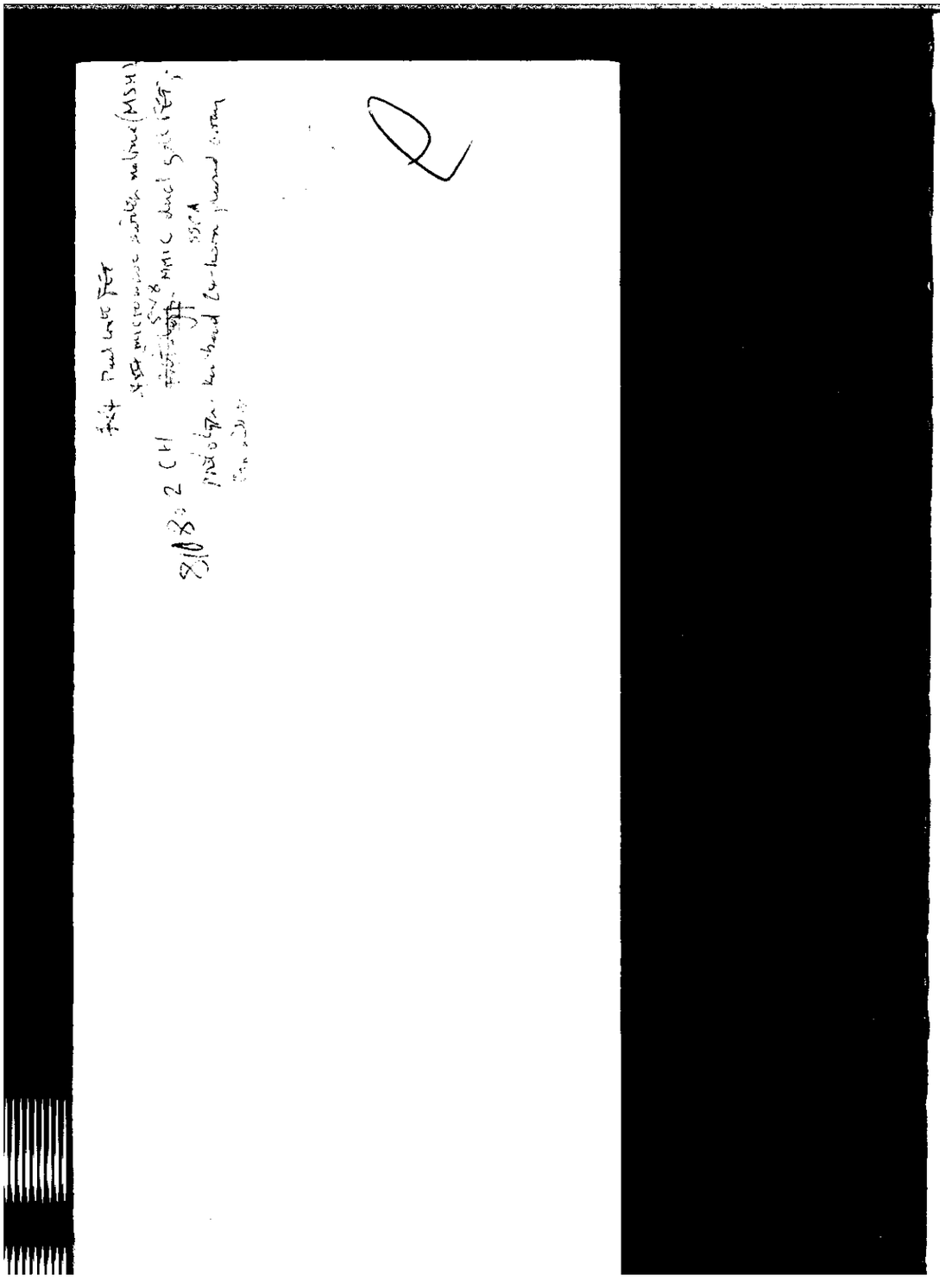


COMSAT

Technical Review

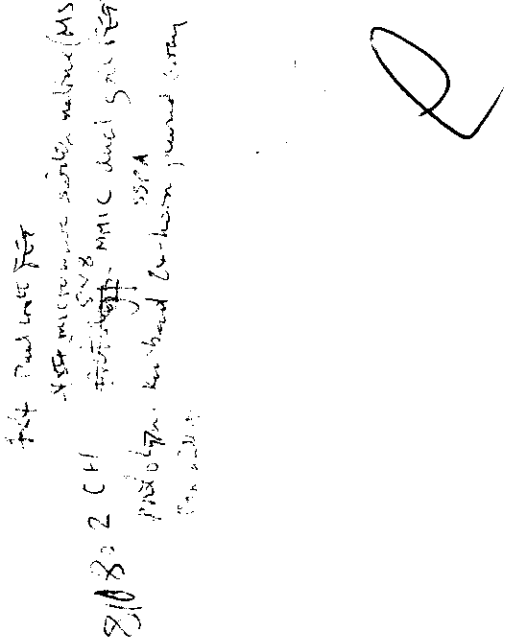
Volume 3 Number 2, Fall 1973

The COMSAT TECHNICAL REVIEW is published twice a year by the Communications Satellite Corporation (COMSAT). Subscriptions including postage, $5.00 U.S. per year; single copies, $3.00 U.S. Bulk rates are available upon request. Remittances for subscriptions should be made payable to the Treasurer, COMSAT, and addressed to COMSAT TECHNICAL REVIEW, Room 5273, 950 L'Enfant Plaza, S. W., Washington, D. C. 20024, U. S. A.



# COMSAT Technical Review

## Volume 3 Number 2, Fall 1973



Index: phase shift keying, power spectra, traveling wave tubes, mathematical models, computerized simulation.

# PSK signal power spectrum spread produced by memoryless nonlinear TWTs

G. ROBINSON, O. SHIMBO, AND R. FANG

## Abstract

This paper develops a combined method of analysis and simulation to obtain the power spectrum of M-ary PSK signals at the output of the TWT with both amplitude and phase nonlinearities. Based on this analysis, the spectra of 4-, 8-, and 16-phase PSK signals have been computed for various symbol rates and satellite input backoff values. The adjacent channel interference caused by the spectrum spreading has also been evaluated.

The computed spectra have been compared with measured spectra of 4-phase PSK signals for one symbol rate and various backoff values. In all cases, the computed and measured spectra are in good agreement. It is found that the power spectrum spread is mainly caused by the amplitude nonlinearity rather than the phase nonlinearity of the TWT amplifier. It is also verified that, at a given symbol rate, the power spectrum spread of 8- and 16-phase signals is not significantly different from that of a 4-phase PSK signal.

---

This paper is based partly upon work performed in COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

## Introduction

The evaluation of the power spectrum spread of PSK signals caused by TWT amplifier nonlinearities is an important problem. Excessive spread produces adjacent transponder (or adjacent channel) interference which may be unacceptable. Sharper filters at the satellite output and/or at the earth station input may be used to keep this undesired interference within acceptable levels. It is well-known that sharper filters will produce larger amounts of intersymbol interference, however. Thus, a compromise is necessary and, to provide an optimum tradeoff between intersymbol interference and adjacent channel interference, an accurate evaluation of the power spectrum spread of PSK signals is just as important as an assessment of the system's performance in the presence of intersymbol interference and Gaussian noise.

This paper develops a combined method of analysis and simulation to obtain the power spectrum of M-ary PSK signals at the output of the TWT with both amplitude and phase nonlinearities. Based on this method, a computer program has been developed and used to compute the spectra of 4-, 8-, and 16-phase PSK signals for various symbol rates and satellite input backoff values. As expected, for a given symbol rate, the power spectrum spread of the PSK signal caused by the TWT nonlinearities decreases as the satellite input backoff increases.

It has been found that the power spectrum spread is primarily caused by the amplitude nonlinearity rather than the phase nonlinearity of the TWT amplifier. It has also been verified that, at a given symbol rate, the power spectrum spread of 8- and 16- phase PSK signals is not significantly different from that of a 4-phase PSK signal. In addition, for a given transponder configuration and a given input backoff, the power spectrum spread is more severe at higher symbol rates, which is not unexpected.

The computed spectra have been compared with measured spectra of 4-phase PSK signals for one symbol rate and various input backoff values. In all cases, the computed and measured spectra are in good agreement.

Once the power spectrum spread is obtained, the computation of adjacent transponder or adjacent channel interference becomes straightforward.

The next section describes the complete transmission system and a mathematical model. The third section presents the method of evaluating the power spectrum spread at the TWT output. The actual numerical computation is outlined in the fourth section, and the formula for calculating the adjacent channel interference from the computed spectrum is given in the fifth section. Then, in the last section, the computed and measured spectra are discussed. Finally, the Appendix shows that the spectrum computation formula derived in this paper reduces to the familiar formula for linear systems when the nonlinear TWT is replaced with a linear one.



## Transmission path model

Figure 1 is a block diagram of the satellite communications chain. Block $F_1$ represents the amplitude and phase characteristics of the filters and the transmission medium between the PSK modulator and the satellite. Blocks $F_2$ and $F_3$ are the satellite input and output filters, respectively. Block TWT represents the measured amplitude and differential phase characteristics of the TWT, and $F_4$ represents the amplitude and phase characteristics of the filters and the transmission medium between the output of $F_3$ and the input of the PSK demodulator.

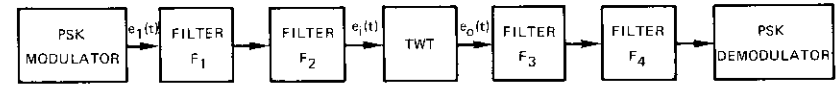



Figure 1. *Block Diagram of the Satellite Communications Chain*

The M-ary coherent PSK signal at the input of $F_1$ can be written as

$$e_1(t) = \sum_{k=-\infty}^{\infty} AW(t - kT) \cos(\omega_c t + \phi_k) \tag{1}$$

where $A$ = carrier amplitude

$W(t)$ = waveform of the transmitting pulse in Figure 2

$T$ = pulse duration ($R_s = T^{-1}$ = symbol rate)

$\omega_c$ = carrier frequency

$\phi_k$ = transmitted symbol selected from the set $\{\pi/M + i(2\pi/M)\}$, $i = 0, \ldots, M-1$, with equal probability.

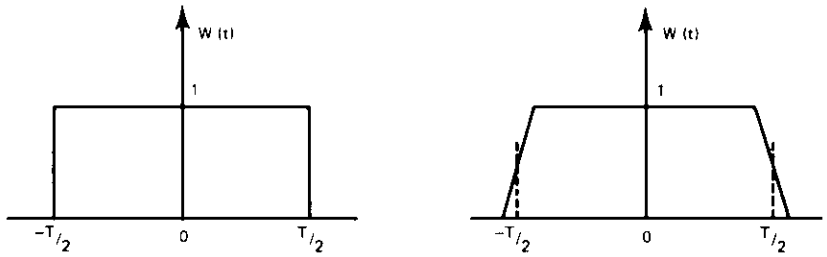



Figure 2. *Typical Waveforms for the Transmitting Pulse*

An alternate form of equation (1) is

$$e_1(t) = Re\left\{e^{j\omega_c t} \sum_{k=-\infty}^{\infty} (a_k + jb_k)\, W(t - kT)\right\} \tag{2}$$

where

$$a_k = A \cos \phi_k \tag{3a}$$

$$b_k = A \sin \phi_k \tag{3b}$$

are the respective in-phase and quadrature random variables of the PSK signal in the time interval $(kT - T/2, kT + T/2)$. For example, in 4-phase PSK, $(a_k, b_k)$ can randomly take one of the four values $(\pm A/\sqrt{2}, \pm A/\sqrt{2})$ with equal probability. The characteristics of the bandpass filters in Figure 1 are not necessarily symmetrical.

Assume that $h(t)$ is the low-pass analog of the impulse response of the filters between the PSK modulator and the TWT. Then the input to the TWT can be written as

$$e_i(t) = \frac{1}{2} Re\left\{e^{j\omega_c t} \sum_{k=-\infty}^{\infty} (a_k + jb_k)\, [W(t - kT) * h(t)]\right\} \tag{4}$$

where the operation $*$ denotes convolution. Equation (4) can be re-written as

$$e_i(t) = x(t) \cos \omega_c t - y(t) \sin \omega_c t \tag{5a}$$

or

$$e_i(t) = \sqrt{x^2(t) + y^2(t)} \cos\left\{\omega_c t + \tan^{-1}\left[\frac{y(t)}{x(t)}\right]\right\} \tag{5b}$$



where

$$x(t) = \sum_{k=-\infty}^{\infty} [a_k R(t - kT) - b_k I(t - kT)] \tag{6a}$$

$$y(t) = \sum_{k=-\infty}^{\infty} [a_k I(t - kT) + b_k R(t - kT)] \tag{6b}$$

and

$$R(t) = \frac{1}{2} Re\, \{W(t) * h(t)\} \tag{7a}$$

$$I(t) = \frac{1}{2} Im\, \{W(t) * h(t)\} \tag{7b}$$

are the respective in-phase and quadrature components of the pulse response of the cascaded filters between the PSK modulator and the TWT.

Define

$$x_1 = x(t), \qquad y_1 = y(t) \tag{8a}$$

$$x_2 = x(t + \tau), \qquad y_2 = y(t + \tau) \tag{8b}$$

and represent the TWT output as

$$e_o(t) = g\left(\sqrt{x_1^2 + y_1^2}\right) \cos\left[\omega_c t + \tan^{-1}\frac{y_1}{x_1} + f\left(\sqrt{x_1^2 + y_1^2}\right)\right] \quad . \tag{9}$$

A comparison of equations (5b) and (9) shows that the function $g(\sqrt{x_1^2 + y_1^2})$ represents the amplitude distortion caused by the AM/AM conversion of the TWT and the function $f(\sqrt{x_1^2 + y_1^2})$ represents the phase distortion caused by the AM/PM conversion. To obtain the power spectrum of $e_o(t)$, the product $e_o(t)\, e_o(t + \tau)$ is formed; the ensemble average of this product is taken on $(a_k, b_k)$, $k = 0, \pm 1, \pm 2, \ldots$, and the time average is taken over $t$; and finally the Fourier transform of the averaged version is taken with respect to $\tau$.

## Power spectrum at the output of the TWT

The product $e_o(t)\, e_o(t + \tau)$ is given by

$$e_o(t)\, e_o(t + \tau) = \frac{1}{2} Re\left[\frac{g(\sqrt{x_1^2 + y_1^2})}{\sqrt{x_1^2 + y_1^2}} e^{jf(\sqrt{x_1^2 + y_1^2})} (x_1 + jy_1) \frac{g(\sqrt{x_2^2 + y_2^2})}{\sqrt{x_2^2 + y_2^2}}\right.$$
$$\left. \cdot\, e^{-jf(\sqrt{x_2^2 + y_2^2})} (x_2 - jy_2)\, e^{-j\omega_c \tau}\right] \tag{10}$$

where the components of $e^{\pm j(2\omega_c\tau)}$ are omitted. Equation (10) can be rewritten as

$$e_o(t)\, e_o(t+\tau) = \frac{1}{2}\, Re\, \{H(x_1, y_1)\, H^*(x_2, y_2)\, e^{-j\omega_c\tau}\} \tag{11}$$

where
$$H[x(t), y(t)] = \frac{g[\sqrt{x^2(t) + y^2(t)}]}{\sqrt{x^2(t) + y^2(t)}}\, e^{jf[\sqrt{x^2(t)+y^2(t)}]} \cdot [x(t) + jy(t)] \quad . \tag{12}$$

Since the random variable pairs $(a_k, b_k)$, $k = 0, \pm 2, \ldots$, for different $k$'s are independently and identically distributed, the ensemble average of equation (11) with respect to $(a_k, b_k)$ is a periodic function of time, with period $T$. Therefore, the power spectrum of the TWT output, $e_o(t)$, is

$$W_o(f) = \int_{-\infty}^{\infty} e^{j\omega\tau}\, d\tau\, \frac{1}{T} \int_{-T/2}^{T/2} \left[ \int \frac{1}{2} \{Re\, [H(x_1, y_1)\, H^*(x_2, y_2) \cdot e^{-j\omega_c\tau}]\}\, dF\, (x_1, y_1, x_2, y_2) \right] dt \tag{13}$$

where $F(x_1, y_1, x_2, y_2)$ represents the probability distribution function of $x_1, y_1, x_2, y_2$. Equation (13) can be rewritten as

$$W_o(f) = \frac{1}{4}\, [P(\omega_c - \omega) + P^*(\omega_c + \omega)] \tag{14}$$

where

$$P(\omega) = \frac{1}{T} \int_{-\infty}^{\infty} e^{-j\omega\tau}\, d\tau \left[ \int_{-\infty}^{\infty} V_o(t) \int H(x_1, y_1)\, H^*(x_2, y_2) \cdot dF\, (x_1, y_1, x_2, y_2)\, dt \right] \quad . \tag{15}$$

In equation (15),

$$V_o(t) = 1, \qquad |t| < \frac{T}{2}$$

$$= 0, \qquad |t| \geq \frac{T}{2} \quad . \tag{16}$$



Transformation of the variables $t$ and $t + \tau$ as $t = \eta$ and $t + \tau = \xi$ yields

$$P(\omega) = \frac{1}{T} \int_{-\infty}^{\infty} e^{-j\omega\xi}\, d\xi \int_{-\infty}^{\infty} V_o(\eta)\, e^{j\eta\omega}\, d\eta \left\{ \int H(x_1, y_1) \cdot H^*(x_2, y_2)\, dF\, [x_1(\eta), y_1(\eta), x_2(\xi), y_2(\xi)] \right\} \tag{17}$$

which can be rewritten as

$$P(\omega) = E_S \left[ \frac{1}{T} \left\{ \int_{-\infty}^{\infty} e^{-j\omega\xi}\, H^*[x(\xi), y(\xi)]\, d\xi \right\} \cdot \left\{ \int_{-\infty}^{\infty} V_o(\eta)\, e^{j\omega\eta}\, H[x(\eta), y(\eta)]\, d\eta \right\} \right] \tag{18}$$

where $E_S$ represents the expectation taken on the set $S$ of random variables $(a_k, b_k)$, $k = 0, \pm 1, \pm 2, \ldots$. To make the numerical computation possible, it is assumed that $R(t)$ and $I(t)$ are limited within a finite time interval and become zero outside this interval. Hence, $x(t)$ and $y(t)$ can be represented by

$$x(t) = \sum_{k=-L}^{L} [a_k R(t - kT) - b_k I(t - kT)] \tag{19a}$$

$$y(t) = \sum_{k=-L}^{L} [a_k I(t - kT) + b_k R(t - kT)] \tag{19b}$$

where $L$ is a finite integer. Therefore, $V_o(\eta)\, H[x(\eta), y(\eta)]$ and $H^*[x(\xi), y(\xi)]$ become disjointed with respect to the random variables $(a_k, b_k)$, $k = 0, \pm 1, \pm 2, \ldots$, when

$$|\xi| > \left(2L + \frac{1}{2}\right) T \tag{20}$$

since these two functions have no common $(a_k, b_k)$ under the condition of equation (20). Equation (18) can then be rewritten as

$$P(\omega) = \frac{1}{T} \int_{-\infty}^{\infty} e^{-j\omega\xi}\, E_S\, \{H^*[x(\xi), y(\xi)]\}\, d\xi \int_{-\infty}^{\infty} V_o(\eta)\, e^{j\omega\eta} \cdot E_S\, \{H[x(\eta), y(\eta)]\}\, d\eta, \quad |\xi| > \left(2L + \frac{1}{2}\right) T \quad . \tag{21}$$

Since $H[x(\eta), y(\eta)]$ is equally likely to take negative or positive values, equation (21) is equal to zero. Therefore,

$$P(\omega) = E'_S\left[\frac{1}{T}\left\{\int_{-\infty}^{\infty} V_{2L}(\xi)\, e^{-j\omega\xi}\, H^*[x(\xi), y(\xi)]\, d\xi\right\}\right.$$
$$\left.\cdot\left\{\int_{-\infty}^{\infty} V_o(\eta)\, e^{j\omega\eta}\, H[x(\eta), y(\eta)]\, d\eta\right\}\right] \quad (22)$$

where
$$V_{2L}(\xi) = 1, \qquad |\xi| < \left(2L + \frac{1}{2}\right) T$$
$$= 0, \qquad |\xi| \geq \left(2L + \frac{1}{2}\right) T \quad (23)$$

and
$S'$ = set of the random variables $(a_k, b_k)$, $k = 0, \pm1, \pm2, \ldots, \pm2L$

$E'_S$ = expectation operator representing ensemble averaging over these $(4L + 1)$ random variables.

It should be noted that, in equation (22), the Fourier transform of $H[x(t), y(t)]$ is taken on two different window sizes. That is, $V_o(t)$ has a length $T$ (the PSK signal pulse duration) and $V_{2L}(t)$ has a length $(4L + 1)\,T$. Therefore, $H[x(\eta), y(\eta)]$ includes only $(a_k, b_k)$, $k = 0, \pm1, \pm2, \ldots, \pm L$, although $H^*[x(\xi), y(\xi)]$ includes all of $(a_k, b_k)$, $k = 0, \pm1, \pm2, \ldots, \pm L, \ldots, \pm2L$. Equations (22) and (14) give the power spectrum of PSK signals at the output of a nonlinear TWT. The result obtained here is not obvious from an intuitive analysis and is quite interesting. It is shown in the Appendix that this result reduces to the familiar form if the TWT is linear.

## Computation of the power spectrum spread at the output of the TWT

The procedure for computing the power spectrum spread at the output of the TWT is as follows:

a. Choose a sequence $(a_l, b_l)$ from the set of random variables $(a_k, b_k)$, $k = 0, \pm1, \pm2, \ldots, \pm2L$, and compute $H[x(t), y(t)]$ of equation (12).



b. Compute the Fourier transforms of $V_o(t)\ H[x(t), y(t)]$ and $V_{2L}(t)\ H^*[x(t), y(t)]$ in equation (22) and multiply them together.

c. Choose another sequence $(a_l, b_l)$ from the set of random variables $(a_k, b_k)$, $k = 0, \pm1, \pm2, \ldots, \pm2L$, and perform the same operation performed in step b. Repeat this procedure for other random sequences of $(a_l, b_l)$.

d. Take the average of these products to obtain $P(\omega)$.

e. Compute the power spectrum $W_o(f)$ at the output of the TWT by using equation (14).

It should be noted that, if $M$, the number of phases for the M-ary PSK signal, and $L$, half the number of intersymbol terms considered, are not large, the exact power spectrum can be obtained by considering all combinations of $(a_l, b_l)$. The total number of combinations is $N_c = (1/2)\, M^{4L+1}$, where the factor $1/2$ is used because the sets $(a_l, b_l)$ and $(-a_l, -b_l)$ result in the same $P(\omega)$, as can be seen from equations (6), (12), and (18). For 4-phase PSK, if $L = 1$, the total number of combinations is $N_c = 512$. For $L > 1$, the number of different sequences that must be considered is very large. Therefore, in practice, the sequences are generated randomly by using a uniform random number generator, and $P(\omega)$ of equation (22) is computed by taking the fast Fourier transform (FFT) on two different window sizes and ensemble averaging over all of the available sequences. It has been verified that power spectra based on as few as $N_c = 50$ random sequences are in very good agreement with measured power spectra of 4-phase PSK signals for $L = 3$.

A general computer program has been developed to compute the power spectrum of PSK signals at the output of the TWT based on the analysis outlined here. This computer program consists of two subprograms. The first computes the pulse response of one or more bandpass filters in cascade, and the second computes and plots the power spectrum spread at the output of the TWT and also evaluates the adjacent channel interference caused by the spectrum spreading.

To compute the TWT output spectrum, the pulse response of the low-pass analog of the bandpass filters before the TWT (the cascade of $F_1$ and $F_2$ in Figure 1) must be known. For practical bandpass filters, the amplitude and phase (or group-delay) characteristics are not necessarily symmetrical with respect to the center frequency. Therefore, the impulse response, $h(t)$, of the low-pass analog has both in-phase and quadrature components [1]. The in-phase and quadrature components of the pulse

response are computed by convolving the complex impulse response with the waveform, $W(t)$, of the transmitting pulse. The pulse response depends on the pulse duration, $T$, and hence on the symbol rate, $R_s = 1/T$. The amplitude and group-delay characteristics of filters $F_1$ and $F_2$ are shown in Figures 3 and 4. Figure 5 shows the computed pulse response of these filters for $R_s = 33.3$ megasymbols/second.

The next step is the computation of $x(t)$ and $y(t)$ for $|t| \leq (2L + 1/2)\,T$. The random sequences $(a_k, b_k)$, $k = 0, \pm 1, \pm 2, \ldots, \pm 2L$, are generated by a uniform random number generator. For M-ary PSK the output, $r$, of the random number generator takes values between 0 and $M - 1$. The corresponding phase, $\phi_k$, is $[\pi/M + r(2\pi/M)]$, and $(a_k, b_k)$ are given by equation (3). (It can be assumed that $A = 1$ without any loss of generality.) A total of $(4L + 1)$ pairs of $(a_k, b_k)$ are generated. Then $x(t)$, $y(t)$, and

$$\rho(t) = \sqrt{x^2(t) + y^2(t)} \tag{24}$$

are computed from equation (6).

To compute $H(t)$ of equation (12), the TWT characteristics must be known. The amplitude and phase nonlinearities of the TWT are expressed as one complex nonlinearity, which is approximated by a Bessel function series with complex coefficients [2]:

$$g(\rho) \exp\{jf(\rho)\} = \sum_{s=1}^{L} b_s J_1(s\alpha\rho) \tag{25}$$

where the functions $g(\rho)$ and $f(\rho)$ represent the respective amplitude and phase nonlinearities of the TWT as functions of the input envelope. The complex coefficients $b_s$ are computed from the measured characteristics of the TWT nonlinearity [3]. Also, $J_1(z)$ is the 1st-order Bessel function of argument $z$, $\alpha$ is an arbitrary constant which scales the input level, and $L$ is the number of complex coefficients. (It has been found that $L = 10$ is adequate for fitting the measured data available.) Figure 6 shows the TWT nonlinearities for a typical INTELSAT IV satellite. Table 1 shows the coefficients for the Bessel expansion of this complex nonlinearity as in equation (25). This table also shows the coefficients of a fictitious TWT with no phase nonlinearity (only AM/AM) or no amplitude nonlinearity (only AM/PM).

The spectrum spread of 4-, 8-, and 16-phase PSK signals has been computed at various symbol rates and TWT input backoff values. These



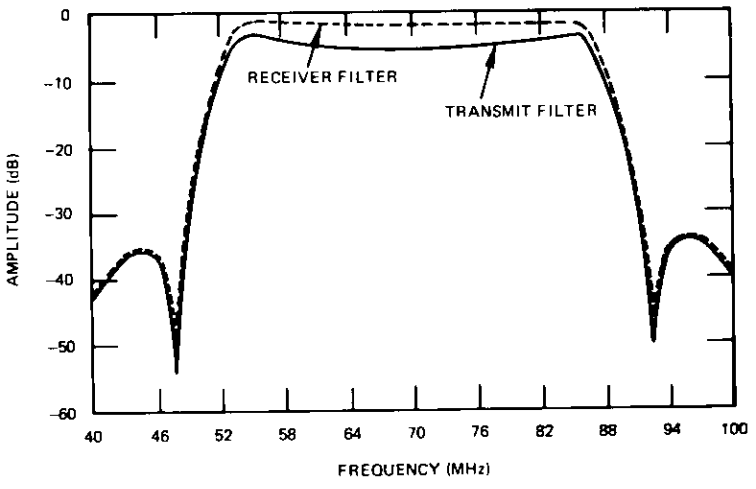



Figure 3a. *Amplitude Characteristics of the Modem (elliptic) Filters (adjusted to 33.3 Mbauds)*

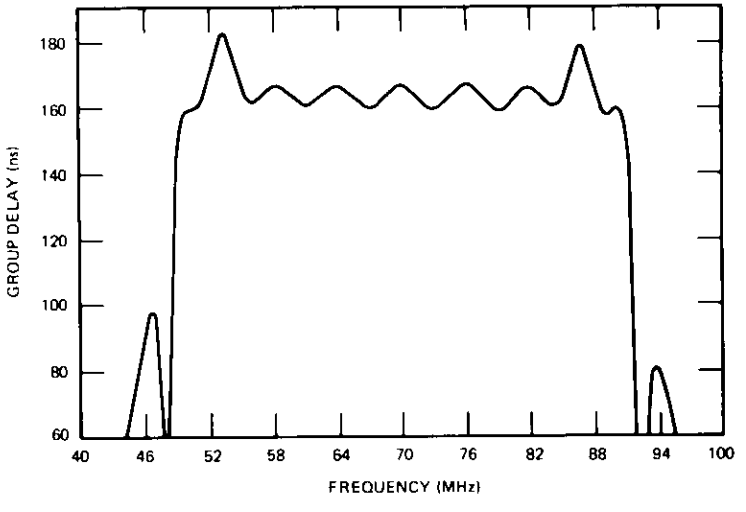



Figure 3b. *Group-Delay Characteristic of Modem Filter $F_1$ in Figure 1 (equalized)*

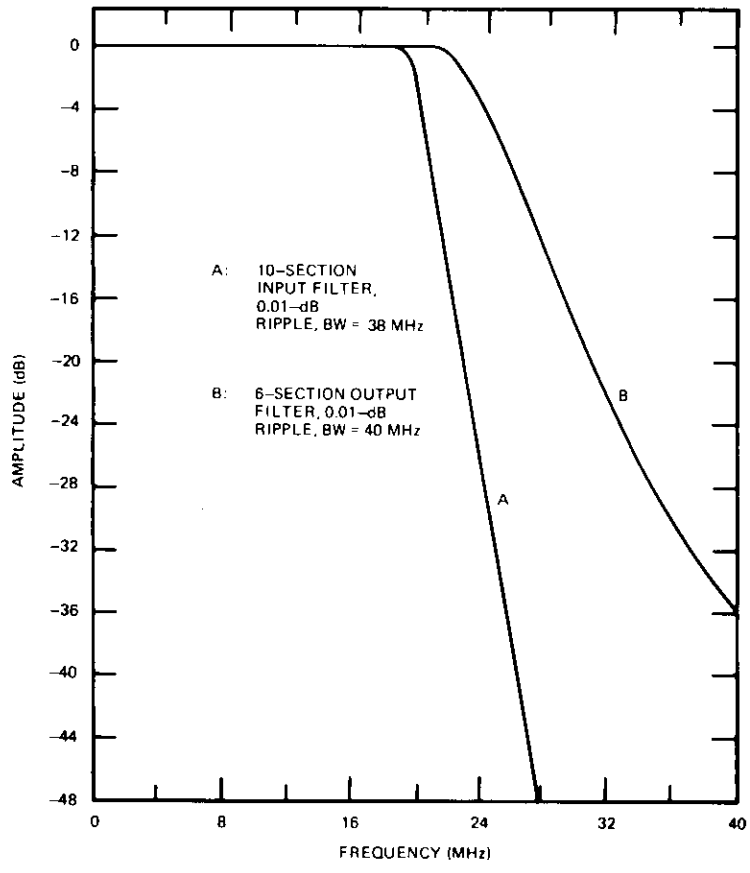



Figure 4a. *Amplitude Characteristics of Satellite Filters*

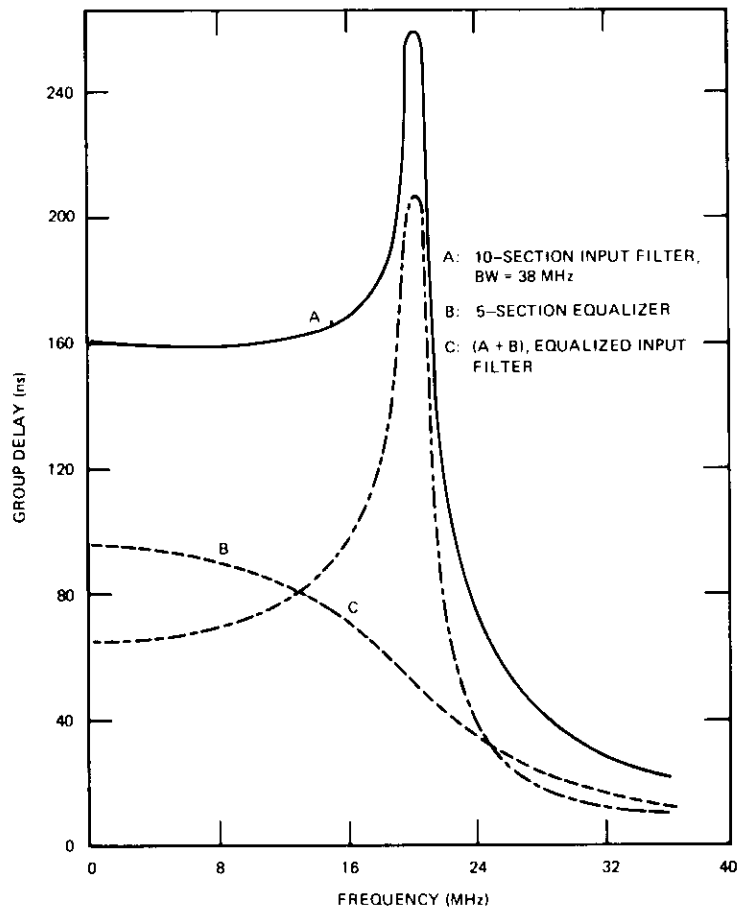



Figure 4b. *Group-Delay Characteristics of Satellite Input Filter*

spectra are shown in Figures 7–9.* All spectra are obtained by averaging over $N_c = 50$ random sequences. The number of intersymbol interference terms considered is 6 ($L = 3$). The case of 0-dB backoff corresponds to the saturation point on the TWT characteristic. Changing the input backoff of the TWT is equivalent to changing $\alpha$ in equation (25). For example, $\alpha = \alpha_0 = 0.7778$ in Table 1 corresponds to the saturation point; changing the value of $\alpha$ to

$$\alpha = \alpha_0 \, 10^{-B_{\text{off}}/20} \tag{26}$$

corresponds to a backoff of $B_{\text{off}}$ in dB.

### *Adjacent channel interference*

To compute the adjacent channel interference caused by the power spectrum spread, the amplitude characteristic of the filters at the output

* In Figures 7–9, 13, and 14, the ordinate represents the power in the 1-MHz frequency band with respect to 0-dB TWT saturation power.

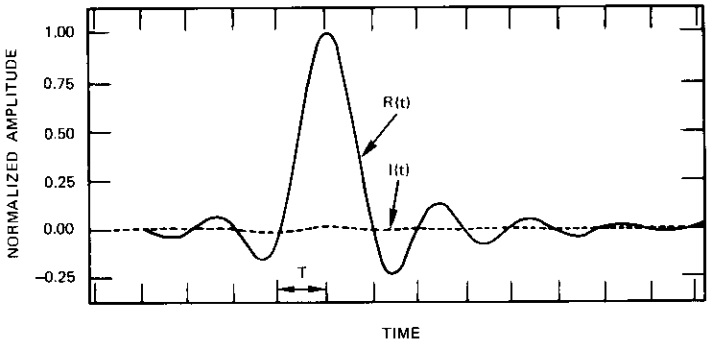



Figure 5. *Pulse Response of Cascaded Filters $F_1$ and $F_2$ in Figure 1 (low-pass analog) for 33.3 Mbauds*

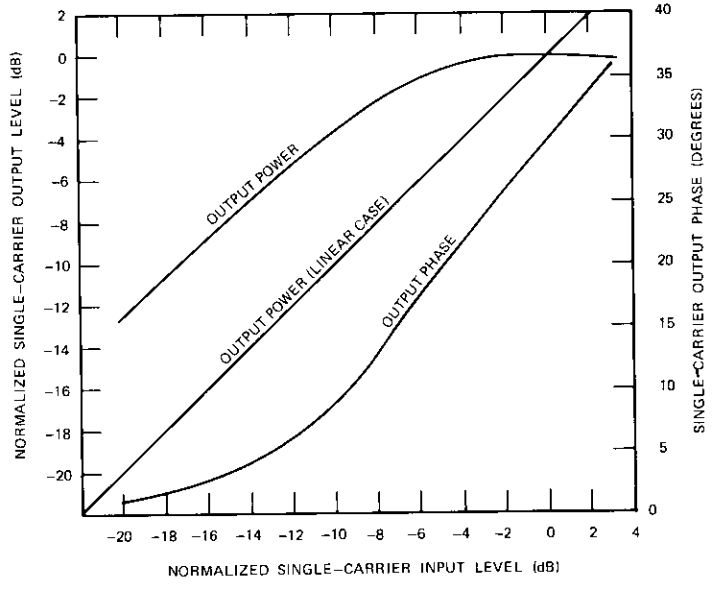



Figure 6. *TWT Nonlinearities*



TABLE 1. COMPLEX BESSEL EXPANSION COEFFICIENTS FOR A TYPICAL TWT ($\alpha = 0.7778$)

| Both AM/AM and AM/PM Conversion | | Only AM/AM Conversion | | Only AM/PM Conversion | |
|---|---|---|---|---|---|
| Re ($b_s$) | Im ($b_s$) | Re ($b_s$) | Im ($b_s$) | Re ($b_s$) | Im ($b_s$) |
| 2.571 | 2.451 | 2.896 | 1.513 | 3.046 | 2.847 |
| 0.0311 | −1.228 | 0.1638 | −1.822 | −0.120 | −1.299 |
| 0.2046 | 0.7837 | 0.0620 | 0.9967 | −0.5593 | 0.3229 |
| 1.007 | −0.2931 | 0.9578 | −0.1430 | 0.6998 | −0.3786 |
| −0.5499 | 0.1939 | −0.4723 | 0.6612 | −0.3750 | 0.4715 |
| 0.1437 | −0.5945 | 0.0892 | −0.5345 | 0.0134 | −0.6516 |
| 0.3895 | 0.6886 | 0.3313 | 0.8352 | 0.1916 | 0.5400 |
| −0.4076 | −0.6207 | −0.3155 | −0.6703 | −0.1857 | −0.3526 |
| 0.2026 | 0.2701 | 0.1268 | 0.3036 | 0.0931 | 0.1345 |
| −0.0535 | −0.0862 | −0.0227 | −0.0628 | −0.0248 | −0.0320 |

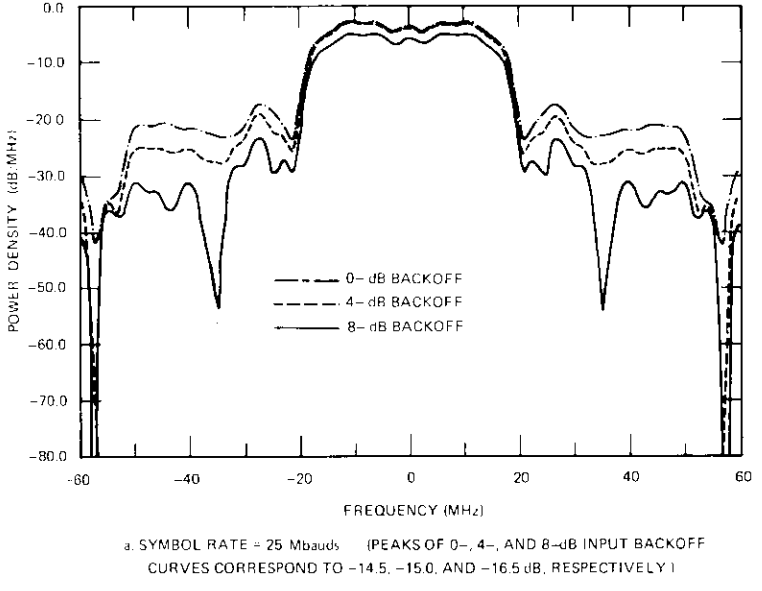
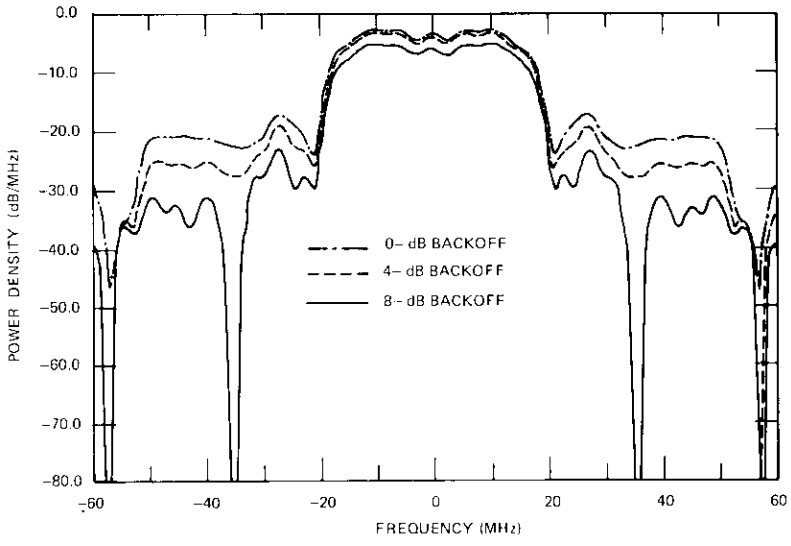
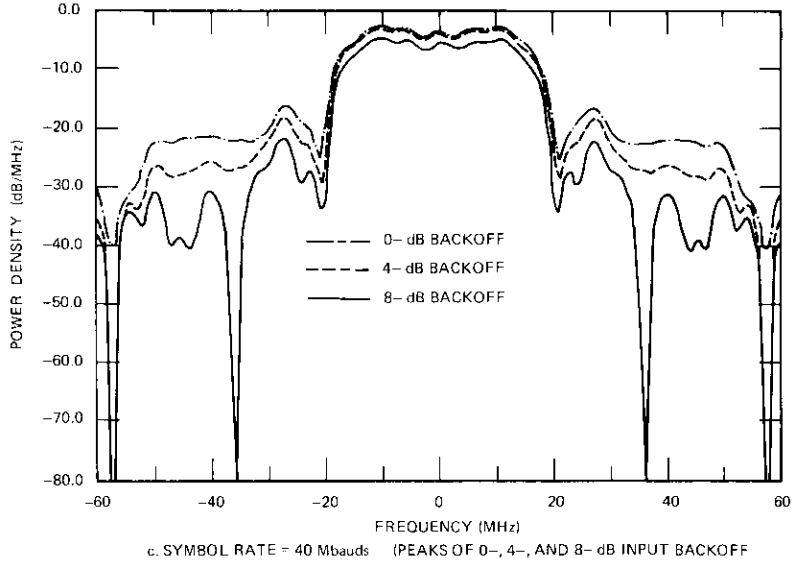



Figure 7. *Power Spectrum of 4-Phase PSK Signal at the Output of the TWT for Various Input Backoffs* (continued)



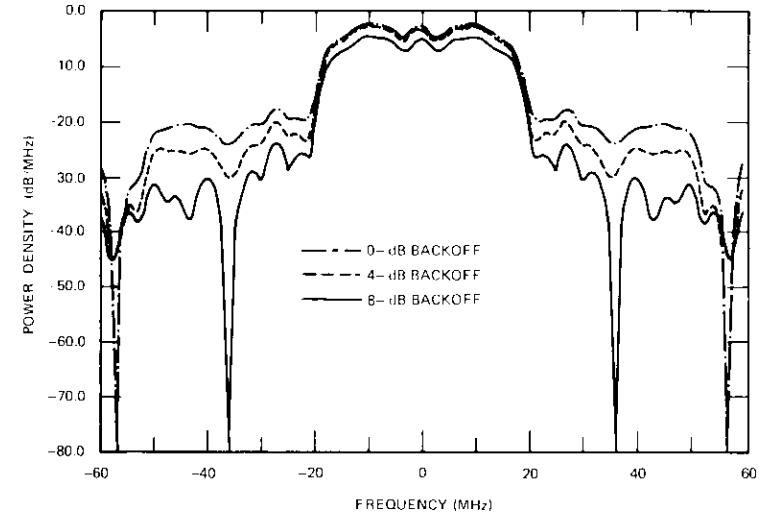

b. SYMBOL RATE 32 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -15.5, -16.0, AND -17.5 dB, RESPECTIVELY.)



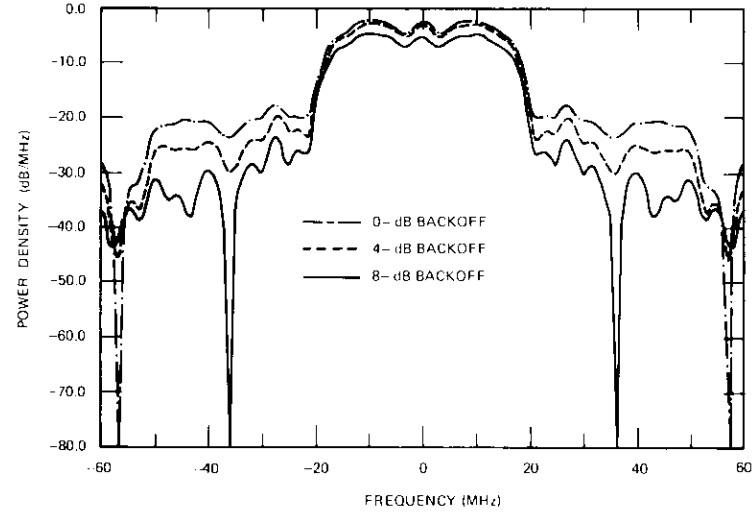

c. SYMBOL RATE = 40 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -16.5, -17.0, AND -18.5 dB, RESPECTIVELY.)

Figure 7 (continued). *Power Spectrum of 4-Phase PSK Signal at the Output of the TWT for Various Input Backoffs*



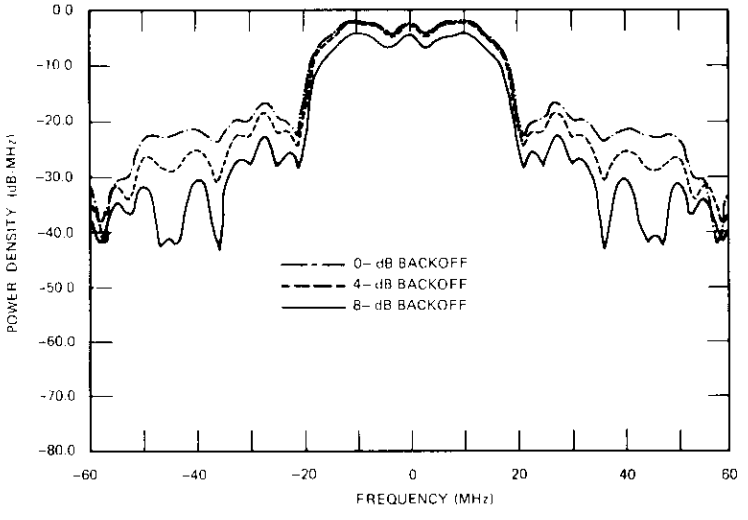



a. SYMBOL RATE - 25 Mbauds (PEAKS OF 0 , 4 , AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -14.5, -15.0, AND -16.5 dB, RESPECTIVELY.)



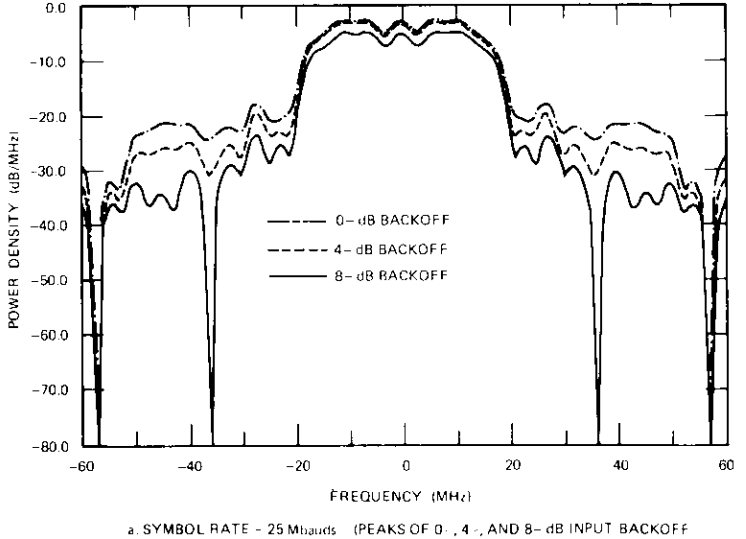

b. SYMBOL RATE 32 Mbauds (PEAKS OF 0 , 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -15.5, -16.0, AND -17.5 dB, RESPECTIVELY.)

Figure 8. *Power Spectrum of 8-Phase PSK Signal at the Output of the TWT for Various Input Backoffs* (continued)




c. SYMBOL RATE = 40 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -16.5, -17.0, AND -18.5 dB, RESPECTIVELY.)

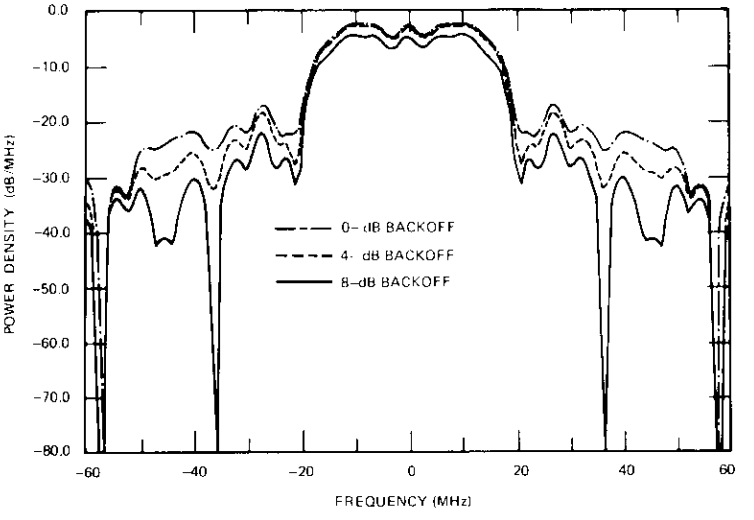
Figure 8 (continued). *Power Spectrum of 8-Phase PSK Signal at the Output of the TWT for Various Input Backoffs*



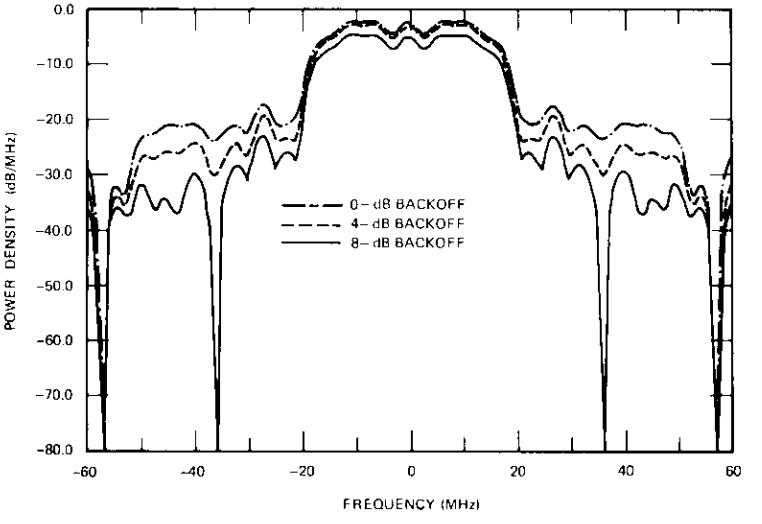

a. SYMBOL RATE = 25 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -14.5, -15.0, AND -16.5 dB, RESPECTIVELY.)

Figure 9. *Power Spectrum of 16-Phase PSK Signal at the Output of the TWT for Various Input Backoffs* (continued)






c. SYMBOL RATE = 40 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -16.5, -17.0, AND -18.5 dB, RESPECTIVELY.)




b. SYMBOL RATE = 32 Mbauds (PEAKS OF 0-, 4-, AND 8- dB INPUT BACKOFF CURVES CORRESPOND TO -15.5, -16.0, AND -17.5 dB, RESPECTIVELY.)

Figure 9c. (continued) *Power Spectrum of 16-Phase PSK Signal at the Output of the TWT for Input Backoffs*

of the TWT must be known. The ratio of the signal power to the interference power in the adjacent channel is given by

$$\frac{\int_{-B/2}^{B/2} W_o(f)\ |F_3(f)|^2\ |F_4(f)|^2\ df}{\int_{\Delta-B/2}^{\Delta+B/2} W_o(f)\ |F_3(f)|^2\ |F_4(f-\Delta)|^2\ df} \tag{27}$$

where $W_o(f)$ = power spectrum at the output of the TWT

$|F_3(f)|$, $|F_4(f)|$ = amplitude characteristics of the filters $F_3$ and $F_4$ in Figure 1, respectively

$\Delta$ = distance between the centers of the adjacent channel and the main channel

$B$ = bandwidth of the main channel.

The inverse of equation (27), in dB, gives the carrier-to-adjacent-channel interference power ratio, $C/I_A$.

The amplitude characteristics of $F_3$ and $F_4$ are given in Figures 3a and 4a, respectively. Table 2 gives the carrier-to-adjacent-channel interference power ratio at various symbol rates and input backoffs for 4-, 8-, and 16-phase PSK based upon the assumptions that the bandwidths of the

TABLE 2. CARRIER-TO-ADJACENT-CHANNEL INTERFERENCE POWER RATIO (IN dB) FOR INTELSAT IV WITH 4-, 8-, AND 16-PHASE PSK SIGNALS

| Number of Phases | Input Backoff (dB) | Symbol Rate (Mbauds) | | |
|---|---|---|---|---|
| | | 25 | 32 | 40 |
| 4 | 0 | 40.24 | 29.05 | 15.42 |
| | 4 | 44.47 | 31.23 | 15.90 |
| | 8 | 51.46 | 34.28 | 16.39 |
| 8 | 0 | 41.10 | 29.93 | 15.34 |
| | 4 | 45.26 | 31.82 | 15.64 |
| | 8 | 50.87 | 34.77 | 16.09 |
| 16 | 0 | 41.37 | 30.45 | 15.90 |
| | 4 | 45.40 | 32.05 | 16.14 |
| | 8 | 50.41 | 34.39 | 16.50 |



main channel and adjacent channel are identical (B = 40 MHz) and that the channel separation, $\Delta$ = 40 MHz.

### *Comparison of the computed and measured spectra*

The computed TWT output spectra have been compared with the measured spectra of 4-phase PSK signals at one symbol rate (25 Mbauds or 50 Mbps) and at various input backoff values. Figure 10 is a block diagram of the measurement setup. Figure 11 shows the measured amplitude and

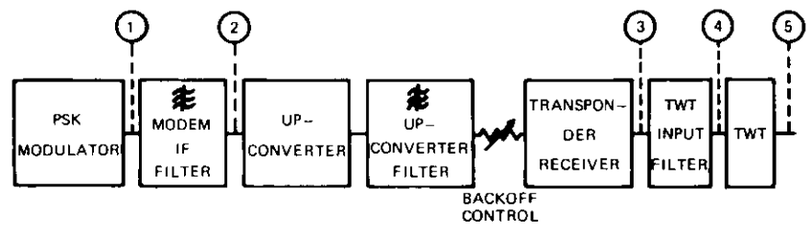



Figure 10. *Block Diagram of the Measurement Setup*

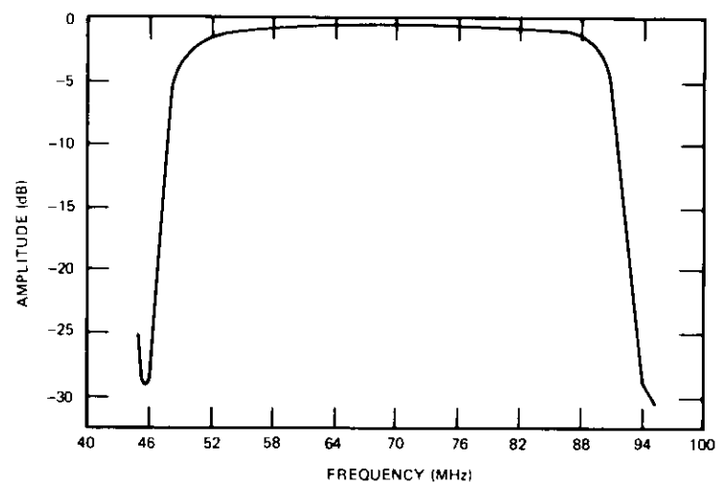



Figure 11a. *Amplitude Characteristic of the Filter Chain in Figure 10*

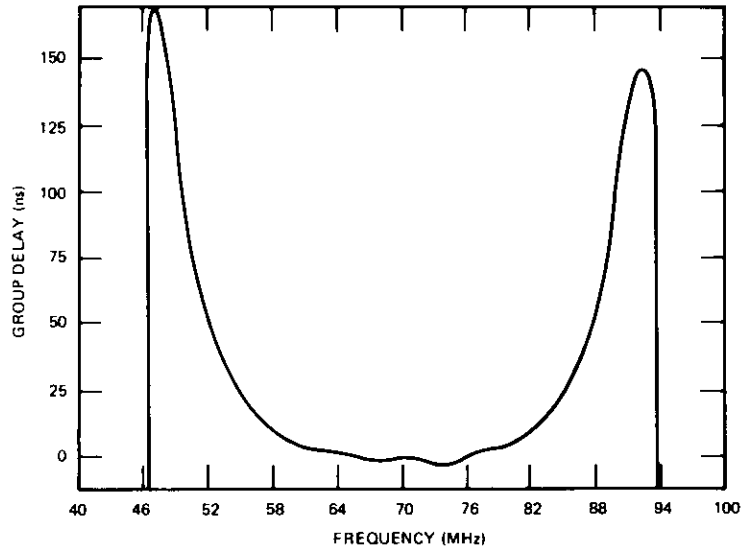



Figure 11b. *Group-Delay Characteristic of the Filter Chain in Figure 10*

group-delay characteristics of the overall input chain between points 1 and 4 in Figure 10. The computed pulse response of the filters between 1 and 4 is shown in Figure 12. Figure 13 shows the computed and measured power spectra at various input backoff values. The input backoff values chosen are −4, −2, 0, 4, 8, and 12 dB. In all cases, the computed spectra

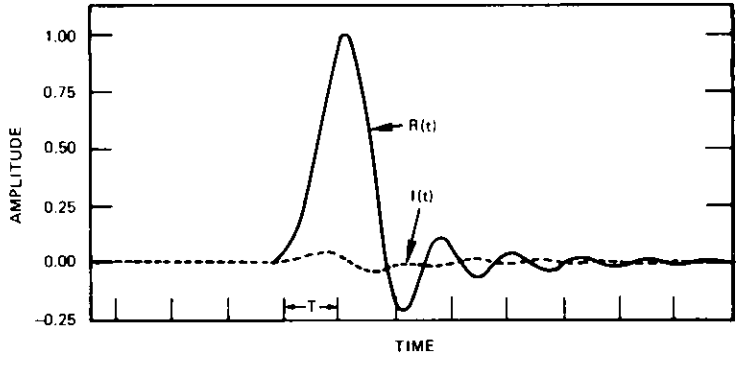
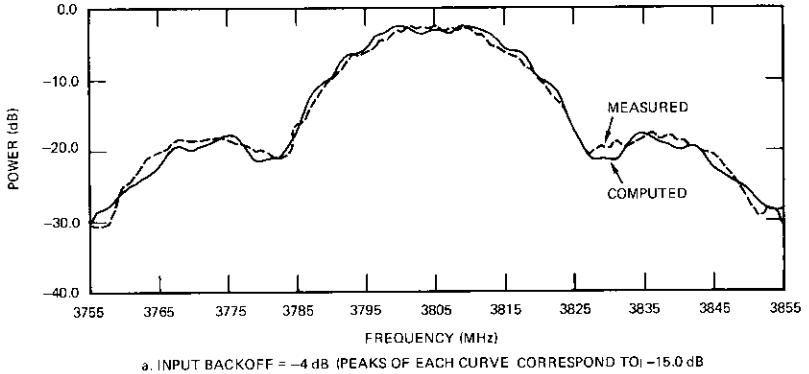


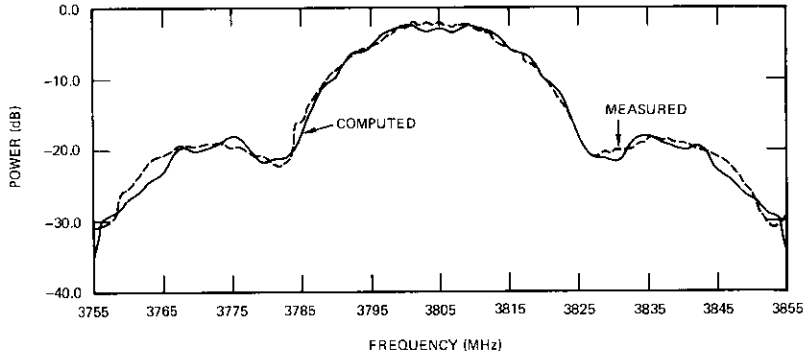

Figure 12. *Computed Pulse Response of the Filter Chain in Figure 10*



are obtained by ensemble averaging over $N_e = 50$ random sequences. Figure 13c also shows the spectrum at 0-dB input backoff, computed by averaging over $N_e = 100$ random sequences. Comparison of the two curves indicates that averaging over about 50 sequences is adequate. In all cases, the measured and computed spectra are in good agreement.




a. INPUT BACKOFF = −4 dB (PEAKS OF EACH CURVE CORRESPOND TO −15.0 dB RELATIVE TO AN UNMODULATED SINUSOIDAL WAVE AT THE SAME INPUT BACKOFF.)

