quad . \tag{6}$$

It follows from equations (3) and (6) that

$$\frac{g_m}{C_{gs}} = -\frac{v_s}{l_g} \tag{7}$$

is independent of the N vs depth profile. Actually, C_{gs} will always be somewhat larger than the value calculated from equation (6) because t varies with position under the gate and the movable depletion layer boundary bulges toward the source and drain. In the work discussed below, C_{gs} is computed as the capacity due to dQ/dV under the gate metal, plus the capacity due to dQ/dV in the bulge toward the source. These details, which affect gain variation with carrier profile, are taken into account by appropriate computer programs in the modeling system.

Other key element values of the AC equivalent network are determined as follows. The dQ/dV in the bulge toward the drain determines the value of the capacitor, C_{gd}. Also associated with the current generation mechanism is the resistor, R_{in}. To change the location of the depletion layer boundary, the volume filled with carriers must change. The resulting motion of the carriers dissipates some input power, which, divided by the square of the AC current through C_{gs}, gives R_{in}, or the charging resistance.

Parasitic elements

The remaining elements of the equivalent network are termed parasitic.

The source-to-gate resistance, R_{sg}, has important effects on performance and may be estimated by the following method. For GaAs, the bare surface depletion layer thickness is the same, or nearly the same, as the depletion layer under the gate when source, drain, and gate are grounded. This allows calculation of the resistance from the source-metal edge to the gate-metal edge in this zero bias condition. This calculation, plus the contact resistance, is a reasonable estimate of R_{sg}. On the other hand, the true definition of R_{sg} is the resistance between the source metal and the point where velocity saturation begins under the gate. This can be expected to vary somewhat with V_{ds} and V_{dg}. Estimates of the magnitude of this variation show that it is usually not important to try to calculate the precise bias-dependent value. But, since R_{sg} is expected to be a bit larger than the above estimate, an approximate correction can be made by adding the resistance under the gate for the zero bias condition to the above estimate.

The values of other parasitic elements can be calculated in a straightforward manner from the geometry of the FET, the semiconductor and gate-metal properties, and the ohmic contact resistance. For example, the capacitor in shunt with the output port, C_{ds}, has capacitance computed from geometric dimensions and the dielectric constant of GaAs. As already noted, a given FET could have a more complicated equivalent network than the one shown in Figure 1, so the modeling system by design allows flexibility in equivalent network configuration.

Modeling system description

The modeling system consists of a group of SPEAKEASY* computer programs, some of which are designed to be called up by other programs [5]. This modular organization makes it fairly easy to refine the model by changing the appropriate programs. Since different FET designs could have different AC equivalent networks and different sets of variable parameters, the modeling system is designed so that a main program can allow the user to name an FET-definition subprogram. This named subprogram may, in turn, call on any subprogram it needs. Figure 4 is the computation flow chart for MODELSYS, an example of a main program. Its operation is discussed in this subsection.

Upon entry into MODELSYS, the user is asked to specify the name of an FET-definition program, whether hard copy is wanted, and whether to initialize
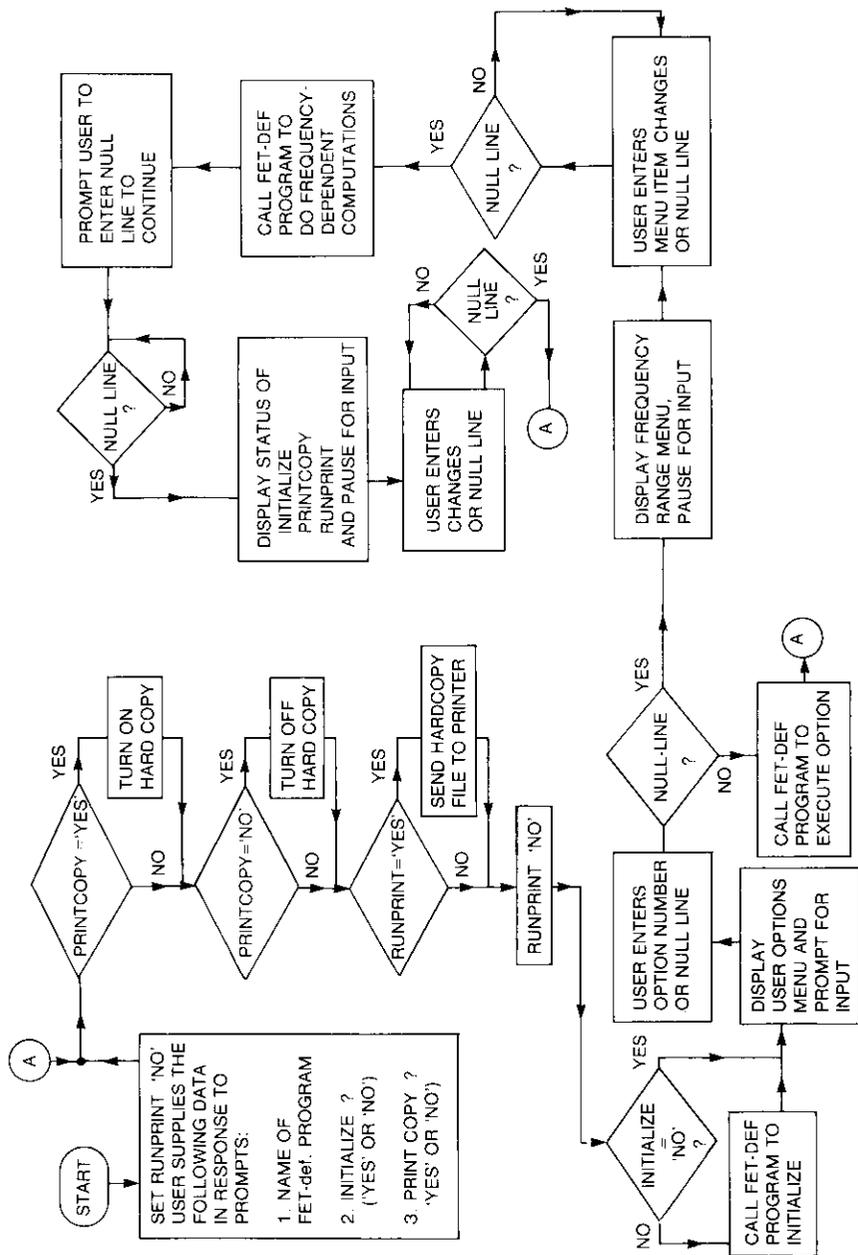
Figure 4. *Flow Diagram of the MODELSYS Program*

parameters. This program specifies a set of FETs that can be generated by variation of the values shown in lists of parameters. It also defines a particular configuration for the equivalent network.

The FET-definition program has several entry points. The first call on it for a given SPEAKEASY session must be via the initialization entry point. Here default values for all parameters are set, a menu of user options is established, and other menus are created. These menus can be arranged in a hierarchy such that entries in a higher menu cause changes in parameter values shown in subordinate ones. For example, material and geometric parameters would be in a higher level menu than equivalent circuit element values. Other entries in the User Option Menu allow specification of what values are to be output. After initialization, control is returned to node A of Figure 4.

At node A, the User Option Menu is displayed and the user keys in an option number. Unless a zero is entered to exit from the program, the option number causes the FET-definition program to be entered at a point corresponding to the option number. Typically the result will be a menu containing parameter values. This would be followed by instructions for changing any of the listed values. After any changes, control is returned to MODELSYS at node A. If a null line is keyed in instead of an option number, MODELSYS calls the FET-definition program for frequency dependant computations and output. Then control is again returned to MODELSYS, which prompts the user to enter a null line to continue (thus giving the user an opportunity to look at some of the results before proceeding). On receiving the null line, the user is given the opportunity to request a change of hardcopy status, to send the hardcopy file to the printer, and to request reinitialization. Control then returns to node A.

Frequency dependent computations

The FET-definition program uses the current parameter values to calculate AC admittances connecting the nodes of its AC equivalent network. The following conventions are followed. Node 0 is always the grounded common node, node 1 the input node, node 2 the output node, and the nodes of the intrinsic FET numbered 3 through 6, as shown in Figure 1. If the equivalent network has additional nodes, these can be assigned arbitrary numbers greater than 6. The elements connected to the extra nodes can only be linear ones such as resistors, capacitors, and inductors. The current generator parameters, g_m and τ, are also calculated.

Next, the $n \times n$ coefficient matrix for the set of simultaneous node current equations is computed. Here, n is the number of nodes minus one. Since the

power-limiting physics requires knowledge of the voltages only at nodes 1 through 5, the $n \times n$ matrix for n greater than six can be mathematically reduced to a 5×5 matrix. Use of the smallest possible matrix results in savings of both computer time and storage space for computations requiring this matrix.

Small signal FETs

For small signal FETs, the only node voltages of any interest are those at nodes 1 and 2. This allows the $n \times n$ matrix to be reduced to a 2×2 matrix known as the two-port Z-matrix. From this, the microwave scattering parameters, stability parameter, maximum associated gain (if it exists for the current parameter values), and associated input and output impedances can be calculated. Fukui's formula [6] can be used to estimate a minimum noise figure from the values of some of the elements of the AC equivalent network, the frequency, and g_m.

Power FETs

The computations required for power FETs are more extensive. In this case the saturated (that is, maximum possible) output power and the output power at which nonlinear effects become important are of interest. These are determined by one or the other of two physical limitations: maximum possible current amplitude and avalanche breakdown at the drain edge of the gate. These limitations can be related to certain node voltages as described below.

MAXIMUM CURRENT

The maximum current is limited by the thickness of the surface depletion layer not under gate metal. This is assumed to be the same as the depletion layer thickness beneath the gate under zero bias conditions ($V_{gs} = 0$ and $V_{ds} = 0$). This current limit is greater than I_{dss}, the saturated current at $V_{gs} = 0$, because the voltage, V, across the thickest part of the depletion layer for $V_{gs} = 0$ is

$$V = I_{dss} R_{sg} + L_s E_s \qquad (8)$$

where L_s = the distance carriers flow at the saturation velocity (approximately l_g)

E_s = the saturation electric field (taken to be 4000 V/cm).

The maximum current will be attained when the difference between V and

V_{gs} reaches zero. The minimum possible current is zero since negative current is impossible for positive drain voltage. The gate voltage at which the current is zero is called the *pinch-off voltage*, V_{po}. This is nearly equal to the punch-through voltage at which the depletion layer punched all the way through the conducting layer. This voltage is easily computed. As illustrated in Figure 1, the current limits are related to the voltage difference between nodes 3 and 5.

AVALANCHE BREAKDOWN

Avalanche breakdown occurs when the electric field intensity in the semiconductor exceeds a critical value which depends on the local density of fixed charges. Usually the field intensity is highest at the drain edge of the gate metal and is determined by the magnitude of the potential drop across the depletion layer at that point. In terms of node voltages, avalanche breakdown will start when $(V_4 - V_3)$, including DC, reaches a certain value. Further increase in this voltage difference causes a very rapid increase in current flowing between these two nodes, thereby preventing useful increase in the output amplitude.

To avoid clipping, V_4 also has a minimum value such that $(V_4 - V_6)$, including DC, is not less than the saturation electric field intensity, E_s, times the distance (taken to be the metal gate length) over which carriers flow at the saturation velocity. Often V_6 is determined only by the current through a linear element connected between node 6 and node 0, e.g., R_{sg}. In this case a minimum value for $(V_4 - V_3)$ can be calculated without actually solving the node equations for V_6.

MAXIMUM SATURATED POWER

For power FETs a large signal linear approximation is used with current clipping caused by the two physical amplitude limitations discussed under maximum current and avalanche breakdown. A large signal value is used for the transconductance; namely, the source-drain current for $(V_3 - V_5) = 0$ divided by the pinch-off voltage, V_{po}. Other parameters of the intrinsic FET are similarly based on a large signal approximation. Because of the use of the large signal linear approximation, the results of the model apply most accurately for signal amplitude at the onset of clipping of the output current wave form. Above this amplitude, current clipping causes the output wave form to approach a square wave as the input amplitude is made very large. Based on Fourier analysis of a square wave, the saturated power should be higher than that based on the model by a factor of at most $(4/\pi)^2$, or approximately 1.62.

RELEVANT NODE VOLTAGES

The relevant node voltages are computed by the following method. Their DC components are easily calculated because there is only one DC current path through the network. The 5×5 matrix discussed above is used to find the AC node voltages at nodes 1 through 5. The nth row of this matrix represents the five coefficients multiplying the five node voltages to give the sum of the currents into the nth node. At nodes 1 and 2 the sums are equal to I_1, and I_2 (the input and output currents), respectively. At all other nodes the current sum must be zero.

The node current sums and the 5×5 coefficient matrix represent 5 equations for 5 unknown AC node voltages. Since the linear approximation is being used, input current amplitude, I_1, may be initially set equal to 1. For any other input current amplitude, the solution is obtained by multiplying the solution for $I_1 = 1$ by the numerical value representing the actual input current.

The output current, I_2, has amplitude and phase represented by a complex number. In principle it could have any value, but consideration is limited to cases in which power flow is inward at node 1 and outward at node 2. Any other case is either unstable or requires a generator instead of a load at the output port.

It is convenient to use the formalism of Cote and Oaks [7] to represent the output currents in terms of a complex number, $u = x + jy$. This is normalized in such a way that input power depends only on x, and the region of positive output power is inside a circle of radius one, centered at the origin. The normalizing factors are derived from the 2×2 Z-matrix found by solving the node equations for the two cases: $I_1 = 1$, $I_2 = 0$; and $I_1 = 0$, $I_2 = 1$ using the subroutine SIMEQ of the SPEAKEASY language. Since a u-value represents an I_2 value for $I_1 = 1$, the Z-matrix can be used to find the corresponding values of V_1 and V_2. The load impendance is found from V_2 and I_2, so a given u-value also corresponds to a particular load impedance. Hence, the computation procedure discussed below constructs a map of maximum power vs load impedance.

COMPUTATION STRATEGY

The starting point of the computation is to construct an array of 331 u-values arranged in a hexagonal pattern with 0.1 spacing inside the unit-radius circle. Any of these found to result in negative input power are eliminated from the array. The remaining values are converted to output currents and the node equations are solved for each of these I_2 values. This requires only

one call to the SIMEQ subroutine for the entire array of I_2 values. To find the maximum linear power for each u-value, it is necessary to find the multiplying factors which give the maximum amplitude allowed by each of the two physical constraints of maximum and minimum current and of avalanche breakdown.

The multiplying factors are found by the following method. The two-dimensional node voltage array returned by SIMEQ has columns corresponding to up to as many as 331 I_2 values and five rows corresponding to node voltages of the lumped element circuit. The complex values of the node voltages represent sinusoidal voltages with the amplitude given by the absolute value. F_1 is the factor related to current clipping due to the voltage amplitude $(V_3 - V_5)$, given by

$$F_1 = \frac{0.5V_{po}}{|V_5 - V_3|} \quad . \tag{9}$$

The factor, F_2, makes the amplitude of $(V_4 - V_3)$ equal to the maximum allowed by avalanche breakdown, $\max V$. The DC operating bias values are chosen to allow F_2 to be as large as possible. F_2 also depends on the RF load line, which is different for each I_2 value. To maximize efficiency, the DC operating point is chosen such that the minimum value of the I_{ds} waveform is at $I_{ds} = 0$, and the minimum value of $V_4 = V_o$. The value of V_o is equal to the DC voltage at node 4 at the onset of saturation for I_{ds} equal to the maximum of the current waveform, including the DC component. For sufficiently small channel heights, this choice also maximizes the power at the onset of nonlinearity due to clipping. In the model used,

$$V_o = \max(I_{ds})R_{sg} + \Delta V \tag{10}$$

where ΔV is the saturation field (taken to be 4,000 V/cm) times the effective gate length. Calculation of R_{sg} is discussed above in the section on FET device physics. The value of $\max(I_{ds})$ is either calculated for the same FET geometry without a gate, or estimated from measurement of saturated I_{ds} of an FET with an open-circuited gate. Note that the chosen DC bias values are not fixed, but depend on both the u-value and the amplitude limitations. The F_2 factor resulting from the above considerations is given by

$$F_2 = \frac{(\max V - DV - V_{po})}{(2|I_2|R_{sg} + |V_4| - |V_3| + |V_4 - V_3|)} \quad . \tag{11}$$

Each column of the node voltage matrix and the corresponding values of

I_1 and I_2 are multiplied by the smaller of F_1 and F_2. For every value of u, this gives the maximum amplitudes for whichever physical limit occurs first. The resulting new values I_1 and I_2, the node voltage matrix, and the DC parameters V_{po} and max(I_{ds}) are used to do the following calculations. For each value of u, input and output power, load impedance, input impedance, DC power, gain, and power-added efficiency can be computed. The search for an optimum value of whatever performance parameter wished is now straightforward. For example, the maximum output power and the associated values can be found as well as the other performance parameters. To improve precision, a second iteration with a 0.01 u-value spacing can be done around the optimum value found in the first iteration.

Example of power FET performance vs channel height

The power FET design shown in Figure 5 and represented by the 9-node equivalent circuit shown in Figure 6 has been chosen for purposes of illustration. Certain element values in the circuit and the device physics limitations depend on channel height and carrier profile. For simplicity, a
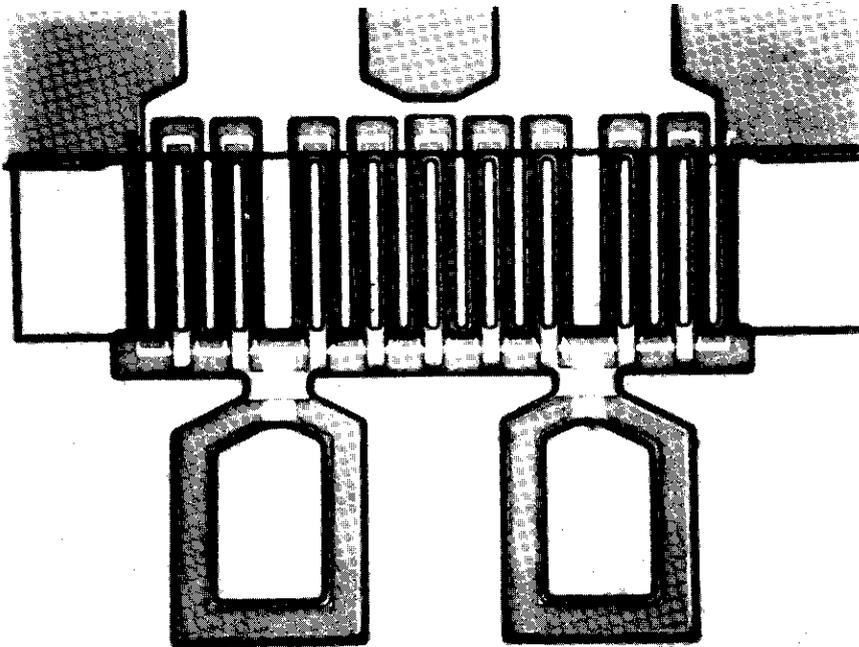

Figure 5. *Photograph of COMSAT-Designed Power FET*
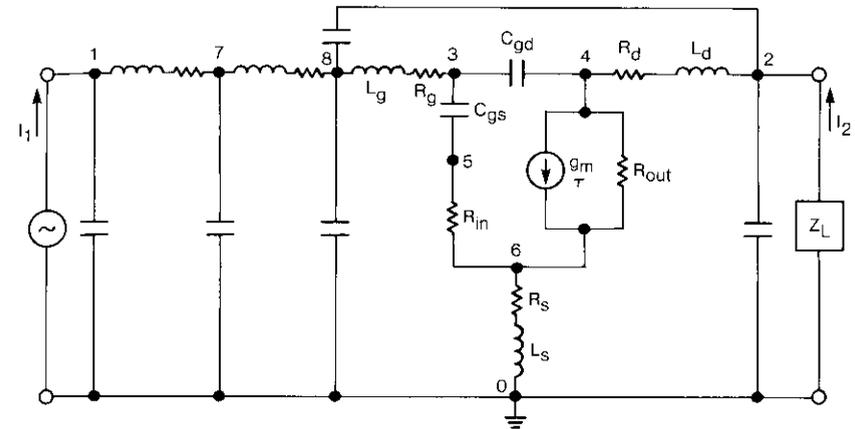
Figure 6. AC *Equivalent Network for One 1.5-mm Width Cell of COMSAT Power FET*

flat carrier profile under the gate is assumed, with a carrier concentration, N, of 2×10^{17}, to a depth equal to the channel height, h, and $N = 0$ at greater depth. The channel height parameter is easily adjusted in the fabrication process by varying the amount of GaAs etched before deposition of the gate metal. The model was used to predict relevant performance parameters at 4.2 GHz as a function of channel height.

The FET-definition program used here is somewhat approximate. R_{out} and various parasitic resistances are included on the parameter list rather than being computed, but the values used were derived from another program, FETRECESS. They were held constant while the channel height was varied. Although this is reasonable for R_{sg} and R_{gd}, R_{out} should vary with channel height. Possible effects of holding R_{out} constant will be discussed in the next section. For this example, the parameter list used by the FET-definition subprogram is shown in Table 1.

Discussion of example results

Figure 7 shows the resulting curve of maximum power output vs channel height. Two experimental points for 1-dB gain compression are also shown for comparison. They are for longer gate lengths than the 0.5-μm value in Table 1, but are valid because gate-length does not affect the maximum power output.

TABLE 1. PARAMETER LIST OF THE FET-DEFINITION SUBPROGRAM

| VARIABLE NAMES	VALUES	UNITS
N	2×10^{17}	cm^{-3}
MOB	3,500	cm^2/Vs
W	0.15	cm
L	5×10^{-5}	cm
h	0.8×10^{-4}	cm
	to 0.26×10^{-4}	
Max$V$	13.2	V
R_{sg}	0.253	Ω
R_{gd}	0.309	Ω
R_s	0.25	Ω
R_{out}	100	Ω
Freq	4.2	GHz

Figure 7. *Maximum Power Output* vs *Channel Height as Computed by the Model*

The shaded area in Figure 7 represents varying degrees of power saturation (*i.e.*, gain compression). The curve at the lower limit of this area represents the maximum linear power. It was computed as discussed in the subsection about power FETs. The upper curve is the square wave limit of 1.62 times the lower.

The computer program also checks to see if higher power could be obtained by choosing the DC operating point such that the I_{ds} waveform maximum is max(I_{ds}), and the minimum is a value determined by the avalanche breakdown limit (not necessarily zero). For the parameters shown in Table 1, this situation occurs when channel heights are greater than 0.24 μm. The plotted power is the higher of the two possibilities.

Associated values for gain and power-added efficiency are given in Figures 8 and 9. No experimental points are shown because the gate length affects these parameters. When the 1.2-micrometer gate length of one of the actual devices is put into the model, both the predicted and actual gain (11.4 dB before any gain compression) agree to within 1 dB. Other associated results are tabulated in Table 2.
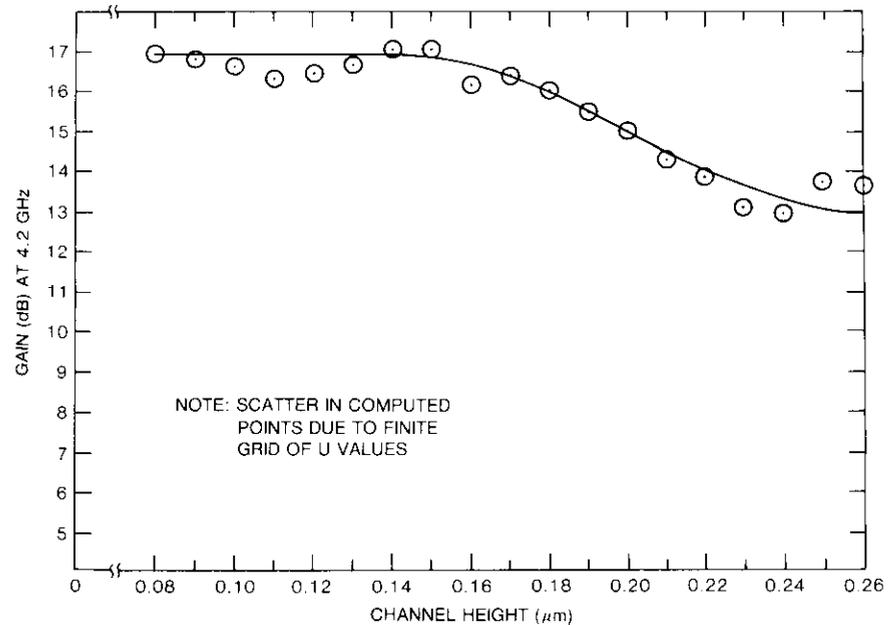

Figure 8. *Computed Gain* vs *Channel Height*
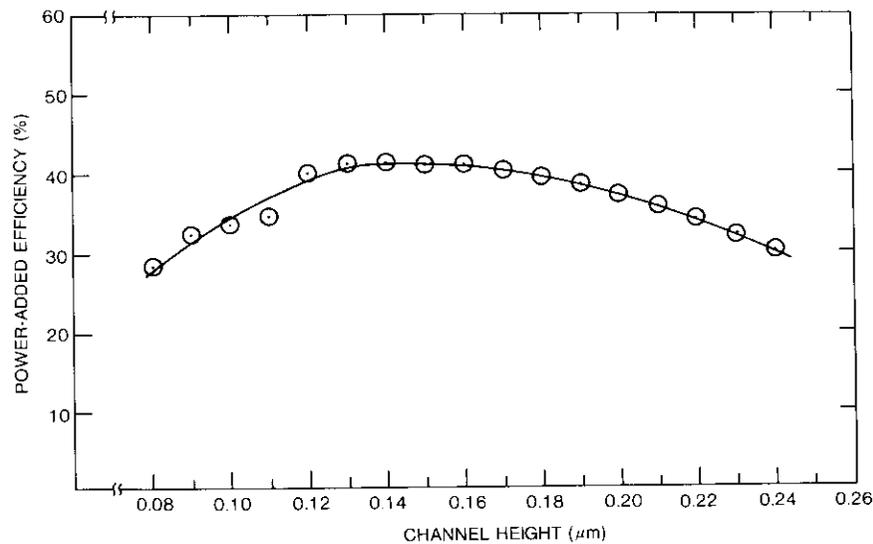
Figure 9. *Power-Added Efficiency* vs *Channel Height*

The most striking feature of the power *vs* channel height curve is the maximum near 0.2 μm. As expected, this is near the maximum predicted using the DC parameters of the FET, which is the channel height that maximizes area of a triangle under an RF load line drawn from the knee of the $\max(I_{ds})$ vs V_{ds} curve to the point of maximum allowed V_{ds} on the V_{ds} axis. At this point the sum of $|V_{po}|$ and the drain voltage equals the gate-drain avalanche breakdown voltage.

As noted above, because R_{out} was held constant, the actual power curve is expected to differ somewhat from the computed one. R_{out} is expected to decrease with increasing h. This would cause the output power, gain, and power-added efficiency to be lower than calculated for the larger channel heights. The effect of channel height on R_{out} may be the basis for the rule-of-thumb that the ratio of gate length to channel height should not be less than 3. The channel height corresponding to this ratio for the case of a gate length of 0.5 μm is indicated in Figure 7. A gate length of about 0.6 μm would make the ratio of 3 coincide with the peak. Of course, the gain for this gate length would be less than the gain for a gate length of 0.5 μm. Therefore the power-added efficiency would also be reduced. The model can be used to optimize a device through variation of several different parameters. For example, if the carrier concentration, *N*, is increased, the model would show that the peak would move to a smaller channel height.

TABLE 2. PINCH-OFF VOLTAGE, MAXIMUM CURRENT, AND LOAD IMPEDANCE FOR MAXIMUM POWER vs CHANNEL HEIGHT

h (μm)	V_{po} (V)	Max I_{ds} (mA)	Z LOAD FOR MAXIMUM POWER (Ω)
0.08	0.2	50.5	60.77 + 50j
0.09	0.4	105	60.46 + 47.85j
0.10	0.7	160	59.99 + 48.74j
0.11	1.0	215	61.48 + 48.83j
0.12	1.3	270	53.58 + 24.48j
0.13	1.7	325	41.51 + 13.34j
0.14	2.05	381	32.24 + 9.51j
0.15	2.5	437	25.99 + 6.50j
0.16	2.9	493	22.32 + 2.59j
0.17	3.4	549	18.13 + 2.63j
0.18	3.8	606	15.19 + 1.60j
0.19	4.4	662	12.85 + 0.69j
0.20	4.9	719	10.93 + 0.12j
0.21	5.5	775	9.20 − 0.60j
0.22	6.1	832	7.71 − 0.65j
0.23	6.7	888	6.37 − 1.10j
0.24	7.4	945	5.17 − 0.63j
0.25	8.1	1,002	7.89 − 0.35j
0.26	8.8	1,057	7.97 − 0.39j

Conclusions

A modeling system for predicting the performance of GaAs power FETs has been described, and its application at COMSAT Laboratories results in predictions which may be used to optimize device designs. For example, for a given carrier concentration, a channel height for maximum output power can be found. Alternatively, if the channel height is fixed, a carrier concentration for maximum power can be found. Instead of maximum power, the modeling system can find maximum power added, maximum power-added efficiency, or the optimum value of some other parameter.

