Get Inside Circular Waveguide with W1GHZ!

QEX

David Sumner, K1ZZ
Publisher

Doug Smith, KF6DX
Editor

Robert Schetgen, KU7G
Managing Editor

Lori Weinberg
Assistant Editor

Peter Bertini, K1ZJH
Zack Lau, W1VT
Contributing Editors

Production Department

Mark J. Wilson, K1RO
Publications Manager

Michelle Bloom, WB1ENT
Production Supervisor

Sue Fagan
Graphic Design Supervisor

David Pingree, N1NAS
Technical Illustrator

Joe Shea
Production Assistant

Advertising Information Contact:

John Bee, N1GNV, *Advertising Manager*
  860-594-0207 direct
  860-594-0200 ARRL
  860-594-0259 fax

Circulation Department

Debra Jahnke, *Manager*
Kathy Capodicasa, N1GZO, *Deputy Manager*
Cathy Stepina, *QEX Circulation*

Offices

225 Main St, Newington, CT 06111-1494 USA
Telephone: 860-594-0200
Telex: 650215-5052 MCI
Fax: 860-594-0259 (24 hour direct line)
e-mail: **qex@arrl.org**

Subscription rate for 6 issues:

  In the US: ARRL Member $22,
  nonmember $34;

  US, Canada and Mexico by First Class Mail:
  ARRL member $35, nonmember $47;

  Elsewhere by Surface Mail (4-8 week delivery):
  ARRL member $27,
  nonmember $39;

  Elsewhere by Airmail: ARRL member $55,
  nonmember $67.

Members are asked to include their membership control number or a label from their *QST* wrapper when applying.

About the Cover

W1GHZ built this slotted line to gather data for his article on **p 37**.

Features

Columns

Jan/Feb 2001 *QEX* Advertising Index

Empirically Speaking

The ARRL Technology Task Force (TTF) will submit its final report to the League's Board of Directors at the January 2001 meeting of the Board. The TTF's report contains recommendations for action based on input from the Technology Working Group (TWG), which filed its comments with the TTF in July 2000. The TTF's recommendations emphasize digital voice modes and software-defined radios. ARRL President Jim Haynie, W5JBP, has initiated a committee to organize a plan of attack for digital voice.

One possible goal of the Digital Voice Committee would be to evaluate and recommend standards for modulation, coding and data formats. One thing stands out about digital voice: There is a marked trade-off between data rate and voice quality. I guess it would be nice if certain aspects of adaptive systems theory could find their way into those recommendations, so that a single system might be usable across the amateur spectrum at a variety of data rates, bandwidths and quality levels. The protocol should be flexible enough to handle future improvements in technology.

Anyone interested in getting or giving more information may contact me at the e-mail address below, or by "Pony Express" at Newington. Many thanks to those of you who sent the TTF and TWG your ideas and suggestions: They made a very significant impact.

In other news: We regret to report that Douglas Page is unable to continue as Contributing Editor. The idea of announcing recently released communications products and other brief items of interest to the technical community still seems appealing. We will be looking, once again, for a way to add that to the list of many small improvements to *QEX/Communications Quarterly*.

Another thing we'd like to improve on is our publishing schedule. Lead-time between acceptance and publication of an article may reach 12 months or so. We'll be trying to reduce that over the coming year. To authors: Your patience is greatly appreciated. The good news is that we kno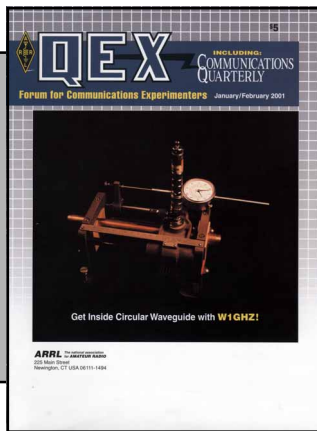w what is coming farther in advance; but with just 64 pages, it is still not easy to predict the content of the next few issues until the last one has been put to bed. We try to run articles in rough-ly the order they were accepted. Other factors often come into play, though, such as making it all fit together, accommodating last-minute changes and so forth. Bob Schetgen, KU7G, and the crew do a heroic job of making those things happen.

In This Issue

Jim Buckwalter, KF6SWC, and his Caltech friends deliver something we've been wishing for: a keyed, switch-mode supply for class-E amplifiers. As good "Tech-ers," they include valuable detail on magnetic design—an aspect of switchers that can mean the difference between success and spectacular disaster. Special thanks to Dave Rutledge, KN6EK, for helping us pull it together. Nat Sokal, WA1HQC, contributes an update on class-E amplifiers themselves. Nat offers improved characterization of their workings and some tips for getting the most out of them. Thanks to William Hagen and IEEE for their gracious assistance.

On the subject of transmission lines: We've seen a lot of wave-reflection discussions recently, and many readers have wondered how to reconcile those with simpler, more-familiar ways of analysis. Steve Best, VE9SRB, donates another thing we've been looking for: a proof of the equivalence of wave mechanics and the well-known transmission-line equation. Aptitude with complex algebra is assumed. This piece is for those who want to get past the hand waving and gain insight into what really makes it all tick. Paul Wade, W1GHZ, looks at circular waveguide. As it turns out, this is a convenient alternative to common, rectangular forms. It may give new meaning to the term "plumber's delight."

We have another article from Robert Brown, NM7M, on LF and MF propagation. He investigates some more interesting, not-so-obvious effects. In "Technical Notes," Peter Bertini, K1ZJH, brings us some information from Steve Hageman about how to build a data acquisition system around your PC. In *RF*, Zack Lau, W1VT, presents high-isolation amplifiers, achieved by means of directional couplers in the feedback networks.—*Happy New Year! de Doug Smith, KF6DX,* **kf6dx@arrl.org**

Wave Mechanics of Transmission Lines, Part 1: Equivalence of Wave Reflection Analysis and the Transmission-Line Equation

Let's see how to derive simple, elegant equations that describe what occurs in transmission lines. The proof starts with an analysis of waves traveling in opposite directions.

By Dr. Steven R. Best, VE9SRB

The general concepts of steady-state transmission-line theory and impedance matching are well documented and well known. However, misunderstandings usually arise when attempting to relate steady-state conditions to the wave reflections occurring within a transmission-line system. To understand the steady-state conditions as they relate to wave reflections, it is necessary to perform a detailed wave-reflection analysis. A wave-reflection analysis is also necessary in order to understand how a steady-state impedance match occurs with any given matching device.

To provide a detailed understanding of wave-reflection mechanics in transmission-line systems, three articles will address the following general issues:

• The physical mechanism that creates wave reflections at either end of a transmission line

48 Perimeter Rd
Manchester, NH 03103
srbest@att.net

• How the steady-state forward and reflected voltage, current and power levels are developed at a transmission-line input
• The relationships between forward, reflected and delivered power in a transmission line
• What happens to reflected voltage, current and power arriving at the output terminals of a transmitter, matching device or match point
• The physical mechanisms that cause an impedance match to occur in various matching devices

This article, Part 1, will provide a general overview of basic transmission-line concepts and a discussion of the relationship between forward, reflected and delivered power in a transmission line. Part 2 will provide a detailed discussion of what happens to the reflected voltage, current and power when it arrives at the output terminals of a transmitter or impedance matching device. Part 3 will provide a treatment of the impedance matching process and a description of the mechanism that causes the impedance match to occur in the steady state.

Comments on the use of Thevenin-Equivalent Circuits in Transmission-Line Analysis

In this and the two subsequent articles, Thevenin-equivalent circuits are used in the analysis of the wave reflections that occur within a transmission-line system. A Thevenin-equivalent circuit consists of a constant-voltage source, V_S, and an equivalent series output impedance, Z_S. This equivalent circuit is not intended to be an accurate physical representation of the output stages of a transmitter. However, the Thevenin-equivalent circuit can be used to illustrate concepts associated with the wave-reflection mechanics that occur within a transmission-line system and specifically at the output terminals of a transmitter. Once these concepts are understood with a Thevenin-equivalent circuit, they can be validly applied to any matching device, practical transmitter or power amplifier, regardless of class or internal configuration.

In the discussions presented in this article, the equivalent output impedance Z_S is intended to represent the impedance seen looking rearward into the output terminals of the transmitter. Physically, this would be the impedance that would be measured by connecting a network analyzer or impedance bridge to the transmitter's output terminals. From the perspective of the rearward-traveling waves in the transmission line, this is the only transmitter impedance that has any physical significance. In any practical transmitter, the impedance seen looking rearward into its output terminals will be a direct function of the configuration of the transmitter's output stages. In a typical solid-state transmitter, these output stages will generally include the power amplifier; a transformer; an output filter (which may be a simple π network); and an SWR bridge or automatic tuning circuit.

Basic Transmission Line Concepts

In any transmission-line circuit, the system antenna is defined as matched to the system transmission line when no reflections exist at the connection point between the transmission line and the antenna. Any antenna having a feed-point impedance ($Z_A = R_A + jX_A$) different from the characteristic impedance ($Z_0 = R_0 + jX_0$) of the transmission line to which it is connected creates an impedance mismatch at the connection point.

When a transmitter is first energized, it develops an initial forward-driving wave (voltage, current and power) at the input to the transmission line. This forward-driving wave travels toward the antenna undergoing attenuation and relative phase delay consistent with the propagation characteristics of the transmission line. Arriving at the antenna, the forward-traveling wave "sees" the antenna impedance as a terminating load. If the antenna impedance is not equal to the characteristic impedance of the transmission line, a rearward-traveling reflected wave is created.

The reflected wave created at the antenna travels rearward toward the transmitter, also undergoing an attenuation and phase delay according to the characteristics of the line. Arriving at the transmitter output, the reflected wave sees the transmitter's effective output impedance as a terminating load. If the transmitter's output impedance is not equal to the characteristic impedance of the transmission line, a forward-traveling, re-reflected wave is created. Assuming that the transmitter is still energized, the re-reflected wave is combined with the forward-driving wave delivered by the transmitter and the process of reflection and re-reflection continues until steady-state conditions are reached.

At either end of a transmission line, the level of reflection or re-reflection is a function of the physical impedance terminating the transmission line. This is true for the forward-traveling wave arriving at the antenna and the rearward-traveling wave arriving at the transmitter output. All of the transmission line's steady-state conditions can be determined from the transmitter's forward-driving voltage, the transmission line's characteristic impedance and propagation factors, the antenna impedance and the transmitter's effective output impedance.

To understand how steady-state conditions relate to these physical properties and the multiple wave-reflection process, we will consider the transmission-line circuit block diagram as presented in Fig 1. The transmitter is represented by a Thevenin-equivalent circuit as a constant-voltage source, V_S, and an effective output impedance Z_S. The antenna is represented as the impedance Z_A. The transmission line connecting the antenna to the transmitter has characteristic impedance Z_0 and is of a physical length L.

The forward-driving voltage delivered by the transmitter to the transmission line is defined as V_I, which can be determined from the transmission line's initial state. The physical impedance that the transmitter sees at the input to the transmission line is the transmission line's characteristic impedance Z_0. Therefore, the forward-driving voltage delivered to the transmission line by the transmitter can be determined through simple circuit theory as follows:

$$V_I = V_S \left(\frac{Z_0}{Z_0 + Z_S} \right) \tag{Eq 1}$$

Two things occur as a wave travels along a transmission line away from its point of origin. First, the wave is attenuated by losses in the transmission line, and second, the wave undergoes a relative phase shift (or delay) as a function of the electrical distance traveled. The level of wave attenuation can be expressed as $e^{-\alpha x}$, where α is the transmission-line attenuation factor expressed in nepers per meter and x is the distance traveled away from the wave's reference point of origin. Nepers per meter can be converted from attenuation measured in decibels/100 feet by multiplying the attenuation by 3.7772×10^{-3}. The level of wave phase delay can be expressed as $e^{-j\beta x}$, where β is simply $2\pi/\lambda$. The electrical wavelength in the transmission line, λ, should be expressed in meters. The level of wave attenuation and phase delay can be combined into a single expression, $e^{-\gamma x}$,
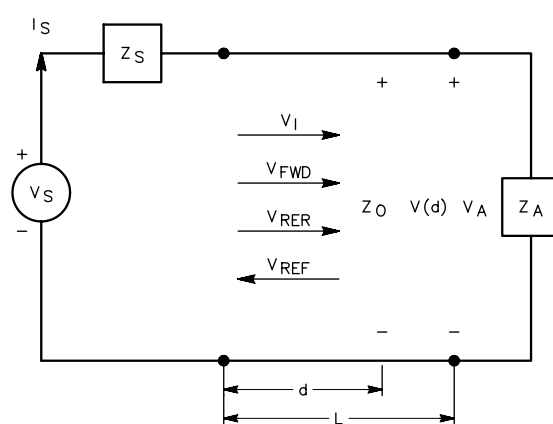
Fig 1—Basic transmission-line circuit block diagram.

where $\gamma = \alpha + j\beta$ is the transmission line's propagation factor. The complex-wave voltage at any point along the length of the transmission line, x, is given by $V e^{-\gamma x}$, where V is the complex-wave voltage determined at the wave's point of origin, $x = 0$. This is true for the forward-traveling wave originating at the input of the transmission line and the rearward-traveling wave originating at the antenna.

Since V_I is the forward-driving voltage at the transmission-line input, the initial forward voltage arriving at the antenna is equal to $V_I e^{-\gamma L}$, where L is total length of the transmission line as illustrated in Fig 1. Assuming the antenna feedpoint impedance Z_A is not equal to the characteristic impedance of the transmission line, Z_0, a rearward-traveling reflected voltage is created at the antenna. The level of voltage reflected at the antenna is a direct function of the transmission line's characteristic impedance and the feedpoint impedance of the antenna. The parameter that is used to determine the level of reflected voltage at the antenna is the antenna reflection coefficient, ρ_A, which can be determined from the following equation:

$$\rho_A = \frac{\dfrac{Z_A}{Z_0} - 1}{\dfrac{Z_A}{Z_0} + 1} \qquad \text{(Eq 2)}$$

The level of reflected voltage developed at the antenna is equal to the incident (forward) voltage arriving at the antenna multiplied by the antenna's reflection coefficient. The total voltage developed at the antenna is equal to the sum of the incident voltage plus the reflected voltage. The antenna's SWR can also be found from the antenna reflection coefficient, ρ_A, using the following equation:

$$SWR = \frac{1 + |\rho_A|}{1 - |\rho_A|} \qquad \text{(Eq 3)}$$

If the magnitude of the antenna reflection coefficient is zero (Z_A is equal to Z_0), no reflected voltage is created at the antenna. In this case, all the total forward power arriving at the antenna is delivered to the antenna. If the mag-

nitude of the reflection coefficient is one (Z_A is an open circuit, a short circuit or purely reactive), the magnitude of the reflected voltage is equal to the magnitude of the incident voltage, and no power is delivered to the antenna. If the magnitude of the reflection coefficient is between zero and one, some power is delivered to the antenna and some power is reflected. This is true for both forward waves arriving at the antenna and reflected waves arriving at the transmitter output. One interesting point occurs when Z_0 is complex. When Z_0 has a reactive component, it is possible to deliver real power to an antenna when the magnitude of the antenna reflection coefficient is one.

The rearward-traveling reflected voltage created at the antenna is equal to the incident forward voltage arriving at the antenna multiplied by the antenna's reflection coefficient ρ_A. The rearward-traveling reflected voltage created at the antenna is therefore equal to $\rho_A V_I e^{-\gamma L}$. When the reflected voltage from the antenna arrives back at the input to the transmission line, it will have traveled the total length of the transmission line, L, and it will have undergone an attenuation and relative phase delay consistent with the propagation factor of the transmission line, γ. The reflected voltage arriving back at the input to the transmission line is therefore equal to:

$$V_{REF} = \rho_A V_I e^{-\gamma L} e^{-\gamma L} = \rho_A V_I e^{-2\gamma L} \qquad \text{(Eq 4)}$$

Arriving back at the input to the transmission line, the reflected voltage sees a terminating impedance equal to the physical impedance seen looking rearward into the transmitter's output terminals. In the case of the Thevenin-equivalent circuit, this impedance is equal to Z_S. If Z_S is not equal to Z_0, a forward-traveling, "re-reflected" voltage, V_{RER}, is developed at the transmission-line input, consistent with the reflection coefficient of the transmitter's output impedance, ρ_S, such that:

$$V_{RER} = \rho_S V_{REF} = \rho_S \rho_A V_I e^{-2\gamma L} \qquad \text{(Eq 5)}$$

In all transmission line systems, the level of voltage or wave reflection at either end of the transmission line is determined from the reflection coefficient of the impedance

$V_I \rightarrow$

$\rho_A V_I e^{-2\gamma L} \leftarrow$

$\rho_S \rho_A V_I e^{-2\gamma L} \rightarrow$

$\rho_S \rho_A^2 V_I e^{-4\gamma L} \leftarrow$

$\rho_S^2 \rho_A^2 V_I e^{-4\gamma L} \rightarrow$

$\rho_S^2 \rho_A^3 V_I e^{-6\gamma L} \leftarrow$

$\rho_S^3 \rho_A^3 V_I e^{-6\gamma L} \rightarrow$

$\rho_S^3 \rho_A^4 V_I e^{-8\gamma L} \leftarrow$

$\rho_S^4 \rho_A^4 V_I e^{-8\gamma L} \rightarrow$

$V_I e^{-\gamma L} \rightarrow$

$\rho_A V_I e^{-\gamma L} \leftarrow$

$\rho_S \rho_A V_I e^{-3\gamma L} \rightarrow$

$\rho_S \rho_A^2 V_I e^{-3\gamma L} \leftarrow$

$\rho_S^2 \rho_A^2 V_I e^{-5\gamma L} \rightarrow$

$\rho_S^2 \rho_A^3 V_I e^{-5\gamma L} \leftarrow$

$\rho_S^3 \rho_A^3 V_I e^{-7\gamma L} \rightarrow$

$\rho_S^3 \rho_A^4 V_I e^{-7\gamma L} \leftarrow$

Fig 2—Multiple-reflection and re-reflection chart.

terminating the transmission line. The "re-reflected" voltage, V_{RER}, adds to the initial forward-traveling voltage, V_I, and the process of reflection and "re-reflection" continues until equilibrium is reached.

To determine the steady-state levels of forward and reflected voltages developed at the input to the transmission line as a function of the transmitter's forward-driving voltage V_I, it is necessary to examine the multiple reflection and re-reflection process in detail. In the steady state, the total forward-traveling voltage developed at the input to the transmission line, V_{FWD}, is equal to V_I plus an infinite number of forward-traveling, re-reflected voltages developed through the processes described by Eqs 4 and 5. A portion of this multiple reflection and re-reflection process is illustrated in Fig 2.

Examining Fig 2 in detail, the total steady-state forward-traveling voltage developed at the input to the transmission line, V_{FWD}, is determined from V_I and the steady-state summation of the re-reflected voltages developed there, as follows:

$$V_{FWD} = V_I + V_I\left(\rho_S\rho_A e^{-2\gamma L}\right) + V_I\left(\rho_S^2\rho_A^2 e^{-4\gamma L}\right) + V_I\left(\rho_S^3\rho_A^3 e^{-6\gamma L}\right) + \ldots$$
$$= V_I\left[1 + \left(\rho_S\rho_A e^{-2\gamma L}\right) + \left(\rho_S\rho_A e^{-2\gamma L}\right)^2 + \left(\rho_S\rho_A e^{-2\gamma L}\right)^3 + \ldots\right]$$
(Eq 6)

which is an equation of the form:

$$V_{FWD} = V_I\left(1 + A + A^2 + A^3 + A^4 + A^5 + \ldots\right)$$
(Eq 7)

where $A = \rho_S \rho_A e^{-2\gamma L}$. The expression in parentheses is an infinite geometric series, which can be written in closed form as $1/(1-A)$. Therefore, the total steady-state, forward-traveling voltage developed at the transmission-line input, V_{FWD}, is given by:

$$V_{FWD} = V_I\left(\frac{1}{1 - \rho_S\rho_A e^{-2\gamma L}}\right)$$
(Eq 8)

From the form of Eq 4, the total steady-state reflected voltage arriving at the transmission-line input, V_{REF}, is determined from the following equation:

$$V_{REF} = \rho_A V_{FWD} e^{-2\gamma L} = V_I\left(\frac{\rho_A e^{-2\gamma L}}{1 - \rho_S\rho_A e^{-2\gamma L}}\right)$$
(Eq 9)

From the form of Eq 5, the total steady-state, re-reflected voltage developed at the transmission-line input, V_{RER}, is determined from the following equation:

$$V_{RER} = \rho_S V_{REF} = V_I\left(\frac{\rho_S\rho_A e^{-2\gamma L}}{1 - \rho_S\rho_A e^{-2\gamma L}}\right)$$
(Eq 10)

The total steady-state, forward-traveling voltage developed at the input to the transmission line, V_{FWD}, can also be determined from the following relationship: $V_{FWD} = V_I + V_{RER}$.

At any point along the length of the transmission line, the total steady-state voltage is the sum of the forward-traveling voltage and the rearward-traveling voltage. The total steady-state voltage at any point d along the length of the transmission line can be found using the following generalized transmission line equation:

$$V_{TOTAL}(d) = V(d) = V_{FWD}e^{-\gamma d} + V_{REF}e^{+\gamma d}$$
(Eq 11)

Since the forward- and rearward-traveling waves only see the characteristic impedance, Z_0, as they travel in the transmission line, the following relationships are always valid:

$$\frac{V_{FWD}}{I_{FWD}} = Z_0$$
(Eq 12)

and

$$\frac{V_{REF}}{I_{REF}} = Z_0$$
(Eq 13)

Therefore, the total steady-state current at any point, d, along the length of the transmission line is given by:

$$I_{TOTAL}(d) = I(d) = \frac{V_{FWD}e^{-\gamma d}}{Z_0} - \frac{V_{REF}e^{+\gamma d}}{Z_0}$$
(Eq 14)

The steady-state line impedance at any point along the length of the transmission line is equal to the ratio of total voltage to total current at that point: $V(d)/I(d)$. At the input to the transmission line, $d = 0$; the steady-state ratio of $V(0)/I(0)$ simplifies to the following:

$$\frac{V(0)}{I(0)} = Z_{IN} = Z_0\left(\frac{V_{FWD} + V_{REF}}{V_{FWD} - V_{REF}}\right)$$
(Eq 15)

Substituting Eq 9 for V_{REF} yields:

$$Z_{IN} = Z_0\left(\frac{V_{FWD} + \rho_A e^{-2\gamma L}V_{FWD}}{V_{FWD} - \rho_A e^{-2\gamma L}V_{FWD}}\right)$$
$$= Z_0\left(\frac{1 + \rho_A e^{-2\gamma L}}{1 - \rho_A e^{-2\gamma L}}\right)$$
(Eq 16)

which is a convenient equation for Z_{IN} given as a function of ρ_A. Substituting Eq 2 for ρ_A and multiplying the top and bottom expressions by Z_A/Z_0+1 yields:

$$Z_{IN} = Z_0\left[\frac{\left(\dfrac{Z_A}{Z_0}+1\right) + \left(\dfrac{Z_A}{Z_0}-1\right)e^{-2\gamma L}}{\left(\dfrac{Z_A}{Z_0}+1\right) - \left(\dfrac{Z_A}{Z_0}-1\right)e^{-2\gamma L}}\right]$$
(Eq 17)

Rearranging terms yields:

$$Z_{IN} = Z_0\left[\frac{\dfrac{Z_A}{Z_0}\left(1 + e^{-2\gamma L}\right) + \left(1 - e^{-2\gamma L}\right)}{\left(1 + e^{-2\gamma L}\right) + \dfrac{Z_A}{Z_0}\left(1 - e^{-2\gamma L}\right)}\right]$$
(Eq 18)

Multiplying the top and bottom expressions by $e^{\gamma L}$ yields:

$$Z_{IN} = Z_0\left[\frac{\dfrac{Z_A}{Z_0}\left(e^{\gamma L} + e^{-\gamma L}\right) + \left(e^{\gamma L} - e^{-\gamma L}\right)}{\left(e^{\gamma L} + e^{-\gamma L}\right) + \dfrac{Z_A}{Z_0}\left(e^{\gamma L} - e^{-\gamma L}\right)}\right]$$
(Eq 19)

Substituting the appropriate hyperbolic functions yields the following final equation for transmission-line input impedance:

$$Z_{IN} = Z_0\left[\frac{\dfrac{Z_A}{Z_0}\cosh(\gamma L) + \sinh(\gamma L)}{\cosh(\gamma L) + \dfrac{Z_A}{Z_0}\sinh(\gamma L)}\right]$$
$$= Z_0\left[\frac{\dfrac{Z_A}{Z_0} + \tanh(\gamma L)}{1 + \dfrac{Z_A}{Z_0}\tanh(\gamma L)}\right]$$
(Eq 20)

This is the familiar transmission-line equation.

The effective steady-state line impedance at any point along the length of the transmission line is not a physical load impedance. It is simply the ratio of total voltage to total current at that point. In the special case where the forward- and rearward-traveling voltages at any one point are of equal amplitude and 180° out of phase, the total voltage will be zero and the steady-state line impedance at that point will be a short circuit. In the special case where the forward- and rearward-traveling currents at any one point are of equal amplitude and exactly in phase, the total current will be zero, and the steady-state line impedance at that point will be infinite. (Note that the forward- and rearward-currents are traveling in opposite directions so that when they are identically equal in amplitude and phase, the net current is zero.) It is important to note that the line impedance in a transmission line is not a mechanism for wave reflection or re-reflection. As the forward- and rearward-waves travel in the line, they see only the transmission line's characteristic impedance and any wave reflection or re-reflection will occur only at the ends of the transmission line (assuming there are no impedance discontinuities within the transmission line).