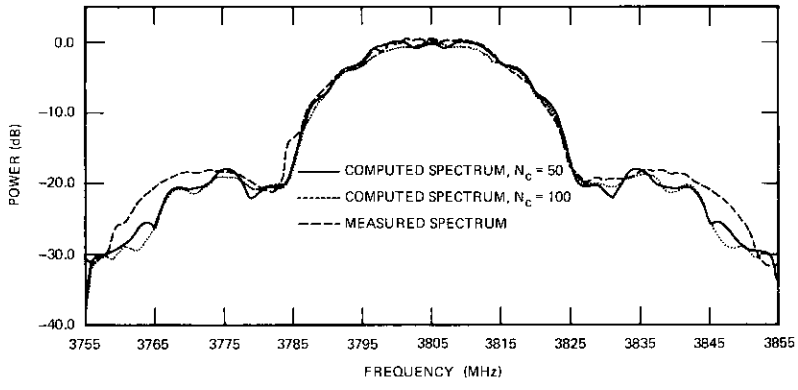



b. INPUT BACKOFF = −2 dB (PEAKS OF EACH CURVE CORRESPOND TO −14.6 dB RELATIVE TO AN UNMODULATED CW AT THE SAME INPUT BACKOFF.)

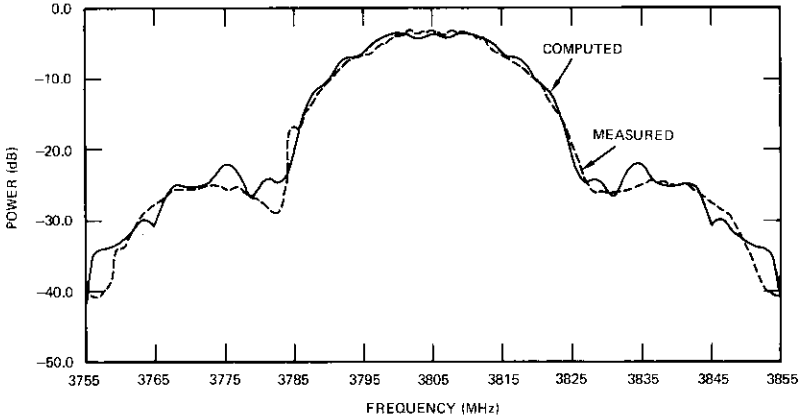
Figure 13. (continued) *Comparison of the Computed and Measured Power Spectra of 4-Phase PSK Signals at the Output of the TWT (bit rate = 50 Mbps)* (continued)

An additional comparison indicates that the computed results in Figure 7a are in good agreement with field tests over the INTELSAT IV-F3 satellite [4].

The three curves in Figure 14 show the TWT output power spectrum




c. INPUT BACKOFF = 0 dB (PEAKS OF EACH CURVE CORRESPOND TO −14.5 dB RELATIVE TO AN UNMODULATED CW AT THE SAME INPUT BACKOFF.)

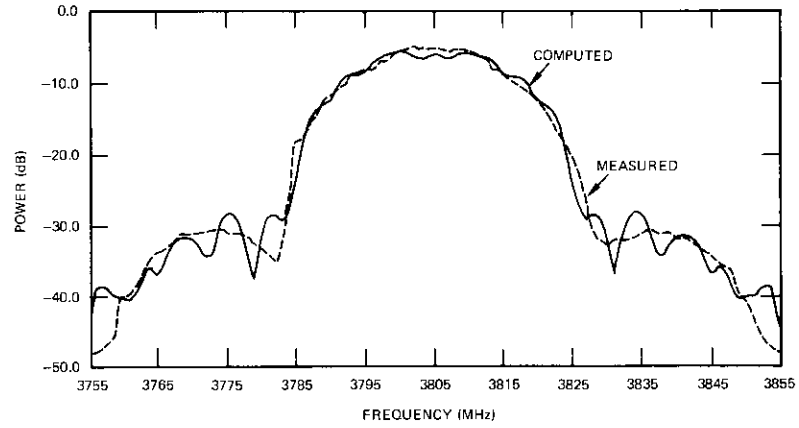



d. INPUT BACKOFF = 4 dB (PEAKS OF EACH CURVE CORRESPOND TO −15.0 dB RELATIVE TO AN UNMODULATED CW AT THE SAME INPUT BACKOFF.)

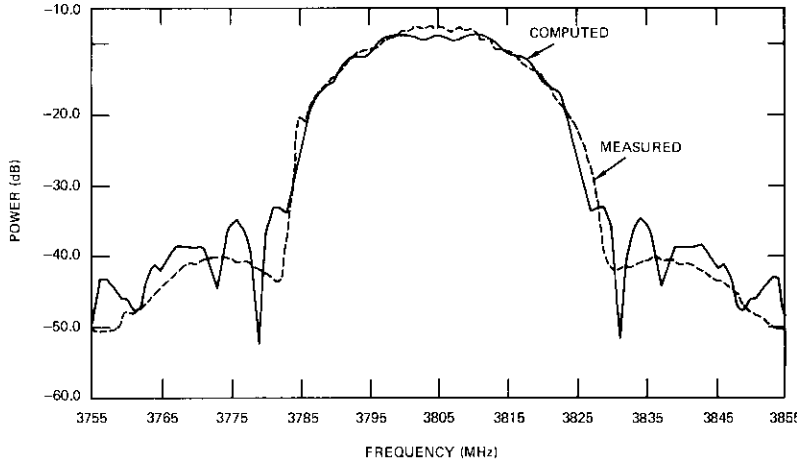
Figure 13 (continued). *Comparison of the Computed and Measured Power Spectra of 4-Phase PSK Signals at the Output of the TWT (bit rate = 50 Mbps)* (continued)



with both amplitude and phase nonlinearities, no phase nonlinearity (only AM/AM), and no amplitude nonlinearity (only AM/PM). These




e. INPUT BACKOFF = 8 dB (PEAKS OF EACH CURVE CORRESPOND TO −16.5 dB RELATIVE TO AN UNMODULATED CW AT THE SAME INPUT BACKOFF.)




f. INPUT BACKOFF = 12 dB (PEAKS OF EACH CURVE CORRESPOND TO −17.3 dB RELATIVE TO AN UNMODULATED CW AT THE SAME INPUT BACKOFF.)

Figure 13 (continued). *Comparison of the Computed and Measured Power Spectra of 4-Phase PSK Signals at the Output of the TWT (bit rate = 50 Mbps)*

spectra are computed by using the Bessel expansion coefficients of Table 1. Comparison of the curves shows that the spreading is caused primarily by the amplitude nonlinearity rather than the phase nonlinearity of the TWT amplifier.

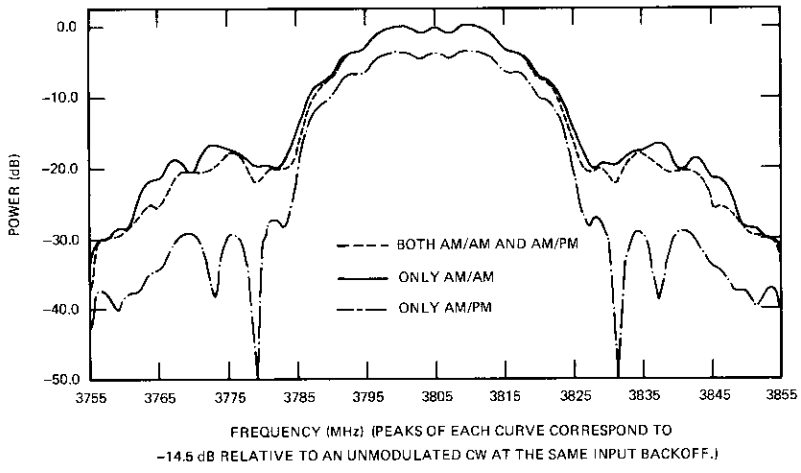



Figure 14. *Computed Power Spectrum of 4-Phase PSK Signal for Various Types of TWT Nonlinearities*


## Acknowledgment

*The authors would like to thank Mr. John Puente who pointed out the necessity of analytical work on the problem of power spectrum spreading of PSK signals. Special thanks are due to Miss Anne Bernstein who put the computer programs developed for this study into a general form which can be used in the performance evaluation of PSK systems. The measured data presented in Figure 13 were contributed by Dr. B. Pontano, whose effort is very much appreciated.*




## References


[1] A. Papoulis, *The Fourier Integral and Its Applications*, New York: McGraw-Hill, March 1962, pp. 131–132.

[2] O. Shimbo, "Effects of Intermodulation, AM-PM Conversion and Additive Noise in Multicarrier TWT Systems," *Proceedings of the IEEE*, Vol. 59, No. 2, February 1971, pp. 230–238.

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

[4] D. Chakraborty and L. Golding, private communications.


## Appendix A. Verification of the spectrum computation formula

To verify the result obtained in equation (22) of the paper, it will be shown that the result obtained for a linear TWT is the same as that obtained previously.

It is assumed that the random variables $(a_k, b_k)$, $k = 0, \pm 1, \pm 2, \ldots$, are mutually independent. In this case, equation (11) is rewritten as $H[x(t), y(t)] = x(t) + jy(t)$, and equation (18) becomes

$$P(\omega) = E_S\left\{\frac{1}{T}\int_{-\infty}^{\infty} e^{-j\omega\xi}\,[x(\xi) - jy(\xi)]\,d\xi \cdot \int_{-\infty}^{\infty} [x(\eta) + jy(\eta)]\,e^{j\omega\eta}\,V_o(\eta)\,d\eta\right\} . \quad \text{(A1)}$$

Since

$$E_S\{[x(\xi) - jy(\xi)]\,[x(\eta) + jy(\eta)]\} = \sum_{k=-\infty}^{\infty} [R(\eta - kT)\,R(\xi - kT) + I(\eta - kT)\,I(\xi - kT)] + j\sum_{k=-\infty}^{\infty} [R(\xi - kT)\,I(\eta - kT) - R(\eta - kT)\,I(\xi - kT)] \quad \text{(A2)}$$

equation (A1) can be written as

$$P(\omega) = \sum_{k=-\infty}^{\infty} \frac{1}{T} \int_{-\infty}^{\infty} R(\xi - kT)\ e^{-j\omega\xi}\, d\xi \int_{-\infty}^{\infty} V_o(\eta)\ R(\eta - kT)$$

$$\bullet\ e^{j\omega\eta}\, d\eta + \sum_{k=-\infty}^{\infty} \frac{1}{T} \int_{-\infty}^{\infty} I(\xi - kT)\ e^{-j\omega\xi}\, d\xi$$

$$\bullet \int_{-\infty}^{\infty} V_o(\eta)\ I(\eta - kT)\ e^{j\omega\eta}\, d\eta + j \sum_{k=-\infty}^{\infty} \frac{1}{T} \int_{-\infty}^{\infty} R(\xi - kT)$$

$$\bullet\ e^{-j\omega\xi}\, d\xi \int_{-\infty}^{\infty} V_o(\eta)\ I(\eta - kT)\ e^{j\omega\eta}\, d\eta$$

$$- j \sum_{k=-\infty}^{\infty} \frac{1}{T} \int_{-\infty}^{\infty} I(\xi - kT)\ e^{-j\omega\xi}\, d\xi$$

$$\bullet \int_{-\infty}^{\infty} V_o(\eta)\ R(\eta - kT)\ e^{j\omega\eta}\, d\eta \quad . \tag{A3}$$

The first sum of equation (A3) can be rewritten as

$$\frac{1}{T} \int_{-\infty}^{\infty} V_o(\eta) \left[ \sum_{k=-\infty}^{\infty} R(\eta - kT)\ e^{j\omega(\eta - kT)} \right] d\eta \int_{-\infty}^{\infty} R(\xi)\ e^{-j\omega\xi}\, d\xi \quad . \tag{A4}$$

From Poisson's formula,

$$\sum_{k=-\infty}^{\infty} R(\eta - kT)\ e^{j\omega(\eta - kT)} = \frac{1}{T} \sum_{k=-\infty}^{\infty} Q\left(\frac{k}{T}\right) e^{j(2\pi/T)k\eta} \tag{A5}$$

where

$$Q(x) = \int_{-\infty}^{\infty} R(\eta)\ e^{j\omega\eta}\, e^{-j\eta x}\, d\eta \quad . \tag{A6}$$

Thus,

$$\int_{-\infty}^{\infty} V_o(\eta) \left[ \frac{1}{T} \sum_{k=-\infty}^{\infty} Q\left(\frac{k}{T}\right) e^{j(2\pi/T)k\eta} \right] d\eta$$

$$= \frac{1}{T} \int_{-T/2}^{T/2} Q(0)\, d\eta = Q(0) = \int_{-\infty}^{\infty} R(\eta)\ e^{j\omega\eta}\, d\eta \quad . \tag{A7}$$



Substituting equation (A7) into equation (A4) yields the first sum of equation (A3):

$$\frac{1}{T} \int_{-\infty}^{\infty} R(\eta)\ e^{j\omega\eta}\, d\eta \int_{-\infty}^{\infty} R(\xi)\ e^{-j\omega\xi}\, d\xi$$

$$= \frac{1}{T} \left| \int_{-\infty}^{\infty} R(t)\ e^{j\omega t}\, dt \right|^2 \quad . \tag{A8}$$

Similarly, the second sum of equation (A3) is $(1/T) \left| \int_{-\infty}^{\infty} I(t) \bullet e^{j\omega t}\, dt \right|^2$ and the remainder of equation (A3) will equal zero. This is the answer expected for a linear TWT.

*Guner S. Robinson was born in Nazilli, Turkey. She received an M.S.E.E. from Istanbul Technical University in 1961. On a Fulbright Scholarship, she received a Ph.D. in Electrical Engineering from the Polytechnic Institute of Brooklyn in 1966 where she became a Research Fellow and a Graduate Assistant.*

*From 1961 to 1968 she was associated with the Department of Electrical Engineering of the Middle East Technical University, Ankara, Turkey. In December 1968, she joined* COMSAT *Laboratories, where she was a member of the Technical Staffs of the Communications Processing and Advanced Studies Laboratories. Since February 1973, she has been a Research Associate at the University of Southern California.*

*Osamu Shimbo, formerly Senior Scientist in the Advanced Studies Laboratory of* COMSAT *Laboratories, holds a Bachelor of Engineering degree from Tohoku University, Japan (1956), and a Doctor of Engineering degree from Hokkaido University of Japan (1965). Before joining* COMSAT, *he was employed by Northern Electric Co., Ltd., General Electric Co., Ltd., England; and Oki Electric Industrial Co., Ltd., Japan. He was an exchange visitor in the Department of Electrical Engineering, Columbia University. He has received awards for achievement in FM and PCM systems analysis.*

*Russell J. F. Fang received a B.S. in Electrical Engineering, National Taiwan University, in 1962 and M.S. and Ph.D. degrees, Stanford University, in 1964 and 1968. Current interests include interference and signal processing problems, feedback communications schemes, and high information rate encoding techniques. Before joining* COMSAT *in 1968, he was employed by Stanford Electronics Laboratories; Stanford Research Institute, Menlo Park, California; and the Chinese Air Force Electronics Division, Pintung, Taiwan. He is a member of IEEE and Institute of Mathematical Statistics.*

**Index: telecommunications, digital techniques, communications satellites, data transmission, time-division multiplexing, phase shift keying.**

# The application of TDMA to the Intelsat IV satellite series

WILLIAM G. SCHMIDT


## Abstract

An extensive series of time-division multiple-access (TDMA) experimental programs has established confidence in the technical feasibility of this transmission technique and the advantages which can be derived from it. This paper outlines the elements of a prototype TDMA system conceptually designed for use with INTELSAT IV satellites and for possible introduction into operational networks.

The terrestrial interface with the TDMA terminal, and the transmit- and receive-side equipment, including the 4-phase PSK modem operating at 60 Mbps, are discussed. The control section relies heavily upon a small processor which performs a number of critical operations. To ensure reliable operation, all common terminal equipment is redundant, and continually monitored "hot standby" capability is provided. The modularity of the TDMA terminal design will permit the operational lifetime to extend through the INTELSAT V era.


## Introduction

In TDMA, nonoverlapping bursts of a carrier are transmitted through the satellite repeater so that many signals can share a common repeater.

This paper is based upon work performed in COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

Such a technique offers a number of advantages over multicarrier FDMA; specifically, it provides a significant increase in capacity and operational flexibility, and it requires much less up-link power control. It also promises service flexibility and cost increments more closely associated with a per-satellite base than a per-link base.

COMSAT began investigations in this area in 1964. Shortly thereafter, the MATE TDMA terminals were developed under INTELSAT sponsorship for use in the first experimental TDMA network operation over the INTELSAT I satellite in 1966 [1]. Other experimental TDMA systems were subsequently tested [2]–[5] or proposed [6], [7]. As a result of this work, the technical feasibility of TDMA has been established and its introduction into INTELSAT's operational networks is being considered. Before this is done, however, certain economic, operational, and technical aspects must be considered. To assist in this effort and to consolidate the various viewpoints into a single body of information, INTELSAT has established a TDMA Working Group, which has met 11 times over a period of more than two years.

One of the most important results of these meetings has been the specification of a prototype TDMA system that will not only be used for field trials with the INTELSAT IV class of satellite, but will also provide a firm design base for operational TDMA terminals with INTELSAT IV satellites and with later generations of INTELSAT satellites. Although INTELSAT is still studying the prototype TDMA specification and considering the approval of a field trial program, the general approach and some of the probable features of the system are of interest.

## *System design objectives and features*

The prototype TDMA system design is intended to provide a service quality and reliability which are equal or superior to those provided by existing FM/FDMA equipment, while yielding significantly higher channel capacities and operational flexibility. The range of services to be accommodated is also enlarged and the latest advances in digital solid-state circuitry will be used to provide these services as economically as possible.

The prototype TDMA system is characterized by a number of features which mark the difference between experimental models and operationally oriented designs. The frame length of this system is 750 microseconds, which is six times the voice channel PCM sampling period of 125 microseconds. This longer frame was chosen for a number of reasons:



a. A longer frame period lowers the TDMA transmission overhead, thereby increasing the system capacity [6]. This capacity increase (which will provide added space segment revenue) is traded off against the cost of additional storage at each terminal.

b. The use of burst-operated high-speed PCM codecs which have output rates equal to the TDMA bit rate is economically attractive. However, past experience with such high-speed codecs at 50 Mbps has indicated that the transient response required of the sample-and-hold circuits in connection with the extremely wideband analog signaling channels yields unacceptable quantization distortion and excessive crosstalk [3], [4]. Since it is likely that TDMA systems will be operated at 500 Mbps during this decade, the difficulties encountered by the burst-operated codecs will be significantly compounded. The use of continuous, low-speed codecs (less than 10 Mbps) with rate-changing buffers will provide the high-quality PCM characteristics required for satellite use, since these same codecs are used for the terrestrial digital transmission systems.

c. To ensure compatible system operation with either 4- or 8-phase PSK modulation, the ratio of TDMA frame to PCM frame must be an integer multiple of three. A value of six was chosen.

In most experimental systems the burst of a preselected station has been used as the reference burst position. If a power failure occurs at that station, another station will replace the reference burst by moving its own burst into the empty reference burst position. All of the bursts in these systems contain traffic data.

In the proposed system, the burst synchronization reference is a special burst, called a "sync burst," which carries no traffic. The traffic-bearing bursts, called data bursts, do not move from their preassigned positions in the frame structure. The sync burst is transmitted as a separate burst by one of a group of previously selected reference stations. Not only does this simplify reference station handover procedures, but it also enables the TDMA system to have a common frame synchronization across the multiple repeaters of an INTELSAT IV satellite. This technique enables a single terminal to operate through several repeaters within one TDMA frame (transponder hopping). Other features of this system are as follows:

a. a high degree of modularity, which will permit its use beyond the INTELSAT IV era;

b. the ability to work in preassignment and time-assignment modes of operation;

c. the ability to process multidestinational subbursts;

d. the ability to provide service over a wide range of terrestrial input signals;

e. the use of minicomputers to provide preprogrammed responses in the control subsystem and flexible monitoring and redundant switching functions; and

f. nonvolatile storage of operating data to ensure minimum interruption of service in case of power system outages.

### *Burst formats*

As indicated earlier, two types of bursts are envisioned for use in the proposed system. The first type, the sync burst, is transmitted by the earth station serving as the primary reference for positional tracking and synchronization. There is no more than one sync burst per repeater. All other bursts in a repeater are data bursts, which carry the normal PCM-coded traffic. The detailed structure of these bursts is shown in Figure 1.

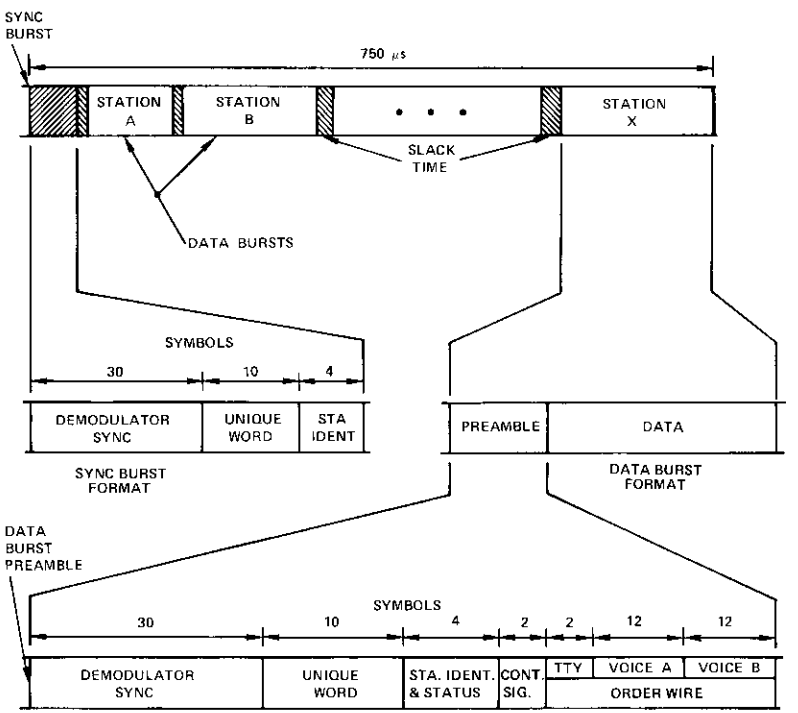



Figure 1. *Burst Formats*



The sync burst, which is the frame positional reference, starts with a 30-symbol pattern of bits which enables the 4-phase coherent PSK demodulators to acquire the carrier reference and regenerate the symbol timing. The next pattern is a 10-symbol (20-bit) correlation codeword (unique word), whose detection identifies the sync burst and uniquely locates the symbol slot, which is generally used as a TDMA frame starting point. This correlation pulse is the basis for the burst synchronization technique employed in the system [8], [9]. The sync burst concludes with a 4-symbol (8-bit) codeword which identifies the transmitting station.

The data bursts are enclosed in a guard time cushion of at least 12 symbols; at least six symbols at the beginning of the burst and six symbols at the end of the burst are allotted to burst allocation. At 60 Mbps, this represents a worst-case burst synchronization tolerance of $\pm 200$ ns, which is much looser than that allowed in most of the recent experimental 50-Mbps TDMA systems [3]–[5]. However, it is consistent with the more conservative design approach which operational systems must embody. The guard time is also a component of the slack time, which is the total unused time between bursts. The burst signal itself starts with a 30-symbol demodulator synchronization pattern followed by a 10-symbol unique word. A 4-symbol (8-bit) codeword identifies the transmitting station in six bits, and two additional bits convey the reference station status of that station. There is one primary reference station (which transmits the sync burst) and two hot standby reference stations, which will replace the primary reference station if it has an outage.

Another 2-symbol time slot in this preamble is devoted to conveying control signaling, which includes the following signaling functions:

a. primary reference station and sync burst replacement,

b. common channel order-wire signaling, and

c. positional data feedback in spot-beam transmission situations if desired.

These signals are contained in 40-bit data blocks which are coded for error detection purposes. These blocks are formed over 10 preambles, and the framing synchronization is provided by transmitting the complement of the unique word every tenth frame.

The remaining 22 symbols of the preamble are dedicated to the normal order-wire requirements of the INTELSAT network, which consists of both teletype and band-limited voice services. Multiple 50-baud service is multiplexed into a 2-symbol slot, and two voice channels are provided in two 12-symbol slots at the end of the preamble. Each voice channel is digitally coded by using 32-kbps delta modulation.

### *Terminal configuration*

Figure 2 is a general block diagram of the terminal equipment. The TDMA terminal has three principal subsystems: TDMA interface modules, common TDMA terminal equipment, and test and maintenance center. The TDMA interface modules format the terrestrial signals or their components into a flexible, efficient block format which forms the basic input to the common TDMA terminal. Because of the large variety of terrestrial signals, as well as the wide range of application and traffic situations which are in many ways unique to a particular station, a large class of optional modules could be specified. However, one station ordinarily would employ only a few of these modules.

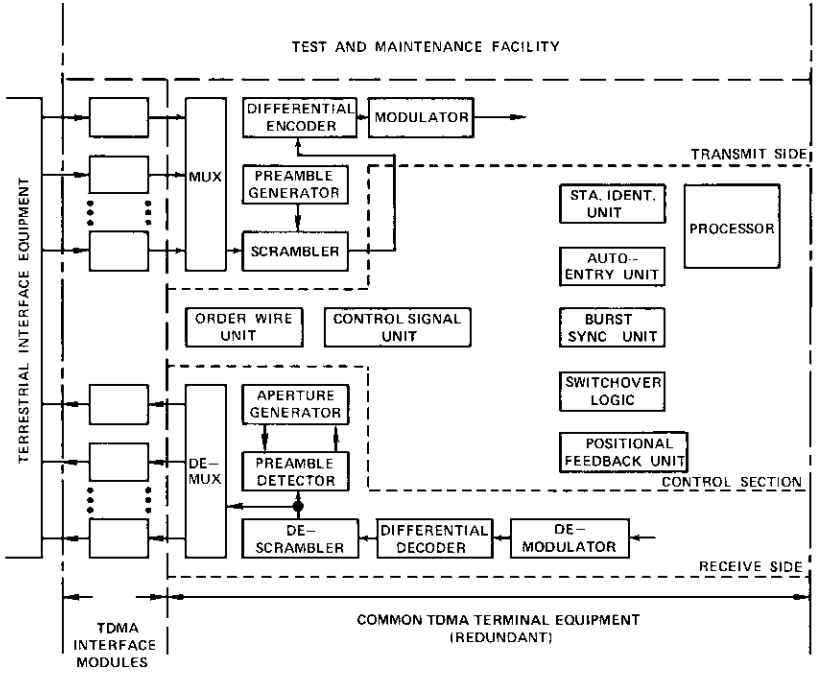



Figure 2. *Block Diagram of the TDMA Terminal Equipment*



The common TDMA terminal consists of three major subsystems: the transmit-side subsystem, the receive-side subsystem, and the control subsystem. The terminal equipment reliability is of prime importance to overall system operation; hence this equipment is completely redundant.

The test and maintenance center helps to troubleshoot and repair off-line common equipment and interface modules. It is envisioned that programmed diagnostic testing capability, which can exercise the common TDMA terminal equipment, each subsystem within a terminal, and each interface module, will be provided. Programmed fault isolation will be provided down to the subsystem level, and in many cases, down to the module level. Beyond that point it will not be practical to have programmed isolation; therefore, troubleshooting and repair will be done by skilled technicians in the more conventional fashion.

#### TDMA interface modules

The inputs which can be employed by PCM/TDMA systems vary considerably in content (i.e., voice, data, visual telephone, or FDM assemblies), mode of operation [i.e., preassignment, demand assignment, or digital speech interpolation (DSI)], bit rate (currently from 1.544 to 8.448 Mbps), and PCM standard (two at the primary level, and currently five at the secondary level). There must be some means of processing these inputs into subbursts which may be easily multiplexed and demultiplexed.

A set of modules, called TDMA interface modules (TIMs), is suggested for this function, in which the terrestrial signals are processed into a format which is optimized for satellite applications. These modules generally have two parts: the first converts the analog signals into digital signals via PCM encoding on the transmit side (and performs the inverse operation on the receive side), and the second converts these digital signals from continuous signals to burst signals by using compression buffers and performs the inverse conversion via expansion buffers. If the terrestrial transmissions are already digital, the PCM equipment is not required, but the compression/expansion buffers may be enlarged to provide the elastic store function associated with pulse stuffing techniques.

There are generally three classes of PCM equipment. The first, called the PCM codec (*co*der-*dec*oder), converts a single analog signal into a PCM bit stream and also performs the inverse operation. The second is the PCM multiplex. This unit individually converts multiple analog signals into digital form and then multiplexes these digital signals into one continuous bit stream. The third class of PCM equipment is the digital multiplex, which multiplexes multiple digital bit streams into one bit stream with a higher bit

rate. A PCM multiplex which has an output bit stream at an approved bit rate corresponding to the "primary" level of a national/international network may be called a primary multiplex. Secondary multiplexes may be of the "digital" or "PCM" variety.

The TIMs fall into three broad categories according to the interface signal being presented to the TDMA terminal:

a. voice frequency interface modules,
b. FDM assembly interface modules, and
c. direct digital interface (DDI) modules.

These classifications do not necessarily correspond to the transmission form of the actual terrestrial facilities. Because of the flexibility and transmission efficiency of the voice frequency input to a TDMA terminal, this category probably has the greatest number of optional modules. However, such advantages are achieved at the cost of a more expensive assembly of mux/demux equipment needed to relate the conventional group and supergroup employed on the CT-ES link to the voice frequency baseband used by the TDMA terminal. Further, the maintenance problem is made more complex because there is a multiplicity of baseband channels rather than one large "channel."

The options specified for voice frequency interface modules are divided into three subcategories. The first and largest subcategory describes candidate options for preassignment operations which may be point-to-point or point-to-multipoint within the TDMA subburst assigned to the signal. The next subcategory is for modules that may be adapted for demand-assignment (variable destination) operation. Since these modules are capable of efficiently using DSI techniques, their signaling aspects are important. However, the TDMA Working Group has decided to tentatively exclude variable destination operation from the prototype TDMA design.

While there are numerous options for voice frequency inputs, the TDMA Working Group has proposed one particular TIM for several reasons. Since the C.C.I.T.T. has been unable to develop a single standard for the primary [10] or secondary multiplex level [11], or for a compression law [12], there are currently multiple standards at the regional level. Hence, the interregional connections, which are usually handled by satellite services, must bear the burden of the "dissimilar connection." Further, there are certain framing, signaling, and housekeeping functions inherent to the terrestrially derived PCM equipment which are not required by a subburst within a TDMA transmission because the information is already



available. Elimination of the bits which perform these functions would permit more efficient subburst formats. In addition, the PCM structure of the subburst should be locked to the TDMA burst framing; such frame lock is not available in terrestrial PCM equipment. Finally, the terrestrially derived multiplex standards employ building blocks of 24 or 30 voice channels so that the basic trunk size is too large for satellite services in which the channel increments are typically one or two voice channels.

The solution envisioned for the foregoing problem is a standard INTELSAT PCM multiplex, which can be expanded in single-channel increments up to a maximum of 60 channels. The basic PCM coding is that prescribed by C.C.I.T.T. recommendation G.711 [12]; since the input signals are at voice band, the digital PCM signals will not propagate beyond the earth station into the national network, which may use a different PCM standard. The performance characteristics of this PCM processing are as follows:

| | |
|---|---|
| sampling frequency: | 8,000 Hz $\pm$ 50 $\times$ $10^{-6}$ |
| load capacity: | +3 dBm0 $\pm$ 0.3 dB |
| compression law: | 13-segment A-law (A = 87.6) |
| quantization: | 256 levels (8-bit) |
| idle channel noise: | −65 dBm0p |
| interchannel crosstalk: | −65 dBm0 |
| quantization distortion: | within 4 dB of theoretical |

Another approach to baseband processing of voice band signals involves DSI. The best known example of speech interpolation technology is the family of TASI (time-assignment speech interpolation) terminals currently being employed over the major intercontinental submarine cable routes [13], [14]. The advantages of DSI over its analog predecessors are its fundamentally digital characteristics, which enable a lower equipment cost per channel, and its improved voice detection circuitry, an example of which is given in Reference 15.

Two competitive classes of DSI equipment have emerged. The first category includes equipment which reflects a digital extension of the principles embodied in TASI [16], [17]; the second category of equipment is more closely related to the highly predictive characteristic of speech [18]. Both approaches have shown such promise in their early developmental stages that the TDMA working group has decided not to specify variable destination demand-assignment equipment for field trial use because DSI

equipment can apparently provide superior service at a lower cost. It is hoped that the field trials, which will test these and perhaps other methods of speech interpolation, will provide the basis for administrations to choose a preferred approach and, hopefully, to determine a single method.

The use of bulk PCM coding of FDM assemblies, especially supergroups and larger assemblies, is particularly advantageous because, until some time in the future, these assemblies will represent the predominant form of terrestrial signal interfacing with the earth station. Hence, the costs associated with the interfacing, PCM coding, and maintenance are minimized. The disadvantage of this approach is that its transmission form is less efficient than a voice frequency input followed by PCM coding and time-division multiplexing. For example, a bit rate of 3,840 kbps will be required to process 60 voice channels by the channel, whereas bulk encoding of a 60-channel supergroup, which could result in a lower quality channel, might require a bit rate of about 5,184 kbps. Both approaches will meet the performance levels recommended by the C.C.I.R. for satellite links, however [19]. While it has been decided by the TDMA Working Group that PCM-coded supergroup equipment will be evaluated in the prototype field trials, this equipment, which requires complex tradeoffs between quantization noise, overload noise, error rate noise, and error correction, is still being investigated. The FDM/PCM codec will operate on the FDM signal positioned in the band from 312 to 552 kHz. The PCM encoding will use bandpass sampling at 576 kHz with linear quantization of 512 levels (9 bits). The PCM code is to be folded binary and the PCM sampling will be synchronized to the TDMA frame (432 samples per TDMA frame). The clipping level of the encoding is to be 21.8 dB relative to one test channel tone level. The quantization noise in any voice channel is to be less than 3,000 pW0p for any supergroup level between +7.6 and −25.0 dB relative to one test channel tone level using white Gaussian noise loading over the supergroup bandwidth.

Because DDI signals are not expected to appear at most earth stations until after 1980, the appropriate equipment has not received much attention. Further, the failure of the C.C.I.T.T. Special Study Group D to agree on a single PCM standard for individual channel characteristics or the primary or secondary multiplex level has placed the burden of interconnecting the dissimilar systems upon intercontinental digital systems, which will be available only by satellite. Because there appears to be no immediate need for DDI equipment and because standardization has not yet been achieved, the DDI aspects of the prototype TDMA system have been given a second priority.



**Transmit/receive-side and control subsystems**

The common TDMA terminal equipment shown in Figure 2 is intended to have three elements which operate as follows:

a. The transmit-side subsystem generally performs the information processing and handling which is appropriate to the transmissions of the terminal. This includes time-division multiplexing of the terminal burst elements, preamble generation, application of an energy dispersal waveform to the mainstream data and, finally, premodulator differential encoding and modulation.

b. The receive-side subsystem performs the inverse operations on the received signals and also performs aperture generation to increase the reliability of burst detection.

c. The control subsystem is responsible for all system restructuring and control, initial acquisition, fast re-entry, burst synchronization, and housekeeping functions required for very reliable terminal and network operation.

The transmit-side subsystem operations are controlled by the multiplexer, in which a nonvolatile content-addressable memory is activated by a symbol slot counter. Hence, the multiplexer not only reads out the contents of the appropriate TIMs at its input ports, but also turns on the preamble generator, in which the burst preamble for a particular frame is stored, and initiates the scrambler so that a pseudorandom sequence is added (in modulo-2 form) to the two parallel bit streams (the P and Q channels). This energy dispersal approach is employed to ensure that the satellite down-link signal does not interfere with terrestrial microwave facilities [20], [21]. Finally, the parallel burst-type bit streams are differentially encoded to aid the demodulator in resolving carrier phase ambiguity and performing data channel identification [22].

The receive-side subsystem demodulates, decodes, and descrambles the received bursts. Then the various information-bearing elements of the burst are distributed to their respective processing units. For example, as shown in Figure 2, the preamble detector selects the preamble from the burst and distributes the station identification code and status, the control signaling channel, the order-wire portions of the received preamble, and certain of its "unique word detected" signals to the control section. These unique word detected signals are also kept within the receive-side subsystem since they are supplied to the aperture generator and to a symbol slot counter in the demultiplex. In the demultiplex the data subbursts of interest are selected out for distribution to the return side of the TIMs.

The control subsystem is the operational heart of the terminal. The major element in this subsystem is a small general-purpose digital computer programmed to perform many functions, which may include the following:

a. order-wire input/output signaling,
b. multiplex/demultiplex format restructuring,
c. control signaling channel word format assembly and error detection encoding and decoding,
d. short-term interrupt response procedures, and
e. standby equipment exercising.

Another vital part of the control subsystem is the equipment used to allow stations to enter the TDMA transmission frame and maintain their relative frame position so that they do not overlap and interfere with any other burst. A station enters the TDMA transmission frame by first demodulating all of the existing bursts and detecting the unique word of the sync burst. This unique word is complemented in a prescribed pattern so that the beginning of a new time-divided superframe is discernible. During this superframe each of the possible stations is given a slot within which it may transmit a low-level ranging signal to synchronize itself in the TDMA frame. After the correct range is determined, the preamble is transmitted at full power and its position is adjusted to achieve final positional accuracy; then the entire burst is transmitted at full power. This entire process requires less than a few seconds. Thereafter, the steady-state burst synchronization technique is used. In case of a short-term power loss, a special technique is employed to provide more rapid re-entry into the transmission frame.

## *PSK modulator-demodulator*

The modem proposed for use in the prototype TDMA terminal equipment employed in an INTELSAT IV global beam will initially operate at about 60 Mbps and employ differentially encoded, 4-phase-coherent, PSK modulation. The differential coding enables the phase ambiguity of the reference carrier to be resolved and also permits data channel determination.

The modulator, which appears in Figure 3, accepts the two parallel data streams in symbol rate bursts, the symbol timing, and a carrier ON/OFF signal in synchronism with the data and provides a modulated carrier burst, which is generally at an intermediate frequency. The signal is then translated into the RF spectrum by the earth station's up-link.



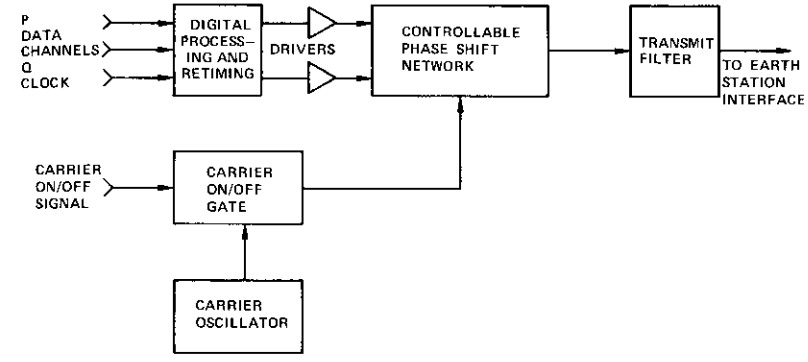



Figure 3. *PSK Modulator Functional Block Diagram*

The demodulator, shown in Figure 4, receives the TDMA modulated carrier burst at IF, recovers the coherent reference carrier and symbol timing for each burst, and uses them to demodulate the remainder of the burst. It then provides an output of two parallel data channels, P and Q, containing the data which have been sequentially transmitted by all stations

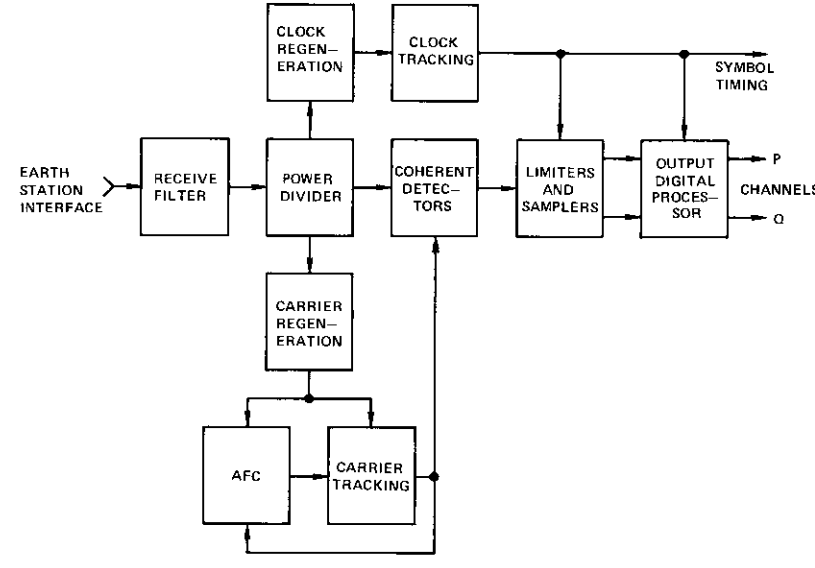



Figure 4. *PSK Demodulator Functional Block Diagram*

in the network to a particular station, and the recovered symbol timing, which is coherent with each channel's data.

A complete discussion of technical considerations and performance tradeoffs is beyond the broad scope of a general article such as this. The effects of factors such as intersymbol interference, AM/PM cochannel and adjacent channel conversion, carrier and timing recovery, and IF selection have been the subject of many discussions in the past and will continue to be discussed in the future [23]-[25]. Hence, the eventual PSK modem must be capable of providing not only exceptionally reliable performance in the burst mode of operation peculiar to these modems, but also a quantitative assessment of those factors which tend to degrade overall performance. These degradation factors will then be the subject of further study.

It appears possible to use an 8-phase modem in the INTELSAT IV spot beams to achieve a very band-limited 105 Mbps. There is insufficient power to support a 16-phase system in its normal band-limited mode of operation, however.

### *Transponder hopping*

A study of algorithms for placing TDMA traffic across the multiple INTELSAT IV transponders usually leads to a first attempt to use one long burst from each station. While this approach minimizes the up-link requirement at each earth station, it also tends to maximize the down-link requirement.

It has been found that the use of many short bursts into judiciously selected transponders makes it possible to significantly reduce the down-link requirement with a comparatively small increase in the up-link requirement. Further, and perhaps even more significantly, it appears that this approach to traffic allocation enables burst positions to be assigned so that simultaneous transmission to multiple transponders or simultaneous reception from multiple transponders may be avoided. Such traffic allocations permit "transponder hopping," i.e., the use of a single TDMA terminal operating across a number of repeaters within a single TDMA frame.

Figure 5 shows the method of allocating transmit bursts and receive bursts of interest across the multiple global-beam repeaters of an INTELSAT IV satellite to a single station. This technique is not restricted to global-beam use, however; it may also be utilized in the spot beams of an INTELSAT IV if appropriate burst synchronization procedures are followed. The transponder hopping terminal equipment configuration is shown in Figure 6. The prototype TDMA system field trials will probably



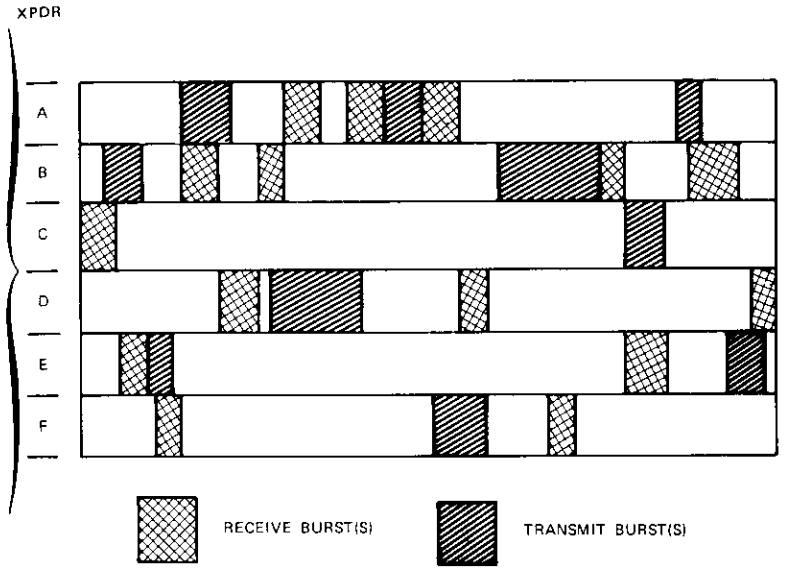



Figure 5. *Possible Transponder Hopping Format*

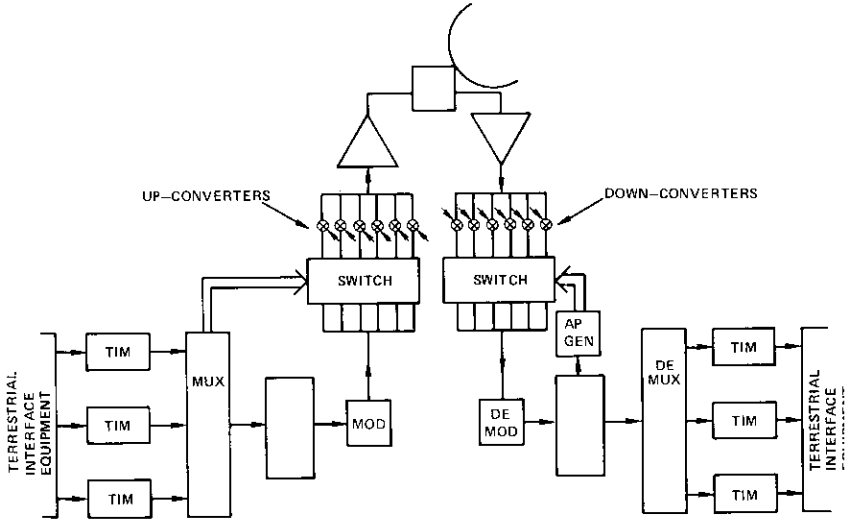



Figure 6. *General Terminal Configuration for Multiple Transmit/Receive (without overlap)*

test and evaluate the performance of this configuration. It is envisioned that selection of the proper up- and down-links will be controlled by memory elements within the multiplex and demultiplex units, but this control can be activated by other timing and control centers in the terminal.

### Redundancy switchover and maintenance

One of the primary functions of the processor in the control section is that of monitoring the operational performance of all of the elements of the common TDMA terminal equipment, i.e., both "on-line" elements and redundant elements in the hot standby condition. Clearly, these elements should have built-in fault and/or performance sensing and monitoring points. These points may be analog signals (to be quantized later), or digital signals. Storing the nominal bounds of these values within the processor, and if necessary, orienting the monitoring procedures in terms of priority, should result in very fast detection of faulty conditions. Periodic printouts of monitoring point values can be easily obtained for logging and maintenance purposes.

After a faulty condition is detected, the usual procedure consists of taking the unit off-line electrically and then advising the test and maintenance (T&M) center of the situation. It is envisioned that the processor in the T&M center will be oriented toward diagnostic fault finding and performance testing. Although the latter function is also performed in the control section processor, it is performed to a much finer degree in the T&M center.

### Acknowledgment

*The very broad scope of the work described in this paper could not be achieved by just one person. Instead, it is the product of years of INTELSAT-supported R&D, the fulfillment of numerous discussions and long hours of work by members of the TDMA Working Group, and the result of the dedication and hard work of COMSAT Laboratories personnel since 1965. To single out any group of people would be an injustice to literally hundreds of people throughout the world whose labor has been focused, directly and indirectly, on the prototype TDMA system; the author is but a convenient spokesman.*



### References

[1] T. Sekimoto and J. G. Puente, "A Satellite Time-Division Multiple-Access Experiment," *IEEE Transactions on Communications Technology*, COM-16, No. 4, August 1968, pp. 581–588.

[2] M. Takada et al., "New PCM-TDMA Satellite Communication System and Variable Destination Channel Control Technique," *Proc. 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 39–50.

[3] W. G. Schmidt et al., "MAT-1: INTELSAT's Experimental 700-Channel TDMA/DA System," *Proc. 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 428–440.

[4] K. Nosaka, "TTT System—50 Mbps PCM-TDMA System with Line Preassignment and TASI Features," *Proc. 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 83–94.

[5] G. Eckhardt, B. Reidel, and H. Rupp, "Ein Flexibles TDMA-System fur 100/50 Mbit/s," *Frequenz*, Vol. 25, 1971, p. 298.

[6] W. G. Schmidt, "An Efficient TDMA System for Use by Emerging Nations with the INTELSAT IV Satellite," *Record of the 1968 EASCON Convention*, Washington, D.C., October 1968.

[7] J. Husted, A. Walker, and G. J. Goubeaud, "A Time-Division Multiple Access System for the Defense Satellite Communications System," *Record of the 1970 EASCON Convention*, Washington, D.C., p. 229.

[8] W. G. Schrempp and T. Sekimoto, "Unique Word Detection in Digital Burst Communications," *IEEE Transactions on Communications Technology*, COM-16, No. 4, August 1968, pp. 597–605.

[9] O. G. Gabbard, "Design of a Satellite Time-Division Multiple-Access Burst Synchronizer," *IEEE Transactions on Communications Technology*, COM-16, No. 4, August 1968, pp. 589–596.

[10] International Telegraph and Telephone Consultative Committee (C.C.I.T.T.), Recommendations G.732 and G.733, Special Study Group D.

[11] International Telegraph and Telephone Consultative Committee (C.C.I.T.T.), Recommendations G.742 and G.743, Special Study Group D.

[12] International Telegraph and Telephone Consultative Committee (C.C.I.T.T.), Recommendation G.711, Special Study Group D.

[13] K. Bullington and J. M. Fraser, "Engineering Aspects of TASI," *Bell System Technical Journal*, Vol. 36, No. 2, March 1969, pp. 353–365.

[14] G. R. Leopold, "TASI-B: A System for Restoration and Expansion of Overseas Circuits," *Bell Laboratories Record*, Vol. 48, No. 10, November 1970, pp. 299–306.

[15] E. Fariello, "A Novel Digital Speech Detector for Improving Effective Satellite Capacity," *IEEE Transactions on Communications Technology*, COM–20, No. 1, February 1972, p. 55.

[16] K. Amano and C. Ota, "Digital TASI System in PCM Transmission," *Record of the International Conference on Communications*, June 1969, p. 34–23.

[17] G. M. Costa, E. Lyghounis, and I. Poretti, "Use of ATIC System to TDMA via Satellite," *Proc. of the 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, p. 39.

[18] S. J. Campanella and J. A. Sciulli, "Speech Predictive Encoded Communications (SPEC)," *Proc. of the Second International Conference on Digital Satellite Communications*, Paris, France, November 1972.

[19] International Radio Consultative Committee (C.C.I.R.), "Allowable Noise Power in the Hypothetical Reference Circuit, Recommendation 353-2, *Green Book*, Vol. IV, Pt. 2, p. 87.

[20] Radio Regulations of the International Telecommunication Union (I.T.U.), "Power Flux Density Limits," Chapter II, Article 7, Section VIII as amended by the *Final Acts of the World Administrative Radio Conference for Space Telecommunications*, Geneva: International Telecommunication Union, January 1972.

[21] A. Tomazawa, "Study of Spectral Energy Dispersion for Use with TDMA," *Proc. of the Second International Conference on Digital Satellite Communications*, Paris, France, November 1972.

[22] E. Cacciamani and C. Wolejsza, "Phase Ambiguity Resolution in a Four-Phase PSK Communications System," *IEEE Transactions on Communications Technology*, COM–19, No. 6, Pt. II, December 1971, pp. 1200–1210.

[23] C. Wolejsza, A. Walker, and A. Werth, "PSK Modems for Satellite Communications," *Proc. of the 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 51–68.

[24] S. Yokohama and J. Noguchi, "Theoretical and Experimental Considerations of the Carrier and the Bit Timing Recovery in the Burst Mode Operation," *Proc. of the 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 95–105.

[25] S. Yokohama and M. Tachibana, "The Optimum Design for Digital Satellite Communications Using Multi-phase Modulation," *Proc. of the 1969 INTELSAT/IEE International Conference on Digital Satellite Communications*, pp. 106–115.



*William G. Schmidt received a B.E.E. from Manhattan College in 1957, and an M.S.E.E. from M.I.T. in 1962.*

*He joined M.I.T.'s Lincoln Laboratory in 1957 and engaged in research in data transmission, error-control techniques, and digital circuitry for both terrestrial and satellite (LES) systems. In 1966, he joined* COMSAT *Laboratories, where he was formerly Director of the Communications Processing Laboratory, responsible for the R&D of new digital communications systems.*

*He is a member of Sigma Xi, Eta Kappa Nu, AIAA, and IEEE. He was* COMSAT'S *delegate to the C.C.I.T.T. Study Groups on PCM Systems and Data Transmission in 1968—1972.*

Index: telecommunication, communications satellites, systems analysis, mathematical models.

# Power balancing in systems employing frequency reuse

J. M. AEIN


## Abstract

This paper addresses the problem of power balancing to equalize interference effects in a satellite communications system employing frequency reuse. The use of power balancing in a simplified mathematical system model leads to a classical eigenvalue problem. Physically admissible solutions are limited to positive eigenvalues with associated eigenvectors having only positive coordinates. The Theorem of Frobenius demonstrates the existence and uniqueness of such solutions. The model is then extended to include thermal noise effects and available power constraints. Finally, the complex effects of nonlinear transponders are discussed.


## Introduction

The expected growth in needed communications satellite capacity in the 4/6-GHz band, coupled with achievements in space technology, is initiating the era of spectrum- or bandwidth-limited satellite system design. The most potent approach to the provision of capacity in a band-

This paper is based upon work performed in COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

limited system is frequency reuse of spatially isolated propagation paths, that is, propagation directionality and polarization. Since spatial isolation is never perfect, systems employing frequency reuse will tend to be limited by intrasystem interference. The objectives of this paper are to exhibit in a simplified way some of the subtleties involved, to establish the complexity of the subject, to provide some initial insight and possible approaches, and to stimulate further interest in the area.

Communications satellite systems employing frequency reuse with conventional FM telephony carriers are examined. The simplified model includes only one satellite input/output port pair per transponder and port coupling coefficients which are linear in power and independent of viewing angle. The problem addressed is that of adjusting the power levels among the various transmitted carriers to equalize the ratio of total noise plus interference to desired carrier power, $(N + I)/C$, over each system port. Baseband performance is directly related to this quantity [1]–[3].

Initially, the up- and down-links are treated separately (single-link case). If thermal noise is neglected, equalization of the $I/C$ ratio over the ports is a classical characteristic value problem with the physical constraint that the only acceptable eigenvalue/eigenvector pairs are positive eigenvalues with associated eigenvectors, all of whose elements are also positive. The eigenvector elements represent the power allocation to the carriers and the eigenvalue is the realized $I/C$ ratio.

The Theorem of Frobenius is used to show that, for suitably structured systems, a physically acceptable solution does exist and is in fact unique. Since the structuring requirement applies to the aggregation of ports into noninteracting groups, it provides physical insight, particularly in the case of a multiband satellite.

Thermal noise and available power constraints on the links are then introduced by utilizing the eigenvector structure of the interference problem. Since an eigenvector can be arbitrarily scaled, thermal noise and power constraints absolutely determine the transmitted power levels and the minimum achievable $(N + I)/C$ ratio.

Next, the end-to-end problem (up-link in cascade with a down-link) is addressed by utilizing linear transponders with variable gain settings. As in the single-link discussion, thermal noise and power constraints are introduced after first considering the problem of interference only. The end-to-end path can always be solved as two single links in cascade; the transponder gain settings are automatically determined by the separately solved up- and down-link transmitted power. The overall $(N + I)/C$ ratio is then the linear sum of the up- and down-link $(N + I)/C$ ratios. For the trivial case of two input/output port pairs, an optimum gain setting which minimizes the total $(N + I)/C$ is achieved by treating the up- and down-links separately. No such result has been realized for more complex systems.



Finally, the nonlinear effects of the transponder output are discussed. For helix TWTAs, there are two principal effects: multicarrier intermodulation noise, which is added to interference and thermal noise components, and a nonlinear output vs input power relationship. Of these two effects, multicarrier intermodulation noise presents the most difficult problem. In this case, only the approach involving single links in cascade appears tractable.

A final observation emphasizes the relative ease with which dimensionality increases in systems employing frequency reuse. Even within the perhaps overly simplified structure adopted here, the dimensionality increases rapidly.