The effect of varying any input parameter on each performance parameter can be conveniently studied. The input parameters have been chosen to relate in a straightforward way to the semiconductor material, the device geometry, and the fabrication process. This makes the modeling system valuable for device design, troubleshooting of the fabrication process, and transforming of FET design and fabrication to specific applications.

Acknowledgments

The author wishes to express his appreciation to A. Muelenberg, H. Hung, A. Ezzeddine, and B. Geller for many helpful discussions. The encouragement and support of H. Huang is also greatly appreciated.

References

[1] R. E. Neidert and C. J. Scott, "Computer Program for Microwave GaAs MESFET Modeling," Naval Research Laboratory Report No. 8561, Washington, DC, 1982.

[2] A. Ezzeddine, private communication.

[3] B. Geller, private communication.

[4] K. Kotzebue, "Microwave Amplifier Design with Potentially Unstable FETs," IEEE Transactions on Microwave Theory and Techniques, Vol. MTT-27, No. 1, January 1979, pp. 1–3.

[5] Private communication, March 1980.

[6] H. Fukui, "Optimal Noise Figure of Microwave GaAs MESFETS," IEEE Transactions on Electron Devices, Vol. ED-26, No. 7, July 1979, pp. 1032–1037.

[7] A. J. Cote and J. B. Oaks, Linear Vacuum-Tube and Transistor Circuits, New York: McGraw-Hill, 1961, pp. 344–375.

Thane Smith received a B.S. degree in physics from the Massachusetts Institute of Technology in 1962, and M.S. and Ph.D. degrees from Carnegie-Mellon University in 1964 and 1968, respectively. He joined COMSAT Laboratories in 1975, and is currently Manager of the Materials and Device Modeling Department of the Microelectronics Division. His responsibilities include designing and modeling GaAs FETs and GaAs MMICs, and coordinating GaAs materials work with the Microelectronics Fabrication Technology Department. Since joining COMSAT, he has been responsible for developing FET and MMIC fabrication capabilities from their inception to near state of the art. He has also been involved in research and development of novel GaAs devices (AMP and SAMP) and epitaxial growth of GaAs. Dr. Smith is a member of Sigma Xi and a Senior Member of IEEE. He has coauthored more than 10 technical papers, presentations, and reports, and shares with P. L. Fleming the patent for the SAMP device.

Experimental measurement of solar absorptance of an Intelsat VI OSR radiator as a function of incidence angle*

W. H. KELLY AND J. H. REISENWEBER

(Manuscript received June 12, 1985)

Abstract

As part of the INTELSAT VI spacecraft development program, INTELSAT funded an experimental effort to determine the solar absorptance of optical solar reflectors (OSRs) as a function of temperature and incidence angle. Using the NASA Jet Propulsion Laboratories' solar thermal vacuum chamber, experimental data were obtained which indicate that the OSR solar absorptance is relatively independent of temperature and that it increases with increasing incidence angle. It was found that the magnitude of the increase in solar absorptance for a body-stabilized spacecraft during peak solar illumination and with a 66.5° incidence angle was about 20 percent larger than the value normally used in analyses based on spectrophotometer measurements. These results help to explain on-orbit temperature anomalies between predictions and flight data. The ability to correctly correlate the spacecraft thermal mathematical model with flight data is extremely important when planning different modes of spacecraft operation or in the event of component failure which results in abnormal spacecraft performance.

Introduction

The interest of INTELSAT and COMSAT in determining the solar absorptance of an OSR radiator at various solar incidence angles and temperatures was prompted by technical questions concerning the correlation of the INTELSAT V spacecraft thermal mathematical model predictions with flight telemetry temperature data and solar beam test results [1]. Furthermore, flight data from the COMSTAR spacecraft analyzed by COMSAT Laboratories indicated a nonconstant value for OSR solar absorptance during the year [2]. Earlier analytical and experimental work identified an increase in OSR solar absorptance at very high incidence angles [3]. INTELSAT provided the funding for the development of an experiment to measure the solar absorptance of an OSR radiator as a function of angle and temperature during the solar beam testing of the INTELSAT VI Thermal Test Model at the NASA Jet Propulsion Laboratory test facility.

Historical background

OSRs are utilized as the primary thermal control surface coating on a spacecraft because of their low and relatively stable solar absorptance and relatively high thermal emittance. Several body-stabilized spacecraft which employed OSRs, including INTELSAT V, have experienced higher than expected beginning-of-life temperatures. The reason for these higher temperatures had been attributed to contamination of the OSRs by either apogee motor exhaust products or initial outgassing of the spacecraft. This explanation has been a source of continuous controversy among thermal engineers because it is difficult to differentiate between an analytical thermal model which has prediction errors and the effect of OSR contamination when correlating flight temperatures. The beginning-of-life temperatures are the only ones which can be correlated with flight data because the OSRs degrade with exposure to the space environment at an unpredictable rate. This project was undertaken to help resolve some of the controversy.

Test sample and fixture description

INTELSAT requested that their radiator manufacturer, Messerschmitt-Bölkow-Blohm (MBB) of Germany, fabricate an OSR radiator sample, or coupon, using INTELSAT VI materials and processes. The 23 mm × 63 mm × 0.2 mm silver/quartz OSRs with metallic deposition coated edges were manufactured by Optical Coatings Laboratories, Inc. For the purpose of this

experiment, the OSRs were bonded to a 0.3 m × 0.25 m × 3.2 mm aluminum substrate with silver filled RTV-S691 adhesive. The interstitial gaps between the 52 OSRs on the coupon are 0.3 mm. The entire aluminum substrate was black anodized to simulate the thermal properties of the graphite fiber reinforced plastic (GFRP) facesheets on the INTELSAT VI radiator.

In order to meet the experimental test objectives, a fixture was designed to rotate the radiator coupon to selected incidence angles, as shown in Figure 1. The radiator coupon is mounted to a steel shaft via conductive isolators and may be rotated by means of a DC stepper motor to any desired angle with respect to the solar beam. The angle of rotation is determined from the output of an angular potentiometer. The temperatures of the radiator coupon, stepper motor, and angular potentiometer are manually controlled by adjusting the power to attached heaters in response to thermocouple readings. The coupon is covered with a multilayered insulation blanket on all sides except the OSR surface in order to minimize heat flow from these surfaces. A custom fabricated resistive heater bonded to the back of the aluminum substrate covers nearly the entire surface for equal heat distribution. The leads of the thermocouples and heater attached to the coupon are swirled prior to exit from the insulation blanket to minimize heat loss. The harness (power and thermocouple wires) from the coupon is wrapped with an insulation blanket and utilizes a guard heater to minimize heat flow from the coupon.

There are three calorimeters on the fixture to aid in determination of the solar intensity that is incident on the radiator coupon. The construction of the calorimeters is illustrated in Figure 2. One calorimeter remains normal to the solar beam throughout the test to detect variations in the solar beam intensity. The disk of this calorimeter is coated with Chemglaze Z306 black paint. A second calorimeter which determines if any heat flux other than direct solar energy is incident on the coupon is coated with the same black paint and rotates with the coupon. A third calorimeter, an aid in determining whether the additional heat flux (if any) on the rotating black calorimeter is solar or infrared radiation, rotates with the coupon and has silver teflon tape on the disk.

The stepper motor used on this experiment is a surplus INTELSAT V solar array drive motor. Electronics equipment allowed the coupon to be rotated clockwise or counterclockwise and then held stationary at a specified angle relative to the solar beam until steady-state temperatures were achieved. The angle of the coupon with respect to the solar beam was determined from the output of a linear rotary potentiometer attached to the fixture shaft. The coupon and the three calorimeters were angularly aligned within 0.25° of each other using a computer-aided coordinate measuring machine whose accuracy is 0.0025 mm over its measuring range.
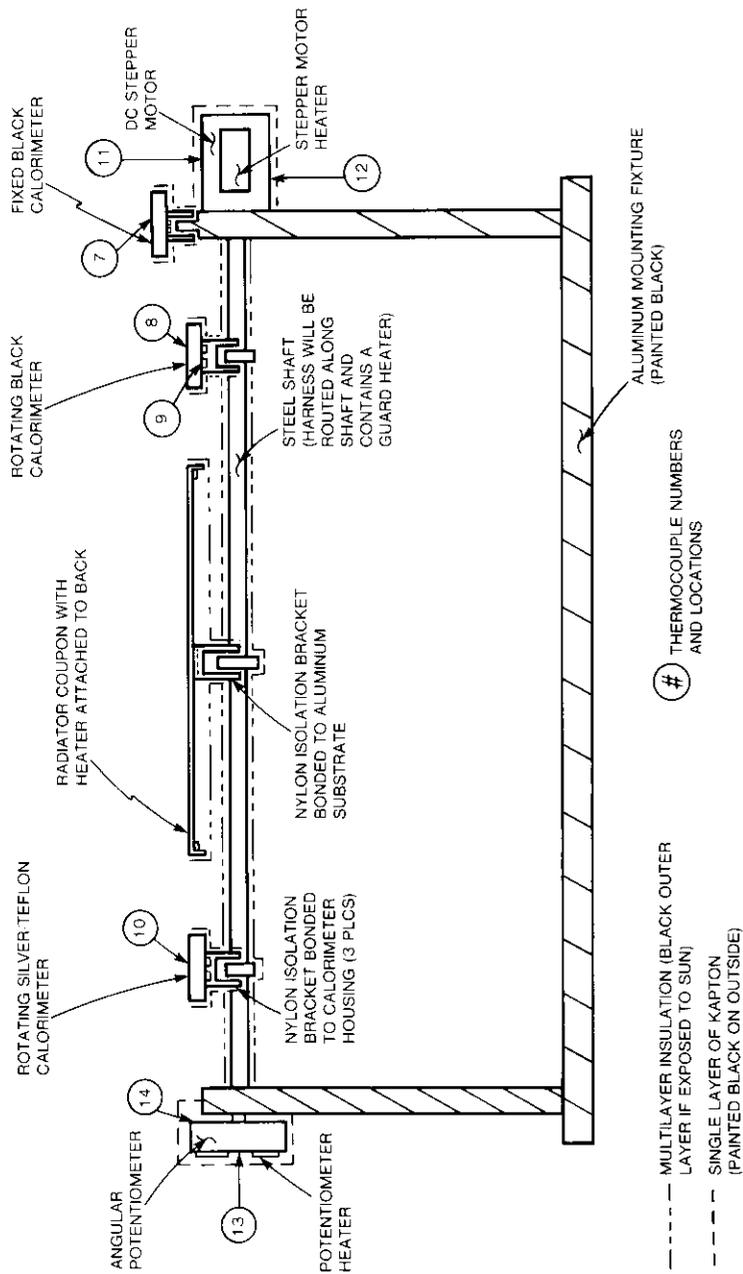
Figure 1. *Schematic of INTELSAT VI Radiator Coupon Experiment*
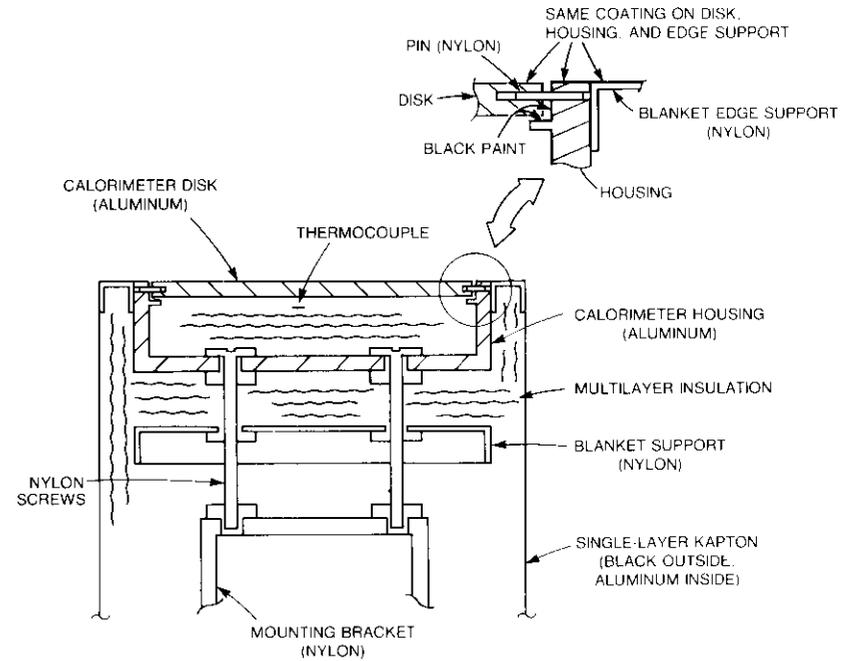
Figure 2. *Coupon Test Fixture Calorimeter Design*

Infrared tests

Prior to subjecting the coupon to solar beam testing, a series of infrared (IR) tests were performed at COMSAT Laboratories to determine the thermal characteristics and calibration of the coupon assembly. These tests were designed to provide the following thermal information:

● The calorimetric emittance of the OSR coupon as a function of coupon temperature.
● The heat loss from the coupon sample to the test fixture as a function of coupon temperature.
● The effective emittance of the multilayer insulation blankets used on the rear surface of the coupon.

The coupon test assembly, OSR insulation blanket, and multilayer insulation test sample are shown in Figure 3. The OSR coupon was placed in a 1-m × 2-m thermal vacuum chamber with liquid nitrogen shrouds and a series of
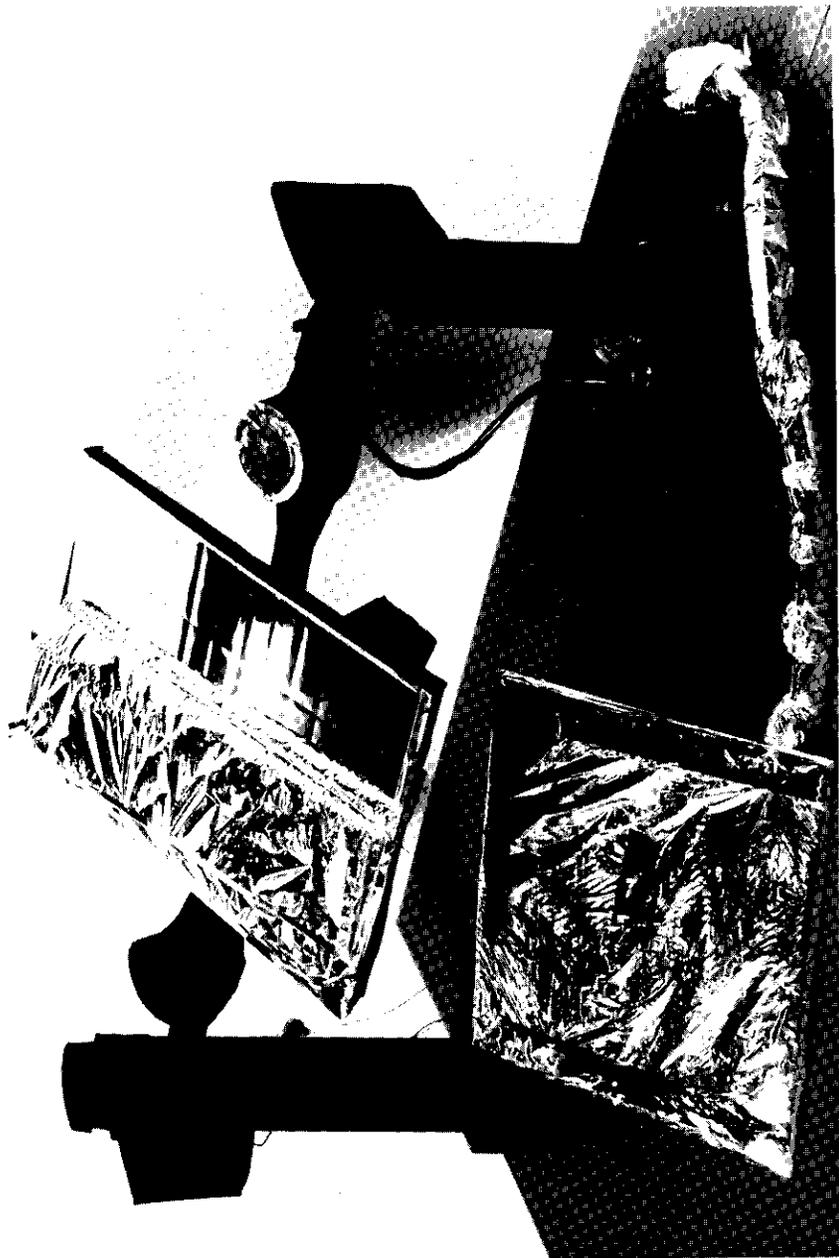
Figure 3. *INTELSAT VI Radiator Coupon Test Fixture and Associated Test Hardware*

test cases were conducted to determine the radiative and conductive heat transfer properties of the coupon over the temperature range of 50°C to −40°C. The test results were used to calculate the following thermal characteristics of the OSR coupon and test fixture:

- The effective emittance of the multilayer insulation blanket used to cover the front face of the OSR coupon is between 0.01 and 0.012 over the temperature range of interest.
- The equivalent effective emittance of the insulation around the OSR coupon (including edge and fixture heat leaks) is in the range of 0.05 to 0.075. The heat leak (both radiative and conductive) of the coupon sample *vs* temperature is presented in Figure 4.
- The OSR coupon calorimetric emittance is a constant value between 0.76 and 0.77 over the temperature range −40°C to 50°C. A plot of emissivity *vs* coupon temperature using best fit plots of the test data is shown in Figure 5. The measured value of emissivity for an individual OSR mirror using an IR reflectometer is 0.76.
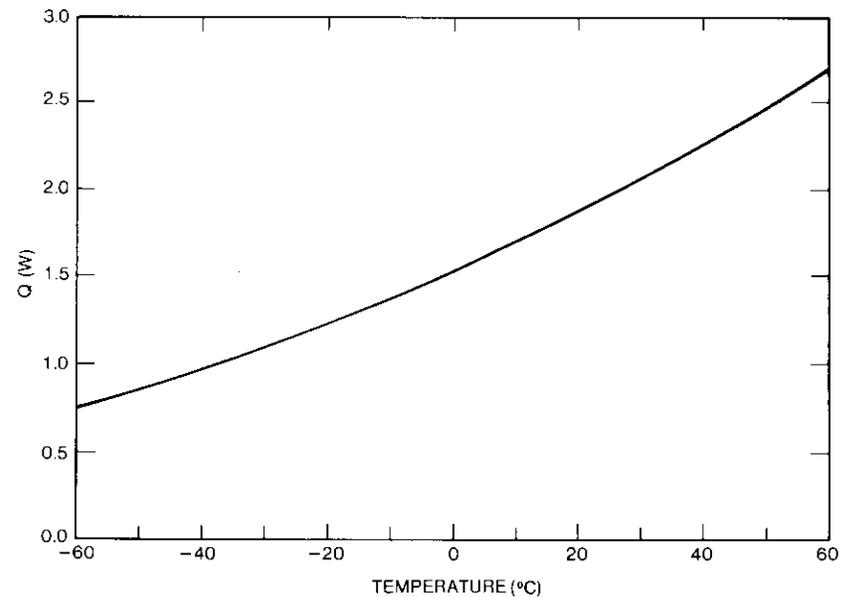


Figure 4. *OSR Insulation Heat Leak*
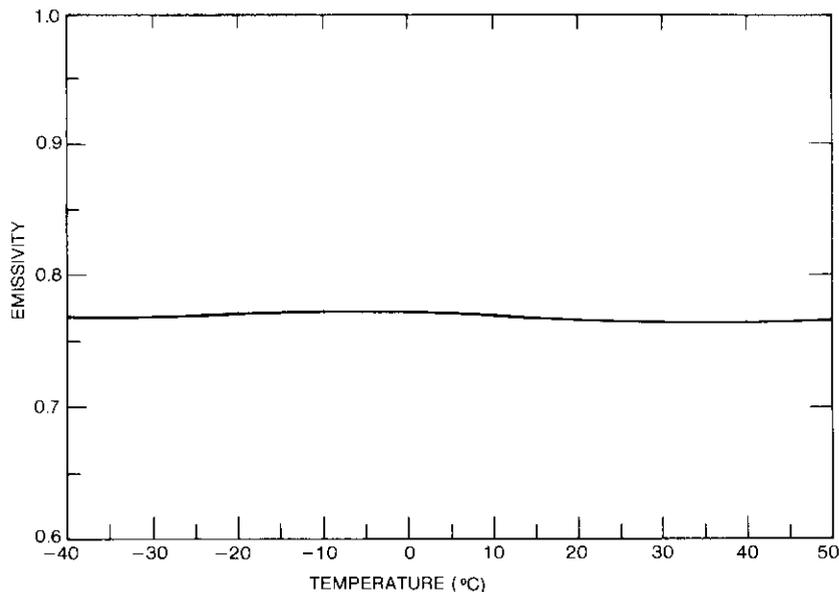
Figure 5. *OSR Coupon Emissivity IR Test Results*

The results from the series of IR tests provide the necessary thermal data to enable the calculation of the heat rejection of the OSR coupon as a function of temperature. Provided that the solar intensity is accurately known, the solar absorptance of the test sample in the solar beam can be determined by using the OSR heat rejection and the heat loss to the test fixture.

Solar beam tests

After the coupon was subjected to the series of IR tests at COMSAT Laboratories, solar beam tests were conducted at NASA's Jet Propulsion Laboratory (JPL).

JPL Test Facility

The JPL space simulator chamber is a stainless-steel cylindrical vessel 8.2 m in diameter and 25.9 m high with a 4.6 m × 7.6 m side-opening door. The walls and floor are aluminum cryogenic shrouds which may be controlled over a temperature range of −195°C to 90°C. The off-axis solar simulation system consists of an array of 37 xenon compact arc lamps of 20 to 30 kW, an integrating lens unit, a penetration window, and a one-piece

collimating mirror. The generated solar beam is reflected down into the test volume by the collimating mirror which is temperature controlled with gaseous nitrogen through a range of −70°C to 90°C.

The chamber test volume is illuminated by a beam of simulated solar energy 5.6 m in diameter with a nominal intensity of one solar constant (intensity is varied to compensate for nonuniformity of beam and seasonal variation). The illumination spectrum is derived from xenon arc lamps, as modified by the simulator optics. Beam uniformity is ±5 percent as measured by a PIN diode detector 6 mm in diameter. The collimation angle varies from ±1° to ±2° and a water-cooled douser is provided to simulate eclipse of the sun.

Location of test fixture

The INTELSAT VI spacecraft test fixture was designed with two massive side frames which supported the spin and tilt mechanisms and the entire spacecraft weight. The radiator coupon fixture and associated liquid nitrogen shrouds were mounted to the right side frame (when viewed from the chamber door) approximately 3.4 m above the floor of the JPL solar beam chamber and at a radial distance of 2.5 m from the center of the solar beam, with an unobstructed view of the chamber walls. A shroud with liquid nitrogen cooling coils blocked the view of the coupon from the INTELSAT VI spacecraft. Thermocouple data from the coupon was transmitted as a part of the normal spacecraft temperature data and stored on the JPL data acquisition system.

The mounting surface of the test fixture was painted black and incorporated liquid nitrogen cooling. The coupon was mounted to the fixture so that rotation from 0° to 90° increased the view to the chamber walls and not to the spacecraft side of the shroud. The location of the test fixture on the side of the main spacecraft spin fixture is illustrated in Figure 6.

After the completion of the first test phase in which the coupon and its fixture were present, an interference developed between the coupon test fixture and one of the spacecraft-deployable antennas. As a result, the entire assembly had to be moved outboard an additional 12.7 cm toward the chamber wall.

Solar Beam Intensity Mapping

The intensity of the solar beam in the JPL chamber was mapped using a PIN diode attached to a long boom which sweeps across the test volume from side to side, generating an analog plot of beam intensity versus position in the chamber. These data were then used to generate contour maps for the entire beam area. The solar intensity at the coupon location is derived by integrating the local contour map over the entire coupon area. This integration
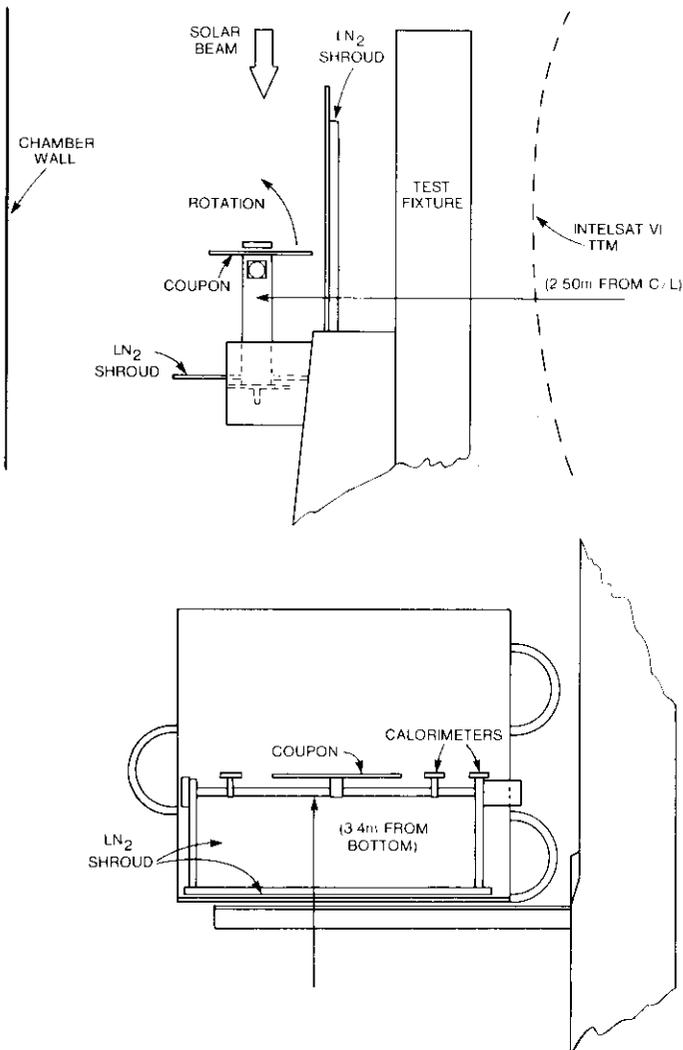
Figure 6. *Location of Radiator Coupon in JPL Solar Simulator*

results in a ratio of the coupon solar intensity to the Kendall calorimeter, which is the intensity standard for the chamber.

The contour plot resulting from the INTELSAT VI spacecraft pretest intensity mapping is shown in Figure 7. The scale represents the percent of solar intensity as measured at the reference Kendall calorimeter, *i.e.*, 1.15 equals
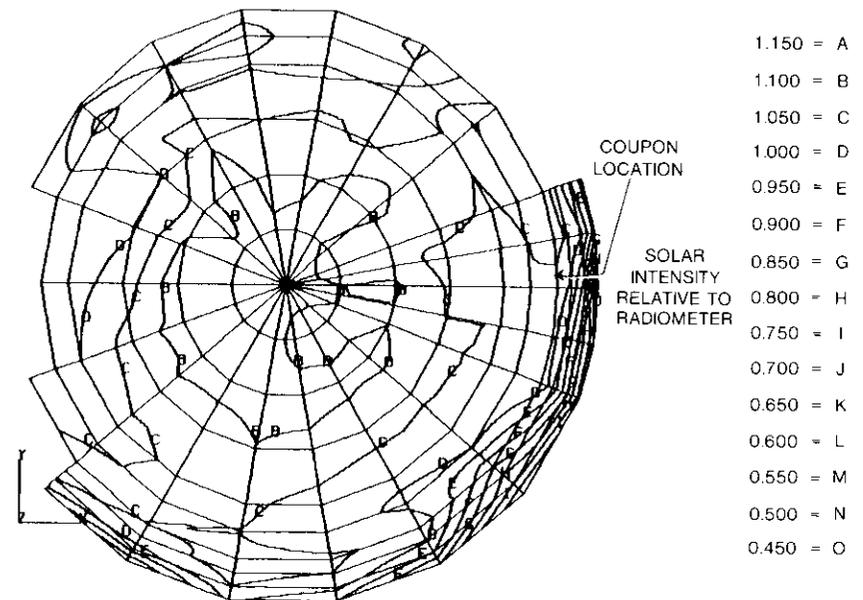
Figure 7. *INTELSAT VI Pretest Intensity Map of JPL Chamber*

an intensity level of 1.15 times the solar constant as determined by the Kendall calorimeter reading. In general, the intensity level at the coupon was on the order of one solar constant when the test fixture was at the original position in the chamber. By moving the fixture outboard, an intensity area with a rapid falloff near the solar beam edge was entered. A representative solar beam intensity contour plot at the original coupon location is shown in Figure 8. Similar contour maps were generated at the final coupon position for both pretest and post-test scan data.