The forward- and rearward-traveling powers at any point, d, along the length of the transmission line are a direct function of the forward- and rearward-traveling voltage and current. If it is assumed that Z_0 is resistive and has no reactive component, the forward- and rearward-traveling power at any point d can be found as follows:

$$P_{\text{FWD}} = \frac{\left| V_{\text{FWD}} e^{-\gamma d} \right|^2}{Z_0} \qquad \text{(Eq 21)}$$

and

$$P_{\text{REF}} = \frac{\left| V_{\text{REF}} e^{+\gamma d} \right|^2}{Z_0} \qquad \text{(Eq 22)}$$

These levels of power would be consistent with those measured using a typical wattmeter.

Finally, the total steady-state voltage developed at the antenna, V_A, is the sum of the total forward and reflected voltages developed there. The voltage V_A is given by:

$$V_A = V_{\text{FWD}} \, e^{-\gamma L} \left(1 + \rho_A \right) = V_I \left[\frac{\left(1 + \rho_A \right) e^{-\gamma L}}{1 - \rho_S \rho_A e^{-2\gamma L}} \right] \qquad \text{(Eq 23)}$$

Summary of the Discussion Thus Far

The four most important aspects of the preceding discussion on basic transmission-line concepts are the following:

1. The forward-driving voltage, current and power delivered by the transmitter to the input of any transmission line are a direct function of the transmission-line characteristic impedance Z_0.
2. As the forward and rearward waves travel in the transmission line, the only physical impedance they see is Z_0. Additionally, the forward- and rearward-traveling waves always travel as a function of $e^{-\gamma x}$, where x is the distance traveled from the wave's point of origin.
3. The levels of voltage, current and power reflection at the ends of a transmission line are direct functions of the physical impedance terminating the transmission line. Reflections developed at the antenna end of the transmission line are direct functions of the antenna impedance. Re-reflections developed at the input to a transmission line are direct functions of the physical terminating impedance seen by the rearward-traveling wave when it ar-

rives at the input to the transmission line. This impedance will be the effective output impedance of the transmitter or matching device connected at the transmission-line input. This is a physical impedance that is measurable by a network analyzer or impedance bridge.
4. The relationships described above for forward- and rearward-traveling voltage, current and power are valid in all cases, independent of match or mismatch in a transmission-line system.

The Relationship between Forward, Reflected and Delivered Power in Equilibrium

In this section, the relationships between steady-state forward, reflected and effective net (delivered) power in the transmission line will be presented. With knowledge of these power levels, the power delivered to the antenna and the power dissipated in the transmission line are easily determined. The relationships for power delivery detailed in this section assume that Z_0 is purely resistive and has no reactive component. These relationships are valid for any transmission line, independent of whether the transmission line input is connected to a transmitter, tuner or other matching device. In the case of a tuner or other matching device, the relationships presented here are valid independent of the existence of an impedance match at the tuner or other matching device input.

Consider the block diagram of Fig 3. The total, steady-state forward power developed at the input to the transmission line is given by P_{FWD}, and the steady-state reflected power is given by P_{REF}. The net power delivered to the transmission line input is given by P_{DEL}.

If the total attenuation of the transmission line over its entire length when connected to a matched load is given by α (dB), then the total forward power arriving at the antenna is:

$$P_f = k P_{\text{FWD}} \qquad \text{(Eq 24)}$$

$$k = 10^{\left(-\frac{\alpha}{10} \right)} \qquad \text{(Eq 25)}$$

The transmission-line attenuation is in decibels, α, is expressed as a positive number. The total steady-state power reflected at the antenna is given by:

$$P_r = \left| \rho_A \right|^2 P_f \qquad \text{(Eq 26)}$$

and the total steady-state reflected power arriving at the input to the transmission line, P_{REF}, is:

$$P_{\text{REF}} = k P_r = k \left| \rho_A \right|^2 P_f = k^2 \left| \rho_A \right|^2 P_{\text{FWD}} \qquad \text{(Eq 27)}$$

The total amount of power dissipated in the transmission line is:

$$P_{\text{LOSS}} = P_{\text{LF}} + P_{\text{LR}} \qquad \text{(Eq 28)}$$

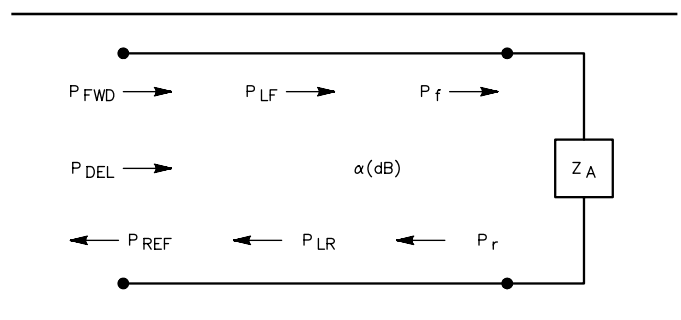where P_{LF} is the power loss occurring the forward direction

Fig 3—Transmission-line block diagram of power delivery.

and P_{LR} is the power loss occurring in the rearward direction. Examining Fig 3, it is evident that the loss in the forward direction is P_{FWD}–P_f and the loss in the rearward direction is P_r–P_{REF}. Substituting the appropriate relationships into Eq 28, the total power dissipation occurring in the transmission line, expressed as a function of the antenna reflection coefficient and the total forward power at the input to the transmission line, is given by:

$$P_{LOSS} = P_{FWD}(1-k)\left(1+k|\rho_A|^2\right) \quad \text{(Eq 29)}$$

Using P_{FWD} as the reference power for calculation purposes, the effective net or total attenuation of the transmission line is:

$$\alpha_F(\text{dB}) = -10\log\left[1-(1-k)\left(1+k|\rho_A|^2\right)\right] \quad \text{(Eq 30)}$$

Using Eq 30, the net attenuation of the transmission line calculates to be a positive number. Eqs 29 and 30 may be used to determine the total level of power dissipation occurring within a transmission line, where the remaining power is equal to the sum of the reflected power at the transmission-line input and the power delivered to the antenna. When considering the total forward power at the transmission-line input, the total power delivered to the antenna can be determined simply from:

$$P_A = P_{FWD}k\left(1-|\rho_A|^2\right) \quad \text{(Eq 31)}$$

In many instances, it is preferable to use the effective net power delivered to the transmission line as the reference for power-delivery calculations. The effective net power delivered to the transmission line, P_{DEL}, is simply the difference between the total forward and reflected powers developed at the input to the transmission line. P_{DEL} is:

$$P_{DEL} = P_{FWD} - P_{REF} = P_{FWD}\left(1-k^2|\rho_A|^2\right) \quad \text{(Eq 32)}$$

The significance of determining P_{DEL} is that it represents the total steady-state power delivered to the transmission-line system. P_{DEL} is equal to the total power delivered to the antenna plus the total power dissipated in the transmission line. Using P_{DEL} as the reference for power-delivery calculations, the total power dissipation occurring in the transmission line is found:

$$P_{LOSS} = P_{DEL}\left[1-\frac{k\left(|\rho_A|^2-1\right)}{k^2|\rho_A|^2-1}\right] \quad \text{(Eq 33)}$$

In this case, the net transmission-line attenuation is given by:

$$\alpha_D(\text{dB}) = -10\log\left[\frac{k\left(|\rho_A|^2-1\right)}{k^2|\rho_A|^2-1}\right] \quad \text{(Eq 34)}$$

and the total power delivered to the antenna is given by:

$$P_A = P_{DEL}\left[\frac{k\left(|\rho_A|^2-1\right)}{k^2|\rho_A|^2-1}\right] \quad \text{(Eq 35)}$$

In using either the total forward power or the effective net power delivered to the input of the line, the total power delivered to the antenna is always a function of the transmission-line attenuation and the antenna's reflection coefficient, ρ_A.

Using Eq 32, the total steady-state forward power at the transmission-line input can be determined from the steady-state net power delivered to the transmission line as follows:

$$P_{FWD} = P_{DEL}\left(\frac{1}{1-k^2|\rho_A|^2}\right) \quad \text{(Eq 36)}$$

It is important to note that this relationship between forward and delivered power is valid in all transmission-line systems. For example, if a tuner or matching device is used between a transmitter and an antenna, the relationship established by Eq 36 is valid in the transmission line connecting the tuner and the antenna independent of whether or not an impedance match exists at the tuner input.

At this point, it is also important to explain the differences between Eqs 30 and 34, both of which can be used to determine the effective attenuation of the transmission line. Eq 30 may be used to calculate the total loss in the transmission line relative to the total forward power at the transmission-line input. The remaining power is the sum of the power delivered to the antenna plus the reflected power at the transmission-line input. Eq 34 may be used to calculate the total loss in the transmission line relative to the effective net power delivered to the transmission line. The remaining power is the power delivered to the antenna.

Summary

This article has presented a detailed description of basic transmission-line concepts and the principles of wave mechanics that occur when a mismatched antenna is connected to the system transmission line. The discussion and equations presented in this article establish the foundation for understanding the nature of wave reflection behavior in transmission-line systems.

The relationships between forward- and rearward-traveling voltage, current and power were developed and sufficient equations were presented relating all of the transmission-line system's physical parameters. All of the transmission-line system's steady-state conditions can be determined from the transmitter's forward-driving voltage, the transmission line's characteristic impedance and propagation factor, the antenna impedance and the transmitter's effective output impedance.

The relationships between forward, reflected and effective net power delivered to the transmission line were presented. Using the equations presented in this article, the power dissipation in the transmission line and the power delivery to an antenna can easily be determined.

Acknowledgement

References

R. A. Chipman. Schaum's Outline Series: *Theory and Problems of Transmission Lines* (McGraw Hill, 1968).

W. C. Johnson, *Transmission Lines and Networks* (McGraw Hill, 1950).

Class-E RF Power Amplifiers

Come learn about this highly efficient and widespread class of amplifiers. Here are principles of operation, improved design equations, optimization principles and experimental results.

By Nathan O. Sokal, WA1HQC
of Design Automation, Inc
ARRL Technical Advisor

This article is based on "Class-E High-Efficiency Power Amplifiers, from HF to Microwave," Proceedings of the IEEE International Microwave Symposium*, June 1998, Baltimore; and "Class- E Switching-Mode High-Efficiency Tuned RF Microwave Power Amplifier: Improved Design Equations,"* Proceedings of the IEEE International Microwave Symposium*, June 2000, Boston; both by Nat Sokal, © IEEE 1998, 2000.—Ed.*

Class-E power amplifiers (See References 1, 2, 3, 4, 5, 6) achieve significantly higher efficiency than conventional Class-B or –C amplifiers. In

4 Tyler Rd
Lexington, MA 02420-2404
NathanSokal@compuserve.com

Class-E, the transistor operates as an on/off switch and the load network *shapes the voltage and current waveforms* to prevent *simultaneous* high voltage and high current in the transistor; that minimizes power dissipation, especially during the switching transitions. In the published low-order Class-E circuit, a transistor performs well at frequencies up to about 70% of its frequency of good Class-B operation (an unpublished higher-order Class-E circuit operates well up to about twice that frequency). This paper covers circuit operation, improved-accuracy explicit design equations for the published low-order Class-E circuit, optimization principles and experimental results. Previously published analytically derived design equations did not include the

dependence of output power (P) on load-network loaded Q (Q_L). As a result, the output power is 38% to 10% less than expected, for Q_L values in the usual range of 1.8 to 5. This paper includes an accurate new equation for P that includes the effect of Q_L.

What Can Class-E Do for Me?

Typically, Class-E amplifiers (see References 1, 2, 3, 4, 5, 6) can operate with power losses smaller by a factor of about 2.3, as compared with conventional Class-B or -C amplifiers using the same transistor at the same frequency and output power. For example, a Class-B or -C power stage operating at 65% collector or drain efficiency (losses = 35% of input power) would have an efficiency of about 85% (losses = 15% of input

power) if changed to Class E (35%/15% = 2.3). Class-E amplifiers can be designed for narrow-band operation or for fixed-tuned operation over frequency bands as wide as 1.8:1, such as 225-400 MHz. (If harmonic outputs must be well below the carrier power, only Class-A or push-pull Class-AB amplifiers can operate over a band wider than about 1.8:1 with only one fixed-tuned harmonic-suppression filter.) Harmonic output of Class-E amplifiers is similar to that of Class-B amplifiers. Another benefit of using Class E is that the amplifier is "designable;" explicit design equations are given here. The effects of components and frequency variations are defined in advance (see Reference 4, Figs 5 and 6, and Reference 7) and are small. When the amplifier is built as designed, it works as expected, without need for "tweaking" or "fiddling."

Physical Principles for Achieving High Efficiency

Efficiency is maximized by minimizing power dissipation, while providing a desired output power. In most RF and microwave power amplifiers, the largest power dissipation is in the power transistor: the *product* of transistor voltage and current *at each point in time* during the RF period, integrated and averaged over the RF period. Although the transistor must sustain high voltage during *part* of the RF period and conduct high current during *part* of the RF period, the circuit can be arranged so that *high voltage and high current do not exist at the same time.* Then the *product* of transistor voltage and current will be low *at all times* during the RF period. Fig 1 shows conceptual "target" waveforms of transistor voltage and current that meet the high-efficiency requirements. The transistor is operated as a switch. The voltage-current product is low throughout the RF period because:

1. *"On" state:* The voltage is nearly zero when high current is flowing, that is, the transistor acts as a low-resistance closed switch during the "on" part of the RF period.
2. *"Off" state:* The current is zero when there is high voltage, that is, the transistor acts as an open switch during the "off" part of the RF period.

Switching transitions: Although the designer makes the on/off switching transitions as fast as feasible, a high-efficiency technique must accommodate the transistor's practical limitation for RF and microwave applications: the transistor-switching times

will, unavoidably, be appreciable fractions of the RF period. We avoid a high voltage-current product during the switching transitions, *even though the switching times can be appreciable fractions of the RF period*, by the following two strategies:

3. The rise of transistor voltage is *delayed until after the current has reduced to zero*.
4. The transistor voltage returns to zero before the current begins to rise.

The timing requirements of 3 and 4 are fulfilled by a suitable load network (the network between the transistor and the load that receives the RF power), to be examined shortly. Two additional waveform features reduce power dissipation:

5. The transistor voltage at turn-on time is nominally zero (or is the saturation offset voltage, V_o, for a bipolar-junction transistor, hereafter, "BJT"). Then the turning-on transis-
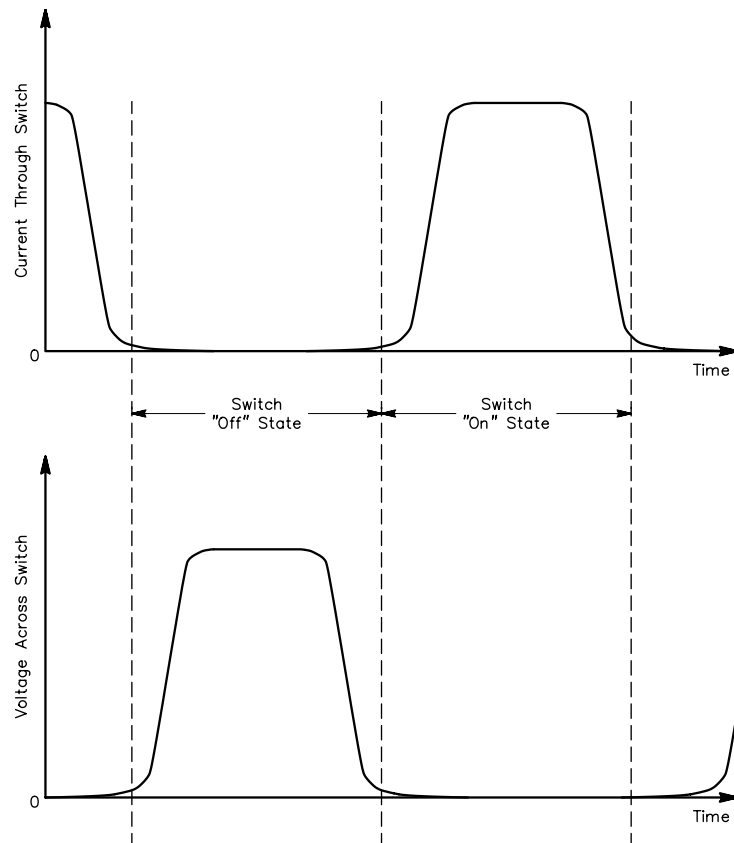

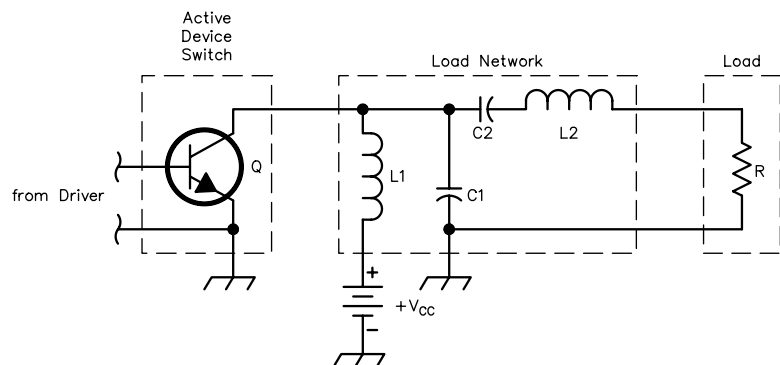**Fig 1—Conceptual "target" waveforms of transistor voltage and current.**

Fig 2—Schematic of a low-order Class-E amplifier.

tor does not discharge a charged shunt capacitance (*C1* of Fig 2), thus avoiding dissipating the capacitor's stored energy (*C1•V2/2*), *f* times per second, where *V* is the capacitor's initial voltage at transistor turn-on and *f* is the operating frequency. (*C1* comprises the transistor output capacitance and any external capacitance in parallel with it.)

6. The slope of the transistor voltage waveform is nominally zero at turn-on time. Then, the current injected into the turning-on transistor by the load network rises smoothly from zero at a controlled moderate rate, resulting in low i^2R power dissipation while the transistor conductance is building-up from zero during the turn-on transition, even if the turn-on transition time is as long as 30% of the RF period.

Result: The waveforms *never* have high voltage and high current *simultaneously*. The voltage and current switching transitions are *time-displaced from each other*, to accommodate transistor switching transition times that can be *substantial fractions of the RF period.* Turn-on transitions may be up to about 30% of the period and turn-off transitions up to about 20% of the period.

The low-order Class-E amplifier of Fig 2 generates voltage and current waveforms that approximate the conceptual "target" waveforms in Fig 1; Fig 3 shows the actual waveforms in that circuit. Note that those actual waveforms meet all six criteria listed above and illustrated in Fig 1. Unpublished higher-order versions of the circuit approximate more closely the target waveforms of Fig 1, making the circuit even more tolerant of component parasitic resistances and non-zero switching-transition times.

*Differences from
Conventional Class B and C*

The load network is not intended to provide a conjugate match to the transistor output impedance. The network design equations come from the solution of a set of simultaneous equations for the steady-state periodic time-domain response of a network (containing non-ideal inductors and capacitors) to periodic operation of a non-ideal switch at the input port, at frequency *f*, to provide (a) an input-port voltage of zero value and zero slope at transistor turn-on time, (b) a first-order approximation to a time delay of the voltage rise at transistor turn-off, and (c) a nearly sinusoidal voltage across the

load resistance *R*, delivering a specified RF power *P* from a specified dc supply voltage V_{CC}.

The transistor's operating locus on the (I_d, V_{ds}) plane is not a tilted straight line (resistance) or a tilted ellipse (resistance + reactance). The operation during the "on" state of the switch is a nearly vertical line whose lower end is at the origin (0, 0); The "off" state of the switch is a horizontal line whose left end is at the origin. By design, the operating locus avoids the remainder of the (I_d, V_{ds}) plane, the region of *simultaneous* high voltage and high current that brings high power dissipation and consequent reduced efficiency. That region is where conventional Class B and C circuits operate.

Analytical and Numerical Derivations of Design Equations

Analytical derivations of design equations for the circuit of Fig 2 can be made only by assuming the current in

$L2$ and $C2$ is sinusoidal. That assumption is strictly true only if the load network has infinite loaded Q (Q_L, defined as $2\pi f L2/R$)[1], and yields progressively less-accurate results for Q_L values progressively lower than infinity. (Q_L is a free-choice design variable,[2] subject to the condition $Q_L \geq 1.7879$—obtained from exact numerical analysis as in References 4 and 6—to obtain the nominal[3] switch-voltage waveform, for the usual choice of the switch "on" duty ratio,[4] D, being 50%.) The amplifier's output power P depends primarily (derivable analytically) on the collector/drain dc-supply voltage V_{CC} and the load resistance R, but secondarily (not derivable analytically) on the value chosen for Q_L. Previously published analytically derived design equations did not include the dependence of P on Q_L. Consequently, the output power is 38% to 10% less than had been expected, for Q_L values in the

[1]Notes appear on page 18.
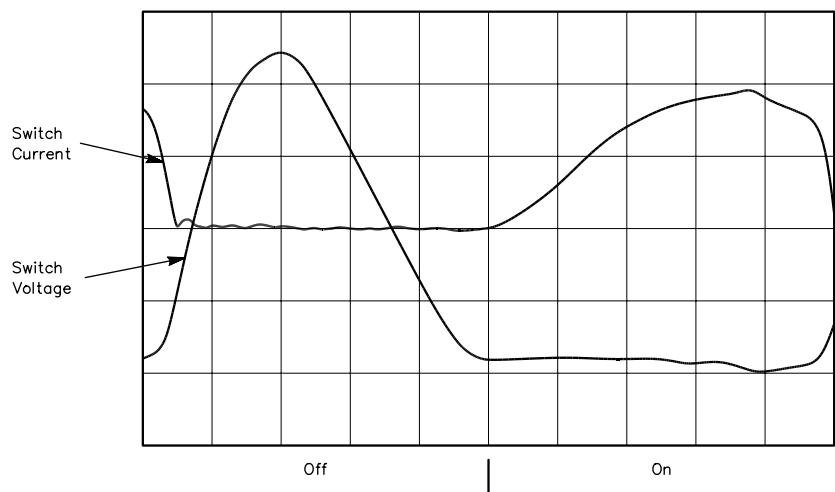
Fig 3—Actual transistor voltage and current waveforms in a low-order Class-E amplifier.

Table 1—Dependence of output power, *C1*, and *C2* on loaded *Q* (Q_L)

Q_L	$\dfrac{PR}{(V_{CC}-V_o)^2}$	$C1•2\pi fR$	$C2•2\pi fR$
infinite	0.576801	0.18360	0
20	0.56402	0.19111	0.05313
10	0.54974	0.19790	0.11375
5	0.51659	0.20907	0.26924
3	0.46453	0.21834	0.63467
2.5	0.43550	0.22036	1.01219
2	0.38888	0.21994	3.05212
1.7879	0.35969	0.21770	infinite

usual range of 1.8 to 5. This paper includes an accurate new equation for P that includes the effect of Q_L. Similar restrictions apply to the analytical derivations of design equations for $C1$, $C2$ and R. However, the needed component values can be found by numerical methods. Table 1 gives normalized exact numerical solutions for output power (hence the needed value of R), $C1$ and $C2$, for eight values of Q_L over the entire possible range from 1.7879 to infinity, for the usual choice of $D = 50\%$. The design equations in the next section are continuous mathematical functions fitted to those eight sets of data. (Having the numerical values of Table 1, readers can derive other mathematical functions to fit the data, if they wish, to substitute for the equations given below.)

Kazimierczuk and Puczko (Reference 5) published a tabulation similar to Table 1 here (using a different mathematical technique, but the two sets of tables agree well; see "Accuracy of Design Equations" below), but they did not include continuous-function design equations based on their tabular data. As a result, a designer using Reference 5 can produce an accurate design at any chosen *tabulated* value of Q_L, but designers lack accurate design information for use at values of Q_L *between* the tabulated values. Avratoglou and Voulgaris (Reference 8) gave an analysis and numerical solutions as graphs but no tables of computed values and no design equations fitted to the numerical results. Precise design values cannot be read from the graphs.

To make accurate circuit designs and advance design evaluations *at any arbitrary value of* Q_L, one needs design equations comprising continuous mathematical functions rather than a set of tabulated values as in Table 1 or Reference 5. The equations should give accurate results, and should be simple enough for designers to easily manipulate. Such equations are given below, for lossless components. The losses are accounted for in References 2, 4, 9, 10 and unpublished notes; the author intends to publish equations for all components of power loss and the resulting collector/drain efficiency. Briefly: Use for P in Eq 6 or 6A the desired output power, divided by the expected collector/drain efficiency and calculate R_{load} from:

$$R_{load} = R + ESR_{L2} + ESR_{C2} + 1.365 R_{on} + 0.2116 ESR_{C1} \qquad \text{(Eq 1)}$$

where R_{on} is the "on" resistance of the transistor. R_{on} is a generic term that represents $R_{DS(on)}$ of a MOSFET or MESFET, or $R_{CE(sat)}$ of a BJT. ESR is the effective series resistance of a reactive component. The expected drain/collector efficiency is approximately

$$\eta_D = \frac{R_{load}}{R_{load} + ESR_{L2} + ESR_{C2} + 1.365 R_{on} + 0.2116 ESR_{C1}} - \frac{(2\pi A)^2}{12} - 0.01 \qquad \text{(Eq 2)}$$

where

$$A = \left(1 + \frac{0.82}{Q_L}\right)\left(\frac{t_f}{T}\right)$$

t_f is the 100%-to-0% fall time of the assumed linear fall of the drain/collector current at transistor turn-off, $T = 1/f$ is the period of the operating frequency, f, and "0.01" allocates about 1% loss of efficiency for the power losses in the dc and RF resistances of the dc-feed choke, $L1$.