## System model

The system mathematical model is made as simple as possible. Shown in Figure 1 are $i = 1, 2, 3, \ldots, n$ simple parallel transponder paths on one satellite. Each path has an input port with up-link peak power density per carrier $p_i$, transponder gain $k_i$, and output peak power density $q_i$. Each path also has normalized up- and down-link thermal noise sources of uniform densities $g_i$ and $h_i$. The input and output ports have mutual isolation coefficients $x_{ij}$, $y_{ij}$, defined as follows. If 1 watt is injected into the $j$th port, $x_{ij}$ watts are inflicted on the $i$th port. Similarly $y_{ij}$ are the output port couplings perceived at an earth terminal. These coefficients must be nonnegative (positive or zero) and should be small. They mathematically represent the isolation achieved by antenna cross polarization and directivity. Since interference effects are being examined and the $i$th path does not interfere with itself, $x_{ii}$ and $y_{ii}$ are identically zero. Thus, input and output isolation matrices $X = \{x_{ij}\}$, $Y = \{y_{ij}\}$, $i, j, = 1, 2, \ldots, n$, with zero main diagonals, are defined. Implicit to this formulation is the linearity on the interference power coupling between ports based on incoherent addition of the RF signals. If 2 watts are injected into the $j$th port, then $2x_{ij}$ watts are inflicted on the $i$th port. Signal-dependent port isolation schemes such as sidelobe cancellation techniques require nonlinear analysis beyond the scope of this work.

The simplification of the model must be emphasized. The parallel path

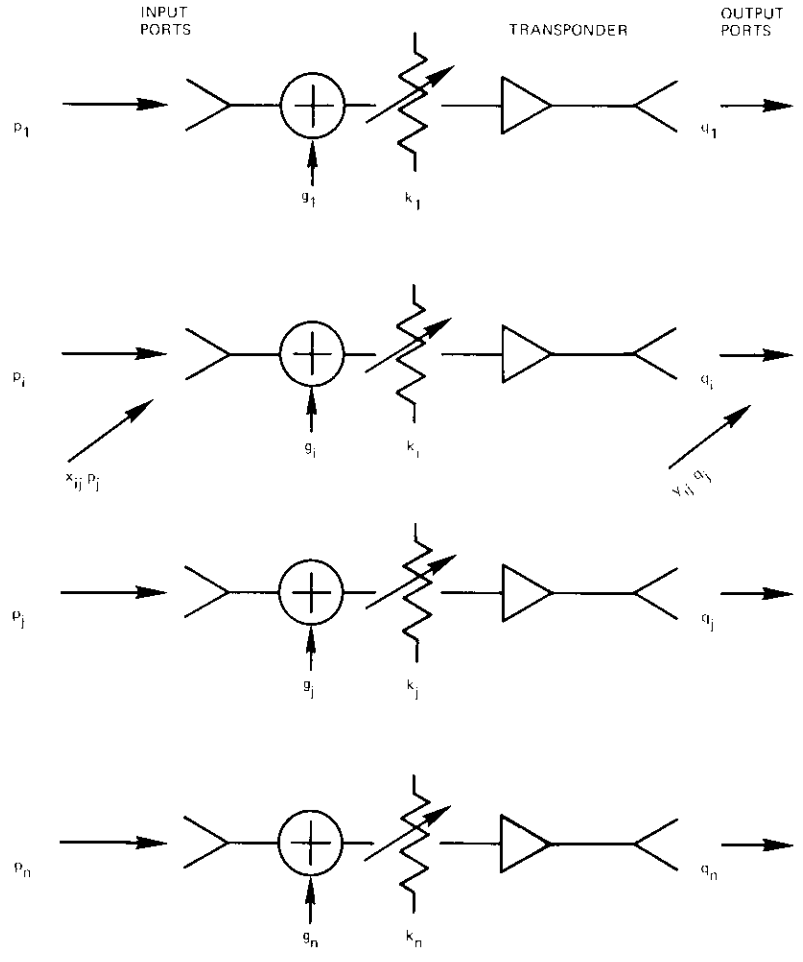



Figure 1. *Equivalent Interference Model*



transponder plan of Figure 1 is a limited representation of a real satellite. Note that only a single input/output port pair per transponder is considered. (The term "port" is used rather than the term "antenna." An antenna with polarizers and frequency filters might be represented by several ports.) Further, since the $X$ and $Y$ matrices are dependent on viewing angles, they need not be the same for all earth stations accessing or receiving from a common port. In addition, contrary to previous assumptions, all earth stations accessing a common port need not be identical.

The effects of interference on an FM telephony carrier can be related to the $(N + I)/C$ ratio. Consequently, this quantity is of principal interest. Carrier power will be allocated in proportion to signal bandwidth to make the results as independent as possible of specific carrier parameters.

Interfering carriers are assumed to be cochannel carriers (i.e., carriers on the same frequency). Although frequency interleaving is certainly advantageous, it is not addressed here. This technique can be equated mathematically to changes in path couplings which in turn generate new $X$ and $Y$ matrices. This then leads to the $X$ and $Y$ synthesis problem (i.e., design), which certainly deserves further effort.

## *Single-link interference*

### Thermal-noise-free problem

Initially thermal noise is ignored and only the up-link portion of the path is considered. Of interest is the summed interference-to-carrier power ratio, $I/C$. Since system customers are usually interested in a standard of performance irrespective of the particular transmission path, it is reasonable to strive for a constant $I/C$ for all paths.

To permit the use of carriers with different numbers of baseband channels and yet preserve the mathematical simplicity of an equal-carrier model, it is assumed that the power assigned to any carrier accessing the $i$th port is

$$P = \sqrt{2\pi}\,\mu f_\mu p_i$$

where $P$ = carrier power for a carrier entering the $i$th port
$\mu$ = rms modulation index of the carrier
$f_\mu$ = highest baseband channel frequency in the carrier.

All carriers entering the $i$th port have the same peak power density, $p_i$, since $\sqrt{2\pi}\,f_\mu\mu$ is the equivalent noise bandwidth of any carrier.

The $I/C$ ratio can be computed by using peak power densities which characterize the carriers entering a common up-link port. This ratio will be the same for all carriers in a common port. If it is assumed that any carrier entering the $i$th port is subjected to cochannel interference from all other ports, then, for $i = 1, 2, \ldots, n$,

$$(I/C)_i = \sum_{j=1}^{n} x_{ij} \frac{p_j}{p_i}$$

where, of course, $x_{ii} = 0$.

For $p_i$, $i = 1, 2, \ldots, n$, the objective is to choose values so that, for each input port, $(I/C)_i$ is a constant value independent of any particular port. Setting $(I/C)_i = \lambda$ and multiplying both sides by $p_i$ for each $i = 1, 2, \ldots, n$ yields the following eigenvalue problem:

$$(X - \lambda I)\, P = 0 \tag{1}$$

where $X$ = interference isolation matrix
$P$ = column vector of input power densities, $p_1, p_2, \ldots, p_n$
$I$ = identity matrix.

The elements of the eigenvector solution represent the peak power densities of carriers accessing each of the $n$ ports needed to achieve a constant $I/C$ which equals the eigenvalue $\lambda$. Every element, $p_i$, of a physically admissible eigenvector must be positive ($P > 0$) and $\lambda$ must be greater than 0.

Within the statement of the problem there is no guarantee that a physically admissible solution, i.e., $\lambda, P > 0$, exists. [If no admissible solution exists for certain $X$ matrices, then the objective of equal $(I/C)_i$ is not achievable.] Since the trace of the isolation matrix is zero ($x_{ii} \equiv 0$), the sum of the eigenvalues must be zero. Consequently, there exist physically unacceptable solutions.

Fortunately, interference matrices are homomorphic to stochastic matrices for which a considerable amount of theory exists. This theory, in particular, the Theorem of Frobenius, establishes the existence of unique admissible solutions. Before the implications of this theorem are explored, it is useful to provide an example of the preceding points.

EXAMPLE

Consider the up-link to a satellite with three spot beams and one global beam using a common frequency band. The three spot beams have a mutual sidelobe rejection ratio, $x_1$. The spot beams have a common polarization which is cross polarized to the global beam. Since the polarization isolation need not be reciprocal, let $x_2$ represent the isolation of the global beam from any of the spot beams and let $x_3$ represent the isolation of any spot beam from the global beam. The $X$ matrix for this example is then



$$X = \begin{pmatrix} 0 & x_1 & x_1 & x_3 \\ x_1 & 0 & x_1 & x_3 \\ x_1 & x_1 & 0 & x_3 \\ x_2 & x_2 & x_2 & 0 \end{pmatrix} \tag{2}$$

where $i = 1, 2, 3$ are the spot beams and $i = 4$ is the global beam.

The characteristic polynomial, $\Delta(\lambda)$,* for the $X$ matrix of equation (2) is

$$\Delta(\lambda) = \lambda^4 - 3(x_1^2 + x_2x_3)\,\lambda^2 - 2(x_1^3 + 3x_1x_2x_3)\,\lambda - 3x_1^2x_2x_3$$

which factors into

$$\Delta(\lambda) = (\lambda + x_1)^2\,(\lambda^2 - 2x_1\lambda - 3x_2x_3) \quad .$$

Consequently, the four eigenvalues of $x$ are

$$\lambda = -x_1, \text{ a double root}$$

$$\lambda = x_1\left[1 \pm \sqrt{1 + \left(\frac{3x_2x_3}{x_1^2}\right)}\right] \quad .$$

Note that $\Sigma\lambda = 0$. Note also that there is only one admissible eigenvalue corresponding to the largest (and only positive) eigenvalue. The corresponding positive eigenvector is given by

$$P = p\left[1, 1, 1, \left(\frac{2\lambda}{3x_2}\right)\right]$$

where $p$ = arbitrary scalar
$\lambda$ = unique positive eigenvalue given previously.

* $\Delta(\lambda)$ = determinant $[X - \lambda I]$.

For a properly defined system, it is doubly fortunate that, not only does a physically admissible solution exist, but that this solution is unique. That is, of all of the eigenvalue/eigenvector pairs, one, and only one, eigenvalue, e.g., $\lambda = r$, a positive number, and its associated eigenvector, $P(r)$, have the physical admissibility property $r > 0$, $P(r) > 0$. The basis for this result is the Theorem of Frobenius for nonnegative irreducible matrices.

The Frobenius Theorem and its implications will now be reviewed from Reference 4. First, two preliminary definitions are presented:

a. A matrix $X$ is nonnegative if every element, $x_{ij} \geq 0$. Interference isolation matrices are nonnegative matrices whose main diagonal elements are all zero.

b. A square matrix $X$ is reducible if there is a permutation* of rows and columns that put it into the form

$$\tilde{X} = \begin{pmatrix} A & O \\ C & B \end{pmatrix}$$

where $A$ and $B$ are square matrices which are not necessarily the same size.

Interference isolation matrices can be reducible; in this case, the system has been improperly defined or is pathologic.

*Theorem of Frobenius* [4]: An irreducible, nonnegative, finite matrix, $X = (x_{ij})$, always has an eigenvalue $\lambda = r > 0$ that is a simple root of the characteristic equation $\Delta(\lambda) = 0$. All of the other eigenvalues satisfy $|\lambda| \leq r$. To this "maximal" eigenvalue, $\lambda = r$, there corresponds a positive eigenvector, i.e., one whose elements are all positive. [Moreover, for some positive integer $h$, $(\lambda^h - r^h)$ is a divisor of the characteristic polynomial $\Delta(\lambda)$. The eigenvalue $\lambda = r, \lambda_2, \lambda_3, \ldots, \lambda_h$ of $X$, taken as a point set in the complex $\lambda$ plane, goes over into itself under a rotation of the plane by the angle $2\pi/h$.]†

* If two rows are permuted, then the corresponding columns must also be permuted. This is equivalent to renumbering ports.

† This portion of the theorem is not used in the following.



A corollary to the Frobenius Theorem states that the positive eigenvalue/eigenvector pair, $r$ and $P(r) > 0$, is unique.

If all of the ports are arbitrarily excited with equal values, $p_i = p$, then $I/C$ on the worst and best ports will be the upper and lower bounds of $r$, respectively. Moreover, in this equal-excitation case, $r$ does not equal its bounds unless all ports have equal $I/C$. Mathematically, $r$ is bounded [4] by the "row sums"

$$\min_{i=1,2,\ldots,n} \sum_{j=1}^{n} x_{ij} \leq r \leq \max_{i} \sum_{j=1}^{n} x_{ij} \tag{3}$$

where neither equality is attained unless all row sums are identical. The eigenvalue maximum principle for irreducible nonnegative matrices is given [4] by

$$r = \max_{P \geq 0} \left\{ \min_{1 \leq i \leq n} \sum_{j=1}^{n} x_{ij} \frac{p_j}{p_i} \right\} . \tag{4}$$

A relatively weighted $(I/C)_i$ objective is easily obtained. That is, given positive numbers $a_i$, $i = 1, 2, \ldots, n$, the objective is to obtain

$$(I/C)_i = \sum_{j=1}^{n} x_{ij} \frac{p_j}{p_i} = \lambda a_i$$

for $i = 1, 2, \ldots, n$. The $a_i$ establish a relative interference-to-carrier level between the ports ($a_1$ could equal 1) and apportion or unbalance the interference effects. The $\lambda$ set the absolute levels. Equation (1) is obtained by modifying the rows of the $X$ matrix to obtain the matrix $X'$, where, for each $i = 1, 2, \ldots, n$,

$$x'_{ij} = \frac{x_{ij}}{a_i} .$$

If $X$ is irreducible, $X'$ is also irreducible; hence all of the previous theoretical results hold except that a different solution is obtained for the $X'$ matrix.

Reducibility is an important issue. Permutation of the rows and columns* makes it possible to put [4] every nonnegative reducible matrix, $X$,

* It should be remembered that, if two rows are permuted, then the corresponding columns must also be permuted.

into the following "normal form," where the submatrices $X_1, X_2, \ldots, X_g, \ldots, X_h$ on the diagonal are irreducible, square, nonnegative matrices:

$$X = \begin{bmatrix} X_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & \\ 0 & 0 & \cdots & X_g & 0 & \cdots & 0 \\ X_{g+1,1} & X_{g+2,2} & \cdots & X_{g+1,g} & X_{g+1} & \cdots & 0 \\ \vdots & & & & & & \\ X_{h,1} & X_{h,2} & \cdots & X_{h,g} & X_{h,g+1} & \cdots & X_h \end{bmatrix}$$

If $g = h$, then $\tilde{X}$ is a diagonal form of irreducible, square, nonnegative matrices. In this case, the submatrices $X_1, X_2, \ldots, X_h$ represent physically noninteracting subsystems whose respective ports are completely isolated. In effect, $X_1, X_2, \ldots, X_h$ represent separate subsystems (in the single-link case), each separately satisfying the conditions of the Frobenius Theorem. Each subsystem, $X_l$, will then have a unique physically acceptable solution $[r_l, P(r_l)]$ for $l = 1, 2, \ldots, h$.

Mathematically, for $g < h$, it is possible to have non-zero off-diagonal matrices. This leads to a pathological physical configuration. Subsystems can still be established as described previously. Any of the first $g$ subsystems are not influenced by any other subsystem. On the other hand, the last $X_{g+1}, X_{g+2}, \ldots, X_h$ subsystems experience unilateral interference from some or all of the first $X_1, X_2, \ldots, X_g$ subsystems according to the non-zero off-diagonal matrices. A physically meaningful mathematical objective has not been formulated for this case.

**Thermal noise**

The effects of thermal noise are now introduced to determine the absolute level* of the vector $P$. In Figure 1, let the up-link thermal noise component on the $i$th port have a value $g_i > 0$ so that the ratio of noise density to desired power density is $g_i/p_i$. The positive constants $g_i$, $i = 1, 2, \ldots, n$, include such factors as port (antenna) gains, path loss, and noise temperature. The ratio of up-link total interference plus noise to desired carrier power for each $i = 1, 2, \ldots, n$ is

$$[(N + I)/C]_i = \frac{g_i}{p_i} + \sum_{j=1}^{n} \frac{x_{ij} p_j}{p_i} \quad . \tag{5}$$

Two approaches are possible. The first simply scales the eigenvector $P(r)$ to reduce thermal noise as desired at the cost of up-link power. This approach presupposes that up-link power is a readily available commodity. The second approach solves equation (5) for a new $P$ vector which makes $(N + I)/C = \rho$ the same for all up-link ports. As will be seen, the smallest $\rho$ value is $r$. Required power varies with $(\rho - r)^{-1}$.

POWER SCALING

Once the unique admissible eigenvalue/eigenvector solution $r$, $P(r)$ has been determined, it is necessary to find the weakest coordinate, $i_o$, of $P(r)$. Specifically, for any given scaling of $P(r)$, the numbers $g_i/p_i(r)$, $i = 1, 2, 3, \ldots, n$, are examined to find the index $i = i_o$ for which the largest $g_i/p_i$ occurs and $P(r)$ is renormalized so that $p_{i_o}(r) = g_{i_o}/(\rho - r)$. Since $p_{i_o}(r)$ must be positive, $\rho$ must be larger than $r$. Then $g_{i_o}/p_{i_o}(r) = (\rho - r)$ is chosen as the largest of all $g_i/p_i(r)$ ratios so that $g_i/p_i(r) \leq \rho - r$. Since $P(r)$ is the eigenvector and the $g_i/p_i(r)$ are bounded by $\rho - r$, the upper bound of equation (5) is $\rho$ for all $i = 1, 2, 3, \ldots, n$:

$$[(N + I)/C]_i \leq (\rho - r) + r = \rho \quad .$$

EXACT SOLUTION

If equation (5) is equated to $\rho$ and then, for each $i$, both sides are multiplied by $p_i$, the following equation is obtained:

$$(\rho I - X)\, P = G \tag{6}$$

where
$I$ = identity matrix
$G$ = column matrix of $\{g_i, i = 1, 2, \ldots, n\}$
$P$ = column matrix of $\{p_i, i = 1, 2, \ldots, n\}$
$X$ = interference isolation matrix $\{x_{ij}\}$.

* Note that an eigenvector can be arbitrarily scaled.

In equation (6), the solution vector, $P(\rho)$, can be close to, but will not equal, the eigenvector $P(r)$. Equation (6) has the formal solution

$$P(\rho) = (\rho I - X)^{-1} G \tag{7}$$

provided that the matrix $(\rho I - X)^{-1}$ exists, and that it is a positive matrix. This latter requirement is needed to ensure that the positive column vector $G$, when multiplied by the matrix $(\rho I - X)^{-1}$, will yield a positive column vector, $P$.

Assume that $B(\rho)$ is the matrix of cofactors of $(\rho I - X)$:

$$b_{i,j}(\rho) = (-1)^{i+j} \det [(\rho I - X); j\text{th row}, i\text{th column deleted}]$$

where det = determinant. Note that $b_{i,j}$ (and hence $B$ and $\Delta$) depends on the values of $\rho$. Then,

$$(\rho I - X)^{-1} = \frac{B(\rho)}{\Delta(\rho)}$$

and

$$P(\rho) = -[\Delta(\rho)]^{-1} B(\rho)\ G \quad . \tag{8}$$

Also note that $P$ depends absolutely on the desired $\rho$ value and, furthermore, that $\rho \neq r$. Moreover, to ensure that $P(\rho)$ is positive, the matrix $(\rho I - X)^{-1}$ must be positive. If $\rho > r$ and $X$ is irreducible [4], $P(\rho)$ will be positive. If $X$ is reducible, the problem should be broken down into its irreducible subsystems for solution.

It is interesting to observe that the solution, $P(\rho)$, of equation (8) also scales as $(\rho - r)^{-1}$. Since $r$ is a simple root of the characteristic polynomial $\Delta(\rho)$, $P(\rho)$ varies with $(\rho - r)^{-1}$ as $\rho \rightarrow r$.

**Down-Link**

The previous discussion has characterized the up-link problem. From Figure 1, it should be clear that, if transponder intermodulation noise is neglected and a relatively unlimited amount of satellite e.i.r.p. is available with respect to thermal noise on the down-link,* the two problems

* This assumption is also implicit to the up-link when no limit is placed on the $p_i$.



are identical. The $P$ matrix goes over to a $Q$ matrix, $X$ to $Y$, and $G$ to $H$. Aside from a change of notation, the down-link problem is the same as the up-link problem. However, the down-link tends to be more power limited than the up-link so that the use of upper limits on available power should be considered. Since the problems are theoretically the same, to maintain continuity of notation and thought, the placement of upper limits on the elements of the $P$ vector will be considered.

Let the positive column vector $\bar{P}$, whose elements are $\bar{p}_i$, $i = 1, 2, \ldots, n$, represent the maximum available power density on each port. Thus, for any vector $P$, it is necessary for $P \leq \bar{P}$, i.e., $p_i \leq \bar{p}_i$ for each $i = 1, 2, \ldots, n$. The presence of thermal noise and the constraint vector $\bar{P}$ determines the minimum achievable $\rho$.

Consider the "solution" obtained by scaling $P(r)$:

$$p_i(r) = \left[\frac{g_{i_o}}{(\rho - r)}\right] \gamma_i$$

where $\gamma_i \leq 1$ denotes the elements of the eigenvector $P(r)$ normalized so that $\gamma_{i_o} = 1$. Then, if $P \leq \bar{P}$ is to be satisfied, for every $i = 1, 2, \ldots, n$,

$$\min (\rho - r) = g_{i_o} \max_i \frac{\gamma_i}{\bar{p}_i} \quad . \tag{9}$$

Equation (9) yields the net achievable thermal contribution over $r$ for a bounded power availability when scaling the eigenvector.

For the exact solution, $P(\rho)$, given in equation (8), the condition $P(\rho) \leq \bar{P}$ is implemented as follows:

$$p_i(\rho) = -\frac{1}{\Delta(\rho)} \sum_{j=1}^{n} b_{ij}(\rho)\, g_j \leq \bar{p}_i, \qquad i = 1, 2, \ldots, n \quad .$$

Consequently, the smallest achievable $\rho$ is that which satisfies

$$\max_i \left\{ -[\Delta(\rho)\, \bar{p}_i]^{-1} \sum_{j=1}^{n} b_{ij}(\rho)\, g_j \right\} = 1 \quad . \tag{10}$$

Equation (10) implies a very significant numerical calculation. An approximate solution can be found if all of the elements of $P$ are large enough

so that $\rho - r$ is a small positive number. In this case,

$$\Delta(\rho) \approx (\rho - r) f(r)$$

$$b_{ij}(\rho) \approx b_{ij}(r)$$

and

$$(\rho - r) \approx \max_{1 \le i \le n} \left\{ [-f(r)\, \bar{p}_i]^{-1} \sum_{j=1}^{n} b_{ij}(r)\, g_j \right\} \quad . \tag{11}$$

All of the preceding also applies to the down-link (when transponder intermodulation effects are omitted) if the quantities $X$, $P$, $P(r)$, $P(\rho)$, and $\bar{P}$ are replaced by $Y$, $Q$, $Q(r)$, $Q(\rho)$, and $\bar{Q}$, respectively.

## *Up- and down-links with linear transponders*

In this section, linear transponders are assumed and the previous results are extended to balance both the up- and down-links. This idealized model is still of practical interest, since single-carrier-per-transponder systems utilize an essentially* linear transponder. The nonlinear effects of transponder intermodulation and power sharing are deferred to the next section.

First, the interference-dominated, noise-free eigenvalue problem is re-examined and then the noise and power bounds are reintroduced. It will be shown that it may be possible to minimize $(N + I)/C$ ratios with respect to transponder gain. Necessary conditions are given.

### Noise-free case

Figure 1 shows the transponder gains $k_1, k_2, \ldots, k_n$, which are defined as

$$k_i = \frac{q_i}{p_i}, \qquad i = 1, 2, \ldots, n$$

$$q_i = k_i p_i \quad .$$

* That is, there is no intermodulation noise. However, the output power versus input power may not be specified simply by the gain constant $k_i$. Instead of $q_i = k_i p_i$, one must use $q_i \equiv q_i(k_i p_i)$, which is a more complex case. However, with modest backoff, the $q_i$ functions are linear. With multiple carriers per transponder, there are intermodulation effects even when the $q_i$ function is effectively linear.



Consider the noise-free, unconstrained problem of choosing input power $P = p_1, p_2, \ldots, p_n$, and gain setting $K = k_1, k_2, \ldots, k_n$ to achieve an equalized net overall $I/C$ ratio:

$$(I/C)_i = \sum_{j=1}^{n} \frac{x_{ij} p_j}{p_i} + \sum_{j=1}^{n} \frac{y_{ij}(k_j p_j)}{(k_i p_i)} = \lambda \quad . \tag{12}$$

It is first necessary to determine whether equation (12) is a meaningful objective in terms of any natural subsystems of ports. If there are no subsystems on either the up- or down-link (i.e., if neither the $X$ nor $Y$ matrix is reducible), then the objective of equation (12) is meaningful. On the other hand, suppose that either the $X$ or $Y$ matrix, but not both matrices, is reducible (i.e., that either the up- or down-link port, but not both, segregates into separate subsystems). Then, the stated objective can be met because, if either $X$ or $Y$ is an irreducible matrix, then their "sum," to be defined later, is also an irreducible matrix and equation (12) can be solved. However, since the up- and down-links can be solved separately (and the $k_i$ values can then be determined), it is possible to alternatively stipulate several objectives, one for each subsystem.

If both the up- and down-link ports contain segregated subsystems of ports (if both $X$ and $Y$ are reducible), the variety of objectives becomes much greater. The exception to this observation is the simple case in which the up- and down-link subsystems exactly match. This discussion on system structure has more than theoretical interest, since it is pertinent to multiband satellites employing cross-strapped frequency reuse.

In equation (12), assume that neither $X$ nor $Y$ is reducible so that there is no ambiguity of objective. One solution to equation (12) is obtained by solving the up- and down-links separately to find $[r_u, P(r_u)]$ and $[r_d, Q(r_d)]$:

$$XP(r_u) = r_u P(r_u)$$

$$YQ(r_d) = r_d Q(r_d) \quad .$$

Then $k_i$ is set equal to $q_i(r_d)/p_i(r_u)$, and the overall solution, $I/C = r_u + r_d$, is obtained.

A more general approach is to rewrite equation (12) in the following form:

$$(Z_K - \lambda I)\, P = 0 \tag{13}$$

where $Z_K$, the $K$ weighted matrix sum of $X$ and $Y$,

$$Z_K = \left\{ x_{i,j} + \left(\frac{k_j}{k_i}\right) y_{i,j} \mid i,j = 1, 2, \ldots, n \right\}$$

is a nonnegative irreducible matrix. Note that the matrix $Z_K$ is controlled by $n - 1$ rather than by $n$ values of $k_i$ because only ratios of $k_j/k_i$ are involved. From the previous section, for every $K > 0$, equation (13) has the eigenvalue/eigenvector solution $r(K)$, $P[r(K)] \equiv P(K)$. Hence, it is possible to seek that value of $K > 0$ which minimizes $r(K) = r(k_1, k_2, \ldots, k_n)$.

Formally, a necessary condition for obtaining a value of $K$ which minimizes $r(K)$ is

$$\frac{\partial r(K)}{\partial k_i} = 0, \qquad i = 1, 2, \ldots, n \quad . \tag{14}$$

Equation (14) actually generates $n - 1$ independent equations on the $k_i$ because, if the solution $K = K^*$ for equation (14) exists, it can be arbitrarily scaled.

It has not been established that equation (14) has a solution, $K^*$, which has all positive elements, $k_i^*$. Even if equation (14) has a positive $K$ solution, $r(K)$ must be convex upwards for this solution to minimize r$(K)$. Although the formulation of equations (13) and (14) is simple, the implied computational effort is quite significant. First, the characteristic polynomial of $Z_K$ must be found and the unique r$(K) > 0$ determined from

$$\Delta(\lambda,K) \equiv \det (Z_K - \lambda I) = 0 \quad .$$

The $n - 1$ independent equations

$$\frac{\partial r(K)}{\partial k_i} = 0$$

must then be solved for a positive solution vector $K$.

For some cases, there is an alternative formulation which exploits the fact that $r(K)$ must be a unique simple root of $\Delta(\lambda,K)$. For these cases there is an equivalent set of necessary conditions which may mitigate some computational complexity:

$$\frac{\partial \Delta(\lambda, K)}{\partial k_i} = 0, \qquad n - 1 \text{ independent equations on } k_i \text{ and } \lambda \tag{15a}$$

$$\Delta(\lambda, K) = 0, \qquad \text{one equation on } k_i \text{ and } \lambda \quad . \tag{15b}$$



Unfortunately, it has not generally been possible to relate the best solution to the value of $K$ chosen by solving the up- and down-link problems independently. That is, no general relationship between $r_u$, $r_d$, and $r^* = r(K^*)$ has been found. For the trivial case of $n = 2$ (2-port system), it is easily shown by direct calculation that $r^* = r_u + r_d$.

**Thermal noise**

If thermal noise is now introduced into the problem, equation (12) is modified as follows:

$$[(N + I)/C]_i = \left[\sum_{j=1}^{n} \frac{x_{ij}p_j}{p_i} + \frac{g_i}{p_i}\right] + \left[\sum_{j=1}^{n} \frac{y_{ij}(k_j p_j)}{(k_i p_i)} + \frac{h_i}{k_i p_i}\right] \tag{16a}$$

or

$$[(N + I)/C]_i = \frac{g_i + (h_i/k_i)}{p_i} + \sum_{j=1}^{n} \frac{[x_{ij} + (k_j/k_i)]\, p_i}{p_i} \quad . \tag{16b}$$

In equation (16a) the up- and down-links are separated so that they can be solved individually (by substituting $q_i = k_i p_i$). From these separate solutions, gain settings $k_i$ are determined by dividing the desired $q_i$ by the desired $p_i$. The net overall $(N + I)/C$ is then the linear sum of the up- and down-link $(N + I)/C$. With this approach,* no attempt is made to optimize gain settings.

If, on the other hand, the formulation of equation (16b) is used, the gain settings can be varied. Stating as an objective

$$[(N + I)/C]_i = \rho, \qquad i = 1, 2, \ldots, n$$

leads to an equation for each gain $K = (k_1, k_2, \ldots, k_n)$:

$$P(\rho) = (\rho I - Z_K)^{-1} (G + H_K) \tag{17}$$

where
$$Z_K = \{x_{ij} + (k_j/k_i)\, y_{ij} \mid i,j = 1, 2, \ldots, n\}$$
$$H_K = \text{column vector } h_i/k_i, \qquad i = 1, 2, \ldots, n$$
$$G = \text{column vector } g_i, \qquad i = 1, 2, \ldots, n \quad .$$

* Note also that in this approach, for single-carrier-per-transponder operation, the nonlinear gain dependence increases the complexity only slightly. The input levels, $k_i p_i$, to the transponder are adjusted so that $q_i(k_i p_i) = q_i'$.

Equation (17) does not assume any power constraint vectors $\bar{P}$ and $\bar{Q}$. Its solution requires $\rho(K) > r(K)$ for each $K$. Although the gain vector $K$ which minimizes $\rho(K)$ need not be the same as that which minimizes $r(K)$, for interference-limited systems it is reasonable to expect that the $K$ which minimizes $r(K)$ will be close to that which minimizes $\rho(K)$.

#### Power Constraints

It is necessary to introduce the effects of the constraints $\bar{P}$ and $\bar{Q}$. From a line of reasoning parallel to that which led to equation (10), it follows that, for each $K > 0$, there is a $\rho(K)$ which solves

$$\max_{1 \le i \le n} \left\{ [-\Delta(\rho, K)\, p_i]^{-1} \sum_{j=1}^{n} b_{ij}(\rho, K) \left( g_j + \frac{h_j}{k_j} \right) \right\} = 1 \qquad (18)$$

where $\Delta(\rho,K)$ = characteristic polynomial of $Z_K$

$b_{ij}(\rho,K)$ = cofactors of $Z_K$ .

For this $\rho(K)$, there is a $P(K) = P[\rho(K)]$ given by equation (17). Consequently, the smallest value of $\rho$, $\rho_{\min}$, is found by minimizing $\rho(K)$ over $K$ subject to the additional limitation that $K$ must be constrained so that* $Q = KP(K) \le \bar{Q}$. This presents a formidable optimization problem, one whose solution is not attempted here.

The preceding theory deals with irreducible up- and down-link interference isolation matrices $X$ and $Y$. As mentioned previously, if either or both $X$ and $Y$ are reducible, the issues become more complicated because of the possibility of establishing alternate $(N + 1)/C$ objectives.

### *Transponder nonlinearity effects*

In this section, a qualitative discussion is presented to introduce a nonlinear element, the saturation effects of the transponder. Two principal effects are of interest when FM multichannel telephony carriers share a transponder. The percentage of power obtained by each of the frequency-divided carriers sharing a transponder and the intermodulation noise generated between them are nonlinear functions of the transponder power backoff. In general, these functions depend on the specific configuration and details of the carriers using the transponder (e.g., relative power, bandwidths, and center frequencies) and the transponder filters and amplifying device(s).

* That is, for each $K > 0$ and corresponding $P(K)$ determined from equations (17) and (18), $K$ must satisfy $k_i p_i(K) \le q_i$ for each $i = 1, 2, \ldots, n$.



For practical application to TWT transponders, the nonlinear input/output transfer functions can be specified to a fair degree of accuracy. These curves can be obtained for a particular TWT and then used in graphic form. Alternatively, the curves can be fitted with a specific function (e.g., polynomial). In the simplified model considered here, it is assumed that there are no "capture" type effects produced by individual carriers in a common transponder. More specifically, it is assumed that, since all input carriers have the same peak power density, $p_i$, they all obtain the same output peak power density, $q_i$. [This need not hold when operating near transponder saturation with disparately sized carriers. In this case, the output power can be unequal and depends on each input and the frequency plan.]

The net useful output power density variable referenced to an earth station receiver for each transponder, $i$, has been $q_i$ with an upper bound, $\bar{q}_i$, which cannot exceed the single-carrier saturated power density, $\hat{q}_i$. In this simplified structure, the backed off transponder output power density, $q_i$, is a nonlinear function of the product of transponder gain, $k_i$, and input power density, $p_i$:

$$q_i \equiv q_i(k_i p_i) \quad .$$

The $q_i$ functions are determined from the input/output transfer functions. If $\Omega_i$ is the sum of the effective RF bandwidths of each carrier sharing the $i$th transponder, $\hat{q}_i$ is equal to the transponder single-carrier saturated power divided by $\Omega_i$. The transponder output backoff factor is then given by the ratio $q_i/\hat{q}_i$. It should be noted that, because of transponder multicarrier sharing loss, $\bar{q}_i$ must be less than $\hat{q}_i$.

Also note that, unless all of the transponders are identical, the nonlinear functions $q_i$ are not the same. For sufficiently small values of $q_i/\hat{q}_i$, the $q_i$ functions become linear; i.e., $q_i \doteq k_i p_i$. In a typical C-band TWT, $q_i$ is linear on $k_i p_i$ when $q_i$ is 5 dB below $\hat{q}_i$. At the price of increased thermal noise, the limiting power density value, $\bar{q}_i$, can be set sufficiently below $\hat{q}_i$ to preserve the linearity on $k_i p_i$.

The preceding discussion has modeled the nonlinear mathematical behavior of the transponder output power. The physical nonlinearity of the devices in the transponder operated with multiple carriers generates additional intermodulation noise which must be added to the external thermal and interference noise. This additional noise has a very important effect on the analysis.

If channel peak loading considerations are ignored and the RF power

allocated to a carrier is apportioned to its rms bandwidth, then the thermal and interference noise degradation does not depend on the actual number of baseband channels, $m$, each carrier contains. On the other hand, the degradation effected by the intermodulation noise does vary in direct proportion to the number of baseband channels on a carrier. Consequently, intermodulation effects are dependent on the individual carrier, even when thermal noise and interference effects are normalized. Primarily because of this dependence on the actual number of baseband channels in a carrier, RF power is not directly apportioned to rms bandwidth in systems not employing frequency reuse. The current practice is to attempt to weaken the dependence on the number of channels by unbalancing the RF power.

In the following, it is assumed that the intermodulation noise can be treated as an equivalent additive source of thermal noise. Moreover, the intermodulation spectrum is assumed to be "white-like." That is, the intermodulation-to-desired-carrier power ratio, $I/C$, is the same (and independent of carrier frequency) for all carriers sharing a common transponder. For carriers having the same peak power density, this assumption has some justification. In more general situations the intermodulation spectrum must be calculated. Since the spectral shape can be strongly dependent on the various input carrier frequencies and power levels, the general problem rapidly becomes intractable for theoretical study.

This problem is one of the deficiencies of the framework proposed in this paper. Up to this point, the mathematical description did not depend on the characteristics (e.g., frequency plan and number of baseband channels) of any particular carrier. The model employed here deals only with variables associated with the ports, i.e., interference isolation, power densities, and gain. Including individual carrier variables enormously increases the dimensionality of the problem. Consequently, in the following it is assumed either that all carriers sharing the same transponder have the same characteristics, or that a worst-case carrier is established for each transponder, and intermodulation noise is adjusted so that it is no worse than it would be if all carriers had the most stressing parameters. This second assumption tends to be rather stringent and to require a larger than necessary transponder backoff.

The problem formulation proceeds as follows. The intermodulation contribution is a function of the backoff variable, $q_i/\hat{q}_i$. If the intermodulation is assumed to be additive with a "white-like" spectrum, the intermodulation-to-carrier power ratio for each transponder, $(IM/C)_i$, $i = 1, 2, \ldots, n$, can be expressed as



$$(IM/C)_i = m_i b_i\left(\frac{q_i}{\hat{q}_i}\right)$$

where $m_i$ = worst-case number of baseband channels per carrier sharing the $i$th transponder

$b_i(\cdot)$ = nonlinear function describing intermodulation power density level versus output backoff.

A typical functional form* for $b_i$ with a helix C-band TWT is

$$b_i = b\left(\frac{q_i}{\hat{q}_i}\right)^{20/9}$$

where $b$ is a constant.

When intermodulation is included, for $i = 1, 2, \ldots, n$,

$$[(N + I + IM)/C]_i = \sum_{j=1}^{n} \frac{x_{ij}p_j}{p_i} + \frac{g_i}{p_i} + \sum_{j=1}^{n} \frac{g_{ij}q_i}{q_i} + \frac{h_i}{q_i} + m_i b_i\left(\frac{q_i}{\hat{q}_i}\right) \qquad (19)$$

where $q_i \equiv q_i(k_i p_i)$ .

Equation (19) exhibits nonlinear behavior through the down-link $q_i$ variables and the backoff $b_i$ variables. Note that, even when $q_i$ is linear on $k_i p_i$, the function $b_i(k_i p_i)$ is still nonlinear and makes a significant contribution.

At this point, no general optimization approach is apparent. A computationally tractable, although suboptimal, approach uses a form of equation (19) in which the up-link, down-link, and backoff problem components are solved separately. The solution of the up- and down-links has been demonstrated previously. Although the down-link variables $q_i$ are nonlinear on $k_i p_i$, the down-link can be solved separately from the up-link. Then the nonlinear dependence on $k_i p_i$ is relatively simple to handle because it affects only the selection of the gain settings. The suggested approach parametrically solves the separated problem with respect to an allowable value of intermodulation noise, $\rho_{IM}$. For each assumed $\rho_{IM}$,

* In the more general case, the function $b_i(\cdot)$ cannot be defined. The intermodulation spectrum must first be calculated for a given input carrier frequency plan and for each carrier input power density. The IM/C ratio can then be obtained for each carrier sharing the $i$th transponder. This process must be repeated for all multicarrier transponders.

a solution to the up- and down-links is obtained, thus determining the ratio of overall noise plus interference plus intermodulation to desired carrier power; the smallest such ratio is the desired solution. A given value of $\rho_{IM}$ will establish upper bound values, $\bar{q}_i$, on the $q_i$. Specifically, the $\bar{q}_i$ are found by solving for a given $\rho_{IM}$, $m_i$, and function $b_i(\cdot)$:

$$m_i b_i\left(\frac{\bar{q}_i}{\hat{q}_i}\right) = \rho_{IM} \quad . \tag{20}$$

Then, with $\bar{Q} = \{\bar{q}_i\}$, the methods described in the section entitled "Single-Link Interference" can be used to obtain the solution, $Q^*$, to the down-link:

$$[\rho_d I - Y]\, Q^* = H, \qquad 0 < Q^* \leq \bar{Q} \tag{21}$$

and the value of the down-link $(N + I)/C$ contribution is $\rho_d$. The up-link problem can be solved with a power constraint $\bar{P}$ determined by the earth station e.i.r.p.:

$$[\rho_u I - X]\, P^* = G, \qquad 0 < P^* \leq \bar{P} \tag{22}$$

and the value of the up-link $(N + I)/C$ is $\rho_u$. The $k_i$ are then found from the functional relationship

$$q_i(k_i p_i^*) = q_i^* \quad . \tag{23}$$

Equations (20) through (23) represent a solution technique with the overall $(N + IM + I)/C$ on each port equal to $\rho_u + \rho_d + \rho_{IM}$.

## Conclusion

This paper has addressed a simplified structure of a frequency reuse satellite system employing conventional FM telephony carriers. Even with the simplified system, meaningful analytical problems are generated and a question is left unresolved. Namely, does an optimum set of transponder gain settings exist and how do they relate to those determined by single-link analysis? It is readily apparent that the dimensionality of the problem increases rapidly, especially if the carriers cannot be treated in an aggregated (by port) fashion. Hence, treatment of the intermodulation effects on multicarrier transponders can present a severe problem.

At least two areas remain unexplored. The first and possibly easier is an extension of the formulation to more than one port pair per transponder. The second is the isolation matrix synthesis problem, especially the use of frequency interleaving.



## Acknowledgment

*The author would like to thank his colleagues at COMSAT Labs for their aid and encouragement. First, he would like to thank E. Podraczky, without whose encouragement this effort would have early terminated. Secondly, he would like to thank N. K. M. Chitre and W. Garner, who educated him in the subject and sensitized him to the complexity of the problem. Finally, he extends his thanks to the RF Transmission Lab for technical interaction and review of the effort.*

## References

[1] R. Hamer, "Radio Frequency Interference in Multichannel Telephony FM Radio Systems," *Proceedings of the IEEE*, Vol. 108, January 1961, pp. 75–89.

[2] V. K. Prabhu and L. H. Enloe, "Interchannel Interference Considerations in Angle Modulated Systems," *Bell System Technical Journal*, Vol. 48, No. 7, September 1969, pp. 2333–2358.

[3] B. A. Pontano, J. C. Fuenzalida, and N. K. M. Chitre, "Interference into Angle-Modulated Systems Carrying Multichannel Telephony Signals," *IEEE Transactions on Communication Technology*, COM-21, No. 6, June 1973.

[4] F. R. Gantmacher, *The Theory of Matrices*, Vol. I and II, New York: Chelsea Publishing Co., 1959.

*Joseph M. Aein received the B.S. and M.S. degrees from M.I.T., in 1958, and the Ph.D. degree from Purdue University in 1962.*

*While at Purdue he was an Instructor in the School of Electrical Engineering. He joined the staff at the Institute for Defense Analyses, Arlington, Virginia, in 1962. During the academic year 1965-1966, he was a Visiting Lecturer in the Department of Electrical Engineering, University of California, Berkeley. In 1972 he joined* COMSAT *Laboratories as a Senior Staff Scientist; in 1973 he rejoined I.D.A. He is a member of Sigma Xi, IEEE, and AIAA.*

Index: telecommunications, bandpass filters, digital techniques, data transmission, optimization design.

# Approach to a near-optimum transmitter-receiver filter design for data transmission pulse-shaping networks

F. Assal


## Abstract

This paper is concerned with the theoretical determination of linear channel transfer functions, and the practical design and physical realization of IF bandpass filters used to approximate optimum signal performance with respect to the lowest bit-error rate for synchronous data transmission.

If an instantaneous sampler is used to identify each received pulse in a bit stream, then the transmitter-receiver filter characteristics may be optimized as follows. Intersymbol interference is minimized by controlling the time-axis crossings of each received pulse. White Gaussian noise is minimized by designing a "matched" filter at the receiver. For a given sample amplitude, the required signal power and white Gaussian noise are further minimized by group-delay equalizing the transmission channel up to the detector input. Finally, bit stream crosstalk is minimized by realizing arithmetically symmetric bandpass filters.

To satisfy these requirements, the design procedure begins with standard filters that are modified in their passbands by selectively adjusting the group delay and the attenuation in the cascaded bridged-tee networks which are added to the original standard filter. Passive, lumped elements are used for fabricating all of the networks at 70-MHz center frequency.

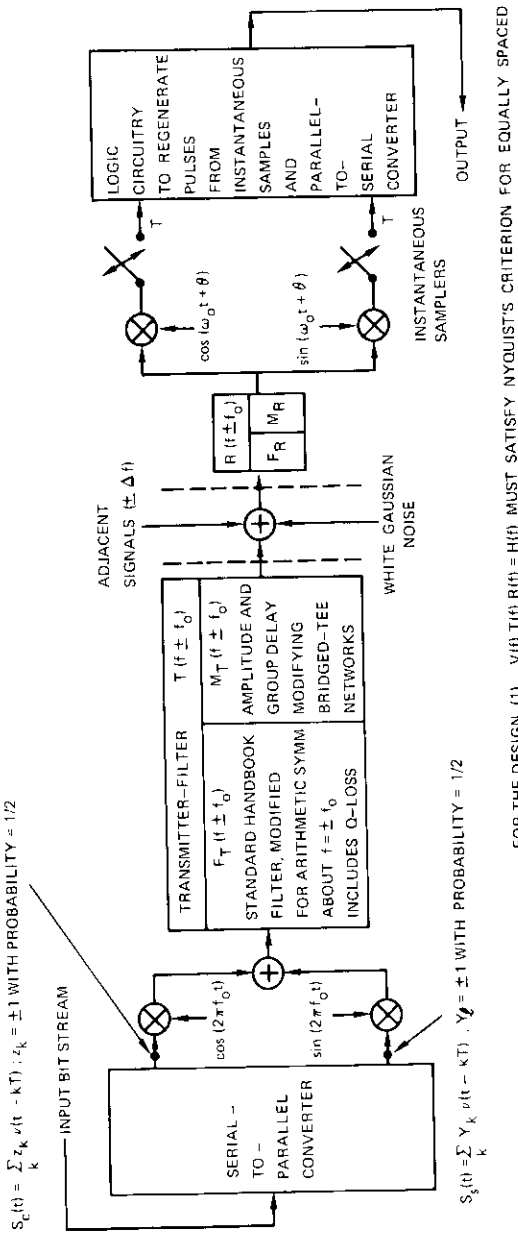



Figure 1. *4-Phase PSK System*



## *Introduction*

In bandwidth-limited synchronous data transmission systems (see Figure 1), a transmitter filter is used to protect adjacent channels and to minimize the required transmitted power. A receiver filter is used to suppress adjacent signals and to minimize white Gaussian noise introduced in an otherwise ideal linear channel.

If an instantaneous sampler* identifies each pulse in a bit stream, intersymbol interference, the most serious (in-band) degradation in data transmission, may be theoretically suppressed by designing cascaded transmitter-receiver filters to satisfy Nyquist's first criterion [1] for equally spaced time-axis crossings in the pulse response.

This paper is concerned with the theoretical determination of optimum transmitter-receiver filter transfer functions and the practical design and physical realization [2] of IF bandpass filters that closely approximate the aforementioned frequency- and time-domain requirements. The design procedure uses standard filters that are modified in their passbands by selectively adjusting the group delay and the attenuation in cascaded bridged-tee [2] networks.

Since the modulation scheme shown in Figure 1 corresponds to a frequency translation for each bit stream, the filtering functions can be performed either at baseband or at IF. In a 4-phase PSK system, this approach has been used to reduce the number of filters from four at baseband to two at IF. However, special attention must be paid to the symmetry [3] of the filters about the carrier frequency, $f_o$, to avoid loss of orthogonality between the carriers and to minimize bit stream crosstalk. Hence, although the derivations in the following sections are carried out for low-pass transforms, bandpass filters will be realized.

## *Optimum transmitter-receiver transfer function design*

In this section, the desired optimum transmitter and receiver filters, $T(f)$ and $R(f)$, respectively, will be derived at baseband as functions of

* Preliminary analyses show that mid-bit sampling is a near-optimum detection method in a bandwidth- and power-limited environment in the presence of intersymbol interference; therefore, mid-bit sampling is utilized as the detection method throughout this paper. It should be noted that the optimization criteria and conclusions would be altered if the detection methods were changed.

the pulse input transform, $V(f)$, and the pulse output transform, $H(f)$, at the detector input so that $H(f)$ satisfies Nyquist's first criterion [1] for equally spaced time-axis crossings. In general, this criterion may be satisfied by an infinite set of output transforms [4]. However, for bandwidth-limited data transmission, the output transform must be negligible for all frequencies above some specified bandwidth, $f_m = f_c + f_e \leq 2f_c$, where $f_c$ is the Nyquist frequency corresponding to a bandwidth-to-baud-rate ratio of unity, and $f_e$ is the excess bandwidth. Then a subset of these transforms is obtained [1], namely,

$$H_d(f) = H(f)\ e^{j2\pi f T_0} = \begin{cases} H_0 + H_1(f), & f < f_c \\ H_1(f), & f_c < f < 2f_c \\ 0, & 2f_c < f \\ H_d^*(-f), & \text{for all } f \end{cases} \tag{1}$$

where $H_d(f)$ is the time-translated output transform which absorbs any transit delay, $T_0$, so that the waveform mid-bit occurs at $t = 0$, and $H_1(f)$ is restricted by the following complex equation:

$$H_1(f_c + \Delta f) = -H_1^*(f_c - \Delta f), \qquad \Delta f < f_c \quad . \tag{2}$$

For some given input pulse transform, $V(f)$, the output transform, $H(f)$, will be approximated by cascading realizable filters, and amplitude and group-delay equalizers. Assume that $F_T(f)$ and $F_R(f)$ are realizable transmitter and receiver filter functions, respectively, chosen from a filter handbook [5] for their out-of-band attenuation selectivities. Also assume that $M_T(f)$ and $M_R(f)$ are realizable amplitude and group-delay equalizers designed to modify the filters $F_T(f)$ and $F_R(f)$, respectively, to approximate optimum digital system performance. The theoretically optimum transmitter-receiver filters are then defined in terms of the realizable transfer functions

$$T(f) \triangleq \epsilon_T(f)\ F_T(f)\ M_T(f) \tag{3}$$

$$R(f) \triangleq \epsilon_R(f)\ F_R(f)\ M_R(f) \tag{4}$$

where $\epsilon_T(f)$ and $\epsilon_R(f)$ are error functions used to compare the realizable functions with the theoretically desirable characteristics. Ideally then,



the exact output transform is

$$H_d(f) = V(f)\ T(f)\ R(f) \quad . \tag{5}$$

Assume that the stopband characteristics of the output transform are

$$H_1(f) = F(f)\ V(f), \qquad f_c \leq |f| \leq 2f_c \tag{6}$$

where the realizable function $F(f)$ is given by

$$F(f) = F_R(f)\ F_T(f)\ e^{j2\pi f T_1} \quad . \tag{7}$$

The passband characteristics are obtained by replacing $f$ with $(f_c + \Delta f)$ in equation (6), equating this result with the expression in equation (2), and replacing $\Delta f$ with $(f_c - f)$. The resulting expression

$$V(2f_c - f)\ F(2f_c - f) = -H_1^*(f), \qquad |f| \leq f_c \tag{8}$$

will then be substituted into equation (1). For

$$H_d(f) = H_0 - [V(2f_c - f)\ F(2f_c - f)]^*, \qquad |f| \leq f_c \tag{9}$$

the constant $H_0$ is obtained by forcing the desired transfer function to be continuous at $f = f_c$ and by equating it to the corresponding expression in equation (6). The output transform is then

$$e^{j2\pi f T_0}\ H(f) = \begin{cases} 2\ Re[V(f_c)\ F(f_c)] \\ \quad - [V(2f_c - f)\ F(2f_c - f)]^*, & |f| \leq f_c \\ V(f)\ F(f), & f_c \leq |f| \leq 2f_c \end{cases} \quad . \tag{10}$$

Bennett and Davey [1] and Lucky, Salz, and Weldon [4] show that the optimum receiver filter amplitude is proportional to the square root of the output pulse transform and inversely proportional to the fourth root of the noise spectrum. Since only white Gaussian noise is considered here,

$$|R(f)| = |H(f)|^{1/2} \quad . \tag{11}$$

Note that this result is independent of phase angle. The input pulse, $V(f)$, the transmit filter, $T(f)$, and the receiver filters, $R(f)$, may have any phase variations provided that $\theta_V(f) + \theta_T(f) + \theta_R(f) = \theta_H(f)$. It is

shown in the appendix that, for optimum operation, the group delay of $H(f)$ must be constant, $|H_R(f)| \leq |H_0|$, and $H_R(f) \leq 0 < H_0$ or $H_R(f) \geq 0 > H_0$.

If the input pulse waveform is selected to be arithmetically symmetric about its mid-bit and if it is advanced sufficiently to absorb the composite filter mid-band delay, then, for optimum operation, $T(f)$ and $R(f)$ must also be real. Hence,

$$R(f) = \begin{cases} [2|V(f_c)|\ |F_T(f_c)|\ |F_R(f_c)| - |V(2f_c - f| \\ \quad \cdot |F_T(2f_c - f)|\ |F_R(2f_c - f)|]^{1/2}, & |f| \leq f_c \\ |F_T(f)|, & f_c \leq |f| \leq 2f_c \end{cases} \tag{12}$$

$$T(f) = \begin{cases} R(f)\ /|V(f)|, & |f| \leq f_c \\ |F_R(f)|, & f_c \leq |f| \leq 2f_c \end{cases} . \tag{13}$$

Equations (12) and (13) indicate that filters $F_T(f)$ and $F_R(f)$ must be group-delay equalized and amplitude modified in the passband. These two operations will be approximated as shown in equations (3) and (4) with bridged-tee networks connected in cascade with the filters. Finally, while the out-of-band attenuation of the receiver filter might ideally be selected so that it is equal to

$$|F_R(f)| = |V(f)|\ |F_T(f)|, \qquad f_c \leq |f| \leq 2f_c \tag{14}$$

since the attenuation of the receiver filter is equal to 3 dB at $f_c$, the receiver noise bandwidth is essentially independent of the particular filter choice. For practical designs it is therefore convenient to select two identical filters; i.e.,

$$F_R(f) = F_T(f), \qquad f_c \leq |f| \quad . \tag{15}$$

In equations (12) and (13), the passband attenuation of the receiver filter, $R(f)$, is almost constant for most of the passband, and the transmit filter in-band modification may be simplified if the pulse duration is reduced. In the limit, as $T \rightarrow 0$, $R(f)$ and $T(f)$ will be identical.

Figure 2 shows the computed characteristics of two cascaded elliptic function filters with a combined excess bandwidth of 14 percent, modified for the optimum transmission of rectangular pulses, $T = 1/2\ f_c$. The measured pulse response (Figure 2b) displays the desired axis crossings of filters designed according to these procedures.



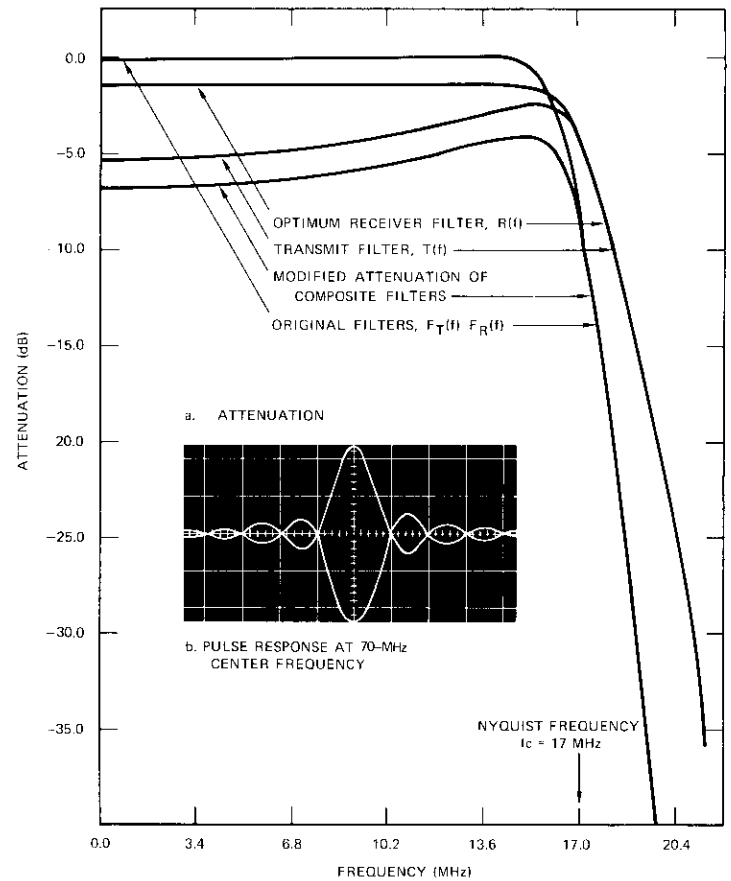
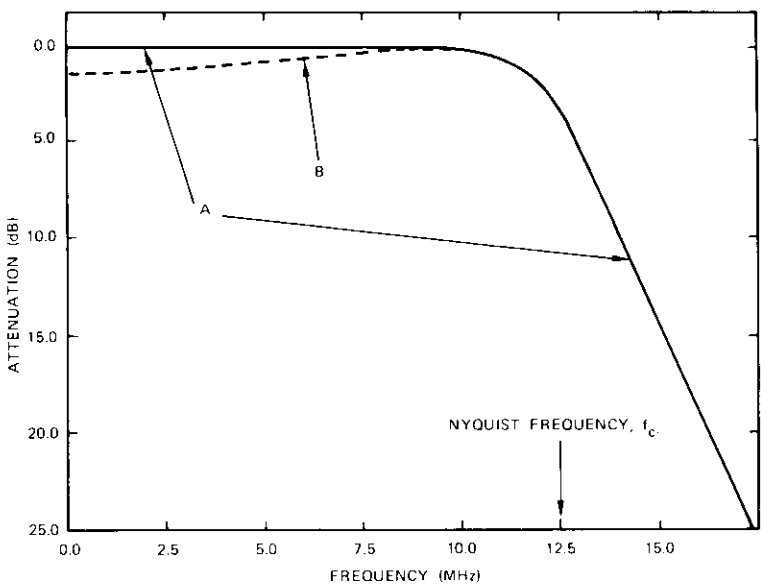



Figure 2. *Optimum Filter Characteristics Using Two 5-Pole Elliptic Function Filters for 68-Mbps, 4-Phase, PSK-PCM Digital Communications*

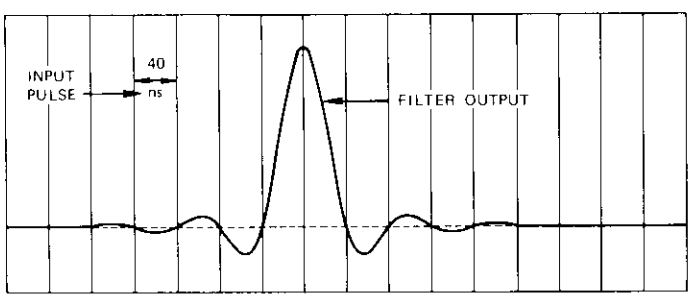
## Bandpass filter design and realization

The filter of Figure 3, which requires a modifying function similar to $M_T(f)$, has been chosen to illustrate the design technique.