Different values of solar intensity used during the INTELSAT VI thermal balance testing required intensity factors for the coupon to be obtained. Because of the size of the INTELSAT VI spacecraft, an integrated average of the solar intensity was used over the spacecraft and this average differed depending upon whether the spacecraft was tilted up +23.5° or down 23.5°, simulating solstice conditions.

An intensity contour map of the chamber was generated after the INTELSAT VI test was completed; it indicated that for the entire beam and coupon location, the post-test intensity was less than the pretest values.

```
1.100 = A
1.070 = B
1.040 = C
1.010 = D
0.980 = E
0.950 = F
SOLAR        0.920 = G
INTENSITY
RELATIVE TO  0.890 = H
RADIOMETER   0.860 = I
             0.830 = J
             0.800 = K
             0.770 = L
             0.740 = M
             0.710 = N
             0.680 = O
```
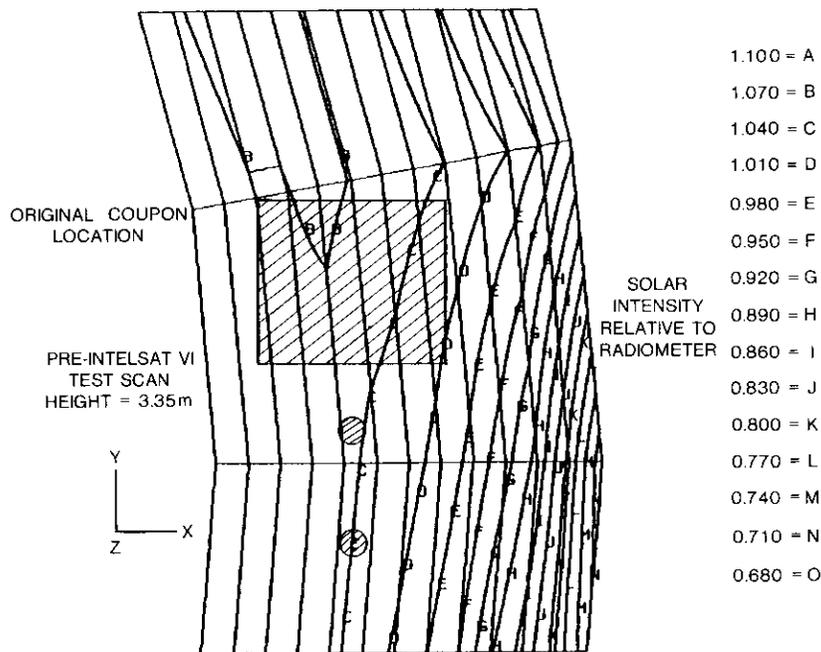
Figure 8. *Contour Plot of Representative Solar Beam Intensity at Original Coupon Location*

Test procedure

The Kapton film heater on the radiator coupon was used to maintain a constant temperature as the coupon was rotated to various solar incidence angles. Temperatures on the coupon for which measurements were made were 50°C, 25°C, 0°C, −25°C, and −40°C. Over 32 steady-state cases were obtained during the entire testing period at JPL. In addition, one point with no solar input (coupon rotated completely away from the solar beam) was obtained to verify the product of emissivity times area for the coupon.

Data recorded included the temperatures of the coupon and guard heaters, heater power, Kendall calorimeter intensity, the angle of the coupon with respect to the solar beam, and the temperatures of the calorimeters on the test specimen. The majority of the data points were obtained at temperatures of 25°C and 0°C because this is the typical spacecraft radiator temperature. Data points were obtained at other temperatures to identify any potential influence of temperature on OSR absorptance.

The solar beam was shut down during those periods when the 72-min.

spacecraft eclipse was being simulated, so data could not be taken continuously. However, much more data were obtained than had been originally planned for the test.

Test sequence

The first series of tests were conducted at a temperature on the radiator coupon of 25°C and with the test fixture at the original position, as shown on chamber drawings (approximately 2.5 m from the centerline of the chamber). A total of six data points were obtained during this test phase.

The interference problem between the INTELSAT VI spacecraft and the test fixture occurred during spacecraft reconfiguration for the next spacecraft test phase. The fixture was moved outboard 12.7 cm and testing was resumed. The data taken during this test phase were also at 25°C, with repeats being made at the 0° solar beam incidence angle (solar beam normal to OSR surface) to establish a correlation between the two fixture locations.

This test phase ended when almost one-third of the xenon arc lamps were destroyed and the test had to be terminated in order for repairs to be made. After the new lamps were installed and the test resumed, a check at the 25°C coupon temperature and 0° solar angle was made prior to proceeding on to the remaining temperature cases. At least seven different measurements of solar absorptance at the 0° incidence angle were obtained to provide a baseline to normalize the other data points so as to account for variations in local solar beam intensity and collimation.

Test data

The data from the solar beam test, together with the results of the COMSAT Laboratories IR tests, were used to compute the coupon OSR solar absorptivity. The solar absorptance, α_s, of the coupon radiator is given by

$$\alpha_s = \frac{\epsilon A_r(\sigma T_r^4 - \sigma T_c^4) + Q_L}{A_r S \cos \theta}$$

where ϵ = coupon emissivity (derived from IR test data)
A_r = coupon radiator area, m²
σ = Stefan-Boltzmann constant, W/m² K⁴
T_r = average temperature of coupon thermocouples, K
T_c = chamber wall temperature, 90 K
Q_L = heat leak from coupon as a function of temperature (derived from IR test data), W
S = solar intensity, W/m², calculated from experiment and chamber calorimeters
θ = solar incidence angle, degrees, (0° angle is solar vector normal to OSR surface).

OSR solar absorptance absolute value is dependent upon the accuracy of determining (by computation or measurement) the solar intensity. The experiment calorimeters were subjected to a series of tests at COMSAT Laboratories to determine both the heat leaks from the calorimeter disks to the test fixture and the radiative properties of the disks. This information was used to augment the measurements of the JPL test facility Kendall calorimeter in deriving values for solar intensity.

Throughout the test sequence, the coupon was rotated back to the 0° position to establish a reference baseline for correlation of all the test data points. The mean value of the OSR solar absorptance at the 0° incidence angle is 0.06, whereas Hughes Aircraft Company computed a value of 0.052 for a single OSR in a xenon arc spectrum [4]. COMSAT's corresponding values of absorptance for the solar spectrum are 0.071 (single OSR) and 0.08 for the OSR radiator.

The computed solar absorptance values for all temperatures are plotted together with a least squares curve fit in Figure 9. Test results demonstrated that solar absorptivity is relatively independent of temperature. However, as the incidence angle increases, there is an increase in solar absorption, as shown by this curve.
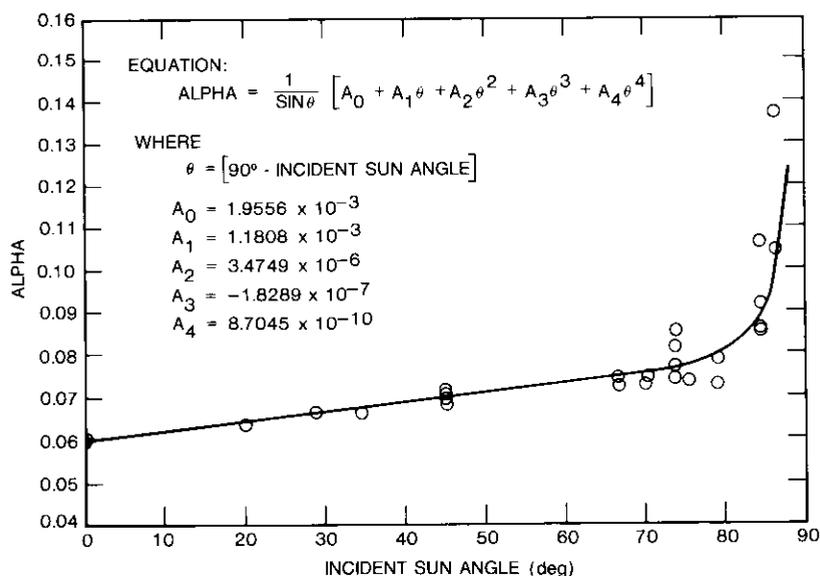

Figure 9. *Solar Absorptance* vs *Incident Sun Angle*

At the point of interest to geosynchronous applications, 66.5°, the increase in absorptance is approximately 20 percent. The values of solar absorptance calculated at incidence angles greater than 75° are difficult to accurately quantify. The sources of error which can enter into the calculations are positioning of the coupon, the collimation and uniformity of the solar beam, and the measurement of small heat inputs. Possible error in solar absorptance calculations increases as the incidence angle approaches 90°. However, the thermal consequences of this increase in probable error are minimal because the projected area receiving solar energy is decreasing towards zero.

Conclusions

The INTELSAT OSR radiator experiment has demonstrated that the solar absorptance of the OSR primary thermal control coating used on the INTELSAT spacecraft does experience an increase of 20 percent at 66.5° incidence angles over the thermal property data measured at normal angles of incidence and furthermore, that solar absorptance is relatively independent of temperature. This increase in solar absorptance at 66.5° is important because it aids in correlation of the thermal mathematical models with beginning-of-life spacecraft flight temperatures.

When used to predict temperatures, an increase in solar absorptance of 0.02 on the initial value of 0.08 (INTELSAT V beginning-of-life value for actual sun) is insignificant compared with the degraded end-of-life value of 0.21 or more. On INTELSAT V, the sensitivity of the spacecraft to a 0.01 change in solar absorptance is approximately 1°C. However, the ability to correctly correlate the spacecraft thermal mathematical model with flight data is of extreme importance when different spacecraft modes of operation are planned or in the event of the failure of certain components which result in the spacecraft operating in an unexpected mode. The ability to accurately predict temperatures for these events could have a direct bearing on the revenues which are generated over the lifetime of a spacecraft.

Evidence of an increase in solar absorptance due to an increase in incidence angle provides the thermal analyst with the means to correlate a thermal model with flight data without allocating the major source of the inaccuracy to an unmeasurable series of events associated with the apogee kick motor (AKM) firing. It is recommended that in the future INTELSAT base their temperature predictions on a beginning-of-life OSR solar absorptance of 0.10 for a body-axis stabilized spacecraft and 0.09 for a spin-stabilized design (solar vector normal to spin axes) instead of the commonly used 0.08. This recommendation does not imply that there are no contamination effects associated with the AKM firing, but it does mean that the assumed magnitude should be reduced.

It should be pointed out that the data collected during this experiment could be improved upon at very high incidence angles. A more uniform solar intensity or better location in the solar beam could yield additional useful data to the spacecraft designer. Furthermore, an attempt was made to duplicate the calorimetric results with spectrophotometer measurements of OSR solar absorptivity as a function of angle. Unfortunately, limitations in spectrophotometer capability precluded obtaining meaningful and consistent results.

Acknowledgments

References

[1] A. L. Spencer et al., "INTELSAT V Thermal Design, Testing and Flight Performance," AIAA and ASME Third Joint Thermophysics, Fluids, Plasma, and Heat Transfer Conference, St. Louis, Missouri, June 1982, AIAA Paper 82-0863.

[2] N. L. Hyman, "Solar Absorptance Degradation of OSR Radiators on the COMSTAR Satellites," in Spacecraft Radiative Transfer and Temperature Control, T. E. Norton, ed., Progress in Astronautics and Aeronautics, Vol. 83, New York, American Institute of Aeronautics and Astronautics, 1982, pp. 212–232, AIAA paper 81-1185.

[3] J. W. Stultz, "Solar Absorptance of Second Surface Mirrors for High Angles of Incidence," paper presented at the AIAA and ASME Thermophysics and Heat Transfer Conference, Boston, Massachusetts, July 1984, AIAA paper 74-0670.

[4] Private communication, June 1985.

William H. Kelly received an A.B. in Physics from West Virginia University in 1964, and pursued graduate studies at Catholic University. He joined COMSAT Laboratories in 1980 and is currently a Member of the Technical Staff in the Spacecraft Technology Division, where his work includes engineering support for the INTELSAT spacecraft, design monitoring and analysis of the DBS spacecraft, and technical support to the Italian government for the ITALSAT spacecraft. Prior to joining COMSAT, he was with Fairchild Industries, where he was a senior member of the Thermal Engineering group. He was responsible for the thermal design and testing of several spacecraft including ATS-6, ISEE-C (recently renamed ICE), and SMM. Mr. Kelly is a Member of AIAA, and the author of numerous technical papers.

John H. Reisenweber received a B.S. from the University of Maryland in 1959, an M.S. from Stevens Institute of Technology in 1964, and took advanced graduate courses at George Washington University, all in mechanical engineering. He joined COMSAT Laboratories in 1974 and is presently Manager of the Thermal Design and Environmental Test Department. His responsibilities include research, development, and support efforts in the areas of spacecraft and terrestrial thermal design. He is also responsible for activities in the Environmental Test Laboratory. Mr. Reisenweber has been involved in the thermal design, analysis, testing, and contractor monitoring on INTELSAT V, INTELSAT VI, ARABSAT, and MARISAT spacecraft. He has participated in spacecraft feasibility studies, and is presently developing the thermal design for a multibeam phased array. Prior to joining COMSAT, he worked in the thermal design departments at RCA and Fairchild. Mr. Reisenweber is a Member of Pi Tau Sigma and AIAA.

Linearized transponder technology for satellite communications

Part I: Linearizer Circuit Development and Experimental Characterization*

D. CAHANA, J. R. POTUKUCHI, R. G. MARSHALEK, AND D. K. PAUL

Abstract

The operational principles and design of transmission and reflection types of predistortion diode linearizers are presented. The performance of four different linearized traveling wave tubes (LTWTS) in multicarrier operation is evaluated experimentally, and the measured performance of a soft-limited LTWT is presented. Experimental results on linearization at an intermediate frequency of a transponder are also given. Some of these results are used in Part II of this paper to perform computer simulations of various LTWT configurations in a transmission channel.

Introduction

When multiple carriers are simultaneously amplified in a transmitter, the intermodulation products that result from amplitude and phase nonlinearities of the high-power amplifier (HPA) become a source of noise in the transmission band. To reduce these nonlinear effects, the operating point of the HPA must be sufficiently backed off relative to the saturation point. This results in a

---

decrease in the DC-to-RF efficiency of the HPA. With linearization, a specified carrier-to-intermodulation ratio (C/I) can be achieved at a higher operating point than would otherwise be permitted. This is possible because the effective amplitude and phase characteristics of the HPA can be modified to become more linear, as shown schematically in Figure 1 for a typical TWT amplifier.
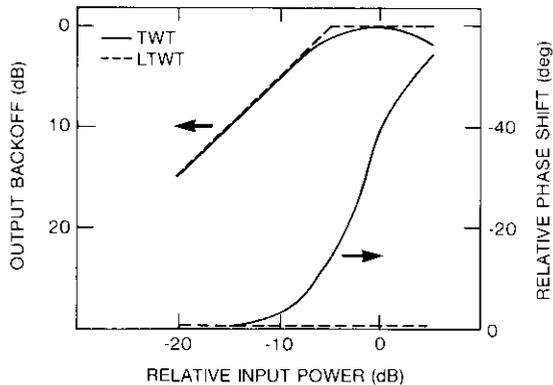

Figure 1. *Typical TWT and Ideal LTWT Transfer Characteristics*

The link quality of a typical satellite system can be quantified by its total carrier-to-noise ratio, $(C/N)_T$. Ignoring second-order sources of noise such as earth station intermodulation noise or cross-polarization interference, the $(C/N)_T$ of a multicarrier system can be expressed as

$$(C/N)_T = \frac{1}{1/(C/N)_U + 1/(C/N)_D + 1/(C/I)}$$

where $(C/N)_U$ and $(C/N)_D$ are the up- and down-link carrier-to-noise ratios, respectively, and C/I is the carrier-to-intermodulation ratio of the transponder. The solid lines in Figure 2 show the three components of $(C/N)_T$ of a typical communications system as functions of the up-link power to the satellite. The variable $(C/N)_U$ is linear with power, $(C/N)_D$ has a nonlinearity which is governed primarily by the TWT of the transponder, and C/I drops nonlinearly with power due to transponder nonlinearity. The resulting $(C/N)_T$ has a maximum at a backed-off operating point and is bounded above by C/I near saturation.

With linearization, C/I performance improves and the associated $(C/N)_T$ curve peaks at an up-link power level closer to that causing saturation, as indicated by the dashed lines. Consequently, for maximum $(C/N)_T$, the
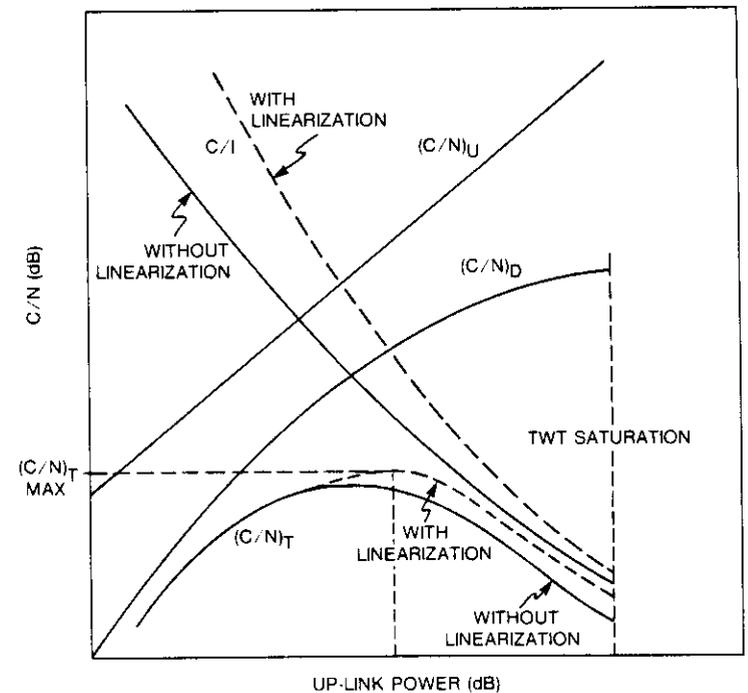
Figure 2. *Typical Link Performance With and Without Linearization*

operating point of the transponder can be moved closer to saturation, with a resulting increase in the effective isotropic radiated power (e.i.r.p.) and in the DC-to-RF efficiency of the TWT. This improvement enables an increase in the ratio of transponder capacity (number of communications circuits) to DC power consumption, if bandwidth is not a limitation. On the other hand, if capacity is not increased, excess e.i.r.p. can be minimized by sizing down the transponder TWT; that is, by using an amplifier with a lower saturated power output rating, P_{sat}. A lower P_{sat} and an operating point closer to saturation also result in lower DC power consumption.

Investigations of linearization have produced three techniques: the feedback amplifier, feed-forward linearization, and predistortion linearization. The feedback technique for TWT amplifiers is not suitable for satellite transponders because it is narrowband [1]. The feed-forward technique has been successfully applied to cable amplifiers and further investigated for use in earth station HPAs [2]. This technique utilizes two TWTs operating in parallel: the main TWT which provides most of the power, and an auxiliary TWT with lower P_{sat}. By adjusting the relative gain and phase between the two amplifiers, it

is possible to cancel selected intermodulation products at the output. The feed-forward approach is not attractive for satellite transponders for two reasons. First, linearized amplifier performance is sensitive to the static phase difference of the two TWTs. Second, the addition of an auxiliary TWT and a power combiner impose a relatively high weight penalty and an increased output loss, respectively.

In the predistortion technique, the amplitude and phase nonlinearities of the TWT are compensated for by predistorting the input signals with gain and phase nonlinearities which are approximately the inverse of the TWT nonlinearities. Predistortion techniques were developed in the early 1970s [3] for earth station HPAS, but were not implemented at that time because of performance limitations. Since then, simpler predistortion schemes have been developed using MIC technology, and diodes [4]. More complex linearizers utilizing nonlinear FET amplifiers have also been designed [5], [6].

Recent developments have shown that the diode predistortion technique is attractive for satellite applications. This linearizer type is simple and compact in design, and exhibits low thermal sensitivity and relatively large bandwidth. Also, unlike the feed-forward approach, this technique produces no added output losses and its performance is independent of absolute phase variations of the TWT.

This paper describes the operating principles of the predistortion diode linearizer and discusses some of its practical limitations. The design and construction of a linearizer on microstrip are presented, and experimental linearization is illustrated. Results are given for four TWTs operating at L-, C-, and K_u-bands, with different transfer characteristics. For each TWT, the transfer characteristics before and after linearization are presented. The associated intermodulation performance is illustrated using the results of two-tone and noise-power-ratio (NPR) tests. The latter measures the C/I at the center of a band uniformly loaded with a noise signal that has a rectangular spectrum. This approximates the condition of a densely filled channel. Measurement results show that the intermodulation performance of a linearized TWT in a selected output backoff range is essentially independent of the transfer characteristics of the TWT.

For QPSK/TDMA operation at or beyond saturation, the desirable TWT characteristics are those of a soft limiter, that is, linear gain below saturation and constant output power beyond saturation, with constant phase throughout these ranges. Theoretical and experimental results have shown that, beyond saturation, the effect of conventional linearization is usually lost and the transfer characteristics of the LTWT resemble those of the TWT [7], [8]. However, as shown experimentally below, by adding a soft limiter before the linearizer, the power output and phase shift of the LTWT can be held

nearly constant for a significant overdrive range. The limited-linearized-TWT (LLTWT) can thus operate in both the backed-off multicarrier and the saturated single-carrier modes with improved performance. Computer simulations of LLTWT configurations and the associated system performance are presented in Part II of this paper [9].

Linearization at the intermediate frequency (IF) of a transmitter has been realized for terrestrial microwave links [10], [11]. Similarly, in a channelized satellite transmitter, one linearizer may be used at the intermediate frequency to predistort the input signals of two or more TWTs. For example, in the INTELSAT V and VI satellites, up-link signals at 6 and 14 GHz which are first converted to a 4-GHz IF could be linearized at this frequency before up-conversion to 11 GHz and retransmission. Experimental results on the linearization of an 11-GHz TWT at a 4-GHz IF are given below. Part II of this paper provides a more detailed evaluation of this scheme, which enables a reduction in the number of linearizers in a transponder.

Principles of operation

The predistortion linearizer offsets the nonlinear gain and phase lag characteristics of a TWT by producing drive-dependent gain expansion and phase advance in cascade with the TWT. Predistortion linearizers can be generally classified as transmission-type (TRL) or reflection-type (RFL). The principle of operation is the same for the two types, but the RFL requires fewer components. Figures 3 and 4 show simplified block diagrams of the TRL and RFL, respectively. In both cases, the input power is divided into two
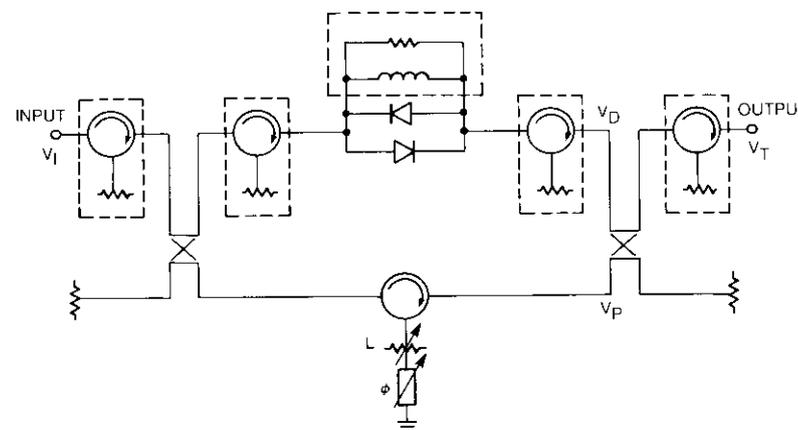

Figure 3. *TRL Schematic Diagram*
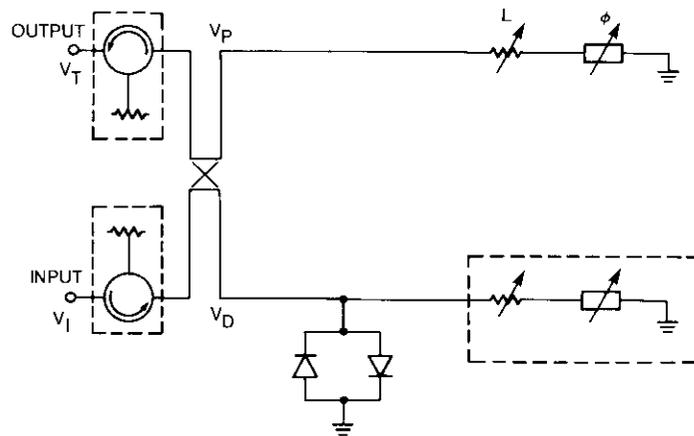
Figure 4. *RFL Schematic Diagram*

branches: a nonlinear branch consisting of two diodes mounted back to back, and a linear branch consisting of an attenuator, L, and a phase shifter, ϕ.

At low drive levels, the dynamic impedance of the diodes is high ($\sim$1 kΩ). Hence, V_D is low in the TRL and high in the RFL (Figure 5a). At high drive levels, the dynamic impedance of the diodes is low ($\sim$50 Ω). Hence, V_D is high in the TRL and low in the RFL (Figure 5b). The output voltage, V_T, is always equal to the vector sum $V_P + V_D$ within a constant. For intermediate drive levels, the ratio of output to input voltage, V_T/V_I, changes nonlinearly, and the insertion phase $\angle V_T - \angle V_I$ increases.

Figures 6a and 6b illustrate the computed change in gain ($\Delta P_{\text{out}}/\Delta P_{\text{in}}$) and the relative changes in insertion phase, respectively, as functions of the input drive level for five combinations of L and ϕ of a TRL. In Figure 6a, the 0-dB slope represents the linear input-output power relationship; that is, where a 1-dB change in input power produces a 1-dB change in output power. At drive levels below -5 and above 15 dBm, the dynamic impedance changes of the diodes are small, and consequently the linearizer is almost linear. For example, if L and ϕ are set to -10 dB and 250°, respectively, the relative slope is close to 0 dB for drive levels up to 2 dBm, and the relative phase advance is 10°. At drive levels between 2 and 15 dBm, there are gain slope variations (both compression and expansion), with $\Delta P_{\text{out}}/\Delta P_{\text{in}}$ between -0.03 and 0.2 dB, and a rapid phase advance from 10° to 50°. At drive levels higher than 15 dBm, the gain expansion and phase advance rates decrease. Hence, the desired linearizer gain expansion and phase advance characteristics can be achieved by a suitable choice of L and ϕ. However, it

ϕ = FIXED PHASE SHIFT BETWEEN THE LINEAR AND NONLINEAR BRANCHES

θ = DRIVE-DEPENDENT INSERTION PHASE

V_T/V_I = DRIVE-DEPENDENT INSERTION GAIN
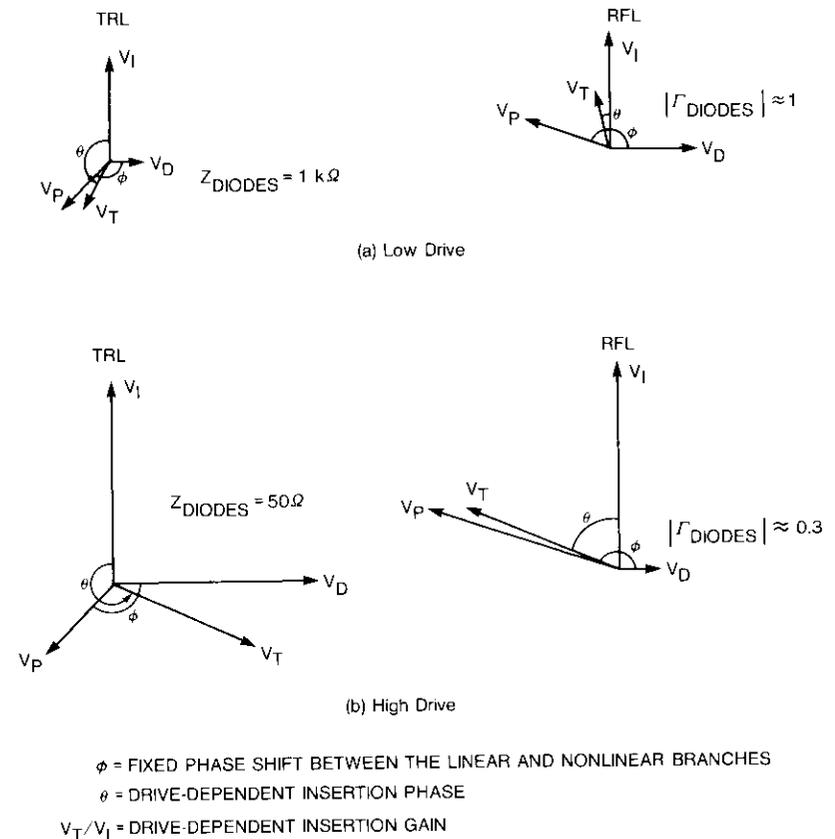
Figure 5. *TRL and RFL: Principles of Operation*

should be noted that the drive range over which this can be accomplished is limited and depends on L, ϕ, and the characteristics of the diode.