Explicit Design Equations

The explicit design equations given below yield the low-order lumped-element Class-E circuit that operates with the nominal waveforms of Fig 3. (Distributed-element circuits

are discussed briefly at the end of the "Applicable Frequency Range..." discussion.) In the equations below, V_{CC} is the dc supply voltage. P is the output power delivered to the load resistance R; f is the operating frequency; $C1$, $C2$, $L1$ (dc-feed choke) and $L2$ are the load network shown in Fig 2. Q_L is the network loaded Q, chosen by the designer as a trade-off among competing evaluation criteria (see Note 2).

In a nominal-waveforms circuit operating with the usual choice of $D = 50\%$, the minimum possible value of Q_L is 1.7879; the maximum possible value is less than the network's unloaded Q. The design procedure is as follows:

$$V_{CC} = \left(\frac{BV_{CEV}}{3.56}\right) SF \qquad \text{(Eq 3)}$$

This includes a chosen safety factor (SF) less than 1, to allow for higher peak voltage resulting from off-nominal load impedance. For example, you could take SF as 80%=0.8. The relationship among P, R, Q_L, V_{CC} and the transistor saturation offset voltage V_o is least-squares fitted to the data in Table 1, over the entire range of Q_L from 1.7879 to infinity, within a deviation of $\pm0.15\%$, by a second-order polynomial function of Q_L:

$$P = \left(\frac{(V_{CC} - V_o)^2}{R}\right)\left(\frac{2}{\left(\frac{\pi^2}{4} + 1\right)}\right) f(Q_L) \qquad \text{(Eq 4)}$$

$$P = \left(\frac{(V_{CC} - V_o)^2}{R}\right) 0.576801\left(1.001245 - \frac{0.451759}{Q_L} - \frac{0.402444}{Q_L^2}\right) \qquad \text{(Eq 5)}$$

Hence:

$$R = \left(\frac{(V_{CC} - V_o)^2}{P}\right) 0.576801\left(1.001245 - \frac{0.451759}{Q_L} - \frac{0.402444}{Q_L^2}\right) \qquad \text{(Eq 6)}$$

Alternatively, a third-order polynomial in Q_L gives a least-squares fit to the data to within -0.0089% to $+0.0072\%$:

$$P = \left(\frac{(V_{CC} - V_o)^2}{R}\right) 0.576801\left(1.0000086 - \frac{0.414395}{Q_L} - \frac{0.577501}{Q_L^2} + \frac{0.205967}{Q_L^3}\right) \qquad \text{(Eq 5A)}$$

Hence:

$$R = \left(\frac{(V_{CC} - V_o)^2}{P}\right) 0.576801\left(1.0000086 - \frac{0.414395}{Q_L} - \frac{0.577501}{Q_L^2} + \frac{0.205967}{Q_L^3}\right) \qquad \text{(Eq 6A)}$$

The *effective* dc-supply voltage is the actual voltage, less the transistor saturation offset voltage, hence $(V_{CC} - V_o)$. V_o is zero for a field-effect transistor. For a BJT, V_o is on the order of 0.1 V at low frequencies, and up to a few volts (depending on transistor fabrication) at frequencies higher than about $f_T/10$.

The design equations for $C1$ and $C2$ that fit the data in Table 1 are given below. The last terms in Eqs 7, 8 and 9 are adjustments to the expressions fitted to the Table 1 data, to account for the small effects of the nonzero susceptance of $L1$.

$$C1 = \frac{1}{2\pi f R \left(\frac{\pi^2}{4} + 1\right)\frac{\pi}{2}} \left(0.99866 + \frac{0.91424}{Q_L} - \frac{1.03175}{Q_L{}^2}\right) + \frac{0.6}{(2\pi f)^2 L1}$$

$$\text{(Eq 7)}$$

$$C1 = \frac{1}{34.2219 f R} \left(0.99866 + \frac{0.91424}{Q_L} - \frac{1.03175}{Q_L{}^2}\right) + \frac{0.6}{(2\pi f)^2 L1}$$

$$\text{(Eq 8)}$$

$$C2 = \frac{1}{2\pi f R} \left(\frac{1}{Q_L - 0.104823}\right)\left(1.00121 + \frac{1.01468}{Q_L - 1.7879}\right) - \frac{0.2}{(2\pi f)^2 L1}$$

$$\text{(Eq 9)}$$

The numerical coefficients in the last terms of Eqs 7, 8, and 9 depend slightly on *L1* and Q_L; those dependencies will be the subject of a planned future article. For the example case of $Q_L = 5$ and the usual choice of X_{L1} being 30 or more times the unadjusted value of X_{C1}, the adjustments for the susceptance of *L1* add 2% or less to the unadjusted value of *C1* and subtract 0.5% or less from the unadjusted value of *C2*. Finally, *L2* is determined by the designer's choice (Note 2) for Q_L, and the value of *R* from Eq 5 or 5A:

$$L2 = \frac{Q_L R}{2\pi f}$$

$$\text{(Eq 10)}$$

Eqs 4, 5, 6, 7, 8, and 9 are more accurate than the older versions in References 1, 2, 4 and 6.

Accuracy of Design Equations

The maximum deviations of Eq 5 from the tabulated values in Table 1 are ±0.15%; those of Eq 5A are –0.0089% and +0.0072%; those of Eq 7 are ±0.13%; and those of Eq 9 are ±0.072%. Kazimierczuk and Puczko (Reference 5) give tables of numerical data (similar to Table 1 here), obtained by a Newton's-method numerical solution of a system of analytical circuit equations they derived, and other useful numerical and graphical data. The tabulated values of *P* in Reference 5 are within –0.13% to +0.47% of the values obtained from the continuous function Eq 5 above. Those differences include (a) the error in the fitting of the continuous function in Eq 5 to the discrete values in Table 1 (±0.15%) and (b) the differences (if any) between the numerical results of Reference 5 and of Table 1 here. Those two sets of tabulated values can be compared directly at only their two values of Q_L in common: infinity (identical results) and 1.7879 (Reference 5 has the same capacitance values and 0.28% lower *P*). The independently computed sets of data here and in Reference 5 agree well (a maximum difference of about 0.3%), giving confidence in the validity of both.

Harmonic Filtering and Associated Changes to Design Equations

The power in Eqs 5 and 5A is the total output power at the fundamental and harmonic frequencies. Most of the power is at the fundamental frequency. The strongest harmonic is the second, with a voltage or current amplitude at *R* of $0.51/Q_L$, relative to the fundamental. For example, with $Q_L = 5.1$, the second-harmonic power is –20 dBc (1% of the fundamental power) without any filtering. Even-order harmonics can be canceled with a push-pull circuit, if desired. In that case, the strongest harmonic is the third, at an amplitude of $0.080/Q_L$ relative to the fundamental, hence –36 dBc (0.025% of the fundamental power) without filtering, for the same example Q_L of 5.1 . In Reference 11, Sokal and Raab give the harmonic

spectrum as a function of the chosen Q_L.[5]

If the circuit includes a low-pass or band-pass filter between *R* and the *C2-L2* branch instead of a direct connection as in Fig 2, the fractions of the output power contained in each of the harmonics will decrease, according to the transmission function of the filter at the harmonic frequencies. As a small side-effect, the total output power and the waveforms of switch voltage and *C2-L2* current will change slightly, requiring small changes to the numerical coefficients in Eqs 6 through 9 above, and in Table 1 and Reference 5. New sets of numerical values can be calculated quickly with the help of a computer program such as *HEPA-PLUS* (Reference 7), which is described briefly below and available from the author's employer.

Optimizing Efficiency

The highest efficiency is obtained by minimizing the *total* power dissipated while the amplifier is delivering a desired output power. That can be done by modifying the waveforms slightly away from the nominal ones shown in Fig 3, allowing *some* of the components of power dissipation to increase, while *other* components of power dissipation *decrease by larger amounts*. For example, allowing the voltage-waveform minimum to be about 20% of its peak value (instead of 0%) increases the *C1*-discharge power loss but reduces the RMS/average ratio of the current waveform and the peak/average ratio of the voltage waveform. Both of those effects can be exploited to obtain a specified output power with a specified safe peak transistor voltage, with lower RMS currents in the transistor, *L1*, *L2*, *C1* and *C2*. That reduces their i^2R dissipations. If their series resistances are large enough, the decrease in their i^2R power losses can outweigh the increase of *C1*-discharge power loss.

The power loss in the transistor R_{on} and in discharging a partially charged *C1* are not functions of the design frequency (*C1* is inversely proportional to frequency, so the product $f(C1 \cdot V^2/2)$ is independent of frequency). For given types of capacitors or inductors, losses in capacitor ESRs (including that in the transistor's C_{out}) increase with design frequency, inductor-core losses increase, and inductor-winding losses decrease.

The optimum trade-off depends on the specific combination of parameter values of the types of components being considered in a particular design. (It does not vary appreciably from one unit to another of a given design.) No explicit analytical method yet exists for achieving the optimum trade-off among all of the components of power loss. Optimization is a numerically intensive task, too difficult to do by explicit analytical methods, but computerized optimization is practical. For example, running on an IBM-PC-compatible computer with a Pentium II/233-MHz processor, *HEPA-PLUS* designs a nominal-waveforms Class-E amplifier in a time too short to observe, simulates the circuit in 0.019 seconds and optimizes the design automatically—according to user-specified criteria—in about 6 seconds. The program uses double-precision computation for accuracy and robustness, yielding the circuit voltage and current waveforms and their spectra, dc input power, RF output power and all components of power dissipation.

Effects of Non-Ideal Components

Many non-ideal characteristics of the circuit components can be included in an analytical solution if the circuit is operating with the nominal switch-voltage waveform, but the task becomes progressively more difficult as one attempts to include more of those effects simultaneously. It

becomes impossible if the circuit is not operating at the nominal-waveforms conditions. *HEPA-PLUS* simulates an expanded version of the Fig 2 circuit in any arbitrary operating condition (nominal or non-nominal waveforms). It includes all-important "real-world" non-ideal characteristics of the transistor, the finite-Q power losses of all inductors and capacitors, and parasitic wiring inductance in series with *C1* and in series with the transistor. Details are available from the author's employer.

Applicable Frequency Range (about 3 MHz to 10 GHz, maybe 11 GHz)

The Class-E amplifier can operate at arbitrarily low frequencies. Below about 3 MHz, one of the three switching-mode Class-D amplifier types might be preferred. Each can be as efficient as the Class-E, with about 1.6 times as much output power per transistor, but with the possible disadvantage that transistors must be used in pairs, versus the single Class-E transistor. Class E is preferable to Class D at frequencies higher than about 3 MHz because it is more efficient, the transistor input port is easier to drive, and Class-E has fewer detrimental effects from parasitic inductance in the output-port circuit.

Low-order Class E amplifiers are useful up to the frequency at which the achievable turn-off switching time is about 17% of the RF period. In a Class-B amplifier, the turn-off transition time is 25% of the period. Therefore a low-order Class-E circuit will work well with a particular transistor at frequencies up to about 17%/25% = 70% of the frequency at which that transistor works well in a Class-B amplifier. (Unpublished higher-order Class-E circuits can operate efficiently at frequencies up to about twice that of the low-order version.)

Class-E circuits have operated at frequencies as high as 8.35-10 GHz (Reference 42). Several microwave designers have reported achieving remarkably high efficiency by driving the amplifier into saturation and using a favorable combination of series inductance with the load resistance (Reference 13) or fundamental and harmonic load impedances (References 14, 15, 16, 17, 18, 19, 20). (The authors of those references found favorable tuning conditions by using an automatic tuner and/or circuit simulation to exhaustively search the multidimensional impedance space for a favorable combination of circuit values, rather than by using explicit design equations.) Secchi (Reference 13) and Mallet *et al* (Reference 14) provided plots of their drain-voltage and collector-voltage waveforms. Inspection of the V_{ds} waveform (Fig 2 in Reference 13) shows a nominal Class-E waveform with $R_{DS(on)}$ = 2.7 V/0.688 A = 3.9 Ω. The waveforms in Fig 2B of Reference 14

Table 2—Example Class-E Power Amplifiers

Frequency	Power	Transistor	Collector or Drain Efficiency/PAE	Organization	Approximate Year	See Reference
0.52-1.7 MHz	44 kW PEP	push-pull MOSFETs	95%	Broadcast Electronics, Inc	1992	34
14 MHz	110 W	International Rectifier IRF540	92%	Design Automation, Inc	1986	36
13.56 MHz, 27.12 MHz	2 kW	MOSFET	90%	Dressler Hochfrequenztechnik	1993	
13.56 MHz	3 kW, 5.5 kW	MOSFET	?%	Advanced Energy Industries, Inc	1992-1997	
27.12 MHz	22 W	International Rectifier IRF510	89-92%	Design Automation, Inc	1991	37
145 MHz	2.58 W	Siliconix VMP4 VMOSFET	96.5%/81.3%*	École Polytech. Féd. Lausanne	1980	32
300 MHz	30 W	push-pull BJTs	89%	Harris RF Communications	1992	39
450 MHz	14.96 W	combine 4 modules MRF873 BJT	89.5%	City Univ. of Hong Kong	1997	30
500 MHz	0.55 W	Siemens CLY5 GaAs MESFET	83%/80%	Univ. of Colorado	1995	23
840 MHz	1.24 W	GaAs MESFET	79%/77%	S. C. Cripps	<1999	40
850 MHz	1.6 W	GaAs MMIC	62.3% PAE	M/A-COM	1994	26
1 GHz	0.94 W	Siemens CLY5 GaAs MESFET	75%/73%	Univ. of Colorado	1995	22, 21
2.45 GHz	1.27 W	Fujitsu FLC30 GaAs MESFET	72% PAE	RCA David Sarnoff Res. Ctr.	1981	13
2.45 GHz†	210 mW	Raytheon RPC3315 MESFET	77%/68%/71%*	Design Automation, Inc	1979	33
5 GHz	0.61 W	Fujitsu FLK052WG MESFET	81%/72%	Univ. of Colorado	1996	12, 23
8.35 GHz	1.41 W	Fujitsu FLK202MH-14 MESFET	64%/48%	Univ. of Colorado	1999	41
10 GHz	100 mW	Alpha Ind. AFM04P2 MESFET	74%/62%	Univ. of Colorado	1999	42

*Overall efficiency = $P_{out}/(P_{dc} + P_{drive})$
†1/20 scaled-frequency model at 122.5 MHz; see Reference 33.

are Class-E, but with an unusually small conduction angle. Higher output power could probably be obtained by increasing the conduction angle and modifying the load-network impedance accordingly. I do not know the operating mode in References 15-20; very likely those amplifiers are distributed-element versions (see below) of Class E, achieved empirically.

Distributed versus Lumped Elements

High-efficiency waveforms similar to those in Figs 1 or 3 can be generated with lumped and/or distributed elements. At a given frequency, the choice depends on the available components and the tradeoffs among their sizes, costs, quality factors and parasitic effects. Amplifiers in References 12, 21, 22, 23, 41 and 42 were transmission-line versions of Class E, operating at 10, 8.35, 5, 2, 1 and 0.5 GHz. The 5-, 2- and 1-GHz circuits were described as having been designed by explicit design procedures, working as expected; they were operated and measured without making any experimental adjustment.

Experimental Results

Table 2 summarizes Class-E performance achieved by amplifiers operating from 44 kW PEP at 0.52-1.7 MHz to 1.41 W at 8.35 GHz and 100 mW at 10 GHz.

Tuning Procedure

Fig 3 shows the nominal Class-E transistor-voltage waveform in the low-order circuit of Fig 2. At the transistor's turn-on time, the waveform has zero slope and zero voltage for an FET or $V_{CE(sat)}$ for a BJT. An actual circuit, or a circuit in *HEPA-PLUS*, can be brought from an off-nominal condition to that nominal-waveform condition by adjusting $C1$, $C2$ and/or $L2$. If R is not already the desired value for the desired output power, it may need adjustment. The desired value of R comes from Eq 6 or 6A after having applied the allowance for parasitic resistances discussed in the last paragraph of "Analytical and Numerical Derivations of Design Equations," above.[6]

After adjusting a matching network (located between the load and the right-hand end of $L2$ in Fig 2) to provide R, there might be residual series reactances in series with R. Any series inductive reactance adds to that of $L2;$ any series capacitive reactance adds to that of $C2$. Then the circuit would operate with an off-nominal V_{CE} waveform and possibly an off-nominal value of output power, because the *effective*

values of $L2$ and $C2$ differ from the design values. To correct for that, the reactances of $L2$ and $C2$ should be reduced by the amounts of the residual reactance at the matching network input. The following text and figures explain how to make those adjustments to the circuit, if needed, *without advance knowledge of the series reactances at the input port of the matching network*. The text is in terms of a BJT; for a FET, substitute "V_{DS}" for "V_{CE}."

The circuit parameters were chosen, via Eqs 2, 3, 4, 5, 6, 7, 8, 9, and 10, to meet a chosen set of requirements. The circuit will operate with the nominal Class-E waveform, while delivering the specified output power at the specified frequency, if the chosen parameter values are installed in the actual hardware. The possible need for tuning results from (a) tolerances on the component values (normally not a problem, because Class E has low sensitivity to component tolerances) and (b) the possibility of unknown-value reactances in series with R (hence, in series with $L2$ and $C2$) after the load resistance has been transformed to the chosen value of R. Those series reactances

require that the reactances of $L2$ and $C2$ be reduced by the amounts of the unknown inserted inductive and capacitive series reactances, but how can we do that when those inserted reactances are unknown?

Fig 4 shows a V_{CE} waveform for an amplifier with off-nominal tuning, with the waveform features labeled for subsequent reference in the text. If we know how changes of $L2$ and $C2$ will affect that waveform, we can adjust $L2$ and $C2$ to meet two criteria at the operating frequency: (a) achieve the nominal V_{CE} waveform of Fig 3 and (b) deliver the specified value of output power.

Fig 5 shows how $L2$ and $C2$ affect the V_{CE} waveform. We know also that increasing $L2$ reduces the output power and vice versa. With (a) an oscilloscope displaying the V_{CE} waveform and (b) a directional power meter indicating the power delivered to the load, we can adjust $L2$ and $C2$ to simultaneously fulfill the two desired conditions (nominal waveform and desired outuput power) *even if the reactances in series with* R *are unknown.*
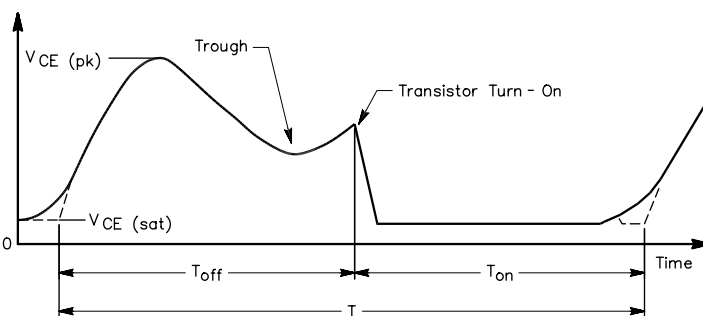
If $C1$ (comprised of the transistor

Fig 4—Typical mistuned V_{CE} waveform, showing transistor turn-on, turn-off and waveform "trough."
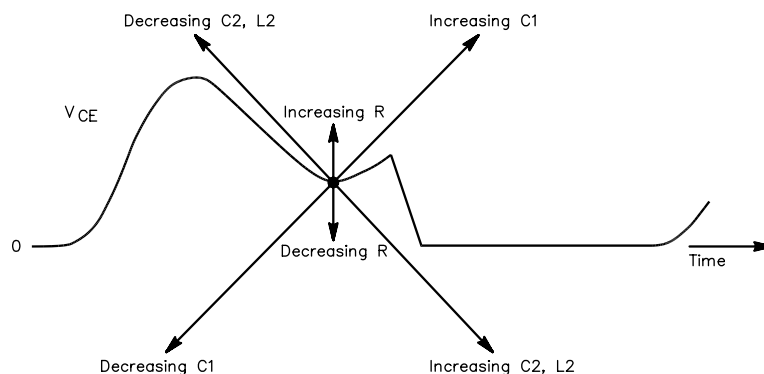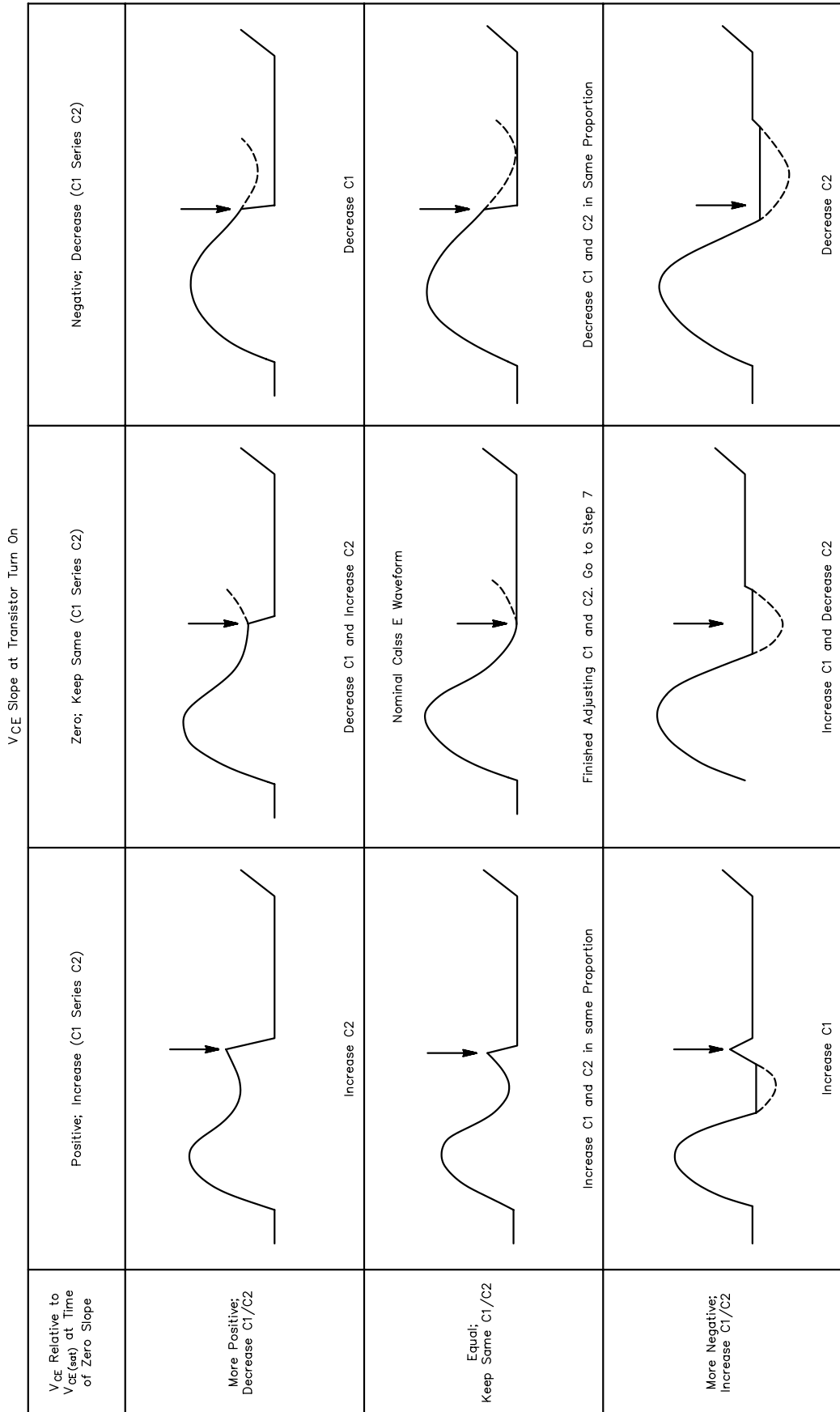

Fig 5—Effects of adjusting load-network components.

V_{CE} Slope at Transistor Turn On

V_{CE} Relative to $V_{CE(sat)}$ at Time of Zero Slope	Positive; Increase (C1 Series C2)	Zero; Keep Same (C1 Series C2)	Negative; Decrease (C1 Series C2)
More Positive; Decrease C1/C2	Increase C2	Decrease C1 and Increase C2	Decrease C1
Equal; Keep Same C1/C2	Increase C1 and C2 in same Proportion	Nominal Calss E Waveform — Finished Adjusting C1 and C2. Go to Step 7	Decrease C1 and C2 in Same Proportion
More Negative; Increase C1/C2	Increase C1	Increase C1 and Decrease C2	Decrease C2

Fig 6—*C1* and *C2* adjustment procedure. The vertical arrow indicates transistor turn-on.

output capacitance and the external capacitor connected in parallel with it) is within about 10% of the intended value, $C1$ will normally not need adjustment. When there is a large deviation from the design value, $C1$ can be adjusted to achieve the nominal V_{CE} waveform, using the information in Fig 5 about the effects of $C1$ on the V_{CE} waveform.