Known procedures [2] are used to derive the elliptic bandpass filter shown in Figure 4. The transformation from low-pass to bandpass leads




A. Computed Attenuation for a 5-Pole Filter (Curve A) and its In Band Modification (Curve B)

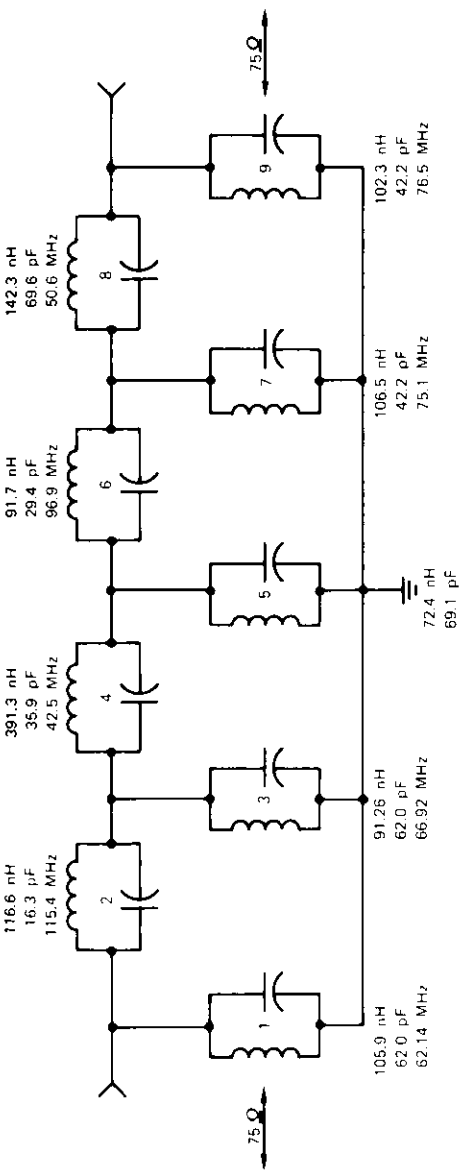



B. Computed Pulse Response for the Modified Filter

Figure 3. *Bandpass Filter Chosen to Illustrate the Design Technique*

to geometrically symmetric characteristics, shown in Figure 5. Arithmetic symmetry is achieved for elliptic function filters by moving the resonant frequencies of the four series tank circuits shown in Figure 4 (corresponding to the infinite rejection points of Figure 5) so that they are arithmetically symmetric about the carrier frequency, $f_o = 70$ MHz. To least perturb the other elements in the circuits, the impedance of these tank circuits is unchanged with respect to the carrier center frequency:






GEOMETRIC CENTER FREQUENCY = 70.0 MHz
EQUIRIPPLE FREQUENCIES = 61.27 MHz 79.97 MHz
BANDWIDTH : CENTER FREQUENCY = 26.7%
INFINITE REJECTION FREQUENCIES - 115.4 MHz 42.5 MHz 96.9 MHz 50.6 MHz
INFINITE REJECTION FREQUENCIES FOR ARITHMETIC SYMMETRY = 106.5 MHz 33.5 MHz 93.2 MHz 46.8 MHz

| MODIFIED SECTION NUMBER | 2 | 4 | 6 | 8 |
|---|---|---|---|---|
| INDUCTANCE (nH) | 104.8 | 765.1 | 83.5 | 191.7 |
| CAPACITANCE (pF) | 21.3 | 29.5 | 34.9 | 60.2 |
| RESONANT FREQUENCY (MHz) | 106.5 | 33.5 | 93.2 | 46.8 |

Figure 4. *"Geometric" Bandpass Filter (5-pole elliptic function filter)*

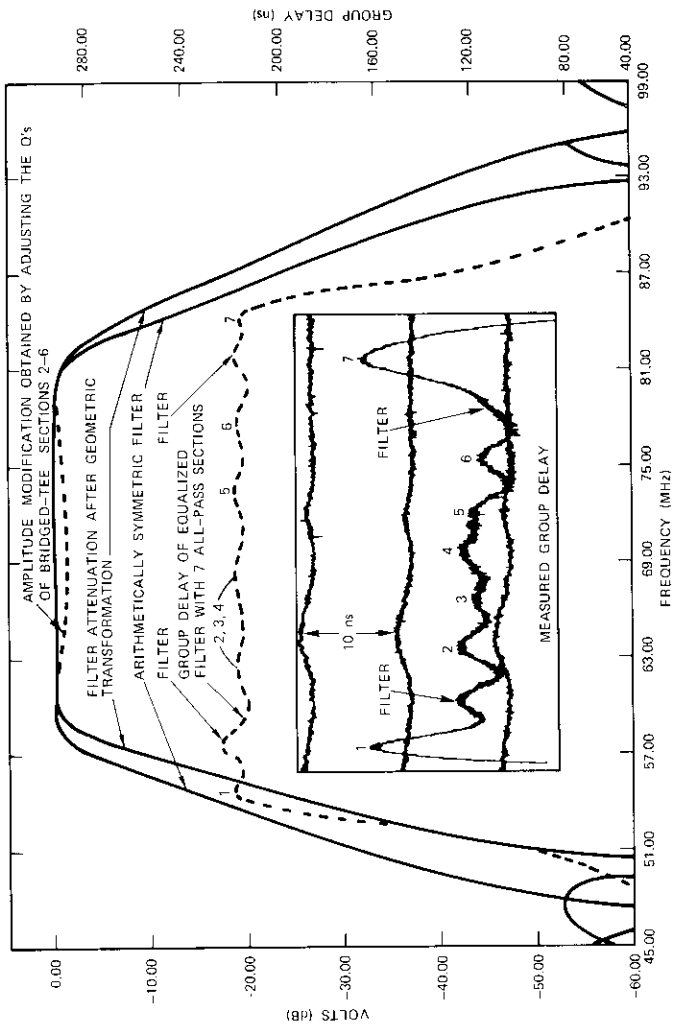



Figure 5. *Filter Characteristics*



$$C_m = C\frac{f^2 - f_o^2}{f_m^2 - f_o^2}, \qquad L_m = \frac{1}{C_m(2\pi f_m)^2} \tag{16}$$

where $C$ and $f$ are the capacitance and the resonant frequency of the original tank circuit, respectively, and $C_m$, $L_m$, and $f_m$ are the modified parameters. The resulting modifications are presented in the table of Figure 4, in the characteristics of Figure 5, and in the measured group delay of Figure 6. The only penalty resulting from these simple corrections is a rise in the maximum VSWR from 1.065 to 1.11.

**Group-delay equalization**

Since the general problem involves a multidimensional space of nonlinear equations, optimization techniques using a digital computer approximate the desired characteristics by minimizing some error function $\epsilon(\omega, \sigma_1, \ldots, \sigma_n, \omega_1, \ldots, \omega_n)$.

The group delay of an all-pass function is given in terms of its real part, $\sigma_n$, and its imaginary part, $\omega_n$:

$$T(\omega, \sigma_n, \omega_n) = \frac{2}{\sigma_n}\left[\frac{1}{1 + \left(\frac{\omega + \omega_n}{\sigma_n}\right)^2} + \frac{1}{1 + \left(\frac{\omega - \omega_n}{\sigma_n}\right)^2}\right] . \tag{17}$$

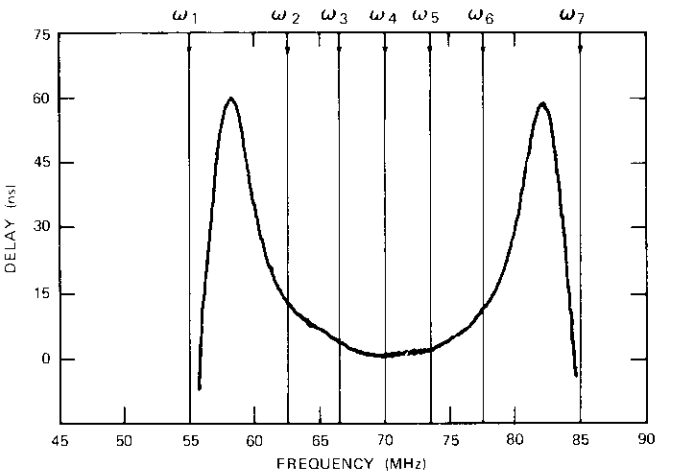



Figure 6. *Measured Group Delay for the Filter of Figures 4 and 5 [Also shown are the imaginary parts for the all-pass initial conditions:* $(\sigma_2, \omega_2)$ *to* $(\sigma_6, \omega_6)$ *are intended to equalize the filter passband;* $(\sigma_1, \omega_1)$ *and* $(\sigma_7, \omega_7)$ *extend the range of equalization to the 15-dB cutoff frequencies.]*

A typical all-pass function group delay is shown in Figure 7. The contribution of $N$ all-pass functions in cascade is

$$T_{\text{equalizer}}(\omega, \sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N) = \sum_{n=1}^{N} T(\omega, \sigma_n, \omega_n) \qquad (18)$$

and the total delay through a filter and an equalizer is

$$T_T(\omega, \sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N) = T_{\text{filter}}(\omega) + \sum_{n=1}^{N} T(\omega, \sigma_n, \omega_n) \quad . \qquad (19)$$

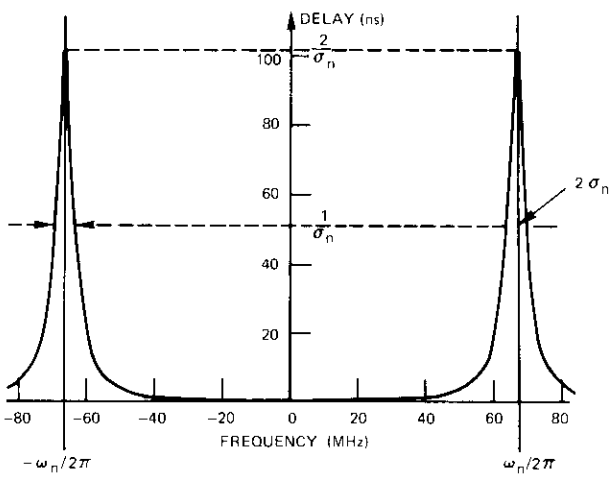



Figure 7. *Delay Characteristic for a Typical All-Pass Function*

The equalization problem requires the determination of a set of $N$ all-pass singularities in $2N$ parameters so that the error function is minimized over a range of frequencies, $B$, about the carrier frequency, $\omega_0$. That is,

$$\epsilon(\omega, \sigma_1, \ldots, \omega_N, \omega_1, \ldots, \omega_N) = W(\omega) \; |T_T(\omega, \sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N) - T_A(\sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N)|^{\gamma} \qquad (20)$$

for $|\omega - \omega_0| \leq B/2$, where $W(\omega)$ is a weighting function, or the inverse of a specification window, describing the relative importance of each



frequency component. The exponent $\gamma$ helps to determine an equalization trajectory; i.e., if $\gamma$ is increased from 2, the optimum solution is likely to approximate a maximally flat group-delay response about $T_A(\sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N)$. $T_A$ is usually the average delay over the specified bandwidth range, $B$:

$$T_A(\sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N) = \frac{1}{B} \int_{\omega_0 - (B/2)}^{\omega_0 + (B/2)} T_T(\omega, \sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N) \, d\omega \quad . \qquad (21)$$

In general, all optimization techniques are basically iterative. For a given set of initial conditions or guesses $(\sigma_n, \omega_n)$, the singularity migration will follow a trajectory in a multidimensional space to reach a local minimum unless the solution diverges or is bounded by one of the external constraints. Then all optimization results are essentially dependent on the choice of the initial conditions $(\sigma_n, \omega_n)$ and the number, $N$, of all-pass functions. Reference 2 describes the procedures for estimating the minimum number of all-pass functions and locating the initial guesses $(\sigma_n, \omega_n)$ in the $s$-plane.

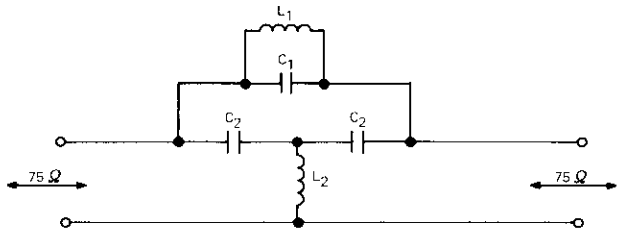
Two techniques have been used to implement numerical solutions to the aforementioned problem on a digital computer. First, with a penalty function used to convert the constrained problem into a sequence of unconstrained minimizations, an efficient optimization routine described in Reference 6 was programmed successfully. The second routine minimized $\epsilon(\omega, \sigma_1, \ldots, \sigma_N, \omega_1, \ldots, \omega_N)$ in equation (20) by sequentially moving each equalizer singularity.

To obtain the channel output waveform, the filter group delay is equalized beyond the 15-dB-attenuation cutoff frequencies. Five all-pass functions equalize the filter passband and two additional functions are used to increase the equalization bandwidth.

Seven bridged-tee networks are used to provide all-pass characteristics to equalize the filter of Figure 4. The design parameters for the seven sections are shown in Figure 8. Each bridged-tee network is used to simulate the 2nd-order all-pass function given by equation (B) in Figure 9.

#### Amplitude modification in the passband

In the bridged-tee networks of Figures 8 and 9, there are five reactive elements, two of which are dependent. It is therefore expected that the



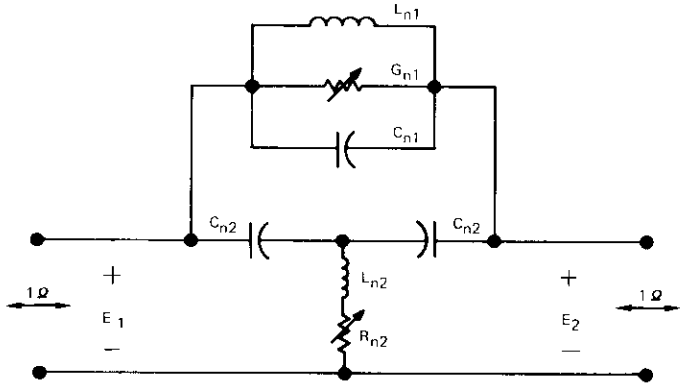

| SECTION NUMBER | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $L_1$ (nH) | 39.42 | 54.45 | 62.10 | 45.96 | 36.12 | 28.75 | 16.50 |
| $C_1$ (pF) | 216.34 | 120.15 | 93.02 | 113.78 | 127.36 | 142.23 | 211.35 |
| $L_2$ ($\mu$H) | 1.23 | 0.689 | 0.539 | 0.652 | 0.725 | 0.807 | 1.193 |
| $C_2$ (pF) | 3.50 | 4.84 | 5.52 | 4.09 | 3.21 | 2.56 | 1.47 |
| FREQUENCY (MHz) | 54.50 | 62.23 | 66.22 | 69.60 | 74.20 | 78.71 | 85.23 |
| PEAK DELAY (ns) | 131.0 | 74.0 | 57.0 | 69.0 | 77.0 | 86.0 | 127.0 |

Figure 8. *Bridged-Tee Networks Used to Group-Delay Equalize the Delay Characteristics of Figure 7 (the computed filter-equalizer group delay is shown in Figure 5)*

characteristic function will be a 4th-order expression [see equation (A) in Figure 9]. The resistive loss in each bridged-tee network causes a migration of the pole-zero in the $s$-plane so that controlled attenuation may be obtained by properly setting the element $Q$'s without changing the group delay of the lossless network, as shown in Figure 10. It should be noted that a $Q$-match must be maintained in the section, i.e., $Q_{TANK} = Q_{T\text{-}SECT}$; otherwise, the bridged-tee network does not present an image termination (Figure 11) and will alter the performance of neighboring sections.

The amplitude modification needed to approximate the design requirements of Figure 3a is obtained by selectively adjusting the $Q$ of the bridged-tee network. The resulting attenuation is shown in Figure 5.






Let $\rho_n^2 = \sigma_n^2 + \omega_n^2;\ \sigma > 0;\ \omega_n^2 > 3\sigma_n^2$

$L_{n1} = 4\sigma_n/\rho_n^2$ $\qquad C_{n1} = (\rho_n^2 - 4\sigma_n^2)/(4\sigma_n\rho_n^2)$

$L_{n2} = 1/(4\sigma_n)$ $\qquad C_{n2} = 2\sigma_n/\rho_n^2$

$$\frac{E_2}{E_1} = \frac{S^4 + a_3S^3 + a_2S^2 + a_1S + a_0}{S^4 + b_3S^3 + b_2S^2 + b_1S + b_0} \tag{A}$$

where $a_0 = \rho_n^4$ $\qquad b_0 = a_0$

$a_1 = 4\sigma_n\rho_n^2\,[G_{n1} + R_{n2}]$ $\qquad b_1 = a_1 + 4\sigma_n\rho_n^2$

$a_2 = 2(\rho_n^2 - 2\sigma_n^2) + 16\sigma_n^2 G_{n1}R_{n2}$ $\qquad b_2 = a_2 + 8\sigma_n^2 + 16\sigma_n^2 R_{n2}$

$a_3 = 4\sigma_n[G_{n1} + R_{n2}]$ $\qquad b_3 = a_3 + 4\sigma_n$

*Note:* If $G_{n1} = 0$, $R_{n2} = 0$, then $a_1 = 0$, $a_3 = 0$. The lossless bridged-tee network and pole-zero cancellations occur at $S = -\sigma_n \pm j\omega_n$. The resulting voltage ratio is reduced to a "bandpass" all-pass function:

$$\frac{E_2}{E_1} = \frac{S^2 - 2\sigma_n S + \rho_n^2}{S^2 + 2\sigma_n S + \rho_n^2} \tag{B}$$

Figure 9. *Transfer Characteristics of Normalized Bridged-Tee Networks*

**Realization**

The elliptic function filter shown in Figure 4 was designed, fabricated, and tuned at the 70-MHz center frequency. The "infinite rejection" elements were modified according to equation (16) to obtain arithmetic

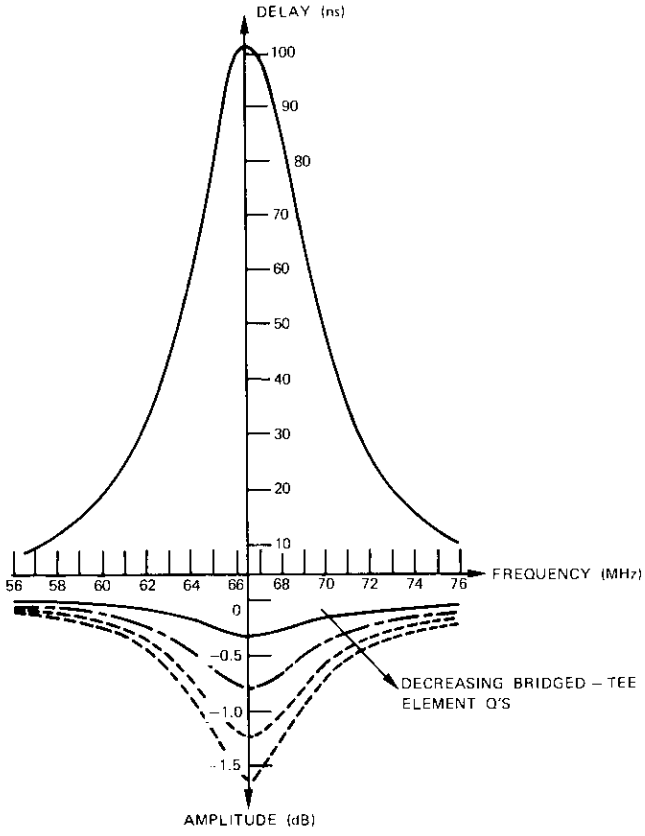



Figure 10. *Realization of Controlled Attenuation by Properly Setting the Element Q's without Changing the Group Delay of the Lossless Network*

symmetry, since the low-pass to bandpass transformation results in geometric symmetry. As a result, the filter return loss dropped from 30 dB to 25 dB. With five all-pass functions in the passband, the arithmetically symmetric filter was group-delay equalized using a digital computer optimization routine, and one section was added on either side of the passband to extend the range of group-delay flatness. The measured group-delay ripple was less than 5 ns for 100 percent of the filter passband and did not exceed 15 ns for filter attenuation frequencies up to 15 dB. These



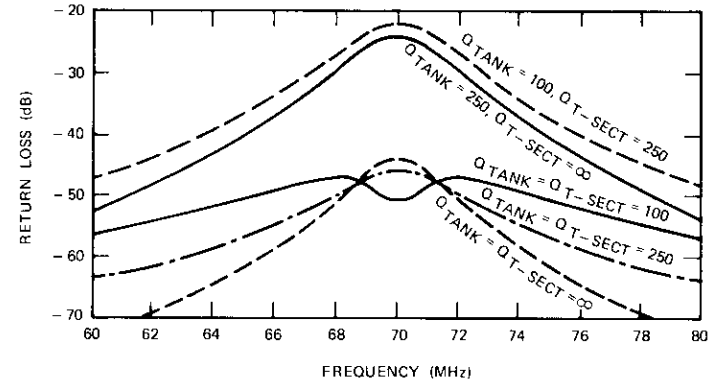



Figure 11. *Computed Return Loss for a Typical Equalizer Section*

all-pass functions were realized with lumped element bridged-tee networks. The inductor $Q$'s were loaded selectively according to equation (A) to modify the passband amplitude of the filter to satisfy the requirements of equation (11), thus obtaining equally spaced time-axis crossings in the rectangular pulse response.

The computed and measured characteristics of the filter are shown in Figure 5. The measured pulse response (Figure 12) is arithmetically symmetrical about its mid-bit, indicating that, in terms of the output pulse, the filter transfer function was perfectly equalized and could be described mathematically by a real part and a constant delay.

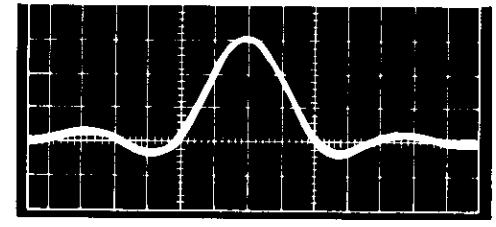


Figure 12. *Measured Pulse Response for the Filter of Figures 3–5, 7, and 8*

## Conclusions

The practical design and physical implementation of transmitter-receiver filtering functions to optimize (with respect to the lowest bit

error) the performance of sequential data transmission through a linear channel perturbed by white Gaussian noise and adjacent frequency-multiplexed channels have been described. The results satisfy both frequency- and time-domain specifications.

The design procedure requires the design of arithmetically symmetric bandpass filters followed by bridged-tee networks for group-delay equalization and amplitude modification. All of these computational steps have been integrated into a single program that takes approximately 4.6 minutes of computations and 17.7 minutes of input/output operations* to design, analyze, and plot the characteristics, and to give fabrication and tuning information for the filter, the bridged-tee networks, and the cascaded circuits.

Nearly optimum transmitter-receiver filters were obtained by designing two identical 5-pole elliptic function filters that were separately group-delay equalized and amplitude modified with seven bridged-tee networks. The computed bit-error rate [7] for these cascaded filters was within 0.4 dB of the theoretical optimum; the major contributor to this small degradation was a small asymmetry in the transmitter filter passband. These filters were later fabricated and tuned with lumped elements; the output at 70 MHz is shown in Figure 2b.

The proposed design techniques have made it possible to minimize the noise level, intersymbol interference, and bit stream crosstalk in 4-phase PSK. The resulting bit-error rate is within 0.4 dB of the theoretical optimum for highly selective filters chosen to suppress adjacent channel interference and to minimize performance degradation caused by multipath effects in a channelized satellite repeater such as INTELSAT IV.

## Acknowledgments

*The author wishes to acknowledge Mr. J. Vergelli for writing the group-delay optimization program, and Mr. L. Parker for writing, modifying, and combining the computer programs used for this task. He also wishes to express his sincere appreciation to Mr. C. Wolejsza for his work on the filters.*

* Input-output operations required to read or store data on disks and/or to write information on a tape for plotting the desired outputs.



## References

[1] W. R. Bennett and J. R. Davey, *Data Transmission*, New York: McGraw-Hill Book Co., 1965, pp. 61–63, 99–110.

[2] F. T. Assal, "Amplitude and Group Delay Equalized Bandpass Elliptic Function Filters," National Electronics Corp., Chicago, Illinois, 1969, pp. 319–324.

[3] A. Papoulis, *The Fourier Integral and Its Applications*, New York: McGraw-Hill Book Co., 1962, pp. 120–136.

[4] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York: McGraw-Hill Book Co., 1968, pp. 45–54.

[5] A. T. Zverev, *Handbook of Filter Synthesis*, New York: John Wiley & Sons Inc., 1967.

[6] L. S. Lasdon and A. D. Waren, "Optimal Design of Filters with Bounded Lossy Elements," *IEEE Transactions on Circuit Theory*, CT-13, No. 2, June 1966.

[7] O. Shimbo and M. J. Celebiler, "The Probability of Error Due to Intersymbol Interference and Gaussian Noise in Digital Communication Systems," *IEEE Transactions on Communication Technology*, COM-19, No. 2, April 1971, pp. 113–120.

## Appendix A. Optimization of digital system performance

A class of pulse output transforms required to satisfy Nyquist's first criterion has been given in equation (1) of the paper; the complex restrictions on these transforms were given in equation (2). Hence, the output waveform of $H_d(f)$ may be obtained as a function of real and imaginary parts. The waveform corresponding to the time response of the real part is arithmetically symmetric about $t = 0$, the pulse mid-bit, while the waveform corresponding to the imaginary part displays odd symmetry about the same time axis. Except at $t = 0$, where the real part waveform displays its maximum, the two waveforms cross the time axis periodically at the sampling instants, although there may be additional crossings. Furthermore, for approximately realizable filters, the contribution of the imaginary part reduces the ringing amplitude for $t < 0$ and increases it for $t > 0$. As a result, the maximum output waveform does not occur at $t = 0$, but at $0 < t = t_M \ll T$, where $T$ is the bit duration and the mid-bit amplitude is unchanged.

In References A1 and A2, it has been shown that, once a pulse output transform satisfying the complex restrictions of equation (2) is chosen, the optimum

receiver filter for a signal perturbed only by white Gaussian noise is given by equation (11). Proper selection of $H_d(f)$, the time-translated output transform, results in further optimization of the system performance, as will be shown in the following.

The indicator of digital system performance is bit-error rate (BER) as a function of carrier-to-noise ratio. In the presence of white Gaussian noise introduced between transmitter and receiver filters, BER is easily computed when the output transform satisfies Nyquist's first criterion for equally spaced axis crossings in the time domain. In this case, BER is given as a function of sample amplitude to noise-root-mean-square ratio, $\sqrt{v^2(t_0)/N}$:

$$\frac{v^2(t_0 = 0)}{N} = \frac{\left\{\int_{-f_c}^{f_c} H_0\, df + \int_{-2f_c}^{2f_c} [H_R(f) + jH_x(f)]df\right\}^2}{2\eta\left[\int_0^{f_c}\left\{|H_0 + H_R(f)|^2 + H_x^2(f)\right\}^{1/2} df + \int_{f_c}^{2f_c} [H_R^2(f) + H_x^2(f)]^{1/2}\, df\right]} \quad \text{(A1)}$$

where $\eta$ is the 2-sided white Gaussian noise spectrum density and $H_1(f) = H_R(f) + jH_x(f)$. To minimize the BER, this ratio is maximized by properly selecting $H_1(f)$. This starting point assumes the output transforms given by equation (1), the complex restrictions of equation (2), and an optimum receiver filter given by equation (11).

In the numerator of equation (A1), since $H_x(f)$ displays odd symmetry about $f = 0$, its contribution to the mid-bit sample amplitude, $v(t_0 = 0)$, vanishes. Similarly, since $H_R(f)$ displays odd symmetry about $f = f_c$ (the Nyquist frequency), its contribution also vanishes. The sample amplitude is then independent of $H_1(f) = H_R(f) + jH_x(f)$ and is given by the area under the curve of an ideal filter,

$$v(t_0 = 0) = \int_{-f_c}^{f_c} H_0\, df = 2H_0 f_c \quad . \quad \text{(A2)}$$

The ratio is then maximized by merely minimizing the noise contribution. This is achieved in two steps. First, since the integrands in equation (A1) are quadratic, the integrals are decreased by perfectly group-delay equalizing the output transform, i.e., $H_x(f) = 0$. Then,

$$\int_0^{f_c} |H_0 + H_R|\, df + \int_{f_c}^{2f_c} |H_R(f)|\, df < \int_0^{f_c} [(H_0 + H_R)^2 + H_x^2]^{1/2}\, df + \int_{f_c}^{2f_c} (H_R^2 + H_x^2)^{1/2}\, df \quad . \quad \text{(A3)}$$



The noise, $N$, is further reduced if

$$|H_0| \geq |H_R(f)|$$

and

$$H_R(f) \leq 0 < H_0 \quad \text{(A4)}$$

or

$$H_R(f) \geq 0 > H_0 \quad .$$

Then

$$\int_0^{f_c} |H_0 + H_R|\, df + \int_{f_c}^{2f_c} |H_R(f)|\, df = \int_0^{f_c} |H_0|\, df \quad \text{(A5)}$$

so that

$$N = 2\eta\, |H_o|\, f_c \quad \text{(A6)}$$

and the ratio in equation (A1) is independent of the particular shaping function $H_R(f)$, namely,

$$\max\left\{\frac{v^2(t_0 = 0)}{N}\right\} = \frac{2\,|H_o|\, f_c}{\eta} \quad . \quad \text{(A7)}$$

A similar analysis will show that the required transmitted power is also minimized if $H_x(f) = 0$ and the inequalities in equation (A4) are satisfied. Therefore output transforms with constant group delay would further optimize the performance of a digital system provided that $H(f)$ satisfies Nyquist's criterion for equally spaced time-axis crossings and the receiver filter is designed according to equation (11).

## References

[A1] W. R. Bennett and J. R. Davey, *Data Transmission*, New York: McGraw-Hill Book Co., 1965, pp. 61–63, 99–110.

[A2] R. W. Lucky, J. Salz, and E. J. Weldon, Jr., *Principles of Data Communication*, New York: McGraw-Hill Book Co., 1968, pp. 45–54.

*François T. Assal received B.E.E. and M.E.E. degrees from City College New York. As a member of the RF Transmission Laboratory of* COMSAT *Laboratories he has participated in transponder design and utilization to maximize satellite communications in bandwidth-restricted and power-limited environments.*

*Prior to joining* COMSAT *in 1969, he held a National Science Foundation Traineeship and was a lecturer at C.C.N.Y. He is a member of Tau Beta Pi and Eta Kappa Nu.*

**Index: synchronous satellite, ATS-F, propagation, electromagnetic wave transmission, earth terminals.**

# ATS-F COMSAT millimeter wave propagation experiment

L. H. WESTERLUND, J. L. LEVATICH, AND A. BUIGE

## Abstract

The ATS-F COMSAT Millimeter Wave Propagation Experiment has been designed to gather statistical data on the attenuation caused by rain at millimeter wave frequencies. These data will be used to determine system design parameters for future communications satellite systems operating at frequencies above 10 GHz. The experiment has 39 ground terminals transmitting at 13.2 or 17.8 GHz to a transponder on board the ATS-F satellite. The transponder retransmits these signals at 4 GHz to a central earth terminal which records their amplitudes once each second.

The data will be analyzed to provide probabilities of attenuation as functions of parameters such as rainfall, location, and time. These probabilities can then be used to determine the required power margins of millimeter wave communications systems. Techniques of overcoming severe attenuation such as site diversity and the use of a spot beam to increase the power level at selected locations will also be evaluated. This paper includes an overall description of the experiment as well as a discussion of system performance and operation.

## Introduction

The ATS-F COMSAT Millimeter Wave Propagation Experiment is designed to gather data on satellite signal attenuation at 13 and 18 GHz caused by atmospheric hydrometeors (primarily rain) associated with

These data will be used to determine the minimum power margins needed for spacecraft communications systems which operate at frequencies above 10 GHz.

At present, 4-GHz systems operate with typical power margins in the range of 3 to 6 dB. As the operating frequency is increased, the attenuation caused by rainfall increases rapidly; hence, unreasonably large power margins are required to obtain the same overall system reliability.

Diversity techniques offer a means of reducing the power margin requirements. The deepest fades are usually the result of intense localized rain cells typical of thunderstorms. Separating two or more stations by some distance makes it possible to reduce significantly the margin required to provide a given grade of service for a specified percentage of the time. To plan such systems intelligently, the joint distribution of attenuation magnitude and duration as a function of separation distance and climatological conditions must be considered.

For operation at frequencies around 13 GHz, a spot-beam technique might be useful. A spot beam could overcome localized fading by temporarily increasing the power margin to a local area. The spot beam could be moved so that all local areas would receive protection for a percentage of the time. However, to evaluate this technique, it is necessary to correlate the attenuation among terminals typically distributed over a large geographic area.

There are a number of theories that can be used with sufficient confidence to predict attenuation caused by hydrometeors in the propagation path if the precipitation parameters used in theoretically derived equations are known. Unfortunately, the meteorological aspects of the problem, such as the distribution of the various drop sizes in 3-dimensional space, the actual shape of drops as a function of size and velocity, and the frequency of occurrence of different concentrations of hydrometeors of a given size, are inadequately defined.

To obtain the necessary system performance information, a measurement program is needed. This program should yield data which are in a form directly applicable to satellite system design, and which are sufficiently accurate and detailed to permit empirical tests of the theories and further scientific analysis.

The COMSAT millimeter wave experiment will yield statistical distributions of signal attenuation from 24 different locations over a significant part of the U.S. Fifteen of these locations will simultaneously provide data at 13 and 18 GHz. This experiment will satisfy the need for a temporal distribution of attenuation data over widely varying climatic and geographic areas. Each location will also be equipped with a rain gauge to gather data for correlation with RF attenuation. This correlation will permit potential extrapolation of the attenuation statistics to other locations. The gathered and analyzed data will help to determine the operational parameters for future satellite communications systems.



## System description

The millimeter wave propagation experiment consists of ground and spacecraft hardware plus the data reduction and analysis equipment required to process the results of the experiment into a manageable and usable form. The general configuration, shown in Figure 1, will consist of 15 dual-frequency transmitting terminals which transmit signals randomly spaced between 13.19–13.20 and 17.79–17.80 GHz. These terminals will be located throughout the eastern half of the U.S. and will be spaced at least 160 km apart. In addition, there will be nine diversity terminals which transmit signals between 17.79–17.80 GHz. Three diversity terminals separated by less than 40 km will be grouped around each of three dual-frequency terminals to provide 4-terminal diversity configurations transmitting at 17.8 GHz. A detailed description of these transmit terminals may be found in Appendix A.

The spacecraft transponder will receive the signals from the transmitting terminals, translate the 13-GHz signals to 4.140–4.150 GHz and the 18-GHz signals to 4.160–4.170 GHz, and then retransmit them. A detailed description of the transponder built by COMSAT is given in Reference 1.

The spacecraft receiving antenna, furnished by the spacecraft contractor, is a dual-frequency parabolic reflector with linear polarization in the east-west plane. Its gain is 25.8 dBi at 13 GHz and 27.6 dBi at 18 GHz. Its boresight will be offset by 2.9° to the east of the spacecraft pointing axis to provide coverage of the ground transmitting sites for each of the two primary spacecraft pointing modes. Figure 2 shows the resulting coverage of the eastern U.S.

The per-carrier power output of the transponder at 4 GHz will be −36 dBW for the diversity experiment's 18-GHz carriers, and −29 and −35 dBW for the 18- and 13-GHz dual-frequency terminal carriers, respectively. The spacecraft's 4-GHz horn antenna, which will serve as the transponder transmit antenna, has a half-power beamwidth of 20° and an on-axis gain of 17 dBi.

The receiving terminal at Andover, Maine, will receive the signals at 4 GHz, separate the 39 individual carriers, and record them for further

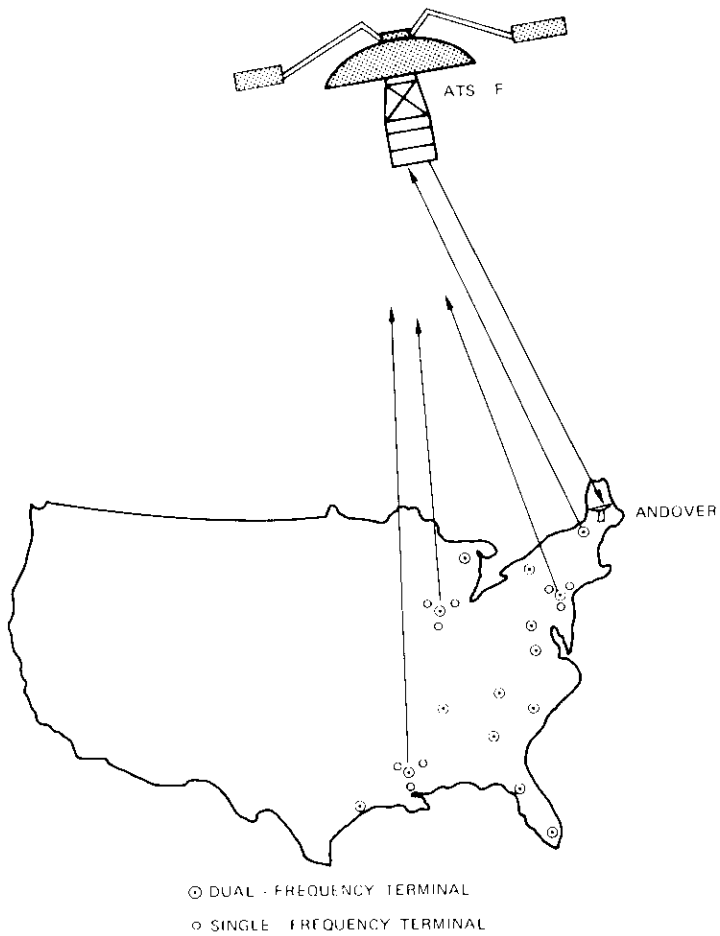



Figure 1. *Propagation Experiment System*

analysis. A brief description of the receive terminal may be found in Appendix A.

The process control computer and its peripheral equipment will control the operations of the receiving equipment and perform the measurements. The power level of each of the received carriers will be scanned once each second by a signal multiplexer. The process control computer




Figure 2. *Spacecraft Receive Antenna Coverage*

will use the gain and noise temperature values determined in the calibration process to calculate the actual received power level of each carrier from these readouts. These results, along with other information, such as identification, carrier number, and time of day, will be recorded on magnetic tape.

The COMSAT propagation experiment has several inherent advantages for the collection of the type of data desired for system design. First, to gather statistical information of any significance in a short period of time, continuous operation is needed. Because the prime power required for the spacecraft transponder is only 11 W, 24-hour-a-day operation is feasible on the multiexperiment, power-limited, ATS-F spacecraft [2]. This low power requirement is achieved by performing the attenuation measurement on the up-link so that the high power requirements are on the ground.

To minimize the 4-GHz satellite output power requirements, a large receiving antenna is used. The down-link is at 4 GHz where the attenuation is insignificant and hence does not contaminate the up-link signal variations.

A second advantage of the COMSAT propagation experiment is the use of a large number of small transmit terminals to permit simultaneous determination of the desired transmission and propagation parameters from many locations. Finally, the concept of the single receiving terminal permits simultaneous collection and recording of the data, with excellent time correlation between signals from widely varying paths. This is very useful in the data analysis program, in which the signals received from different terminals will be correlated.

The data gathered from the 15 widely separated, dual-frequency sites will be used for correlating signal attenuation among any number of terminals. Determining the number of terminals which simultaneously experience fades at a given fade level will assist in evaluating the possibility of providing extra power (e.g., via spot beam) for communications links suffering attenuation loss because of precipitation in the path.

In addition, the single-frequency transmitting terminals will be used to evaluate the smaller scale (less than 40-km) spatial distribution of severe attenuation caused by intense rainfall. These terminals will be arranged to yield a number of separation distances. Correlation analyses of attenuation at these terminals, taken two at a time, will reveal the separation with the lowest probability of simultaneous high attenuation over two satellite links.

The duration of the experiment will be sufficient to permit a good statistical comparison of the measured attenuation at a site and general meteorological parameters such as rainfall rate, number of thunderstorm days, and total precipitation.

## *System performance*

The desired fade measurement range of the experiment must be sufficient to provide the information needed to evaluate alternative schemes for overcoming rain attenuation in an eventual millimeter wave satellite communications system. Estimates of the state-of-the-art indicate that a future millimeter wave satellite communications system might economically provide less than a 10-dB margin against attenuation caused by frequently occurring moderate rain rates. It also appears feasible to use a steerable high-gain antenna on the satellite combined with a higher power transmitter tube to provide an additional 20-dB margin on a particular transmission path. Therefore, to evaluate the feasibility of this scheme, an attenuation measurement range of 30 dB at 18 GHz over the transmission paths from the 15 widely separated transmitting locations has been established. Since the loss at 13 GHz will be less than the corresponding loss at 18 GHz, a measurement range of 20 dB at 13 GHz has been established.



Over the transmission paths from the closely spaced transmitters, a somewhat smaller measurement range is needed. It is expected that switchover to a diversity terminal will occur before the attenuation over a particular transmission path reaches the 10- to 15-dB range. Thus, a 20-dB attenuation measurement range has been specified for the transmission paths that will be used to assess the feasibility of earth terminal diversity.

Since the signals are translated in frequency and amplified in the spacecraft transponder, they are subjected to nonlinearities which generate intermodulation cross products, primarily in the TWT. These cross products may fall on or near a wanted signal and cause erroneous data during severe signal attenuation. However, the most efficient use of down-link power occurs when the TWT is operated near saturation, which is also the region of large nonlinearities which produce intermodulation cross products [3]. Therefore, the TWT operating point has been chosen to achieve a compromise between intermodulation level and output power per carrier.

Theoretical calculations indicate that, with an input operating point of $-7$ dB with respect to saturation, the errors caused by amplitude nonlinearity will be less than $\pm 0.5$ dB. The intermodulation products generated will be $>6$ dB below the noise level of the receiving system (Figure 3).

A substantial increase in the carrier-to-intermodulation ratio is also obtained by randomly spacing the signals over the band. This random spacing spreads the intermodulation products so that they do not coincide with the signal frequencies.

Based on the preceding requirements, as well as on size and power constraints, the RF link power budgets in Tables 1 and 2 were developed as the optimum compromise for the overall system. Each entry is shown at its nominal or clear sky value. The dynamic range is the range, or depth of fades caused by rainfall, which the experiment is capable of measuring. Since several of the parameters, e.g., the position of the satellite with respect to antenna patterns, vary with time, this range also varies with time. The effect of these variations and other tolerances has been calculated separately and is shown as a tolerance on the overall dynamic range of the system.

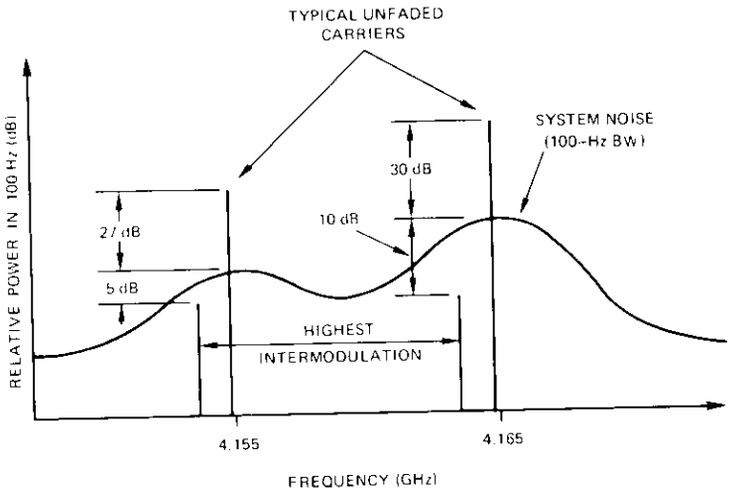



Figure 3. *Representative Down-Link Spectrum*

The experiment will be controlled and conducted from the central receiving terminal at Andover, Maine. Personnel of this terminal will monitor the performance of all ground transmit terminals and the spacecraft transponder and will initiate corrective actions if necessary. Also associated with the central receiving terminal will be one of the dual-frequency ground transmitters, modified to include two 1.8-m parabolic reflectors. These reflectors will be mounted on the horn antenna and boresighted with it so that they will always be properly aimed at the satellite.

Remotely controlled attenuators will be inserted in this system to maintain the terminal's output power at the level provided by the standard terminals. If there is rain attenuation in this path, sufficient attenuation will be removed from the ground link to keep the level of these two control carriers at the spacecraft receiver at the clear sky value. These two carriers will also be used as the autotrack beacons and phase-lock references (see Appendix A). If attenuation on this path becomes too severe, these functions will be temporarily transferred to some other path which is not experiencing a severe fade. This control terminal will also allow end-to-end calibration of the overall system at any time during the experiment.



TABLE 1. UP-LINK PERFORMANCE PARAMETERS

| Parameter | Attenuation Experiment 13 GHz | Attenuation Experiment 18 GHz | Diversity Experiment 18 GHz |
|---|---|---|---|
| Transmitter | | | |
| Power (dBW) | +14 | +14 | +7 |
| Antenna Gain (dB) | 32.5 | 32.5 | 32.5 |
| e.i.r.p. (dBW) | 46.5 | 46.5 | 39.5 |
| Free-Space Propagation Loss (dB) | −206.2 | −209.8 | −209.8 |
| Average Antenna Misalignment | | | |
| Transmitter (dB) | −1.5 | −1.5 | −1.5 |
| Satellite Receiver (dB) | −0.7 | −1.5 | −1.5 |
| Total (dB) | −2.2 | −3.0 | −3.0 |
| Satellite | | | |
| On-Axis Antenna Gain (dB) | 25.7 | 28.6 | 28.6 |
| Receive Power (dBW) | −136.2 | −137.7 | −144.7 |
| Noise Power Density (dBW/Hz) | −194.5 | −194.5 | −194.5 |
| Up-Link Performance | | | |
| $C/N_o$ (dB-Hz) | 58.3 | 56.8 | 49.8 |

TABLE 2. 4-GHZ DOWN-LINK PERFORMANCE PARAMETERS

| Parameter | Attenuation Experiment 13 GHz | Attenuation Experiment 18 GHz | Diversity Experiment 18 GHz |
|---|---|---|---|
| Satellite | | | |
| Transponder Gain (dB) | 101 | 109 | 109 |
| Power/Carrier (dBW) | −35.2 | −28.7 | −35.7 |
| Antenna Gain (less loss) (dB) | 16.0 | 16.0 | 16.0 |
| e.i.r.p. (dBW) | −19.2 | −12.7 | −19.7 |
| Free-Space Propagation Loss (dB) | −196 | −196 | −196 |
| Antenna Misalignment (dB) | 0 | 0 | 0 |
| Earth Terminal G/T (dB/°K) | 37 | 37 | 37 |
| Down-Link Performance | | | |
| C/N (dB-Hz) | 50.4 | 56.8 | 49.9 |
| System Performance | | | |
| $C/N_o$ Total (dB-Hz) | 49.7 | 53.9 | 46.8 |
| Receive System Bandwidth (dB) | 20 | 20 | 20 |
| Minimum C/N for Retaining Phase Lock (dB) | 3 | 3 | 3 |
| Measurement Dynamic Range (dB) | 26.7 (+1.2 to −2.4) | 30.8 (+1.9 to −2.7) | 23.8 (+1.9 to −2.7) |

## Data collection

The data to be used in the analysis will include propagation data from the central receiving terminal, telemetry and orbital data from NASA, rain gauge data from each of the transmitter sites, and calibration data from the power monitors on each transmitter. Data processing (Figure 4) comprises two separate functions: data reduction and data analysis. For the rain gauge and transmitter power data, data reduction simply consists of manual scanning of the stripcharts. The rain gauge data are placed on computer cards for entry into the computer and further analysis in conjunction with the propagation data. Transmitter power levels are manually scanned for significant changes and outages.

The major effort in the data reduction task consists of calibration, quality checks, and editing of the propagation data. Because of the large amount of data involved, it is necessary to edit the data as early as possible, deleting the majority of the data which contain no fades. However, many experiment parameters, such as satellite motion across antenna beams and transponder temperature changes, may cause data variations which the computer may interpret as fades to be saved for analysis. Therefore, the data are corrected for these variations before they are analyzed.

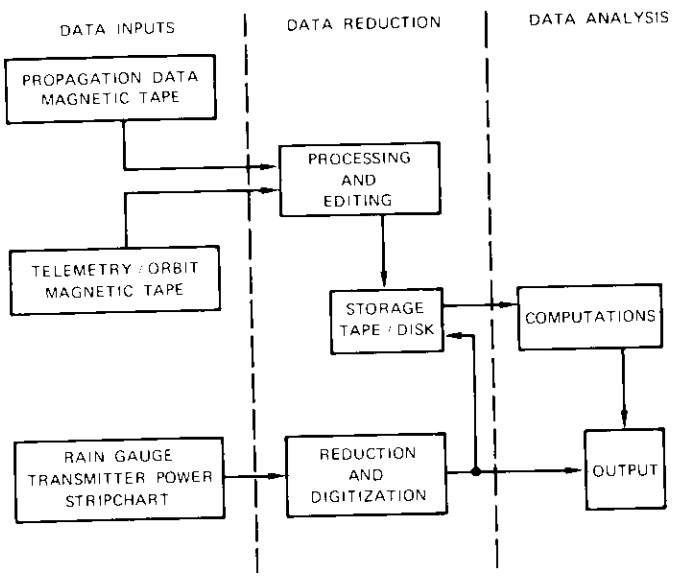



Figure 4. *Data Flow Chart*



Once the raw data have been edited and corrected, the actual data analysis is performed. All analyses are statistical and typically result in graphs of attenuation probability as a function of such parameters as rainfall, time, and location.

One-dimensional probability-density functions and cumulative distribution functions will be generated for the attenuation experienced by each of the 39 signals. These are the most basic statistics needed to determine the reliability for a given fixed margin against rain attenuation.

A fade duration histogram will also be prepared. This histogram shows the distribution of the durations of various depths of fade, indicating the time duration of service interruptions for a given fixed margin against attenuation.

Further, a joint fading probability for the diversity terminals will be established. This analysis will be performed on the signals from the three diversity clusters which have a total of four transmission paths to the satellite. For any two of these transmission paths, it will be necessary to determine the percentage of time during which both paths simultaneously experience various fading levels. Six such analyses will be required at each of the three locations, including all combinations of the four transmission paths, taken two at a time, to yield the best diversity terminal arrangement for overcoming periods of high attenuation.

Similarly, the joint fading probability will be calculated for the 15 widely separated terminals. This computation, which determines the percentage of time that the attenuation at one selected site exceeds various levels, will indicate the percentage of time during which a high-e.i.r.p. satellite spot beam will be available to overcome attenuation at that site.

In addition to these analyses of the digitized signal amplitude data, the rain gauge data from each of the 24 transmitting locations will be manually analyzed to yield cumulative distribution functions for the rain rate at each location. This information will be used to extrapolate the measured attenuation data to other locations where rain data may be available. In addition, attenuation as a function of rain and rain rate will be determined for special events.

## Acknowledgment

*The authors wish to give due credit to Mr. E. Podraczky, who conceived the overall experiment and was instrumental in its materialization; to Mr. J. Kiesling, who first introduced the concept of the up-link measurement; and to Dr. C. Mahle and Mr. E. Steinbrecher, who are responsible for*

*the transponder and receiving terminal, respectively. This work was supported by the National Aeronautics and Space Administration under Contract NAS5-21616.*

## References

[1] "The ATS-F COMSAT Propagation Experiment Transponder," edited by A. Berman, *COMSAT Technical Review*, Vol. 3, No. 2, Fall 1973.

[2] H. L. Gerwin, "ATS F&G Pioneer Application of Space Technology," *Journal of Environmental Sciences*, Vol. 15, April 1972, p. 12.

[3] O. Shimbo, "Effects of Intermodulation, AM-PM Conversion and Additive Noise in Multicarrier TWT Systems," *Proceedings of the IEEE*, Vol. 59, No. 2, February 1971, p. 230.

## Appendix A. Transmitting terminals and receive system

### Transmitting terminals

The two types of transmitting terminals used in this experiment are dual-frequency terminals and single-frequency diversity terminals. These terminals are essentially identical except for power output and frequencies. As shown in Figure A-1, the basic configuration consists of a small parabolic-reflector antenna (positioned manually) with a prime focal feed, a power amplifier, a frequency generator, and power monitoring system. A rain gauge, a stripchart recorder, and an auxiliary power system are also provided to record the power outputs and rain rates at each transmitter site.

Each dual-frequency transmitting terminal is assigned two independent frequencies, one between 13.19 and 13.20 GHz, and one between 17.79 and 17.80 GHz. The diversity terminals, which are assigned frequencies in the 17.79- to 17.80-GHz band, will be interleaved with the dual-frequency terminals. The half-power beamwidth of the fixed antennas is 4° by 4° at both 13 and 18 GHz with a manual pointing accuracy of 0.1°. The antennas are linearly polarized in the same plane for both frequencies.

The power output of the power amplifiers is 25 W at 13 and 18 GHz for the dual-frequency terminals and 5 W at 18 GHz for the diversity terminals. The output stability is ±0.5 dB for a 72-hour period. The frequency stability is better than $1 \times 10^{-7}$ over a temperature range of −40°C to +45°C with a power source variation of ±5 percent in frequency and/or ±10 percent in voltage.

A rain gauge will be installed at each terminal to record rain rates of 5 to 150 mm/hr, with a resolution of 2 mm/hr. A stripchart recorder, which is capable of operating at a speed of 120 mm/hr, will record the transmitter output power and the accumulated rainfall.



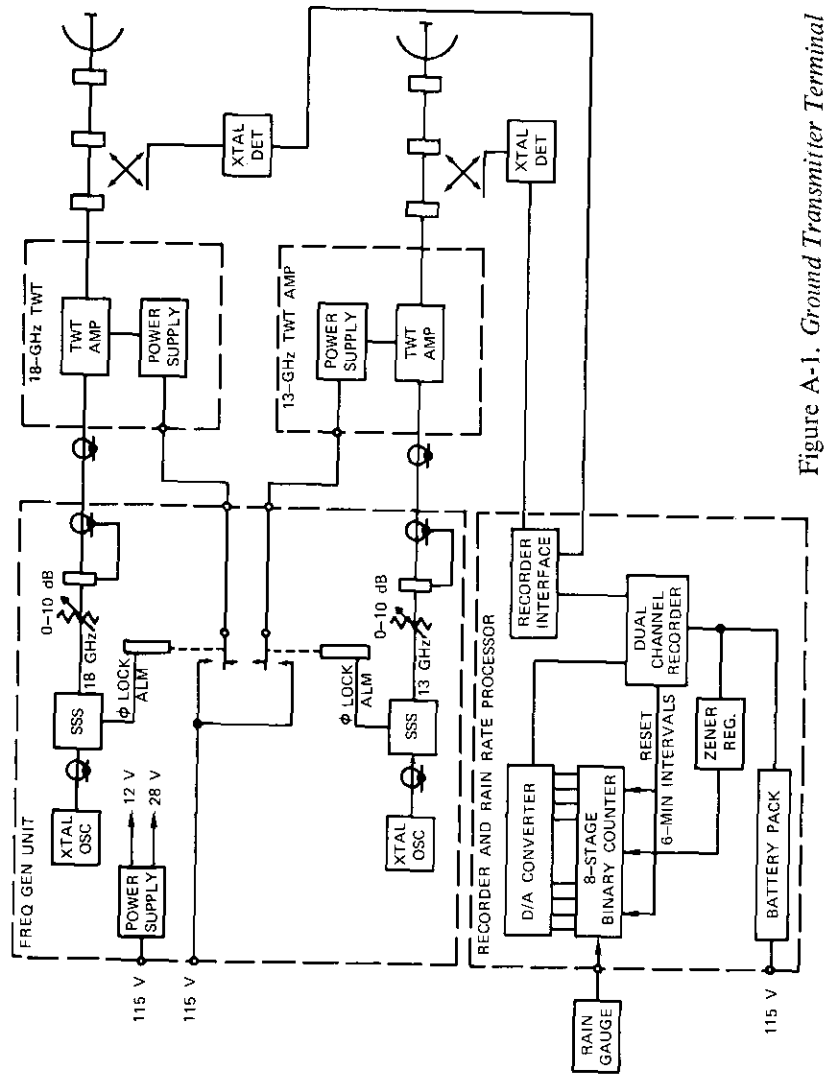



Figure A-1. *Ground Transmitter Terminal*

To obtain information on the distribution of high-intensity precipitation attenuation over the greatest possible area, the 15 terminals transmitting at 13 and 18 GHz will be located in various parts of the U.S. Probable sites are Boston, Massachusetts; Columbus, Ohio; Starkville, Mississippi; Andover, Maine; Ithaca, New York; Detroit, Michigan; Philadelphia, Pennsylvania; Washington, D.C.; Nashville, Tennessee; Asheville, North Carolina; Fayetteville, North Carolina; New Orleans, Louisiana; Atlanta, Georgia; Tampa, Florida; and Miami, Florida. These sites are expected to represent most of the climatic regions having significant precipitation in the U.S.

Three of these sites, Boston, Columbus, and Starkville, will also be part of the experiment to evaluate diversity receiving systems.

Since the terminals are located over such a wide geographic region, often in rather remote areas, they are designed for highly reliable unattended operation. They are designed to withstand temperature ranges from $-40°C$ to $+45°C$, winds over 50 mph, and rain rates of 150 mm/hr while operating fully within specifications. To meet these requirements, the terminals are housed in a 1.2- x 1.2- x 2.1-m fiberglass-constructed, weatherproof equipment enclosure with a self-contained heating system to maintain the internal temperature above 10°C during the winter, and a self-ventilating exhaust fan for the summer.

The design parameters of the terminals have been dictated primarily by the requirements of the experiment. The antenna beamwidth has been specified to accommodate the satellite motion, the effective radiated power to meet the dynamic range of the experiment, and the frequency stability, perhaps the most important parameter, to accommodate the receiving system. The receiving system is expected to operate in a 100-Hz loop bandwidth, placing an extremely rigid short-term stability specification on the frequency generators. Similarly, to permit unattended operation of the transmitters, the long-term stability of the generators must be sufficient to keep the carriers within the tracking range of the receiver for a period of not less than six months.

Previous experience indicates that long-term drift rates of $1 \times 10^{-8}$ per day and $1 \times 10^{-6}$ per year can be achieved through the use of phase-lock solid-state oscillators in which crystals in the 100-MHz region are used as the reference frequency. It has been determined that the short-term requirements should be attainable with similar devices. Thus, the basic design of the frequency generators calls for solid-state oscillators, operating in the 1.0-GHz region, to be phase locked to crystal oscillators operating around 100 MHz, followed by an appropriate multiplier chain.