The input drive range required in order to obtain the desired linearizer characteristics is lower for diodes with low forward voltage, V_F (bias required for 1-mA forward current). In the computed results shown in Figure 6, a diode with a V_F of 0.45 V was assumed.

In addition to producing the desired shape of gain and phase nonlinearities, it is also necessary to match the linearizer absolute output power to the TWT drive requirements. Therefore, an attenuator is usually needed for TWT drive level adjustment.
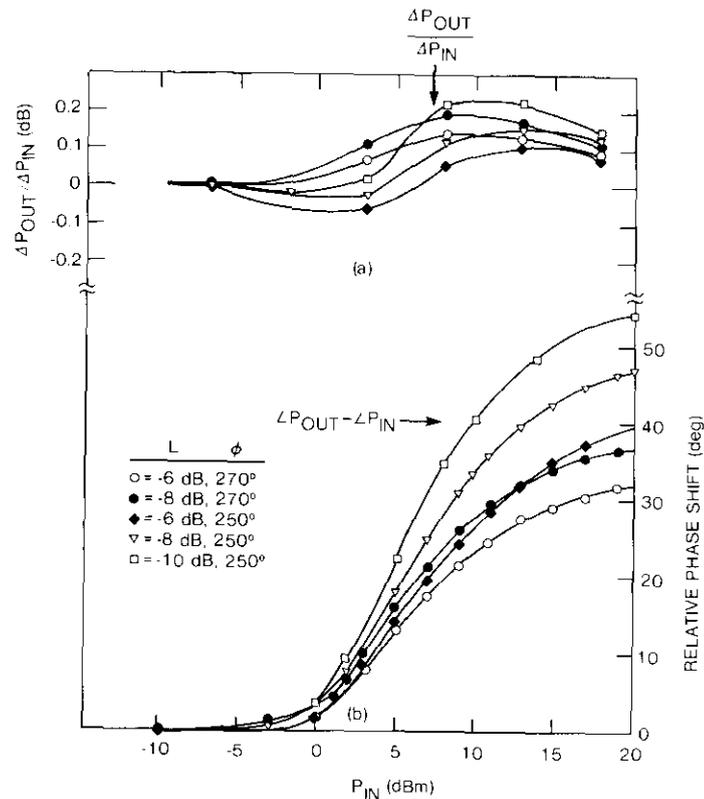
Figure 6. *Computed Variations of TRL Relative Gain and Phase Shift as a Function of Parameters L and φ*

Linearizer design and fabrication

Figure 3 is a schematic diagram of a typical TRL consisting of an input isolator, a power divider, linear and nonlinear branches, an output power combiner, and an output isolator. The linear branch consists of an attenuator and a phase shifter coupled by a circulator, whereas the nonlinear branch consists of two Schottky barrier diodes mounted back to back between two isolators. Schottky barrier diodes are desirable for their low forward voltage, and hence low drive requirements. Although the junction capacitance, C_j, of these diodes is low compared to that of other diodes, it nevertheless varies with input drive and produces an undesired, drive-dependent phase shift in addition to a frequency-dependent phase shift. These effects can be neutralized

over the desired operating frequency and drive level range by shunting the diodes with a suitable low-Q inductor, as shown.

Figure 4 is a schematic diagram of an RFL. As with the TRL, the input power is divided by a hybrid coupler and applied to linear and nonlinear branches. The linear branch consists of a transmission line terminated by an attenuator and a phase shifter. The nonlinear branch consists of two back-to-back shunt-mounted diodes. Here, C_j can be conveniently resonated with a variable attenuator and a phase shifter in cascade, as shown. In addition, circulators are not needed in each branch, and are provided only at the input and output ports of the hybrid coupler for high return losses. The reflected power from these branches is recombined in the hybrid coupler.

Although the performances of the two linearizers are similar from an analytical viewpoint, the RFL is more attractive for spacecraft applications because of the smaller number of isolators and couplers needed. Furthermore, as shown in Figures 3 and 4, the diodes of the RFL are more accessible and their parasitic elements, including C_j, can be more easily tuned out. Figure 7 is a photograph of an experimental MIC C-band RFL developed at COMSAT Laboratories. The unit is fabricated on a 25-mil alumina substrate. Power division and combining are performed by the 3-dB Lange coupler; however, a branch line coupler could be used as well, with similar performance in a 15-percent bandwidth. The linear branch phase shift and attenuation are
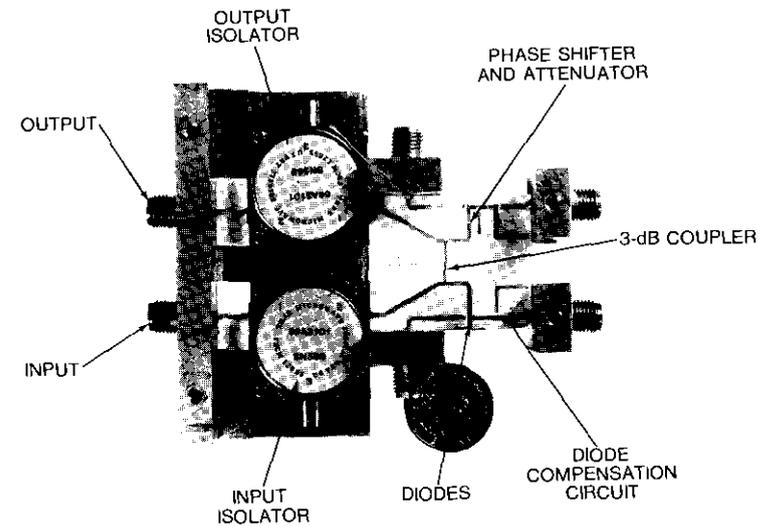
Figure 7. *4-GHz Linearizer*

provided by an open-ended microstrip line. The Schottky barrier diodes (V_F = 0.3 V and C_j = 0.13 pF) are terminated with a complex impedance whose real part is approximately 50 Ω. For tuning purposes, several small segments of transmission lines are provided which can be bonded to the main transmission lines as needed. To meet a 160-MHz bandwidth specification, the transmission line lengths are limited to a half-wavelength or less. The above design approach was also used to develop experimental L- and K_u-band linearizers.

Experimental results

This section presents the experimental results of L-, C-, and K_u-band TWTs and LTWTs. The results include transfer characteristics, two-tone intermodulation performance, and NPR performance.

Test setup

Figure 8 shows the experimental setup used for measuring transfer characteristics and intermodulation performance. For single-carrier transfer characteristics and two-tone intermodulation measurements, the outputs of one or two tone generators at 70 and 71 MHz were up-converted to the appropriate frequency band, filtered, and amplified before being applied to the subsystem under test. For the NPR measurements, the band-limited (316 to 12,360 kHz) and notched (6,300 kHz) output of a noise generator was first up-converted to 70 MHz, filtered, and further up-converted to the appropriate frequency band. The NPR was computed from the ratio of two noise-power measurements at the notch filter center frequency, with the filter switched into and out of the circuit. With the resolution and video bandwidths of the HP8566 spectrum analyzer set to 30 kHz and 30 Hz, respectively, and using the internal averaging capability to average 20 samples, NPR measurement repeatability was within ±0.1 dB.

To ensure measurement accuracy, the linearity and noise performance of the test system alone were adjusted up to the input of the subsystem under test. First, linearity was established using a two-carrier measurement. Next, total noise performance was established using an NPR test. The NPR data were later used to separate the performance of the subsystem under test from the measured overall system performance. For example, a measured NPR of 28 dB in a test system characterized by an NPR of 38 dB would be corrected by +0.4 dB.
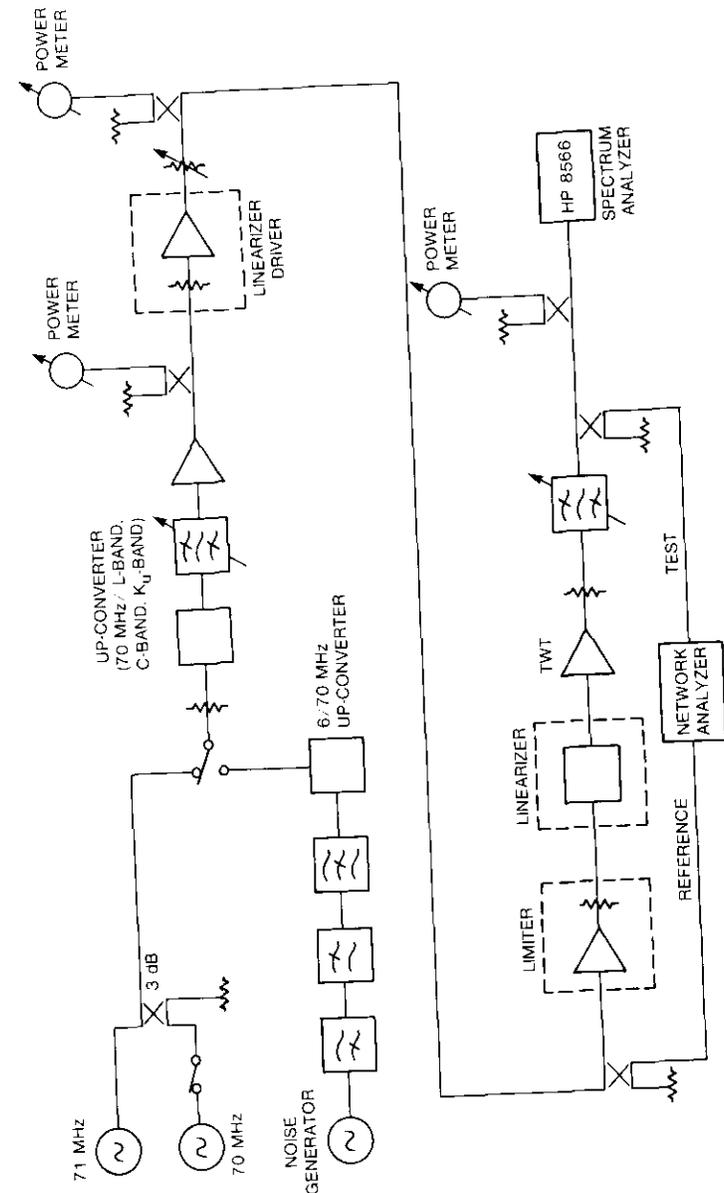
Figure 8. General Experimental Setup

447

Figure 9. *TWT1 and LTWT1 Transfer Characteristics*

Figure 10. *C/I and NPR of TWT1 and LTWT1*

performance. Figure 10 gives the measured two-tone C/I for third-order products and midband NPR as a function of TWT effective output backoff.* The predicted $(C/I)_3$ and NPR, also shown in this figure, are in very good agreement with the measured values.

A COMSAT Laboratories RFL was used in cascade with the TWT, and the combination was optimized as described above for operation about a 22-dB NPR. A 7-dB output backoff operating point was chosen so that the LTWT would be optimized for an NPR of approximately 20 to 26 dB. The TWT and LTWT transfer characteristics are compared in Figure 9. With linearization, improved gain linearity and phase flatness were observed for drive levels of up to 2-dB input backoff. The relative phase shift of the LTWT at saturation was $-16°$, compared with $-42°$ for the TWT. The intermodulation performances of the TWT and LTWT at 3.8 GHz are described in Figure 10 and Table 2.

*Total output power less the intermodulation product power, relative to single-carrier saturation.

TABLE 2. NPR OF TWT1 BEFORE AND AFTER LINEARIZATION

NPR (dB)	OUTPUT BACKOFF (dB)		
	TWT1	LTWT1	IMPROVEMENT WITH LINEARIZATION
15	4.8	3.1	1.7
20	8.3	4.9	3.4
25	11.0	7.0	4.0

At operating points sufficiently below the optimized level (*e.g.*, at 14-dB output backoff with NPR > 30 dB), the improvement in linearization is significantly diminished. However, for typical satellite links, NPR improvement at levels above 30 dB would affect overall system performance only

marginally. The measured LTWT NPR variation with frequency is also illustrated in Figure 10. The output backoff variation over a 160-MHz band centered at 3.8 GHz is less than 0.3 dB for NPR ≤ 28 dB.

TWT2 PERFORMANCE

A TRL built by Thomson CSF was optimized with TWT2 at 5-dB output backoff. As shown in Figure 11, phase response is improved with less than 5° variation from −20 to −4 dB drive relative to saturation. As indicated in Table 1, the normalized 1-dB gain compression point of this TWT is higher than is typical. In fact, the 5-dB output backoff is well within the linear gain region. Nevertheless, Figure 11 shows that the LTWT gain response is degraded in the 6- to 2-dB output backoff range. Gain is slightly compressed from 6- to 4-dB output backoff, and slightly expanded from 4- to 2-dB output backoff. Finally, between 2- and 1-dB output backoff, the LTWT gain shows some linearity improvement as compared to the TWT.
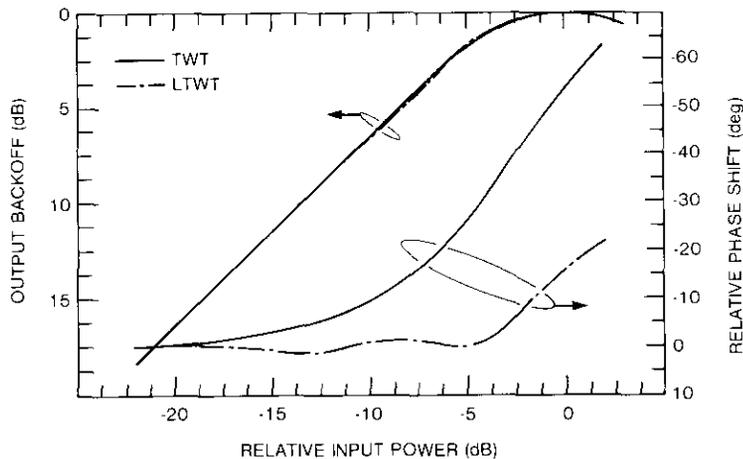
Figure 11. *TWT2 and LTWT2 Transfer Characteristics*

The observed LTWT gain performance is probably caused by the following phenomena:

a. In multicarrier operation, the drive power envelope is generally not constant because of the random spectral characteristics of the transmitted band. The degree of power envelope variation about the nominal level (i.e., the peak-to-rms ratio of the transmission band)

increases with the number of carriers. The NPR test conditions under which the LTWT was optimized create the extreme case of an infinite number of carriers occupying the band and a theoretically infinite peak-to-rms ratio. The effective NPR performance of the LTWT was therefore influenced by the shape of the transfer characteristics, not only at the nominal (or rms) 5-dB output backoff, but also in the operating range above and below it. This effect was also illustrated by computer simulation results which indicated that the NPR performance at a given drive level is significantly influenced by the nonlinearities which exist at drive levels as much as 6 dB above or below the nominal level. In the case of the LTWT, the best NPR at 5-dB output backoff was achieved by optimizing both the phase performance and the net effect of gain compression and expansion, as outlined above for the 6- to 1-dB output backoff range.

b. It would have been desirable to eliminate the gain variations observed between 6- and 2-dB output backoff. However, as shown in Figure 6, linearizer gain and phase characteristics are interrelated, and when phase is advanced by 40° or more, gain compression cannot be totally avoided. Figure 11 showed that the maximum phase advance produced by the linearizer was indeed greater than 40°, and therefore gain compression inevitably occurred.

The intermodulation performances of the TWT and LTWT at 4 GHz are given in Figure 12 and Table 3.

TWT3 PERFORMANCE

A COMSAT Laboratories RFL was optimized with TWT3 at a 2.3-dB output backoff in order to achieve the best intermodulation performance near an NPR of 13.5 dB. Figure 13 shows the measured TWT and LTWT transfer characteristics. The LTWT phase characteristics exhibited significant improvement, with a maximum variation of 16° from 17-dB input backoff to 2-dB overdrive, and with only a 3° variation within a 5-dB range about the optimized input backoff. The gain slope at this operating point is increased from 0.6-dB output variation before linearization to 0.8 dB after linearization, for 1 dB of input variation. At very low drive levels, where the linearizer is essentially linear, the TWT and LTWT curves merge as expected. Near 10-dB input backoff, the LTWT gain is compressed relative to the TWT gain for the same reasons that were enumerated in discussing the performance of TWT2.

The intermodulation performances of the TWT and LTWT at 1.5 GHz are given in Figure 14 and Table 4. It can be seen that the NPR of the LTWT below 5-dB output backoff is inferior to that measured with the two previous
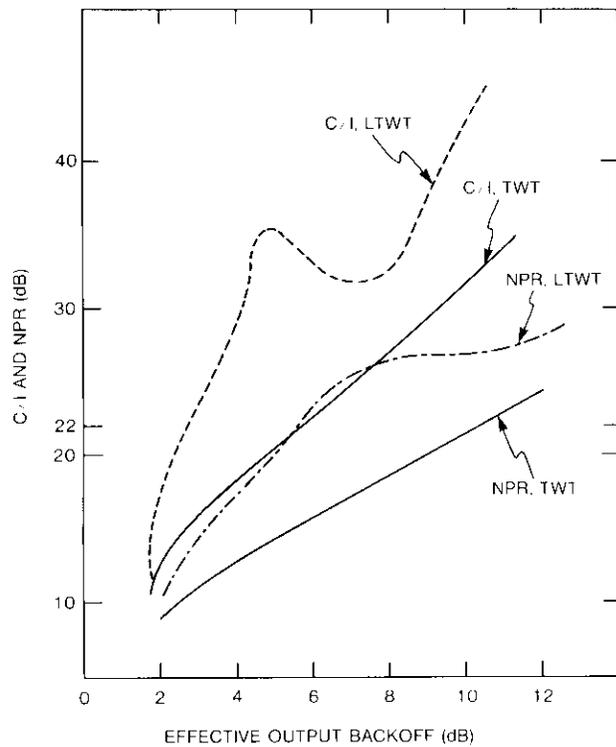
Figure 12. *Measured Two-Carrier C/I and NPR of TWT2 and LTWT2*

TABLE 3. NPR OF TWT2 BEFORE AND AFTER LINEARIZATION

NPR (dB)	OUTPUT BACKOFF (dB)		
	TWT2	LTWT2	IMPROVEMENT WITH LINEARIZATION
15	5.5	3.2	2.3
20	9.0	5.0	4.0
25	12.5	7.1	5.4

LTWTs. Since this LTWT was optimized at a distinctly higher operating point of 2.3-dB output backoff, as compared with 7 and 5 dB in the two previous cases, the NPR improvement of this LTWT at low drive levels was expected to be inferior. Furthermore, the gain and phase nonlinearities of this TWT in the backed-off mode are more severe than those found in the two previous cases.
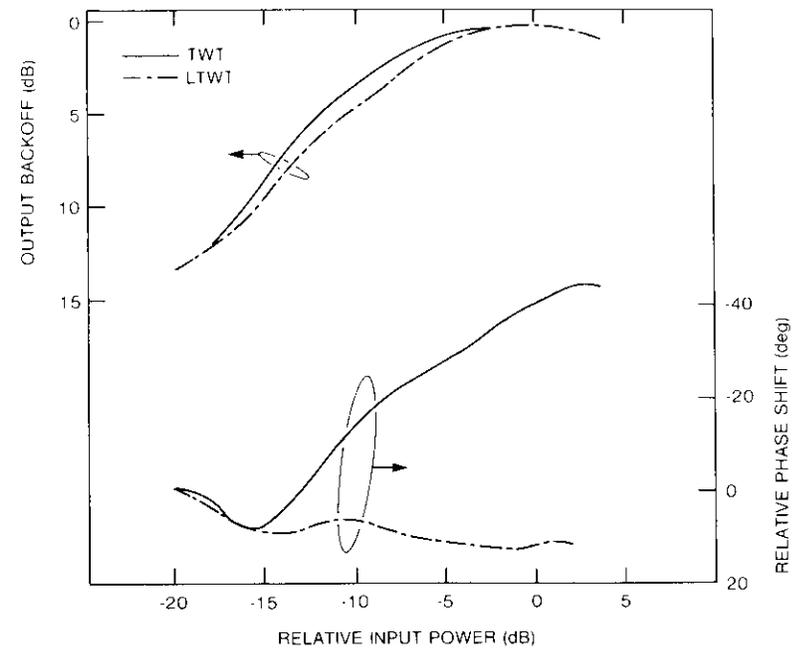
Figure 13. *TWT3 and LTWT3 Transfer Characteristics at 1.5 GHz*

TWT4 PERFORMANCE

TWT4 is the type of TWT used in INTELSAT K_u-band transponders. Figure 15 shows the measured and computer-approximated transfer characteristics of this TWT at 11.58 GHz. The measured and computed third-order, two-tone C/I and the NPR are illustrated in Figure 16. The computed values of $(C/I)_3$ and NPR are also shown, and are in very good agreement with measurements at output backoff levels above 9 dB.

A COMSAT Laboratories RFL was optimized with this TWT at a 7-dB output backoff in order to achieve the best NPR performance in the range of 20 to 26 dB. The TWT and LTWT transfer characteristics are compared in Figure 15. With linearization, the relative phase shift at saturation is reduced from $-38°$ to $-4°$, and the gain linearity is improved in the 8- to 4-dB input backoff range. A gain expansion in the output backoff of 1.2 dB per 1 dB of change in the input backoff is also noted between 16- and 10-dB input backoff.

The intermodulation performances of the TWT and LTWT are given in Figure 15 and Table 5.
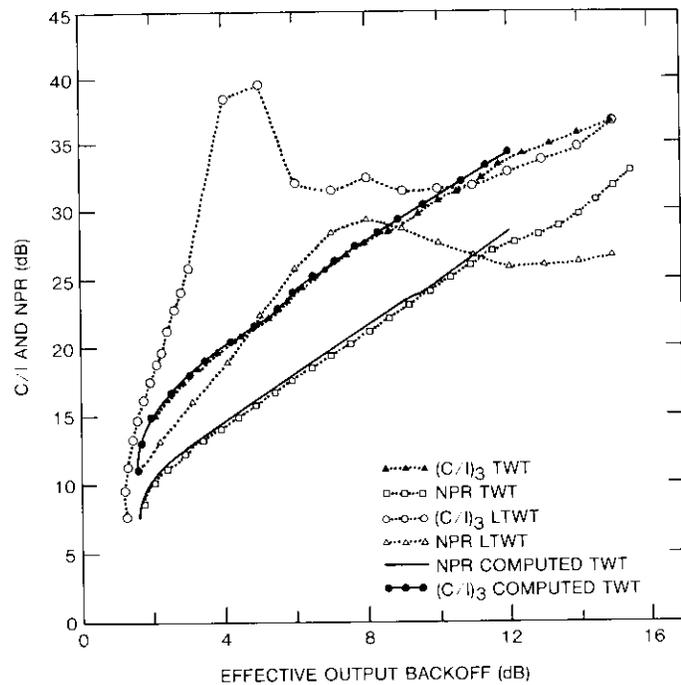
Figure 16. *Two-Tone C/I and NPR of TWT4 and LTWT4*

ments show that NPR performance near saturation is better than that of the LTWT.

LIMITER CHARACTERISTICS

For experimental evaluation of the LLTWT, a limiter was used as the driver for the linearizer. The limiter consisted of a 4-stage C-band FET amplifier which was tuned for best phase flatness throughout the drive range, including oversaturation. The transfer characteristics of the limiter measured at 3.92, 4, and 4.08 GHz are shown in Figure 17. Since the phase characteristics at 4 GHz showed the least variation over a significant drive range, linearization was performed at this frequency.

LLTWT OPTIMIZATION AND PERFORMANCE

For this exercise, TWT1 and the associated linearizer were used. Before integration with the limiter, the LTWT was tuned at 4 GHz. The limiter was then added before the linearizer, and the LLTWT was optimized for best NPR
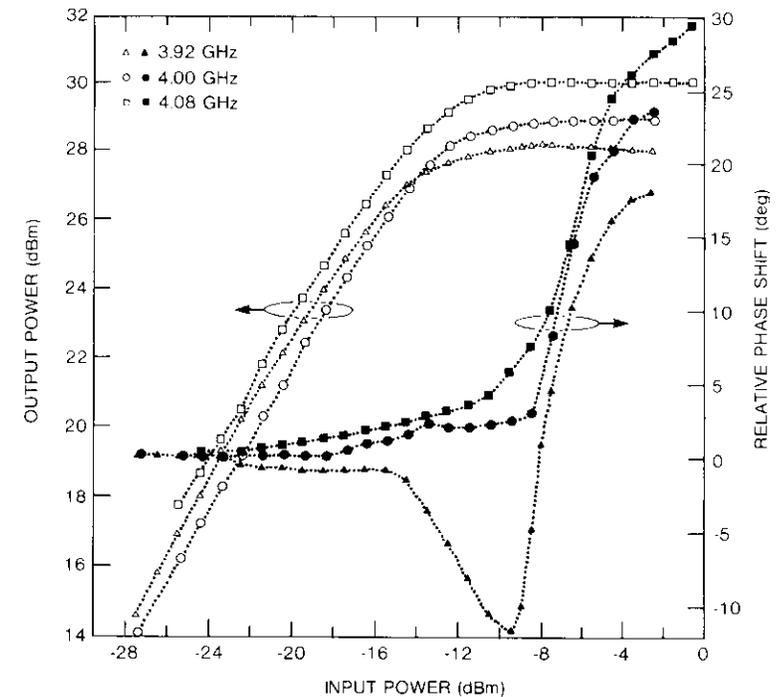
Figure 17. *4-GHz Limiter Transfer Characteristics*

near saturation by adjusting the power levels at the limiter and linearizer inputs.

Figure 18 illustrates the transfer characteristics of the LLTWT. The LTWT and LLTWT have similar gain characteristics at drive levels above 6-dB input backoff. At higher drive levels up to saturation, the LLTWT gain is more compressed than that of the LTWT because of limiter gain compression in the same range. Beyond saturation, the LLTWT output is constant and equal to the saturation value, but the phase varies. Phase behavior can be interpreted as follows: at a low drive level (below 8-dB input backoff) the linearizer and TWT phase shifts are offset within 2°, while the limiter phase is 0°. In the drive range between 8-dB input backoff and 2-dB overdrive, the TWT phase lag exceeds the phase advance of the linearizer, while the phase of the limiter is still near 0°. Consequently, the LLTWT undergoes a net negative phase shift.

At 2-dB overdrive and above, the limiter inhibits further phase variations of the TWT or the linearizer, and the phase response of the LLTWT is governed
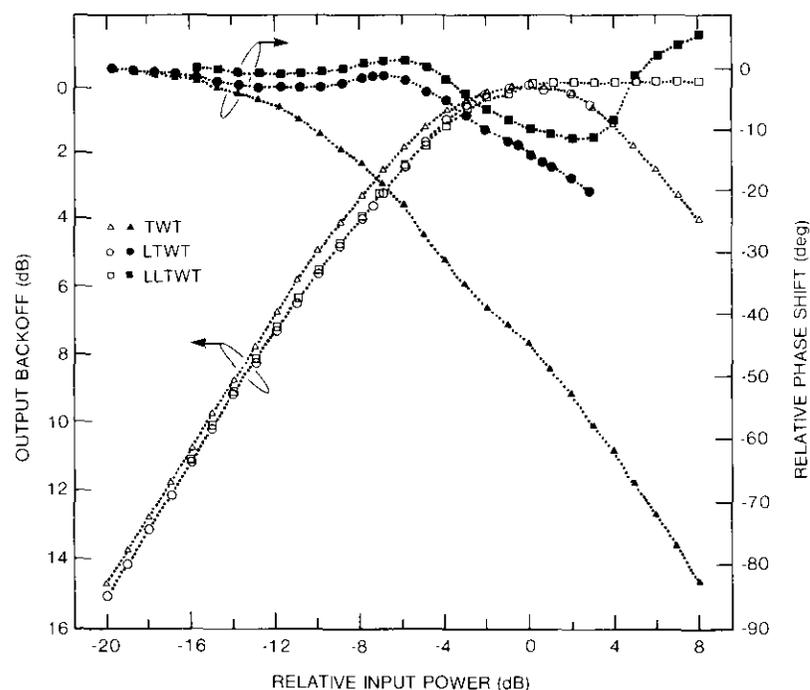
Figure 18. *Transfer Characteristics of TWT1, LTWT, and LLTWT at 4 GHz*

Figure 19. *NPR of TWT1, LTWT, LLTWT, and Ideal Limiter*

by the limiter phase advance. The relative phase shift at saturation, $\Delta\phi_{sat}$, for TWT1, LTWT1, and LLTWT1 was observed to be 42°, 16°, and 10°, respectively. Figure 19 illustrates the NPR of the TWT, LTWT, and LLTWT at 4 GHz, and of an ideal soft limiter. As expected, the LTWT and LLTWT are similar in performance at operating points below 6-dB output backoff. At higher operating points, the soft-limiting mechanism improves the NPR. For example, at NPR values of 13.5 and 11 dB, the LLTWT output backoffs are 2.4 and 1.7 dB, respectively, compared to 2.7 and 2.3 dB with the LTWT.

Table 6 summarizes the intermodulation performance of the TWT at 4 GHz before and after linearization and soft limiting.