In that case, the three components $C1$, $C2$ and $L2$ can be adjusted to achieve three conditions simultaneously at the operating frequency: desired output power, transistor voltage of $V_{CE(sat)}$ just before transistor turn-on and zero slope of the V_{CE} waveform just before turn-on. The following diagrams and text explain how to adjust $C1$, $C2$, $L2$ and R (if desired) to adjust the shape of the V_{CE} waveform.

Changes in the values of the load-network components affect the V_{CE} waveform as follows, illustrated in Fig 5:
1. Increasing $C1$ moves the trough of the waveform upwards and to the right.
2. Increasing $C2$ moves the trough of the waveform downwards and to the right.
3. Increasing $L2$ moves the trough of the waveform downwards and to the right.
4. Increasing R moves the trough of the waveform upwards (R is not normally an adjustable circuit element).

Knowing these effects, you can adjust the load network for nominal Class-E operation by observing the V_{CE} waveform. (Depending on the settings of the circuit component values, the zero-slope point and/or the negative-going jump may be hidden from view, as in some of the waveforms in Fig 6. If that occurs, the locations of those features on the waveform can be estimated by extrapolating from the part of the waveform that can be seen.) The adjustment procedure is:
1. Set R to the desired value or accept what exists.
2. Set $L2$ for the desired $Q_L = 2\pi f L2/R$ or accept what exists.
3. Set the frequency as desired.
4. Set the duty ratio (T_{on}/T) to the desired value (usually 50%), with V_{CC} set to approximately 4 V. If the transistor turn-on is visible on the V_{CE} waveform (as in Fig 4), measure the duty ratio. Otherwise, observe the V_{BE} waveform and assume that turn-on occurs when the positive-going edge of V_{BE} reaches +0.8 V and turn-off occurs when the negative-going edge of V_{BE} reaches 0 V.

5. Observe the trough of the V_{CE} waveform:
   A. At the zero-slope point: What is the voltage relative to $V_{CE(sat)}$, more positive, more negative or equal?
   B. At transistor turn-on: What is the slope, positive, negative or zero?

   If these points are unobservable because they lie below the 0 V axis, the voltage at zero slope is "more negative." Estimate the slope at turn-on by extrapolation of the waveform.

   If the voltage at zero slope is unobservable because transistor turn-on occurs before zero slope is reached, the slope at turn-on is "negative." Estimate the voltage at zero slope by extrapolation of the waveform.

   If you cannot estimate the V_{CE} or the slope by extrapolation, assume that V_{CE} is "equal" or that the slope is "zero."
6. Adjust $C1$ and/or $C2$ as shown in Fig 5, and in expanded form in Fig 6.
7. If V_{CC} is now the desired value, go to Step 8. If V_{CC} is less than the desired value, increase V_{CC} by up to 50% and readjust the duty ratio, $C1$ and $C2$ as needed. (The V_{CC} increase will decrease the effective value of C_{CB}, causing the effective value of $C1$ to be reduced. Therefore, $C1$ will need to be increased slightly.)
8. For a final check of the adjustments, increase $C1$ slightly to generate an easily visible marker of transistor turn-on: the small negative-going step of V_{CE}. Verify that the duty ratio is the desired value (usually 50%) and that the waveform slope is zero at turn-on time. Now return $C1$ to the value that brings the waveform to $V_{CE(sat)}$ at turn-on time (and also eliminates the marker).

Gate- and Base-Driver Circuits

If one takes a simplistic view, driver-stage design is less important than that of the output stage. The reasoning is that the driver power level is lower than that of the output stage, by a factor equal to the gain of the output stage—typically a factor of 10 to 100. *That simplistic view is not correct*, because the output transistor will not operate as intended if its input is not driven properly. If the output transistor does not operate as intended, the output stage will not operate as intended, either. The resulting output-stage performance might or might not be acceptable. The output-stage

transistor will operate properly as a switch, as intended, if its input port (gate-source of an FET or base-emitter of a BJT) is driven properly by the output of its driver stage. The driver stage must provide the output specified below. (Symbols for FETs are used below; you can convert to BJT symbols if you wish.)

1. *It must provide enough "off" bias* during the "off" interval to maintain the drain or collector current at an acceptably small value. If you are willing to tolerate a power loss of x fraction of the normal dc-input power due to non-zero "off"-state current, the drain or collector current during the "off" interval can be up to

$$I_{D(off)} = x \bullet I_{DD}\left(\frac{1}{1-D}\right) \qquad (\text{Eq } 11)$$

where I_{DD} is the dc current drawn from the V_{DD} dc drain-voltage supply, and D is the output-transistor's "on" duty ratio (usually 0.50, but it can be any value you choose and provide for in the choice of R, L and C values in the load network).

Example: If you are willing to tolerate 1% additional power consumption from the V_{DD} voltage supply caused by the non-zero "off"-state current, if I_{DD} is 5 A and if D is the usual value of 0.50, you can tolerate an "off"-state drain current of 0.01 (5 A) (1/(1–0.50)) = 0.1 A or 100 mA. That's easy to meet. Consider the International Rectifier IRF540 (rated at 100 V, 28 A). It is specified for 0.25 mA maximum at V_{GS} = 0 and V_{DS} = 80 V at T_J = 150°C, a factor of 400 smaller than the 100 mA you are willing to accept in this example.

2. *It must provide enough "on" drive* during the latter 75% of the "on" interval to maintain a low-enough R_{on}. You can choose what is "low enough" for your purposes: Refer to Eq 2 and make R_{on} a small-enough fraction of R_{load} to yield a collector/drain efficiency that you consider satisfactory. Why is it "the latter 75% of the 'on' interval"? The current $i(t)$ during the first 25% of the "on" interval is small enough that $[i(t)]^2 R_{on}(t)$ can be acceptably small for a fairly high $R_{on}(t)$ because the small $i(t)$ during the first 25% of the "on" interval causes an even smaller $[i(t)]^2$ (the square of a small number is even smaller).

3. *It must provide enough turn-off drive* to turn-off the drain or collector current from 100% to 0% in a fall-time, t_f, fast enough to make the turn-off power dissipation an acceptably small fraction of the output power. That fraction is

$$\frac{(2\pi A)^2}{12}$$

where

$$A = \left(1 + \frac{0.82}{Q_L}\right)\frac{t_f}{T}$$

and $T = 1/f$ is the period of the operating frequency, f. Choose the acceptable fraction of the output power to be dissipated during the non-zero turn-off switching time. Then calculate the required drain- or collector-current-fall time t_f that must result from the "enough turn-off drive." Then provide sufficient turn-off drive to accomplish your chosen objective, according to the characteristics of the chosen output transistor. (That is the subject of an intended future publication.)

For example, if you are willing to have the turn-off power dissipation ($P_{\text{diss,turn-off}}$) be 6% of the output power, and if $Q_L = 3$, the allowable value for

$$\frac{t_f}{T} = \frac{\sqrt{12\left(\dfrac{P_{\text{diss, turn-off}}}{P}\right)}}{2\pi\left(1 + \dfrac{0.82}{Q_L}\right)} \quad \text{(Eq 13)}$$

$$= \frac{\sqrt{12(0.06)}}{2\pi\left(1 + \dfrac{0.82}{3}\right)} = 0.106$$

That is, t_f can be 10.6% of the period.

4. *It must provide enough turn-on drive* to turn-on the output transistor fast enough to make power dissipation during the turn-on switching acceptably small. That has never been a problem with any of the drivers I have seen. Most driver circuits turn the transistor "on" and "off" with about the same switching times. If the more-important turn-off switching time is fast enough, the accompanying turn-on switching time will be more than fast enough.

The input-port characteristics of BJTs, MOSFETs and MESFETs are so different that a different driver circuit should be used for each type of transistors.[7] I intend to publish a future article that discusses details of driver circuits meeting criteria 1 through 4 for MOSFETs, MESFETs and BJTs. A brief summary of methods for driving MOSFETs or MESFETs follows. The polarity descriptions assume N-channel or NPN; reverse the polarity descriptions for P-channel or PNP.

The best gate-voltage drive is a trapezoid waveform, with the falling transition occupying 30% or less of the period. (Trade-off: Shorter turn-off transition times yield less power dissi-

pation in the output transistor during turn-off switching, but greater power consumption of the driver stage. For both MOSFETs and MESFETs, the optimum drive minimizes the sum of the output-stage power dissipation and the driver-stage power consumption.) The peak of the drive waveform should be safely below the MOSFET's maximum gate-source voltage rating. For MESFETs, it should be less than the gate-source voltage at which the gate-source diode conducts enough current to cause either of two undesired effects: (a) metal migration of the gate metalization at an undesirably rapid rate (making the transistor operating lifetime shorter than desired) or (b) enough power dissipation to reduce the overall efficiency more than the efficiency is increased by the lower dissipation in the lower $R_{\text{DS(on)}}$ that results from a higher upper level of the drive waveform. The lower level of the trapezoid should be low enough to result in a satisfactorily small current during the transistor's "off" state, discussed in requirement 1 above.

A sine wave is a usable (but not optimum) approximation to the trapezoid waveform described above. To obtain an output-transistor "on" duty ratio of 50% (usually the best choice, but a larger or smaller duty ratio can be used if appropriate component values are used in the load network), the zero level of the sine wave should be positioned slightly above the FET's turn-on threshold voltage.

It is a better approximation to remove the part of the sine-wave that goes below the V_{GS} value that ensures fully "off" operation, replacing it with a constant voltage at that V_{GS} value. This reduces the input-drive power by slightly less than 50%, almost doubling the power gain of the output stage. A planned subsequent article will discuss in detail a simple circuit that generates such a waveform.

Acknowledgements

The author thanks Professor Alan D. Sokal of the Physics Department, New York University, for many helpful discussions and for producing the numerical solutions in Table 1 and the initial set of equations that fit the data in Table 1; John E. Donohue, formerly of Design Automation, for computing the coefficients of $f(Q_L)$ in Eq 4 to fit the data in Table 1, yielding Eqs 5 and 6; and Dr. Richard Redl of ELFI S. A., for computing the improved-accuracy functions in Eqs 5A, 7 and 9 that fit the P, $C1$ and $C2$ data of Table 1.

Notes

[1] Most papers on the Class-E amplifier of Fig 2 (including this one) define Q_L as $2\pi f\, L2\,/R$. A few papers, for example, Reference 3, define Q_L as $(1/(2\pi fC2R)$. Kazimierczuk and Puczko (Reference 5, to their credit) give *both* values in their tabulations, as Q_L and as Q_1, respectively.

[2] The choice of Q_L involves a trade-off among operating bandwidth (wider with lower Q_L), harmonic content of the output power (Reference 11, lower with higher Q_L) and power loss in the parasitic resistances of the load-network inductor $L2$ and capacitor $C2$ (lower with lower Q_L).

[3] The nominal switch-voltage waveform has zero voltage and a zero slope at the time the switch will be turned on. References 1, 2, 3, 4 and papers by other authors, referred to that nominal waveform as the "optimum" waveform, a misnomer. That waveform is "optimum" for yielding high efficiency in the case of a switch with negligibly small series resistance. If the switch has appreciable resistance, however, the efficiency can be increased by moving away slightly from the nominal waveform, to a waveform whose voltage at the switch turn-on time is of the order of 20% of the peak voltage. No analytical optimization procedure yet exists, but the circuit can be optimized numerically, by a computer program such as *HEPA-PLUS*, discussed briefly in this paper (see Reference 7).

[4] Beware: A few publications define D as the fraction of the period that the switch is *off*.

[5] Updates to Reference 11: (a) Delete the column in Table 1 for $Q_L = 1$ because Q_L must be ≥1.7879 to obtain the nominal Class-E collector/drain-voltage waveform in the circuit described in References 1, 2, 3, 4, 5, 6, when the switch duty ratio D is 50%. (b) In Eq 4, change the factor 1.42 to 1.0147, the factor 2.08 to 1.7879 and the factor 0.66 to 0.773. (c) Recalculate the numerical values of I_n/I_1, using Eq 4 with the revised factors.

[6] The 1997 two-part *QST* article by Eileen Lau, KE67VWU, et al, about 300-W and 500-W 40 meter transmitters (Reference 43), discussed tuning in Part 2, but without a description of how to adjust the load-network components to obtain the nominal Class-E voltage waveform, as is included here under "Tuning Procedure."

[7] In the early 1980s, I made a driver circuit that would drive a BJT or a MOSFET interchangeably, with no change needed in the driver or in the PA input circuit. That driver was used in a Class-E demonstrator circuit, so that a person evaluating Class-E technology could insert either type of transistor for test purposes and observe that the changes of PA output power and efficiency were almost unnoticeably small, with any of 20 transistors of different type numbers and manufacturers, some BJTs and some MOSFETs. Some of those people, accustomed to working with conventional Class-C power amplifiers, were astonished when they witnessed the results of that test.

REFERENCES

1. N. O. Sokal and A. D. Sokal, "High-Efficiency Tuned Switching Power Amplifier," *US Patent 3,919,656*, Nov 11, 1975 (now expired, includes a detailed technical description).

2. N. O. Sokal and A. D. Sokal, "Class E—A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers," *IEEE Journal of Solid-State Circuits*, Vol SC-10, No. 3, pp 168-176, June 1975. (The text of Reference 1 cut to half-length; retains the most-useful information. Text corrections available from N. O. Sokal.)

3. F. H. Raab, "Idealized Operation of the Class E Tuned Power Amplifier," *IEEE Transactions on Circuits and Systems*, Vol CAS-24, No. 12, pp 725-735, Dec 1977.

4. N. O. Sokal and A. D. Sokal, "Class E Switching-Mode RF Power Amplifiers— Low Power Dissipation, Low Sensitivity to Component Tolerances (Including Transistors), and Well-Defined Operation," *Proceedings of the 1979 IEEE ELECTRO Conference, Session 23*, New York, New York, 25 April 1979. Reprinted in *RF Design*, Vol 3, No. 7, pp 33-38 and 41, July/Aug 1980 (includes plots of efficiency versus variations of all circuit parameters and efficiency versus frequency with Q_L as a parameter).

5. M. K. Kazimierczuk and K. Puczko, "Analysis of Class E Tuned Power Amplifier at any Q and Switch Duty Cycle," *IEEE Transactions on Circuits and Systems*, Vol CAS-34, No. 2, pp 149-159, Feb 1987.

6. N. O. Sokal, "Class E High-Efficiency Switching-Mode Power Amplifiers, from HF to Microwave," *1998 IEEE MTT-S International Microwave Symposium Digest,* June 1998, Baltimore, Maryland, CD-ROM IEEE Catalog No. 98CH36192 and also *1998 Microwave Digital Archive, IEEE Microwave Theory and Techniques Society,* CD-ROM IEEE Product # JP-180-0-081999-C-0.

7. *HEPA-PLUS* computer program, available from author's employer, Design Automation.

8. Ch. P. Avratoglu and N. C. Voulgaris, "A New Method for the Analysis and Design of the Class E Power Amplifier Taking into Account the Q_L Factor," *IEEE Transactions on Circuits and Systems*, Vol CAS-34, No. 6, pp 687-691, June 1987.

9. F. H. Raab and N. O. Sokal, "Transistor Power Losses in the Class E Tuned Power Amplifier," *IEEE Journal of Solid-State Circuits*, Vol SC-13, No. 6, pp 912-914, Dec 1978.

10. N. O. Sokal and R. Redl, "Power Transistor Output Port Model for Analyzing a Switching-Mode RF Power Amplifier or Resonant Converter," *RF Design*, June 1987, pp 45-48, 50-53.

11. N. O. Sokal and F. H. Raab, "Harmonic Output of Class-E RF Power Amplifiers and Load Coupling Network Design," *IEEE Journal of Solid-State Circuits*, Vol SC-12, No. 1, pp 86-88, Feb 1977.

12. E. W. Bryerton, W. A. Shiroma, and Z. B. Popovic, "A 5-GHz High-Efficiency Class-E Oscillator," *IEEE Microwave and Guided Wave Letters*, Vol 6, No. 12, Dec 1996, pp 441-443. (300 mW to external load at 5 GHz at 59% conversion efficiency, additional RF output was used for input-drive to oscillator.)

13. F. N. Sechi, "High Efficiency Microwave FET Amplifiers," *Microwave Journal*, Nov 1981, pp 59-62, 66. (Several "saturated Class B and Class AB" amplifiers at 2.45 GHz, using several types of GaAs MESFETs: 0.97 W at 71% PAE, 1.2 W at 72% PAE, 1.27 W at 72% PAE. The V_{ds} waveform in Fig 2 is a low-order Class-E waveform with apparently $R_{DS(on)}$ = 2.7 V/0.688 A = 3.9 Ω. All drain-current waveforms are sinusoidal; that seems to be inconsistent with the non-sinusoidal drain-voltage waveforms. Perhaps the bandwidth of the current-sensing instrumentation was sufficient to display only the fundamental component of the probably non-sinusoidal current waveforms.)

14. A. Mallet, D. Floriot, J. P. Viaud, F. Blache, J. M. Nebus, and S. Delage, "A 90% Power-Added-Efficiency GaIn P/GaAs HBT for L-Band and Mobile Communication Systems," *IEEE Microwave and Guided Wave Letters*, Vol 6, No. 3, pp 132-134, March 1996. (Fig 1 is well annotated with the HBT parameter values, but it omits values for C_{be} and C_{bc}.)

15. S. R. Mazumder, A. Azizi, and F. E. Gardiol, "Improvement of a Class-C Transistor Power Amplifier by Second-Harmonic Tuning," *IEEE Transactions of MTT*, Vol MTT-27, No. 5, pp 430-433, May 1979. (800 mW output at 865 MHz, 53.3% collector efficiency, coupled-TEM-bar circuit. In a similar paper at the 9th European Microwave Conference, Sep 1979, the same authors reported 64% collector efficiency at 800-mW output at 850 MHz.)

16. J. J. Komiak, S. C. Wang, and T. J. Rogers, "High Efficiency 11 Watt Octave S/C-Band PHEMT MMIC Power Amplifier," *Proceedings of the IEEE 1997 MTT-S International Microwave Symposium*, Denver, Colorado, June 8-13, 1997, IEEE Catalog No. 7803-3814-6/97, pp 1421-1424. (17 W at 5.1 GHz, 54.5% PAE, harmonic tuning.)

17. J. J. Komiak, L. W. Yang, "5 Watt High Efficiency Wideband 7 to 11 GHz HBT MMIC Power Amplifier," *Proceedings of the IEEE 1995 Microwave and Millimeter-Wave Monolithic Circuits Symposium*, Orlando, Florida, May 15-16, 1995, IEEE Catalog No. 95CH3577-7, pp 17-20.

18. W. S. Kopp and S. D. Pritchett, "High Efficiency Power Amplification for Microwave and Millimeter Frequencies," *1989 IEEE MTT-S Digest*, IEEE Catalog No. CH2725-0/89/0000, pp 857-858.

19. Bill Kopp and D. D. Heston, "High-Efficiency 5-Watt Power Amplifier with Harmonic Tuning," *1988 IEEE MTT-S Digest*, pp 839-842. [Twelve FETs in parallel produced (from Table 3) 5.27 W output (apparently 0.27 W of that is lost in power-combining network) at 10 GHz with 35.3% PAE (Abstract says 5 W at 36% PAE). This results from an exhaustive search for the best combination of impedance and frequency. The circuit was built with distributed elements.]

20. L. C. Hall and R. J. Trew, "Maximum Efficiency Tuning of Microwave Amplifiers," *1991 IEEE MTT-S Digest*, IEEE Catalog No. CH2870-4/91/0000, pp 123-126. [Circuit simulations of optimum design found by exhaustive search of 12-dimensional parameters space; the resulting design appears to be higher-order Class F with third-harmonic resonator (Class F3).]

21. T. Mader, M. Markovic, Z. B. Popovic, and R. Tayrani, "High-Efficiency Amplifiers for Portable Handsets," *Conference Record, IEEE PIMRC'95* (Personal, Indoor and Mobile Radio Communications), Sep 1995, Toronto, Ontario, Canada, IEEE publication 0-7803-3002-1/95, pp 1242-1245. (Class E, 0.94 W at 1 GHz, at 75% drain efficiency, 73% PAE, Siemens CLY5 GaAs MESFET.)

22. T. B. Mader and Z. B. Popovic, "The Transmission-Line High-Efficiency Class-E Amplifier," *IEEE Microwave and Guided Wave Letters*, Vol 5, No. 9, Sep 1995, pp 290-292. (0.94 W at 1 GHz at 75% drain efficiency, 73% PAE; 0.55 W at 0.5 GHz at 83% drain efficiency, 80% PAE; Siemens CLY5 GaAs MESFET.)

23. T. B. Mader, "Quasi-Optical Class-E Power Amplifiers," PhD thesis, 1995, University of Colorado, Boulder, Colorado. (Operates Class E with transmission lines: 0.55 W at 0.5 GHz at 83% drain efficiency, 80% PAE from Siemens CLY5 MESFET; 0.61 W at 5 GHz at 81% drain efficiency, 72% PAE from Fujitsu FLK052WG MESFET. Four of the latter fed into a quasi-optical power combiner gave 2.4 W at 5.05 GHz at 74% efficiency, 64% PAE.)

24. T. Sowlati, C. A. T. Salama, J. Sitch, G. Rabjohn, and D. Smith, "Low Voltage, High Efficiency GaAs Class E Power Amplifiers for Wireless Transmitters," *IEEE Journal of Solid-State Circuits*, Vol 30, No. 10, pp 1074-1080, Oct 1995. The same authors and almost-identical title and text appear in *Proceedings of the IEEE GaAs IC Symposium*, Philadelphia, Pennsylvania, Oct 18-19, 1994, IEEE Catalog No. 0-7803-1975-3/94, pp 171-174. [24 dBm = 0.25 W output at 835 MHz, at >50% power-added efficiency using integrated impedance-matching networks (PAE would be 75% with hybrid matching networks), from an 8.4-mm^2 GaAs IC at 2.5 V dc.]

25. T. Sowlati, Y. Greshishchev, C. A. T. Salama. G. Rabjohn, and J. Sitch, "Linear Transmitter Design Using High Efficiency Class E Power Amplifier," *Conference Record, IEEE PIMRC'95* (Personal, Indoor and Mobile Radio Communications), Sep 27-29, 1995, Toronto, Ontario, Canada, IEEE publication 0-7803-3002-1/95, pp 1233-1237. (24 dBm = 251 mW at 835 MHz, 65% PAE.)

26. J. Imbornone, R. Pantoja, and W. Bosch, "A Novel Technique for the Design of High Efficiency Power Amplifiers," European Microwave Conference, Cannes, France, Sep 1994. [32.1 dBm = 1.6 W output at 850 MHz, at 62.3% power-added efficiency, from an 18-mm^2 GaAs IC (output stage and driver stage) with high-Q lumped elements, at 5 V dc. Simulated V_{ds} and I_d waveforms for optimized output stage are Class E with $V_{turn-on}/V_{pk}$ = 4.9 V/27.4 V = 18%, as discussed here under "Optimizing Efficiency."]

27. K. Siwiak, "A Novel Technique for Analyzing High-Efficiency Switched-Mode Amplifiers," *Proceedings of the RF Expo East '90*, Nov 1990, pp 49-56. [Higher-order Class E with third-harmonic resonator (Class F3).]

28. C. Duvanaud, S. Dietsche, G. Pataut, and J. Obregon, "High-Efficiency Class-E GaAs FET Amplifiers Operating with Very Low Bias Voltages for Use in Mobile Telephones at 1.75 GHz," *IEEE Microwave and Guided Wave Letters*, Vol 3, No. 8, pp 268-270, Aug 1993. [Higher-order Class E with third-harmonic resonator (Class F3).]

29. R. M. Porter and M. L. Mueller, "High Power Switch-Mode Radio Frequency Amplifier Method and Apparatus," *US*

Patent 5,187,580, Feb 16, 1993. (Class E with substantial voltage at turn-on, as under "Optimizing Efficiency" here.)

30. Y-O Tam and C-W Cheung, "High Efficiency Power Amplifier with Traveling-Wave Combiner and Divider," *International Journal of Electronics*, Vol 82, No. 2, pp 203-218, 1997. (Class E 450 MHz/5 W with 89.4% collector efficiency. The outputs of four such amplifiers are combined with a traveling-wave power combiner, yielding 14.96 W output at 89.5% collector efficiency.)

31. J. E. Mitzlaff, "High Efficiency RF Power Amplifier," US Patent 4,717,884, Jan 5, 1988. [1.6 W at 76% drain efficiency at 840 MHz. This circuit exhibits at least 1.5 W output with at least 74% efficiency over a 50-MHz band centered at 840 MHz (6% bandwidth). Its operation is described as Class-F operation, but appears to be high-order Class-E circuit with lumped and transmission-line resonators. The patent shows transistor voltage and current waveforms for three "prior-art" circuits, but not for the circuit covered by this patent. It gives a detailed explanation of how to synthesize the load network to produce desired input-port impedance versus frequency.]