The short-term frequency stability has been specified in terms of a power spectral density. It is characterized in the bandwidth from 0 to 30 Hz by a $K/f^3$ distribution so that

$$10 \log \frac{P_T}{P(10)} \geq 30 \text{ dB}$$



where $P_T$ is the total carrier power and $P(10)$ is the power measured in a 1-Hz bandwidth 10 Hz from the carrier. Beyond 30 Hz, the power spectral density has been specified as shown in Table A-1.

TABLE A-1. POWER SPECTRAL DENSITY SPECIFICATION BEYOND 30 HZ

| Frequency Separation from the Carrier | Power Spectral Density Below the Carrier |
|---|---|
| 30 Hz $\leq$ df $<$ 100 Hz | $>$ 45 dB |
| 100 Hz $\leq$ df $<$ 1 kHz | $>$ 55 dB |
| 1 kHz $\leq$ df $<$ 10 kHz | $>$ 65 dB |
| 10 kHz $\leq$ df $<$ 100 kHz | $>$ 75 dB |
| 100 kHz $\leq$ df | $>$ 80 dB |

Figure A-2 shows the power spectral density of a typical frequency generator compared with the specification. To perform the measurement, two frequency generators were mixed together and the beat frequency was fed to a narrowband wave analyzer in which the power spectral density was measured in a 1-Hz bandwidth.

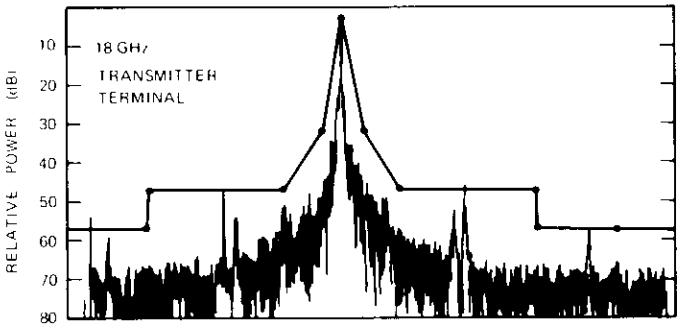



Figure A-2. *Power Spectral Density of a Typical Frequency Generator Compared with the Specification*

**Receive system**

The receiving terminal, to be located at the Andover, Maine, earth terminal, will use the horn antenna and associated tracking and 4-GHz receiving equipment. As shown in Figure A-3, the receive system consists of three sections: receiving terminal, signal processing, and data acquisition. The calibration unit will be installed between the receiving antenna and the low-noise parametric amplifier to permit calibration of the entire system.

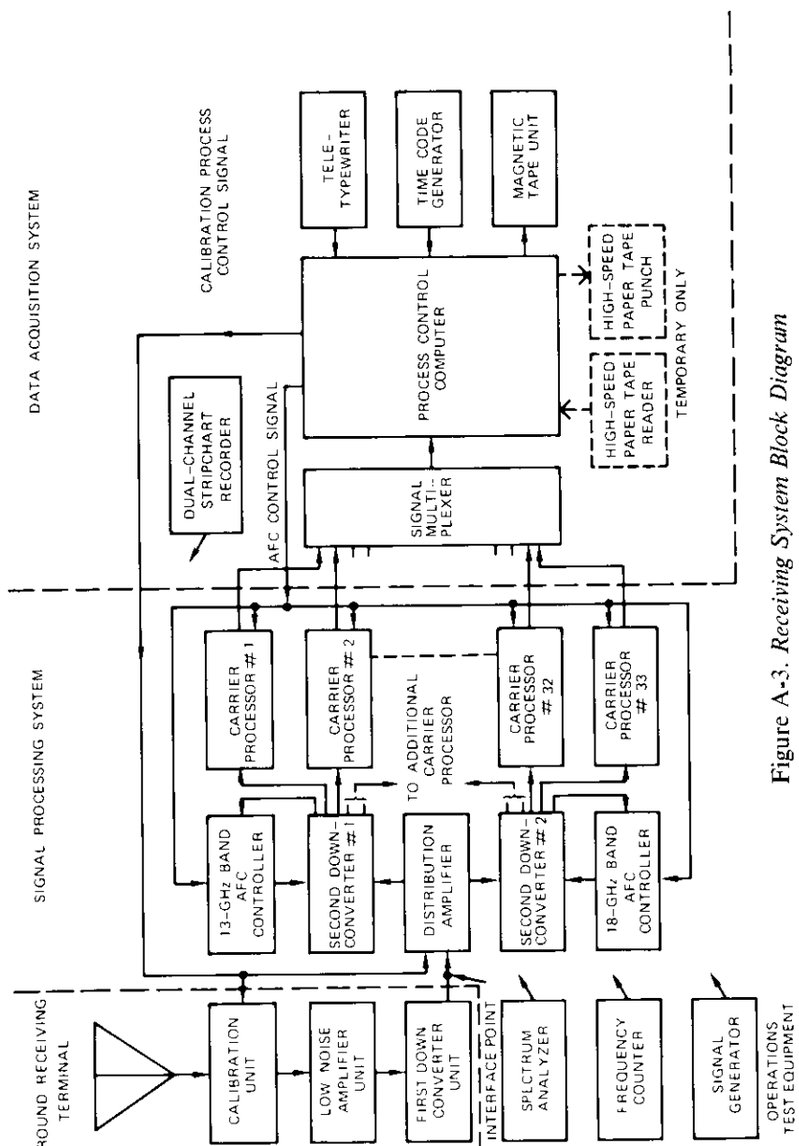



Figure A-3. *Receiving System Block Diagram*



The first down-converter will convert the amplified signals to the 70-MHz range. The 70-MHz signals will then be amplified and distributed by the distribution amplifier to the second down-converter. The local oscillator signal for the second down-converters will be supplied by 13- and 18-GHz AFC controllers, which will follow any shifts of the two bands caused by the frequency converters in the satellite.

A distribution amplifier will split the single 30-MHz-wide receive signal into two 10-MHz-wide bands. One band will contain the 13-GHz transmitted carriers and the other band will contain the 18-GHz transmitted carriers. The 10-MHz bands will be converted to 20 MHz and then amplified to drive the carrier processors, which convert the signal to DC and filter out noise. The system bandwidth will be determined by a crystal filter with a 100-Hz bandwidth in the carrier processor.

*Lawrence H. Westerlund is a Member of the Technical Staff of the Special Projects Division of* COMSAT *Laboratories. Since 1971, he has also served as Assistant Project Manager for the* COMSAT *ATS-F Propagation Experiment. Previous experience at* COMSAT *Labs includes research performed in the Propagation and Space Physics Branches. Before joining* COMSAT *in 1968, he was a Research Associate at Rice University, where he received a B.A. in Physics (1964) and a Ph.D. in Space Science (1968).*

*Julius L. Levatich is Manager of the Propagation Department of* COMSAT *Laboratories and Project Manager and NASA Principal Investigator for the ATS-F Propagation Experiment. Prior to joining* COMSAT *in 1966, he was employed by the MITRE and Bendix Corporations. He received a B.E.E. from R.P.I. in 1955 and an M.S. from Johns Hopkins in 1963. He is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi. He is also a Registered Professional Engineer and a senior member of IEEE.*

*Anthony Buige received a B.S. degree in mathematics from St. Francis College in 1959. He spent seven years with the NASA Goddard Space Flight Center, where he was the Manager of the Data Acquisition Section in the Tracking and Data Systems Directorate. He joined* COMSAT *Laboratories in 1967. He was formerly Manager, Laboratory Simulation, in the Systems Division, engaged in transmission modeling and system simulation programs. He is a member of IEEE and Delta Epsilon Sigma.*

**Index: synchronous satellites, ATS-F, electromagnetic wave transmission, radio transponders.**

# The ATS-F COMSAT propagation experiment transponder

A. L. BERMAN, EDITOR *


## Abstract

This paper describes a satellite transponder for use in the ATS-F COMSAT Propagation Experiment. In operation, this transponder receives 39 continuous wave carriers transmitted in the 13.2- and 17.8-GHz bands, frequency translates, and then retransmits in the 4-GHz band. The hardware design includes several novel features. Three-stage tunnel diode amplifiers, implemented with microwave integrated circuit technology, provide the major portion of the repeater gain. Image terminated frequency translation mixers incorporating dual-mode filters to minimize weight and size have been used. Transponder reliability is highlighted throughout the design. Final flight qualification testing has verified that all performance requirements are conservatively met.


## Introduction

The ATS-F COMSAT Propagation Experiment repeater has been designed to satisfy the performance requirements dictated by the overall ex-

* The material reported herein is the result of the collaborative work of many people. The principal contributors are A. Atia, A. Berman, R. Bounds, R. Cooperman, Y. Lee, C. Mahle, R. Stegens, R. Strauss, J. Talcott, M. Wachs, and A. Williams.

periment plan described in "ATS-F COMSAT Millimeter Wave Propagation Experiment" by L. Westerlund, J. Levatich, and A. Buige [1]. The hardware for the experiment consists of a spacecraft transponder, 15 ground transmitter stations, and one receiving station. The transmitter stations transmit a total of 39 continuous wave carriers, 15 in the 13.2-GHz band and 24 in the 17.8-GHz band. The spacecraft transponder translates these carriers to the 4-GHz band, amplifies, and retransmits them to a central receiving earth station. The received signal strength of each carrier is recorded and correlated with rain gauge readings from the remote sites.

This paper describes the design and performance of the spacecraft transponder. The requirements of this transponder, determined from the overall experiment plan, are presented in Table 1. The experiment plan calls for accurate measurement of carrier power level over a 30-dB dynamic range. In practice, the lower limit of this range is set by the noise figure, while the upper limit is determined by the linearity of the repeater. The frequency translation jitter specification is derived from considerations arising from the narrowband tracking loop used to receive the carriers. The power consumption, size, and weight constraints are imposed by the ATS-F satellite environment.

TABLE 1. TRANSPONDER REQUIREMENTS

| | |
|---|---|
| Input (receive) Frequencies | 13.19–13.20 GHz, 17.79–17.80 GHz |
| Output (transmit) Frequencies | 4.14–4.15 GHz, 4.16–4.17 GHz |
| Overall Gain (input flange to output flange) | 100–110 dB |
| Gain Stability vs Temperature | ±1 dB (+5°C to +35°C) |
| Noise Figure | <10.5 dB |
| Output Power | +18 dBm |
| Frequency Translation Jitter (short term) | < 10-Hz deviation at rates <1 Hz |
| Power Consumption | <14 W |
| Size | |
| Receiver | 27.9 cm × 22.8 cm × 16.5 cm |
| Transmitter | 30.5 cm × 22.3 cm × 17.4 cm |
| Maximum Weight | <16.8 kg |



## *Overall repeater description*

A repeater consisting of a receiver and transmitter satisfying the requirements of Table 1 is shown in Figure 1. The receiver performs the required input filtering, frequency translation, preamplification, and bandpass filtering. The transmitter provides high-level amplification to the required output power level.

### Receiver

Each receiver chain consists of a low-conversion-loss mixer followed by a 3-stage tunnel diode amplifier (TDA). The TDA supplies approximately 35 dB of gain and sets the receiver noise figure. A 10-MHz-wide bandpass filter limits the receive noise bandwidth at the output of the receiver. The local oscillator (LO) signal required for each of the receiver mixers is obtained from independent crystal oscillator-multiplier chains. Two receiver chains are provided for each received band to enhance the experiment reliability.

### Transmitter

The transmitter combines the outputs of both channels at 4 GHz and amplifies them in either of two high-level amplifier chains. A 3-stage TDA provides approximately 35 dB of gain, while an output traveling wave tube (TWT) supplies 50 dB of gain.

### Major repeater subsystems

The design of the major subsystems is described in the following subsections. The unique features are emphasized.

#### MIXERS

The mixers incorporate filters utilizing dual-mode operation, i.e., two electrical cavities in one physical cavity. This technology has resulted from an ongoing program at COMSAT Laboratories to minimize filter weight and size without degrading performance.

The first active stage in each receiver chain is the Schottky barrier diode mixer. The 13- and 18-GHz units are shown in Figure 2. In each case, a minimum conversion loss is necessary to meet the noise figure performance requirements. Providing the mixer diode with a reactive termination at the sum, image, and regenerated signal frequencies makes it possible to reduce the conversion loss below the 3-dB minimum predicted for a resistively matched termination mixer [2].

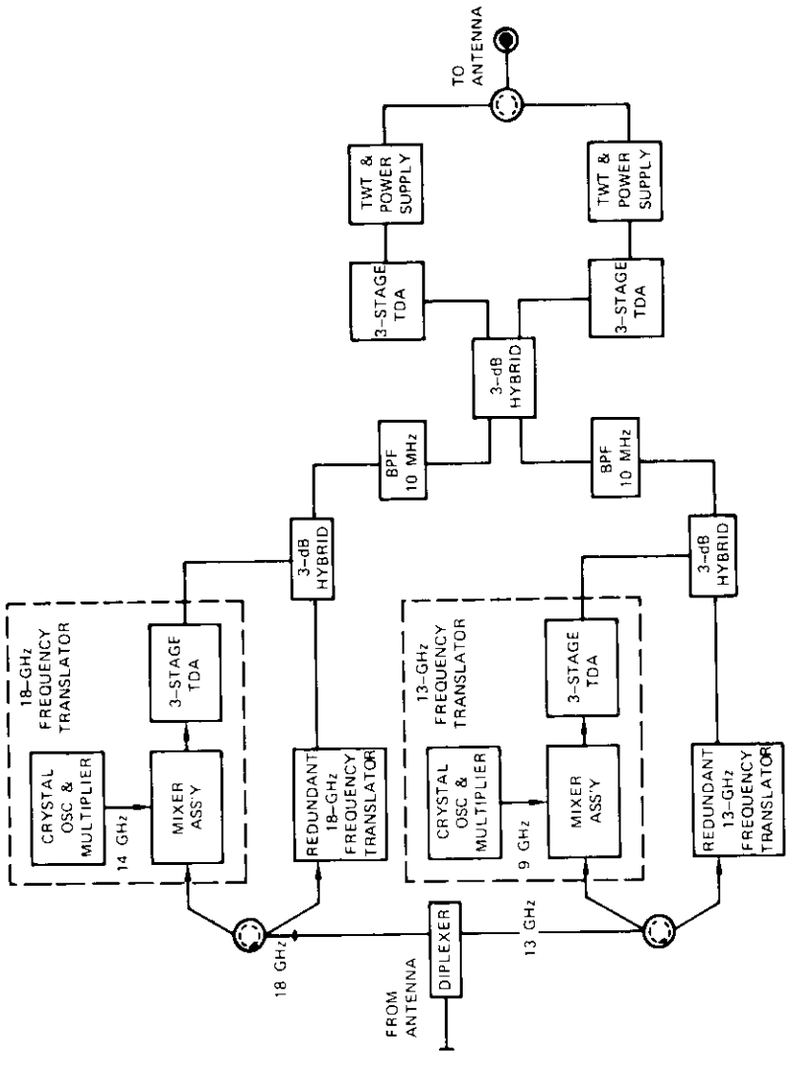



Figure 1. *Overall Repeater Block Diagram*



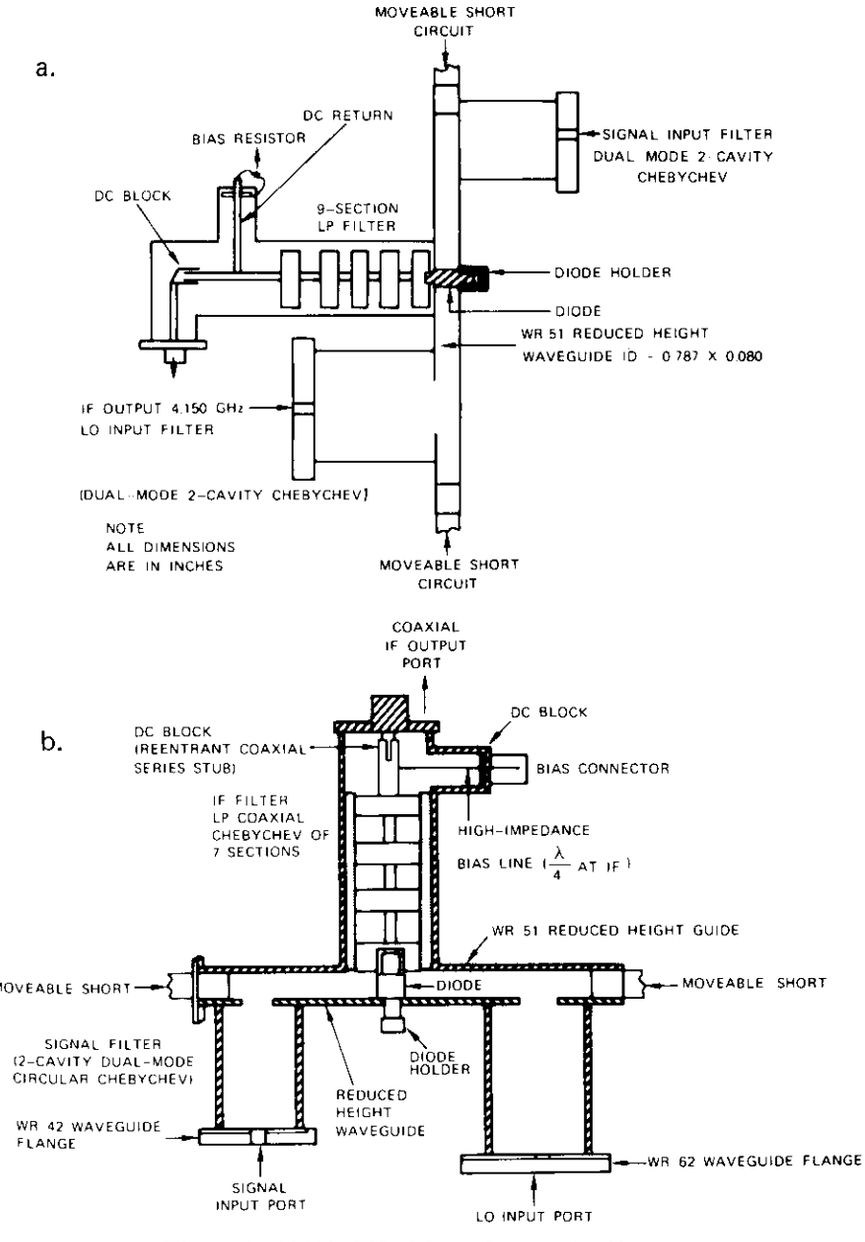



Figure 2. *13/18-GHz Mixer Schematic Diagram*

To maximize the conversion efficiency of the design, it is necessary to minimize the principal losses [3], which include the following:

a. mismatch loss at either the signal or IF port, resulting in less signal being delivered to the diode junction;

b. loss of signal power due to the presence of series resistance and junction capacitance in the diode;

c. losses related to the current-voltage characteristic of the diode junction; and

d. losses caused by the signal and IF filters.

Minimization of these losses requires proper matching of signal and IF sections, optimization of LO drive for the particular diode selected, and presentation of a reactive termination to the diode at frequencies other than the intermediate, LO, and signal frequencies. The diode should be selected for high reliability, minimum noise figure in the frequency band of interest, low series resistance, and large reverse breakdown voltage. Furthermore, the diode package should have a self-resonance frequency higher than the signal frequency. A type HP5082-2723 diode has been chosen.

The mixer equivalent circuit is shown in Figure 3. The Schottky barrier diode is placed perpendicular to the broad wall of a reduced height waveguide to facilitate impedance matching at the signal and LO frequencies. The broad wall dimension has been chosen so that the image is well below the guide cutoff frequency and is therefore reactively terminated.

The LO and input signals are introduced via the broad wall of the guide, from opposite sides of the diode, by using 2-pole filters constructed from single dual-mode circular cavities [4]. Such a filter design significantly reduces the volume and weight of each mixer. Short circuits are provided at both ends of the waveguide. A low-pass filter with a cutoff frequency of 4.5 GHz is used at the output. This filter allows the 4-GHz difference frequency to pass with minimum loss and, at the same time, provides a short circuit for the image, LO, signal, and sum frequencies.

To minimize signal filter losses and to avoid tuning variations caused by temperature changes, the filters have been designed to have 100-MHz bandwidths although the signal bandwidth is only 10 MHz. These wide bandwidths also permit the use of lightweight aluminum. The IF filters have been designed in the manner described in Reference 5. Additional mixer design considerations are given in Appendix A.



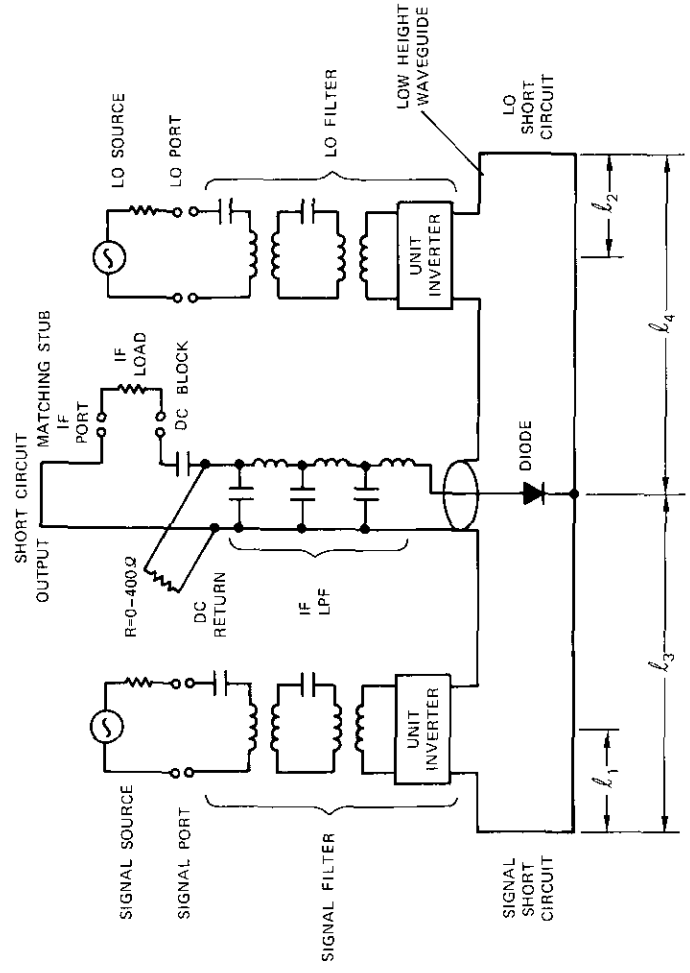



Figure 3. *Mixer Equivalent Circuit*

Mixer performance is summarized in Table 2. The low conversion loss indicates that the image and sum frequencies are reactively terminated. The conversion loss vs LO drive level characteristics are presented in Figure 4.

TABLE 2. MIXER PERFORMANCE

| | 13-GHz Mixer | 18-GHz Mixer |
|---|---|---|
| Waveguide Dimensions | 2 cm × 0.2 cm | 1.3 cm × 0.18 cm |
| Cutoff Frequency | 7.5 GHz | 11.5 GHz |
| Signal Frequency | 13,195 ± 5 MHz | 17,795 ± 5 MHz |
| Signal Filter Loss | 0.3 dB | 0.4 dB |
| LO Frequency | 9,050 MHz | 13,630 MHz |
| LO Power | 10 ± 3 dBm | 6 ± 2 dBm |
| Image Frequency | 4,905 ± 5 MHz | 9,465 ± 5 MHz |
| Return Loss, All Ports | >15 dB | >15 dB |
| Typical Conversion Loss | 3 dB | 2.5 dB |
| Typical Noise Figure | 3 dB | 2.5 dB |
| Weight | 212.6 g | 149.6 g |

During the tuning process, it was observed that the noise figure could increase markedly without appreciably changing the conversion loss. A similar phenomenon had been previously investigated [6] for mixers having equal terminations at the signal and image ports. This effect is primarily attributed to the presence of a nonlinear capacitance in the diode's equivalent circuit. The presence of this nonlinear capacitance, which is pumped by the LO, can, under certain conditions, result in a net conversion gain. Thus, any thermal noise generated in the diode can be amplified parametrically and appear at the IF port, contributing to a noise figure which is larger than that caused by the mixer conversion loss. However, this condition can be avoided if the nonlinear diode resistance and capacitance are pumped in opposite phases, which is possible if the proper reactances are introduced at the diode terminals.

The conspicuousness of this phenomenon in the case of 18-GHz mixer is probably due in part to the fact that the DC return path is a short circuit and in part to the low conversion loss obtained. Nevertheless, for this mixer, it was always possible to find one particular tuning at which the noise figure and conversion loss were in good agreement.



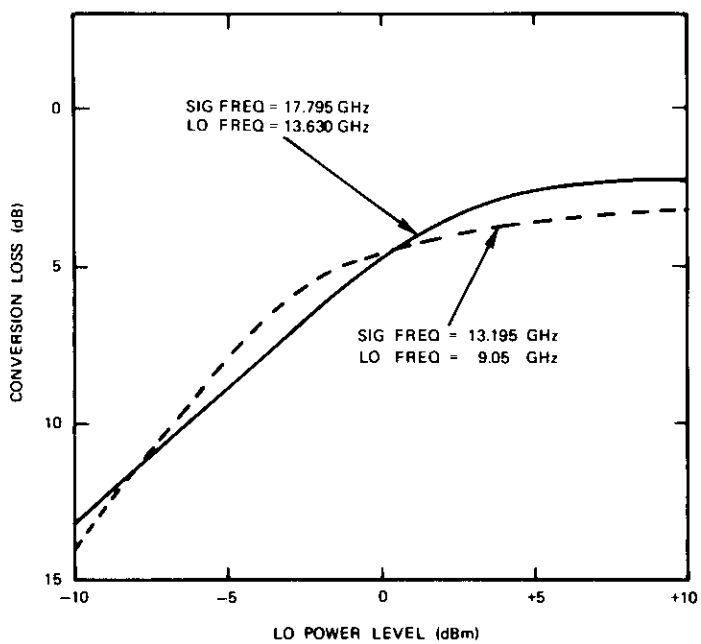



Figure 4. *Flight 13/18-GHz Mixer*

TUNNEL DIODE AMPLIFIER

The design of the TDA is based on a microwave integrated circuit (MIC) approach using microstrip on alumina. Mechanical ruggedness and unit-to-unit uniformity make this implementation attractive for spacecraft applications.

Three-stage TDAs are used to provide low-noise, low-level amplification at 4 GHz. In addition to minimizing the noise figure while providing the correct gain, the design logic has considered the elimination of spurious oscillations and temperature sensitivities.

A germanium tunnel diode with a peak current of 2.5 mA has been selected for the design as a compromise between low noise figure and moderate linear output power capabilities. The admittance vs frequency characteristic of the diode determines the electrical design of the amplifier module and circulator as follows:

a. The capacitive reactance of the diode must be resonated in-band.

b. The circuit must provide a resistive load impedance (smaller than the negative diode impedance) to set the in-band gain.

c. At frequencies outside of the band (i.e., DC to 3.5 GHz and 5.2 to 20 GHz), the circuit must provide a positive resistance with a magnitude greater than that of the diode's negative resistance.

The design is complicated by the large variation with frequency exhibited by the circulator impedance.

The tunnel diode matching and stabilizing circuits are included on a 2.54- x 2.54- x 0.06-cm alumina substrate, as shown in Figure 5. Figure 6 shows the design of the stabilizing network. Two pseudo-lumped networks provide series resonance at 4.15 GHz and form a 2-section band-rejection filter when separated by one-quarter wavelength. Since the input of the 2-section filter (at plane B-B) is a short at 4.15 GHz, the length of lines $L_A$ and $L_B$ is chosen so that these lines will resonate with the diode capacitive susceptance. Line $L_A$, a 90° transformer at 14 GHz, causes the shunt conductance loading the diode above 4 GHz to increase, thereby following the increase in the device's negative conductance.

Because of the lower limitation imposed upon line impedance by the MIC format, two stabilizing and tuning networks are used in the TDA design to double the stabilizing conductance at all frequencies. Stagger tuning the length of line $L_A$ in each network makes it possible to achieve maximum stabilization conductance over a broad range near 14 GHz so that the diode is stabilized in the range where a large negative resistance is exhibited.

Since the net conductance of the circuit is positive at all frequencies outside the range of 3.5–5.2 GHz, no reflection gain and hence no external oscillations can exist. This condition is referred to as conditional terminal plane stability. The effect of circulator loading must, however, be included to ensure stable operation.

Final stability is determined by interconnecting the amplifier to the circulator and checking the Nyquist stability criteria for the entire assembly. A computer technique is used to perform this operation. The program computes the reflection gain of the amplifier at all frequencies of interest. (Theoretical performance is 12-dB gain with a 1-dB bandwidth of 200 MHz.) Stability is analyzed at selected frequencies from DC to 20 GHz by determining the limiting circulator load reflection coefficients which violate the Nyquist criteria. The limiting reflection coefficients define the return loss specification of the circulator ports.



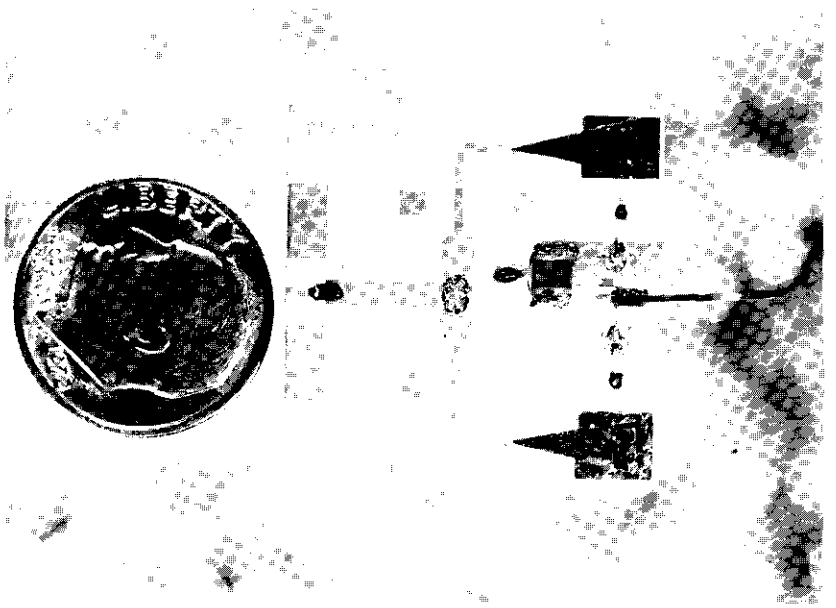


Figure 5. *ATS-F Amplifier Substrate*

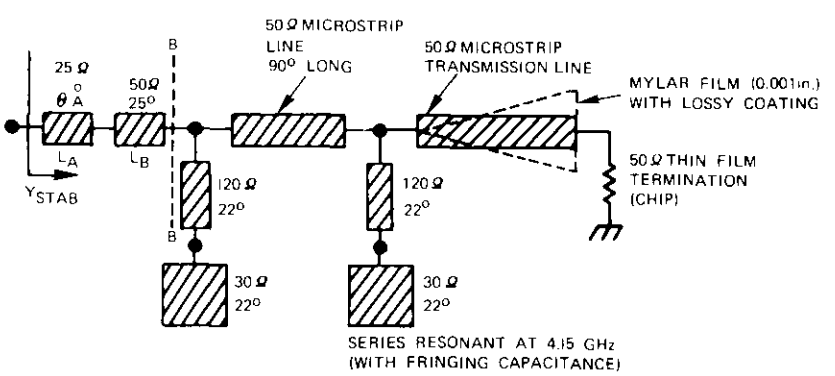



Figure 6. *TDA Stabilizing Network (line lengths are in degrees at 4.15 GHz)*

To eliminate the possibility of interaction between TDA modules, a 9-port, 7-junction circulator has been designed to provide 60-dB isolation between ports. To guarantee the isolation performance, RF-absorbing ferrite mode suppressors have been included to eliminate parallel-plate mode propagation. The three amplifier port voltage standing wave ratios (VSWRs) are less than 1.2:1 in-band and comply with the load reflection requirements determined by the stability analysis.

Circulator characteristics, e.g., match and isolation, were made relatively constant over the qualification temperature range by selecting a ferrite material (Transtech type G1002) which exhibits small variations with temperature, and by using Alnico VIII biasing magnets. A more detailed design computation is presented in Appendix B, and the measured performance of the circulator is summarized in Table 3.

TABLE 3. MEASURED CIRCULATOR PERFORMANCE

| | |
|---|---|
| Gain | 26–35 dB* |
| 1-dB Bandwidth | 90 MHz |
| Center Frequency | 4,150 MHz |
| Noise Figure | <5.5 dB |
| Gain Variation over 0°C–50°C | $<\pm 0.5$ dB |
| Weight | 529 g |

* The gain was set at 26 dB, 33 dB, and 35 dB for amplifiers in the 13-GHz channel, the 18-GHz channel, and the transmitter, respectively.

### TUNNEL DIODE AMPLIFIER BIAS SUPPLY

Each TDA module requires a +3-V bias supply with a high degree of line regulation, low ripple, and constant voltage with time and temperature variation. A high-efficiency switching regulator has been designed to provide 75 mW from the +28-V line with losses of only 121 mW.

### BANDPASS FILTERS

Both the 13- and 18-GHz channels incorporate a filter (at 4 GHz) to limit the noise bandwidth. The filter bandwidth represents a tradeoff between the frequency plan required to provide a flat frequency response over the bandwidth used by the carriers and the input power backoff of



the traveling wave tube amplifier (TWTA). (A bandwidth which is too wide will saturate the TWT with noise.) The temperature range specified for the transponder requires a filter constructed of invar.

The dual-mode filter is similar to that described in the mixer section [4]. Each of the filters consists of a circular waveguide cavity excited by two orthogonal $TE_{111}$ modes coupled by a screw at a 45° angle from the direction of polarization of each mode. The input and output couplings are provided by extending the center conductors of the coaxial line inside the cavity, as shown in Figure 7; hence, waveguide-to-coaxial transitions

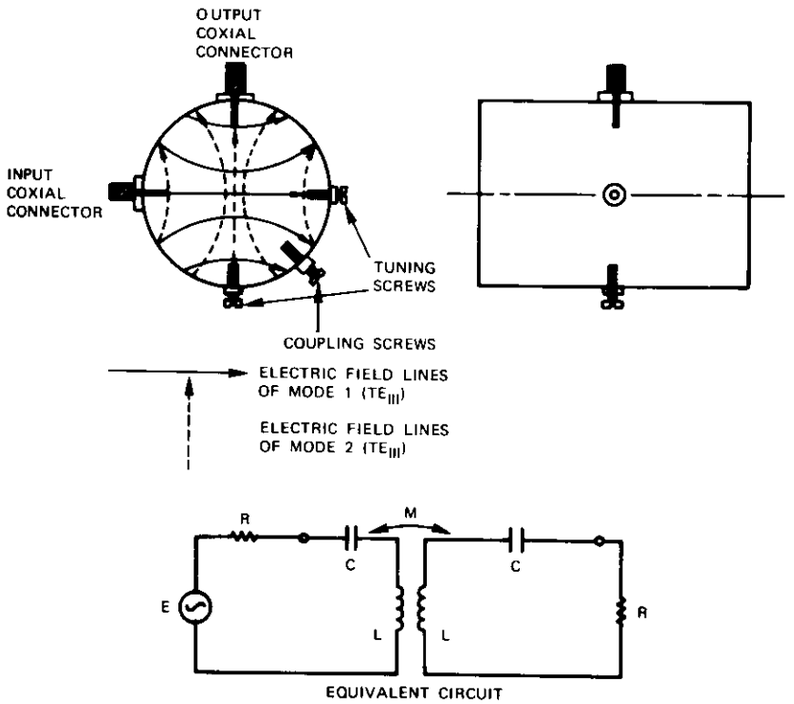



Figure 7. *Modes in the 2-Cavity Filter and Its Equivalent Circuit*

are unnecessary. To conserve further space and reduce the independent mechanical supports, the two filters are constructed as one integral piece.

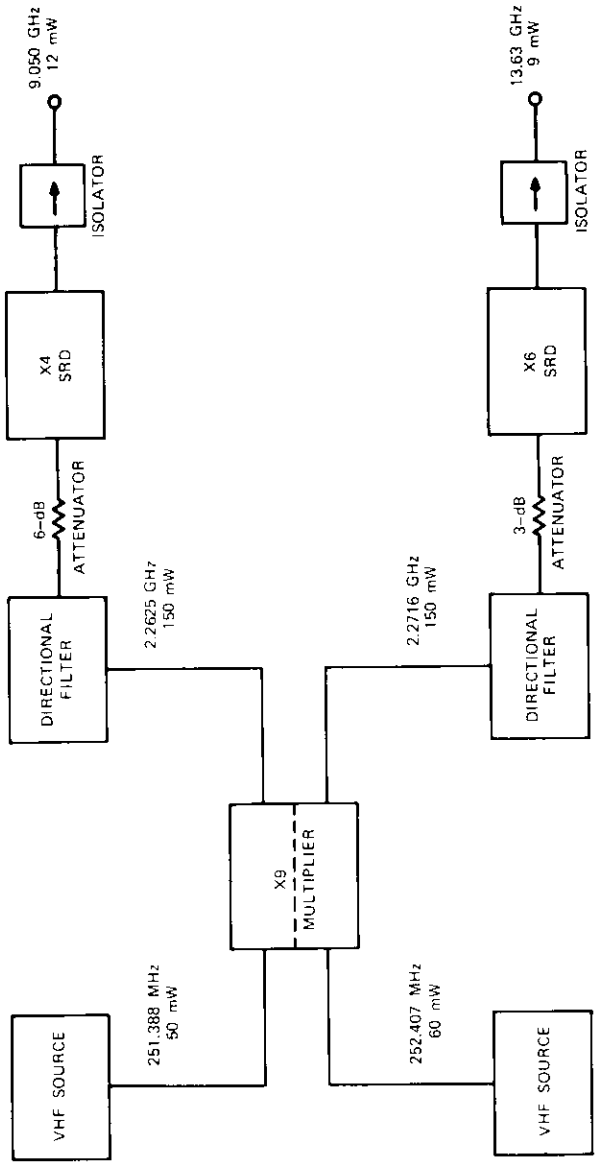



Figure 8. *LO Chains*



LOCAL OSCILLATOR CHAIN

The transponder requires LO sources at 9.05 and 13.63 GHz with a nominal level of 10 mW and a frequency stability of $\pm 4$ ppm over a temperature range of $-10°C$ to $+65°C$. The design of the LO, shown in Figure 8, is straightforward and emphasizes immunity to spurious oscillation.

The first stage of the LO consists of an ultra-stable quartz crystal (5th overtone) oscillator, buffer amplifier, and transistor doubler. A limiting stage minimizes level variation with temperature, and an output filter eliminates spurious outputs. The output is $+17 \pm 0.2$ dBm over 0°C to $+50°C$. Stability is within 2 ppm from $-10°C$ to $+65°C$. (No oven is used.) The oscillator phase noise was measured as described in Appendix C.

A transistor X9 frequency multiplier provides an output of more than 100 mW at approximately 2.25 GHz. The multiplier consists of two transistor triplers, each having $\sim 3$ dB of gain. Lumped element matching at the input (nominally 250 MHz) and in the interstage networks (nominally 750 MHz) minimizes the overall size. The circuit is unconditionally stable for any combination of drive level and frequency, DC voltage, and temperature. The devices typically operate at 10 percent of rated dissipation. At an output level of +21 dBm (levels up to +24 dBm are possible) at 2.25 GHz, the X9 multiplier has an overall DC-to-RF efficiency of 20 percent.

A directional filter at the output of the X9 multiplier provides a wideband 50-ohm match while attenuating unwanted multiplication products. This filter consists of an MIC ring resonator and 50-ohm terminations. It has been designed with a new mode compensation technique [7] applied to the two parallel coupled directional coupler sections of the microstrip traveling wave loop directional filter.

The compensated-design filter circuit has been photo-etched on a 2.54- x 2.54- x 0.1-cm alumina substrate ($\epsilon_r = 9.7$) and has two 50-ohm film resistor internal terminations. It has shown return losses greater than 20 dB up to 4 GHz, which is a substantial improvement over the performance of an uncompensated-design filter. The insertion loss is 1.4 dB at the band center of 2.25 GHz, and the 3-dB bandwidth is approximately 90 MHz.

The final frequency multiplication is performed by a single X4 or X6 step recovery diode multiplier.

The multipliers feature a conversion efficiency of 1/N or better at the design frequency, and stability over the ±250-MHz input bandwidth for power levels from 0–5 dB nominal, thereby demonstrating a state-of-the-art performance.

The LO chain is open-circuit stable at any phase.

### TELEMETRY AND COMMAND MODULE

The telemetry and command (T&C) module is the interface between the experiment and the spacecraft telemetry, command, and power systems. It operates on the spacecraft command signals to control power switching within the experiment and to properly condition telemetry sensor signals.

The telemetry section consists of the signal conditioners and the voltage reference (for thermal sensors) required to properly process the sensor data. It accepts signal inputs from various sources within the experiment, and conditions and buffers these signals before presenting them as outputs to the redundant spacecraft telemetry encoders. There are two types of signal sources: digital status bits and analog voltages derived from thermal sensors.

The command section, which consists of the command interface circuits, command driver circuits, and command inhibit circuit, accepts redundant spacecraft command decoder outputs and generates control signals for the experiment in response to these commands. These responses take the form of signals to switch power to the appropriate redundant circuitry and control pulses to switch the latching circulators and TWTAs.

The power section consists of a DC/DC converter, power disconnect circuit, and distribution relays for power switching. It accepts a regulated 28-V input from a 14-W spacecraft load-interface-controller (LIC). It provides overload protection, voltage conversion, and switching power to the various experiment components. It also provides the required isolation between power, telemetry, and command power supplies. Figure 9 is a block diagram of the command and telemetry units.

## *Transponder integration and test*

An important guideline throughout the transponder development was the objective of solving as many potential problems as possible before or during the engineering model integration to minimize difficulties in the prototype and flight model. Each of the engineering model circuits was



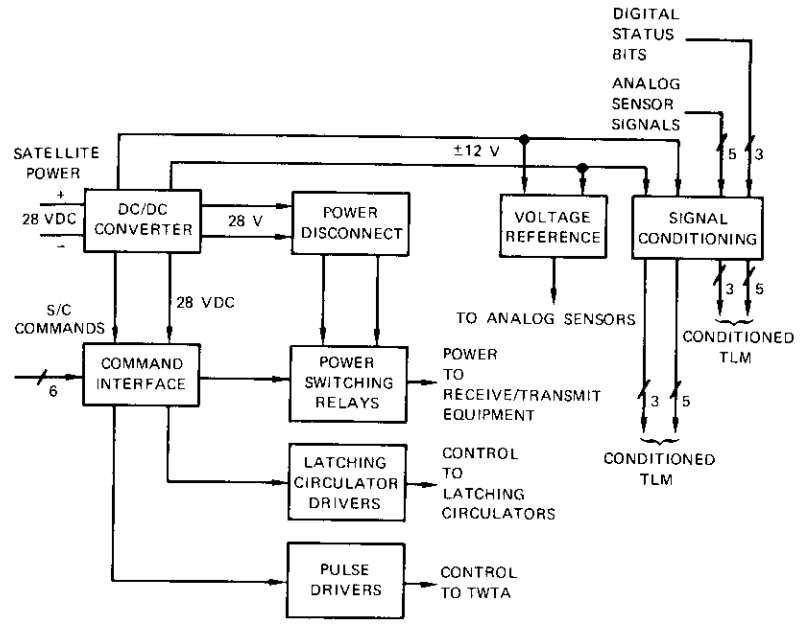



Figure 9. *Telemetry, Command, and Power Block Diagram*

therefore developed until its electrical performance (e.g., swept frequency response, gain, linearity, noise figure, power output, matching, stability, and RF leakage) satisfied predetermined criteria. Potential interaction susceptibility received particular attention to avoid unforeseen integration problems.

After each circuit was brought into compliance with performance requirements and after all measurements verified expected operating characteristics, environmental tests were performed. The program included qualification level vibration and temperature tests over an extended range with the device in operation (i.e., DC and RF signals applied).

Only after the device demonstrated acceptable performance variation during environmental tests was it integrated into the engineering model transponder. This model was then measured on the bench, during vibration and in a thermal vacuum chamber, to verify that the overall performance could be predicted from the performance of the individual subassemblies during unit test. In this way, all major design problems could be solved

during development and integration of the engineering model. Only minor structural problems were encountered during the transponder vibration test. One such problem was solved when the structure supporting the 13-GHz mixers in the receiver was stiffened with three additional brackets after the first round of vibration tests.

No significant design changes were required between engineering model, prototype, and flight model. The subassemblies for the prototype and flight models underwent a similar test cycle, i.e., bench tests followed by live environmental tests and additional bench tests. After integration, formal test programs were conducted for qualification and acceptance of the prototype and flight model transponders, respectively. The transponder met all pertinent specifications including the design goal of ±1-dB gain variation over a +5°C to +35°C temperature range (see Figure 10). No malfunctions were experienced. Table 4 summarizes the performance of the flight model transponder. Figure 11 shows the delivered flight transponder.

TABLE 4. FLIGHT TRANSPONDER PERFORMANCE

| | Specification | Measurement |
|---|---|---|
| Overall Gain | | |
| at 13,195 MHz | 101 dB | channel AA: 101.1 dB<br>channel AB: 101.8 dB<br>channel BB: 102.4 dB<br>channel BA: 101.8 dB |
| at 17,795 MHz | 110 dB | channel AA: 109.4 dB<br>channel AB: 109.6 dB<br>channel BB: 109.4 dB<br>channel BA: 109.9 dB |
| Noise Figure | <10.5 dB | channel 13A: 9.25 dB<br>channel 13B: 8.98 dB<br>channel 18A: 10.09 dB<br>channel 18B: 9.78 dB |
| LO Frequencies | 9.05 and 13.63 GHz | channel 13A: 9.050003 GHz<br>channel 13B: 9.050001 GHz<br>channel 18A: 13.630002 GHz<br>channel 18B: 13.630002 GHz |
| DC Power | <14 W | 10.9 W |
| Weight | <16.8 kg | 14 kg |



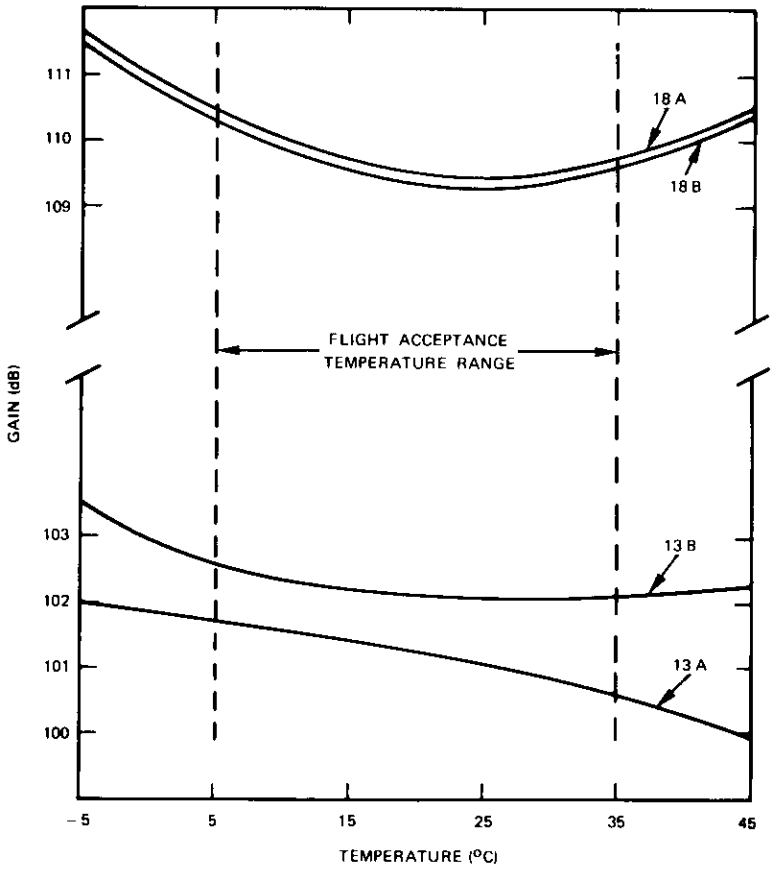



Figure 10. *Flight Transponder, Gain vs Temperature*

## Reliability/quality assurance program

As an integral part of the transponder development cycle, a rigorous parts procurement and quality assurance program was maintained from design inception to delivery of the space-qualified flight transponder. The primary functions of this program were specification of purchased flight materials and parts, in-house inspection, and maintenance of technical quality assurance records and traceability of flight equipment.

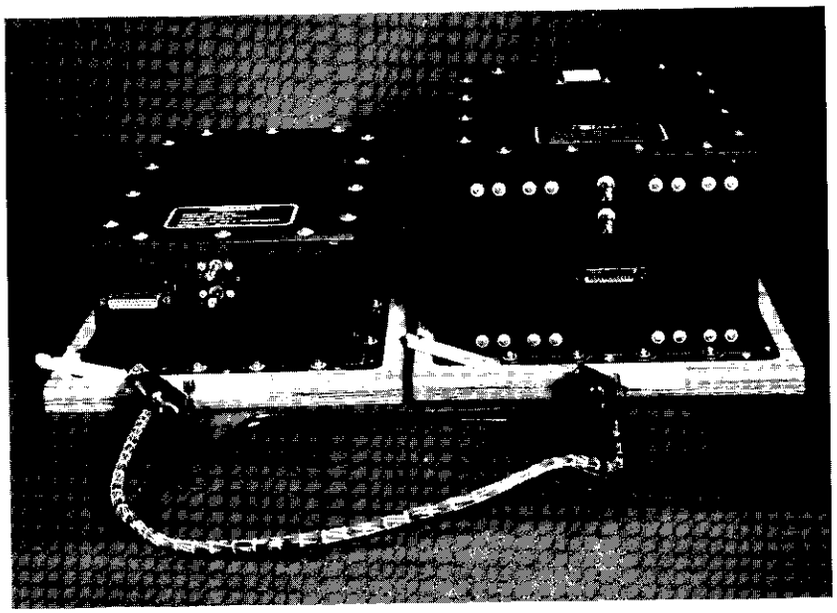


Figure 11. *Flight Transponder*

The majority of the more than 120 high-reliability electronic part types selected were "standard" flight parts found on the published NASA (GSFC) Preferred Parts List. For these parts, existing government specifications were used. New specifications were prepared and approved by NASA for "nonstandard" flight parts and subassemblies.

As an example, the germanium tunnel diode specification included innovations in operating point characterization and screening tests which materially contributed to the reliability of the diodes. Based on studies of tunnel diode characterization techniques, rigid limits were imposed upon peak ($I_p$) and valley ($I_v$) currents, the valley voltage ($V_v$) was specified, and the vendor was required to report the negative resistance and terminal voltage at the operating point as a reference for adjusting the TDA circuits. A high-temperature stabilization bake (100°C for 72 hours) was incorporated to screen out those diodes with excessive built-in mechanical stresses. Step-stress testing [8] showed that these stresses can cause gradual diode degradation, particularly in $I_v$. All units with $\Delta I_v$ greater than $\pm 10$ percent for the bake period were rejected.



When necessary, flight qualification testing of many parts was conducted by their vendors. Since the MIC form of the TDA had no known spaceflight history, COMSAT subjected it to a qualification program. A great deal of testing and process development was required to qualify the alumina substrate MIC TDAs for flight. Serious problems with thick-and thin-film chip resistors and their epoxy attachment techniques were the subject of intensive construction analysis efforts.

It was found that, in two cases, the 330-ohm thick-film chip resistor used to set the diode bias had changed value. (In one case, the resistor demonstrated increased resistance and erratic behavior; in the second case, the resistor demonstrated decreased resistance.) In a third case, one or more of the 50-ohm thin-film chip matching resistors had increased in value.

Investigation of the thick-film resistor failures revealed that the lot of flight resistors was mixed; samples of Pd/Ag and Bi/Ru thick-resistor films included reglazed units. Although all units passed "burn-in," some had probe cracks or bubble cracks in the thin glass protective layer. Both failures occurred in the Pd/Ag-type resistor with propagated glass cracks. A glass chip broken off from one resistor explained its erratic variation and increase in resistance. The second malfunction was more subtle. In this case, the resistance decreased from 325 ohms to 290 ohms and held this lower value through "stripping" from the substrate and some pragmatic immersion tests. Finally, exposure to 60°C high humidity plus epoxy hardener vapors for 16 hours further decreased the resistance to 275 ohms. It is hypothesized that the combination of interface stresses, permeation of moisture at elevated temperature, glass crack propagation, and exposed Pd/Ag thick film can result in the type of resistance changes which were observed.

The 50-ohm thin-film resistor malfunctions were traced to the bonding procedure used. Analysis indicated that the primary cause of failure was the method of melding the silver and clear epoxies during mounting. The resultant epoxy matrix provided a significantly reduced conductive contact area, combined with interface (epoxy/resistor and epoxy/substrate) stresses. The net result was an increase in resistance from 50 ohms to 61 ohms produced by a slight "lifting" at the interface. The interface stresses resulted from marked differences in the participating thermal coefficients of expansion and the moisture absorbing (slight swelling) property of the clear epoxy material.

These investigations resulted in the following corrective actions:

a. changes in process specifications to preclude mixing of clear and silver epoxies;
b. improved epoxy mixing methods, formulas, and dispensing;
c. elimination of the high-humidity paint cure;
d. addition of electrical test points;
e. revision of the thick-film resistor purchase specification and procurement of a new lot of Bi/Ru resistors; and
f. incorporation of all of these changes in the flight transponder.

As part of the transponder assembly, the documentation necessary to maintain control was developed. This documentation included R&QA plans; process specifications; test plans; assembly and inspection procedures; photographs of assemblies (used as working standards); parts, materials, and processes lists; and a released drawing set. Other records included malfunction and discrepancy reports, material flow records, and transponder qualification and acceptance test logs for end item delivery. The R&QA system ensured traceability of parts and materials throughout the assembly phase of the transponder.

## Conclusion

A satellite transponder to be used in the ATS-F Propagation Experiment has been constructed. While the transmission bandwidth requirement dictated by the experiment is limited to 10 MHz, each of the subsystems has been designed so that it can be readily expanded for wideband communications use. In addition, the MIC implementation of the TDAs will permit the incorporation of additional stabilizing network sections to increase bandwidth. The basic design and implementation techniques, however, will remain the same. Similarly, the mixers can be adapted to communications use in the 20- and 30-GHz bands. The mixer bandwidth capabilities can be fully exploited through the use of temperature-stable invar construction. Thus, the propagation experiment serves two purposes. First, all of the experiment requirements are conservatively satisfied. Second, many new design techniques and technologies directly applicable to communications satellite transponders are demonstrated.



## Acknowledgment

*The authors wish to acknowledge the significant contributions of many people who participated in this program. They especially wish to acknowledge W. Getsinger for his guidance in the design of the multiplier and directional filter and his assistance in computer modeling of the microwave circuits; P. Schrantz for structural design; H. Flieger and G. L. Albright for thermal design; P. Varadi, L. Sparrow, and M. Sandor for material analysis and chemical processing; P. Fleming for tunnel diode and other microwave parts specifications and assessment evaluation; and L. Pollack for his overall guidance and encouragement.*

## References

[1] L. Westerlund, J. Levatich, and A. Buige, "ATS-F COMSAT Millimeter Wave Propagation Experiment," *National Telecommunications Conference*, Houston, Texas, December 4–6, 1972.

[2] K. M. Johnson, "X-Band Integrated Circuit Mixer with Reactively Terminated Image, *IEEE Transactions on Electron Devices*, ED-15, No. 7, July 1968.

[3] "The Hot Carrier Diode: Theory, Design and Application," Hewlett Packard Application Note No. 907, May 15, 1967.

[4] A. E. Atia and A. E. Williams, "Narrow-Bandpass Waveguide Filters," *IEEE Transactions on Microwave Theory and Techniques*, MTT-20, No. 4, April 1972, pp. 258–265.

[5] G. L. Matthaie, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching Networks and Coupling Structures*, New York: McGraw-Hill Book Co., Inc., 1964.

[6] C. A. Liechti, "Down-Conversion Using Schottky-Barrier Diodes," *IEEE Transactions on Electron Devices*, ED-17, No. 11, November 1970, pp. 975–983.

[7] Y. S. Lee, "Mode Compensation Applied to Parallel Coupled, Microstrip Directional Filter Design," to be published.

[8] A. G. Revesz, J. Reynolds, and J. Lindmayer, "New Aspects of Failure Mechanisms in Germanium Tunnel Diodes," *Solid State Electronics*, Vol. 14, No. 11, November 1971, pp. 1137–1147.

## Appendix A. Additional mixer design considerations

The diode impedance (approximately 120 ohms resulting from the reactive termination of sum and image frequencies) is matched to the 50-ohm output by using a 2-section stub tuner. The mixer diode is self-biased by using an RF choke, and the DC return is completed by a short circuit for the 18-GHz mixer and by approximately 400 ohms for the 13-GHz mixer. The sum frequency is terminated by the diode package.

To optimize the mixer, the spacings of the signal and LO filters and short circuits with respect to the diode must be determined so that all three ports of the mixer are well matched, and the minimum conversion loss and noise figure are achieved. These conditions can be met when the image, sum, and higher order products are properly terminated. Simple 1st-order analysis can be used to determine the starting values of the spacings $l_1$, $l_2$, $l_3$, and $l_4$ shown in Figure 3 of the paper. These distances are found by choosing integers $n_1$, $n_2$, $n_3$, and $n_4$ such that

$$l_1 = \frac{n_1 \lambda g_s}{2}$$

$$l_2 = \frac{n_2 \lambda g_{LO}}{2}$$

$$l_3 = \frac{(2n_3 + 1)\lambda g_{LO}}{4}$$

$$l_4 = \frac{(2n_4 + 1)\lambda g_s}{4}$$

where $\lambda g_s$ and $\lambda g_{LO}$ are the guide wavelengths at the center frequencies of the signal and LO filters, respectively. The spacings are then perturbed experimentally so that the mixer characteristics are improved.