The measured transfer characteristics and NPR indicate that the addition of the experimental limiter did indeed bring LLTWT performance closer to that of the ideal limiter. For better performance, special attention should be given to the phase flatness of the limiter beyond saturation [13].

TABLE 6. NPR PERFORMANCE OF TWT1, LTWT1, AND LLTWT1

NPR (dB)	OUTPUT BACKOFF (dB)		
	TWT1	LTWT1	LLTWT1
20	8.3	4.9	4.9
15	4.8	3.1	2.9
12	2.7	2.4	2.1
11	2.4	2.3	1.7

Linearization at an Intermediate Frequency

In a transponder with a multibeam transmitter, the TWTs can be linearized directly at the transmit frequencies or alternatively at their respective intermediate frequencies. For example, in the INTELSAT V and VI spacecraft the received signals from the hemi and zone beams at 6 GHz, and spot beams

at 14 GHz, are down-converted to 4 GHz. They are then transmitted, either at this frequency or at 11 GHz after up-conversion. This common IF configuration allows interconnection between various beams. For a 14/11- or 6/11-GHz link, the 11-GHz TWTs can be linearized at either the 4-GHz IF or the 11-GHz transmit frequency. The impact of IF linearization on system performance and reliability is discussed in Part II of this paper.

Experimental results are presented for linearization of a K_u-band TWT (TWT4) with a C-band reflection linearizer followed by a C/K_u-band up-converter (4/11-GHz LTWT). The C-band linearizer, the same as that used for TWT1, was optimized for best NPR at 11.58 GHz and 7-dB output backoff. Figure 20 illustrates the intermodulation performance of the 4/11-GHz LTWT, and, for comparison, the performance of the TWT. The intermodulation performance of the 4/11-GHz LTWT is close to that achieved with direct linearization at 11 GHz for NPR values below 23 dB. For example, the output backoffs for an NPR of 15 and 20 dB are 2.6 and 4.4 dB, respectively, compared to 2.8 and 4.3 dB with direct linearization.
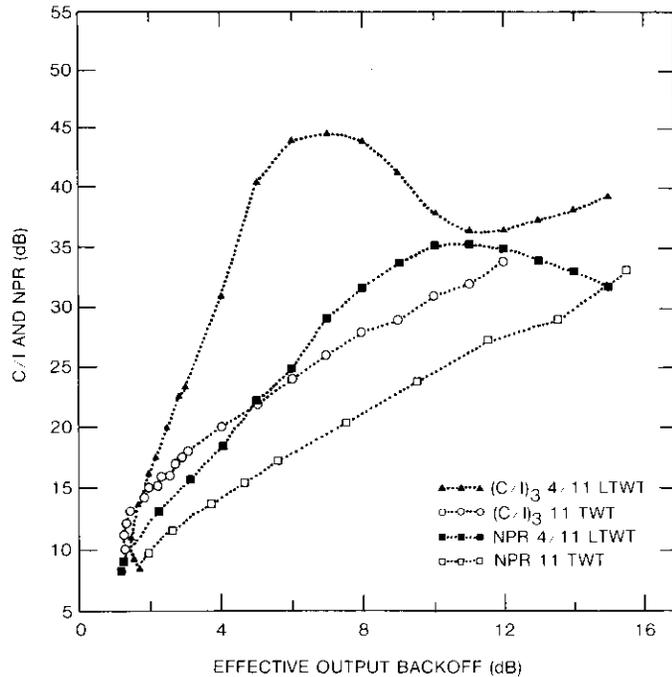
Figure 20. *Two-Tone C/I and NPR of 11-GHz TWT and 4/11-GHz LTWT*

Above an NPR of 23 dB, the 4/11-GHz LTWT exhibited better performance. In this case, the improvement is attributed to a smaller ripple in the phase response of the 4/11-GHz LTWT, compared to that of the 11-GHz LTWT. To verify this conjecture, NPR performance was assessed by a software simulation for one soft limiter with a constant phase and another with a 2.5° peak-to-peak phase ripple. The second case showed an NPR degradation of 5 dB, compared to the first case, in the NPR range near 30 dB. Figure 21 compares the amplitude characteristics of the TWT and the 4/11-GHz LTWT. As expected, the 4/11-GHz LTWT exhibits improved linearity near saturation.
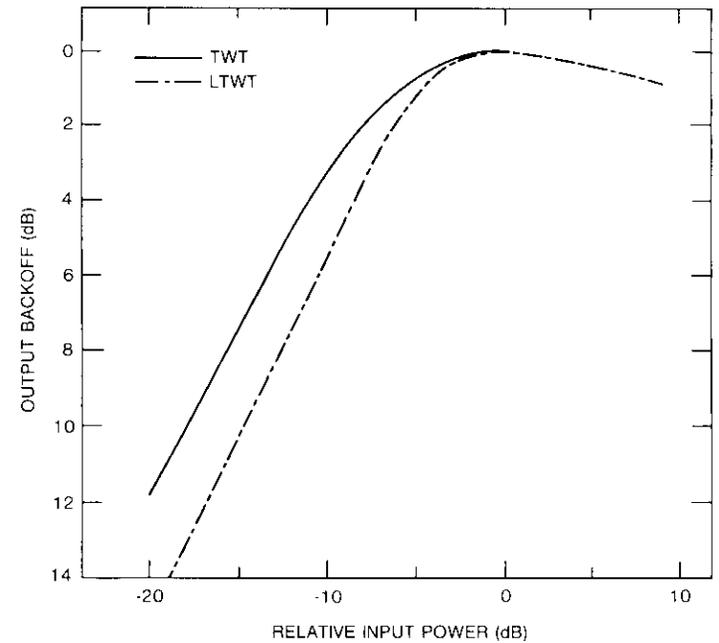
Figure 21. *Gain Characteristics of 11-GHz TWT and 4/11-GHz LTWT*

In conclusion, 4/11-GHz linearization/up-conversion is feasible. Furthermore, as shown in Figure 10, a 4-GHz linearizer has adequate bandwidth to cover two adjacent 72-MHz channels of the INTELSAT 11-GHz down-link.

Summary and Conclusions

The experiments reported here have dealt with predistortion linearization of TWTs that have a variety of gain and phase transfer characteristics. As discussed earlier, optimization of the linearizer is critically dependent on the TWT transfer characteristics. However it is important to note that, with the exception of TWT3, the improvement in output backoff at an NPR of 20 dB (a typical NPR for multicarrier transponders) was between 3.3 and 4 dB.

TWT3 was found to have noticeably different transfer characteristics (both nonlinear amplitude and phase ripple) in the backed-off mode, which may explain the 7.5-dB output backoff improvement at a 20-dB NPR. Closer to saturation, at an NPR of 15 dB, the output backoff improvement with linearization is 1 to 2.3 dB, depending on the TWT. In this region, the transfer characteristics of TWT3 are closer to those of the other TWTs, as reflected by the corresponding output backoff improvement of 1.6 dB.

Linearization experiments have also shown that the amount of improvement achieved depends on the output backoff selected for optimization. For example, optimization of TWT1 was first performed at 5-dB output backoff and yielded an NPR of 20 dB. However, the NPR dropped below 20 dB as the output level was backed off below −7 dB. New optimization performed at 7-dB output backoff yielded a 26-dB NPR at that point, and a monotonic NPR performance improvement from saturation to 15-dB output backoff. Additionally it was observed that, in the linear operating range of the LTWT, the NPR is highly sensitive to a ripple in the phase response. This observation was confirmed by a computer simulation which showed that a small phase ripple of 2.5° peak-to-peak will lower the NPR in the linear operating range (25 to 30 dB) by approximately 5 dB.

It was experimentally confirmed (TWT1 and TWT4) that sufficiently broadband operation (e.g., 160-MHz bandwidth at 4 GHz) of the LTWTs is readily achieved by using the microstrip tuning provisions of the linearizers.

The intermodulation performance results of a linearized C-band TWT preceded by a GaAs FET soft limiter were also presented. These results show an additional backoff improvement of 0.3 dB at 12-dB NPR, which increases for lower NPR (0.6 dB at 11-dB NPR). Linearization can also be successfully accomplished at an intermediate frequency band that differs from that of the TWT. As one example, it was shown that the performance of a K_u-band TWT linearized at C-band is essentially the same as that of the TWT linearized directly at K_u-band.

All of the experimental results presented apply to the nominal performance of linearized TWTs at room temperature and in the absence of TWT aging. Studies have shown that some performance degradation can be expected

during the lifetime of the LTWT and with temperature variations [14]. These effects are accounted for in Part II of this paper, which evaluates system performance.

References

[1] C. C. Hsieh and E. Strid, "An S-Band High-Power Feedback Amplifier," IEEE MTT-S International Microwave Symposium, San Diego, California, June 1977, *Digest*, pp. 182–184.

[2] P. M. Bakken, "Technical Considerations for the Linearization of High Power Amplifiers in Satellite Ground Terminals," Space Technical Center, The Hague, Technical Memorandum STC-509, 1976.

[3] A. Standing, "An Active Phase and Amplitude Correction Device for Reducing Intermodulation Produced by TWTs and Klystrons," IEE Conference on Earth Station Technology, London, October 1970.

[4] C. Bremenson and D. Lombard, "Linearisation du Canal de Transmission par Satellite pour un Systeme en Acces Multiple par Repartition dans le Temps," INTELSAT/IECE/ITE 3rd International Conference on Digital Satellite Communications, Kyoto, Japan, November 1975, *Proceedings*, pp. 144–151.

[5] G. Satoh and T. Mizuno, "Impact of a New TWTA Linearizer Upon QPSK/TDMA Transmission Performance," *IEEE Journal on Selected Areas in Communication*, Vol. SAC-1, No. 1, January 1983, pp. 39–45.

[6] M. Kumar and J. Whartenby, "TWTA Linearizer Using Dual Gate MESFET," IEEE MTT-S International Microwave Symposium, San Francisco, California, 1984, *Digest*, pp. 314–315.

[7] "Assessment of Predistortion Linearizer Technology," COMSAT Laboratories Final Report for INTELSAT Contract INTEL-317, October 1983.

[8] "Investigation of Linearized Transponders for INTELSAT VI Space-craft," COMSAT Laboratories Final Report for INTELSAT Contract INTEL-317, June 1984.

[9] Y. S. Lee, I. Brelian, and A. Atia, "Linearized Transponder Technology for Satellite Communications, Part II: System Simulation and Perform-ance Assessment," *COMSAT Technical Review*, Vol. 15, No. 2A, Fall 1985 (this issue).

[10] T. Nojima and Y. Okamoto, "Predistortion Nonlinear Compensator for Microwave SSB-AM System," International Conference on Commu-nications, Seattle, Washington, June 1980, *Conference Record*, pp. 33.2.1–33.2.6.

[11] A. Egger, M. Horn, and T. Vien, "Broadband Linearization of Microwave Power Amplifiers," 10th European Microwave Conference, Warszawa, Poland, September 1980, *Proceedings*, pp. 490–494.

[12] J. C. Fuenzalida, O. Shimbo, and W. L. Cook, "Time-Domain Analysis of Intermodulation Effects Caused by Nonlinear Amplifiers," *COMSAT Technical Review*, Vol. 3, No. 1, Spring 1973, pp. 89–143.

[13] C. Baughman and J. Chin, "GaAs FET Limiting Amplifier Designed for Low AM-to-PM Conversion," IEEE MTT-S International Micro-wave Symposium, Dallas, Texas, June 1982, *Digest*, pp. 268–270.

[14] D. Cahana, "Practical Aspects of Predistortion Diode Linearizers for Satellite Transponders," 14th European Microwave Conference, Liège, Belgium, September 1984, *Proceedings*, pp. 112–117.

David Cahana received a B.S.E.E. from the Technion-Israel Institute of Technology in 1979, and an M.S.E.E. from Rensselaer-Polytechnic Institute, New York, in 1980. Since joining the Microwave Systems Department at COMSAT Laboratories in 1980, he has worked in microwave circuit design and in satellite communications system studies. His circuit design work has included development of a 12/4-GHz dual-gate FET mixer, low-noise and low-power K-band amplifiers, and a new modeling technique for spiral microstrip inductors. Systems-related work has included the development of a transponder configuration for the second-generation INMARSAT system and a baseline for an in-traffic communications system monitor for the INMARSAT tran-sponder; the development of diode predistortion linearizers for spacecraft TWT amplifiers at L-, C-, and K_u-bands; experimental and computed evaluation of the multicarrier performance of linearized TWTs; and system performance assessments for typical INMARSAT and INTELSAT multicarrier links. Mr. Cahana has been with Mikrokim, Israel, since August 1985.

Janaki R. Potukuchi received his baccalaureate degree in mathematics from Andhra University, Waltair, India, and his B.S., M.S., and Ph.D. degrees in electrical engineering from the Madras Institute of Technology, the Indian Institute of Technology, and North Carolina State University in 1963, 1965, and 1974, respectively. From 1966 to 1969, he was a lecturer in the Department of Electronics and Communications Engineering at Osmania University, Hyderbad, India. During 1976, he was a Principal Engineer at Fairchild Industries in Germantown, Maryland, where he was responsible for the design and development of low-noise GaAs FET amplifiers.

Dr. Potukuchi joined COMSAT Laboratories in 1976 as a Member of the Technical Staff. He is presently Manager of the Microwave Systems Department, and is responsible for supervising research and development in the areas of satellite payloads, earth stations, and state-of-the-art monolithic and hybrid microwave integrated circuits.

Robert G. Marshalek received the B.E.S. and Ph.D. degrees from The Johns Hopkins University in 1976 and 1982, respectively, both in electrical engineering. From September 1976 to May 1981, he was a Research and Teaching Assistant at The Johns Hopkins University. He joined COMSAT Laboratories in October 1981 as a Member of the Technical Staff in the Optical Commu-nications Department of the Microwave Technology Division. He has been involved in investigations of wavelength-division-multiplexed optical links, the use of fiber optics in a digitally multiplexed earth station, and fiber optic local area networks. His current responsibilities include detailed studies of optical intersatellite link technology, and contributions to the design and testing of modular hardware for prototype space optical links. Dr. Marshalek is a member of IEEE, SPIE, Phi Beta Kappa, Tau Beta Pi, and Eta Kappa Nu.

Dilip K. Paul received a B.S. with Honors in electronics and telecommunications in engineering from Jadavpur University, India, in 1961; an M.S. in UHF and microwave engineering from the Indian Institute of Science (I.I.Sc.) in 1962, and a Ph.D. in electrical engineering from the University of Rhode Island in 1974. He joined COMSAT Laboratories in 1981, and is presently Manager of the Optical Communications Department where his R&D work includes integrated optics in ferroelectric and GaAs-based materials, and undersea, terrestrial, and space optical links. Prior to joining COMSAT, he held senior research/teaching faculty positions at the Indian Institute of Technology, Kanpur; I.I.Sc.; the University of Maryland; Harvard University; and IBM's Thomas J. Watson Research Center. His research spans antenna beam shaping, optical waveguides, and thin-film materials and devices.

Dr. Paul holds two patents on display devices, and has authored numerous research publications. He was awarded the Institution of Electronics and Telecommunications Engineers (IETE) prize in 1967 for his work on optical waveguides. He is cited in Who's Who in Technology Today and in Marquis International Who's Who in Optical Sciences and Engineering for contributions in the areas of thin-film technology and fiber optics. He is a Fellow of the IETE, and a member of the Society of Photo-Optical Instrumentation Engineers (SPIE), the Optical Society of America, and Sigma Xi. He is Chairman of the communications working group for Fiber-Optic Components, Communications, and Sensors (FOCCAS), SPIE.

Linearized transponder technology for satellite communications

Part II: System simulation and performance assessment*

Y. S. LEE, I. BRELIAN, AND A. ATIA

(Manuscript received November 15, 1984)

Abstract

This paper describes the simulation analysis of a diode linearizer and linearized TWTAS based upon experimental results described in the companion paper (Part I) which appears in this issue. Simulation results of channel transmission performance of various linearized TWT INTELSAT VI transponder configurations for multicarrier traffic as well as single carrier TDMA traffic are presented. Predistortion-type diode linearizers with and without soft limiters were evaluated at C- and K_u-bands to linearize TWTAS either directly or at a 4-GHz IF frequency for K_u-band channels. The performance evaluation included beginning-of-life and end-of-life channel characteristics of the linearized transponders. Based on the statistical TWT transfer characteristics, the lower bound of linearized transponder channel performance was assessed and other implementation issues of the linearized spacecraft transponders were discussed.

Introduction

Intermodulation analyses of nonlinear amplifiers for frequency division multiple access (FDMA) satellite systems have been reported by Fuenzalida

et al. [1] and Saleh [2]. The transmission performance limitation due to traveling wave tube (TWT) nonlinearity of a conventional nonregenerative satellite transponder has been well-known. The amplitude and phase nonlinearity of a high-power amplifier, that is, a TWT amplifier (TWTA), limits the useful output power level far below its maximum (saturation) capability for multicarrier transmission. The output backoff requirement of a TWTA has been determined mainly from the highest intermodulation noise allowable for various types of traffic. Increased output backoff causes reduced DC-to-RF conversion efficiency of the TWTA.

To improve the linearity of TWTAS, various techniques have been developed; namely, the feed-forward method, negative feedback method, and predistortion technique. Among these, the predistortion technique is most practical for satellite TWTS as described in Part I of this paper, providing less complexity with low mass and power requirements. Implementation of predistortion linearizers includes a number of known circuit configurations, either transmission or reflection types using microwave diodes or GaAs field-effect transistors (FETS). GaAs FET predistortion linearizers for earth station applications have been developed and field tested by various manufacturers [3].

Spacecraft TWTAS can be linearized on a per-channel basis. The bandwidth requirement of a predistortion linearizer is typically 40 to 80 MHz, which can be readily achieved with the current linearizer technology discussed in Part I [4].

This paper describes the computer simulation of the diode linearizer and linearized TWTAS (LTWTAS). Simulation analysis results are presented for various configurations of LTWTAS using a reflection type of diode linearizer both in C- and K_u-bands. The basic simulation results were verified through laboratory measurements of experimental hardware components, as reported in Part I. Predistortion linearizers with and without soft limiters were also evaluated. Their transmission performance predictions are described for multicarrier traffic as well as for single TDMA carrier traffic in the interference-limited environment. The linearized transponder characteristics at beginning of life (BOL) and end of life (EOL) described in this paper provide the basis for assessment of the power and weight savings that can be realized by using linearizers. A number of system implementation issues such as the use of a single IF linearizer for each of the channels are also discussed.

Simulation analyses of the diode linearizer and linearized TWTAs

The reflection type of diode linearizer circuit is shown in Figure 4 in Part I. A computer simulation program was developed for this type of linearizer using the diode network model shown in Figure 1 of this paper.
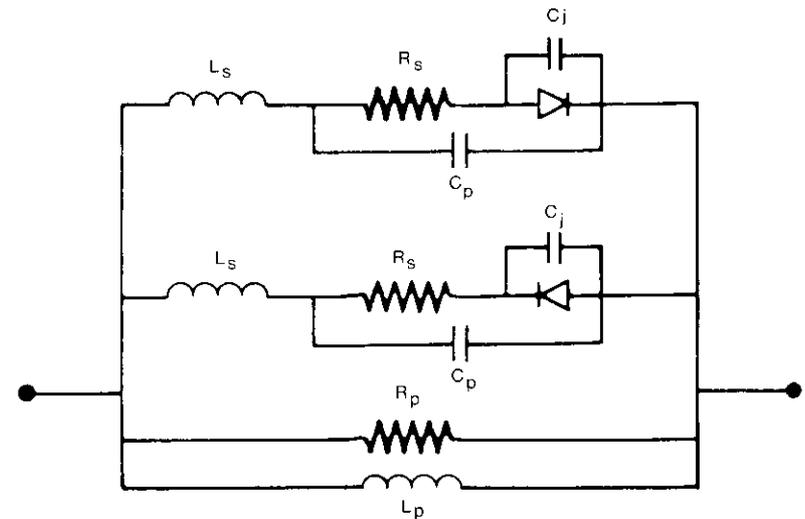
Figure 1. *Diode Network Model*

Simulation model

The impedance termination consisting of a transmission line element and Z_o (50 Ω) in the nonlinear branch was simulated with R_p and L_p (lumped circuit elements) within the narrow frequency bandwidth of interest. The program uses time domain analysis to compute the steady state waveforms of the voltages and currents existing in the network. The differential equations of the circuit are solved with zero initial conditions. The time variable is sampled and the differential operators are replaced by recurrent relations. The Fourier fundamental component of the diode voltage is computed from the initial conditions. Then the initial conditions of the next RF cycle step are taken as the results of the previous step to compute the next fundamental voltage. This iteration process is repeated until either the variation of the fundamental voltage component over two consecutive periods is less than some prescribed tolerance, ϵ, in which case the steady-state solution is obtained, or a specified maximum number of iterations has been reached. For a given set of circuit element values and a given frequency, the transfer characteristics of the linearizer are obtained by computing the output power and phase over the specified range of linearizer input power levels.

The individual transfer characteristics of a linearizer and a TWTA are used to compute the overall LTWTA transfer characteristics. The linearizer transfer

characteristics (output power and phase *vs* input power) denoted by functions f_1 and f_2 and the TWTA transfer characteristics denoted by functions g_1 and g_2 are given in Figure 2. The overall LTWTA transfer characteristics are found by interpolating the g_1 and g_2 by cubic spline functions and evaluating these functions at values corresponding to linearizer output.
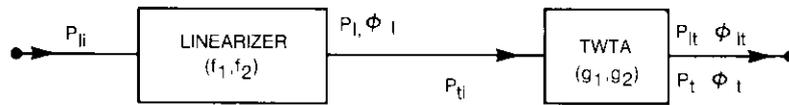
UNIT	TRANSFER CHARACTERISTICS
LINEARIZER	$P_l = f_1 (P_{li})$ $\phi_l = f_2(P_{li})$
TWTA	$P_t = g_1(P_{ti})$ $\phi_t = g_2(P_{ti})$
LTWTA	$P_{lt} = g_1(f_1(P_{li}))$ $\phi_{lt} = f_2(P_{li}) + g_2(f_1(P_{li}))$

Figure 2. *Linearizer TWTA Configuration*

Analysis results

The transfer characteristics of a C-band (3.8 GHz) simulated and measured linearizer are shown in Figure 3a. Measured data were taken as part of the experimental verification described in Part I. Because the simulation of the linearizer and external circuits, including hybrids and others, is assumed to be ideally matched to all of the harmonic components and the fundamental frequency taken as the response function, there is a difference between the simulated and measured linearizer characteristics, as shown in Figure 3a. Figure 3b shows the computed transfer characteristics of the LTWTA when this linearizer is cascaded with a C-band TWTA that has a transfer characteristic shown in Figure 9 of Part I. The curves labeled 1 in Figure 3b are the overall measured characteristics of the cascade of the linearizer and the TWTA. The curves labeled 2 are the computed characteristics of the cascade using the simulated transfer characteristics of the linearizer and the measured transfer characteristics of the TWTA. In the curves labeled 3, the measured linearizer

(a) 3.8-GHz Linearizer

(b) 3.8-GHz LTWTA
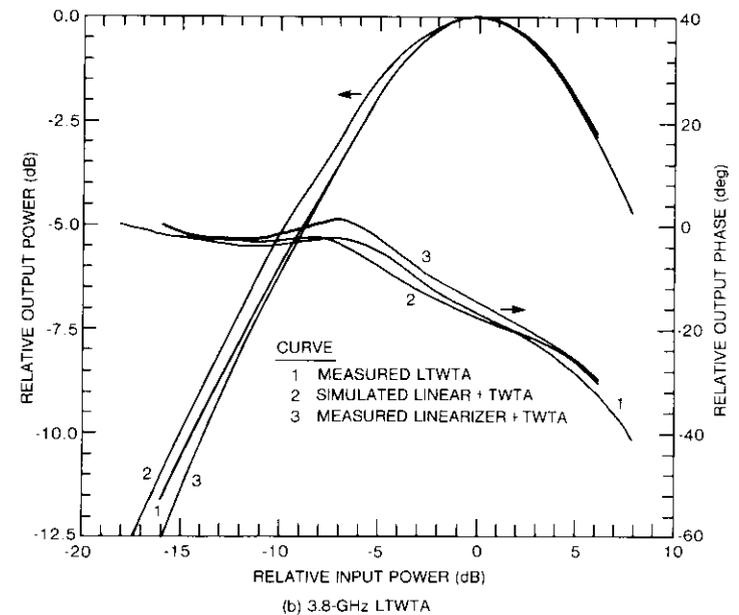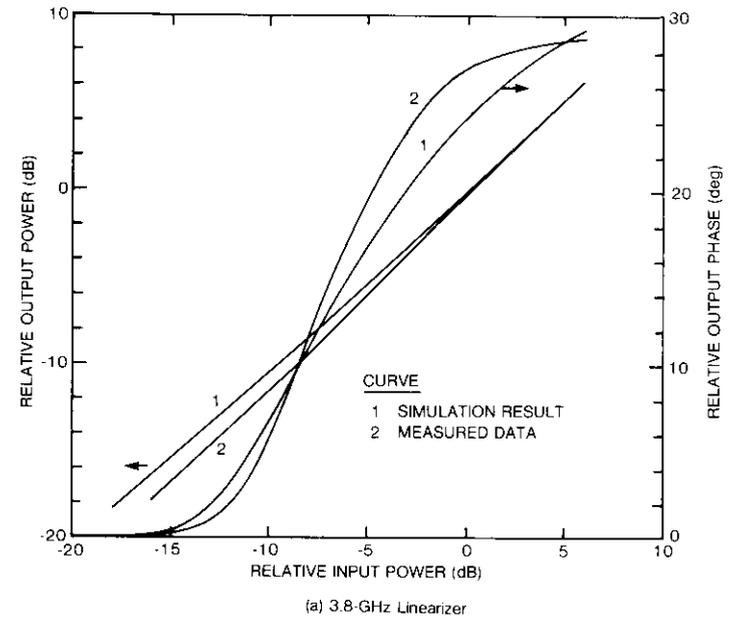
Figure 3. *Comparison of Simulated and Measured Transfer Characteristics*

characteristics are used together with the measured TWTA characteristics to compute the overall transfer characteristics.

The computed or measured overall transfer characteristics defined in Part I were used for computations of intermodulation products and noise power ratio (NPR) by the method described in Reference 1 for the evaluation of multicarrier transmission performance. The computed NPR *vs* output backoff (OBO) is shown in Figure 4. Curve 1 in Figure 4 shows the computed NPR using directly measured transfer characteristics of the linearizer-TWTA combination. Curves 2 and 3 are the computed and measured NPR *vs* OBO for the TWTA alone and for the LTWTA, respectively. Curve 4 is the computed NPR *vs* OBO for the simulated linearizer cascaded with the measured TWTA.
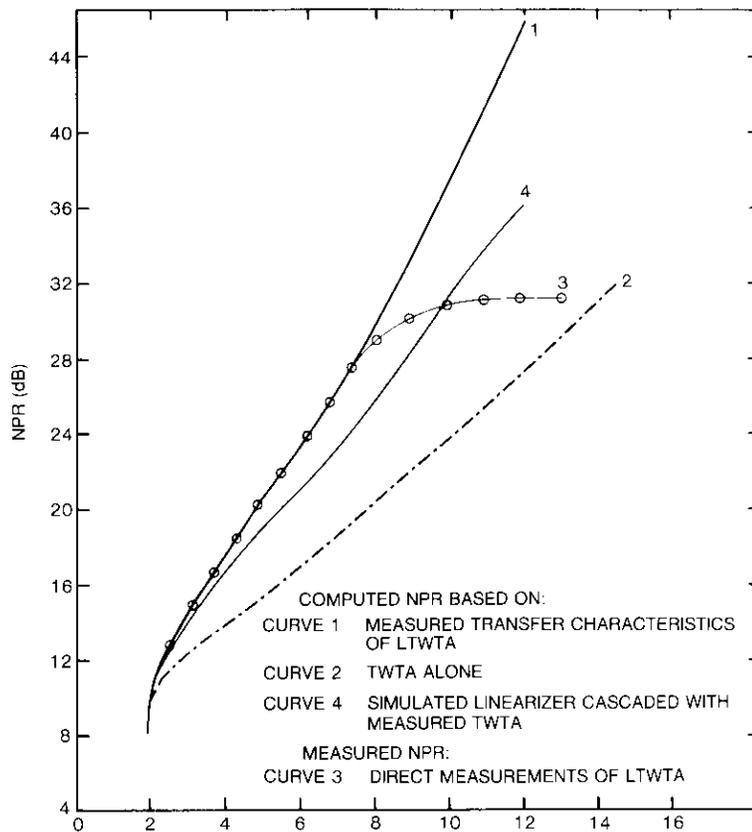
Figure 4. *Computed and Measured NPR Characteristics of C-Band 3.8-GHz LTWTA*

Comparing curves 1 and 3 shows that the computed and measured NPRs of the LTWTA are in good agreement up to NPR values of 27 dB, which corresponds to an OBO of 7.3 dB. Comparing curves 3 and 4, a noticeable difference is observed (0.7 dB OBO at an NPR of 22 dB) between the direct, measured NPR of the LTWTA and the computed NPR of the simulated linearizer cascaded with measured TWTA. This difference is due to the nonidentical transfer characteristics of the simulated and measured linearizer, as shown in Figure 3a. The results of the simulation analysis given in this section have been used as guidelines in the development of the linearizer network.