32. M. Kessous and J.-F. Zürcher, "Amplificateur VHF en Classe E Utilisant un Transistor à Effet de Champ (FET) VMOS de Puissance" (VHF Class-E Amplifier Using VMOS Power FET), *AGEN-Mitteilungen* (Switzerland), No. 30, pp 45-49, Oct 1980. [2.58 W output at 145 MHz with 96.5% drain efficiency, 81.3% total efficiency = $P_{out}/(P_{dc} + P_{drive})$, using Siliconix VMP-4.]

33. N. O. Sokal, "Design of a Class E RF Power Amplifier for Operation at 2.45 GHz, and Tests on a Scaled-Frequency Model at 122.5 MHz" (1/20 frequency), Oct 1979, unpublished report of Design Automation for Project 4198. [We used a Raytheon RPC3315 GaAs MESFET that

was intended for 2.45 GHz. Initial tests were made with frequency scaled-down by factor of 20, all inductance and capacitance (including transistor capacitances and expected wiring parasitic inductance) scaled-up by factor of 20. All resistances, voltages and currents were at intended final values. Results were 210 mW output, 77% drain efficiency, 24 mW input drive, 9.4 dB power gain, 71% overall efficiency = $P_{out}/(P_{dc} + P_{drive})$, 68% PAE.]

34. D. W. Cripe, "Improving the Efficiency and Reliability of AM Broadcast Transmitters through Class-E Power," National Association of Broadcasters annual convention, May 1992, seven pages.

35. S. Hinchliffe and L. Hobson, "High Power Class-E Amplifier for High-frequency Induction Heating Applications," *Electronics Letters*, Vol 24, No. 14, pp 886-888, July 7, 1988. (Greater than 550 W at 3-4 MHz with more than 92% efficiency across the band; 450 W at 3.3 MHz at 96% efficiency from 104 V dc, IRF450 MOSFET.)

36. R. Redl and N. O. Sokal. "A 14-MHz 100-Watt Class E Resonant (dc/dc) Converter: Principles, Design Considerations and Measured Performance," Power Electronics Conference, San Jose, California, Oct 1986. (This Class E dc/dc converter had 87% efficiency at 100 W dc output. An IRF540 RF power stage supplied an estimated 105 W at 91.4% efficiency because of an estimated 5-W loss in the coupling transformer and rectifier associated with the 100-W dc load.)

37. N. O. Sokal and Ka-Lon Chu, "Class-E power Amplifier Delivers 24 W at 27 MHz, at 89-92% Efficiency, Using One Transistor Costing $0.85," *Proceedings of RF Expo East*, Tampa, Florida, Oct 1993, pp 118-127, RF Expo West, San Jose, California, March 1993 (not in *Proceedings*). [This used International Rectifier (89%) and Harris Semiconductor (92%) IRF510

SMPS MOSFETs. The Harris device had a slightly larger die, lower $R_{DS(on)}$ and higher efficiency. The silicon-gate R_g (about 1-2 Ω, but never specified) was borderline-acceptable at 27.12 MHz for $i_g^2 R_g$ input-drive power. P_{drive} varies as f^2; performance would have been quite acceptable at 13.56 MHz.]

38. N. O. Sokal and I. Novak, "Tradeoffs in Practical Design of Class-E High-Efficiency RF Power Amplifiers," *Proceedings of RF Expo East*, Tampa, Florida, Oct 1993, pp 100-117, and RF Expo West, San Jose, California, March 1993.

39. P. J. Poggi, "Application of High Efficiency Techniques to the Design of RF Power Amplifier and Amplifier Control Circuits in Tactical Radio Equipment," *Proceedings of MILCOM '95*, San Diego, California, Nov 5-8, 1995, pp 743-747.

40. S. C. Cripps, *RF Power Amplifiers for Wireless Communications*, (Norwood, Massachusetts: Artech House, 1999), ISBN 0-89006-989-1, pp 170-177. [Fig 6.19 on p 176: GaAs MESFET, 840 MHz, 79% efficiency at 1.24 W output, 15 dBm (31.6 mW) input, power gain 1.24 W/0.0316 W = 39.2 amplification factor, 15.9 dB.]

41. M. D. Weiss, M. H. Crites, E. W. Bryerton, J. F. Whitaker, and Z. Popovic, "Time-Domain Optical Sampling of Switched-Mode Microwave Amplifiers and Multipliers," *IEEE Transactions on MTT*, Vol 47, No. 12, pp 2599-2604, Dec 1999.

42. M. D. Weiss and Z. Popovic, "A 10 GHz High-Efficiency Active Antenna," *1999 IEEE MTT-S International Microwave Symposium Digest*, June 13-19, 1999, Anaheim, California. It is available as file TU4B_5.PDF on CD-ROM IEEE Catalog No. 99CH36282C.

43. E. Lau, KE67VWU, et al, "High-Efficiency Class-E Power Amplifiers," *QST*, Part 1 May 1997, pp 39-42; and Part 2 Jun, pp 39-42. □□

A Keyed Power Supply for Class-E Amplifiers

*Supply 500 W for your Class-E power amplifier
with a compact switcher that costs under $100.*

By Jim Buckwalter, KF6SWC; John Davis, KF6EDB; Dragan S. Maric;
Kent Potter, KC6OKH; and David Rutledge, KN6EK

Power amplifiers have always
been popular projects for Ama-
teur Radio operators.[1, 2, 3, 4] Ac-
cordingly, interest in power supplies
for these amplifiers continues to
grow.[5, 6, 7] Switching power supplies
are much lighter and more compact
than conventional linear power sup-
plies, and potentially less expensive.
Switching supplies offer high efficien-
cies and require smaller transformers,
allowing for a compact design. The
disadvantage of these circuits is the
RFI generated by the switched-mode

[1]Notes appear on page 26.

Jim Buckwalter, KF6SWC,
jimb@ece.ucsb.edu
John Davis, KF6EDB,
John.F.Davis@jpl.nasa.gov
Dragan Maric, DMARIC1@irf.com
David Rutledge, KN6EK,
rutledge@caltech.edu
Kent Potter, KC6OKH,
c/o David Rutledge

operation of the supply. However, the
design presented here reduces this
problem with filters and demonstrates
the potential of switching power sup-
plies for Class-E amplifiers.

This supply design operates at
100 kHz and can provide up to 500 W
to an amplifier in CW operation (see
Note 1). At key-down, the power out-
put ramps linearly from zero volts to
full output—between 40 and 120 V—
in 7 ms. The supply operates with an
efficiency of 85%. It does not require
a fan for CW operation, but would
require cooling for FSK. The ripple is
around 1 V at full output power, or
1% of the dc output voltage. The sup-
ply weighs only three pounds and
costs about $100. These characteris-
tics offer an attractive design for the
power requirements of a Class-E
amplifier.

Amplifier Developments

In 1997, Eileen Lau, KE6VWU, and
Kai-Wai Chiu, KF6GHS, introduced

high-power, Class-E amplifiers to the
amateur community that they had de-
veloped in undergraduate projects at
Caltech (see Note 1). These amplifiers
produce up to 500 W on the 40-meter
band and use the popular NorCal40A
QRP kit as the driver. Class-E amplifi-
ers have extremely high efficiencies, in
the range of 90%, so that they can oper-
ate without a cooling fan. In 1998,
Raab and Roberts developed a 700-W,
160-meter amplifier (see Note 2). In
1999, Der Stepanians and Rutledge
built a 200-W, 30-meter amplifier (see
Note 3). These amplifiers take advan-
tage of the power MOSFETs commonly
used in power supplies, making the am-
plifiers quite inexpensive. A complete
set of parts for the Caltech amplifiers is
available for only $50 (see Note 4). For
CW operation, however, the amplifiers
need a keyed 120-V power supply. The
original power supply used by Lau *et al*
was heavy and expensive because it had
an isolation transformer, a variable
auto transformer and a large choke (see

Note 5). Recently, switching power supplies have been developed for amateurs to build by Manfred Mohinweg, XQ2FOD (see Note 6). Our supply is designed to be simple to construct for an amateur with little experience in power-supply design.

Circuit Description

The design of an off-line, isolated, switched-mode supply involves two separate circuits: an ac/dc rectifier and a dc/dc power converter. Control and feedback circuits are also included in the design, as shown in Fig 1.

The power supply is intended for operation from a 120-V, 60-Hz wall supply. A power-entry module (P1) and common-mode choke (consisting of C1, C2 and L1) reduce harmonics coupled to the power line and avoid disrupting the operation of other circuits and equipment (see Fig 2). A 6-A fuse protects the operator and circuit from accidental failures. A bridge rectifier, U1, converts the 60-Hz ac to an unregulated dc voltage. For primary dc filtering, two 2200-μF electrolytic capacitors, C3 and C4, reduce the input-voltage ripple entering the power converter to about 7 V. The output voltage ripple is further reduced with feedback.

We must calculate the copper and ferrite losses of the transformer. Ideally, a well-designed transformer will have equal copper and ferrite losses to give an optimized total loss. In this case, the ferrite losses were 0.76 W and the copper losses were 1.03 W, for 1.79 W of dissipation. The transformer's efficiency is partly a result of the circuit topology, which maintains a zero-dc balance across the transformer; the transformer does not need to be gapped. The design procedures and equations are given at the end of this paper.

Since the switching action of the MOSFETs cuts off the current through the transformer abruptly, leakage in-

ductance in the transformer generates voltage spikes during the time intervals when both transistors are off. These spikes can couple into the control circuitry and may destroy the MOSFETs. To reduce this spiking, a lossy snubber network (R1 and C7) is placed in parallel with the transformer.

A full-wave diode pair, U2, rectifies the transformer voltage. Again, resistive snubber networks are added across the anode and cathode to prevent the inductive voltage spikes from destroying these devices. L2 (250-μH) and C12 (10 μF) form a second-order filter to remove switching harmonics, while providing a reasonable keying response. The dc resistance of the Caltech Class-E amplifiers is between 25 and 50 Ω when keyed. R6 and C13 dampen the transient response of the converter and ensure stability over the range of output voltages.

The controller circuit chosen for this project in the Motorola MC34025, a dual, totem-pole output, pulse-width-modulation chip. The chip, U4, incorporates an error-amplifier and soft-start mechanism requiring as little external circuitry as possible, and it costs under $4. Switching frequency is determined by R13 and C20. Pulse transformers T2 and T3 isolate the control circuit from the primary side of the converter. Feedback is used to allow the output to follow the keying signal while reducing the output voltage ripple at 120 Hz. The feedback compensation consists of a voltage divider (R6, R7, C12 and C13) and an op amp (U3) with gain determined by R9, R10 and C18. The internal op amp of the controller chip is configured as a differential comparator with R11, R12, R13 and R14.

The dc/dc half-bridge converter consists of two power capacitors (C5 and C6) and two MOSFETS (Q1 and Q2) placed in a bridge configuration with a power transformer connected between the capacitors and the MOSFETs. The

Fig 1—Block diagram for the keyed switching power supply.
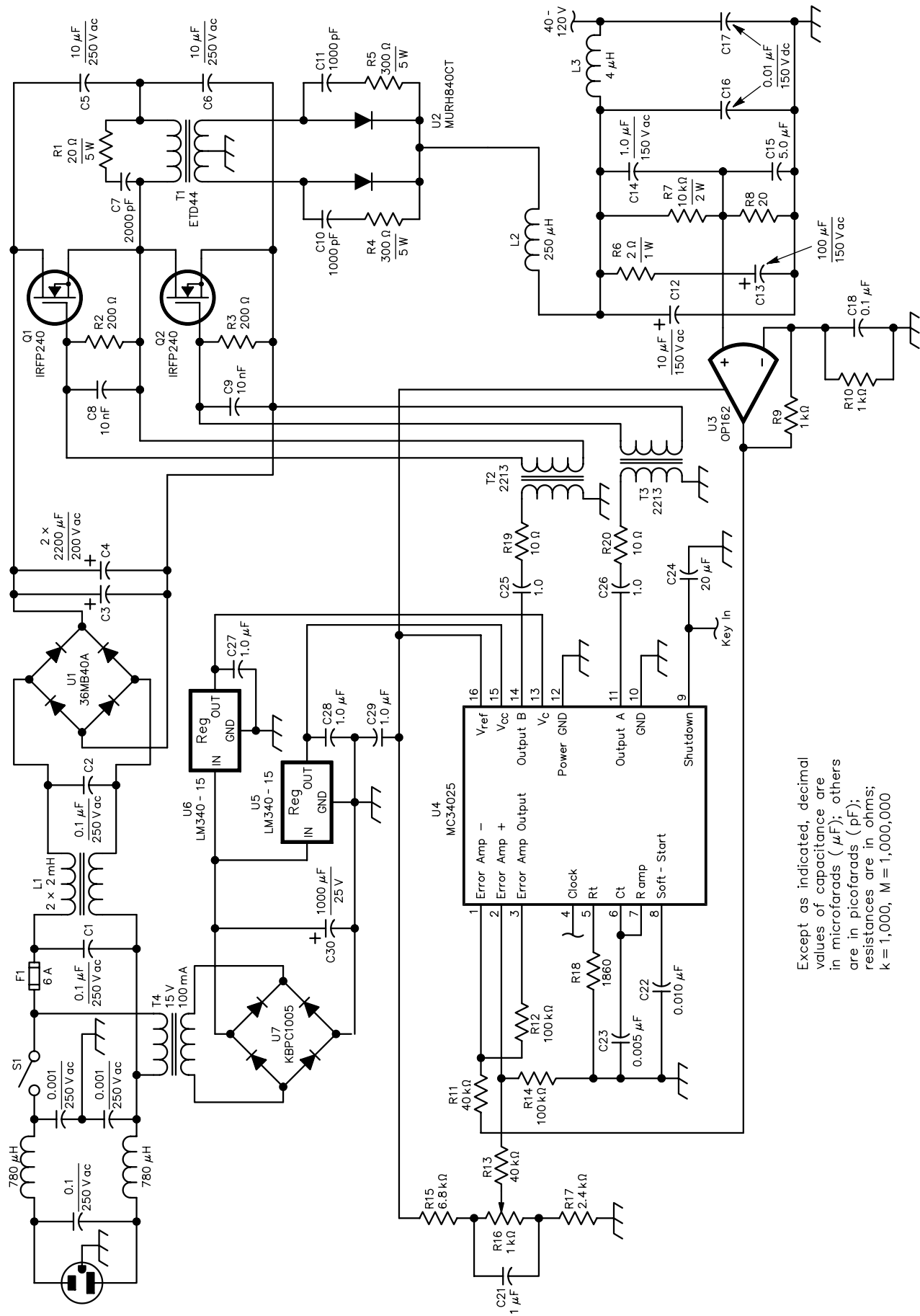
Except as indicated, decimal values of capacitance are in microfarads (μF); others are in picofarads (pF); resistances are in ohms; k = 1,000, M = 1,000,000

MOSFETs chosen for his project are IRFP240s. This MOSFET has a maximum V_d of 200 V, and it handles a drain-to-source current of 20 A. The high-speed MOSFETs generate interfering harmonics, but proper EMI filtering in the output stage of the supply reduces these.

The advantage of the half-bridge converter is that current flow from the hot line to the neutral line is through a power capacitor, preventing saturation of the transformer core. The high current, however, stresses the capacitors in the bridge network. Film capacitors are chosen because of their low equivalent series resistance. A dc resistance is sometimes placed in parallel with C5 and C6 to ensure dc balance at the transformer. In this design, we have relied on our PWM controller to provide the balance and removed the resistors to improve the efficiency of the converter. These resistors can be easily added with attention paid to power loss.

The transformer design follows the convention outlined in Robert Erickson's *Power Electronics*.[8] An ETD-style core is employed because the round cross-section of the central arm allows the bobbin and wire coils to fit closely against the magnetic material (see Fig 3). This reduces leakage inductance and copper loss. The turns ratio is chosen to provide the correct maximum output voltage when the converter operates at maximum duty cycle. To develop 120 V at the output, a turns ratio of 2.5:1 is required. The design of the transformer requires a center-tapped secondary winding (Fig 2). The switching frequency, input voltage, and material loss characteristics determine the core geometry of the transformer. This allows us to choose a core, in this case the ETD-44, which will handle the power requirements. From the physical specifications of the core size, we may calculate the peak flux density in the magnetic material. For the ETD-44, the flux density is calculated to be about 0.12 teslas (T). From this flux density, we can derive the proper wire size and number of turns. The primary needs 15 turns of #14 AWG magnet wire and the secondary needs 37 turns of #18 AWG. Litz wire was not used in this design to simplify construction. With the copper requirements reached, it was not deemed necessary.

Performance

The power supply was tested with the 7-MHz, 500-W amplifier design developed by Lau *et al* (Note 7). To analyze the impact of the supply on the amplifier, the RF output was measured at the carrier frequency, 7 MHz. The voltage rises to the output level in about 7 ms for 10 nF at C22 and 22 µF at C24. C22 controls the turn-on time of the chip while C24 creates a delay for the keying signal. The supply oscillator is not audible when the supply and amplifier are not transmitting.

Fig 3—The ETD44 can isolate 500 W at a switching frequency of 100 kHz. The round central arm of the core improvres efficiency. Each piece of the core measures 4.4×1.5×2.25 cm.
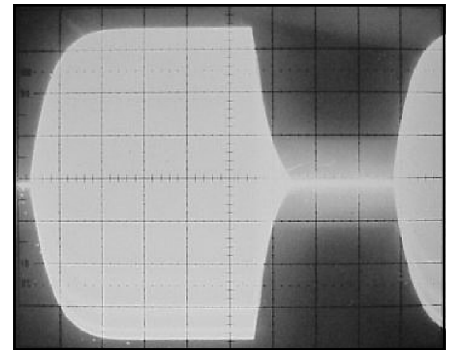
Fig 4—A keying waveform with a 7-MHz class-E amplifier at 500 W output. The horizontal axis is 10 ms per division. The rise and fall times are about 10 ms but can be changed with C22 and C24.
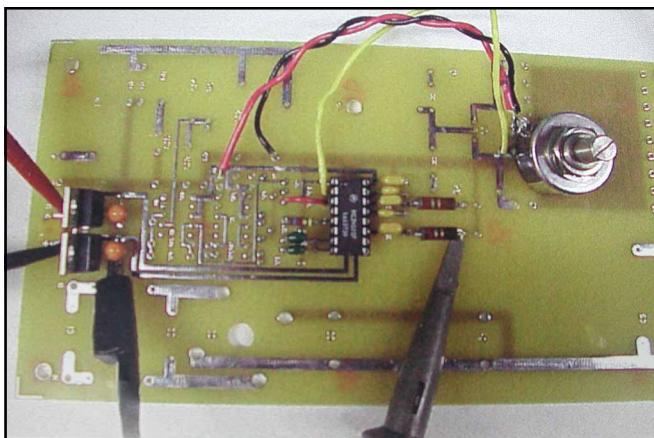
(A)
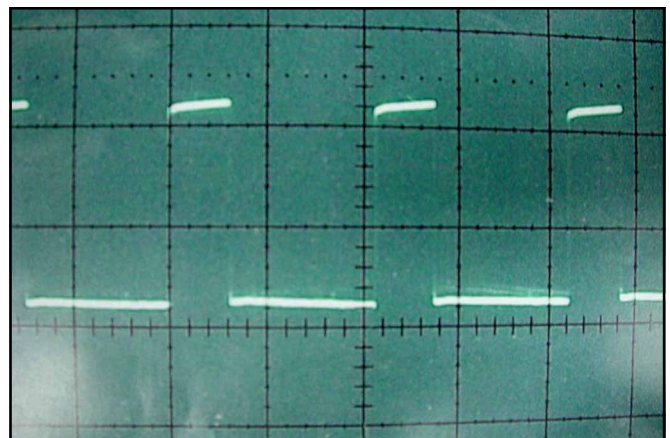
(B)

Fig 5—Pulse-width-modulator (PWM) circuit construction and testing. The horizontal axis of the waveform is 5 ms per division. The vertical axis is 5 V per division.

Construction and Tune-Up

The power circuit is constructed piece-wise, allowing the builder to test blocks of the circuit as they are constructed. Construction of this project requires a multimeter, oscilloscope, frequency counter, 17-V dc supply and a Variac. The PWM circuit is an appropriate place to start. Place the PWM chip, U4, on a solder-tail header. R18 and C23 determine the switching frequency. Use a 2-kΩ potentiometer for R18 to allow adjustment of the switching frequency. C22 determines the soft-start speed of the chip's turn-on; 10 nF gives a rise time of about 7 ms. Install feedback components R11 through R14 and C20. R15, R16 and R17 control the PWM output. At this point, place a jumper wire across C24 to ground pin 9 of U4. Install bypass capacitors C27, C28 and C29. Install the gate drive components, C25, C26, R19 and R20.

Install voltage regulators U4 and U5 with C30 and C31 at the output of the regulators to improve transient response. Fig 4 shows what the PC board looks like at this point. Currently, the 17-V source will power the control circuitry. The operation of the PWM circuit can be verified by inspecting the output pulses with an oscilloscope at R19 and R20. The waveforms at each pin should look similar, with amplitude of about 12 V and a frequency of 100 kHz. Varying the control potentiometer should change the pulse width from 10% to 50% of the switching period.

Build the half-bridge circuitry next. Carefully bolt (#6-32 hardware) MOSFETs Q1 and Q2 be to a heat sink with a Kapton pad, to prevent electrical connection between the FETs and the heat sink. The heat sink is 3×4$^1/_8$ inches, with 1-inch fins (HS50-3.0 from RF Parts, with a thermal resistance of 2°C/W and no fan). The placement height of the MOSFETs depends on your chassis, and the PC board determines the distance between the MOSFETs. The Kapton pad, manufactured by Bergquist[9], has a thermal resistance of 0.2°C/W and a 6-kV breakdown rating. Poor thermal connection will lead to early device failure. The MOSFETs will be soldered to the board once the board is fitted to the chassis. Install snubber circuits R2, R3, C8 and C9, as well as the isolation transformers T2 and T3. Construct the isolation transformers on two 2213 pot-core bobbins. The primary and secondary sides should be wrapped with 30 turns of #24

Fig 6—Construction of the center-tapped secondary coil. Two #18 AWG coils are connected in series.
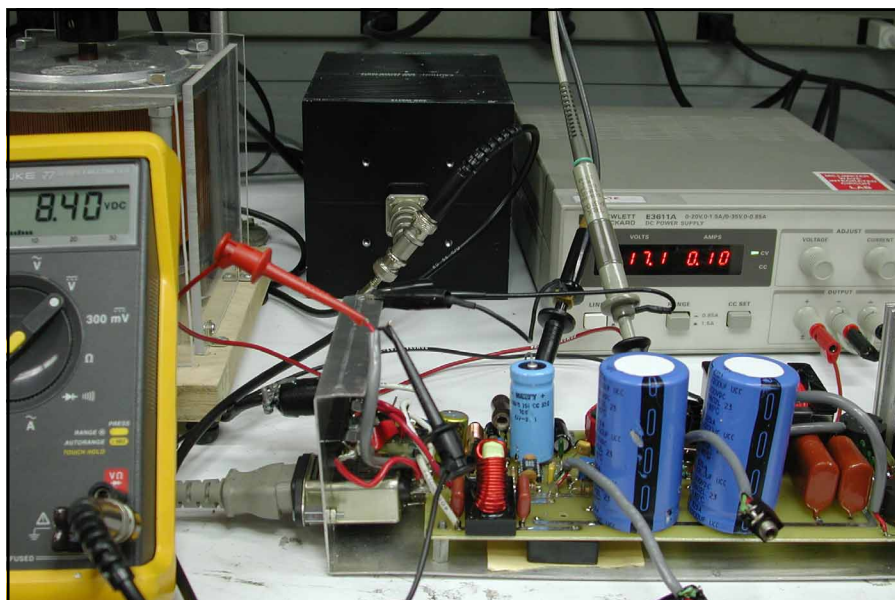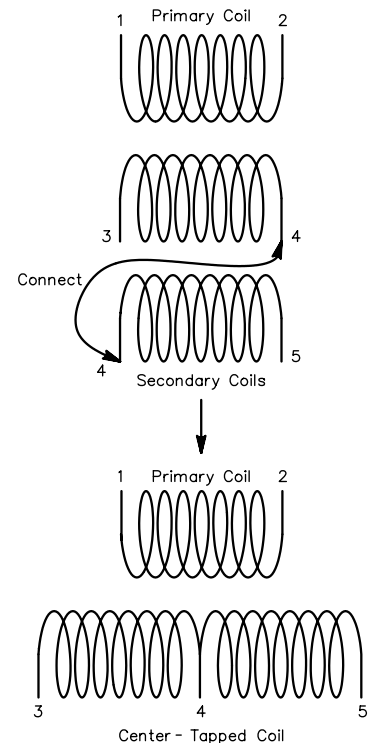
Fig 7—Low-power tests of the converter. At left, the setup for the test. Above, the waveforms of the inductor input voltage and the output voltage ripple. The horizontal axis is 2 μs per division. The vertical axis is 10 V per division for the upper trace and 0.2 V per division for the lower trace.

AWG magnet wire. Use two four-foot lengths of the wire, and wrap the primary and secondary coils simultaneously. Trim the wires when the wrapping is complete, then remove the enamel at the ends of the wire.