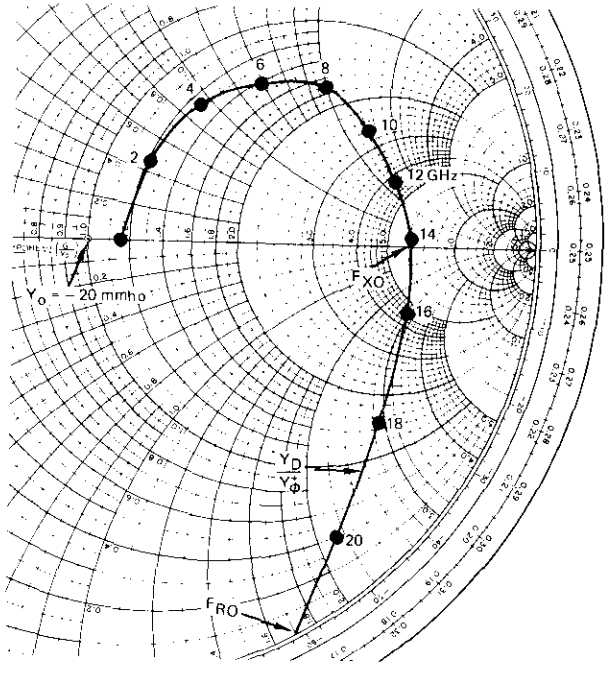
## Appendix B. Tunnel diode amplifier design

### Diode

A germanium tunnel diode with a peak current of about 2.5 mA has been selected for use in the amplifier. Germanium has been chosen because it has been previously qualified for spaceflight use. Since the "micropill" package is 0.05 cm high, it is compatible with a 0.06-cm-thick alumina MIC substrate.

Figure B-1 shows the equivalent circuit of the diode and the variation of its terminal admittance with frequency. A Nyquist analysis [B1] shows that the diode is both open- and short-circuit unstable. A large troublesome negative





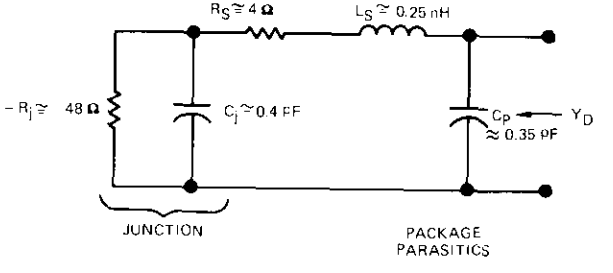




Figure B-1. *Equivalent Circuit and Terminal Admittances for a Typical ATS-F 2.5-mA Tunnel Diode*

conductance exists in the 12- to 16-GHz region as a result of the condition $f_{r_0} < f_{R_0}$.

Near 4 GHz, the diode is modeled by the circuit of Figure B-2. The amplifier design proceeds by synthesizing networks which provide the following characteristics:

a. resonance of the capacitance, $C_T$, at 4.15 GHz;
b. a load impedance (~25 ohms) which is consistent with a gain of approximately 12 dB at 4.15 GHz; and
c. proper loading of the diode at all other frequencies (DC to 20 GHz) to prevent spurious oscillations.

**Gain-bandwidth limitation**

The overall objective of a 90-MHz 1-dB bandwidth for each 3-stage assembly indicates that each TDA module must have a bandwidth of about 170 MHz. If it is assumed that a lumped inductor is used to resonate the circuit of Figure B-2, the reflection gain (Figure B-3) can be determined as follows:

$$\Gamma_{\text{Plane A-A}} = \frac{y_1 - y_2^*}{y_1 + y_2^*} = \frac{(G_s + G_D) + j(\omega C_T - 1/\omega L_r)}{(G_s - G_D) - j(\omega C_T - 1/\omega L_r)} \quad . \tag{B1}$$

An expression for the 1-dB bandwidth can then be derived:

$$\frac{\Delta\omega_{1dB}}{\omega_0} = \frac{G_D}{\omega_0 C_T (G_{v0} - 1)} \tag{B2}$$

where $G_{v0}$ is the voltage gain at resonance. For $C_T = 0.8$ pF and $G_D = 40$ ohms, equation (B2) predicts a bandwidth of 1,600 MHz; clearly the required 170-MHz bandwidth presents no problem. This is important, since circuit designs which guarantee stability greatly degrade bandwidth.

The inductive susceptance of Figure B-3 cannot be achieved by using a lumped element at 4 GHz; instead, some form of distributed network must be used. The design requires that the network which resonates with $C_T$ must also provide out-of-band loading to prevent spurious oscillations.

**Stabilizing network**

Examination of the Nyquist stability criteria confirms that open-circuit stability can be achieved for the diode of Figure B-1 by adding shunt conductance at the diode terminals at frequencies removed from 4.15 GHz. Short-circuit stability cannot be achieved without the further addition of a series resistance; this approach is considered impractical, however. Fortunately, the Nyquist criteria are sufficient but not necessary for stability because the complex behavior of the circulator load impedance, which cannot be modeled and therefore cannot be included in any Nyquist analysis, guarantees stability.



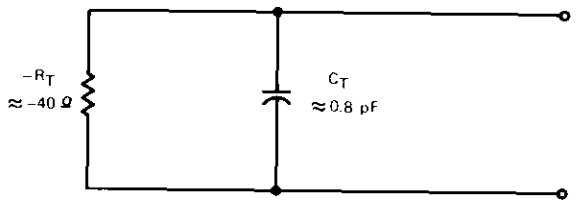



Figure B-2. *Equivalent Circuit of a 2.5-mA Tunnel Diode Near 4 GHz*

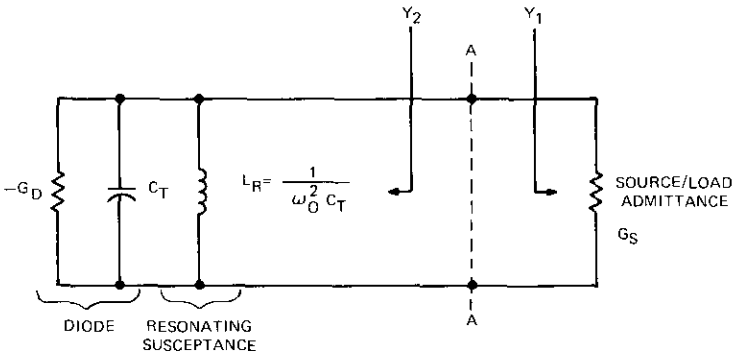



Figure B-3. *Circuit for Calculation of 1-dB Bandwidth Limitation*

Figure B-4 shows the form of the stabilizing network chosen to provide an out-of-band shunt conductance to the diode. Its success depends upon the following factors:

a. realization of an all-band load (DC to 20 GHz);
b. synthesis of a bandstop filter at 4.15 GHz which has no other (spurious) stopbands below 20 GHz; and
c. tailoring of the lossless network, $N_T$, so that $\text{Re}[Y_{\text{STAB}}]$ exceeds the diode's negative conductance at all frequencies removed from 4.15 GHz.

In an MIC format, a frequency-independent load may be obtained by using thin-film resistors (good to about 4 GHz) combined with extremely lossy transmission lines. These are easily synthesized on microstrip lines using lossy film overlays. Bandstop filters having no spurious responses below 20 GHz can be designed only by using "pseudo-lumped" series resonant networks consisting of low- and high-impedance (e.g., 20- and 120-ohm, respectively) transmission lines.

These lines, whose behavior approximates that of capacitors and inductors, differ from a $\lambda/4$ open-circuited transmission line in that they have no resonance at $3f_o$.

The computer analysis used to verify overall stability proceeds by assuming an amplifier load reflection coefficient defined by

$$\Gamma_L = |\Gamma_L| \angle\theta_L$$

(see Figure B-5), where

$$0 \leq |\Gamma_L| \leq 1.0$$

and

$$0 \leq \theta_L \leq 2\pi \quad .$$

Admittance $Y_L$ of Figure B-5 is calculated from

$$Y_L = 0.020 \frac{1 - \Gamma_L}{1 + \Gamma_L}$$

and admittance $Y_i$ is calculated while the frequency is swept from 0 to 20 GHz. Stability is violated when admittance $Y_i$ has a negative real part and a simultaneous zero imaginary part, indicated by a sign change in $\mathrm{Im}(Y)$.

Since the angle $\theta_L$ of $\Gamma_L$ cannot be guaranteed, while its magnitude $|\Gamma_L|$ (circulator return loss) can be measured, a plot of maximum allowable return loss vs frequency can be compared with measured circulator data. Figure B-6 presents measured and computed performance demonstrating unconditional stability.

**Reference**

[B1] H. A. Watson, *Microwave Semiconductor Devices and Their Circuit Applications*, New York: McGraw-Hill Book Co., Inc., 1969.



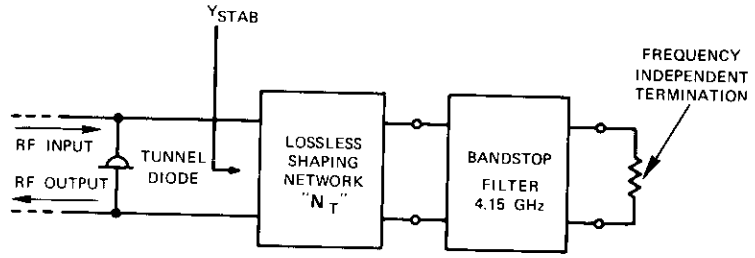



Figure B-4. *Configuration for a Stabilizing Network Which Guarantees Open-Circuit Stability*

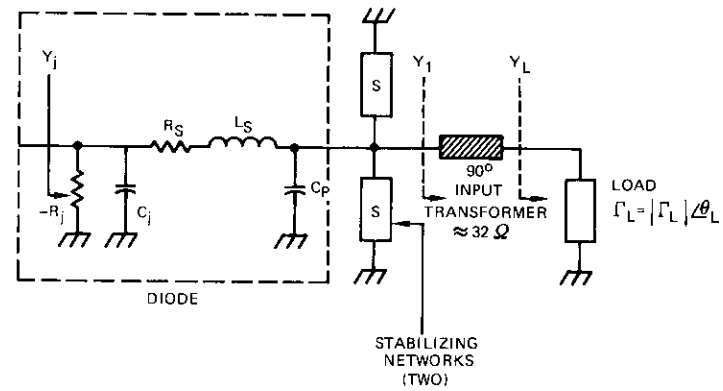



Figure B-5. *Numerical Technique for Testing the Stability of the TDA Module for a Random Load Defined by* $\Gamma_L$

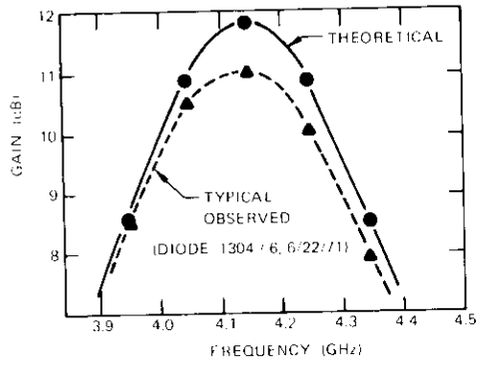



Figure B-6a. *Theoretical Gain-Frequency Curve of ATS-F TDA*

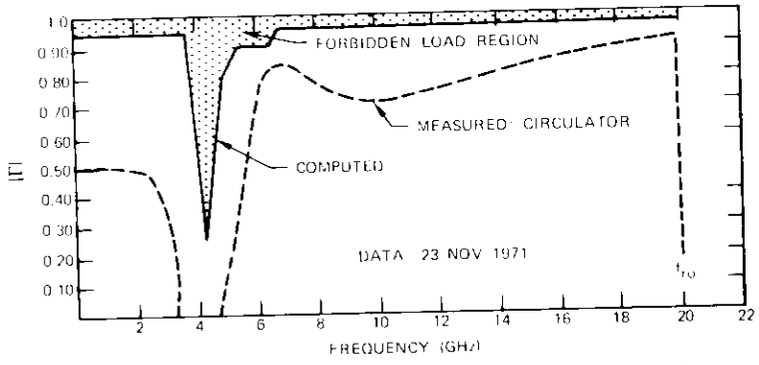



Figure B-6b. *Maximum Allowable Stable Load Reflection Coefficient vs Frequency Compared with Actual Circulator Measurement*



## *Appendix C. Oscillator phase noise measurement*

Oscillator phase noise is measured in two steps. Two identical oscillator chains, including crystal sources and multipliers to the 2.25-GHz range, are fed to a phase detector. The output voltage, proportional to the phase difference between the two sources, is then recorded. A digital computer is used to perform a Fourier analysis of the recorded waveform. The spectrum of phase noise "baseband" thus computed is presented in Figure C-1.

The RF spectrum of the noise modulated oscillator is obtained by slightly shifting the frequency of one of the oscillators and recording the audio frequency mixing component. The RF spectrum obtained by Fourier analysis is presented in Figure C-2.

Examination of the data reveals that the significant phase noise components are at frequencies below 1 Hz. The RF bandwidth caused by phase noise is approximately 10 Hz. The deviation caused by noise may be determined from the Carson's Rule bandwidth:

$$\text{deviation } \Delta f_{pK} = \left(\frac{\text{RF bandwidth}}{2}\right) - \text{modulating frequency}$$
$$= 5 - 1$$
$$= 4 \text{ Hz} \quad .$$

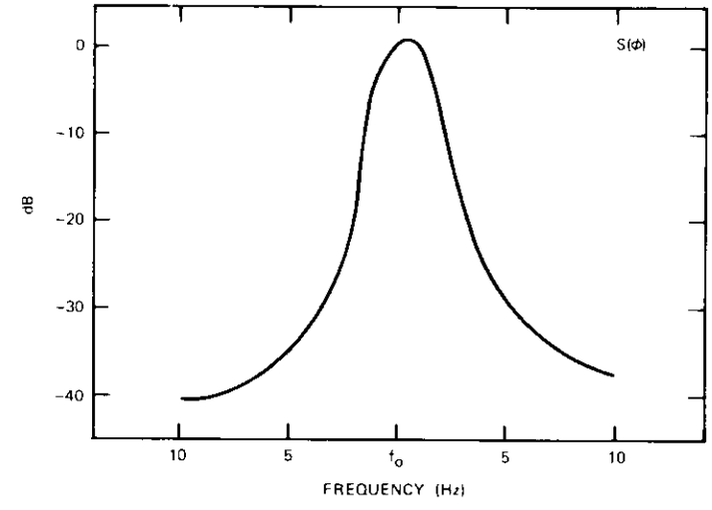



Figure C-1. *Phase Noise "Baseband" Obtained by Fourier Analysis*

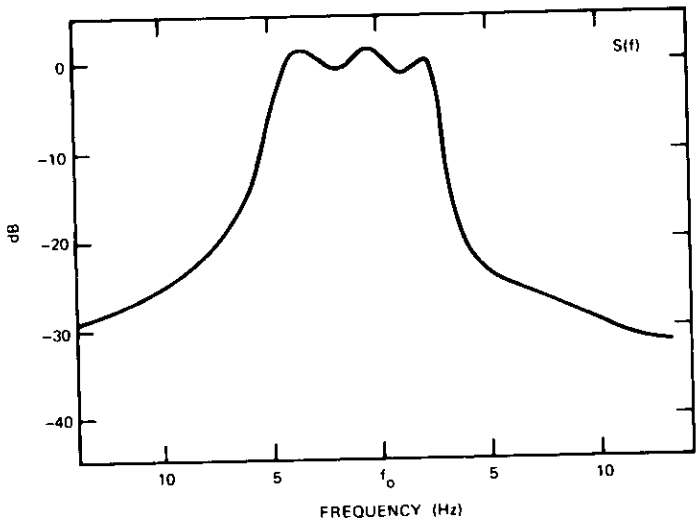



Figure C-2. *RF Spectrum Obtained by Fourier Analysis*



*Arnold L. Berman received the B.S.E.E. from M.I.T. in 1957. Prior to joining* COMSAT *in 1964, he performed communications system analysis for meteorological and communications satellite systems at the Astro-Electronics Division of RCA, Princeton, New Jersey. At* COMSAT *he initially participated in the Interim C.C.I.R. Conference (Monaco, 1965). He later participated in the development of the* INTELSAT II, III, *and* IV *communications systems and the* COMSAT *ATS-F Propagation Experiment transponder. He is presently Director of the RF Transmission Laboratory of* COMSAT *Laboratories.*

*A. E. Atia is a member of the Technical Staff, Transponder Department, RF Transmission Laboratory, Technology Division,* COMSAT.

*R. W. Bounds is a member of the Technical Staff, Electric Power Department, Spacecraft Laboratory, Technology Division,* COMSAT.

*R. S. Cooperman is Manager, Satellite Control Electronics Department, Spacecraft Laboratory, Technology Division,* COMSAT.

*Y. S. Lee is a member of the Technical Staff, Microwave Circuits Department, RF Transmission Laboratory, Technology Division,* COMSAT.

*C. E. Mahle is Manager, Transponder Department, RF Transmission Laboratory, Technology Division,* COMSAT.

*R. E. Stegens is a member of the Technical Staff, Microwave Circuits Department, RF Transmission Laboratory, Technology Division,* COMSAT.

*R. Strauss is Manager, Reliability and Quality Assurance Department, Applied Sciences Division,* COMSAT.

*J. W. Talcott is a member of the Technical Staff, Reliability and Quality Assurance Department, Applied Sciences Division,* COMSAT.

*M. R. Wachs is a member of the Technical Staff, Transponder Department, RF Transmission Laboratory, Technology Division,* COMSAT.

*A. E. Williams is a member of the Technical Staff, Transponder Department, RF Transmission Laboratory, Technology Division,* COMSAT.

Index: radio transmission, antennas, polarization (waves), depolarization, mathematical models.

# The orthogonalization of polarized fields in dual-polarized radio transmission systems

R. W. KREUTEL

## Abstract

Coupling between a pair of nominally orthogonal radio waves in a dual-polarized transmission system results in interference which limits bandwidth utilization efficiency. In this paper, some of the more common types of polarization coupling are examined. Polarization coupling is represented by a transmission matrix, which is then reduced to its equivalent diagonal form by an orthogonal transformation. This transformation is identified with a set of simple lossless networks. It is shown that the basic form of these networks is appropriate to common types of polarization coupling. Finally, some fundamental limitations imposed on transmission efficiency by field coupling are cited.

## Introduction

The channel capacity of a bandwidth-limited radio transmission system can be theoretically doubled by transmitting and receiving information on each of two orthogonal wave polarizations. If the two waves are not perfectly orthogonal, however, interference results. Therefore, the utility of the dual-polarized transmission is contingent on achieving a satisfactory level of orthogonality. In principle, perfect polarization orthogonality can be realized; however, in practice, because of antenna design limitations,

antenna misalignment, transmission media anomalies, and other difficulties, perfect orthogonality can only be approximated.

In this paper, the coupling between the waves of a dual-polarized transmission system is described in terms of a 2 × 2 transmission matrix. It is demonstrated that matrix operations which orthogonalize the transmission matrix can be implemented by adding simple hybrid networks at the sending and receiving ends of the transmission system.

## Mathematical model of the transmission system

For the present analysis, the dual-polarized transmission system is defined in very general terms. At the sending end of the system, two orthogonally polarized waves, $x_1$ and $x_2$, are generated. A corresponding set of waves, $y_1$ and $y_2$, arrives at the receiving end of the system. The relationship between the waves at the sending and receiving ends is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{1}$$

The square matrix in equation (1) is the transmission $[T]$ matrix of the system. Its off-diagonal terms represent coupling between the two wave polarizations and include all sources of coupling (i.e., antennas and medium).

It is evident that, to orthogonalize the process described by equation (1), the $T$ matrix must be diagonalized so that $y_1$ is a function of $x_1$ alone and $y_2$ is a function of $x_2$ alone. This diagonalization can be effected by performing operations involving premultiplication and/or postmultiplication [1]. In the present work, $T$ has been reduced to its diagonal form by means of an orthogonal transformation (or more generally, for a complex $T$ coefficient, a unitary transformation) because the networks required to effect this type of transformation are necessarily lossless.

In the following sections the orthogonalizing operation is performed for two classes of special importance: a symmetric $T$ matrix and a Hermitian $T$ matrix. In each case, the $T$ matrix is identified with the class of polarization coupling it represents. A linear dual-polarized system is used as an example, but the analysis is equally appropriate to any dual-polarized system. In each of the cases examined, it is assumed that $T_{11} = T_{22}$. Since the transmission gains on the two nominally orthogonal channels are ordinarily identical, this should be a reasonable assumption.



## The symmetric T matrix

The symmetric $T$ matrix corresponds to the case in which $T_{21} = T_{12}$. Hence, equation (1) becomes

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{12} & T_{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{2}$$

The $T$ matrix in equation (2) is diagonalized by the orthogonal transformation

$$Q^t T Q = T_D \tag{3}$$

where the superscript $t$ denotes the transpose and $T_D$ is the diagonalized form of $T$. The diagonal elements of $T_D$ are simply the eigenvalues of $T$, which are given by the roots of the characteristic equation

$$|T - \lambda I| = 0 \tag{4}$$

where $I$ is the unit matrix.

The $Q$ matrix is composed of columns which are each eigenvectors of $T$. These eigenvectors are the solutions to the following equations:

$$Te_1 = \lambda_1 e_1 \tag{5a}$$

$$Te_2 = \lambda_2 e_1 \tag{5b}$$

where $e_1$ and $e_2$ are the eigenvectors corresponding to the eigenvalues $\lambda_1$ and $\lambda_2$. It follows that equation (3) must take the general form

$$\begin{bmatrix} e_1^t \\ e_2^t \end{bmatrix} [T] [e_1 \; e_2] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{6}$$

Substitution of equation (6) into equation (2) yields the equations of the orthogonalized transmission system:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{12} & T_{11} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} T_{11} - T_{12} & 0 \\ 0 & T_{11} + T_{12} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{7}$$

The operators in equation (7) can be identified with lossless sum and difference hybrid networks inserted at the receiving and sending ends, respectively. Figure 1 is a schematic representation of the system described by equation (7).

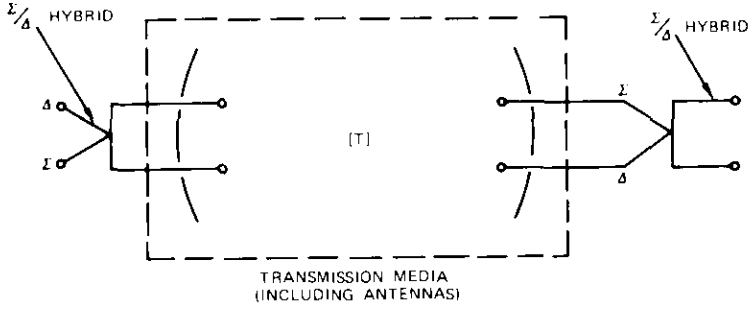



Figure 1. *Orthogonalized Dual-Polarized Transmission System (symmetric T)*

The various types of depolarization which can be represented by a symmetric $T$ matrix are indicated in Figure 2. All three of these depolarized waves can be compensated by the system described by equation (7) and shown in Figure 1. In each case the fields are perfectly linear and orthogonal at the sending end of the system and arrive at the receiving end as shown. As indicated, a reasonably broad range of commonly encountered types of depolarization is included. A typical example is depolarization resulting from antenna misalignment.

### The Hermitian T matrix

The Hermitian $T$ matrix is defined by $T_{12} = T_{21}^*$ and real principal diagonal terms. Thus, a transmission system with Hermitian-type polarization is characterized by the following equations:



$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12}^* \\ T_{12} & T_{11} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{8}$$

The system described by equation (8) can be orthogonalized by performing a unitary transformation on $T$:

$$Q^{t*}\, T\, Q = T_D \tag{9}$$

where, as in the previous section, the operation is completely defined in terms of the eigenvalues and eigenvectors of $T$. Thus, equation (9) can be rewritten as

$$\begin{bmatrix} e_1^{t*} \\ e_1^{t*} \end{bmatrix} [T] [e_1\ e_2] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{10}$$

where $\lambda_1$ and $\lambda_2$ are the eigenvalues of $T$, and $e_1$ and $e_2$ are the corresponding eigenvectors. Finally, evaluation of the terms in equation (10) results in the orthogonalized form of equation (8):

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{e^{j\psi}}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{e^{j\psi}}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12}^* \\ T_{12} & T_{11} \end{bmatrix} \begin{bmatrix} \frac{e^{-j\psi}}{\sqrt{2}} & \frac{e^{-j\psi}}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} T_{11} - |T_{12}| & 0 \\ 0 & T_{11} + |T_{12}| \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{11}$$

where $\psi$ is the argument of $T_{12}$.

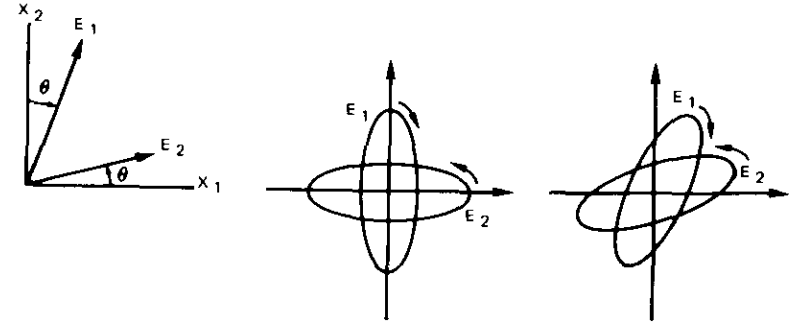



Figure 2. *Types of Depolarization Characterized by Symmetric Transmission Matrix*

The premultiplying and postmultiplying operators in equation (11) can be identified with a network that consists of simple lossless sum and difference hybrids with a phase shift of $\psi$ radians inserted in one of the output arms, as depicted in Figure 3. Also shown in Figure 3 is an additional phase shifter of $\beta$ radians. Although this phase shifter is not essential to the operation in equation (11), its significance will become clear in the following discussion.

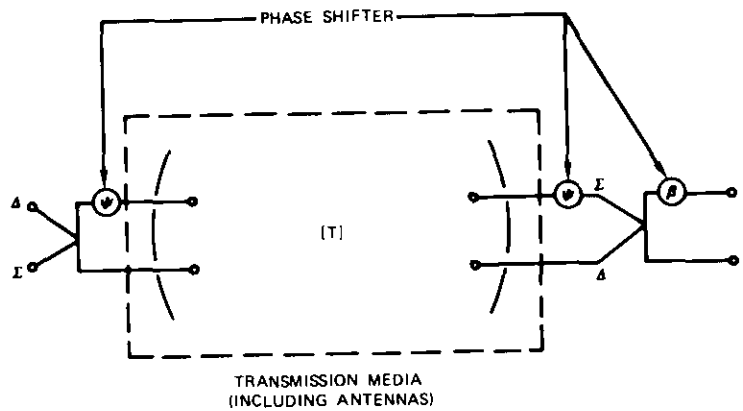



Figure 3. *Generalized Form of Orthogonalized Dual-Polarized Transmission Network*

Figure 2 has summarized the types of depolarization which are characterized by a symmetric transmission matrix. It should be noted that the symmetric matrix generally encompasses the case of two waves which originate as two orthogonal linearly polarized fields, but which are depolarized to yield two nonorthogonal elliptically polarized fields of the opposite polarization sense. It is desirable to broaden the class of depolarization which can be compensated by including the counterpart of the aforementioned case, namely, the case in which the two orthogonal linearly polarized fields are depolarized to yield nonorthogonal elliptically polarized fields of the same polarization sense. The representation of equation (11) provides compensation for only that special case in which the major axes of elliptically polarized fields of the same sense are perpendicular.

The $T$ matrix for the coupling between a pair of nonorthogonal elliptically polarized waves of the same sense has the general form



$$[T] = \begin{bmatrix} A\, e^{-j\alpha} & B\, e^{-j\gamma} \\ B\, e^{+j\gamma} & A\, e^{+j\alpha} \end{bmatrix} . \tag{12}$$

If the first column of $[T]$ is multiplied by $e^{j2\alpha}$ and $e^{j\alpha}$ is factored out,

$$[T] = e^{j\alpha} \begin{bmatrix} A & B\, e^{-j(\gamma+\alpha)} \\ \mathrm{B}\, e^{j(\gamma+\alpha)} & A \end{bmatrix} . \tag{13}$$

Equation (13) is in the desired Hermitian form and can be diagonalized by using the operation of equation (11). The significance of the $\beta$ radian phase shifter in Figure 3 is now clear. The value of $\beta$ is set equal to $2\alpha$ radians to "force" the coupling matrix for elliptically polarized waves of the same sense into Hermitian form so that orthogonalization according to equation (11) can be effected.

## General orthogonalizing network

The circuit shown in Figure 3 represents the generalized form of the orthogonalizing network. Clearly the network in Figure 1 is a special case of the network in Figure 3, for which $\beta = \psi = 0$.

For convenience, the types of depolarization which can be compensated, together with the associated transmission matrix and circuit parameters, are tabulated in Figure 4, which includes a broad range of depolarization.*

It is possible to achieve complete orthogonalization of the transmission system if two requirements are met. First, the fields must be made symmetric with respect to the reference axes. This can be accomplished by rotating the axes. Second, the depolarized fields must have the same axial ratio. If they do not, incomplete orthogonalization results in a residual level of interference. For example, if the depolarized waves of Figure 4b or 4c have axial ratios of 29 and 30 dB, respectively, the system can be partially orthogonalized to yield a wanted-to-unwanted signal ratio of approximately 55 dB. Of course, if the axial ratios are equal, the wanted-to-unwanted signal ratio will be infinite.

* The origin of the depolarization in a practical system is a complex subject beyond the scope of this paper. Cases listed in Figure 4 can be assumed to be representative of typical practical systems.

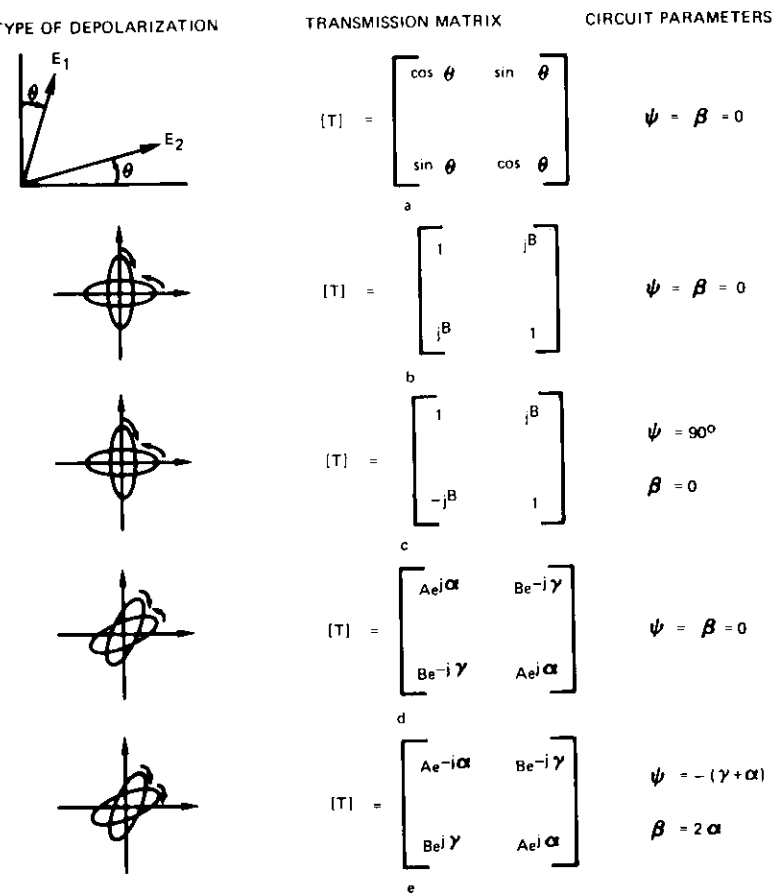



Figure 4. *Types of Depolarization Which Can Be Orthogonalized with the Network of Figure 3*

### Power conservation

If the generalized matrix formulation of the dual-polarized transmission system [equation (1)] is written so that antenna gain and path loss are normalized, then the transfer efficiency* of the network can be evaluated. For example, if the transmission system conserves power, i.e., if the power at the sending end equals the power at the receiving end, then the normalized transmission matrix must be unitary. It can easily be shown that this will be true if and only if the depolarized fields are orthogonal. In the more general case, the transmission matrix will not be unitary, and a more complete representation of the network must be defined to determine the efficiency of transmission and the distribution of power in the system.

A dual-polarized radio transmission system is actually a 4-port network; hence, its complete characterization requires a $4 \times 4$ matrix representation. A complete representation can be achieved by using scattering matrix format:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \left[\begin{array}{cc:cc} & & & \\ & \Gamma & & T \\ & & & \\ \hdashline & & & \\ & T' & & \Gamma' \\ & & & \end{array}\right] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \tag{14}$$

where $a_n$ and $b_n$ represent the waves incident on and emerging from the $n$th port, respectively. The partitioned $T$ matrix is identical (except for normalization) to the $T$ matrix in equation (1). The $\Gamma$ matrices represent reflection and/or cross-coupling at the sending or receiving ports.

It has been postulated that antenna gain and path loss have been factored out of the transmission equations, and, further, that the transmission medium is lossless. Consequently, the scattering matrix in equation (14) must be unitary and this property can be used to evaluate unknown coefficients. Then the transmission efficiency of the system can be determined. This operation will be performed for the example outlined in the next section.

### Example

To illustrate the foregoing concepts, the case of Figure 4a will be examined in some detail. This is a case of considerable practical importance. It is clear that transmission and cross-coupling in this system are proportional to $\cos\theta$ and $\sin\theta$, respectively. Consequently, the procedure outlined in the previous section is used and the scattering matrix of the system is constructed as follows:

$$[S] = \left[\begin{array}{cc:cc} b & c & a\cos\theta & a\sin\theta \\ c & b & a\sin\theta & a\cos\theta \\ \hdashline a\cos\theta & a\sin\theta & b & c \\ a\sin\theta & a\cos\theta & c & b \end{array}\right] . \tag{15}$$

* It is assumed that the transmission medium is lossless.

The coefficients $a$, $b$, and $c$ are determined by invoking the unitary property in equation (15) and solving the resulting set of equations. Hence,

$$|S| = \frac{1}{\sqrt{1+|\sin 2\theta|}} \begin{bmatrix} j\sqrt{\frac{\sin 2\theta}{2}} & -j\sqrt{\frac{\sin 2\theta}{2}} & \cos\theta & \sin\theta \\ -j\sqrt{\frac{\sin 2\theta}{2}} & j\sqrt{\frac{\sin 2\theta}{2}} & \sin\theta & \cos\theta \\ \cos\theta & \sin\theta & j\sqrt{\frac{\sin 2\theta}{2}} & -j\sqrt{\frac{\sin 2\theta}{2}} \\ \sin\theta & \cos\theta & -j\sqrt{\frac{\sin 2\theta}{2}} & \sqrt{\frac{\sin 2\theta}{2}} \end{bmatrix} . \quad (16)$$

The network appropriate for orthogonalizing the system described by equation (16) is shown in Figure 1 (or Figure 3 with $\psi = \beta = 0$). Equation (7) is the resulting system equation. If it is assumed that $\theta$ is positive, substitution from equation (16) into equation (7) yields

$$y_1 = \frac{\cos\theta - \sin\theta}{\sqrt{1+\sin 2\theta}} x_1 = x_1 \tan(45° - \theta) \quad (17a)$$

$$y_2 = \frac{\cos\theta + \sin\theta}{\sqrt{1+\sin 2\theta}} x_2 = x_2 \quad . \quad (17b)$$

The differential channel gain given by $\tan(45° - \theta)$ is identical to a result reported elsewhere [2], [3]. However, in this case, it is clearly not a consequence of dissipation. It is interesting to note that all of the loss is incurred in one channel. If $\theta$ is negative, equations (17a) and (17b) simply interchange.

It can easily be seen that the implicit differential gain in equation (17) is a consequence of reflection at the $x_1$ port. The magnitude of the reflected field at that port is

$$|\Gamma| x_1 = \sqrt{\frac{2\sin 2\theta}{1+\sin 2\theta}} x_1 \quad . \quad (18)$$

Finally, to prove that power is conserved, equations (17) and (18) are used to obtain



$$\begin{aligned} P_{OUT} &= \frac{1}{2}|y_1|^2 + \frac{1}{2}|y_2|^2 + \frac{1}{2}|\Gamma x_1|^2 \\ &= \frac{1}{2}\left[\left(\frac{1-\sin 2\theta}{1+\sin 2\theta}\right)|x_1|^2 + |x_2|^2 + \left(\frac{2\sin 2\theta}{1+\sin 2\theta}\right)|x_1|^2\right] \\ &= \frac{1}{2}|x_1|^2 + \frac{1}{2}|x_2|^2 \end{aligned} \quad (19)$$

which is, of course, equal to the power input to the system.

## Conclusion

A broad class of polarization couplings in a dual-polarized transmission system can be represented by either a symmetric or Hermitian transmission matrix. It has been shown that a system so characterized can be perfectly orthogonalized by an orthogonal equivalence transformation and, further, that such a transformation can be effected by adding simple lossless networks at the sending and receiving ends of the transmission system. Of course, the orthogonalizing process completely eliminates polarization coupling and its degrading effects on transmission quality.

The orthogonalization process is summarized as follows. The transmission system is represented by a transmission matrix, and the eigenvectors of that matrix are formed. These eigenvectors correspond to a pair of characteristic polarizations which (for symmetric and Hermitian systems) are orthogonal. The compensating networks simply adjust the transmission polarizations so that they correspond precisely to the characteristic polarizations of the system. Since the characteristic polarizations are orthogonal, it follows that the system is orthogonalized.

Of particular interest is the fact that, except for the type of polarization shown in Figure 4e, the orthogonalizing networks are completely independent of the implicit polarization coupling in the transmission matrix if the class of transmission matrix (i.e., symmetric or Hermitian) is known. Further, these networks, comprising sum and difference hybrids, are lossless. For the depolarization shown in Figure 4e, the orthogonalizing networks remain independent of the magnitude of the polarization coupling, but each requires a phase shifter, $\psi$, that is dependent on the transmission matrix. If $\psi$ is time variant, dynamic control of the phase shifters is required.

In the general case, the depolarization in the system results in nonorthogonal fields. It has been shown that the orthogonalization of such a system necessitates loss. While this is not surprising, it should be noted that

this loss need not be dissipative. As in the example cited in this paper, the orthogonalizing loss may be reflective. This may be advantageous for many applications, e.g., a low-noise receiving system. In fact, because the orthogonalizing networks are lossless, any orthogonalizing loss incurred as a result of the processes outlined in this paper will be nondissipative. Further, based on transmission efficiency calculations which assume equal power input to each orthogonal transmitting port, it is postulated that these processes result in minimum orthogonalization loss.

## References

[1] F. B. Hildebrand, *Methods of Applied Mathematics*, Englewood Cliffs, N.J.: Prentice-Hall, 1961, Chapter 1.

[2] T. S. Chu, "Restoring the Orthogonality of Two Polarizations in Radio Communications System I," *Bell System Technical Journal*, Vol. 50, No. 7, November 1971, pp. 3063–3069.

[3] T. S. Chu, "Restoring the Orthogonality of Two Polarizations in Radio Communications System II," *Bell System Technical Journal*, Vol. 52, No. 3, March 1973, pp. 319–329.

*Randall W. Kreutel, Jr., received B.S.E.E. and M.S.E.E. degrees from Northeastern University. He joined* COMSAT *in 1966 as a Member of the Technical Staff. As Manager of the Antenna Department he is responsible for research and development of antennas and antenna-related equipment. He was formerly a Research Engineer with the Antenna and Microwave Lab, Sylvania Electronic Systems.*

*He is a senior member of IEEE, a member of Eta Kappa Nu, and of Commission I of the International Scientific Radio Union.*

**Index: varactor diodes, frequency converters, earth terminals, telecommunication.**

# A 6-GHz broadband varactor up-converter

R. L. SICOTTE

## Abstract

This paper provides design and performance data for an upper-sideband up-converter which is useful in satellite communications earth terminals. This up-converter not only has the high output level (>50 mW) typical of varactor up-converters, but also has a wide bandwidth (500 MHz) at all three ports, a characteristic previously associated only with low-level mixer-type up-converters. The design has been achieved by using a heavily overdriven varactor in a low-inductance mount and matching into a contiguous band diplexer.

## Introduction

High-level up-converters with broadband signal, pump, and upper-sideband (USB) ports simplify the design of earth terminals for satellite communications. The broad instantaneous bandwidth of the pump-USB ports provides a system flexibility which makes it possible to up-convert wideband signals to a number of up-link channels by selecting a suitable local oscillator (pump) frequency. The high-level output reduces the gain required of the transmit power amplifier.

Current earth terminals use broadband up-converters which employ varistor devices such as point-contact or Schottky-Barrier diodes as the nonlinear element, usually in a 4-diode bridge circuit. Usable output

power levels are typically limited to 1 mW; hence, a 2-stage traveling wave tube amplifier is necessary to attain the required transmit power level. The use of high-power up-converters is desirable to eliminate the complexity of a high-gain amplifier carrying information. The required high-power pump can be provided by solid-state amplifiers operated at saturation for maximum efficiency. Although varactor up-converters* have been employed for high-level operation [1], their use results in some sacrifice in bandwidth [2] because the varactor is a nonlinear capacitor which, when pumped, has an average capacitance which must be tuned. Typical bandwidths have been 1 percent.

This paper deals with a new varactor up-converter design for wideband, high-level operation. It introduces a new varactor circuit design technique based on broadband matching of a pumped, mounted varactor diode to a diplexer filter network. This technique is applied to the design and realization of an up-converter having the following characteristics:

| | |
|---|---|
| signal center frequency: | 855 MHz |
| signal bandwidth: | 100 MHz at −1 dB |
| pump frequency: | 5.07–5.57 GHz |
| USB frequency: | 5.925–6.425 GHz |
| | minimum at ±1 dB |
| power output: | 30 mW |
| signal power: | 10 mW |
| pump power: | 400 mW minimum. |

## *Design considerations*

### Stability

The primary design criterion for varactor circuits is stability. Two major types of instability, parametric and "jump," may occur. Parametric instability is characterized by RF oscillation, usually occurring at frequencies below one of the driving (i.e., signal or pump) frequencies. These oscillations modulate the input and output signals and appear as multiple sidebands on a spectrum analyzer. They are initiated by tuned idlers, which cause a negative resistance to appear at the difference frequency between an idler and the driving frequency, as in a parametric amplifier.

* The conventional distinction between varistor and varactor up-converters is that varistor circuits have broader bandwidths and operate at a low level, while varactor circuits have narrower bandwidths and operate at a high level.



When the negative resistance is equal to or greater than the terminating resistance at that frequency, oscillations occur. The circuit conditions for oscillation have been calculated by Parker and Grayzel [3] for a frequency doubler. Parametric instabilities are avoided by eliminating undesirable idlers and resistively terminating the diode at frequencies where oscillations may occur. This is accomplished practically by designing very simple, compact circuitry around the diode to eliminate idler formation and by resistively loading any idlers that occur.

Jump instability is characterized by a low-frequency relaxation oscillation in resistively biased varactor circuits or an abrupt turn-off or turn-on of power in fixed bias circuits. According to the literature [4], this phenomenon is peculiar to circuits containing nonlinear reactances. It becomes less severe and eventually disappears when the loaded $Q$ of the nonlinear circuit is lowered sufficiently.

### Bandwidth

There are four basic design requirements for broadband up-converters:

a. all tuning elements must be located close to the diode,
b. the diode impedance must have a low loaded $Q$,
c. a low-inductance mount must be chosen, and
d. terminations must be controlled at harmonic frequencies.

These requirements also contribute to the stability of the design.

Figure 1 is a simplified block diagram of the up-converter. The circuit, which consists of a pump-USB diplexer and a signal arm comprising a low-pass filter, quarter-wave transformer, and signal tuning element, provides for diode tuning at signal, pump, and USB bands.

Conning [5] has analyzed the operation of high-level varactor up-converters for an arbitrary drive level, $M$, and capacitance law, $\gamma$. The normalized impedance and power, plotted in Figures 2, 3, and 4 vs drive level, $M$, for abrupt junction ($\gamma = 1/2$) and punch-through ($\gamma = 0$) varactors, are, respectively,

$$Z(\omega_k) = (K_R - jK_X)\frac{S_{\max}}{\nu\omega_k} \tag{1}$$

$$P(\omega_k) = K_p \nu\omega_k \frac{V_B^2}{S_{\max}} \tag{2}$$

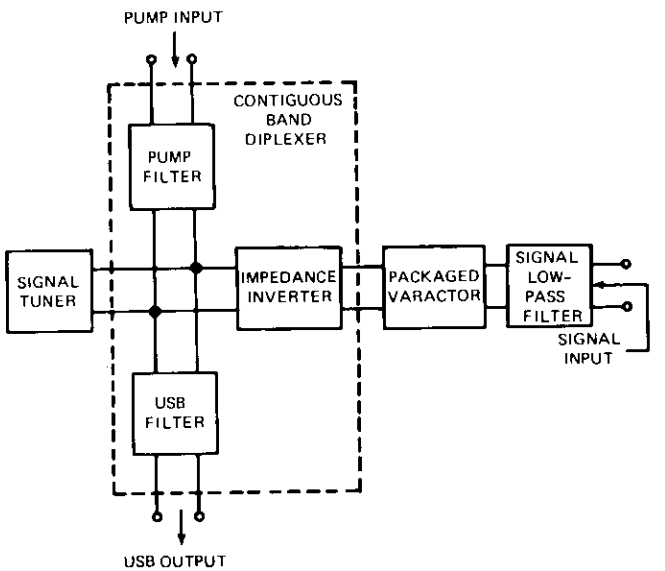



Figure 1. *Block Diagram of Varactor Up-converter*

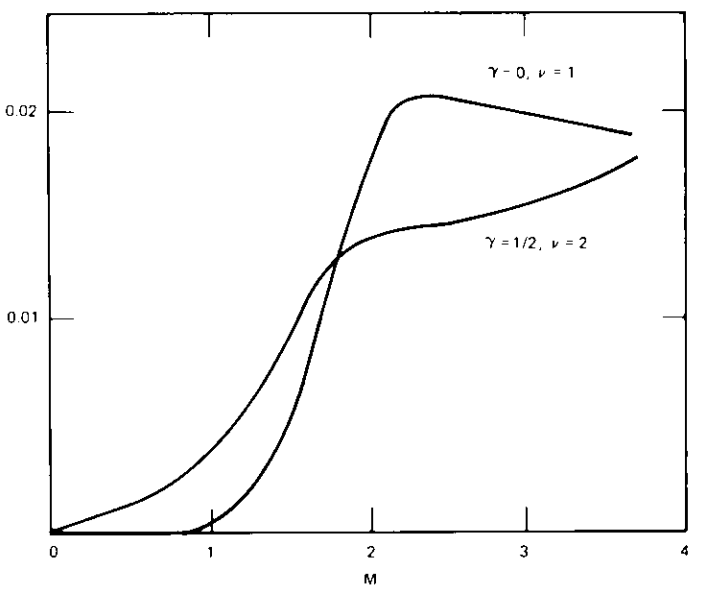



Figure 2. *Power Levels for Lossless Varactor Up-converter*



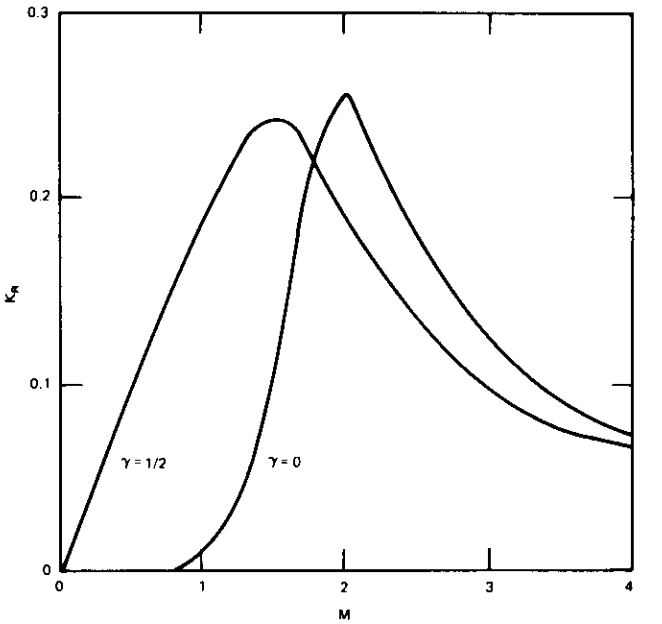



Figure 3. *Real-Part Impedance Levels for Lossless Varactor Up-converter*

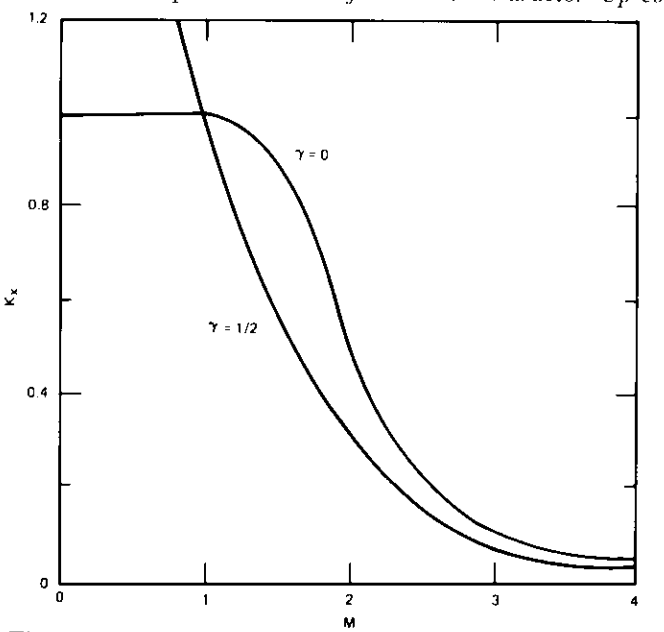



Figure 4. *Imaginary-Part Impedance Levels for Lossless Varactor Up-converter*

where $\nu = \dfrac{1}{1-\gamma}$

$\gamma$ = capacitance law of the diode, defined in Reference 5

$S_{max} = \dfrac{1}{C_{min}}$

$C_{min}$ = minimum capacitance swept out by the RF voltage across the varactor

$V_B$ = breakdown voltage

$K_R$ = coefficient of the real part of the impedance

$K_X$ = coefficient of the imaginary part of the impedance

$K_p$ = power coefficient

$M$ = drive level parameter, defined in Reference 5.

The average elastance,

$$S_o = K_X S_{max} \tag{3}$$

is a useful design parameter. The loaded $Q$ of the pumped varactor is

$$Q_L = \frac{K_X}{K_R} \quad . \tag{4}$$

Figure 5 is a plot of $Q_L$ vs $M$. To achieve a low loaded $Q$, it is desirable to operate at drive levels, $M > 2$. It should be noted that drive level is often associated with power level. This is indeed the case when $M$ is defined as [5]

$$M = 1 - \frac{q_{min}}{q_{max}} \tag{5}$$

where $q_{max}$ is the charge at the breakdown voltage and $q_{min}$ the peak forward driven charge, which is negative for the overdriven varactor.

A more general definition of $M$ is obtained when $q_{max}$ is defined as the peak reverse driven charge:

$$q_{max} = C_{min} V_{pk} \tag{6}$$

where $|V_{pk}| \leq |V_B|$ and $C_{min}$ are the peak negative voltage and associated minimum capacitance, respectively. The composite charge waveform comprises the charge flowing through the diode at the three frequencies



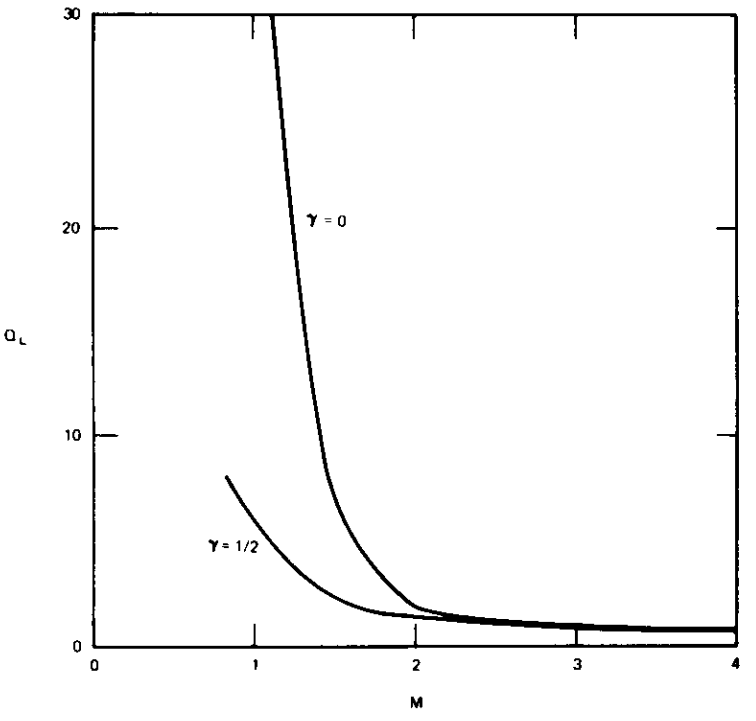



Figure 5. *Loaded Q vs Drive Level Parameter*

of interest. The peak-to-peak charge excursion is dependent upon the power levels at these three frequencies. The ratio of the peak-to-peak charge determines the drive level, however. If $q_{max}$ is fixed as the charge at the breakdown voltage, then the ratio $q_{min}/q_{max}$ becomes power dependent. The definition in equation (6) assigns an arbitrary value to $q_{max}$ so that the ratio $q_{min}/q_{max}$ can change independent of power level. Hence, this definition disassociates drive level from power level and allows the selection of drive level on the basis of other considerations such as loaded $Q$.

There is a maximum practical drive level, $M$, for efficient operation. As the value of $M$ increases, the diode's forward drive level also increases and the diode tends to become more inefficient because of losses resulting from carrier recombination under high forward injection levels. The amount of forward drive consistent with efficient operation depends upon the type of varactor used. For example, the exceptional forward charge storage capabilities of silicon punch-through varactors permit very high

forward drive levels. On the other hand, gallium arsenide varactors have poor forward charge storage capability at 6 GHz and would have poor efficiency at high forward drive levels. Hence, a silicon punch-through varactor having $C_{min} = 1\ pF$ and a drive level, $M = 2.5$, have been chosen for this application. The drive level choice is arbitrary; it yields a value of $Q_L \approx 1$, and an average capacitance, $C_o = 5\ pF$.

The inductance of the diode package is important because it determines the reactance slope of the tuned circuit formed with the diode average capacitance. The reactance slope must match the negative reactance slope of the diplexer.

Figure 6 is an assembly drawing of the up-converter. The unit is fabricated in 7-mm coaxial transmission line with the exception of the 3.5-mm section enclosing the diode.

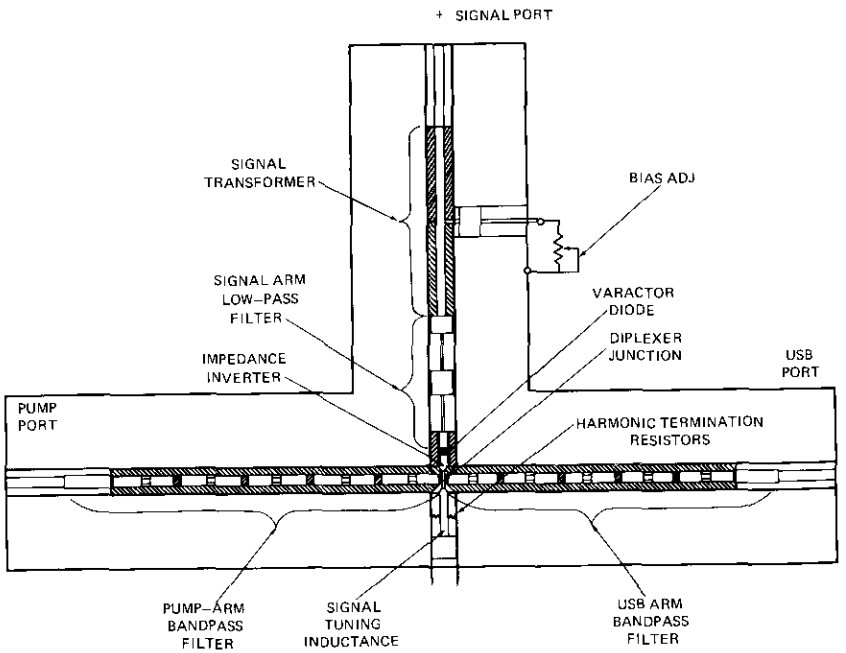



Figure 6. *Upper-Sideband Up-converter Circuit*

The equivalent circuit of a packaged varactor mounted in series with a coaxial line [6] is shown in Figure 7a. Parameters $C_A$, $L_d$, and $C_B$ are internal package parameters and depend solely on package design. The



inductance, $L_a$, is external to the package and depends on the mounting. The diode junction impedance is a function of diode parameters $C_{min}$ and $\gamma$ and the drive level, $M$. For a standard micropill diode package mounted in series with a 3.5-mm coaxial line, the equivalent circuit in Figure 7b provides a good approximation across the pump-USB band.

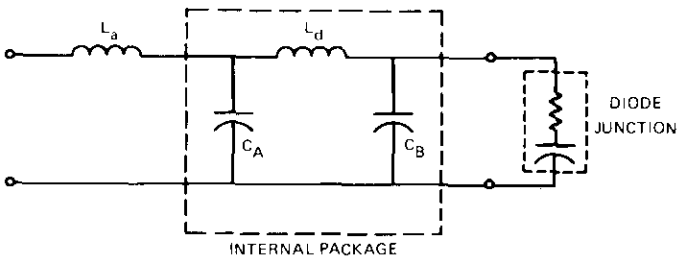



Figure 7a. *Packaged, Mounted Diode Equivalent Circuit*

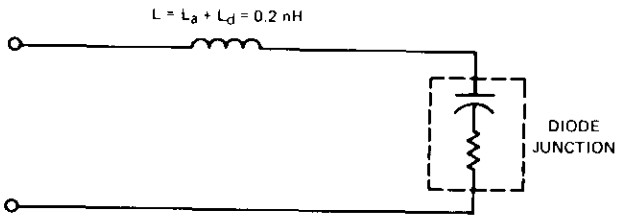



Figure 7b. *Simplified Equivalent Circuit*

The reactance slope for a lumped series $L$-$C$ circuit at the resonant frequency, $f_o$, is

$$\left.\frac{dx}{df}\right|_{f = f_o} = 4\pi L \qquad (7)$$

or $2.5\Omega$/GHz for the packaged device. The reactance slope is modified by adjusting the inductance, $L_a$, which is a function of the diameter of the coaxial line [6]. The resonant frequency of the packaged device (4.7 GHz)

is raised to the pump-USB diplexer mean frequency ($f_o$ = 5.7 GHz) without affecting the reactance slope by adding external series capacitance. The external capacitance is obtained from the stopband reactance of the signal low-pass filter.

The design of a contiguous band diplexer derived from singly terminated prototype filters is described in Section 16.04 of Reference 7. For the design employed here, parallel-connected bandpass filters are coupled to the diode through an impedance inverter [7]. The computed, normalized, driving-point impedance for the lumped element equivalent circuit of the diplexer is shown in Figure 8. In the passband, the real part of the impedance has a Chebychev equiripple response, while the imaginary part has a negative frequency slope and passes through zero at band center. The positive reactance slope, formed principally by the lead inductance and average capacitance of the pumped diode, is used to cancel the negative sloping response of the diplexer and to effectively tune the diode across the pump-USB band.