Alternative approach

An alternative approach based on frequency domain analysis of the diode linearizer was formulated as part of the simulation study. This approach (summarized in Appendix A) uses a general method to compute the steady-state transfer characteristics of any network containing a number of memoryless nonlinear elements, linear, lumped, or distributed elements, without explicitly performing time domain analysis. The general method of extracting the nonlinearities into external ports of a linear multiport network is applied to obtain the overall response. Specifically, the linearizer is treated as a four-port network composed of an input port, an output port, and two other ports to which the nonlinear diodes are connected. The impedance matrix of the linear part of the linearizer circuit is computed and used with the nonlinear model of the diodes to compute all components (fundamental and harmonics) of the voltage across the diodes as well as the circuit transfer characteristics. This approach merits further investigation.

Linearized transponder configurations

The transponders can be linearized in a number of ways. In this paper, the following two configurations are discussed:

- *Configuration I.* Linearizer matched to each TWTA, as shown in Figure 5.
- *Configuration II.* Linearizer at 4-GHz IF only, as shown in Figure 6.

Assessment of configurations

Configuration I requires both C-band and K_u-band linearizers and a 160-MHz bandwidth K_u-band linearizer for the INTELSAT VI spacecraft [5] channel (9–12). On the other hand, Configuration II needs only 4-GHz IF linearizers for all TWT transponder channels, including K_u-band spot beam
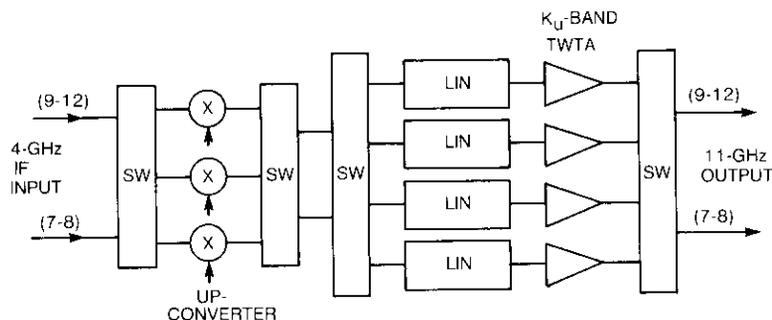
Figure 5. *Configuration I: Linearizer Matched to Each TWTA*
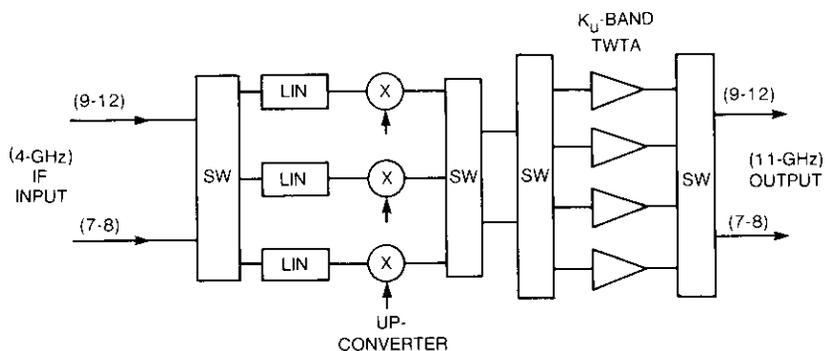
Figure 6. *Configuration II: 4-GHz IF Channel Linearizer for K_u-Band TWTAs*

channels. Linearizers may be inserted before the redundancy switch in Figure 6. In this case, the linearizer bandwidth requirement is determined by each channel bandwidth, *i.e.*, 80 MHz for C-band and 160 MHz for K_u-band channels. In Configuration II shown in Figure 6, linearizers after the redundancy switch provide acceptable levels of redundancy (3-for-2) for the linearizer and up-converter combination. However, the 4-GHz IF linearizer design must function over a larger tolerance range due to parameter variations in redundant TWTAS, up-converters for K_u-band channels, and redundancy switches.

The feasibility of Configuration II was assessed through comprehensive evaluation of parameter variations in TWTAS, as described in Appendix B, and was experimentally verified.

Figure 7 shows computed NPRs for each configuration of various combinations of limiter, linearizer, and TWTA. The ideal limiter and measured LTWTA configuration (curve 4) show about 0.2 dB of improvement in OBO as compared to the actual measured limiter and measured LTWTA configuration (curve 1) up to an NPR of 22 dB. Comparing curves 2 and 3 for another configuration, it should be noted that substitution of the actual measured limiter with the ideal limiter improves the OBO by about 0.6 dB at an NPR of 22 dB.

Various combinations of a linearizer and a soft limiter were also evaluated experimentally. Table 1 presents a summary of the results for an effective
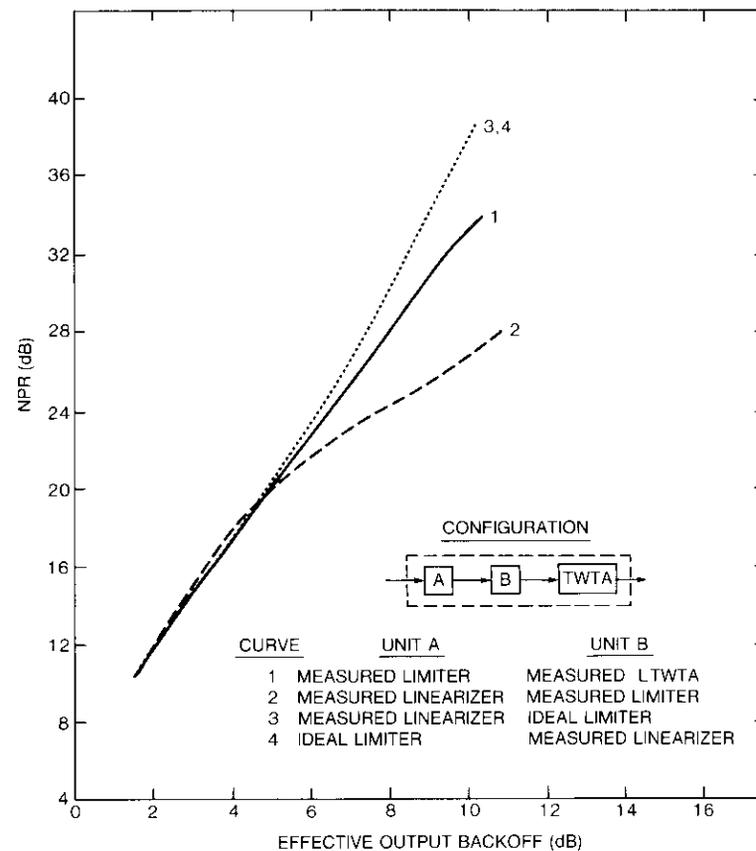
Figure 7. *Computed NPR Characteristics of Various Cascaded Configurations of Limiter, Linearizer, and TWTA at 4 GHz*

OBO requirement to meet the prescribed NPR for a number of configurations at 4 GHz. A GaAs FET amplifier is the soft-limiting linearizer used. When the limiter is ideally envelope limiting without AM/PM conversion, the order of the linearizer and limiter yields identical subsystem performance, providing the upper bound of soft limiting LTWTA performance, as shown in the last column of Table 1 (also curves 3 or 4 in Figure 7).

TABLE 1. NPR VS OUTPUT BACKOFF (dB) FOR VARIOUS COMBINATIONS (AT 4 GHz)

NPR (dB)	TWTA ONLY	SOFT-LIMITING LINEARIZER			
		LINEARIZER PLUS TWTA	LIMITER PLUS LINEARIZER PLUS TWTA	LINEARIZER PLUS LIMITER PLUS TWTA	LINEARIZER PLUS IDEAL LIMITER PLUS TWTA
26	11.7	7.0	7.4	9.5	6.8
20	8.3	4.9	5.0	5.1	4.9
16	5.4	3.4	3.5	3.3	3.4
12	2.8	2.3	2.1	2.0	2.0

For frequency division multiple access (FDMA) traffic, an NPR of 20 dB is considered typical for link performance. Table 1 shows that a combination of a linearizer plus TWTA is the preferred configuration, providing a 3.4 dB higher useful output power than the case of the TWTA only. The combination of linearizer plus TWTA with ideal limiter provides almost identical NPR performance as the linearizer plus TWTA combination, except for a small improvement in low values of the NPR (below 12 dB) region that is unacceptable for FDMA anyway. For this reason, a linearizer plus TWTA combination is preferred for FDMA traffic.

The insertion of linearizing elements in the transponder channel changes the total small signal gain of the channel. To maintain zero (dB) insertion gain for the existing transponder channel, a simple arrangement for setting gain must be included. The range of gain setting should be within 5 dB for any of the above configurations. Actual gain adjustment for each channel can be controlled within 0.5-dB-to-1-dB incremental steps, depending on the operational requirements.

Impact of linearized transponder performance for INTELSAT VI spacecraft

Linearized INTELSAT VI transponders can improve transmission performance of FDMA traffic in one of two ways. In the first instance, for a prescribed

minimum value of NPR for the FDMA channel, the LTWTA can be operated at a reduced OBO, providing a higher useful RF output. Table 2 lists typical BOL and 10-year EOL OBOs required for 20 dB of NPR. The nominal OBOs of TWTA and LTWTA in C- and K_u-bands are based on measured data, and the tolerances for BOL and EOL 10-year performance are derived from the results of simulation analyses, as described by Cahana [6]. The BOL tolerance includes the effects of frequency band, temperature, and variance due to TWTA parameter spread which is applicable only for the case of 4-GHz IF channel linearizer implementation. An IF channel linearizer must operate with any combination of the redundant K_u-band TWTAs and up-converters.

TABLE 2. BOL/EOL OUTPUT BACKOFF REQUIRED FOR A NPR OF 20 dB

PARAMETERS	C-BAND		K_u-BAND	
	TWTA	LTWTA	TWTA	LTWTA
Output Backoff (dB)				
Nominal	8.3	4.9	7.6	4.3
Frequency Band		0.1		0.1
Temperature		0.1		0.1
TWTA Variance		0.4*		0.5**
BOL Tolerance		0.6		0.7
10-Year Aging		0.5		0.5
Improvement in Output Backoff (dB)				
BOL		2.8*		2.6**
		(3.2)		(3.1)
EOL (10 yr)		2.3*		2.1**
		(2.7)		(2.6)

* Worst case for a fixed 4-GHz linearizer for any of the C-band TWTA channels.
** Worst case for IF (4-GHz) channel implementation.

The BOL improvements in OBO are 3.2 dB and 3.1 dB for C-band and K_u-band, respectively, in the case of dedicated linearizers to each TWTA. A 10-year aging reduces the improvement margin by 0.5 dB and provides EOL improvement in OBO by 2.7 dB and 2.6 dB for C-band and K_u-band, respectively. However, 4-GHz IF channel linearizers for Configuration II, because of redundant TWTA parameter spread, could decrease the output margin by 0.2 dB and 0.5 dB in the worst-case K_u-band channels for an NPR of 16 dB and 20 dB, respectively. Appendix B presents the statistical parameter variations and NPR performance prediction of C- and K_u-band TWTAS.

In the second instance, if the LTWTA operates at the same OBO as the transponder with only a TWTA, the linearized transponder provides improved NPR performance. Table 3 shows that BOL/EOL NPR 10-year improvement at fixed values of OBO for any case exceeds 5 dB for C-band and 2.5 dB for K_u-band. The NPR variation due to K_u-band redundant TWTA parameter spread is larger than the C-band case. This indicates that linearizing K_u-band transponders using 4-GHz IF channel linearizers requires a tighter TWTA specification for amplitude and phase transfer characteristics to maintain reduced NPR variations among the redundant K_u-band TWTAS.

TABLE 3. BOL/EOL NPR PERFORMANCE AT CONSTANT OUTPUT BACKOFFS

PARAMETERS	C-BAND		K_u-BAND	
	TWTA	LTWTA	TWTA	LTWTA
NPR (dB)				
Output Backoff (dB)	8.3	8.3	7.6	7.6
Nominal NPR	20.0	29.5	20.0	28.0
Frequency Band		0.3		0.5
Temperature		0.5		0.5
TWTA Variance		2.2*		3.0**
BOL Tolerance		3.0		4.0
10-Year Aging		1.5		1.5
Improvement in NPR (dB)				
BOL		6.5*		4.0**
		(8.7)		(7.0)
EOL (10 yr)		5.0*		2.5**
		(7.2)		(5.5)

* Worst case for a fixed 4-GHz linearizer for any of the C-band TWTA channels.
** Worst case for 4-GHz IF channel implementation.

TDMA performance in a linearized transponder

The transmission performance of a band-limited 120-Mbit/s quadrature phase-shift keying (QPSK) signal was evaluated for a number of different INTELSAT VI transponder configurations through computer simulation analyses of the transponder channel(s) using a channel modeling program [7], [8]. The transponder channels, including linearized TWTAs with and without soft limiting, were simulated to compute bit error rate (BER) vs energy per bit to noise power density ratio (E_b/N_o) for the TDMA carrier. The computer simulation results provide the technical basis for a comparative evaluation of

linearized transponders vs conventional transponders to assess predicted link performance improvement and the associated costs for optimum transponder configuration in the INTELSAT TDMA system.

The simplified simulation model of an INTELSAT VI transponder channel, including the nonlinear subsystem, is shown in Figure 8. The ideal 120-Mbit/s QPSK modulator is simulated with a random bit sequence PSK generator followed by a Nyquist cosine rolloff filter which includes an $x/\sin x$ amplitude compensation for the rectangular waveform bit stream. The transponder channel includes INTELSAT VI input- and output-multiplexer channel filters and a nonlinear subsystem. Measured filter performance data (both amplitude and group delay responses) of typical input and output multiplexers were used in the channel filter simulation.
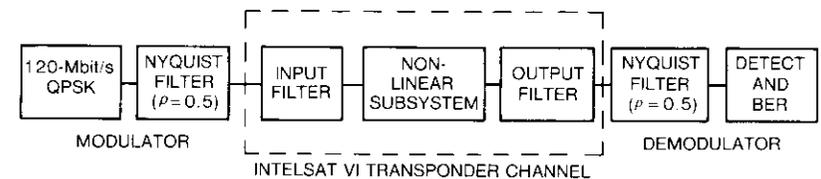
Figure 8. *Simulation Model of an INTELSAT VI Transponder Channel With Nonlinear Subsystem*

The nonlinear subsystem was simulated for each of the following cases:

a. TWTA only,
b. linearizer plus TWTA, and
c. linearizer with ideal limiter plus TWTA.

The nonlinearity characteristic of each subsystem was represented by measured or computed overall transfer characteristics (i.e., relative output amplitude and phase vs input drive level) under a memoryless device assumption.

An ideal QPSK demodulator consists of a Nyquist cosine rolloff filter plus $x/\sin x$ compensator, a midbit sampler, and BER calculator. The BER module assumes additive white Gaussian noise in the theoretical BER calculation.

As shown by curve 1 in Figure 9, the BER performance of a linear Nyquist channel provides a theoretical limit which corresponds to the ideal modem back-to-back performance. Figure 9 also shows BER vs E_b/N_o performance for each of the three transponder configurations when TWTAs are driven to saturation. A soft-limiting linearized TWTA transponder provides about 0.9 dB improvement in E_b/N_o at a BER of 1×10^{-6}, compared to the other two cases: a conventional TWTA transponder and a transponder with an LTWTA without soft limiting.
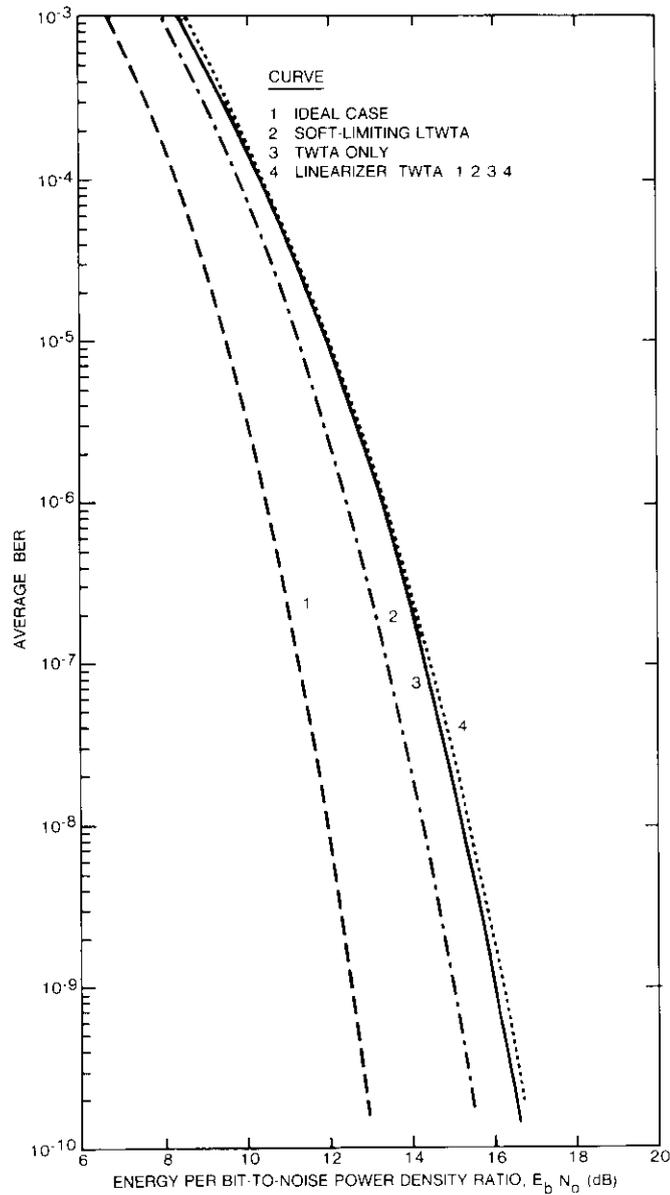
Figure 9. *BER Performance of INTELSAT VI Transponders for 120-Mbit/s QPSK Transmission*

The BER performance *vs* input backoff of the three transponder configurations, including the INTELSAT VI input and output filters, is compared in Figure 10 and shows E_b/N_o degradation from the ideal linear channel
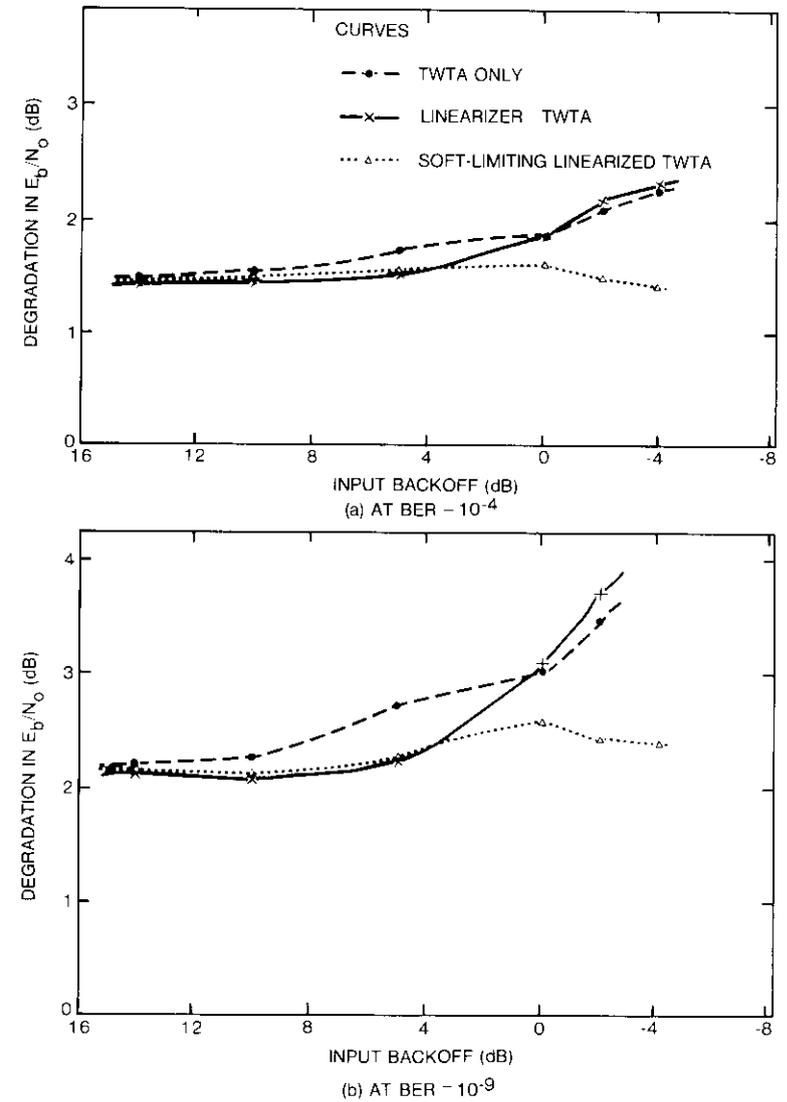
Figure 10. *BER Degradation in INTELSAT VI Transponder Channel (3-4) With Nonlinear Subsystems*

performance when the BER equals 10^{-4} and 10^{-9}. This figure indicates that the soft-limiting LTWTA transponder improves the channel performance when the transponder is driven near to or beyond saturation. However, the LTWTA transponder without soft limiting degrades the channel performance more than the conventional transponder when driven beyond saturation.

The TDMA signal transmission performance was also computed in a Nyquist filter channel which included the same nonlinear subsystems without the input and output filters shown in Figure 8. The results of this computation presented in Figure 11 indicate BER performance characteristics similar to those shown in Figure 10a, within a 0.3-dB margin in E_b/N_o in INTELSAT VI transponder channels.

These analyses provide an assessment of an LTWTA channel performance with reference to a conventional linear channel. Because the Nyquist channel filters are not necessarily optimum in the nonlinear channel, the optimum filter design for the heavily nonlinear repeater and the corresponding performance improvement of a linearizer should be further investigated in the future.

The adjacent channel interference (ACI) caused by the spectral sidelobe regrowth of the TDMA signal of limited bandwidth in a conventional TWTA transponder was evaluated in the simulation analysis using three contiguous TDMA channels. For INTELSAT VI spacecraft, the ACI effect at the transponder output multiplexer is negligible at saturation, typically less than 0.02 dB of degradation in E_b/N_o at a BER equal to 1×10^{-9}. The transponder output multiplexer prevents any BER degradation associated with spectral spreading. However, the spectral spreading effect can be significant for the earth station high-power amplifiers (HPAs) with no output filters [9].

Available experimental data on QPSK/TDMA transmission measurements through a soft-limiting linearized (FET linearizer and TWTA combination) transponder show that the overall link improvement of the soft-limiting TWTA transponder over the TWTA alone (both operating at saturation) is about 0.9 dB in E_b/N_o at a BER of 10^{-6}, with up- and down-link interferences of 19.9 dB and 20.2 dB, respectively, and a linear up-link HPA [9]. The simulation result is supported reasonably well by the measurements described in Satoh and Mizuno [9]. The soft-limiting LTWTA transponder provides overall link E_b/N_o improvement by about 1 dB at a BER equal to 10^{-6} when the TWTA is driven to or beyond saturation. When the transponder is overdriven (beyond -4 dB input backoff, for example), a soft-limiting LTWTA provides increased link margin by more than 1 dB compared to the conventional transponder for the INTELSAT TDMA transmission.

The current INTELSAT TDMA system employs forward error correction codes (FECs), i.e., (128, 112) BCH block coding for the conventional TWTA
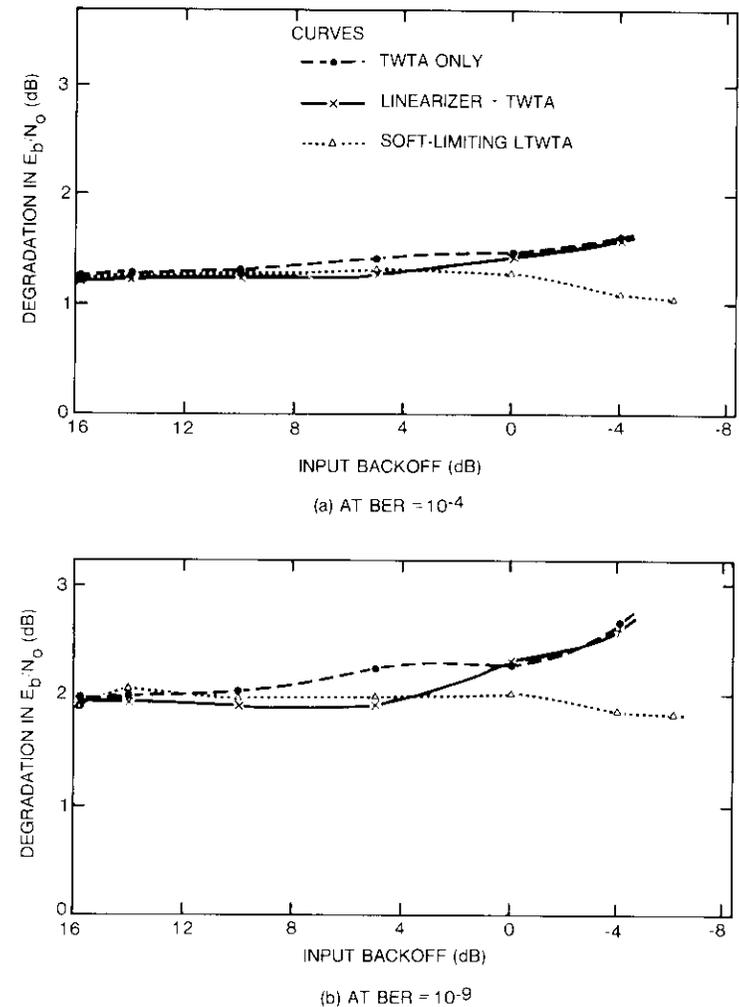
Figure 11. *BER Degradation in Nyquist Channel With Nonlinear Subsystems*

transponder channel. Table 4 lists the FEC coding gain, which shows about 3 dB gain in E_b/N_o for a BER equaling 10^{-6} of the INTELSAT QPSK/TDMA transmission [10]. If the soft-limiting LTWTA transponder could provide at least the same 3 dB or greater link margin improvement, then simply by avoiding FEC the traffic capacity could be increased by one-seventh of the current capacity that is supported by the conventional transponder. However,

the link margin improvement is limited to only about 1 dB with a soft-limiting LTWTA transponder.

TABLE 4. FEC FOR 120-Mbit/s QPSK TDMA INTELSAT SYSTEM*

BER	E_b/N_o (dB)	CODING GAIN (dB)
10^{-4}	6.47	2.00
10^{-5}	7.00	2.60
10^{-6}	7.64	2.86

* (128, 112) BCH block coding.

Implementation considerations

A comparative assessment of linearity performance of a solid-state power amplifier (SSPA), TWTA, and LTWTA is discussed in this section.

Figure 12 shows the single-tone output phase shift vs input backoff of C-band and K_u-band LTWTAS in comparison to the corresponding INTELSAT VI TWTA and SSPA requirements. The output phase linearity of the LTWTA is the same as that of SSPA for input backoffs not exceeding 6 dB. However, the LTWTA shows a larger phase shift characteristic than SSPA for large input
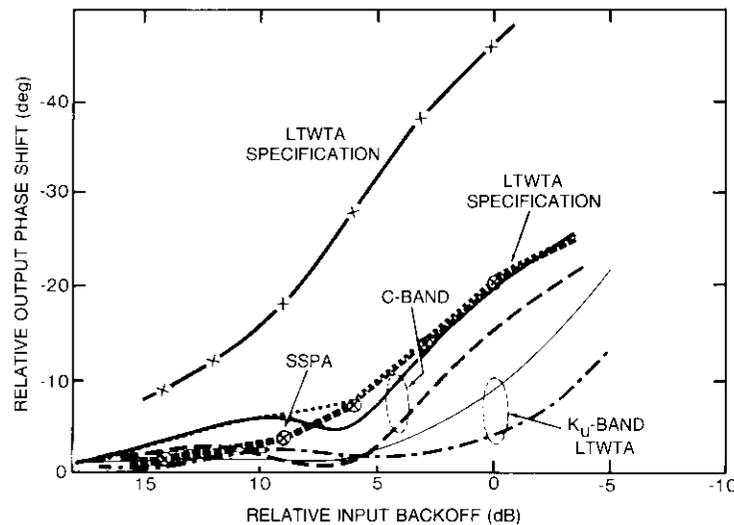
Figure 12. *Output Phase Shift vs Input Backoff*

backoffs exceeding 6 dB. The corresponding maximum allowable AM/PM transfer coefficient is shown in Table 5, based on measurements conducted with two carriers having an amplitude difference of up to 20 dB, with the larger carrier amplitude modulated at 1 dB. The AM/PM transfer coefficient of an LTWTA is larger than that of an SSPA due to ripple components observed in the LTWTA output phase vs input drive characteristic.