Solder the transformers in place ensuring that the polarities of the primary and secondary coils agree. Bolt the transformers to the board with nylon screws and nuts. The complete gate-drive circuit can now be checked. After the PWM drives are connected to the pulse transformers with #22 AWG coated wire, the waveform across the gate and source pins of the MOSFET should look similar to the waveform observed when the drive circuit was not connected. Check that the pulse heights are greater than 10 V and that the pulse width is never greater than 50% of the switching period by varying the control potentiometer setting. Further, the circuit should draw about 100 mA at maximum pulse width.

Install C1, C2, L1 and U1 and ensure that U1 is connected according to the polarity of the bridge rectifier. Install C3, C4, C5 and C6 to complete the half-bridge circuit.

Construction of the main isolation transformer (T1) is the most time-consuming element. The primary coil should be a single layer, close-wound, with 15 turns of #14 AWG magnet wire on the ETD-44 bobbin. Magnet wire was chosen over Litz wire to simplify the design. Secure the winding by wrapping electrical tape over the primary turns. Notice that the half bridge requires two secondary coils with a 2.5:1 turns ratio. The secondary coils should be wound with 37 turns of #18 AWG magnet wire. With two five-foot pieces of the wire, wrap both coils at the same time. Snip the excess wire and remove the protective coating on the magnet wire. The secondary side of the transformer should have two separate windings at this point. Since the circuit requires these two coils to be formed into a center-tapped configuration, use a multimeter to identify the separate coils. Connect the coils in series as shown in Fig 6. Again, wrap the visible wire with electrical tape. Note the orientation of the transformer on the PC board and ensure that the coil wires will extend to the proper connections on the PC board. This connection forms the ground of the secondary circuit. Insert the magnetic pieces into the bobbin and attach them with the metal springs.

Attach the secondary components. Install the diode pair (U2) and the snubber circuit consisting of R4, R5, C10 and C11. U2 will need a small clip-on heat sink. Output inductor L2 is wound on a 2616 pot core. The inductor core is gapped to prevent the dc current from saturating the core. The core is wound with 27 turns of #18 AWG wire. Bolt the inductor to the PC board with a nylon screw and nut. Install output-filter capacitor C12 along with R6 and C13. Finally, install the feedback-compensation components, R7, R8, C14 and C15, as well as the RF filter, C16, C17 and L3.

At this point, the circuit needs to be set into a chassis. The power-entry module is bolted to the chassis as well as a fuse holder and a Cinch-Jones female connector for connection to the amplifier. The board is bolted to the chassis in all corners. The heat sink is mounted so that the MOSFETs fit into their footprints on the PC board. Once the chassis and heat sink are secure on the chassis, solder the MOSFETs to the board. Connect the power-entry module and the Cinch-Jones connector to the board with #14 AWG wire.

The circuit is now ready for an open-loop test. Since the power supply is not a resonant device, there is not a lot of "tuning" to be done. Testing the device is critical, however. A small mistake may result in a destroyed board if the unit is plugged directly into the ac line. Use a Variac to slowly increase the input amplitude of the 60-Hz power. Connect a 25-50 Ω resistor to the output. Be sure of the power rating of your load and test the supply only to the load's power rating. If you do not have a suitable load, order a 25-Ω, 5-W resistor. This allows you to test to about 10 V.

Connect the 17-V supply to the control circuit and increase the Variac setting to 5 V (RMS). Use an oscilloscope to observe the waveforms at the input and output of L2. You should see a series of rectangular pulses with overshoot on the pulses. At the output of the inductor, the pulses should disappear and the dc output should be clear. The switching ripple is apparent if the voltage scale is reduced. Varying the duty cycle changes the output voltage.

Once this test has been satisfied, install the remaining components. T4 is a 60-Hz transformer that can deliver 15-20 V at 100 mA to the control circuit. Install U7 and C32 to rectify and filter the 60-Hz ac. Next, incorporate the feedback into the design. Install U3, R9, R10, C18 and C19. The circuit now operates as a closed-loop system. Set the Variac to about 100 V (RMS).

Avoid running the converter for long periods in continuous operation. Monitor the 120-Hz voltage ripple and ensure that it is about 1 V, maximum, over the range of operation. If all the previous checks have been confirmed, remove the jumper across C24 and install the stretcher circuit presented by Lau *et al* (Note 5). The stretch time required for proper operation is about 4 ms. Attach two #22-AWG wires to a jack for your key. You can connect the power supply directly to the 120-V, 60-Hz line with the amplifier as the load. Ensure that the positive and negative connections of the supply and amplifier are correct. The control pot allows output voltages between 40 and 110 V at 25 Ω.

Acknowledgement

Notes
[1]E. Lau, K-W. Chiu, J. Qin, J.F. Davis, K. Potter, D. Rutledge, "High-Efficiency Class-E Power Amplifiers: Part 1," *QST*, May 1997, pp 39-42.
[2]T. Roberts and F. Raab, "Class-E Power Amplifier and Digital Driver for 160 Meters," *Communications Quarterly*, Fall 1998, pp 9-19.
[3]T. der Stepanians and D. Rutledge, "10-MHz Class-E Power Amplifier," *Communications Quarterly*, Winter 1999, pp 31-39.
[4]A package of amplifier parts, including PC board, components, connectors, heat sink and chassis is available at cost from Puff Distribution, Department of Electrical Engineering, MS136-93, Caltech, Pasadena, CA 91125. Price: $50 for US order, $60 foreign. The price includes sales tax and shipping by surface mail. Payments must be by check or money order to "Caltech-Puff Distribution." Foreign checks must be drawn on a bank with a US branch office. Please provide your Amateur Radio call sign and specify which amplifier (300 or 500 W) you want. For more information, contact Dale Yee at **yee@systems .caltech.edu**, or fax a form from **www .systems.caltech.edu/EE/Faculty/ rutledge/puffform.html**.
[5]E. Lau, K.-W. Chiu, J. Qin, J. F. Davis, K. Potter, D. Rutledge, "High-Efficiency Class-E Power Amplifiers: Part 2," *QST*, June 1997, pp 39-42.
[6]M. Mohinweg, "A 13.8-V, 40-A Switching Power Supply," *QST*, Dec 1998, pp 37-41.
[7]W. Sabin, "Power Supply for a MOSFET Power Amplifier," *QEX*, Mar/Apr 1999, pp 50-54.
[8]R. W. Erickson, *Fundamentals of Power Electronics*, (New York: Chapman & Hall, 1997).
[9]The Bergquist Company, 18930 West 78th St, Chanhassen, MN 55317; tel 952-835-2322, 800-347-4572, fax 952-835-4156 (sales), 952-835-0430 (corporate); e-mail **madelinef@bergquistcompany.com**; **www.bergquistcompany.com/**.

Appendix: Transformer Design Details

Design Variables	Equations	Values
Copper resistivity, ρ		1.724×10^{-6} Ω-cm
Turns ratio, n		2.5
Duty cycle, D	$\dfrac{V_{in} D}{2 f_s}$	0.9 (max)
Primary V–s, λ		675 µV-s
Where f_s = switching frequency		

$$I_{primary} + n \bullet I_{secondary} (\text{RMS})$$

$$I_{primary} = (I_{load})(n)(D)^{\frac{1}{2}}$$

I_{tot} 16.5 A

$$I_{secondary} = \frac{(I_{load})(n)(1+D)^{\frac{1}{2}}}{2}$$

P_{diss} estimate		1.8 W

P-material Specifications

Window-fill factor, K_u		0.4
Core-loss exponent, β		2.7
Core-loss coefficient, K_{fe}		11.98 W/(cm³T$^\beta$)

Core Size Requirements

Core geometry size, K_{gfe}	$\dfrac{\rho(\lambda \ I_{tot})^2 (K_{fe})^{\frac{2}{\beta}}(10^8)}{4 K_u (P_{tot})^{\frac{2+\beta}{\beta}}}$	0.030

ETD 44 Specifications

Core cross-sectional area, A_c		1.74 cm²
Core window area, W_a		2.13 cm²
Mean length per turn, MLT		7.62 cm
Magnetic path length, l_m		10.3 cm

Design Equations

Peak flux density, B_{max}	$\left(\dfrac{\left(\rho(\lambda \ I_{tot})^2 (MLT)(10^8)\right)}{2(K_u)(W_a)(l_m)(K_{fe})(\beta)(A_c)^3} \right)^{\frac{1}{\beta+2}}$	0.124 T

(P material saturates at 0.2 T and B_{max} is well below this point.)

Primary turns, $n1$	$\dfrac{\lambda (10^4)}{2 B_{max} A_c}$	15.7 turns
Secondary turns, $n2$	$(n)(n1)$	38 turns
$\alpha 1$	$\dfrac{I1}{I_{tot}}$	0.58
$\alpha 2$	$\dfrac{(n)(I2)}{I_{tot}}$	0.42
Primary wire size, A_{w1}	$\dfrac{(\alpha 1)(K_u)(W_a)}{n1}$	0.0314 cm²
Secondary wire size, A_{w2}	$\dfrac{(\alpha 2)(K_u)(W_a)}{n2}$	0.0101 cm²

Unfortunately, A_{w1} is rather large and AWG #14 with 0.0202 cm² is chosen (while accepting greater losses). A_{w2} is correspondingly scaled to AWG #18 with 0.0082 cm².

Power Dissipation Check

P_{fe}	$(K_{fe})(B_{max})^\beta (A_c)(l_m)$	0.76 W
P_{cu}	$\dfrac{(\rho)(MLT)(n1 \bullet I)^2}{K_u W_a}$	1.03 W
P_{tot}	$P_{fe} + P_{cu}$	1.79 W

More on Atmospheric Ozone and Low-Frequency Propagation

Explore how LF radio signals reach your station—through chemistry and sunlight.

By Robert R. Brown, NM7M

Ozone is a trace constituent in the lower atmosphere, but it is formed after the photo-dissociation of oxygen molecules above 25 km and then carried downward by atmospheric mixing processes. Beyond that, ozone is transparent to visible light, but it shows very strong absorption in the ultraviolet (UV) part of the spectrum and is responsible for keeping harmful solar radiation from reaching the earth's surface. That absorption contributes to the heating of the "statosphere" and the rise in temperature with height in the upper stratosphere. Also, ozone is of meteorological importance as it is may serve as a tracer to show atmospheric circulation.

A recent article[1] pointed out that atmospheric ozone also represents a meteorological factor that affects low-frequency and 160-meter propagation. In that study, 55.5-kHz signals were monitored daily, and the interference of the sky wave and ground wave on a one-hop path was used to show the lowering of the LF reflection region at sunrise. That transition starts from a night-time position at about 90 km, determined by the formation of negative ions in the lower D-region and goes to the day-time level below 75 km with full solar illumination.

The LF study found the transition to be delayed by about 15-20 minutes from what one would expect if visible radia-

tion were considered responsible for detaching electrons from the negative ions at the start of the sunrise descent. That means the shadow cast by the solid earth is not responsible for the delay in the build-up of ionization in the lower D-region. Instead, the delay can be understood in terms of the opacity of the ozone layer to UV radiation and the release of electrons from negative ions that results when UV finally reaches the region with the lowering of the ozone shadow at sunrise.

This article summarizes other features that were brought out when a year-long LF study, March 1998 to March 1999, was completed. In addition, further discussion is given regarding other aspects of 160-meter propagation resulting from negative ions in the lower D region.

1105 27th St #AW110
Anacortes, WA 98221
bobnm7m@cnw.com

[1]Notes appear on page 36.

Experimental Details

The present study used 55.5-kHz signals on a N-S path from the Navy Station NPG at Dixon, California (38.4° N, 121.9° W) to Guemes Island, Washington (48.5° N, 122.6° W). The signals consisted of a one-hop sky wave and a ground wave. Recordings showed the result of interference between the two waves. As such, the recordings depended on the amplitudes of the sky waves and ground waves as well as their initial phase difference, while the time variations in signal strength resulted from the phase change between the two waves as the reflection region of the LF sky wave descended from about 90 km to 75 km with sunrise on the region.

The nighttime LF reflection region is characterized by a large decrease in electron density, some three orders of magnitude[2] descending from 90 km to 60 km. Of particular importance to the present discussion is the ledge between 90 km and 80 km. There the decrease in electron density is about an order of magnitude in a few kilometers, a distance less than one full wavelength (5.5 km) at 55.5 kHz. The decrease in electron density is attributed to the formation of negative ions at the lower altitudes and electrons are freed again by photo-detachment when the sun rises on the region.

Being on a small island, the receiving site was relatively free from man-made noise. There were a few instances when the recordings were not satisfactory, due to atmospheric noise as well as line noise and occasional power failures from wind storms. The records (data points every 15 seconds for about 15 minutes) typically consist of a sunrise signature, which is a decrease in signal strength, followed by a near-complete recovery (shown in Fig 1).

The data in Fig 1 are from June 15, 1998 and illustrate the features noted above. The decrease in signal strength extended over about 15 minutes, from 1150 UTC to 1205 UTC instead of from 1125 UTC, when the start of sunrise at the 90 km level would be expected from the shadow of the solid earth. In that time, the solar depression angle at the midpoint of the path (43.45° N, 122.2° W) went from about −6° to −4° and the signal variation amounted to −2.0 dB. In the year of data recording, however, there were both faster and slower sunrise signatures, as shown in Figs 2 and 3. Those are due, in part, to seasonal differences in the rate of change of the solar altitude seen at D-region heights. Over a year's time the signal-strength variations at sunrise ranged from −0.4 dB to −3.7 dB, with an average of −1.3 dB.

Interpretation

As noted above, the timing of the sunrise signatures can be understood from the ozonosphere controlling the time when the LF reflection region is lowered with sunrise. Turning to details of the interpretation, the vertical distributions of ozone discussed in the literature are quite varied, ranging from
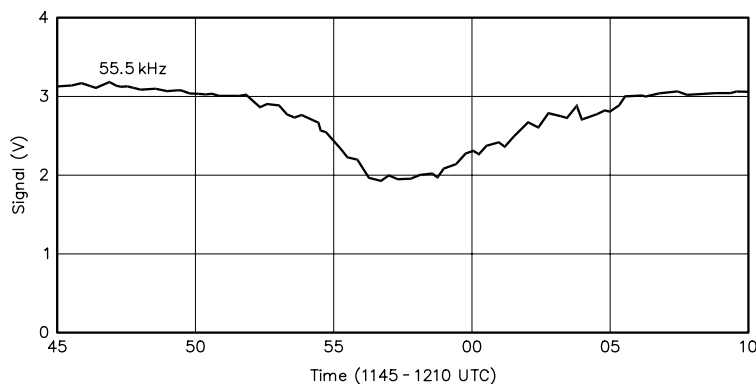
Fig 1—Sunrise signature of NPG on June 15, 1998.
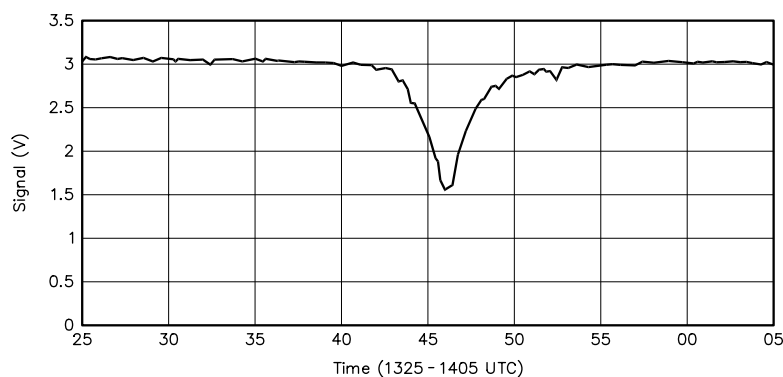
Fig 2—An example of a fast sunrise signature on October 2, 1998.
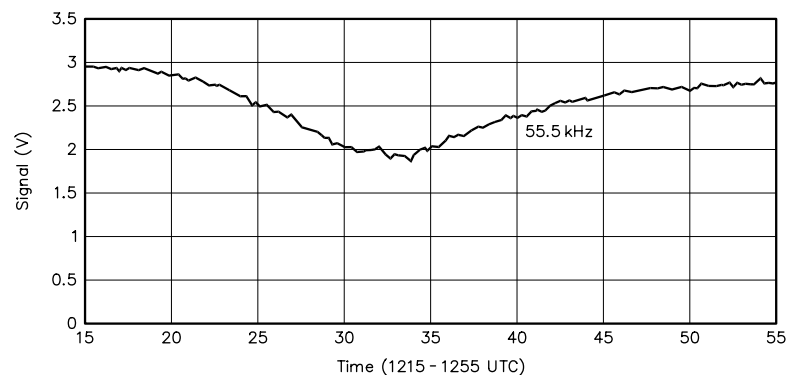
Fig 3—An example of a slow sunrise signature on May 7, 1998.

those having a well-defined peak to others, on occasion, with a broad, flat maximum. For the present discussion, the form of a Gaussian function was used as a reasonable representation of the ozone distribution (see Note 2). It assumes a maximum density at a 25-km altitude and a standard or mean-square deviation of 10 km from the location of the peak value. Thus, from the known properties of the Gaussian function, 68% of the ozone content would lie within ±10 km (1.0 standard deviation) of the peak of the ozone distribution. By the same token 95% lies within ±20 km (2 standard deviations) of the peak of ozone concentration.

The total ozone content in a vertical column of 1 cm^2 cross-section can be expressed in terms of the thickness that the layer would occupy if the pressure and density were reduced to standard values (NTP) throughout the layer. The meteorological unit is termed a *Dobson*, 0.001 atm-cm, and 300 DU (Dobson Units) corresponds to an ozone layer of 3-mm thickness. In that regard, the annual variation of the ozone content at midlatitudes[3] ranges from about 365 DU in the spring to 285 DU in the fall, with an average of about 325 DU. In addition, it should be noted that the deviations of daily values of ozone content from the monthly means have been observed[4] to be both large and striking.

For the matter at hand, the ozone content in the average vertical atmosphere is sufficient to block solar UV below 300 nm wavelength from reaching ground level. That being the case, an ozone column along an oblique line of sight, from the sun to the LF reflection region, was considered to be opaque to solar UV if an integration of its density along the line of sight gave a content equal to that for a vertical atmosphere (325 DU). That method was used to examine the role of atmospheric ozone by using the Earth and a spherical ozone layer, along with a straight-line solar-ray path penetrating the ozone layer, as in Fig 4. While not to scale, the figure illustrates how ray paths pass through different densities and regions in the distribution, depending on the distance of approach of the ray to the earth.

The same physical situation may be represented in plane geometry by transforming the straight ray paths to curved paths relative to a plane ozone layer and in 111-km steps along the earth's surface. One example is given in Fig 5; that figure is for ray paths to the 90-km level, the altitude where sunrise starts at the midpoint of the path for

NPG signals. It shows changes in the ray paths for solar angles going from –10° to –4°relative to the horizon, corresponding to the advance of time toward ground sunrise (0°).

The only detail of the ozone layer given in that figure is a horizontal line at the height of the peak, taken as 25 km altitude and extending over a considerable horizontal distance. The
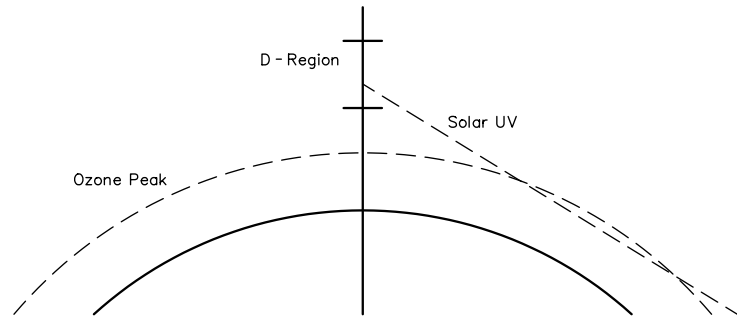
Fig 4—Illustration (not to scale) showing solar UV blocked from the D-region by the ozone layer.
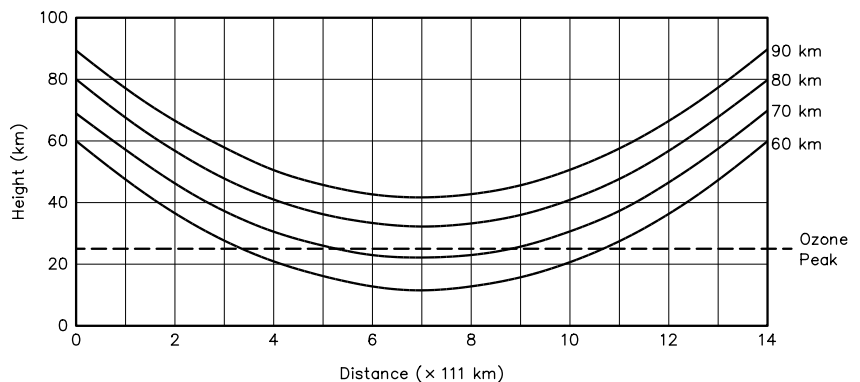
Fig 5—Solar-ray paths at 7° solar depression when transformed to plane geometry. The height for peak ozone density is given by the horizontal line.
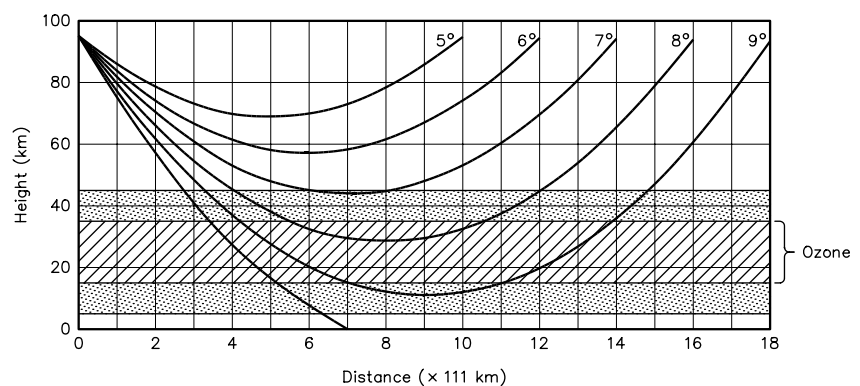
Fig 6—A more detailed, schematic representation of the ozone distribution. The diagonal pattern shows where the bulk (68%) of the ozone layer is located, and the two other regions each include about 14% of the layer.

figure also shows how the ray paths cross the level of peak concentration along the horizontal extent of the distribution. The sunrise effects are better understood if a more detailed, schematic representation of the ozone distribution is used, as shown in Fig 6. There, the diagonal pattern shows where the bulk (68%) of the ozone layer is located, at altitudes between 15 and 35 km. The two other regions with a vertical pattern are more-dilute portions of the layer, and each includes about 14% of the layer.

Like Fig 5, Fig 6 shows the ray paths for solar depression angles going from –10° to –4° toward sunrise (0°). The short ray path for –10° is given without markers in the figure and is blocked by the earth at about 777 km from the left of the figure. The other ray paths are later in real-time, with paths for –9° to –8° passing through the denser portions of the ozone layer while the path at –7° goes through the more-dilute portions of the ozone layer. Similar figures for 80 km, 70 km and 60 km show the same features, except that the paths that pass through the more-dilute portions of the layer are later in time (corresponding to angles of –6° and –5°, respectively) and the distances of closest approach to the Earth are shifted to the left of the figure by 111 and 222 km, respectively.

At this point, it is important to notice that in addition to the daily sunrise signatures, roughly 10% of the days showed other significant signal changes about 15-20 minutes before the peak of the daily sunrise signatures. Those signal variations were faster than typical sunrise signatures, taking about two minutes to reach peak values and decaying in a comparable period. The size of the variations ranged from being just detectable to – 2.2 dB and, on occasions, they were sometimes noted on successive days. In addition, their distribution in solar-depression angle ranged from –7° to –8°, with an average of –7.5°, as shown in Fig 7. A typical example of a precursor is shown in Fig 8, and a precursor in association with its own sunrise signature is shown in Fig 9.

In regard to their magnitudes, they were comparable to longer events attributed to the lowering of the LF reflection region by photo-detachment and are presumed to be of the same origin.