The real part of the diplexer impedance is adjusted by choosing appropriate characteristic impedances for the filters and impedance inverter to match the average real part of the diode impedance across the pump-USB

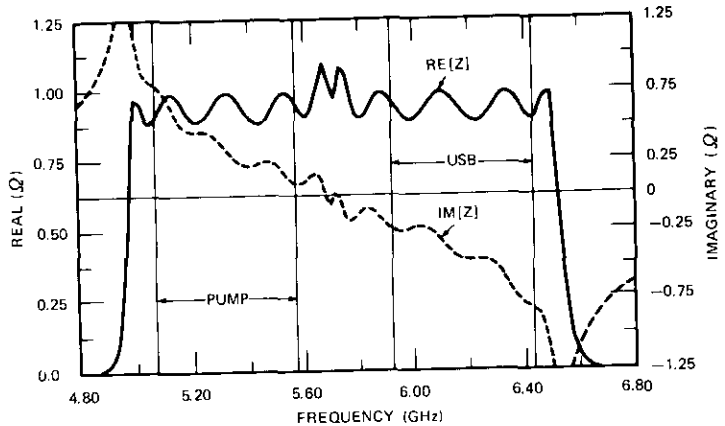



Figure 8. *Computed Normalized Driving Point Impedance for Lumped Equivalent Diplexer*



band. It is a function of the relative power levels at the pump and signal frequencies. The impedances given by equation (1) are based on the choice of power level ratios which are equal to the frequency ratio to obtain optimum conversion efficiency. For good transmission linearity, however, it is desirable to increase the pump-to-signal power ratio, hence decreasing the real part of the pump impedance while increasing the real part of the USB impedance [5]. This tends to offset the differences in the real part of the pump and USB impedances determined by equation (1).

## *Design realization*

Figure 9 is a photograph of the up-converter. It is constructed in 7-mm coaxial line using the split-block technique. The circuit is complete as shown and was designed to match from and into 50Ω. Extensive modeling and computer-aided microwave circuit analysis (GCP-CSC) [8] are

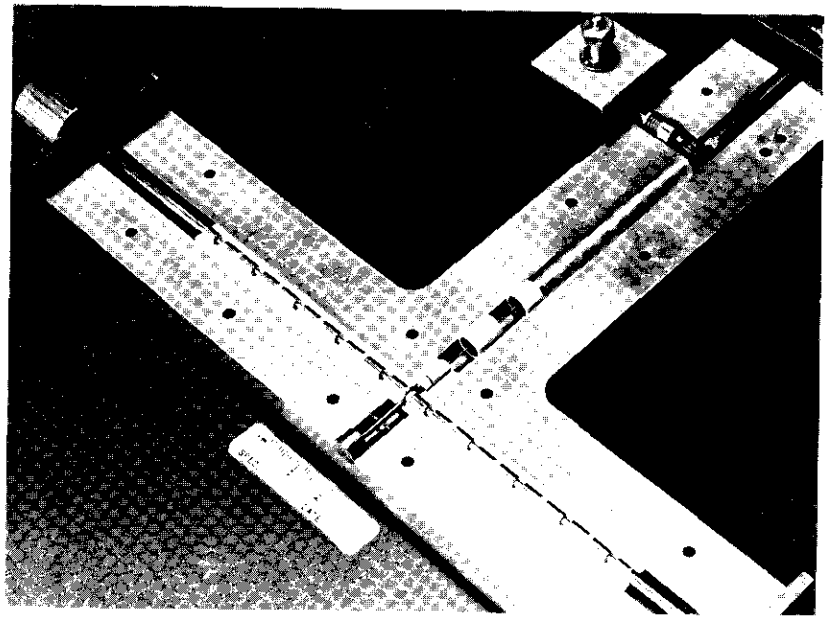


Figure 9. *Up-converter Circuit*

required to arrive at a configuration which is physically realizable and electrically functional before the initial design can be made.

The low-pass filter, which is based on a 0.1-dB Chebychev prototype, is a 5-section stepped impedance design having a cutoff frequency of 1.2 GHz and a characteristic impedance of 32Ω. Signal tuning is accomplished by selecting the appropriate impedance of the short-circuited line (see Figure 6), which is one-quarter wavelength at 5.7 GHz. The 10K shunt resistors* halfway along the line are chosen to be transparent in the passbands of interest. In addition, they have sufficiently low impedance at 10–12 GHz to provide a uniform high impedance to the diplexer junction, which is one-quarter wavelength from the resistors at 11.4 GHz, across the second harmonic band of the pump frequency.

The pump-USB diplexer filters use capacitive coupled resonators, each approximately one-half wavelength long at the filter center frequency. The frequency responses of these resonators are made aperiodic by inserting a high-impedance section at mid-length. Hence, the second passband of each filter is raised to a frequency above that of the pump-USB second harmonic band. The characteristic impedance of the pump and USB filters is 35Ω. The design of aperiodic resonator filters having predictable second passbands is described in Appendix A.

The diplexer is connected to the diode through a 14.5Ω transmission line one-quarter wavelength long at 5.7 GHz, which serves as an impedance inverter. (The actual line is shorter, since it is partly absorbed in the diplexer junction.) The junction design is critical since it influences the adjacent capacitive couplings as well as the impedance inverter. The design details are given in Appendix B.

The computed driving point impedance for the distributed element diplexer, shown in Figure 10, compares well with the measured value. The presence of distributed elements, primarily the diode-diplexer inverter, affects the shape of the reactance curve. A comparison of Figures 8 and 10 indicates that the reactance slope is smaller and more nonlinear with frequency for the distributed element design. The packaged diode reactance slope closely matches that shown in Figure 10.

* British Radio Engineers type RKL-2.



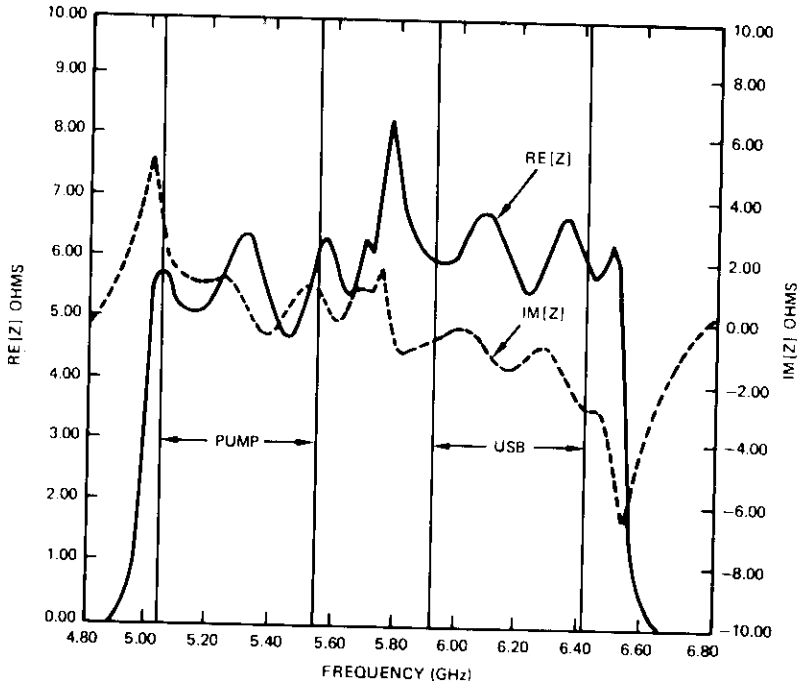



Figure 10. *Computed Driving Point Impedance for the Distributed Element Diplexer*

## Measured performance

### Frequency response

The pump-USB frequency response at various signal levels is shown in Figure 11 for a constant signal frequency of 855 MHz and a constant pump power of 500 mW. This measurement is obtained by driving the up-converter with a constant-level, swept-frequency pump source and a constant-frequency, variable-level signal source. The measured gain is 4.8 dB

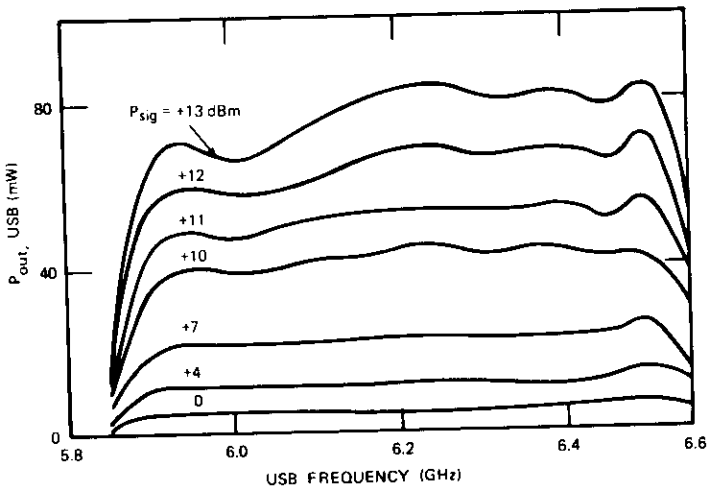



Figure 11. *Pump-Upper-Sideband Frequency Response*

minimum and 5.35 dB maximum across the operating USB band of 5.925–6.425 GHz at a signal level of 10 mW.

The signal-USB frequency response at various signal levels is shown in Figure 12 for a constant pump frequency of 5.4 GHz and a pump level of 500 mW. Although the design signal bandwidth was 100 MHz centered at

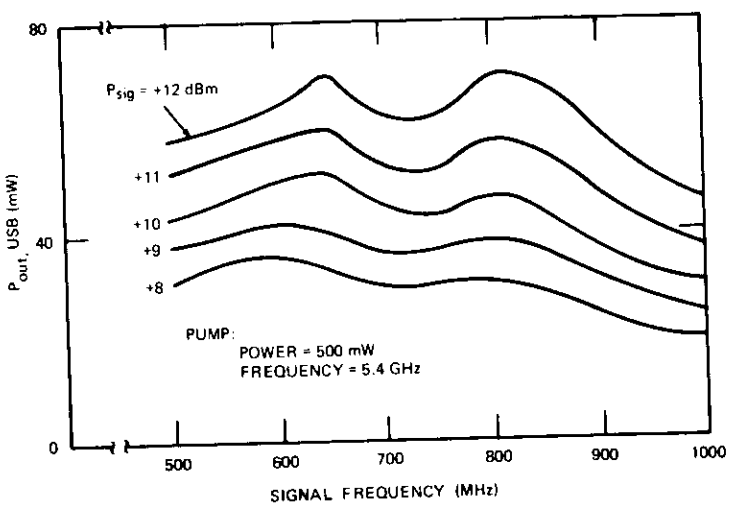



Figure 12. *Signal-Upper-Sideband Frequency Response*



855 MHz, the measured bandwidth was 500 MHz at −1 dB centered at 750 MHz. The large signal bandwidth is attributable to both the low loaded $Q$ of the diode at the signal frequency and the broadband matching at the USB. The maximum gain slope is 0.017 dB/MHz at an input level of 10 mW.

#### Amplitude transfer characteristic

The static amplitude transfer characteristic for a pump frequency of 5.425 GHz and a signal frequency of 855 MHz is shown in Figure 13. It is typical of those measured in the operating band. The transfer characteristics indicate the amplifying properties of the upper sideband up-converter. Linearity at a particular signal level can be improved by raising the pump power.

#### AM/PM conversion

AM/PM conversion measurements are plotted in Figure 14 vs pump power at five pump frequencies for a signal level of +10 dBm. The measurements are made by using a small signal technique described by Laico [9]; i.e., two signals, separated by $\Delta f$ and having a power ratio of 20 dB, are fed into the signal port of the up-converter. The resulting output spectrum contains the large signal (carrier) and two or more sidebands also separated by $\Delta f$. The amplitude ratio of each sideband to the carrier contains AM/PM conversion as well as amplitude compression information. It can be seen that the AM/PM conversion decreases with increasing pump level.

#### Intermodulation distortion

The 3rd-order intermodulation (IM) product amplitude relative to the carrier is plotted in Figure 15 vs signal level for a pump at 5.36 GHz and 470 mW. The carrier/intermodulation (C/I) level in normal operation at a signal level of +10 dBm is −9 dB. Since present earth terminal design employs one carrier per up-converter, IM distortion is of little concern. For multicarrier operation, however, the signal-to-pump power ratio must be set to achieve the desired IM distortion level.

## *Conclusions*

The feasibility of realizing a varactor up-converter covering the pump band from 5.07–5.57 GHz and the output band from 5.925–6.425 GHz has been established. A broad instantaneous bandwidth at all three ports

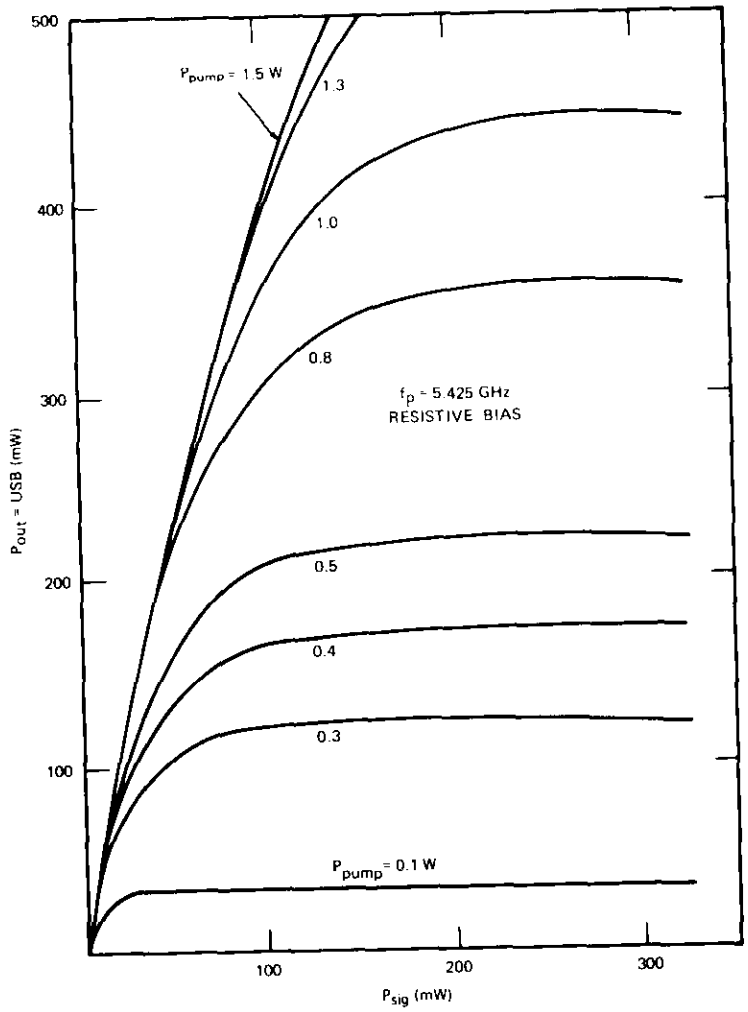



Figure 13. *Transfer Characteristics*

is achieved by operating the varactor at high drive levels for low loaded $Q$ in a low-inductance mount and matching the positive reactance slope of the mounted diode with the negative slope of a contiguous band diplexer.

Measured results confirm the validity of the design technique and also describe the effect of operating levels on the transmission characteristics of the up-converter.



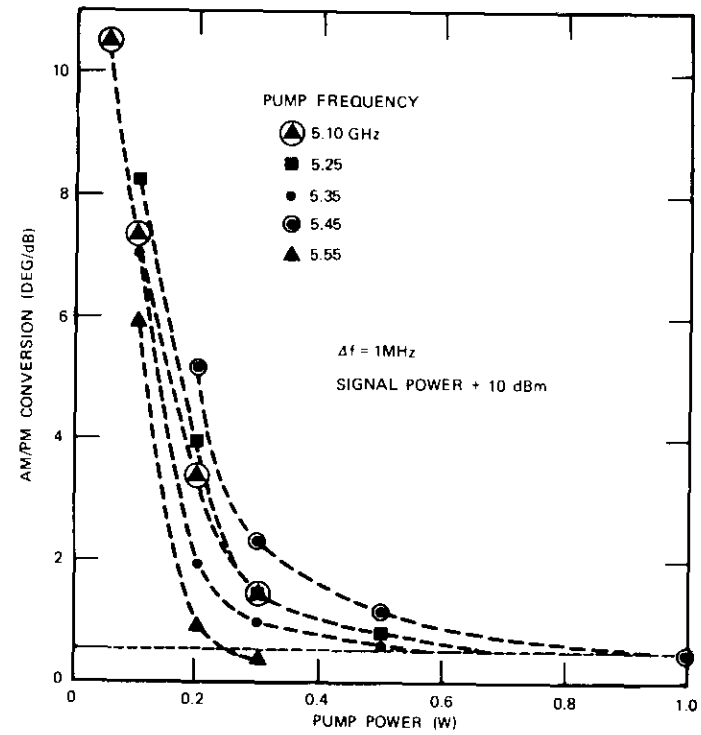



Figure 14. *Up-converter AM/PM Conversion Measurements Obtained by Using Laico's Method*

## Acknowledgment

*The author wishes to thank Mr. William J. Getsinger for his many helpful suggestions, especially those concerning the aperiodic filter design.*

## References

[1] O. Giust, "TH-3 Microwave Radio System: Modulators," *Bell System Technical Journal*, Vol. 50, No. 7, Section II, September 1971, pp. 2155 2173.

[2] H. J. Pratt, W. J. Ince, and R. L. Sicotte, "High Efficiency Varactor Upper-sideband Upconverter," *Proceedings of the IEEE*, Vol. 53, No. 3, March 1965, p. 105.

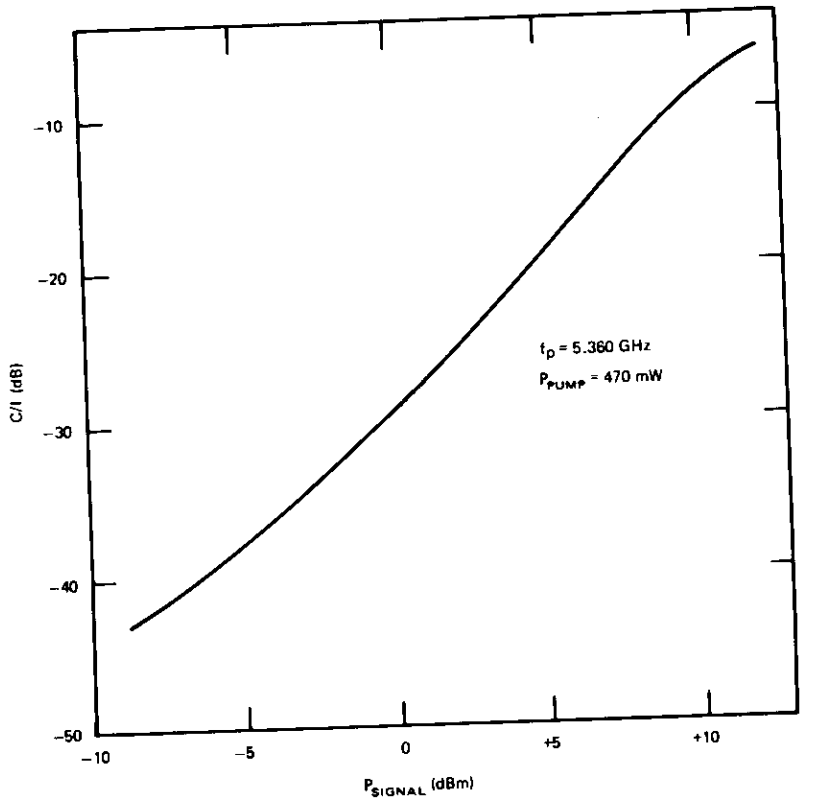



Figure 15. *3rd-Order Intermodulation Product Amplitude Relative to the Carrier vs the Level of each Signal*

[3] A. I. Grayzel and D. Parker, "Conditions for Parametric Oscillations in an Abrupt Junction Non-Overdriven Varactor Doubler," *Proceedings of the IEEE*, Vol. 56, No. 9, September 1968, p. 1589.

[4] J. C. McDade, "Jump Phenomenon in Varactor Diode Circuits," Diamond Ordnance Fuze Laboratories, TR-1008, Department of the Army, AD 269919, 10 January 1962.

[5] S. W. Conning, "An Analysis of the High-Power Varactor Up-converter," *Proceedings of the IREE Australia*, Vol. 27, No. 7, July 1966, pp. 163–173.

[6] W. J. Getsinger, "The Packaged and Mounted Diode as a Microwave Circuit," *IEEE Transactions on Microwave Theory and Techniques*, MTT-14, No. 2, February 1966, pp. 58–69.



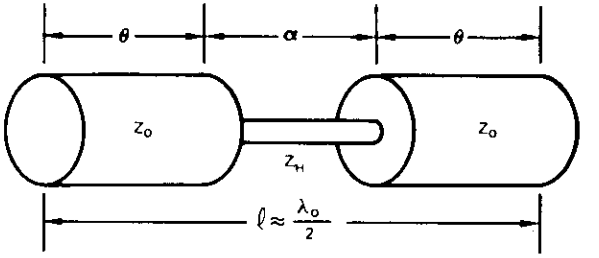
[7] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching Networks and Coupling Structures*, New York: McGraw-Hill, Inc., 1964.

[8] W. J. Getsinger, T. J. Celi, and R. C. Trushel, *Reference Manual for GCP-CSC*, COMSAT Laboratories, October 1970.

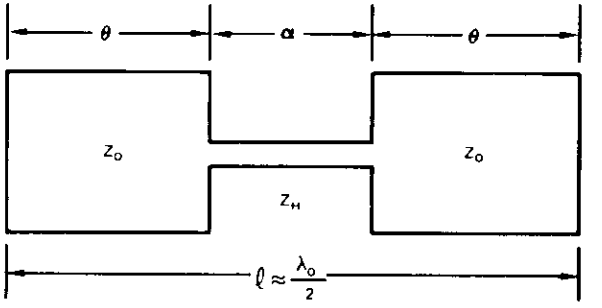
[9] J. P. Laico et al., "A Medium Power Traveling Wave Tube for 6,000-Mc Radio Relay," *Bell System Technical Journal*, Vol. XXXV, No. 6, Section IV, November 1956, pp. 1285–1346.

## Appendix A. Design of capacitive-gap-coupled transmission line filters with aperiodic resonators

The proposed filter is a modification of the capacitive-gap-coupled transmission line filter described in Section 8.05 of Reference A1. Specifically, the basic half-wave resonator is modified, as shown in Figure A-1, to shift the filter's second passband upward in frequency so that it occurs above the second harmonic of signals in the primary passband.




a. *Pictorial for Coaxial Line*




b. *Pictorial for Stripline or Microstrip*

Figure A-1. *Aperiodic Resonator*

The resonator is symmetrical about its midpoint. Hence, half of the resonator, with a short circuit at its midpoint, is used for analysis at the fundamental resonance. The input susceptance is given by

$$B = \frac{1}{Z_o}\,\frac{-Z_o + Z_H \tan(\alpha/2)\tan\theta}{Z_H \tan(\alpha/2) + Z_L \tan\theta} \tag{A1}$$

and the fundamental resonance occurs when

$$\frac{Z_o}{Z_H} = \tan\left(\frac{\alpha}{2}\right)\tan\theta \tag{A2}$$

where $\theta$ is less than $\pi/2$ and $2\theta + \alpha$ is less than $\pi$. The second resonance, obtained by assuming that half of the resonator is open circuited, occurs when

$$\frac{Z_o}{Z_H} = -\frac{\tan M\theta}{\tan M(\alpha/2)} \quad . \tag{A3}$$

In equation (A3), the design parameter, $M$, from which $\theta$, $\alpha/2$, and $Z_o/Z_H$ are derived, must be greater than or equal to 2. A relationship between $\alpha$ and $\theta$ can be found for a particular $M$ by combining equations (A2) and (A3). The resulting transcendental equation is readily solved on a programmable desk calculator using iteration techniques.

Another resonator parameter important to filter design is the slope parameter. The resonator susceptance slope parameter* is obtained by taking twice the derivative of the input susceptance of the shorted half-resonator given in equation (A1):

$$b = \frac{\omega_o}{2}\left[2\frac{dB}{d\omega}\right]\Bigg|_{\omega_o}$$

$$= \frac{1}{Z_o}\,\frac{(\alpha/2)R + \theta\,[R^2\sin^2(\alpha/2) + \cos^2(\alpha/2)]}{[R\sin(\alpha/2)\cos\theta + \sin\theta\cos(\alpha/2)]^2} \tag{A4}$$

where $R = Z_H/Z_o$.

* The resonant frequency and slope parameter for this type of resonator are dependent to a small degree on the load impedance. That dependence is neglected here. Although this results in a less accurate approximation of the center frequency and ripple characteristics, this loss of accuracy was not considered critical to the application of interest here.



The design equations from Figure 8.05–1 of Reference A1 can now be rewritten as follows:

$$\frac{J_{o,1}}{Y_o} = \sqrt{\frac{\omega b/Y_o}{g_0 g_1 \omega_1'}} \tag{A5}$$

$$\frac{J_{k,k+1}}{Y_o} = \frac{b}{Y_o}\,\omega\,\frac{1}{\sqrt{g_k g_{k+1}\omega_1'}} \tag{A6}$$

$$\frac{J_{N,N+1}}{Y_o} = \sqrt{\frac{\omega b/Y_o}{g_n g_{n+1}\omega_1'}} \tag{A7}$$

where $b$ = slope parameter from equation (A4)
$Y_o$ = characteristic admittance of the filter
$\omega$ = fractional bandwidth.

Also,

$$\frac{B_{j,j+1}}{Y_o} = \frac{J_{j,j+1}/Y_o}{1 - (J_{j,j+1}/Y_o)^2} \quad . \tag{A8}$$

Each $\theta_j$ and $\theta_{j+1}$ adjacent to the coupling $B_{j,j+1}/Y_o$ must be shortened by

$$\Delta\theta_j = -\frac{1}{2}\left[\tan^{-1}\frac{2B_{j,j+1}}{Y_o}\right] \quad . \tag{A9}$$

The following is a design example for the pump frequency bandpass filter. The design is based on a singly terminated Chebychev prototype filter having the following characteristics:

$$N = 5$$
$$Z_o = 34.5\Omega$$
$$f_o = 5.33\ \text{GHz}$$
$$\text{bandwidth} = 0.700\ \text{GHz}$$
$$L_{AR} = 0.5\ \text{dB}$$
$$M = 2.6$$

It has been determined from equations (A2) and (A3) that $\theta = 65°$, $\alpha = 20°$, $M = 2.608$, and $Z_o/Z_H = 0.378$ are the required resonator parameters, yielding a center section impedance, $Z_H = 91.23\Omega$. The computed impedance of the filter terminated with a $34.5\Omega$ load is shown in Figure A-2. The second passband begins at 12.8 GHz, which is $12.8/4.8 = 2.66$ times the frequency of the beginning of the fundamental passband.

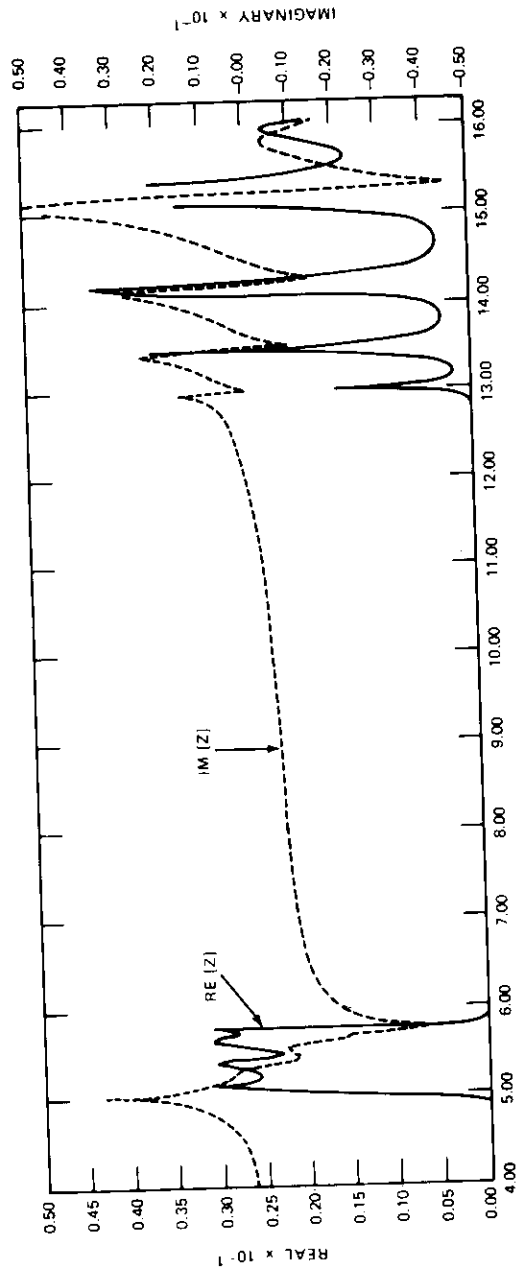



Figure A-2. *Computed Impedance of a Chebychev Filter Terminated with a 34.5Ω Load*



**Reference**

[A1] G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching Networks and Coupling Structures,* New York: McGraw-Hill, Inc., 1964.

## *Appendix B. Diplexer junction design*

Figure B-1 is a sketch of the diplexer junction with the signal return section removed. The electrical equivalent circuit is shown in Figure B-2.

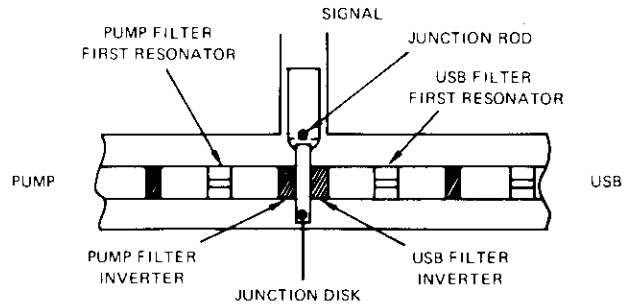



Figure B-1. *Diplexer Junction*

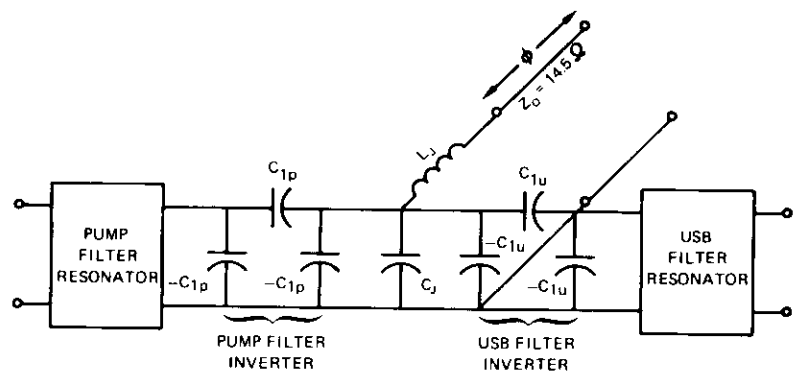



Figure B-2. *Equivalent Circuit*

The parasitic terms $L_J$ and $C_J$ arise from the rod coupling to the junction disk and the disk capacitance, respectively. The negative capacitances $C_{1p}$ and $C_{1u}$

are required to form the filter inverters. The junction is formed by choosing a disk diameter to achieve an inductance $L_J$ and a capacitance $C_J'$:

$$C_J' = C_J - C_{lu} - C_{lp} \tag{B1}$$

The junction can be described as a semilumped transmission line with characteristic impedance $Z_o$ and phase length $\theta$,

where

$$Z_o = \sqrt{\frac{L_J}{C_J}} \tag{B2}$$

and

$$\theta = \sin^{-1}\left[\omega_o \sqrt{L_J C_J}\right] \tag{B3}$$

The disk diameter is chosen to yield $Z_o$ equal to the diode-diplexer inverter characteristic impedance. The length of the inverter transmission line, $\phi$, is calculated by subtracting the value of $\theta$ obtained in equation (B3) from 90°.

*Raymond L. Sicotte received a B.S.E.E. degree from Northeastern University in 1962 and an M.S. degree from the University of Connecticut in 1964. From 1962 to 1963 he was employed by Sylvania Electronic Systems, and in 1964 he became a member of M.I.T.'s Lincoln Laboratory. He joined* COMSAT *Laboratories in 1969. He is presently a Member of the Technical Staff of the RF Transmission Laboratory working on solid-state microwave circuit development for advanced communications systems.*

**Index: synchronous satellites, earth orbits, ephemerides, mathematical models, polynomials.**

# The use of Chebychev polynomials for satellite ephemerides

A. J. CORIO


## Abstract

This article describes the use of Chebychev polynomials to represent satellite ephemerides. This technique requires only half of the computer running time needed to perform a 5-minute tabulation of six position and velocity coordinates. A 1-day ephemeris is stored on disk as 150 numbers, which is one-tenth of its former size. The maximum errors in interpolated position and velocity for synchronous orbit are less than 1 m and $10^{-4}$ m/s, respectively. Both the required equations and a sample application to an INTELSAT IV satellite are included.


## Introduction

For routine maintenance of the global communications satellite system, large quantities of technical information must be predicted by computer programs which require a time history of a satellite's position. To obtain the satellite's coordinates at any time, all significant forces acting on the satellite are used to perform a time-consuming numerical integration of its orbit from some earlier epoch. Because many different computer runs often require the same satellite position history, repeated integration of the same orbit is required. Therefore, satellite ephemerides are created by numerically integrating each satellite's orbit once, and then storing the

An earlier version of the satellite ephemeris files contained the six inertial Cartesian components of position and velocity tabulated at 5-minute intervals. They spanned a period of 20 days and occupied 50 tracks of disk storage per satellite. The search for a new, more compact ephemeris technique was motivated by this large storage requirement.

Several techniques using longer tabulation intervals were developed and evaluated. Any candidate technique was required to produce exact values of position and velocity at equally spaced tabulation points. If desired, these exact position and velocity components at some known time (a particular tabulation point) could be used to initiate numerical integration of the orbit with no initial error.

A development in Chebychev polynomials has proven to be most effective. This technique uses 25 equally spaced data points to obtain coefficients for a 24th-order Chebychev polynomial expansion. Each of the six position and velocity components is expanded separately. In addition to producing exact values at tabulation points, this method achieves greater accuracy at intermediate points. It also reduces computer running time and required computer disk storage.

The remaining sections of this article describe the highlights of this ephemeris technique. The next section contains a brief description of the mathematics involved. The following section discusses the use of the required computer programs. Operational advantages and performance characteristics are discussed in the conclusion.

## *Equations*

Chebychev polynomials have certain properties which make them especially useful for polynomial representation of tabular data. These properties are discussed in Reference 1. Only the pertinent relationships will be included here.

The first two Chebychev polynomials are defined as follows:

$$T_0(x) = 1 \tag{1}$$

$$T_1(x) = x \quad . \tag{2}$$

Higher ordered polynomials are obtained from the 3-term recurrence relationship

$$T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x) \quad . \tag{3}$$



A function of $x$, $f(x)$, with values of $x_i$, $i = 1, n$, can be expressed as a linear combination of Chebychev polynomials. The values of $x$ must be scaled so that they are between $-1$ and $+1$, however. Then,

$$p(x) = \sum_{j=0}^{m} b_j T_j(x) \quad , \qquad m \leq n - 1 \tag{4}$$

where the $b$'s are the constant coefficients of the Chebychev polynomials. The values of these coefficients can be computed by minimizing the sum of the squares of the errors (least squares). The least squares criterion produces the normal equations

$$\sum_{j=0}^{m} b_j \sum_{i=1}^{n} T_k(x_i)\, T_j(x_i) = \sum_{i=1}^{n} f(x_i)\, T_k(x_i) \quad . \tag{5}$$

There are $m + 1$ of these equations over the index $k$.

The polynomial products in equation (5) are obtained from the following relationship:

$$2T_j(x)\, T_k(x) = T_{j+k}(x) + T_{j-k}(x) \quad , \qquad j > k \quad . \tag{6}$$

The matrix form of equation (5) is solved by using Cholesky's square root method for symmetric factorization. This technique, described in Reference 2, basically consists of matrix manipulation to restructure the coefficient matrix to upper triangular form. The solution minimizes the sum of the squares of the differences between the known functional values and the polynomial values at the tabular points. If the numbers of coefficients and tabular points are equal, the polynomial produces the exact functional values at these points. Because of the need to obtain exact coordinate values at sample points, the exact fit is chosen.

For the particular case of an exact fit, there is no difference between an ordinary polynomial and an expansion in Chebychev polynomials. This can be verified by inspecting equations (1) through (4). The expansion, $p(x)$, is the sum of polynomials of increasing order. If the order of the highest ordered Chebychev polynomial equals the order of the ordinary polynomial, then the coefficients of the ordinary polynomial can be obtained by combining the coefficients of equal powers of $x$ in the Chebychev polynomials of equation (4). Since only one ordinary polynomial can

represent the sample points exactly, both must be algebraically equivalent.

The use of equally spaced sample points causes large variations in the interpolation error; these errors are small near the center and large near the end points of the interval. Better tabular data spacing would be achieved by using the zeros of the $m + 1$st Chebychev polynomial, as discussed on p. 253 of Reference 1. This spacing would produce approximately equal interpolation errors in all portions of the interval. The zeros, however, are not equally spaced, but are sparse near the interval center and crowded at the ends. Their values are calculated from the arc cosine definition of the Chebychev polynomials:

$$T_k(x) = \cos(k \arccos x) \quad . \tag{7}$$

For example, let the highest ordered Chebychev polynomial be $T_{24}$. Then the zeros of $T_{25}$ are

$$x_i = \cos\left(i - \frac{1}{2}\right)\frac{\pi}{25} \quad , \qquad i = 1, \quad 25 \quad . \tag{8}$$

## Implementation

Coefficients of the Chebychev polynomials are computed from Cartesian position and velocity components obtained by numerically integrating the orbit. This integration is performed in earth-centered inertial Cartesian coordinates with the $x$-$y$ plane in the earth's equator, the $x$ axis along the first line of Aries, and the $z$ axis positive north. The resulting components are quite accurate since the integration of the equations of motion includes all significant forces.

Encke's method is used to formulate the equations of motion, which are integrated by using a 4th-order Runge-Kutta algorithm. Universal formulations are used to compute the coordinates of the conic reference orbit to avoid the near singularities of geosynchronous orbits (inclination and eccentricity near zero). Reference 3 describes the integration algorithm, and Reference 4 describes Encke's method and universal orbit formulations.

Initially, the 24th-order Chebychev polynomial expansions were obtained from 25 sets of position and velocity components spaced at hourly intervals. Therefore, the expansions represented a 24-hour time period with exact values obtainable at hourly intervals. However, the accuracy of the interpolated values varied widely, as discussed in the preceding section. This divergence was overcome by restricting the use of the polynomial expansion to the central half of its total span. In actuality, the time span was increased to 48 hours and the tabulation interval to two hours, thereby retaining 25 data points. The polynomial expansions thus obtained were valid only during the central 24-hour period, where the maximum interpolation errors were uniformly small.



For this paper, interpolation error is defined as the difference between the position or velocity components computed by the polynomial expansion and those obtained from numerical integration. The polynomial expansion is an exact fit to a sufficient number of data points and represents those points precisely. There is no a priori reason why it should represent the intermediate values determined by numerical integration. The exactness of the polynomial fit is attributable to the order of the polynomial, which is considerably higher than necessary to represent the significant variations in the position or velocity components. (Equivalently, the sampling frequency is higher than the frequency of any significant force term affecting the satellite.) Therefore, the sampled points provide sufficient information to adequately represent the continuous behavior of the position and velocity components. For synchronous orbits, the maximum difference between those positions determined by numerical integration and those obtained from the polynomial expansion is 1 m. For velocity the value is $10^{-4}$ m/s.

This technique, with minor changes, is also used for transfer orbits. Because of the wide range of position and velocity magnitudes encountered in the highly eccentric transfer orbit, the interpolation error varies considerably; it is much worse near perigee. To accommodate this divergence with the same basic programs and still preserve sufficient accuracy throughout the orbit, the time intervals were changed. The length of time represented by a set of polynomial coefficients was reduced from one day to two hours. The data point separation was accordingly reduced to 10 minutes. At apogee the resulting errors were considerably less than in synchronous orbit, and at perigee they were somewhat larger than in synchronous orbit, but still acceptable. Since a transfer orbit file is necessary only during and shortly before a launch, the emphasis was placed on optimizing synchronous orbit ephemerides.

Implementation of the Chebychev polynomial ephemeris files was facilitated by using three existing subroutines in IBM's Scientific Subroutine Package (SSP): DAPCH, DAPFS, and DCNPS. DAPCH constructs the normal equations for Chebychev polynomial coefficients, and DAPFS obtains the solution. DCNPS is a very efficient subroutine which

uses a backward iteration scheme to evaluate the polynomial for a given argument. Even though it is not necessary to use a least squares formulation to obtain the coefficients of the Chebychev polynomials, the availability of existing programs warrants its use. The savings are realized in the shortened ephemeris file reading time and decreased disk storage space. The programs are described in Reference 5.

As an example, consider the inertial $x$ coordinate of INTELSAT IV F-4 stationed over the Pacific Ocean at 174° east longitude in near-geosynchronous orbit. Table 1 contains the coefficients of the Chebychev polynomials used to obtain the inertial $x$ coordinate from 0 hr GMT on March 17, 1973, to 0 hr GMT on March 18, 1973. The surrounding 12-hour segments are plotted in Figure 1 since they contributed to the coefficients. Figure 2 is a plot of the difference between the polynomial expansion's $x$ coordinate and the integrator's $x$ coordinate. Note the convergence to an absolute value of less than 1 m within the central 24-hour span.

TABLE 1. COEFFICIENTS OF CHEBYCHEV POLYNOMIALS FOR INERTIAL $x$ COORDINATE (INTELSAT IV F-4, 1200 GMT, MARCH 16, 1973, TO 1200 GMT, MARCH 18, 1973)

| Order | Coefficient (m) |
|---|---|
| 0 | $-0.92597054 \times 10^{7}$ |
| 1 | $0.34860881 \times 10^{7}$ |
| 2 | $-0.23955071 \times 10^{8}$ |
| 3 | $0.40126335 \times 10^{6}$ |
| 4 | $-0.25838487 \times 10^{8}$ |
| 5 | $-0.62540717 \times 10^{7}$ |
| 6 | $0.23104196 \times 10^{8}$ |
| 7 | $0.26709765 \times 10^{7}$ |
| 8 | $-0.61496570 \times 10^{7}$ |
| 9 | $-0.49690572 \times 10^{6}$ |
| 10 | $0.85342817 \times 10^{6}$ |
| 11 | $0.52661226 \times 10^{5}$ |
| 12 | $-0.75815534 \times 10^{5}$ |
| 13 | $-0.35568938 \times 10^{4}$ |
| 14 | $0.17832437 \times 10^{4}$ |
| 15 | $-0.27794478 \times 10^{3}$ |
| 16 | $-0.20320295 \times 10^{4}$ |
| 17 | $-0.16829076 \times 10^{3}$ |
| 18 | $-0.11241256 \times 10^{4}$ |
| 19 | $-0.10875161 \times 10^{3}$ |
| 20 | $-0.61717093 \times 10^{3}$ |
| 21 | $-0.40535123 \times 10^{2}$ |
| 22 | $-0.18527560 \times 10^{3}$ |
| 23 | $-0.84512777 \times 10^{1}$ |
| 24 | $-0.41552232 \times 10^{2}$ |



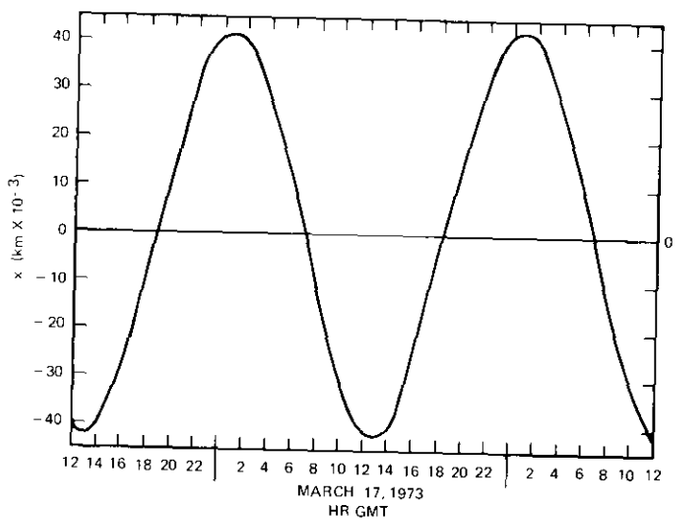



Figure 1. *Inertial x Coordinate* (INTELSAT IV F-4, *1200 GMT, March 16, to 1200 GMT, March 18, 1973*)

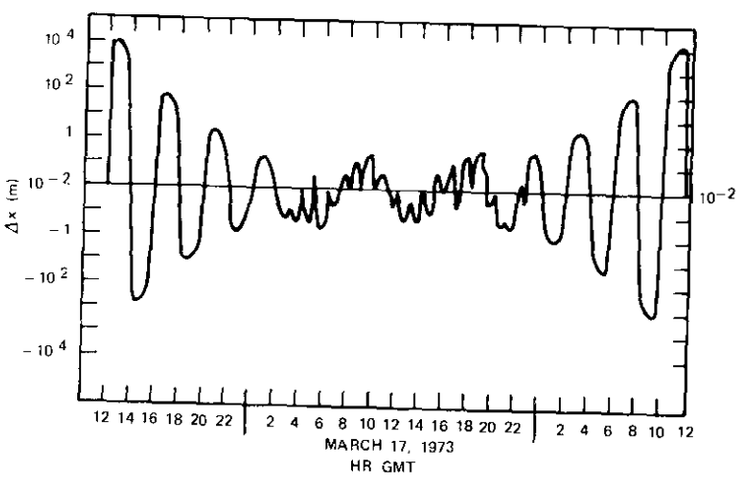



Figure 2. *x Coordinate Difference* (INTELSAT IV F-4, *1200 GMT, March 16, to 1200 GMT, March 18, 1973*)

## Conclusion

Use of the new ephemeris files and operating programs has resulted in substantial savings in disk storage and running time; the size of a satellite's 20-day ephemeris file has been reduced from 50 tracks to 5 tracks. Several similar runs for both ephemeris file systems indicate that running times for the file writing and reading programs have been reduced by one-third and one-half, respectively.

In contrast to many almanac ephemerides, these tabulated coefficients do not require supplementary interpolation factors; they contain all of the information needed for interpolation. The interpolation is carried out with a maximum error which does not exceed 1 m in position or $10^{-4}$ m/s in velocity for synchronous orbits.

## References

[1] R. W. Hamming, *Numerical Methods for Scientists and Engineers*, New York: McGraw-Hill Book Co., Inc., 1962, Chapter 19.

[2] Mario G. Salvadori and Melvin L. Baron, *Numerical Methods in Engineering*, New York: Prentice-Hall, 1955, 2nd edition, pp. 23–28.

[3] J. B. Scarborough, *Numerical Mathematical Analysis*, Baltimore: The Johns Hopkins Press, 1958, 4th edition, pp. 314–319.

[4] Richard H. Battin, *Astronautical Guidance*, New York: McGraw-Hill Book Co., Inc., 1964, pp. 49–52, 188–190.

[5] "System/360 Scientific Subroutine Package," 360A-CM-03X, *Version III Programmer's Manual*, 4th edition, H20-0205-3, pp. 199, 260–265.

*Anthony J. Corio received a B.S. in Physics from Manhattan College in 1959 and an M.S. in Physics from Syracuse University in 1965. Since 1967 he has been a Member of the Technical Staff in the Astrodynamics Department of the Celestial Mechanics Division of* COMSAT. *His duties include development of mathematical methods for determining and controlling satellite orbital and attitude motion. He was previously employed by Melpar, Inc. and General Electric.*

Index: communications satellite, transponders, modulation.

# Application of hybrid modulation to FDMA telephony via satellite

G. R. WELTI

## Abstract

This paper describes a promising new form of hybrid modulation which requires less power and less bandwidth per channel than angle modulation systems. Theoretical calculations show that the voice channel capacity of a satellite transponder can be approximately doubled by using 2-pulse amplitude-and-phase modulation instead of the present FM/FDMA method.

## Introduction

Hybrid modulation systems combine two forms of modulation to reduce the power and bandwidth requirements for data or waveform transmission. A number of studies describing various forms of hybrid modulation [1]–[6] have been reported. These systems are based on the concept of "twisted modulation" [7]–[9].

COMSAT's investigations of hybrid modulation began in 1970 when it was realized that this technique would be potentially beneficial to spectrum conservation. Initial studies concentrated on 1-pulse amplitude-and-phase modulation (1P-APM) [10]. The results of these studies were applied to an

This paper is based upon work performed in COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

investigation of the potential benefits of 1P-APM for future communications satellite systems. It was found that the inherent spectral savings of 1P-APM could not be fully realized in a satellite using state-of-the-art antennas because there was insufficient isolation between beams to permit frequency reuse. Moreover, a severe power demand was imposed on the satellite as a result of the high transponder linearity required by 1P-APM.

Recently, attention was focused on 2-pulse APM (2P-APM). This version of hybrid modulation relaxes the isolation requirements intended to prevent cochannel interference; it also requires less total RF power per channel and less transponder linearity than 1P-APM. A program for implementing a 2P-APM modem is currently underway.

This paper describes a 2P-APM system designed specifically for efficient FDM/FDMA telephone transmission. Basic performance calculations show that 2P-APM requires less power and less bandwidth per channel than angle modulation systems. Use of 2P-APM instead of the present FM/FDMA method in a typical satellite transponder would roughly double its voice channel capacity. For example, a spot-beam transponder with a usable bandwidth of 35 MHz and a saturated e.i.r.p. of 29.6 dBW could provide 1,680 voice channels using earth stations with G/T = 40.7 dB/°K.

## Description of 2-pulse APM

Assume that the modulator shown in Figure 1 has an input $s$, and that this signal is Gaussian with a standard deviation $\sigma_s$. Let $s^* = s/\sigma_s$.

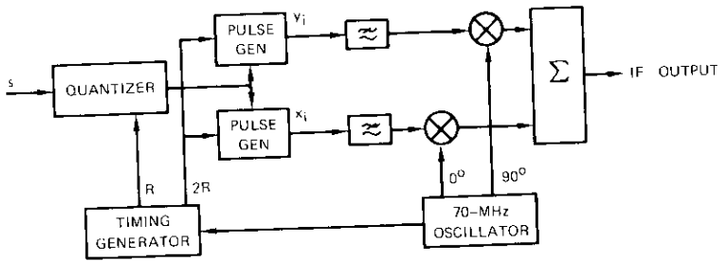



Figure 1. *Modulator*

The input is sequentially sampled at a rate of $R$ samples per second. The sampled input signal is approximated by the nearest of seven possible voltage levels (the coarse levels, $s_C$). A first IF pulse, having one of seven



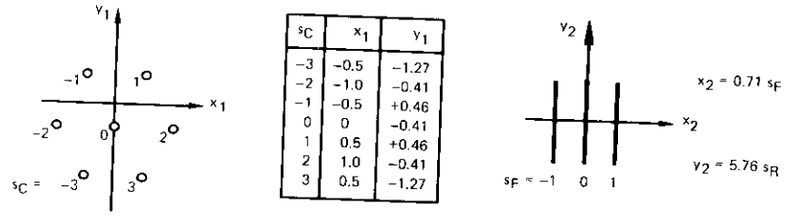
possible amplitude-and-phase combinations to represent this coarse level, is transmitted. The difference between the actual signal sample and the coarse approximation is then represented by the closest of three possible levels (the fine levels, $s_F$). Finally, the remainder ($s_R$), or the difference between the actual signal and the second approximation, is also measured.

A second IF pulse is formed by combining a cosine and a sine wave. The amplitude of the cosine wave has one of three possible voltages representing the fine levels, and the sine wave is directly proportional to the remainder. Thus, the amplitude and phase of the second pulse depend on both the fine level and the measured remainder. More precisely, $s_C$ is the integer in the set $\{-3, -2, \ldots, 3\}$ which minimizes $|s^* - s_C|$, $s_F$ is the integer in the set $\{-1, 0, 1\}$ which minimizes $|s^* - s_C - s_F/3|$, and $s_R = s^* - s_C - s_F/3$. Hence, two IF pulses, $f_1(t)$ and $f_2(t)$, are generated for each sample of the input signal:

$$f_1(t) = x_1 \cos \omega t + y_1 \sin \omega t, \qquad 0 < 2Rt < 1$$

$$f_2(t) = x_2 \cos \omega t + y_2 \sin \omega t, \qquad 1 < 2Rt < 2$$

where $x_1$, $y_1$, $x_2$, and $y_2$ are related to $s_C$, $s_F$, and $s_R$ as illustrated in Figure 2.




| $s_C$ | $x_1$ | $y_1$ |
|---|---|---|
| −3 | −0.5 | −1.27 |
| −2 | −1.0 | −0.41 |
| −1 | −0.5 | +0.46 |
| 0 | 0 | −0.41 |
| 1 | 0.5 | +0.46 |
| 2 | 1.0 | −0.41 |
| 3 | 0.5 | −1.27 |

Figure 2. *Signal Mapping for FDMA*

The signal mapping shown in Figure 2 resulted from repeated trial calculations of $S/N$ improvement and threshold location such as those presented in the next two sections. This process was intended to minimize the average required $C/N$ for frequency-division multiplex telephony. The scale used in Figure 2 is arbitrary; it defines a unit of distance as the distance between signal loci in the $x_1$, $y_1$ plane.

The NRZ pulse trains are low-pass filtered in two separate baseband filters. The filter outputs amplitude modulate two quadrature 70-MHz carriers, which are added to produce an IF output. A numerical example

illustrating this process is as follows. Suppose that $s = 1.4\ \sigma_s$. Then,

$$s^* = 1.4$$
$$s_C = 1$$
$$s_F = 1$$
$$s_R = 0.067 \quad .$$

From Figure 2,

$$x_1 = 0.5$$
$$y_1 = 0.46$$
$$x_2 = 0.71$$
$$y_2 = 0.385 \quad .$$

Hence, the IF pulse waveforms are

$$f_1(t) = 0.5 \cos \omega t + 0.46 \sin \omega t, \qquad 0 < t < 1/2R$$
$$f_2(t) = 0.71 \cos \omega t + 0.385 \sin \omega t, \qquad 1/2R < t < 1/R \quad .$$

The coherent demodulator is shown in simplified form in Figure 3. The incoming 70-MHz IF signal, which passes through an automatic voltage control amplifier, is demodulated and low-pass filtered with matched filters. The filter outputs are sampled at rate $2R$, and the detected levels are converted to signals $s_C$, $s_F$, and $s_R$. Samples of baseband signal $s = s_C + s_F/3 + s_R$ are reconstituted by a digital-to-analog converter. A continuous output is produced by a final low-pass filter. Amplitude, phase, and pulse timing are recovered by feedback tracking loops. (Accurate performance of these loops is important, and imposes a difficult design challenge.)

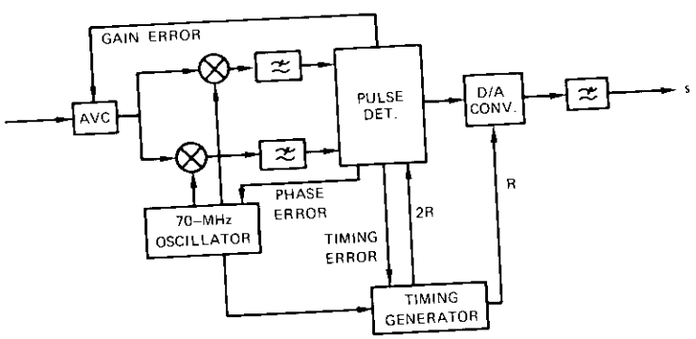



Figure 3. *Demodulator*



## Signal-to-noise ratio improvement

A method for calculating the signal-to-noise ($S/N$) ratio improvement for 1P-APM has been described previously [1]. This method can be generalized for $n$-pulse APM by assuming that the signal mapping function is the mapping of a line into a $2n$-dimensional signal space. Let $V = \{x_1, y_1, x_2, y_2, \ldots, x_n, y_n\}$ be a vector in the signal space, and let $x_i$, $y_i$ represent the in-phase and quadrature components of the $i$th pulse, respectively. The average total energy received for a given baseband sample $s$ is therefore equal to the average of $V \cdot V = \Sigma_{i=1}^{n} (x_i^2 + y_i^2)$, while the average noise energy is $2n\sigma_n$. Hence, the average carrier-to-noise ($C/N$) ratio is

$$\frac{C}{N} = \frac{\sum_{i=1}^{n} (\overline{x_i^2} + \overline{y_i^2})}{2n\sigma_n^2} \quad . \tag{1}$$

Above threshold, the baseband $S/N$ ratio is $\sigma_s^2/\sigma_n^2$, where $\sigma_s^2$ is the variance of $s$, the distance along the mapping contour. Therefore, the $S/N$ improvement factor is

$$\frac{2n\sigma_s^2}{\sum_{i=1}^{n} (\overline{x_i^2} + \overline{y_i^2})} \quad . \tag{2}$$

For the mapping functions described in Figure 2, $n = 2$ and $\sigma_s = 5.76$. To compute the carrier power, it is convenient to approximate the actual probabilities of $x_2$ and $y_2$ with a uniform probability density along the contour in the $x_2$, $y_2$ plane. In the $x_1$, $y_1$ plane exact probabilities associated with the Gaussian baseband signal must be used. Therefore,

$$\overline{x_1^2} = 2[(p_1 + p_3)(0.5)^2 + p_2] \tag{3a}$$

$$\overline{y_1^2} = (p_0 + 2p_2)(0.40)^2 + 2p_1\left(\frac{\sqrt{3}}{2} - 0.40\right)^2 + 2p_3\left(\frac{\sqrt{3}}{2} + 0.40\right)^2 \tag{3b}$$

$$\overline{x_2^2} = \frac{2}{3}(0.71)^2 \tag{3c}$$

$$\overline{y_2^2} = \frac{1}{3}(0.96)^2 \tag{3d}$$

where $p_i = \text{erf}(i + 0.5) - \text{erf}(i - 0.5)$. The resulting $S/N$ improvement above threshold is equal to a power ratio of 122.7, which corresponds to 20.9 dB.