TABLE 5. AM/PM TRANSFER COEFFICIENT

INPUT BACKOFF* (dB)	AM/PM TRANSFER COEFFICIENT (deg/dB)		
	TWTA	LTWTA	SSPA
Overdrive 3	–	8	–
0	8	5	2
3	9	4	2
6	9	3	2
9	8	3	2
12	5	2	1
14	3	1.5	1
Greater than 14	3	1.5	1

* Relative input drive referenced to single-carrier saturation.

Table 6 lists NPR vs input backoff performance characteristics. The NPR of the LTWTA channel shows improvement approaching that of an SSPA in Table 6. Figure 13 shows measured NPR vs effective OBO that can be used for link design application for C-band and K_u-band LTWTAS.

TABLE 6. NOISE-POWER RATIO

INPUT BACKOFF (dB)	NOISE-POWER RATIO (dB)		
	TWTA	LTWTA	SSPA
Overdrive 3	5	4.5	
0	8	8	9
3	9	10	11
6	11	13	
(7)			15
9		14	19
(10)			21
12	17	25	
(13)			27
14		19	
>14	>19		
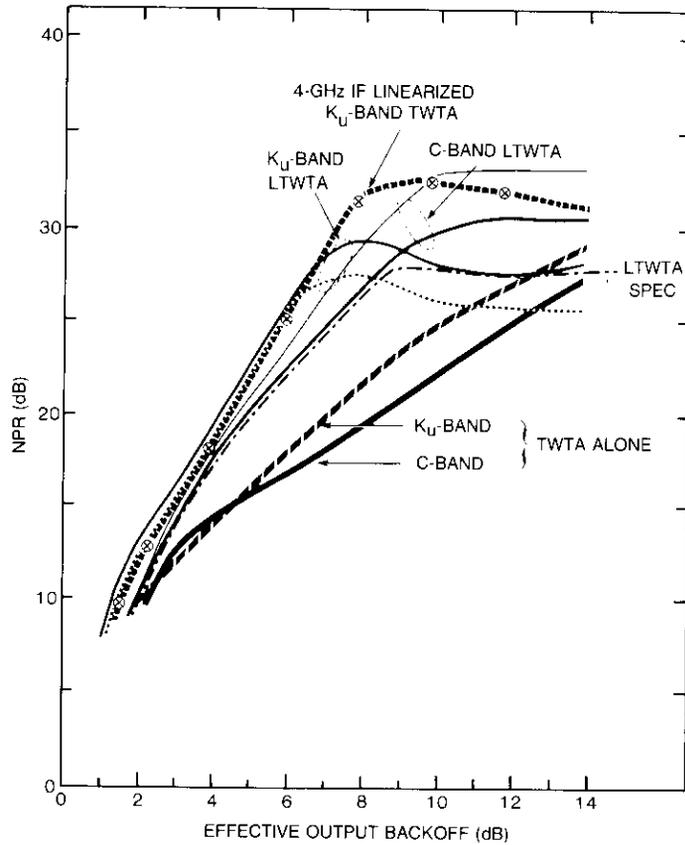
Figure 13. *NPR vs Effective Output Backoff*

For typical INTELSAT FDMA traffic for hemi-to-hemi coverage, the overall link calculation of Figure 14 shows that an LTWTA transponder can provide a linear output power increase of 2.2 dB for total system $C/(N+I)$ of 13 dB.

There are several implementation advantages of an LTWTA transponder. First, the saturated output power rating of the TWTA is reduced. An OBO improvement of 2.2 dB, for example, allows a TWTA output reduction of about 40 percent. Second, the prime power requirement is reduced, providing increased DC-to-RF conversion efficiency for an on-board HPA. A linearizer, including a preamplifier, requires about 2 W per unit and weighs 0.8 kg. When a 10-W TWTA channel is linearized, the calculations shown in Table 7 indicate that the prime power saving is about 2 W for a total system
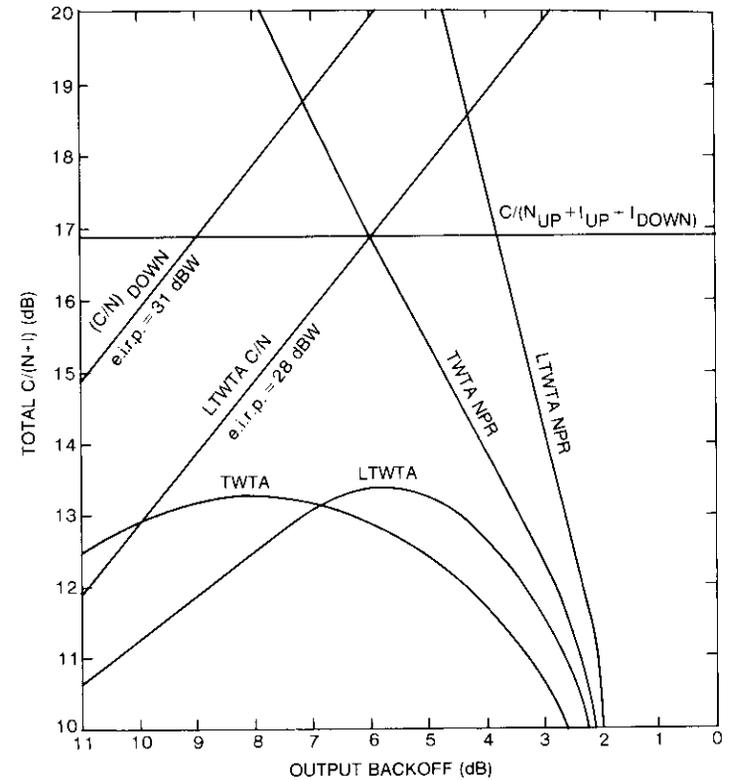
Figure 14. *Total System $C/(N + I)$ vs Output Backoff for Measured TWTA and LTWTA (Occupied bandwidth = 64.8 MHz)*

TABLE 7. TWTA LINEARIZATION COMPARISONS

PARAMETER	TWTA	LINEARIZED TWTA
RF Output Power at Saturation (W)	10.0	6.0
Output Backoff (dB)	6.4	4.2
Linear RF Output (W)	2.3	2.3
TWT Efficiency (%)		
At Saturation	50.0	50.0
At Operation	18.6	27.5
Prime Power (W)	12.3	8.3
Linearizer Power (W)	0.0	2.0
Total Power Requirement (W)	12.3	10.3
Power Saving (W)	0.0	2.0

$C/(N+I)$ of 13 dB (see also Figure 14). Third, the linearization reduces the mass of the HPA subsystem because the prime power requirement is less. And finally, using linearizer technology with an existing TWTA achieves a high level of linear output power. This avoids the need for new TWT development for some applications.

The BER performance advantage of a soft-limiting LTWTA for a 120-Mbit/s QPSK signal transmission over a 72-MHz channel of INTELSAT VI transponder is shown in Figure 15. For each case, input backoff conditions are noted in parentheses and typical BER specifications for channels employing conventional TWTAS and SSPAS are also indicated.

The gain slope of the driver and linearizer does not exceed 0.006 dB/MHz over a 160-MHz bandwidth. Experimental data showed that the gain flatness of a linearized TWTA is also acceptable.

Conclusions

Comprehensive computer-aided simulation analyses of the linearizers and the overall INTELSAT VI transponder transmission channels were performed and verified through experimental evaluations. The simulation analyses provided necessary performance bounds that are difficult to measure directly, such as linearized transponder performance variations due to the aging effect of TWTAS, the tube-to-tube parameter spread among redundant TWTAS in the transponder, and the effect of temperature.

The performance improvement of LTWTAS was assessed for various types of INTELSAT traffic. For FDMA carriers, NPR was used as the main transmission performance criterion. For a 120-Mbit/s QPSK/TDMA carrier, BER was computed for the transponder channels operating under various input drive conditions.

Two different implementation approaches to the linearized transponders were used for measurements and computer simulation: the linearizer matched to each TWTA, and the 4-GHz IF channel linearizers for C-band as well as for K_u-band TWTAS. For each configuration, various combinations of linearizers, limiters, and TWTAS were considered. The performance limitation of the soft-limiting LTWTA was assessed using an ideal envelope limiter in the channel simulation model. The experimental and simulation results presented in this paper indicate the following:

- A linearizer plus TWTA combination without soft limiting is preferred to other combinations for FDMA traffic (quantitative estimates of improvements are discussed in the following paragraphs).
- For TDMA carriers, a linearizer plus soft-limiter plus TWTA combination
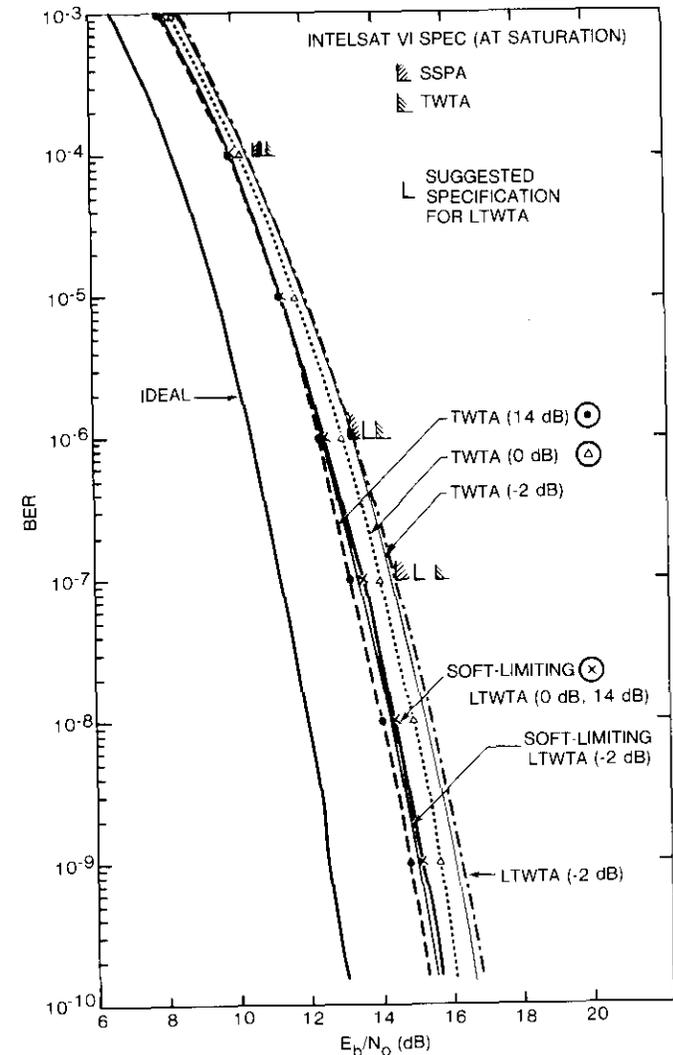
Figure 15. *120-Mbit/s QPSK Signal Transmission Performance of INTELSAT VI Transponders*

improves BER performance by almost 1 dB in E_b/N_o for the TWTA operating near saturation in the interference-limited environment.
- Linearizing all transponders at a 4-GHz IF for each channel is feasible. About a 0.5-dB margin in OBO at an NPR of 20 dB is needed because

of redundant TWTA parameter variation. For a K_u-band transmit channel, the IF channel linearizer approach is a viable option if TWTAs can be selected for smaller variance of transfer parameter spread.

The incorporation of linearizers on board INTELSAT VI provides quantitative performance improvement in four ways. First, in an FDMA network operating at a prescribed NPR, the LTWTA can be used at a reduced OBO, providing higher useful output power by 3.2 dB at BOL and 2.8 dB at a 10-year EOL. This improvement can be used to increase the overall system capacity for small carriers such as SCPCs if they are not already bandwidth limited.

Second, when the linearized transponder is used at the same OBO as the TWTA-only case, the NPR performance is improved from 3 to 5 dB, yielding improvement in transmission quality. For large carriers such as FM/TV and FDM/FM, the transmission quality may be the major concern, because of the bandwidth-limited transmission capacity situation.

Third, in the case of TDMA (120-Mbit/s QPSK), a linearized transponder improves the digital transmission BER performance by reducing intersymbol impairment due to nonlinearity of the TWTA in a band-limited channel. However, the improvement in E_b/N_o by 1 dB with a soft-limiting TWTA is not sufficient to increase actual TDMA traffic capacity (e.g., by eliminating FEC coding). The margin obtainable with the linearizing transponder alone cannot allow TDMA transmissions from Standard A earth stations to smaller INTELSAT Standard B earth stations. However, if the up-link transmit power control is required during rain fades in K_u-band, the soft-limiting linearized transponder allows the on-board TWTA to be operated far into an overdrive region (exceeding 4 dB) under a clear-sky condition, providing some operational advantage.

And finally, the out-of-band emission due to spectrum spreading of a PSK signal in an INTELSAT VI channel is reduced with a linearized transponder. However, the improvement in E_b/N_o is insignificant for linearizing the transponder alone. The out-of-band emission from earth station HPAs with no output filters is a more serious factor, causing ACI degradation. Linearized earth station HPAs can provide further improvement when the link is served with linearized on-board transponders.

Various configurations of linearized transponders were evaluated in this paper, and it is concluded that there are no major technical problems associated with linearizing the transponders.

Acknowledgments

References

[1] J. C. Fuenzalida, O. Shimbo, and W. L. Cook, "Time-Domain Analysis of Intermodulation Effects Caused by Nonlinear Amplifiers," *COMSAT Technical Review*, Vol. 3, No. 1, Spring 1973, pp. 89–143.

[2] A. Saleh, "Intermodulation Analysis of FDMA Satellite Systems Employing Nonlinear TWTs," *IEEE Transactions on Communications*, COM-30, No. 5, May 1982, p. 1233.

[3] M. Kumar, J. C. Whartenby, and H. J. Wolkstein, "Predistortion Linearizer Using GaAs Dual Gate MESFETs for TWTA and SSPA Used in Satellite Earth Stations and Transponders," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-33, No. 12, December 1985, pp. 1479–1489.

[4] D. Cahana et al., "Linearized Transponder Technology for Satellite Communications, Part I: Linearizer Circuit Development and Experimental Characterization," *COMSAT Technical Review*, Vol. 15, No. 2A, Fall 1985 (this issue).

[5] A. Ghais et al., "Summary of INTELSAT VI Communications Performance Specifications," *COMSAT Technical Review*, Vol. 12, No. 2, Fall 1982.

[6] D. Cahana, "Practical Aspects of Predistortion Diode Linearizers for Satellite Transponders," 14th European Microwave Conference, Liège, Belgium, September 10–13, 1984, *Proceedings*, pp. 112–117.

[7] L. C. Palmer, "Computer Modeling and Simulation of Communications Satellite Channels," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-2, No. 1, January 1984, pp. 89–102.

[8] A. Hamid, S. R. Baker, and W. Cook, "A Computer Program for Communications Channel Modeling and Simulation," *COMSAT Technical Review*, Vol. 13, No. 2, Fall 1983, pp. 355–383.

[9] G. Satoh and T. Mizuno, "Impact of a New TWTA Linearizer Upon QPSK/TDMA Transmission Performance," *IEEE Journal on Selected Areas in Communications*, Vol. SAC-1, No. 1, January 1983, pp. 39–45.

[10] J. S. Snyder and T. Muratani, "Forward Error Correction for Satellite TDMA in the INTELSAT V Era," AIAA Eighth Communications Satellite Systems Conference, Orlando, Florida, April 20–24, 1980, *Proceedings*, pp. 674–683.

Appendix A. Steady-state analysis of a nonlinear network

Consider a two-port network (Figure A-1) whose input port (1) is driven by a sinusoidal source of voltage e_o, angular frequency ω, and internal impedance R_g, and
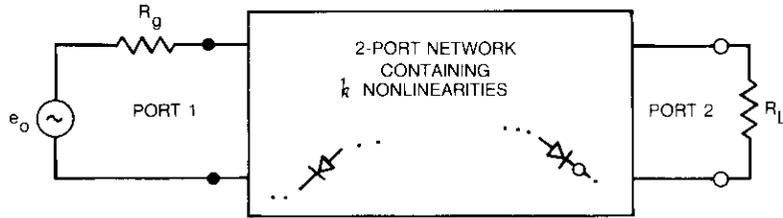

Figure A-1. *Two-Port Network Containing k Nonlinear Memoryless Two-Terminal Elements*

whose output port (2) is terminated in a load impedance R_L. The network consists of linear elements (lumped or distributed) and a number (k') of nonlinear two-terminal memoryless elements. These nonlinear memoryless elements are defined by their instantaneous voltage-current relationships:

$$V_i(t) = F_i[I_i(t)] \tag{A-1a}$$

or

$$I_i(t) = G_i[V_i(t)] \tag{A-1b}$$

where $V_i(t)$ and $I_i(t)$ are the instantaneous voltage across or current through the element i; $F_i(\bullet)$ and $G_i(\bullet)$ are single valued functions which are the inverse of each other. The network is redrawn in Figure A-2 as a linear k-port network L with the nonlinear elements extracted in ports 3, 4, . . . , k, where $k = k' + 2$, with ports 1 and 2 as the input and output ports. The linear k-port network will be assumed to possess an impedance matrix Z. The assumption of the existence of the impedance matrix of the k-port linear network is almost always satisfied in practice; however, if it does not exist for a specific network, the analysis that follows can always be carried out using other network parameters, such as the admittance, hybrid, or scattering parameters of the network. It is assumed that the network has reached steady state due to the driving source e_o and the problem is to determine the voltages and currents at each of the ports. Since the source is sinusoidal with angular frequency ω, all the steady state port voltages and currents will be periodic with period $T = \omega/2\pi$. Thus,
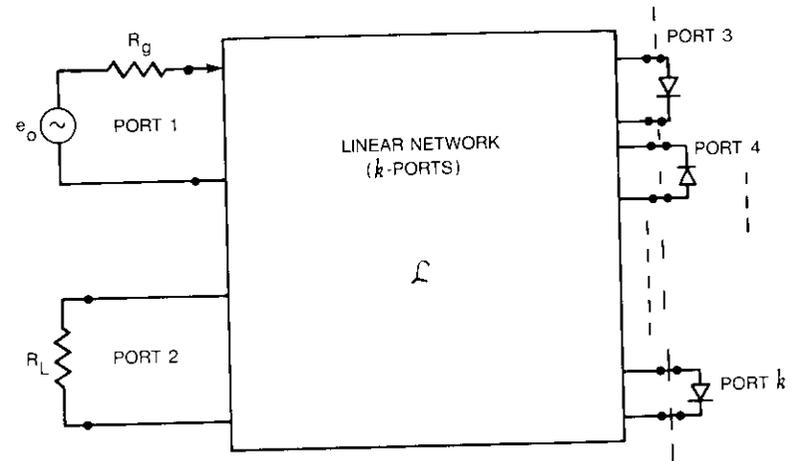
Figure A-2. *Network of Figure 1, Redrawn as a k-Port Linear Network ℓ, With Ports 3, . . . , k Terminated in the $k' = k - 2$ Nonlinearities*

the total port voltage $V_i(t)$ and current $I_i(t)$ of port i can be expanded into Fourier series expansions of the form

$$V_i(t) = \text{Re} \sum_{n=0}^{N} v_{i,n} e^{jn\omega t} \tag{A-2a}$$

$$I_i(t) = \text{Re} \sum_{n=0}^{N} i_{i,n} e^{jn\omega t} \tag{A-2b}$$

where $i = 1, 2, \ldots, k$. N is chosen large enough to include all the necessary harmonics that may exist in the network.

The terminal conditions for each of the ports can be written as follows:

$$\underline{v}_n = Z(n\omega)\underline{i}_n \tag{A-3}$$

$$\begin{bmatrix} i_{1,n} \\ \\ i_{2,n} \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{R_g} & 0 \\ \\ 0 & -\dfrac{L}{R_L} \end{bmatrix} \begin{bmatrix} v_{1,n} \\ \\ v_{2,n} \end{bmatrix} + \begin{bmatrix} \dfrac{\delta(n-1)}{R_g} e_o \\ \\ 0 \end{bmatrix} \tag{A-4}$$

$$i_{i,n} = \frac{1}{2T} \int_0^T G_{i-2} \left[\sum_{m=0}^{N} \text{Re}(v_{i,m} e^{jm\omega t}) \right] e^{-jn\omega t} dt \quad , \tag{A-5}$$

$$i = 3, 4, \ldots, k$$

where $\delta(n - 1) = 0$ for n different from 1, and 1 for n equal to 1, $Z(n\omega)$ is the value of the impedance matrix at angular frequency $n\omega$, and $\underline{v}_n$ and $\underline{i}_n$ are vectors whose ith components are equal to $v_{i,n}$, and $i_{i,n}$, respectively.

The solution to the sets of equations (A-3) through (A-5) can be carried out iteratively using the following algorithm:

a. Start by an initial set: index $\ell = 0$, $\underline{v}_n^{(\ell)} = \underline{0}$ and the corresponding $\underline{i}_n^{(\ell)} = (e_o\delta(n - 1)/R_g, 0, 0, 0)$.

b. Increment ℓ by one, i.e., set ℓ to $(\ell + 1)$.

c. Use equation (A-3) to determine the ℓth iterate $\underline{v}_n^{(\ell)} = Z(n\omega)\underline{i}_n^{(\ell-1)}$.

d. Determine the ℓth iterate for the current vector $\underline{i}_n^{(\ell)}$ by using equations (A-4) and (A-5).

e. Check if $\|\underline{i}_n^{(\ell)} - \underline{i}_n^{(\ell-1)}\| < \varepsilon_1$ and $\|\underline{v}_n^{(\ell)} - \underline{v}_n^{(\ell-1)}\| < \varepsilon_2$, where ε_1 and ε_2, are allowable errors. If true, then stop.

f. Check if $\ell > \ell_{max}$, the maximum allowable number of iterations. If true, then stop.

g. Go to step b.

This algorithm should converge to the steady state values of $\underline{v}_n$ and $\underline{i}_n$ in a reasonable number of iterations. It is important to notice that the impedance matrix Z of the linear network L need only be computed once for each of the frequencies $(n\omega)$ where $n = 0, 1, 2, \ldots, n$. The major part of the computations involved in the algorithm are in the calculations of the currents $\underline{i}_n$ from equation (A-5), through the nonlinear elements. Since these calculations are basically a form of Fourier integrals, they can be computed very efficiently using fast Fourier transform integration.

Appendix B. Parameter variations and NPR performance of statistical TWTAs

To simulate the effect of tube-to-tube variation on performance of the LTWTA, the test results of INTELSAT V TWTA transfer characteristics were used in a statistical averaging model. In the C-band, a sample size of twenty 8.5-W TWTAS were considered. Table B-1 lists the C-band TWTA group. At each input drive level, the mean and standard deviation (σ) of both output power and phase were computed. Figure B-1 shows the statistical transfer characteristics of the sample mean and $\pm 3\sigma$ deviations from the mean.

TABLE B-1. SERIAL NUMBERS OF C-BAND 8.5-W TWTAS

PART NUMBER 278382-07 (CHANNEL 7)	PART NUMBER 278382-11 (CHANNEL 11)
526, 545, 577, 560, 578, 580, 586, 591, 601, 564	524, 540, 548, 556, 559, 567, 572, 575, 590, 600
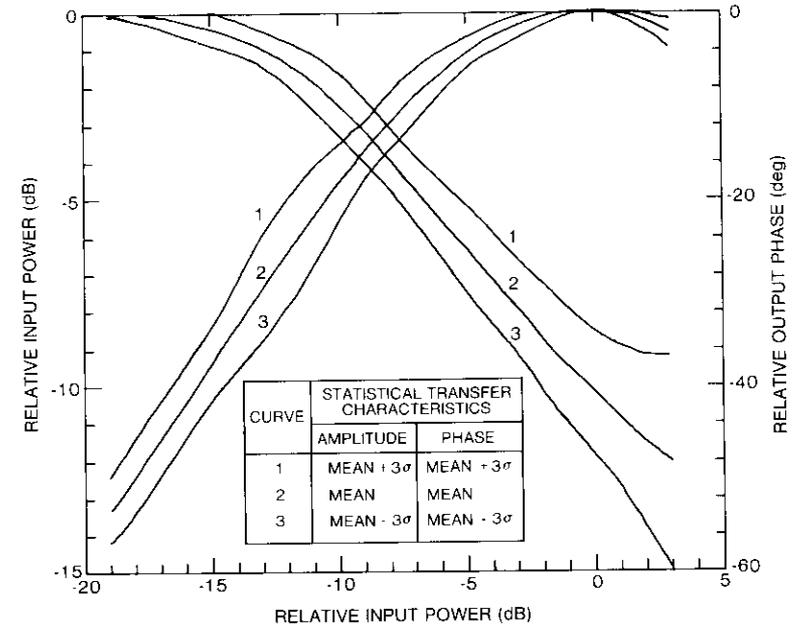
Figure B-1. *Parameter Variation of TWTAs (20 C-Band 8.5-W TWTAs)*

Figure B-2 represents the computed NPR performance of the statistical LTWTAS for 3σ amplitude and 3σ phase spreads from the mean. The analysis is for a single linearizer optimized for a measured TWTA: the linearizer was not optimally adjusted for any of the statistical TWTA characteristics. Figure B-2 shows that the phase spread among TWTAS is the dominant factor in NPR performance. For an NPR lower than 24 dB, a 3σ variation in amplitude transfer characteristic alone does not cause any NPR degradation from that of the mean TWTA. However, the 3σ spread in phase transfer characteristic causes an output backoff increase by about 0.3 dB from 5.1 dB (corresponding to the mean phase) at 20 dB of NPR.

For the K_u-band (11 GHz), a sample size of sixteen 8.5-W TWTAS listed in Table B-2 was considered. The transfer characteristics of the mean and $\pm 3\sigma$ deviations of both phase and amplitude from the mean were computed and plotted in Figure B-3. The NPR performance of the linearized K_u-band statistical TWTAS is shown in Figure B-4. The parameter variation of the 11-GHz TWTAS is larger than the C-band case, and the resulting NPR performance window is larger than C-band.
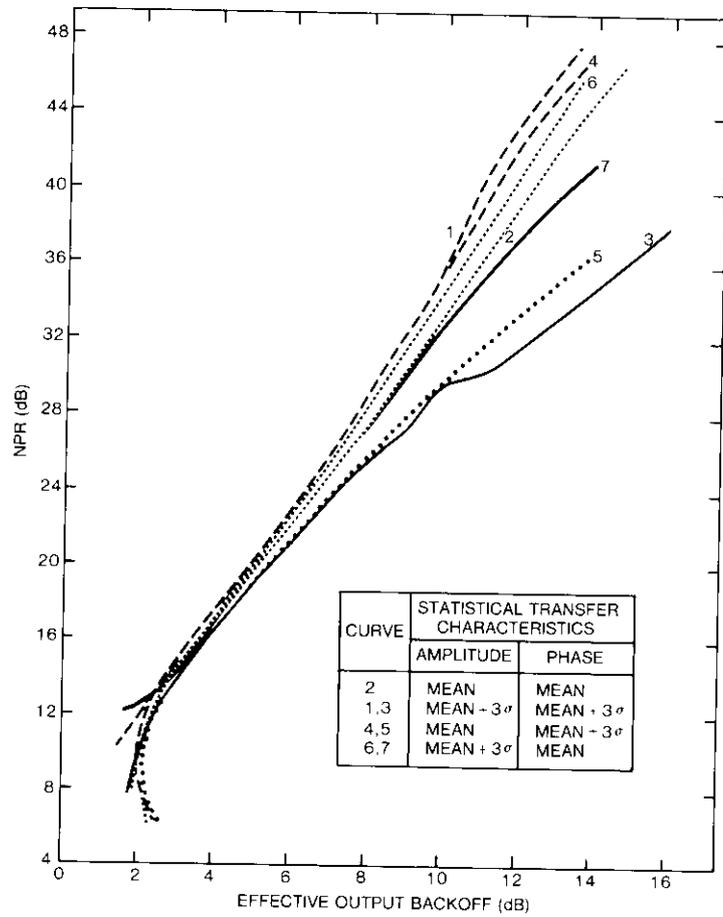
Figure B-2. *Computer NPR Characteristics of Linearized Statistical TWTAs (20 C-Band 8.5-W TWTAs)*

TABLE B-2. SERIAL NUMBERS OF K$_u$-BAND TWTAs

PART NUMBER 1	PART NUMBER 5
18, 179, 13, 32, 171, 30	65, 42, 12, 40, 34, 62, 15, 66
Total: 6 Units	Total: 10 Units (2 test runs of No. 12 and No. 62)
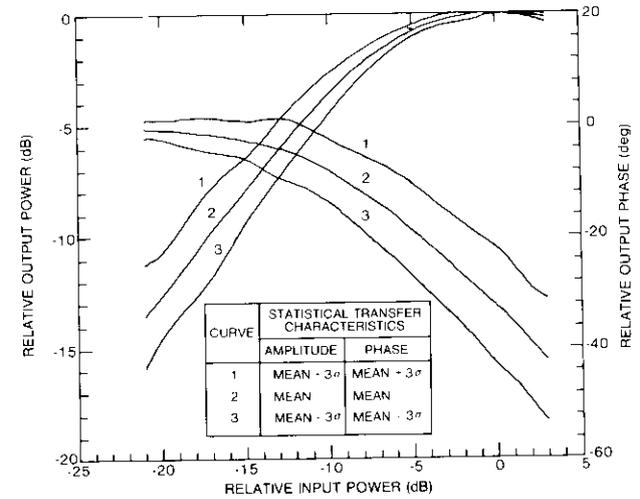
Figure B-3. *Parameter Variation of TWTAs (16 K-Band TWTAs)*
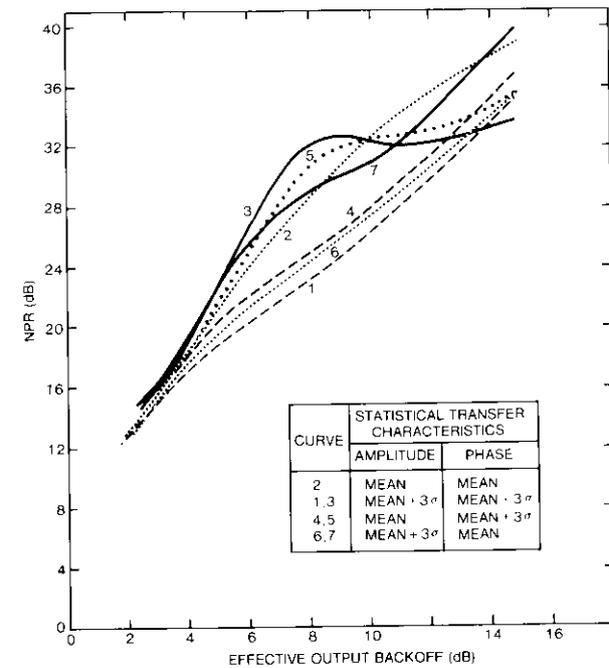
Figure B-4. *Computed NPR Characteristics of Linearized Statistical TWTAs (16 K-Band TWTAs)*
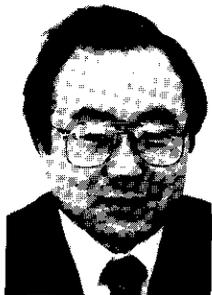
Young Soo Lee received a B.S. and M.S. in electronics engineering from Seoul National University, and an S.M. (E.E.) and Degree of Electrical Engineer from the Massachusetts Institute of Technology in 1969. From 1969 to 1979, he was with the Microwave Laboratory at COMSAT Laboratories, specifically involved in the R&D of microwave integrated circuits and systems, including spaceflight hardware development for the COMSTAR Beacon and the ATS-F projects. He has also been a Staff Scientist in the Optical Communications Laboratory, where he conducted research in optical transmission technology applicable to satellite communications systems. He is currently a Senior Staff Scientist in the Satellite System Design group of COMSAT Technical Services Division, with responsibility for the planning, analysis, and design of advanced communications satellite systems. Mr. Lee is a Senior Member of IEEE and a member of the Optical Society of America.