Ozone Distribution

Earlier, Reid[5] questioned whether the ozone layer played a role in sunrise/sunset variations of ionospheric

absorption during Polar Cap Absorption (PCA) events. There, extensive calculations were carried out and ultimately showed that photo-detach-ment of electrons from the negative ion of molecular oxygen, in conjunction with the shadow of the Earth, could not account for the variations
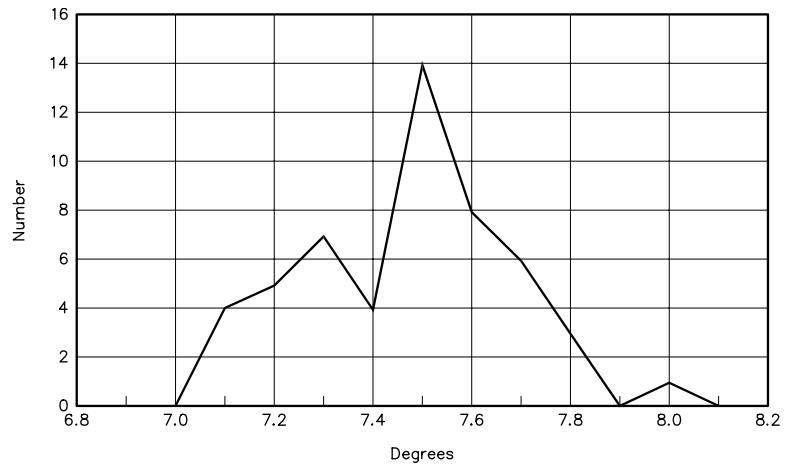
Fig 7—Distribution of the solar-depression angles for precursor events.
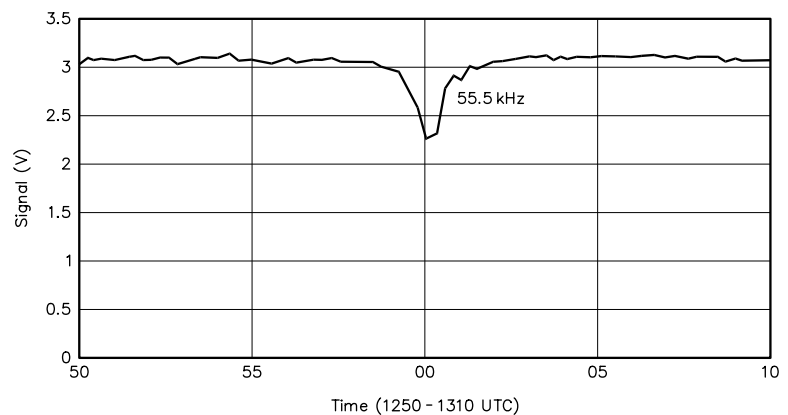
Fig 8—A typical example of a precursor event.
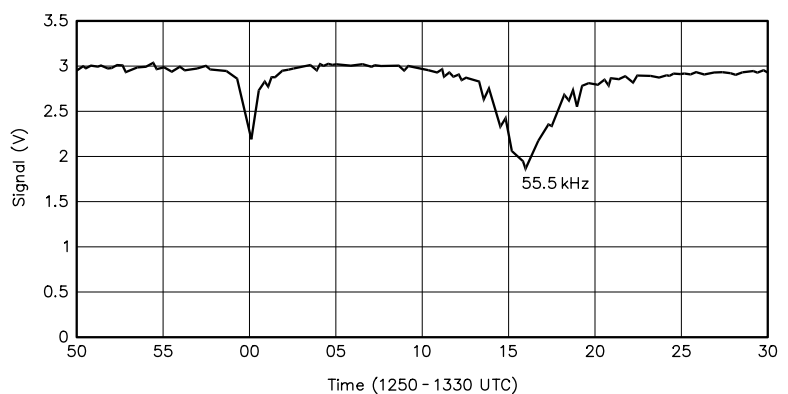
Fig 9—Sunrise signature of NPG on September 7, 1998 showing a precursor and main signature.

in absorption. After that detailed, exhaustive analysis, it was left that another ion (with a larger electron affinity and requiring UV light for photo-detachment) was probably involved and was an ion with which the ozone layer would play a major role in the day/night variations.

The present LF observations also show that the shadow of the Earth is not involved in the sunrise variation of conditions in the lower D region. That is quite evident from the fact that the sunrise signature of NPG signals is about 15-20 minutes later than would be expected from solar radiation being blocked by the solid Earth. Instead, it is possible to account for the delay just by using the known properties and extent of the ozone layer. That shows again, as Reid (Note 5) demonstrated, that the negative ion of molecular oxygen is not the controlling negative ion in the lower D region as atmospheric ozone is transparent to visible radiation.

At this point, consideration should return briefly to the precursors noted earlier, as in Fig 7. Some 39 of those events were observed during the year-long recording of LF sunrise signatures. The very limited spread of solar-depression-angle values is rather remarkable when one considers that it was obtained while observing the effects of radiation controlled by the properties of a fluid medium, not the hard, sharp boundary of the solid Earth. The question then becomes just how that sort of result is obtained and the answer proves to be both simple and complex.

Sunrise on the region of interest begins around 90 km and illumination proceeds down toward the lower D region. As shown by the ray paths in Fig 6, sunrise at the 90-km level is blocked until the ray paths for solar UV reach the outer fringes of the ozone layer—at an angle of about 7° below the horizon—for the ozone distribution used there. That goes a long way to explain the range of solar depression angles of precursors in Fig 7.

Once started, though, precursor signal variations do not continue on the path to become early examples of LF sunrise signatures, like those in Figs 1, 2 and 3. Instead, they end after about the same time interval as it takes them to appear. Since the ray paths in Fig 6 move upward with the advance of time, there was a brief time interval during which photo-detachment took place. Then UV radiation could get through the dilute, upper parts of the ozone layer, but after that,

ozone again blocked UV from reaching the reflection region.

The time when UV first got through the top of the ozone layer could have involved lower-than-normal ozone concentrations there, either a fluctuation of spatial or temporal origin. Of course, with the data in hand, it is not possible to decide which of the two was involved. The ray paths in Fig 6 show that a precursor requires a lower-than-normal ozone density in a bubble or tunnel over several hundred kilometers at the top of the ozone layer. The small number of precursor events in a year's time indicates the probability of that circumstance occurring. Nonetheless, the limited range of the solar angles during precursor events suggests that the extent of the vertical distribution of ozone used here is fairly common.

Beyond the question of vertical extent, a static ozone distribution over the horizontal distances in Fig 6 is quite unrealistic. Spatial and temporal variability of the layer, noted earlier, over that vast expanse could well be the reason why the LF precursors end after their brief appearance. In any event, vertical motions of ozone at the top of the layer seem more reasonable than horizontal motions as that transport is part of the circulation pattern that distributes ozone in the atmosphere.

Negative Ions

The discussion up to this point has dealt with the effects of ozone (a minor constituent of the atmosphere at low altitudes) and how it casts a shadow on the lower D region at sunrise. The delay in lowering the LF reflection region cited earlier indicates that visible radiation, blocked by the shadow of the Earth, is not responsible for the detachment of electrons from the negative ion population. In terms of ion species, that means that the negative ion of molecular oxygen, a constituent in the D and E regions, is not involved to any extent. Its electron affinity is too small (only 0.45 eV), and electrons would be detached with the arrival of quanta in the infrared and visible portions of the spectrum.

While there is a large body of knowledge about positive ion reactions in the D region, the same is not true of negative ions. Earlier, the negative-ion reactions of oxygen species in the atmosphere were the starting point of discussions of absorption in PCA events.[6, 7] Thus, it was thought that negative ions were formed in the D region by the attachment of electrons to different forms of oxygen:

$$O + e \rightarrow O^- + h_\nu \qquad \text{(Eq 1)}$$
$$O_2 + M + e \rightarrow O_2^- + M \qquad \text{(Eq 2)}$$
$$O_3 + e \rightarrow O^- + O_2 \qquad \text{(Eq 3)}$$
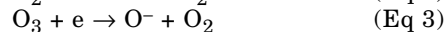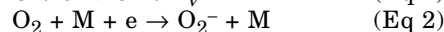
The second reaction is a three-body process and M is a neutral constituent, either N_2 or O_2, with different rate coefficients for each possible collision partner (see Note 2).

While photo-detachment dominates during daylight, there are different types of collisional detachment that operate day or night:

$$O^- + O \rightarrow O_2 + e \qquad \text{(Eq 4)}$$
$$O_2^- + O \rightarrow O_3 + e \qquad \text{(Eq 5)}$$
$$O_2^- + N_2 \rightarrow O_2 + N_2 + e \qquad \text{(Eq 6)}$$
$$O_2^- + O_2 \rightarrow O_2 + O_2 + e \qquad \text{(Eq 7)}$$
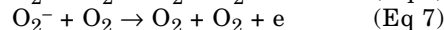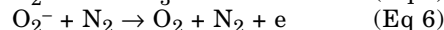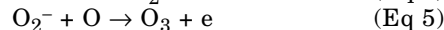
But Reid (Note 5) was unable to show the importance of negative ions of molecular oxygen in the sunrise and sunset variations of PCA events. With that, the discussion becomes more speculative, moving toward other reactions between those negative ions and the various minor atmospheric constituents such as O_3, CO_2, NO and H.

While ozone is formed above a 25-km altitude by chemical reactions that follow the photo-dissociation of molecular oxygen, it is then carried downward by atmospheric mixing. It then becomes part of the mixture of gases that result from human activity at lower altitudes. That same mixing carries upward carbon monoxide (CO) emissions from automobiles, nitrogen compounds from the extensive use of chemical fertilizers and water vapor from evaporation of the oceans. In the presence of solar radiation, those chemicals could be the basis from which more complex negative ions develop and populate the D region. Any greater electron affinity, up to 4.5 eV, would go to help understand the UV and ozone effects noted in the LF monitoring at sunrise.

Sunset

In spite of that statement, there remain problems with the negative ions in the D region and, more importantly, how they affect propagation at low frequencies and the 160-meter band. In specific terms, the present study was extended from the sunrise hours to those around sunset, looking for similar effects on LF propagation from the ozone layer. Thus, at first glance, it was thought that since the ozone layer delayed the onset of absorption at sunrise, perhaps it would do just the opposite: speed up the decrease of absorption with sunset. With LF observation intervals at sunset similar to those at sunrise, though, sunset signatures were just not found in number and amplitude

comparable to those at sunrise. That calls for some explanation, and it must be given in terms of photo-detachment of negative ions in the D region.

About the only conclusion that can be reached at this time is that there is a continuous variation in the types of negative ions in the lower D region after the sun sets. Thus, the negative ions present at sunrise are the result of long hours of darkness, while the negative ions that form after sunset follow hours of sunlight on the region.

Therefore, the negative ions at sunrise have enjoyed many hours of darkness in which to develop. Negative ions at sunset are not quite as fortunate; they have to start from scratch as the sun goes down. Nevertheless, they do develop, mainly as atomic oxygen converts to ozone after sunset. So there is a period of growth and change after sunset, starting negative ions on reaction chains that ultimately lead to the sunrise ionosphere.

The operative word in that last sentence is "change." Nowadays, the negative ions that develop after sunset are thought to result from ion chemistry involving the minor constituents noted earlier—ozone, carbon dioxide, nitric oxide and hydrogen from water vapor. Those molecules can combine with the simple negative ions of oxygen after sunset and develop into bicarbonate ions (HCO_3-) and nitrate ions (NO_3-), which hold their electrons closely. Fig 10 illustrates a detailed view[8] of negative ion reactions that remove electrons from the D region with the aid of those minor constituents in the atmosphere.

The left-hand side of the figure can be considered as the start (at sunset), with electrons attaching to molecular oxygen and ion reactions then proceeding toward the right-hand side, ultimately giving the stable ions in the dawn ionosphere. The progress through that complicated chain of reactions depends on the local availability of those minor constituents, from the exhaust of autos, use of nitrogen fertilizer and water vapor from the oceans and lakes.

Close examination of Fig 10 shows the various steps in which O_3, CO_2, NO and H are involved. However, their actual availability is dependent on the vertical and horizontal transport processes in the atmosphere. Thus, ion progress (from left to right in Fig 10) may be subject to bottlenecks when a minor constituent is missing, say a low density of ozone or nitric oxide, here or there, now and then.

All that means that the lower ionosphere, where 160-meter propagation takes place, is not a uniform region. Indeed, going from the sunset terminator to the sunrise terminator, the slow development of the negative ions which hold electrons closely will have ramifications for low-frequency and 160-meter propagation.

For that set of circumstances, negative ions species changing from one with a low electron affinity at sunset to another with a large electron affinity at sunrise, collisional detachment of electrons at sunset would yield a significant electron population in the D region. Thus, any strong gradient in electron density associated with the presence of

negative ions would be absent initially. Only as atomic oxygen decays with the halt of photo-dissociation of O_2 would large numbers of negative ions begin to form, albeit slowly, and then change the height of the reflection region by electron attachment.

Any changes in LF or VLF propagation that result from the rise of the reflection region would be slow at sunset, in contrast to the rapid changes at sunrise from the photo-detachment
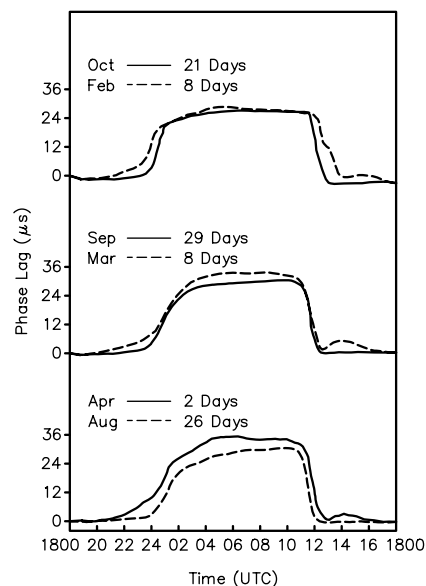
Fig 11—Diurnal phase changes of NBA VLF signals from Panama, as observed in Boulder, Colorado (after Davies, Note 9).
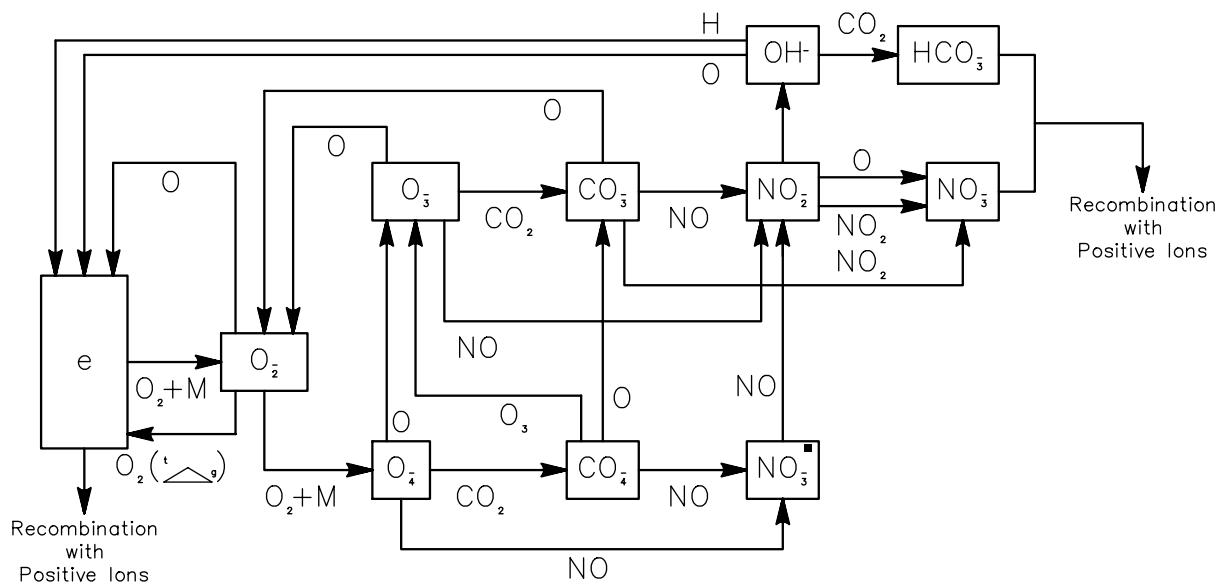
Fig 10—Schematic diagram of negative-ion reactions (after Reid, Note 8).

that requires UV. In that regard, a good illustration is seen in Fig 11, after a figure in Davies' book[9] (1989), where the diurnal phase changes of NBA VLF signals from Panama were observed in Boulder, Colorado. The change in phase around sunset is much slower than that at sunrise and that asymmetry can be attributed to differences in the electron affinity of negative ions in the two regions.

160-Meter Propagation

For 160-meter propagation, absorption in the D region is the biggest concern and a growing concentration of massive, negative ions there, with a large electron affinity, serves to lower its electron density and reduces the absorption of any waves passing through the region. The question then is how rapidly after sunset does the more-stable negative-ion population develop along a radio propagation path? Experimentally, that could be estimated from absorption of a 160-meter beacon's signals from sunset to sunrise.

But leaving time or distance scales out of the discussion for the moment, one must consider the propagation mode in effect and the number of D-region traversals and ground reflections that are involved. To the extent that conventional Earth-ionosphere hops take place along a path, absorption effects would be less on any wave traversals of the D region, which are closer to the sunrise terminator than those closer to the sunset terminator, where ionospheric electrons are held less strongly by negative ions. This leads to yet another circumstance, in addition to signal ducting,[10] which favors propagation on paths going from west to east.

Beyond that geographical consideration, it should be noted that the concentrations of minor neutral constituents that are involved in negative-ion chemistry, such as ozone, carbon dioxide, nitric oxide and hydrogen, are highly variable. That is the case as they are created and destroyed in various regions and are linked to circulation through vertical and horizontal transport in the lower atmosphere. Thus, like the delay in absorption at sunrise due to the role of ozone (see Note 1), there will be some variability in the electron/negative-ion distribution along a dark path, and ionospheric absorption will be quite variable as well. As a result, the meteorology of minor atmospheric constituents is seen to play an important role in 160-meter propagation.

More on Atmospheric Ozone

Concerning the meteorology of minor atmospheric constituents, their concentrations are not measured as frequently nor as widely as the other atmospheric parameters, say temperature and pressure. But in the present instance, where the reflection region of NPG's signals is lowered as UV penetrates the ozone layer, it is possible to use NPG as a beacon of sorts, looking at daily data, to examine spatial and temporal features of ozone where nothing else is available.

It should be noted, however, that the parts of the ozone layer in question are several hundred kilometers east of the path midpoint. In addition, solar directions change with the time of year, being more northerly in summer and southerly in winter. With that caveat, the results of the year-long study of LF signals can be used to look at ozone variations as well as their implications for 160-meter propagation and considered as new, fresh information.

In that regard, the basic data are the magnitudes (in decibels) of sunrise signatures, the times of greatest decrease in signal strength, the corresponding solar-depression angles at the path midpoint and the shapes of the signature curves. The distribution of magnitudes of sunrise signatures is shown in Fig 12, along with the line representing a linear regression of the data. For 353 data points, the maximum and minimum values for the sunrise signatures were 3.76 and 0.2 dB, respectively, while the median value was 1.23 dB with a standard deviation of 0.69 dB. The distribution of the data points relative to the regression line do not suggest any strong correlation and
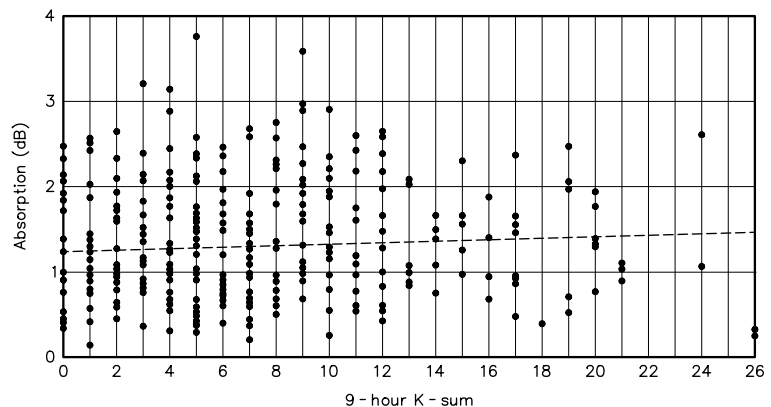
Fig 12—Point scatter-diagram for NPG signals and 9-hour Kp-sums from records of NOAA magnetometers.
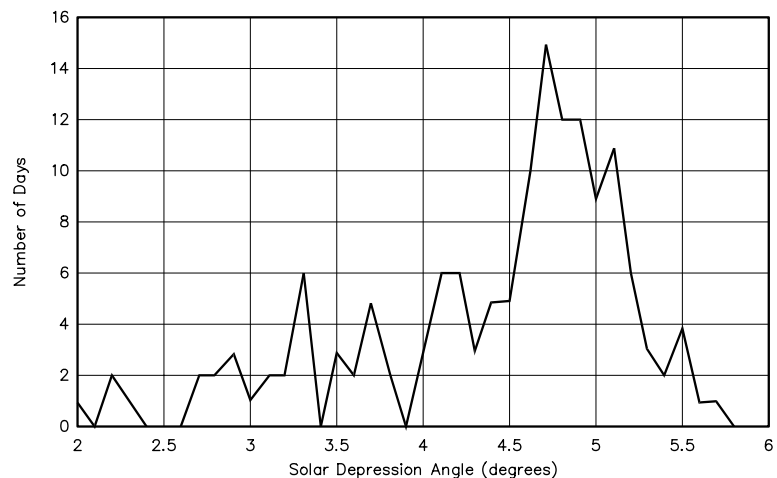
Fig 13—Distribution of solar depression angles.

the statistics bear out that point, the correlation coefficient being only 0.22.

The solar depression angles were calculated for data points above the average value, the aim being to use observations where a time could be obtained without too much in the way of uncertainty. That was not always the case as shapes of the signatures varied throughout the year and time data were not always accurate enough when the signatures were broad. But for observations where times and solar-depression angles were readily obtained from those data points above the average value in Fig 12, the distribution of angles is given in Fig 13.

The curve shows a strong peak just below 5°, as well as a broad range of depression angles down to about 2°, but that is a summary figure for the year-long period of the study. A more informative display of the year's data is given in Fig 14, where highest and lowest depression angles are given for each week throughout the year. That figure shows a seasonal dependence for depression angles, greater angles between mid-March and mid-November than for the rest of the year.

While depression angles were smaller in the winter, the variability of the data points was much greater during winter than in summer. In particular, the spread of data points in winter was about ±1.0°, compared to ±0.5° in other parts of the year. In addition, the trends in Fig 14 are more indicative of the heights reached by the upper parts of the ozone distributions. Thus, the small angles and greater variability in winter months are characteristic of distributions of ozone where the upper part extends to greater altitudes than those with larger angles and less variability in summer.

The time durations of the sunrise signatures for the LF signals also differed with season, being longer in summer than in winter. In addition, there were differences in the shapes of the sunrise signatures. Thus, there were three distinct shapes: one with a symmetrical decrease in signal strength with time and two different, asymmetric shapes. One where the signal strength decreased slowly at first, but then with a faster recovery; and the other just the opposite, with a fast decrease and a slow recovery in signal strength.

From mid-March to mid-November, the shapes noted most often were either symmetrical or with a slow recovery, only a few having a slow onset. The asymmetric shapes with a slow onset were found more frequently in the winter months, when the spread in depression angles was greater, but a few of those shapes were also noted in the summer. Those few, however, were rather striking. They involved large changes in the depression angle, from fairly steady background values around 5° and going down to 3.0° or 3.5°, meaning there were sudden increases in heights reached by the upper parts of the ozone distributions. A search of historical weather data failed to indicate any obvious causes for those changes, though.

An effort was made to better understand the shape and the qualitative features of the sunrise signatures by using model calculations for the ozone layer. Thus, a Gaussian height distribution of ozone like that used earlier (see Note 10) to interpret the delay in sunrise signature was extended, beyond just its opacity due to the total mass of the ozone, to now include the exponential absorption of UV along the line of sight from the sun.

For that calculation, the absorption coefficient of ozone[11] was considered over the UV range, where the principal photo-detachment of negative ions would be expected. There, the absorption coefficient peaks at 140/cm-atm at 255 nm, between low values at 200 nm and 300 nm, and has an average of 75/cm-atm.

In the model calculations, the height of the peak ozone concentration ranged between 20 and 35 km. The width or standard deviation of the Gaussian distribution was varied from 6 to 12 km, and the total ozone content of the distribution taken from 0.24 to 0.36 cm-atm. With those ranges, calculations for D-region heights showed that the intensity of UV during sunrise was far more affected by changes in the height or width of the ozone layer than the total ozone content.

In regard to the last point, the total ozone content at middle latitudes in the northern hemisphere (see Note 2) shows a maximum in March of about 460 DU and drops to a minimum of about 280 DU in November. A long-term study of total ozone cited by Craig (see Note 3) shows that same annual variation but variations from monthly means that were rather large in January through March and then declined, becoming much smaller from June through November.

It is interesting to note that the weekly values of solar-depression angles in Fig 14 are more in phase, as it were, with those variations of monthly means of total ozone content than the ozone content itself. Of course, the data in that figure is only from one year of observation. Another year of study is now in progress, and it will be interesting to compare results between the two years and the variability of total ozone content.

As for the shapes of the signatures, they depend on the rate of phase change between the sky wave and ground wave as the reflection region is lowered at sunrise. The lowering results from solar UV that gets to the D region at the midpoint of the path by advancing closer and closer, as in Fig 5, through the upper reaches of the ozone layer. So the more ozone encountered, the slower the lowering and vice-versa.

With that, those signatures with more rapid change in signal strength at the outset represent cases where the top of the ozone layer was either lower
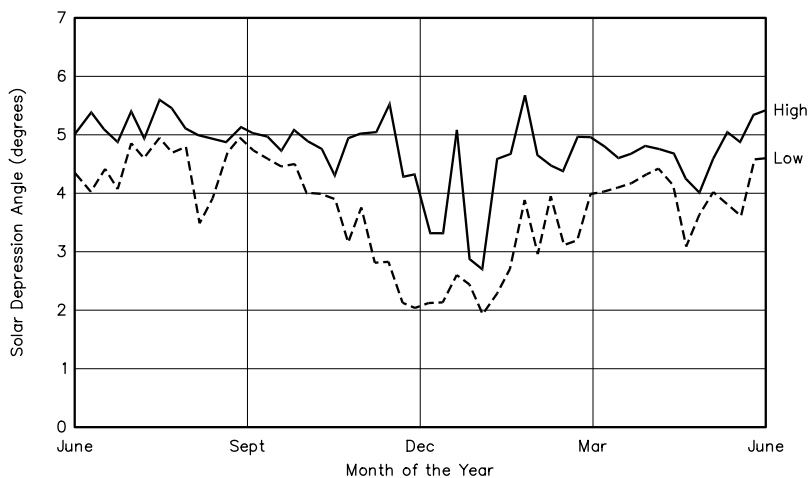
Fig 14—Weekly values of solar depression angles.

or more dilute at far distances, when first encountered by the rising solar UV, than in close to the midpoint, when last encountered. By the same token, slow changes in signal strength represent cases where the top of the ozone layer was either higher or more dense at far distances than in close to the midpoint. Those signatures have implications for the onset of photo-detachment of electrons and ionospheric absorption, but it should be stressed again that they are more sensitive to the height and width of the ozone layer than the total ozone content at points along the optical path.