### Threshold location

When the $C/N$ ratio is below threshold, the effects of decision errors in the receiver become significant. For the mapping in Figure 2, decision errors can occur when the absolute value noise voltage exceeds 0.5 on the first pulse or 0.355 on the second pulse. The resulting baseband noise power can be calculated by multiplying the various transition probabilities by the corresponding error power. In practical cases each transition probability is almost exactly equal to the product of the probability of sending a particular symbol times the appropriate error function. The fact that transitions are not independent can be ignored. Moreover, the transitions between non-neighboring symbols can be neglected. Therefore, the threshold noise power for the mapping is almost exactly

$$\begin{aligned} N_T &= 2\,\text{erfc}\left(\frac{0.5}{\sigma_n}\right)[p_0(1 + 2^2 + 3^2) + p_1(1 + 1 + 2^2) \\ &\quad + p_2(1 + 1 + 2^2) + p_3(1 + 3^2 + 6^2)] \\ &\quad + \text{erfc}\left(\frac{0.355}{\sigma_n}\right)\left\{\frac{1}{3}\left(\frac{1}{3}\right)^2 + \frac{1}{3}\left[\left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right] + \frac{1}{3}\left(\frac{1}{3}\right)^2\right\} \\ &= 14.34\,\text{erfc}\left(\frac{0.5}{\sigma_n}\right) + 0.1481\,\text{erfc}\left(\frac{0.355}{\sigma_n}\right) \quad . \end{aligned} \tag{5}$$

The noise power thus obtained is normalized with respect to the baseband signal power. To determine the location of the threshold, it is necessary to combine $N_T$ with the noise power, $N_S$, contributed by noise components aligned with the signal locus. The normalized value of $N_S$ is

$$N_S = \left(\frac{\sigma_n}{\sigma_s}\right)^2 = \frac{\sigma_n^2}{5.76^2} \quad . \tag{6}$$

The combined overall baseband noise power ratio ($NPR$),*

$$NPR = \frac{122.7\,(C/N)}{1 + (N_T/N_S)} \tag{7}$$

* The ratio of the mean absolute power level of the equivalent uniform spectrum baseband signal to the baseband noise power level.



is plotted in Figure 4 as a function of the $C/N$ given by equation (1). Threshold occurs at $C/N = 13.5$ dB and $NPR = 33.4$ dB. Also shown in the figure is the rate distortion bound for a bandwidth expansion factor of 4. At threshold, the $C/N$ ratio is about 6 dB from this bound.

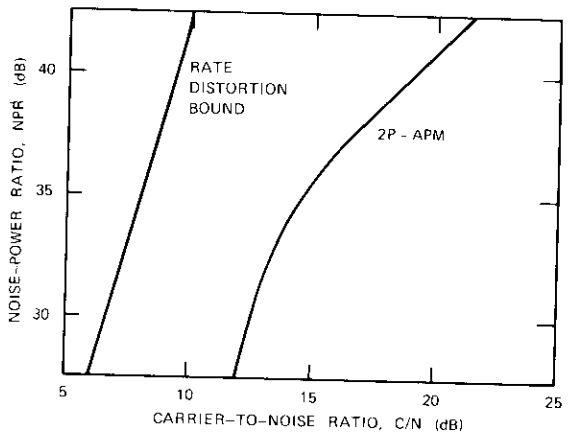



Figure 4. *2P-APM Signal-to-Noise Improvement*

### Bandwidth requirement

The power spectrum, $\Phi(f)$, of the baseband input to the modulator may be considered to be uniform between $f_1$ and $f_2$, where $0 < f_1 < f_2$. The spectrum associated with periodic samples of the baseband depends on the sampling waveform. In the limit, as the duration of the sampling pulses approaches zero, this spectrum becomes

$$\lim \frac{1}{2n + 1} \sum_{i=-n}^{n} \Phi(f + iR) \tag{8}$$

where $R$ is the sampling rate. To prevent spectral overlap, $R$ must exceed $2f_2$.

In the demodulator, the spectrum characterized by equation (8) must be low-pass filtered to eliminate the terms for which $i \neq 0$. This filtering, which is performed by the final output filter, must not distort the spectrum between $f_1$ and $f_2$, but must stop the band above $R - f_2$. The rolloff

region of the output filter thus extends from $f_2$ to $f_2 + (R - 2f_2)$. In practice, the width of the rolloff region, $R - 2f_2$, is about 20 percent of the top baseband frequency, $f_2$. Therefore,

$$R = 2.2f_2 \quad . \tag{9}$$

The RF spectrum is governed by the value of $R$ and the choice of pulse-shaping filter in the modulator. In an FDMA application, this filter must be designed to permit close spacing of adjacent carriers entering the satellite. The use of Nyquist filters with small rolloff factors is appropriate for this purpose. Use of a Nyquist filter for transmitting $2R$ pulses per second confines the RF spectrum to the range $f_i \pm (1 + \rho)\ R$, where $f_i$ is the $i$th carrier frequency, and $\rho$ is the rolloff factor. Adjacent carrier interference is eliminated by using matched Nyquist filters at the demodulator input provided that $|f_{i+1} - f_i| > 2(1 + \rho)\ R$. Therefore, the minimum occupied bandwidth for 2P-APM is about $4.4(1 + \rho)\ f_2$. To conserve the RF spectrum, it is desirable to choose a baseband signal structure that minimizes $f_2$, and to select the smallest practical rolloff factor, $\rho$. For example, two voice telephony supergroups (a total of 120 voice channels) can be confined to a 492-kHz baseband spectrum and transmitted with a 15-percent-rolloff filter in an occupied bandwidth of 2.5 MHz. This corresponds to a spectrum capacity of 48 channels per megahertz.

### Ideal comparisons with other systems

Figure 5 compares the ideal performance of various modulation systems, including 1- and 2-pulse APM. Bandwidth and power requirements are shown relative to the requirements for single-sideband modulation. Also shown are contours of constant RF $C/N$ ratio, and Shannon's rate-distortion bound for a Gaussian baseband $NPR$ of 32 dB, which is appropriate for large FDM telephone multiplex assemblies. All realizable systems lie above this bound.

The following assumptions are used in Figure 5. For FM, APM, and AM, a 10-dB peak factor is assumed, and $NPR = 32$ dB. The FM curve includes a 4-dB pre-emphasis gain and shows the Carson's Rule bandwidth. The APM and PSK bandwidths are the noise bandwidths obtained by sampling at the Nyquist rate. The 1P-APM performance is based upon the signal mapping described in Reference 10, while the PSK performances are based on 8-bit coding and a bit-error rate of $10^{-4}$. The peak power values for PSK give the approximate envelope power at the output of a Nyquist filter which can be exceeded with a probability of $10^{-5}$.



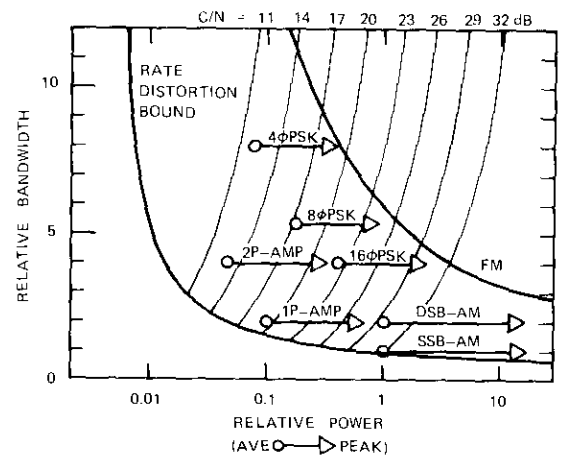



Figure 5. *Modulation Systems Comparisons*

Figure 5 does not include practical system degradations, which vary among modulation systems. Nevertheless, it may be used as an approximate guide for comparing well-engineered FDMA systems. It should be noted that 2P-APM uses less power and less bandwidth than any of the angle modulation systems.

### Practical FDMA link performance

The required overall $NPR$ for a 120-channel telephone multiplex assembly is 32.9 dB [11]. Figure 4 shows that this requires an equivalent $C/N$ ratio of 13.5 dB. This $C/N$ must be achieved in the presence of earth station equipment noise, interference from terrestrial systems in the same band, rain attenuation effects, intermodulation distortion in the satellite transponder, interference from adjacent carriers, cochannel interference caused by frequency reuse in the satellite, and nonideal modem implementation.

A noise budget for these impairments is given in Table 1. Effects of transponder nonlinearities are included in item 4. Items 3, 4, and 5 add up to 3,400 pW, which corresponds to a total $C/N$ ratio of $(C/N)_T = 18.2$ dB. The difference between this value and the 13.5-dB $C/N$ requirement represents a margin of 4.7 dB for items 1, 2, and 6–9.

As shown previously, the required total $C/N$ ratio $(C/N)_T$, including margins, is 18.2 dB. This ratio, which combines up-link, down-link, and

TABLE 1. FDMA NOISE BUDGET FOR 2P-APM

| Item | Noise Source | pW Relative to 1-mW Test Tone | C/N (dB) |
|---|---|---|---|
| 1 | Earth Station Equipment | 1,500 | 21.7 |
| 2 | Terrestrial Interference | 1,000 | 23.5 |
| 3 | Up-Link Noise | 140 | 32.0 |
| 4 | Intermodulation Noise | 1,560 | 21.6 |
| 5 | Down-Link Noise | 1,700 | 21.2 |
| 6 | Cochannel Interference | 900 | 24.0 |
| 7 | Adjacent Channel Interference | 200 | 30.5 |
| 8 | Modem Implementation | 2,000 | 20.5 |
| 9 | Rain | 1,000 | 23.5 |
| | Combined Value | 10,000 | 13.5 |

satellite intermodulation noise, can be provided by a typical satellite FDMA spot-beam transponder using the transmission parameters listed in Table 2. These transmission parameters were calculated by using the computational methods developed by Fuenzalida, Shimbo, and Cook [12]. The total transponder capacity for this example is 1,680 channels, or about twice the capacity achieved with multicarrier FM/FDMA in a transponder with the same parameters.

TABLE 2. FDMA TRANSMISSION PARAMETERS FOR 2P-APM

| | |
|---|---|
| Earth Station e.i.r.p. per 120-Channel Carrier | 78 dBW |
| Up-Link Frequency | 6 GHz |
| Satellite G/T | −11.6 dB/°K |
| Saturation Flux Density | −60 dBW/m² |
| Satellite Input Backoff | 13.5 dB |
| Satellite Output Backoff | 6.8 dB |
| Satellite Saturated e.i.r.p. | 29.6 dBW |
| Transponder Usable Bandwidth | 35 MHz |
| Number of 120-Channel Carriers | 14 |
| Down-Link Frequency | 4 GHz |
| Earth Station G/T | 40.7 dB/°K |

## Conclusion

This paper has described a form of hybrid modulation which appears to be suitable for FDM/FDMA transmission. A program for developing and testing this technique is underway. Additional studies of hybrid modulation systems for other forms of transmission will be undertaken.



## References


[1] M. Pent, "Hybrid P.C.M.-F.M. Communication System," *Electronics Letters*, Vol. 2, No. 3, March 1966, pp. 121–123.

[2] P. C. J. Hill, "Hybrid P.C.M.—An Experimental Video System," *Electronics Letters*, Vol. 5, No. 16, August 7, 1969, pp. 352–354.

[3] D. D. McRae, C. J. Palermo, and M. G. Pelchat, "Analog Coding," *1970 International Symposium on Information Theory*, Noordwijk, The Netherlands, June 15–19, 1970.

[4] D. D. McRae, "Performance Evaluation for a New Modulation Technique," *IEEE Transactions on Communication Technology*, COM-19, No. 4, August 1971, pp. 431–445.

[5] J. Saltz, J. R. Sheehan, and D. J. Paris, "Data Transmission by Combined AM and PM," *Bell System Technical Journal*, Vol. 50, No. 7, September 1971, pp. 2399–2419.

[6] K. Kawai, S. Shintani, and H. Yanagidaira, "Optimum Combination of Amplitude and Phase Modulation Scheme and Its Application to Data Transmission Modem," *Proc. International Communications Conference*, June 1972, pp. 29-5 to 29-11.

[7] V. A. Kotel'nikov, *The Theory of Optimum Noise Immunity*, New York: Dover Publications, Inc., 1960.

[8] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*, New York: John Wiley and Sons, Inc., 1965.

[9] D. J. Sakrison, *Communication Theory: Transmission of Waveforms and Digital Information*, New York: John Wiley and Sons, Inc., 1968.

[10] G. R. Welti, "Pulse Amplitude-and-Phase Modulation," *Second INTELSAT/IEE International Conference on Digital Satellite Communications*, Paris, November 1972.

[11] International Radio Consultative Committee (C.C.I.R.), *XIIth Plenary Assembly, New Delhi, 1970*, Vol. IV, Part 2, pp. 87–88, Geneva: International Telecommunications Union, 1970.

[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.

*George R. Welti received B.S. and M.S. degrees in Mechanical Engineering in 1947 and 1948, respectively, and an M.S. degree in Electrical Engineering in 1955 from M.I.T. He has been Manager of the Advanced Studies Laboratory at* COMSAT *Laboratories since 1968. He was previously employed by Westinghouse Electric Corporation; DECO Electronics, Inc.; Raytheon Co.; M.I.T.; and Sperry Gyroscope Company. He holds a number of patents in fire control, radar, and communications.*

# CTR Notes

## A microstrip balanced transistor amplifier with collector-base feedback for 0.6—1.1 GHz

C. B. COTNER

### Introduction

A preliminary study of the modulation methods and their parametric requirements for the COMSAT Laboratories unattended earth terminal led to the decision to use an intermediate frequency centered on 0.855 GHz, with a bandwidth of 500 MHz. In addition, the objective of unattended operation necessitated reliability and configuration requirements which would permit continuity of service without an elaborate system of monitoring, spares, and switching. This note describes a transistor amplifier designed to accommodate these requirements.

### Design objective

In the normal mode of operation, when an interruption occurs in a receive or transmit chain, it is necessary to switch to standby equipment. An alternative, which avoids information disruption and sensitive monitoring and switching, is parallel redundancy. For this mode of operation, the probability of complete interruption may be made low enough so that emergency repair is unnecessary. Parallel redundancy is achieved with a balanced configuration in which two identical stages are cross-connected at both input and output by quadrature hybrids (see Figure 1).

It may be shown [1] that a transistor amplifier with three balanced stages using parallel redundancy has a mean time to failure which is up to three times that of an amplifier with three single stages using switched redundancy. This figure does not take into account the failure probabilities involved in the necessary sensing and switching equipment associated with switched redundancy.

Although reliability is the primary concern here, use of the amplifier with parallel redundancy results in additional benefits, such as broadband

*C. B. Cotner is a member of the Technical Staff, RF Transmission Laboratory, Technology Division, COMSAT.*

input and output return losses better than 20 dB and superior dynamic range and intermodulation performance. With such excellent return losses, cascading of stages is facilitated. Further, these modular balanced stages may be used as building blocks to "tailor" the amplifier gain for a particular application.

## Amplifier design and construction

### DESIGN

The design effort concentrated upon equalizing the gain in each transistor circuit of the balanced pair and obtaining broadband performance in the hybrids at input and output. The transistor (Hewlett Packard 35821E) was selected on the basis of its proven performance and characteristics, and its availability in a stripline package. Of the several techniques investigated for broadband gain equalization [2]-[5], the most promising was collector-base feedback, which also proved to be simple to implement. Hence, the required gain equalization was obtained without using multiple pole matching networks or accurately tuned circuits.

Computer programs were used to select feedback network element values compatible with the transistor parameters. The resulting balanced, common (grounded) emitter amplifier is shown in Figure 1.

*Quadrature Hybrids.* A critical component in the amplifier design was the single, quarter-wavelength, interdigitated hybrid, which has been described by Lange [6]. A sample hybrid using an improved cross-bonding technique was constructed for the frequency range of interest; its performance is shown in Figures 2 and 3. Based upon these results, this hybrid was simply rescaled slightly in frequency for the amplifier design.

### CONSTRUCTION

Thin film techniques are used throughout the amplifier with a ceramic substrate (99.5-percent-pure aluminum oxide) 0.050 inch thick. With this substrate thickness, 50Ω transmission lines require conductors approximately 0.050 inch wide, which are physically compatible with many chip components. Tolerance problems in etching the hybrid are also less severe than for thinner substrates.

The transistors are mounted in holes through the alumina substrate, minimizing the lead length between the emitter and ground plane. Lead connection is accomplished by gap welding. Chip components are bonded in place using a conductive epoxy. The necessary cross-bondings in the hybrid are protected by small drops of nonconductive epoxy.



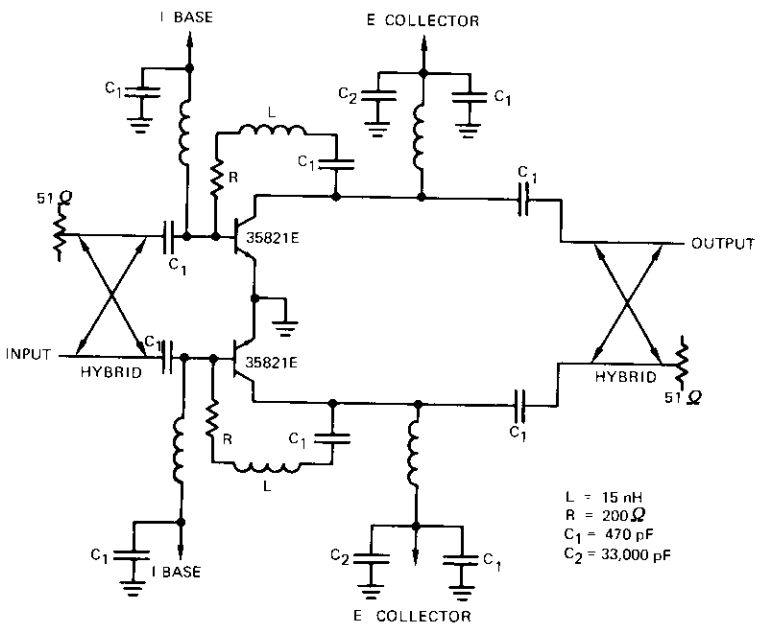



Figure 1. *Balanced Amplifier Schematic Diagram*

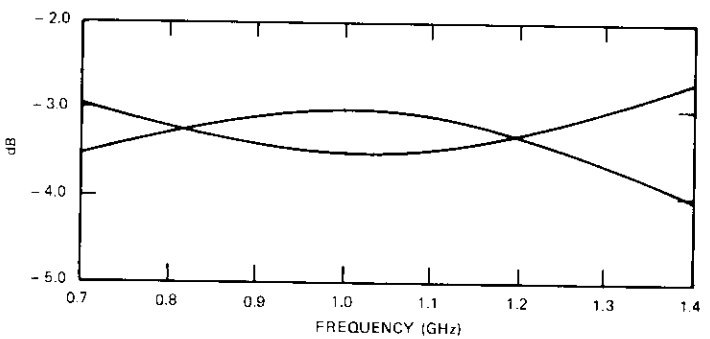



Figure 2. *Interdigitated Hybrid Power Split*

Bias lines are meander patterns whose overall length is $\lambda/4$ at the center frequency of 0.855 GHz. The required hybrid terminations are uncompensated 51Ω chip resistors with return losses of 28.5 dB or more in this frequency range.

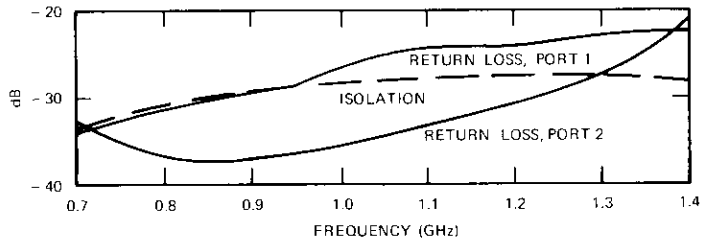



Figure 3. *Return Loss and Isolation for Interdigitated Hybrid*

**Results**

Figure 4 shows measured and calculated amplifier gain versus frequency performance for a single amplifier stage and a single-ended "half stage," respectively. Return losses obtained in the balanced stage are shown in Figure 5. The difference between measured and calculated gain is

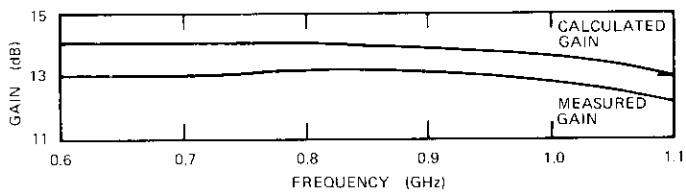



Figure 4. *Gain vs Frequency*

attributed to hybrid insertion loss, mismatch loss to the fourth port of the hybrid, and other circuit losses. As expected, no stability or gain ripple problems are experienced in cascading three of these stages. Figure 6 shows the measured 1- and 2-carrier transfer characteristic of three cascaded stages. Figure 7 shows a particular physical configuration used in the unattended earth terminal down-converter.

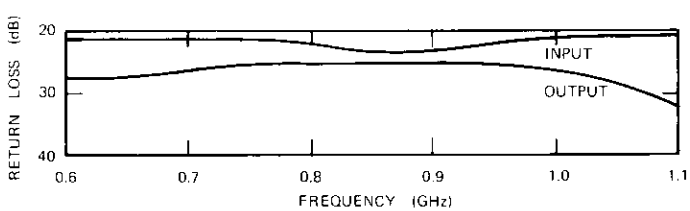



Figure 5. *Return Loss vs Frequency*



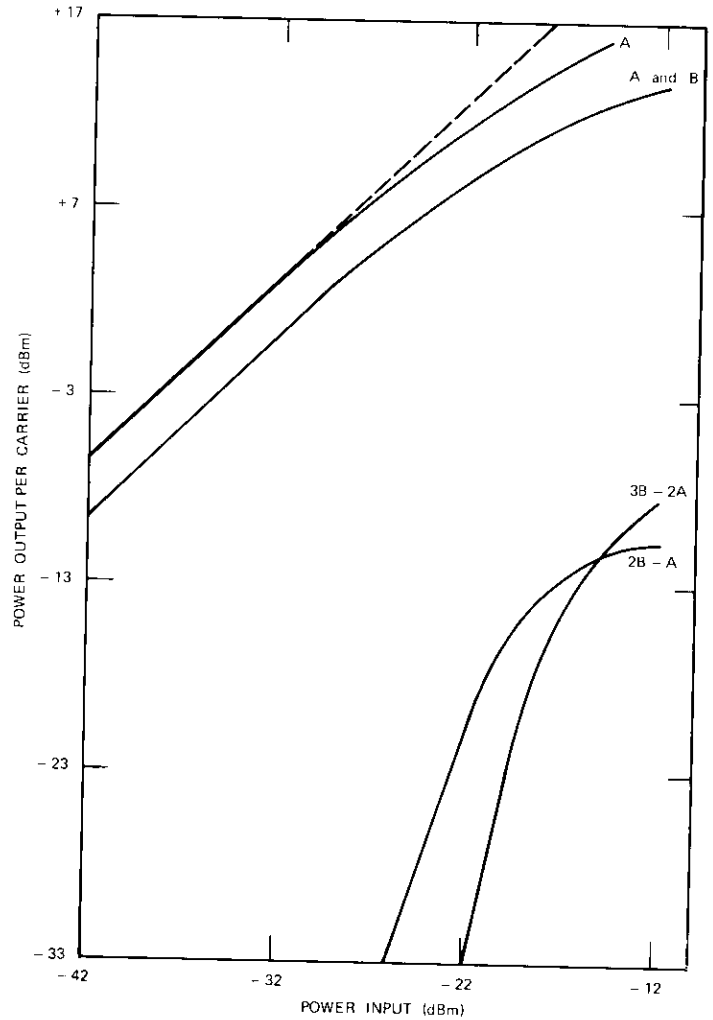



Figure 6. *Transfer Characteristics of 3-Stage Balanced MIC Transistor Amplifier (carrier A = 855 MHz, carrier B = 845 MHz)*

**Acknowledgments**

*The writer wishes to thank Messrs. Louis Pollack and Laurence Gray for their encouragement in the development of this amplifier. Thanks are*

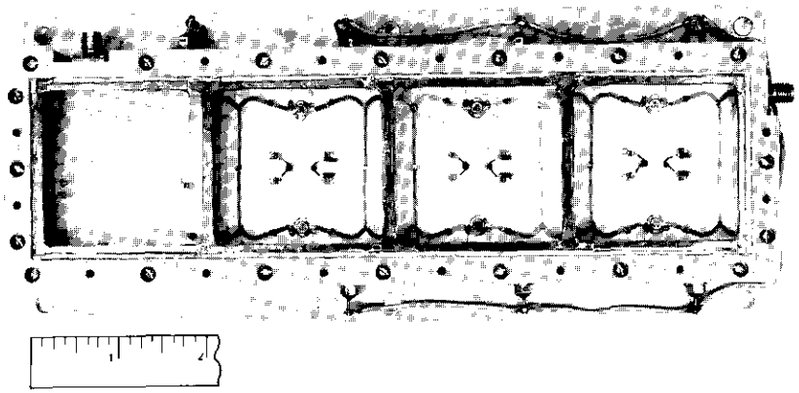


Figure 7. *Mixer and 3-Stage Preamplifier*

*also extended to Dr. William Sones, Mr. Arthur Standing, and Mr. Ronald Stegens for their technical suggestions, and to Dr. A. Atia for the use of his computer program. Much of the computer analysis and actual construction was done by Messrs. Nils Jespersen and Richard Perry.*

**References**

[1] ARINC Research Corp., *Reliability Engineering*, edited by W. H. Von Alven, Englewood Cliffs, N.J.: Prentice-Hall, 1964.

[2] R. S. Engelbrecht and K. Kurokawa, "A Wide Band Low Noise L-Band Balanced Transistor Amplifier," *Proceedings of the IEEE*, Vol. 53, No. 3, March 1965, pp. 237–247.

[3] T. E. Saunders and P. D. Stark, "An Integrated 4-GHz Balanced Transistor Amplifier," *IEEE Journal of Solid State Circuits*, SC-2, No. 1, March 1967, pp. 4–10.

[4] W. H. Frochner, "Quick Amplifier Design with Scattering Parameters," *Electronics*, Vol. 40, No. 21, October 16, 1967.

[5] S. C. Chao, "Computer Aided Design of Broadband UHF Transistor Amplifiers," *The Microwave Journal*, Vol. 12, No. 7, July 1969, pp. 78-82.

[6] J. Lange, "Interdigitated Stripline Quadrature Hybrid," *IEEE Transactions on Microwave Theory and Techniques*, MTT-17, No. 12, December 1969, pp. 1150–1151.

# *Chemical storage of hydrogen in Ni/$H_2$ cells*

M. W. Earl and J. D. Dunlop

## Introduction

Under INTELSAT support, COMSAT Laboratories has pioneered the development of nickel/hydrogen batteries for possible replacement of nickel/cadmium batteries as the energy storage system in communications satellites. Potential advantages of the nickel/hydrogen battery in terms of energy density, cycle life, and trouble-free operation have been discussed in a previous paper [1].

In normal operation hydrogen gas is generated on charge and consumed on discharge. The overall electrochemical cell reaction is

$$NiOOH + \tfrac{1}{2} H_2 \uparrow \underset{\text{charge}}{\overset{\text{discharge}}{\rightleftharpoons}} Ni(OH)_2 \quad .$$

For optimum energy density, the operating pressure is from 3 to 33 atm. This note describes a new approach in which hydrogen is stored as a rare earth metal hydride, thereby reducing the operating pressure to a few atmospheres.

## Concept

J. H. N. van Vucht et al. have reported on the reversible absorption of hydrogen by the intermetallic compound $LaNi_5$ [2]. Their isotherms (Figure 1) show that, at 21°C, $LaNi_5$ absorbs hydrogen at a relatively low pressure to form the hydride $LaNi_5H_6$ whose hydrogen density is nearly twice as high as that of liquid hydrogen.

These properties make $LaNi_5$ extremely interesting for use in the $Ni/H_2$

This note is based upon work performed in COMSAT Laboratories under the sponsorship of the International Telecommunications Satellite Organization (INTELSAT). Views expressed are not necessarily those of INTELSAT.

*M. W. Earl is a Senior Technician, Energy Storage Department, Physics Laboratory, Applied Science Division, COMSAT.*

*J. D. Dunlop is Manager, Energy Storage Department, Physics Laboratory, Applied Science Division, COMSAT.*

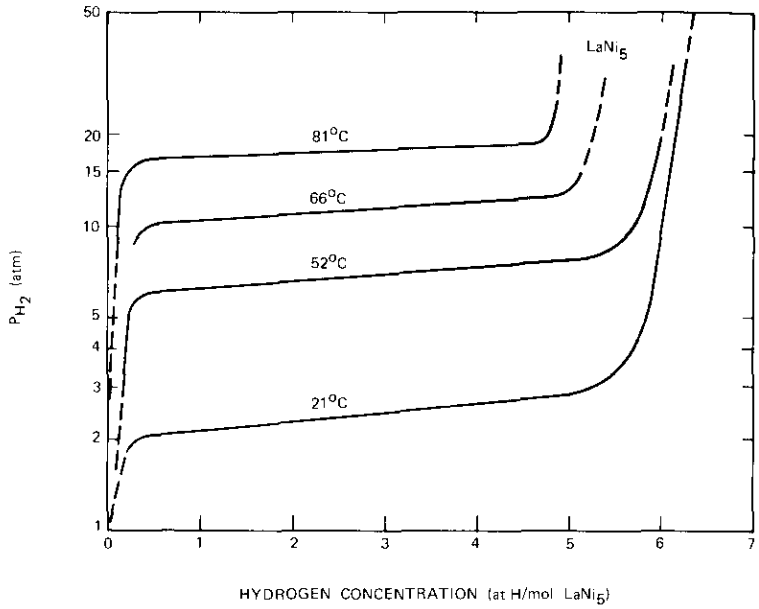



Figure 1. *Pressure-Composition Isotherms of Hydrogen Gas in Equilibrium with Absorbed Hydrogen in $LaNi_5$* [2]

cell.* In the new approach described here, hydrogen is stored as a hydride during charging and desorbed during discharge:

$$LaNi_5H_6 \underset{\text{charge}}{\overset{\text{discharge}}{\rightleftharpoons}} LaNi_5 + 3H_2\uparrow$$

Hence, the cell operating pressure is reduced to approximately 1 atm [3].

Commercial $LaNi_5$ was activated by high pressure absorption—vacuum desorption cycles at high temperature according to the procedure described by J. J. Reilly and R. H. Wiswall [4]. The hydride formed was $LaNi_5H_{1.8}$. Although the hydrogen content of this hydride was lower

*Chemical storage of hydrogen is not restricted to the nickel/hydrogen couple, but is equally applicable to other metal/hydrogen couples (e.g., silver/hydrogen) and in general to many nonelectrochemical hydrogen gas systems.



than had been previously reported, it was deemed sufficient for initial investigations.

**Experimental work**

To demonstrate the feasibility of this concept, an experimental nickel/hydrogen cell with hydride storage (Figure 2) has been built. The electrode stack has a sintered nickel hydroxide positive electrode, a polypropylene separator, and a platinum catalyzed hydrogen negative electrode. This stack is assembled in a heavy walled pressure vessel with a stainless steel reactor which contains 5.29 g of $LaNi_5$. Hydrogen absorption by this amount of $LaNi_5$ to form $LaNi_5H_{1.8}$ corresponds to 1.55 Ahr of capacity. The electrochemical cell capacity is 1.6 Ahr.

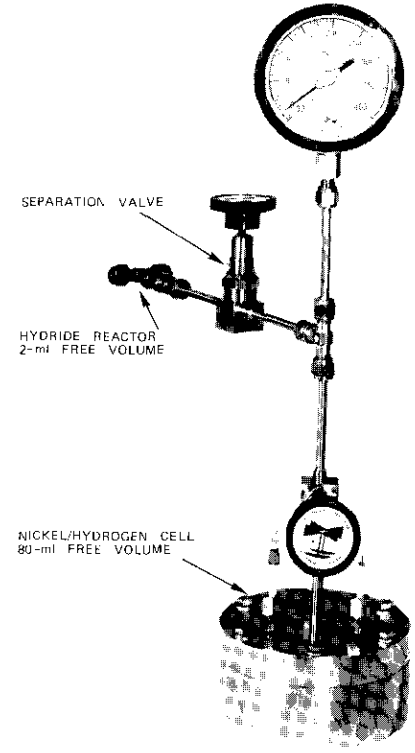



Figure 2. *Experimental Nickel/Hydrogen Cell with Hydride Reactor Attached*

Over 1,000 charge-discharge cycles have been completed on this cell at room temperature. Throughout the test, the pressure range during cycling was consistently 2.5–6 atm per cycle, indicating that the $LaNi_5$ had not been contaminated by $O_2$, $H_2O$, or KOH. This is important since it has been reported that $O_2$ and $H_2O$ vapor may inhibit hydrogen absorption by $LaNi_5$ [5]. The $Ni/H_2$ cell environment contains less than 1 percent oxygen, which is apparently less than the contamination level for this particular application.

After 1,000 one-hour cycles the cell was completely discharged and the voltage and pressure vs time were recorded. The hydride was then isolated from the cell by closing the valve to the reactor containing the $LaNi_5$. Finally, the cell was recharged and discharged, and the pressure without the hydride was recorded (Figure 3). A comparison of the cell operating pressures with and without the hydride demonstrates that the cell operating pressure is significantly reduced by using the hydride. The pressure at the beginning of discharge is reduced by a factor of three. The amount of pressure reduction is a function of the electrochemical capacity, cell free volume, and quantity of $LaNi_5$, as well as the degree of hydride activation.

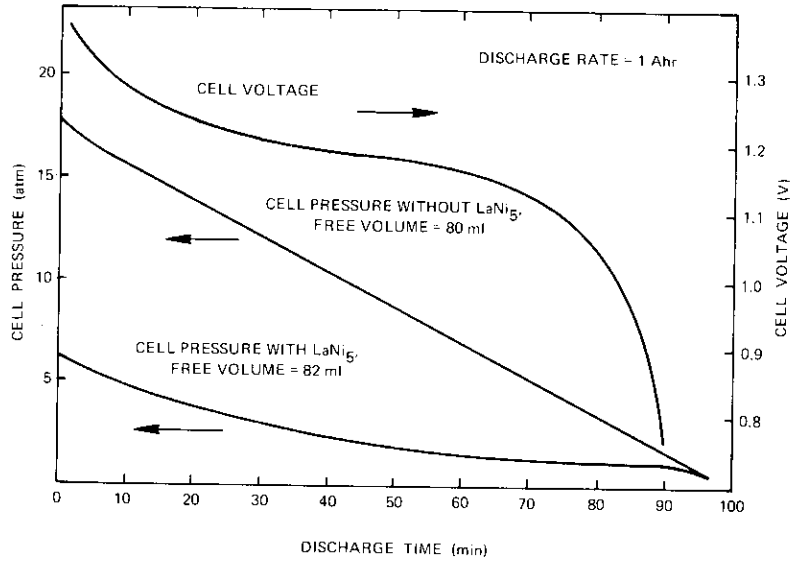



*Figure 3. Cell Pressure and Voltage for Complete Discharge with and without $LaNi_5$*



According to Figure 1 the hydrogen pressure should assume a value of about 3 atm instead of the observed range of 2.5 to 6 atm. However, the absorption of hydrogen in $LaNi_5$ is an exothermic process (7.2 kcal/mol $H_2$), whereas desorption is an endothermic process. This nonisothermal behavior partially explains the gradual change in pressure during a cycle and the higher pressure observed at the end of charge, i.e., at the beginning of discharge.

This experiment clearly demonstrates the hydride's usefulness in reducing the operating pressure of nickel/hydrogen cells. As a result, prismatic cell configurations with their associated ease of construction, volume reduction, better electrolyte management, and structural integrity become feasible and preferable to cylindrical structures. An experimental prismatic cell is being designed to incorporate these advantages as a further step in the exploitation of this concept.

## Conclusion

Test results demonstrate the feasibility of using $LaNi_5$ hydride to reduce the operating pressure of a nickel/hydrogen cell without altering its high cycle-life expectancy. The advantages of this concept are as follows:

a. elimination of concerns associated with operation at high hydrogen pressures;

b. cell volume reduction of almost 50 percent; and

c. simplification of cell pressure vessel design through prismatic-type cell construction, thereby eliminating electrolyte loss problems, improving the electrode stack design to better withstand shock and vibration, and simplifying battery packaging.

## References

[1] J. F. Stockel et al., "A Nickel/Hydrogen Secondary Cell for Synchronous Orbit Application," *Seventh Intersociety Energy Conversion Engineering Conference*, San Diego, California, September 1972, *Proceedings*, p. 87.

[2] J. H. N. van Vucht, F. A. Kuijpers, and H. C. A. M. Bruning, "Reversible Room-Temperature Absorption of Large Quantities of Hydrogen by Intermetallic Compounds," Philips Research Report, Vol. 25, 1970, pp. 133–140.

[3] J. Dunlop, M. Earl, and G. Van Ommering, U.S. patent pending.

[4] J. J. Reilly, and R. H. Wiswall, "The Reaction of Hydrogen with Alloys of Magnesium and Copper," *Inorganic Chemistry*, Vol. 6, 1967, p. 2220.

[5] R. H. Wiswall, Jr., and J. J. Reilly, "Metal Hydrides for Energy Storage," *Seventh Intersociety Energy Conversion Engineering Conference*, San Diego, California, September 1972, *Proceedings*, p. 1342.

# *Frequency reuse in collocated earth and terrestrial stations*

H. DODEL AND B. PONTANO

## Introduction

A recent study [1] postulated the feasibility of frequency reuse at an earth station in the fixed-satellite service and a collocated terminal of an associated radio-relay interconnect link. This concept hinges on the degree of coupling that must be expected between the terminals of a large earth station antenna and the relatively large aperture of a collocated radio-relay antenna when these facilities are located back to back, each in the other's near field.*

To obtain experimental corroboration of the predicted feasibility, a test was undertaken at the Cayey, Puerto Rico, earth station. This note will present a brief summary of the measurement procedures and the results, verifying the practicability of such collocation at common frequencies.

## Coupling measurements

### FIELD MAPPING

To obtain a general picture of the radiation in the rear sector of the Cayey earth station antenna, the level of a transmitted signal received by a small aperture standard-gain horn was mapped over a wide range of relative geometries. For each location, the plane of polarization and pointing direction of the standard-gain horn were adjusted for maximum coupling (maximum received signal). Typical measurement results shown in Figure 1 agree with those obtained from similar tests performed at the Etam earth station several years ago [2], as well as with recent theoretical calculations. With the exception of a slight increase in coupling near the

*The term "near field" is applied to all geometrical configurations for which a large radiating aperture cannot be considered to constitute a point source. Neither the Fraunhofer nor Fresnel approximations may be applied to these configurations.

*H. Dodel is a member of the Technical Staff, Spectrum Utilization Department, Systems Engineering Division, COMSAT.*

*B. A. Pontano was Manager, Laboratory Simulation, Systems Division, COMSAT Laboratories.*

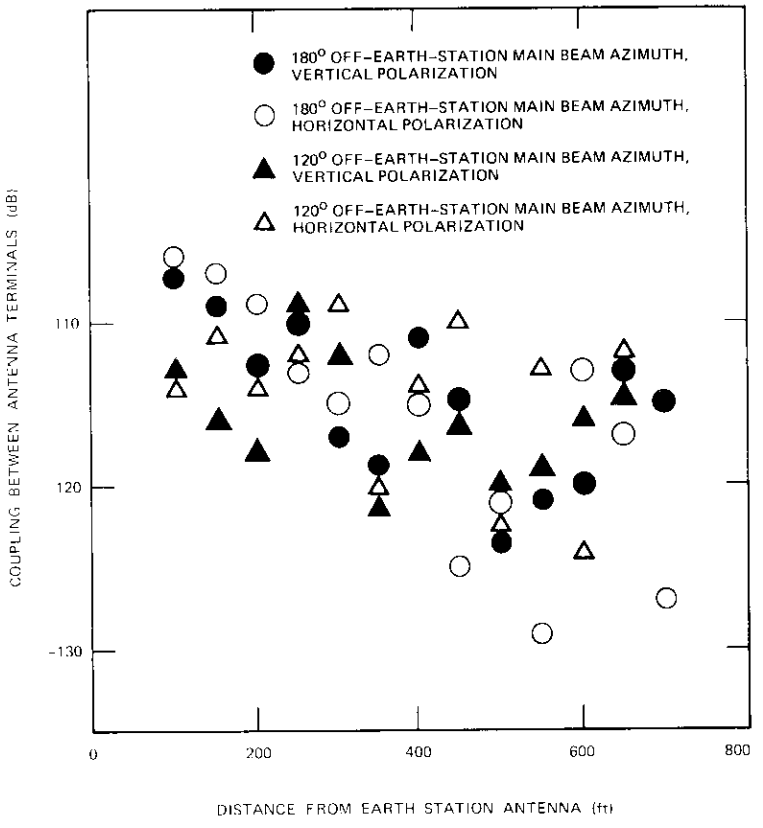



Figure 1. *Coupling Between the Terminals of the Earth Station Antenna and the Terminals of the Standard-Gain Horn as a Function of Distance*

rear focal axis of the earth station antenna, there appeared to be little variation in maximum coupling over a range of about 150°.

COLLOCATED RADIO-RELAY ANTENNA

A 3-meter-diameter horn reflector antenna whose main beam was directed at various angles away from the earth station antenna was used to replace the standard-gain horn at several of the previously mapped locations. Figure 2 compares the coupling measurements for this antenna with those obtained by using the standard-gain horn at the same locations. For the 3-meter horn antenna, the maximum coupling measurement showed less variation at different locations and little sensitivity, over a



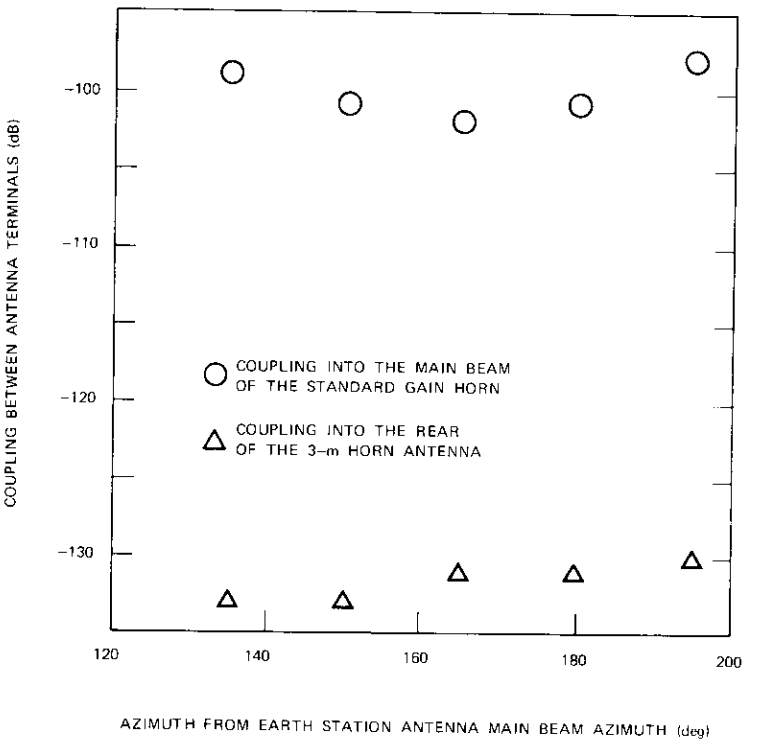



Figure 2. *Comparison of the Coupling from the Earth Station Antenna (distance from the earth station = 100 ft)*

large angular range, to variations of the pointing azimuth at each location. When the plane of polarization was rotated, the coupling variation was of the same order of magnitude as that of the standard-gain horn (about ±3 dB) and different geometries showed different preferred polarization planes.

There is evidence that much of the coupling was caused by wide-angle ground scatter from the region in front of the 3-meter antenna and beneath its main beam. In moderate rain, coupling tended to decrease somewhat relative to that observed in the absence of rain. Over longer time periods with the same geometry, coupling varied over an appreciable range, but peaked consistently near the values of Figure 2.

**Link interference tests**

To assess the effects of coupling in an actual operational configuration, a radio-relay link was set up. This test configuration, shown in Figure 3,

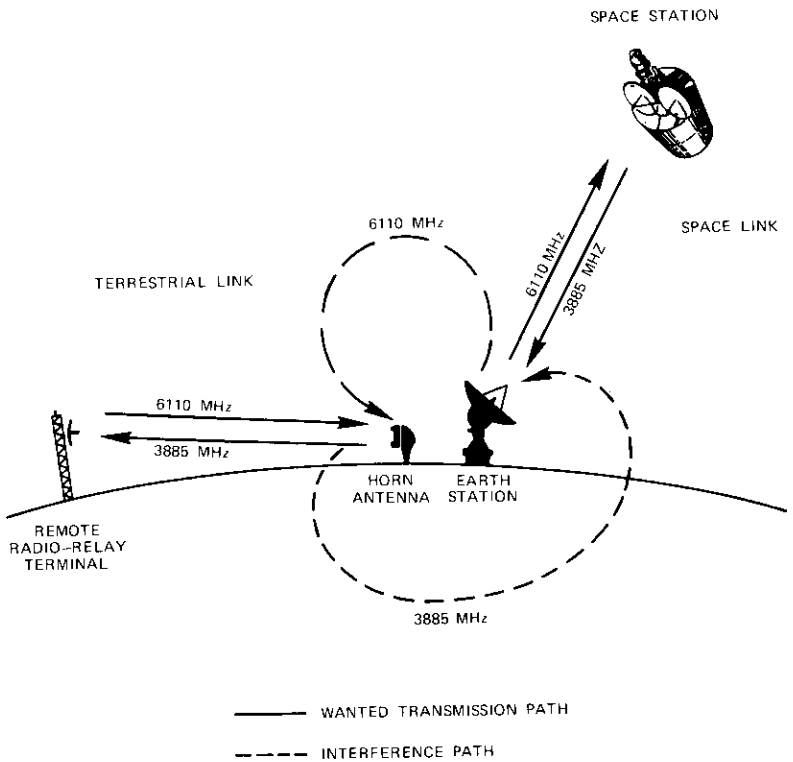



Figure 3. *Test Configuration*

used a frequency plan which was believed to result in minimum implementation constraints on the terrestrial link. The collocated 3-meter horn antenna was deliberately placed where relatively strong coupling had previously been measured.

Noise power ratio (NPR) measurements were made for a wide variety of space and terrestrial link carrier sizes and transmission parameters. Standard INTELSAT IV performance parameters were used in the space link. As expected, the strongest interference occurred from emissions of the collocated horn antenna into the satellite signal received at the earth station (in the 4-GHz band). Measurements of wanted-to-unwanted carrier ratios, NPR, and coupling for a range of capacities and transmission parameters were consistent and in good agreement with expected values.



TABLE 1.

| Space Link | | Terrestrial Link | | | Coupling Factor (dB) | Measured C/I (dB) | Space Link $(S/N)_w$ (dB) | |
|---|---|---|---|---|---|---|---|---|
| Channels | Modulation Index | Transmit Power (dBW) | Channels | Modulation Index | | | Calculated | Measured |
| 432 | 1.50 | −5.0 | 1,800 | 0.49 | 142 | 33.2 | 51.6 | 50.2 |
| 432 | 1.50 | −10.0 | 960 | 1.10 | 135 | 30.8 | 51.6 | 50.8 |
| 432 | 1.50 | −10.0 | 432 | 1.50 | 131 | 27.0 | 51.2 | 50.7 |
| 252 | 1.54 | −5.0 | 1,800 | 0.49 | 132 | 19.3 | 51.6 | 51.1 |
| 252 | 1.54 | −5.0 | 432 | 1.50 | 132 | 19.2 | 51.6 | 52.1 |
| 252 | 1.54 | −5.0 | 252 | 1.54 | 132 | 19.2 | 50.7 | 50.7 |
| 24 | 2.55 | −10.0 | 1,800 | 0.49 | 127 | 10.9 | 50.4 | 50.9 |
| 24 | 2.55 | −10.0 | 432 | 1.50 | 129 | 12.2 | 49.4 | 49.9 |
| 24 | 2.55 | −10.0 | 252 | 1.54 | 130 | 14.0 | 49.8 | 49.9 |

Coupling measurements made during these tests showed variation over a 15-dB range. This variation apparently resulted from changes in such environmental factors as wind, rain, fog, humidity, and preferred plane of polarization.

The effects of interference on the terrestrial link at 6 GHz were investigated with a 972-channel carrier on the space link and an 1,800-channel carrier on the terrestrial link. With an e.i.r.p. of only 26 dBW transmitted by the remote terminal (received at the collocated terminal at a level of −68 dBW), the wanted-to-unwanted carrier ratio was 30 dB. The voice channel interference noise power was 1,000 pWp, which agrees closely with predictions and corroborates the contention that 6-GHz interference can be kept to negligible values by using suitably modified transmission parameters. Table 1 summarizes the results.

Video signals were transmitted using INTELSAT IV transmission parameters on both links. For a variety of live and test signals, the test team detected no impairment on either link.

### Conclusions

The test results confirmed the feasibility of frequency reuse at an earth station and its associated radio-relay interconnect link. Further, the tests provided data on coupling between two large aperture antennas, each in the other's near field.

### References

[1] J. B. Potts, "Feasibility of Collocating a Radio-Relay Station with a Sharing Earth Station." *COMSAT Technical Review*, Vol. 2, No. 1, Spring 1972.

[2] International Radio Consultative Committee (C.C.I.R.), "Earth-Station Antenna for the Fixed-Satellite Service," Draft Report 390–2 (Rev. 72), Conclusions of the Interim Meetings of Study Group 4, Part 2, 1972, p. 108.

# *Vitreous oxide antireflection films in high-efficiency solar cells*

A. G. REVESZ

Significant improvement in the short-wavelength response and fill factor of silicon solar cells has been recently achieved by combining a major change in the contact grid geometry with a very shallow junction and a new antireflection film. Under simulated sun illumination, the conversion efficiency of this new (so-called "violet") cell is 13 to 14 percent [1]. In this note, the fundamental considerations underlying the new antireflection film are described and some experimental data are given. Details will be published later.

All attempts to use the antireflection coatings employed in conventional solar cells, i.e., $SiO_x$ and $TiO_x$, with the violet cell have failed because light absorption in these films is high in the very region of the spectrum where the violet cell exhibits improved sensitivity. The use of such coatings resulted in an unacceptably low cell conversion efficiency. Therefore, to implement the basic concept of the violet cell, it became necessary to replace the conventional antireflection coating.

In addition to the obvious requirements, i.e., a refractive index of 2.0 to 2.5 and complete transparency at wavelengths longer than ~300 nm, the film should be prepared in the noncrystalline state, but with a reasonably high degree of short-range order. Noncrystallinity is important since grain boundaries in a polycrystalline film cause light scattering and hence decrease the transparency. Because impurities can easily migrate along grain boundaries, these boundaries are also responsible for unwanted interactions with the junction region; hence, the stability of the antireflection film may be impaired. It is well-known that dielectric films used in capacitors [2] and semiconductor devices (especially $SiO_2$ film in MOS-type devices [3]) must be noncrystalline to achieve the best possible properties. On the other hand, it is desirable for the noncrystalline structure to have a reasonable degree of short-range order; otherwise light absorption is increased (because of unsaturated bonds), while stability and reproducibility are decreased.

These two classes of noncrystalline solids, i.e., solids with and without short-range order, have been designated vitreous and amorphous, re-

*Akos G. Revesz is Manager, Electrophysical Devices Department, Applied Science Division, COMSAT.*

spectively. Striking representatives of the former are $SiO_2$ glass and, especially, noncrystalline $SiO_2$ films grown by thermal or anodic oxidation of silicon [4]. Vacuum-deposited $SiO_x$ (silicon monoxide) is an example of the amorphous class. Silicon dioxide films prepared by vacuum deposition or sputtering are intermediate; that is, they are between $SiO_x$ and thermally or anodically grown $SiO_2$ films. This points out the importance of the preparation method in obtaining well-defined noncrystalline oxide films; a well-controlled growth process, such as that which occurs in anodic or thermal oxidation of the corresponding metal and in some chemical deposition processes, usually results in a higher degree of short-range order than vacuum deposition or sputtering.

Since the refractive index of $SiO_2$ is too low for antireflection films, other dielectrics which can be prepared in the vitreous state have been sought. Consideration has been restricted to oxides, since they generally have the greatest tendency to form noncrystalline structures (because of the "bridging" role played by oxygen) and also since they are the most stable inorganic compounds. Because of their band gap and refractive index values, as well as their use in capacitors [2] and MIS devices [5], $Ta_2O_5$, $Nb_2O_5$, $HfO_2$, and $ZrO_2$ have been selected as candidates. The potential use of $HfO_2$ and $ZrO_2$ has not been explored to as great an extent as that of $Ta_2O_5$ and $Nb_2O_5$. From the viewpoint of final performance as antireflection films, there is practically no difference between $Ta_2O_5$ and $Nb_2O_5$. However, there are some significant differences in the process parameters, and the choice between them is dictated by the particular technology employed.

Various methods (details of which will be published later) can be used for preparing oxide films of suitable quality. Under proper conditions, these methods result in films that are very uniform and exhibit bright interference colors. As an example, electron diffraction and electron microscope investigations of $Ta_2O_5$ films reveal that they are noncrystalline and that they lack any discernible morphological features. However, if some inadvertent contamination has occurred during their preparation, oxide films usually exhibit some structure under the optical microscope, indicating the beginning of crystallization. In addition, in this case the fill factor of the solar cell generally deteriorates. This observation confirms the importance of noncrystallinity.

The refractive index of the oxide films was determined by ellipsometry at a wavelength of 546.1 nm. It was generally observed that the values for each oxide depend on the preparation conditions, ranging for instance from 2.202 to 2.259 for $Ta_2O_5$. Typical values for $Nb_2O_5$ and $HfO_2$ are



2.4 and 2.3, respectively. Both $Ta_2O_5$ and $Nb_2O_5$ films are completely transparent above a wavelength of 300 nm. Similar to that of crystalline $Ta_2O_5$ [6], the infrared spectrum of $Ta_2O_5$ films did not exhibit any absorption below a wavelength of 11 $\mu$m, indicating that no suboxides were present. (Ta suboxides absorb in the 8.5- to 11-$\mu$m range [7].) The density of the $Ta_2O_5$ films was 8.3 $gcm^{-3}$; this value is close to the bulk density of crystalline $Ta_2O_5$, namely, 8.77 $gcm^{-3}$.

These observations indicate that properly prepared $Ta_2O_5$ and $Nb_2O_5$ films are structurally well-defined; that is, the short-range order in these films and the properties determined by it are similar to those of the crystalline structure. Hence, these films can be considered to be vitreous solids. In contrast, sputtered Ta oxide film can be classified as amorphous, since it exhibits marked absorption in the visible spectrum. This observation confirms the importance of short-range order in noncrystalline oxide films and the preference for growth methods rather than vacuum deposition or sputtering.

There are several means of producing the $Ta_2O_5$ or $Nb_2O_5$ antireflection film in a pattern which will not interfere with the current collector grid. The various technological steps employed for this purpose must be compatible with other processes (e.g., contact fabrication) and should not degrade the junction properties. Thus, in contrast with the coating process of conventional solar cells, the application of the vitreous oxide antireflection films is an integral part of the violet cell technology, and the film itself is an integral part of the cell structure. In this respect, the role of this film in the solar cell is somewhat similar to that of vitreous $SiO_2$ films in MOS devices.

Typical characteristics of 4-$cm^2$ violet cells made with the new antireflection film (determined under simulated AM0 sun illumination) are as follows. The open-circuit voltage is ~600 mV, the short-circuit current is ~160 mA, the fill factor is 79 percent, and the conversion efficiency (with respect to total area) is ~13.5 percent. More details are given in [1].

**Acknowledgments**

*J. Lindmayer participated in many discussions and was responsible for the overall solar cell effort. J. Allison collaborated in the experimental work. J. Reynolds cooperated in the ellipsometric measurements. The infrared and visible spectra were determined by T. Kirkendall and C. Maag, respectively. The electron diffraction and electron microscope investigations were performed by K. R. Lawless (University of Virginia). The author is indebted to all.*

**References**

[1] J. Lindmayer and J. Allison, "An Improved Silicon Solar Cell—the Violet Cell," *Conference Record of the Ninth IEEE Photovoltaic Specialists Conference*, Silver Spring, Maryland, pp. 83–84; "The Violet Cell: An Improved Silicon Solar Cell," *COMSAT Technical Review*, Vol. 3, No. 1, Spring 1973, pp. 1–21.

[2] P. J. Harrop and D. S. Campbell, "Selection of Thin Film Capacitor Dielectrics," *Thin Solid Films*, Vol. 2, No. 4, November 1968, pp. 273–292.

[3] A. G. Revesz, "Noncrystalline Silicon Dioxide Films on Silicon: A Review," *Journal of Non-Crystalline Solids*, Vol. 11, 1973, pp. 309–330.

[4] A. G. Revesz, "Noncrystalline Structure and Electronic Conduction of Silicon Dioxide Films," *Physica Status Solidi*, Vol. 24, 1967, pp. 115–126.

[5] K. H. Zaininger and C. C. Wang, "Thin Film Dielectric Materials for Microelectronics," *Proceedings of the IEEE*, Vol. 57, September 1969, pp. 1564–1570.

[6] N. T. McDevitt and W. L. Baun, "Infrared Absorption Study of Metal Oxides in the Low Frequency Region (700–240 cm)," *Spectrochimica Acta*, Vol. 20, 1964, pp. 799–808.

[7] H. Kihara-Morishita, T. Takamura, and T. Takeda, "Infra-red Reflectance Spectra (IRRS) of Anodic Oxide Films Formed on Pure Tantalum," *Thin Solid Films*, Vol. 6, No. 3, September 1970, pp. R29-R31.

# Translations of Abstracts

## Etalement du spectre de puissance de signaux MDP produit par des T.O.P. non linéaires sans mémoire

G. Robinson, O. Shimbo, et R. Fang

**Sommaire**

La présente contribution expose une méthode combinant l'analyse et la simulation pour obtenir le spectre de puissance de signaux MDP multiphases à la sortie du T.O.P. en présence de non-linéarités d'amplitude et de phase. Sur la base de cette analyse, les spectres de signaux MDP à 4-, 8- et 16-états ont été calculés pour diverses rapidités de modulation et différentes valeurs de recul de puissance à l'entrée du satellite. Le brouillage causé dans le canal de transmission adjacent par l'étalement du spectre a également été évalué.

Les spectres calculés ont été comparés avec les spectres mesurés de signaux MDP à 4-états pour une rapidité de modulation et différentes valeurs du recul d'entrée. Dans chaque cas, on trouve une bonne concordance entre les spectres calculés et les spectres mesurés. On constate que l'étalement du spectre de puissance est essentiellement dû à la non-linéarité d'amplitude, plutôt qu'à la non-linéarité de phase de l'amplificateur à T.O.P. On démontre en outre que, à une rapidité de modulation donnée, l'étalement du spectre de puissance de signaux à 8- et à 16-états ne diffère pas de manière significative de celui d'un signal MDP à 4-états.

## Application de l'AMRT à la série des satellites INTELSAT IV

W. G. Schmidt

**Sommaire**

Une série exhaustive de programmes expérimentaux portant sur le système d'accès multiple par répartition dans le temps (AMRT) a permis de démontrer la faisabilité technique de ce mode de transmission, de même que les avantages qui peuvent en découler. La présente communication décrit les éléments d'un système prototype AMRT conçu aux fins d'utilisation avec les satellites INTELSAT IV et aux fins d'une éventuelle introduction dans les réseaux opérationnels.

On analyse la jonction terrestre avec le terminal AMRT, ainsi que l'équipement d'émission et de réception, notamment le modem MDP quadriphase