Iraj K. Brelian received the B.S.E.E. and M.S.E.E. degrees in 1973 and 1974, and an M.S.E. in computer information and control engineering in 1976, all from the University of Michigan. From 1977 to 1979, he was with the Burroughs Corporation, where he was involved in the design and testing of computer switching power supplies. From 1979 to 1981, he worked for the Space Communications Company, providing system analysis on the Tracking and Data Relay Satellite System for NASA. Since joining COMSAT in 1981, he has been engaged in reliability and orbital life extension activities for INTELSAT VI, technology assessment for improved orbit/spectrum utilization, microwave filter computer simulation design, investigation of frequency spectrum sharing problems between fixed satellite and other radio services, investigations of linearized transponders for INTELSAT VI, spacecraft subsystems characteristics, and analysis of spacecraft saturation flux density and gain step requirements. He is presently a Senior Satellite Systems Analysis Engineer in the Satellite Systems Design group of COMSAT Technical Services Division.

Ali E. Atia received a B.S.E.E. from Ain Shams University, Cairo, Egypt, and an M.S.E.E. and a Ph.D. from the University of California at Berkeley. Prior to joining COMSAT in 1969, he held various research and teaching positions at both universities. As a Senior Scientist in the Microwave Laboratory at COMSAT Laboratories, he made significant contributions to several satellite programs, including INTELSAT, ARABSAT, and AUSTRALIASAT. He was also involved in COMSAT's NASA ATS-F Propagation Experiment and the COMSTAR Beacon Experiment. He is presently a Senior Director of Systems and Software Design for COMSAT Technical Services, responsible for the analysis and evaluation of current and future satellite systems. Dr. Atia is a Senior Member of IEEE and an Associate Fellow of AIAA.

CTR Note

Parallel implementation of linear feedback shift register circuits

T. INUKAI AND J. S. SNYDER

Introduction

Serial feedback shift register (FSR) circuits are frequently used in digital systems to perform functions such as pseudorandom sequence generation, data scrambling and descrambling [1], error detection [2], and error correction [3],[4]. However, implementation of a serial FSR circuit in a high-speed application such as a time-division multiple-access (TDMA) satellite communications system may require not only a large amount of power, but also a complex circuit design. On the other hand, a parallel FSR structure [5],[6] often allows the use of a lower speed logic family, thus offering a higher level of integration, lower cost, and reduced power dissipation. Furthermore, a parallel form is more suitable for microprocessor-based equipment in which data are commonly bussed in parallel bytes.

A simple procedure is developed for converting a known serial FSR to an equivalent parallel configuration. The serial structure is first described by a matrix equation and then elementary operations are performed that yield a matrix description of the corresponding parallel circuit. Although a variation of this matrix approach has been used in the computer industry [5],[6], this conversion technique has not been generally

T. Inukai is a Senior Research Scientist in the Network Technology Division at COMSAT Laboratories.

J. S. Snyder is a Staff Scientist in the Image Processing and Digital Techniques Department of the Communications Techniques Division at COMSAT Laboratories.

known to communications engineers. Moreover, it is proven here that only one matrix rather than two is needed to fully specify the parallel connections. Most importantly, the practical benefit of this approach is that all matrix elements may be obtained by a simple hand procedure based on inspection of the serial register.

General feedback shift register circuit

The general serial FSR (Figure 1) implements polynomial division of an input sequence by a generator or characteristic polynomial

$$g(x) = x^n + p_1 x^{n-1} + p_2 x^{n-2} + \ldots + p_{n-1} x + 1 \quad .$$

Since only the binary case is considered here, the hardware implementation of each product in the generator polynomial is simply a connection or an open circuit, depending on whether the corresponding coefficient, p_i, is a binary 1 or 0. The binary values of the input data bits are considered to be coefficients of a data polynomial which is subjected to ordinary long division with addition performed modulo two (*i.e.*, each adder is an exclusive-OR gate).
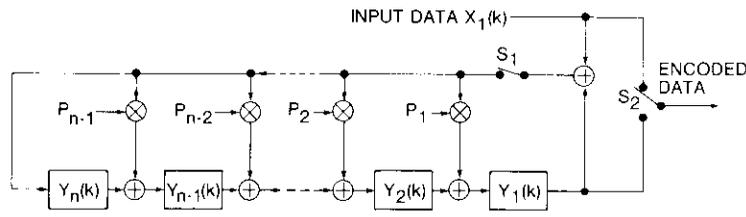

Figure 1. *Serial Feedback Shift Register Circuit*

Input data enter the FSR (Figure 1) with switch S_1 closed and are also steered directly to the output by switch S_2. After a block of input data has been completely entered, the register contains the remainder from the division of the input polynomial by the generator polynomial. Switch S_1 is then opened, and S_2 is switched to output this remainder.

Adding the input directly to the feedback signal from the first stage is equivalent to multiplying the input polynomial by the term x^n. This premultiplication causes the remainder to be already in the register and ready for readout after the last input bit has been entered. Mathematically, the circuit calculates the remainder polynomial in

$$\frac{x^n f(x)}{g(x)} = q(x) + \frac{r(x)}{g(x)}$$

where $f(x)$ = input polynomial
  $g(x)$ = generator polynomial
  $q(x)$ = quotient polynomial
  $r(x)$ = remainder polynomial.

The state of the serial FSR at time $t = k$ consists of its contents, $y_1(k)$, $y_2(k)$, . . . , $y_n(k)$. In matrix form, the recursive relationship between successive states may be described by

$$\begin{bmatrix} y_1(k) \\ y_2(k) \\ \cdot \\ \cdot \\ \cdot \\ y_{n-1}(k) \\ y_n(k) \end{bmatrix} = \begin{bmatrix} p_1 & 1\,0\,0\ldots0\,0 \\ p_2 & 0\,1\,0\ldots0\,0 \\ & \cdot \\ & \cdot \\ & \cdot \\ p_{n-1} & 0\,0\,0\ldots0\,1 \\ 1 & 0\,0\,0\ldots0\,0 \end{bmatrix} \begin{bmatrix} y_1(k-1) \\ y_2(k-1) \\ \cdot \\ \cdot \\ \cdot \\ y_{n-1}(k-1) \\ y_n(k-1) \end{bmatrix} + \begin{bmatrix} p_1 \\ p_2 \\ \cdot \\ \cdot \\ \cdot \\ p_{n-1} \\ 1 \end{bmatrix} x_1(k-1) \qquad (1)$$

or, more compactly, by

$$[y(k)] = T_c[y(k-1)] + T_x x_1(k-1) \quad . \tag{2}$$

Note that the left column of the matrix T_c, sometimes called the companion matrix, consists of the coefficients of $g(x)$, as does the column matrix T_x. If an alternate serial FSR structure is used in which feedback is taken from various stages, added, and applied only to the nth stage, the coefficients of $g(x)$ appear along the bottom row of T_c instead. If no input signal is present, as in a pseudorandom sequence generator, the term containing T_x is absent.

The corresponding parallel configuration (Figure 2) employs m parallel inputs and is clocked at a rate $1/m$ that of the serial version. A simple procedure for generating the n parallel output sequences, $y_i(km)$, for $2 \le m \le n$ will be developed from operations on the matrix T_c.

The serial register state at time $t = km - m$ is

$$[y(km-m)] = \begin{bmatrix} y_1(km-m) \\ y_2(km-m) \\ \cdot \\ \cdot \\ \cdot \\ y_n(km-m) \end{bmatrix} \quad . \tag{3}$$

Applying equation (2) repeatedly m times to equation (3) yields the state of the register at time $t = km$ in terms of its state m units of time earlier, as

$$[y(km)] = T_c^m[y(km-m)] + T_p[x_1(km-m)] \tag{4}$$

where

$$T_p = [(T_c^{m-1}T_x)\,(T_c^{m-2}T_x)\,\ldots\,(T_cT_x)\,(T_x)] \tag{5}$$

Figure 2. *Parallel Feedback Shift Register Circuit*

and

$$[x_1(km - m)] = \begin{bmatrix} x_1(km - m) \\ x_1(km - m + 1) \\ \cdot \\ \cdot \\ \cdot \\ x_1(km - 2) \\ x_1(km - 1) \end{bmatrix} . \qquad (6)$$

The state of the parallel implementation after each clock period is identical to the state of the serial register after m clock periods. Thus, designing the corresponding parallel FSR circuit consists of determining T_c^m and T_p. The matrix T_c^m defines the connections providing feedback from the previous state, and T_p provides the connections from the m parallel inputs. However, the matrix T_c^m need not be calculated explicitly because it may be fully specified from a knowledge of T_p [6], as will be proven. Furthermore, a procedure for rapidly determining the elements of T_p will be developed.

Determining T_c^m and T_p

The relationship between T_c^m and T_p may be established by first writing T_c in the form

$$T_c = \begin{bmatrix} T_x & \begin{matrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ & & . & & & \\ & & . & & & \\ & & . & & & \\ 0 & 0 & . & . & . & 1 \\ 0 & 0 & . & . & . & 0 \end{matrix} \end{bmatrix} = \begin{bmatrix} (T_x)_{n,1} & \begin{matrix} I_{(n-1),(n-1)} \\ - - - - - \\ O_{1,(n-1)} \end{matrix} \end{bmatrix}$$

That is, the first column of T_c is identical to the column matrix, T_x, and the remaining columns form an $(n - 1) \times (n - 1)$ identity matrix, $I_{(n-1),(n-1)}$, plus a row of zeros, $O_{1,(n-1)}$.

In this format, T_c^2 may be written as

$$T_c^2 = T_c T_c = \begin{bmatrix} T_c T_x & T_x & \begin{matrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ & & . & & & \\ & & . & & & \\ 0 & 0 & . & . & . & 1 \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & . & . & . & 0 \end{matrix} \end{bmatrix} .$$

In general,

$$T_c^m = \begin{bmatrix} T_c^{m-1} T_x & T_c^{m-2} T_x & . . . & T_c T_x & T_x & \begin{matrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & . & . & . & 0 \\ & & . & & & \\ & & . & & & \\ & & . & & & \\ 0 & 0 & . & . & . & 1 \\ 0 & 0 & . & . & . & 0 \\ & & . & & & \\ 0 & 0 & . & . & . & 0 \end{matrix} \end{bmatrix}$$

or

$$T_c^m = \begin{bmatrix} (T_p)_{n,m} & \begin{matrix} I_{(n-m),(n-m)} \\ - - - - - \\ O_{m,(n-m)} \end{matrix} \end{bmatrix} ; m \le n . \qquad (7)$$

That is, the first m columns of T_c^m form the $n \times m$ matrix, T_p; the upper right-hand corner consists of an $(n - m) \times (n - m)$ identity matrix, $I_{(n-m),(n-m)}$; and the lower right-hand corner contains m rows of zeros, $O_{m,(n-m)}$. Consequently, only the matrix T_p needs to be determined to fully specify the parallel FSR connections given by equation (4).

In the parallel configuration, T_p specifies the connections from the m parallel inputs. The elements of T_p may be readily determined from the serial version [7] by noting that shifting in a single binary 1 gives the contribution to the state of the FSR at time $t = k$ of an input at time $t = k - 1$. Furthermore, shifting the serial register i additional times yields the contribution to the state of the FSR at time $t = k$ of an input at time $t = k - 1 - i$.

The state of the serial FSR immediately after entering the single binary 1 constitutes the mth (i.e., rightmost) column of T_p, which is denoted by T_c in equation (5). The column is ordered such that the element composed of stage $y_1(k)$ (Figure 2) is at the top. The state after the second shift specifies column $m - 1$, which corresponds to the term T_cT_x in equation (5). The serial FSR state after m such shifts consists of the elements of column 1 of T_p, which represents the term $T_c^{m-1} T_x$. Use of this procedure is clarified in the following example.

Example

The simplified matrix approach to parallel FSR design, based on equations (4) and (7) and the rapid technique for determining T_p, may be illustrated by considering the (127, 112) modified BCH code used for error correction [8],[9] which would ordinarily be implemented in serial form (Figure 3). The encoder divides an input sequence of 112 information bits by the generator polynomial

$$g(x) = x^{15} + x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^7 + x^2 + x + 1 \quad .$$

The remainder of this calculation forms 15 check bits, which are appended to the information bits to form a complete code block.

INPUT DATA $X_1(k)$
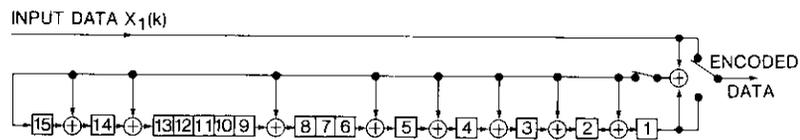
Figure 3. *Serial (127, 112) BCH Encoder*

For a bus-oriented system with 8-bit parallel data, it is more convenient to convert the serial structure to a parallel form. The parallel configuration is determined by evaluating the matrices T_p and T_c^m for $m = 8$. The matrix T_p is determined by applying a single 1 at the data input of the serial register and shifting eight times, with the state recorded after each shift. The results shown in Table 1 are written by inspection of the serial register (Figure 3).

TABLE 1. SUCCESSIVE STATES RESULTING FROM APPLYING A SINGLE BINARY INPUT 1 AND SHIFTING EIGHT TIMES

SHIFT NUMBER	SHIFT REGISTER STAGE														
	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
1	1	1	1	0	0	0	0	1	0	0	1	1	1	1	1
2	1	0	0	1	0	0	0	1	1	0	1	0	0	0	0
3	0	1	0	0	1	0	0	0	1	1	0	1	0	0	0
4	0	0	1	0	0	1	0	0	0	1	1	0	1	0	0
5	0	0	0	1	0	0	1	0	0	0	1	1	0	1	0
6	0	0	0	0	1	0	0	1	0	0	0	1	1	0	1
7	1	1	1	0	0	1	0	1	1	0	1	1	0	0	1
8	1	0	0	1	0	0	1	1	1	1	1	0	0	1	1

Thus, T_p is given by

$$T_p = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad .$$

From equation (7), T_c^8 may immediately be written as

$$T_c^8 = \left[\begin{array}{cccccccc|ccccccc} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] .$$

From equation (4), the matrices T_p and T_c^8 describe the parallel FSR connections (Figure 4). T_p provides the connections from the eight parallel inputs, and T_c^8 defines the feedback connections from the previous state.

The resulting parallel (127, 112) BCH encoder accepts a sequence of 112 input bits in the form of 8-bit parallel bytes. After the last byte is clocked into the parallel encoder, the register contents $(y_1, y_2, \ldots, y_{15})$ comprise the desired check bits, which are identical to those produced by the serial encoder.
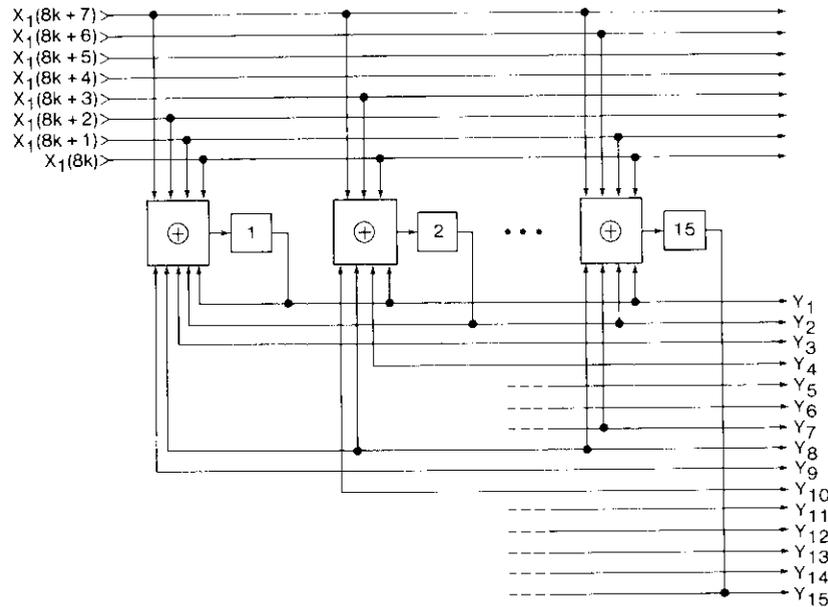
Figure 4. *Parallel (127, 112) BCH Encoder*

An 8-bit parallel implementation of a complete error-correction coder/decoder employing the (127, 112) modified BCH code of the example has been developed at COMSAT Laboratories for use in the 120-Mbit/s INTELSAT TDMA system [9]. A major benefit of the 8-bit parallel structure is that the codec clock rate needs to be only 15 MHz for 120-Mbit/s system operation. This significantly slower clock rate allows the use of Schottky transistor-transistor logic (TTL) circuits instead of emitter-coupled logic (ECL) devices. The use of TTL decreases chip cost and power consumption dramatically, while eliminating the need for ECL termination resistors.

This form of parallel architecture does not involve repeating a function eight times, but instead involves implementing a single function which operates on 8-bit parallel data. The parallel FSR contains the same number of register elements (Figure 4) as the serial structure, while employing an exclusive-OR function ahead of each register stage. Parallel-to-serial and serial-to-parallel transformations are eliminated in both the encoder and decoder since the parallel structure is directly compatible with the system data bus. Moreover, the 8-bit data path facilitates use of an efficient RAM data buffer in the decoder instead of numerous shift registers. The total codec employs only 127 TTL integrated circuits, as compared with the 210 ECL chips in another design—a 40-percent reduction in chip count.

Conclusion

A general approach to the design of parallel FSR circuits based on their serial representation has been developed. This method evaluates the T_p matrix to obtain the

parallel input connections and T_c^m to obtain the parallel feedback connections. The T_c matrix specifies the serial register feedback connections and m is the degree of parallelism desired.

This parallel design method is practical because no actual matrix operations are required. T_p may be rapidly hand calculated by a procedure based on a single binary 1 which is shifted m times in the serial FSR. T_c^m may then be immediately written by inspection of T_p. This technique considerably simplifies the process of designing parallel FSR circuits, and the resulting parallel structure often permits such system advantages as lower clock rates and greatly reduced chip counts.

References

[1] J. E. Savage, "Some Simple Self-Synchronizing Digital Data Scramblers," *Bell System Technical Journal*, Vol. 46, No. 2, February 1967, pp. 449–487.

[2] W. W. Peterson and D. T. Brown, "Cyclic Codes for Error Detection," *Proceedings of the IRE*, January 1961, pp. 228–235.

[3] S. Lin, *An Introduction to Error-Correcting Codes*, Englewood Cliffs, New Jersey: Prentice-Hall, 1970.

[4] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, Cambridge, Massachusetts: MIT Press, 1972.

[5] M. Y. Hsiao and K. Y. Sih, "Serial to Parallel Transformation of Linear-Feedback Shift-Register Circuits," *IEEE Transactions on Electronic Computers*, Vol. EC-13, No. 12, December 1964, pp. 738–740.

[6] A. M. Patel, "A Multi-Channel CRC Register," Spring Joint Computer Conference, Atlantic City, New Jersey, May 18–20, 1971, *AFIPS Conference Proceedings*, Vol. 38, pp. 11–14.

[7] N. Glover, *Practical Error Correction Design for Engineers*, Data Systems Technology, 1047 East Ninth Avenue, Unit 102, Broomfield, Colorado, 80020, 1982, pp. 211–214.

[8] J. S. Snyder and T. Muratani, "Forward Error Correction for Satellite TDMA in the INTELSAT V Era," AIAA 8th Communications Satellite Systems Conference, Orlando, Florida, April 20–24, 1980, *Proceedings*, pp. 674–683.

[9] C. J. Wolejsza, Jr., J. S. Snyder, and J. M. Kappes, "120-Mbit/s TDMA Modem and FEC Codec Performance," to be published, *COMSAT Technical Review*, Vol. 15, No. 2, Part B, Fall 1985.

Translations of Abstracts

Modulateur MDPQ à 4 GHz pour débits binaires élèves

K. KARIMULLAH ET R. G. EGRI

Sommaire

On a mis au point un modulateur à déplacement de phase à quadrature (MDPQ), circuit intégré hyperfréquence fonctionnant à 4 GHz, pour les vitesses de transmission de données de 120 Mbit/s. La phase et l'amplitude mesurées sont équilibrées à $\pm 1°$ et $\pm 0,15$ dB, respectivement, sur une largeur de bande de 200 MHz. Le temps de montée mesurée d'une transition typique d'état de phase est inférieur à 1,5 ns. Ces caractéristiques garantissent de bonnes performances à des débits binaires mêmes supérieurs à 120 Mbit/s.

La modulation MPDQ est obtenue au moyen de deux modulateurs à déplacement de phase binaire (MDPB) équilibrés en parallèle, ainsi que d'un circuit hybride de 3 dB à 90° pour combiner leurs signaux de sortie. La concentration des modulateurs MDPB s'inspire d'une approche différente, et chacun présente un équilibre d'amplitude excellent de $\pm 0,1$ dB et un équilibre de phase de $\pm 0,5°$ sur la largeur de bande de 200 MHz. Le modulateur MPDQ est étalé sur un substrat d'aluminium mesurant 4,5 sur 6,35 cm environ, et il est directement compatible avec les circuits d'attaque à porte logique (ECL) ordinaires à émetteurs couplés. Ses dimensions, sa consommation d'énergie et son poids permettent de l'utiliser dans un modem régénérateur de satellite.

Algorithme permettant de tracer les contours pour la preparation de diagrammes de rayonnement d'antenne

K. D. FULLETT ET V. E. RIGINOS

Sommaire

On trouvera ici une description des algorithmes utilisés pour produire les tracés de contours appliqués aux diagrammes de rayonnement d'antenne obtenus pendant les essais sur orbite de satellites de télécommunications. Ces contours sont utilisés lors de la planification opérationnelle et pour vérifier que les satellites sur orbite respectent bien les spécifications. Ces algorithmes ont été mis au point pour permettre de tracer les contours de couverture d'antenne en utilisant des données non-uniformes et à distribution clairsemée en azimut et en élévation dans les coordonnées d'antenne de satellite, conditions sur lesquelles les algorithmes classiques de génération de contour

Tecnologia de transpondedores linealizados para las telecomunicaciones por satelite

Parte I: Perfeccionamiento del circuito linealizador y caracterización mediante experimentos

D. CAHANA, J. R. POTUKUCHI, R. G. MARSHALEK Y D. K. PAUL

Abstracto

Se presentan los principios operativos y el diseño de los linealizadores de diodos de predistorsión del tipo de transmisión y reflexión. Por medio de experimentos se evalúa el funcionamiento de cuatro tubos de ondas progresivas linealizados diferentes (LTWT) para trabajar con portadoras múltiples, y se mide el rendimiento de un LTWT con limitador gradual. También se muestran los resultados del experimento sobre linealización en una frecuencia intermedia de transpondedor. Algunos de estos resultados se utilizan en la Parte II de este estudio para realizar simulaciones de computadora con distintas configuraciones de LTWT en un canal de transmisión.

Parte II: Simulación de sistemas y evaluación del funcionamiento

Y. S. LEE, I. BRELIAN Y A. ATIA

Abstracto

Se describe el análisis de simulación de un linealizador de diodos y de amplificadores de tubos de ondas progresivas (TWTA) linealizados, conforme a los resultados obtenidos en los experimentos descritos en el artículo complementario (Parte I) que aparece en esta edición. También se presentan los resultados de la simulación del funcionamiento de transmisión del canal bajo distintas configuraciones de transpondedor con TWT linealizado para cursar tráfico por portadoras múltiples, así como tráfico TDMA por una sola portadora. Se evaluaron varios linealizadores de diodos del tipo de predistorsión, con y sin limitador gradual, en las bandas C y K_u, para linealizar los TWTA, bien sea directamente o en una frecuencia intermedia a 4 GHz para canales de banda K_u. En la evaluación del funcionamiento se incluyeron las características al comienzo y al final de la vida útil del canal de transpondedores linealizados. Sobre la base de las estadísticas referentes a las características de transferencia del TWT, se determinó el límite inferior de funcionamiento del canal de transpondedor linealizado y se examinaron otras cuestiones sobre la puesta en servicio de los transpondedores de satélite linealizados.

Corrigendum

In the paper "Pseudo-Bit-Error-Rate Measurement for 120-Mbit/s TDMA" by J. S. Snyder and W. J. Hersey, *COMSAT Technical Review*, Vol. 14, No. 2, Fall 1984, the following revisions should be made:

1. On page 293, the last sentence should read:

   For a large number of observations, the probability is P_c (known as the degree of confidence) that the estimate will differ from the true mean by less than K_c standard errors of the mean [10], where P_c, the area under the Gaussian curve between abscissa points $\pm K_c$ σ-units from the mean, is given by

   $$P_c = 1 - 2Q(K_c) .$$

2. In Table 2 (page 298), the quantity in the leftmost column should be labeled n and its units should be 10^6 symbols, as shown below:

TABLE 2. PREDICTED MEASUREMENT ERROR AT $P_e = 1 \times 10^{-5}$ for θ = 8° WITH 2° PHASE ERROR (95-PERCENT CONFIDENCE)

n (10^6 SYMBOLS)	P_c	$\left(\dfrac{\Delta P_p}{P_p}\right)_{int}$	$\left(\dfrac{\Delta P_p}{P_p}\right)_{db}$	$\dfrac{\Delta P_p}{P_p}$	$\dfrac{\Delta P_e}{P_e}$
0.7	0.95	0.145	0.096	0.174	0.234
0.7	0.90	0.122	0.081	0.147	0.197
1.0	0.95	0.122	0.096	0.155	0.208
1.0	0.90	0.102	0.081	0.130	0.175
1.4	0.95	0.103	0.096	0.141	0.189
1.4	0.90	0.086	0.081	0.118	0.159
2.0	0.95	0.086	0.096	0.129	0.173
2.0	0.90	0.072	0.081	0.109	0.146

3. On page 310, the correct equation for the variance of the sampling distribution of the mean is

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} .$$