Conclusion

The study of LF propagation at sunrise and sunset has shown that atmospheric ozone, and probably other minor constituents in the atmosphere, play an important role in negative-ion formation in the lower D region. While the data presented here is limited to those two extremes in a day, it seems fairly clear that the lower D region is not uniform in its properties at night. Thus, the present results suggest that a significant electron density, which contributes to ionospheric absorption of 160-meter signals, is present after sunset takes place. Nevertheless, the transition to the time around dawn, when negative ions play a more important part in removing free electrons from the D region, may be variable. Moreover, it may not be predictable, given the role of transport processes in distributing minor atmospheric constituents.

While transport processes play a major role in the global distribution of ozone, notice that the brief precursor events, as in Fig 7, probably reflect such processes, but on a smaller scale. Of course, satellite measurements such as those from the Solar Backscatter UV (SBUV) detector give data for ozone profiles on a global scale; but, because of its rapid motion, the satellite cannot observe brief events in the ozone density. That problem has existed before in high altitude research, satellites often miss details of geophysical phenomena that near-stationary vehicles, such as high-altitude balloons, are able to detect. By that token, the present study using radio propagation provides a type of information about the ozone layer that is not available from any other source at the present time. The area of concern is at altitudes too high for balloons, too low for satellites and the phenomena are too infrequent for the use of sounding rockets.

Turning to propagation on the 160-meter band, it's of great interest in the winter months. In that regard, the LF signatures with lower solar-depression angles as well as greater variability in angles, noticed in connection with the data in Fig 14, suggest that ozone effects on ionospheric absorption at those times can be quite pronounced. They may actually be more variable than suggested earlier (see Note 1), before 1998-1999 winter data was in hand. It will be interesting to compare that aspect of these observations with the results from a study of dawn enhancements in the MF frequency range (Hall-Patch, private communication).

Finally, the present study was conducted in the period from April '98 through March '99. Considering that the weather patterns in that year were highly influenced by the El Niño/La Niña processes, the question arises as to whether the variability of ozone effects in the winter months between the equinoxes are typical or, perhaps more likely, reflect the unusual circumstances of the times. In that regard, an attempt will be made to obtain another year of LF data to provide an answer to that question.

Notes

[1] R. Brown, NM7M, "Atmospheric Ozone, a Meteorological Factor in Low-Frequency and 160 Meter Propagation," *Communications Quarterly*, Spring 1999, Vol. 9 No. 2, p 97.

[2] A. Brekke, *Physics of the Upper Polar Atmosphere*, (New York: Wiley and Sons, 1997).

[3] R. Craig, *The Upper Atmosphere*, (San Diego: Academic Press, 1965).

[4] M. Salby, *Fundamentals of Atmospheric Physics*, (San Diego: Academic Press, 1996).

[5] G. Reid, "A Study of Enhanced Ionization Produced by Solar Protons during a Polar Cap Absorption Event," *Journal of Geophysical Research*, 1961, Vol. 66, p 4071.

[6] D. Bailey, "Abnormal Ionization in the Lower Ionosphere Associated with Cosmic-Ray Flux Enhancements," *Proceedings of the IRE*, 1959, Vol 47 p 255.

[7] R. Brown, and R. Weir, "Ionospheric Effects of Solar Protons," *Arkiv for Geophysik*, (Royal Swedish Academy of Science, 1961), Bd. 3, Nr. 21, p 523.

[8] G. Reid, "Ion Chemistry of the D-region," *Advances in Atomic and Molecular Physics*, Vol. 12, (San Diego: Academic Press), 1976.

[9] K. Davies, *Ionospheric Radio*, (London: Peter Peregrinus, Ltd, 1989).

[10] R. Brown, "Signal Ducting on the 160 Meter Band," *Communications Quarterly*, Spring 1998, p 65.

[11] R. Whitten and I. Poppoff, *Fundamentals of Aeronomy*, (New York: J. Wiley, 1971).

Bob Brown was first licensed as W6PDN in 1937. He held KA6PTT, N7DGZ before his present call. He holds a doctorate in physics from the University of California. He has been an instructor of physics at Princeton University and a professor at the University of California at Berkeley. He retired in 1982.

Bob has done research on high-energy cosmic rays, solar proton and auroral electron bombardment of the polar atmosphere, ionospheric absorption at high latitudes and auroral X-rays in conjugate regions of the geomagnetic field. He has published over 80 scientific papers in refereed journals from 1944 through 1982 and a review article, "Electron Precipitation in the Auroral Zone," in Space Science Reviews *(1966).*

He has conducted several propagation columns in Amateur Radio publications: "Propagation and DX" in QRP Quarterly, *"Over the Horizon" in* The Canadian Amateur *and "Propagation" in* Worldradio. *He has written* Long-Path Propagation *(published privately in 1992) and* The Little Pistol's Guide to HF Propagation *(Worldradio Books, 1995). Bob's current interest is the role of atmospheric effects in 160-meter propagation.* ☐☐

Understanding Circular Waveguide—Experimentally

Old-fashioned microwave engineering.

By Paul Wade, W1GHZ

Waveguide is an excellent microwave transmission line, with low loss and predictable performance, usable at any frequency by choosing the proper dimensions. The most common type of commercial waveguide is precision rectangular tubing, which is only affordable on the surplus market. Elliptical waveguide is also used commercially for microwave transmission line. Many microwave structures, particularly antennas, have a round cross section and are better suited to circular (cylindrical) waveguide. Unfortunately, commercial circular waveguide is rare and unlikely to

161 Center Rd
Shirley, MA 01414
w1ghz@arrl.net

be found surplus. As luck would have it, ordinary copper water pipe works just fine and is universally available at low cost. In particular, $^3/_4$-inch copper pipe is perfect for 10 GHz.

Most microwave design today is done with the aid of computers. However, only a few programs handle electromagnetics with the capabilities required for circular waveguide, and their prices are somewhere between a fancy car and a new house. Solving the problems without good software involves some difficult math, and the solution is probably only approximate. The remaining alternative is old-fashioned empirical microwave engineering.

The ham's favorite design technique, reverse engineering (copying something that works!) is made difficult by the lack of commercial circular-

waveguide examples. Reference books have extensive information on rectangular waveguide but very little information on circular waveguide. The pioneering work on waveguides was done by George Southworth before World War II, but I only recently located a copy of his book.[1] I was afraid that I had duplicated some of his work, but that's a good way to learn. On first reading, it appears that he did a lot of work on waveguide fundamentals, but very little on waveguide-to-coax transitions, probably because good coaxial cable and connectors were not available prior to the wartime development.

Circular Waveguide

We all know that electromagnetic

[1]Notes appear on page 48.

waves travel through space—that's what radio is all about. They can also travel inside a hollow pipe of any shape; if the dimensions are right, the pipe makes a very low-loss transmission line, much better than any coaxial cable. In order for the waves to travel with low loss, the pipe dimensions must be large enough for the lowest-order waveguide mode, the TE_{11} mode, to propagate. In circular waveguide, the cutoff wavelength for this mode is $1.706{\times}D$ (diameter) so the minimum waveguide diameter is 1/1.706, or 0.59 λ. The diameter of the copper water pipe I used is nominally $^3/_4$-inch, type M, which has a larger inner diameter than other types. The typical inner diameter is 0.81 inches, but that may vary slightly because this is not precision tubing. Thus, the cutoff wavelength is 1.38 inches, so the minimum frequency is 8.55 GHz. Clearly, 10 GHz is comfortably above the minimum.

Moving in the other direction, a large waveguide diameter would permit additional higher-order waveguide modes to propagate. While the additional modes also propagate with low loss, they often arrive at the far end with different phase, so that they interfere with the TE_{11} wave and we are unable to extract them without losses. The next mode, TM_{01}, needs a minimum diameter of 0.76 λ to propagate, setting the maximum operating frequency without any additional modes. For the $^3/_4$-inch pipe, this upper frequency limit is 11.08 GHz, but it isn't a hard limit like the lower cutoff frequency. At higher frequencies, the waveguide still propagates energy, it's just difficult to predictably couple that energy efficiently.

Thus, a hollow round pipe is an excellent waveguide for wavelengths between 0.59 and 0.76 times the inside diameter. Standard USA $^3/_4$-inch copper water pipe, type M, has an inner diameter of 0.70 λ at 10.368 GHz, so it is ideal for 10-GHz operation. It is readily available in almost any hardware store at a cost much lower than that of coaxial cable suitable for VHF use. Copper water pipe does come in other versions, but type M is preferable since it has the largest inner diameter.

Measurement

Our design style is "old-fashioned empirical microwave engineering." "Empirical" is a fancy word for "cut and try," but it is only engineering if we make measurements, record data and try to understand the results.

The most important measurement we

need is impedance in the waveguide. Today, impedance measurements are made using a network analyzer, preferably one that is automated, with computer control and error corrections. There aren't any waveguide network analyzers; they are all based on coaxial transmission lines. For the most popular standard sizes of rectangular waveguide, good coaxial transitions and calibration kits are available to allow network-analyzer measurements, with the computer correcting errors caused by the transitions. Of course, none of this is available for circular waveguide—if we had a quality coax transition to copy, we'd be one giant step closer to using circular waveguide.

Before network analyzers, microwave impedance measurements were made using a slotted line. A narrow longitudinal slot in the outer conductor of a coaxial line does not interrupt any current flow in the line, so it has no effect. A small probe may be inserted in the slot to measure the voltage in the line, and moved to measure the voltage at other points along the line. If the line is mismatched, the voltage varies in a pattern referred to as a standing wave. Normally, we measure the ratio between the minimum and maximum voltages of the standing wave and call it the standing wave ratio, or SWR. Using a slotted line is becoming a lost art, but the technique is covered pretty well in a recent book by Pozar.[2]

In waveguide, a slot will have no effect if it does not interrupt any current. In rectangular waveguide, this location is easy to find. It's in the center of the broad wall. In circular waveguide, there is no obvious orientation; we must

orient the guide so that the E-field is symmetrical around the slot and the probe is parallel to the E-field.

I built a slotted line for circular waveguide by cutting a longitudinal slot in a piece of $^3/_4$-inch copper pipe. To fit the pipe to a surplus slotted-line carriage, I made a pair of plywood blocks. The carriage is designed for interchangeable line sections of coax or different sizes of rectangular waveguide, so the circular section had to fit the same mounting points for the probe to travel in the slot correctly. Fig 1 is a photograph of the slotted line. Later, I found a sketch[3] of Southworth's slotted line. The line section was similar, but without the advantage of a surplus carriage, he had to build a sliding mechanism as well.

To feed RF energy to the slotted line, I started with a surplus coax-to-WR-90 rectangular waveguide section. I then made a rectangular-to-circular transition by hammering one end of a $^3/_4$-inch copper pipe until it fit into a WR-90 waveguide flange. This makes a good transition if the shape is a long, smooth taper. I added a WR-90 isolator between the coax transition and the tapered section to absorb any reflected power so that the signal generator would not change frequency or power output due to loading. Finally, at the input to the slotted section of pipe, I added a septum (a flat plate across the diameter of the pipe) perpendicular to the probe; only energy polarized parallel to the probe will propagate past the septum. This polarization is important, so that the E-field is parallel to the probe—otherwise the probe and slot might upset the fields and convert en-
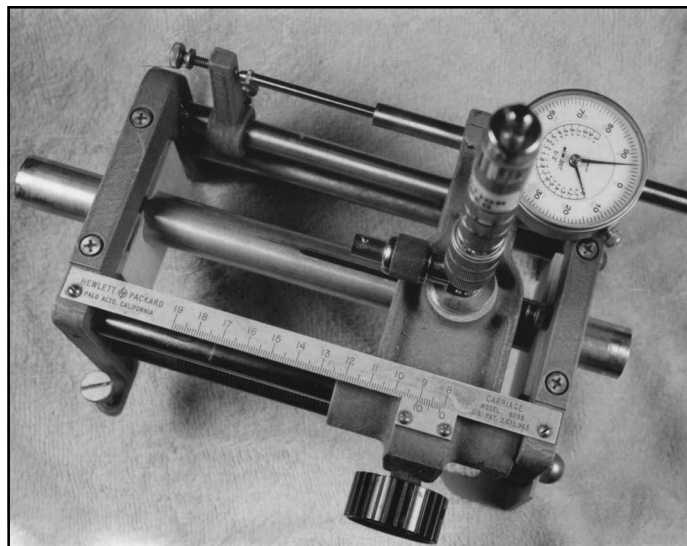
Fig 1—
Homebrew slotted line for circular waveguide mounted in a surplus carriage.

ergy to unwanted waveguide modes, with unpredictable results and strange measurements.

The voltage in the slotted line is sampled by the probe inserted through the slot; if the probe is inserted too far, it will affect the fields in the waveguide and produce erroneous readings. On the other hand, a deeper probe produces more output voltage, so less RF power is necessary for the measurement. The proper probe depth is found experimentally, by increasing the depth until the measured SWR starts to change, then backing off.

The probe assembly shown in Fig 1 contains a diode detector with a tuning mechanism; when it is adjusted for resonance, much more detected voltage is available. The output from the detector goes to a standing-wave meter such as an HP415; other manufacturers made similar instruments. The meter is an ac voltmeter tuned to 1 kHz, so the RF source must be AM modulated at 1 kHz; most signal generators have this capability.

Once the slotted line was working, I quickly discovered two things:

1. The wavelength at 10 GHz is nearly twice as long in ³/₄-inch waveguide as it is in free space.

2. This means that the probe must travel near the end of the slot to measure a full wavelength.

As the probe approached the ends of the slot, I could see that there was an effect. A well-matched horn antenna had a low indicated SWR with the probe near the center of the slot, but a higher indication when it was near the ends. Since commercial rectangular-waveguide slotted sections taper the end of the slot to a point, I used a tapered file to trim the ends of the slot to smooth the response.

Wavelength in the waveguide is measured by shorting the end of the waveguide with a flat plate; this produces a standing-wave pattern with a null every $\lambda/2$ from the short. I knew that guide wavelength, λ_g, would be longer than the wavelength in free space, λ_0, (phase velocity is greater than the speed of light), but hadn't realized how much longer. Fig 2 shows λ_g versus frequency; as the cutoff frequency is approached, λ_g increases dramatically, while λ_0 increases linearly with decreasing frequency.

Impedance is measured and calculated graphically by plotting SWR and phase on a Smith chart. Phase is measured by the location of the standing-wave minimum voltage on the slotted line. The distance between two stand-ing-wave minimums is $\lambda_g/2$. With a short circuit on the end of the slotted line, we can locate two voltage nulls on the line to provide the reference points for our measured location. A classic Smith chart has a wavelength scale around the perimeter that is used to plot phase; the circumference of the chart equals $\lambda/2$.

On a Smith chart, the impedance is *normalized* to the characteristic impedance of the transmission line. Common coaxial lines have characteristic impedances near 50 Ω. For waveguide, we use *wave impedance* rather than characteristic impedance. The wave impedance for TE modes in circular waveguide is calculated as:

$$Z_0 = Z_{fs}\left(\frac{\lambda_g}{\lambda_0}\right) \qquad \text{(Eq 1)}$$

where Z_{fs} is the impedance of free space, 377 Ω. From Fig 2, the guide wavelength, λ_g is longer than the free-space wavelength λ_0, so our circular waveguide impedance is greater than 377 Ω. Unlike coaxial transmission lines, the impedance varies with frequency. At 10.368 GHz, Z_0 is about 650 Ω; however, it is about 1130 Ω at 9 GHz and 580 Ω at 11.2 GHz. For our purposes, the exact impedance does not matter, as long as we can match it empirically and achieve a low SWR. In fact, I did not calculate Z_0 until after I had completed all the experimental work described here.

For phase measurements, the slotted line includes a vernier scale to measure distance traveled; the scale is quite accurate if used carefully. However, once I realized that many slotted-line
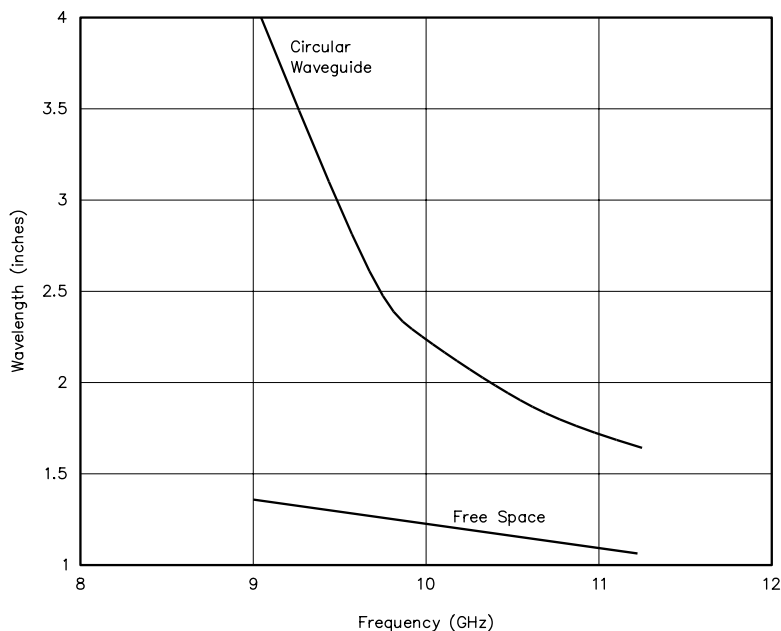
Fig 2—Wavelength does not vary linearly in ³/₄-inch circular waveguide.
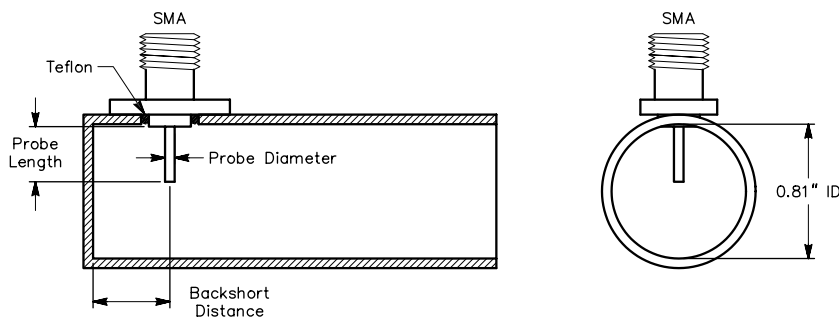
Fig 3—Construction of a ³/₄-inch pipe waveguide transition.

measurements would be required, I added a dial indicator (shown in Fig 1) to the slotted line. The dial indicator, together with a dial caliper to measure probe dimensions and backshort (I'll explain this term shortly) distances, speeds up measurements significantly. Unfortunately, inexpensive dial calipers and indicators read in inches only, rather than the metric units preferred for microwave work, so we will stick to inches. Anyway, it would be silly to refer to "3/4-inch pipe" in metric units.

Finally, we need a matched load. K2RIW reports carving a tapered point on a broomstick to make a load for circular waveguide, but a horn antenna with a long taper is known to provide a decent match. Therefore, I made a simple conical horn from copper flashing. The measured SWR is about 1.14:1.

Coaxial Transition

Our ultimate goal is to make better antennas using circular waveguide, but the antennas must connect to equipment that uses coaxial cable for interconnections. I had already built several feed antennas for dishes that I wanted to test, but first I needed a good reproducible coax transition to connect them. The simplest coax transition extends the center conductor of the coax as a radial probe in the waveguide, as shown in Fig 3. The end of the waveguide behind the probe ends in a short circuit referred to as a *backshort*.

My previous attempts at building transitions from coax to circular waveguide were not always successful. A number of designs have been published for transitions at lower frequencies, but with widely varying dimensions. The ones I have tried were not always well matched, and some were very critical. On closer examination, some of them are feeding mismatched antennas (such as "coffee can" feeds or open circular waveguides with typical SWRs of 2:1), so they are probably adjusted to a spe-

cific mismatch, rather than providing a matched transition.

On the other hand, coax transitions to rectangular waveguide have been more successful, probably because many of them start with dimensions of commercial transitions. I have adjusted some by making measurements with a rectangular-waveguide slotted line. The measurements suggest that the three variable dimensions in a waveguide transition (shown in Fig 3) all interact. There are combinations of probe diameter, probe length and distance to the backshort that transform the impedance of the waveguide to a desired coax impedance, usually 50 Ω. A textbook I consulted long ago suggested that there is an optimum probe length and backshort distance for each desired impedance, but the inductance of the probe must be compensated by changes in the length and backshort distance. Since the inductance is a function of probe diameter, the calculations only provide a rough approximation. Previous experiments at lower frequencies showed that the approximation was not very good.

Since calculations seemed inadequate, experimentation seemed like a good alternative. I built a fully adjustable coax transition, shown in Fig 4. A hexagonal plumbing fitting provides flat sides for an SMA panel-mount connector, so the probe length and diameter can be changed quickly. The connector is held by two screws and the Teflon dielectric extends through the pipe wall so that the probe starts at the

inside wall of the waveguide, a reproducible position. A sliding backshort provides a full adjustment range. At low frequencies, sliding finger stock is used to provide an adjustable short-circuit, but the fingers are too long to provide a good short at 10 GHz. The alternative, shown in Fig 4, is stepped quarter-wave sections: the first section is a sliding fit in the pipe, which provides a very low impedance. It is followed by a small-diameter high-impedance section, then another low-impedance section. The large sections have such low impedances that it doesn't really matter if they make contact; the result looks like a short circuit in the waveguide.

My strategy was to make some measurements with different probe diameters and find a good combination of dimensions for each diameter. I then planned to make some bandwidth measurements to see which combination was least critical, so it could be easily reproduced and scaled to other frequencies.

To choose a range of probe diameters, I looked at published designs. Low-frequency designs, for instance 1296 MHz, often use a thin probe of #14 AWG wire. At 10 GHz, this scales to perhaps 0.012 inches, which isn't very substantial. Published designs for higher frequencies use probes as large as 0.093 inches, which scales to about a 3/4-inch diameter at 1296 MHz. This wide range led me to try a range of diameters, from 0.010 to 0.062 inches, at 10 GHz.

For each probe diameter, I measured
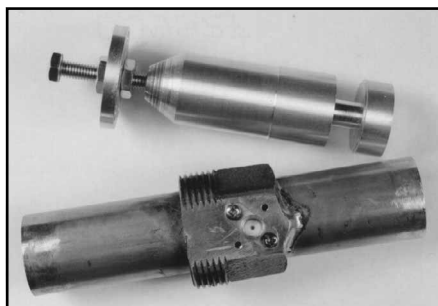
Fig 4—Adjustable circular waveguide-to-coax transition with a sliding short circuit.
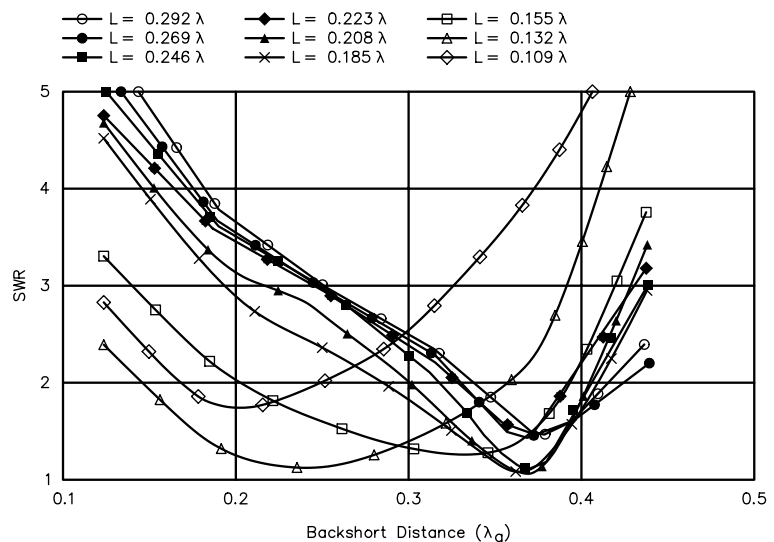
Fig 5—SWR versus backshort distance (in λ_g) for a probe diameter of 0.040 inches.

the waveguide impedance with the SMA connector terminated in a good 50-Ω termination. A low SWR in the waveguide is a good transition. I made the impedance measurement with the backshort distance moving in λ/8 increments, starting with long probe lengths that I trimmed in small increments.

The plan was simple: plot this data, spot a trend and zero in on optimum combinations. The first plot was SWR only, shown in Fig 5. This is what we would be able to measure with a network analyzer or a directional coupler. Perhaps you can spot a trend in Fig 5, but I only found it confusing!

So much for simple plans, it was time for the Smith chart. In my previous work with rectangular waveguide, plotting a few points by hand was sufficient to understand what was going on. However, I now had dozens of data points and no clear idea of which of them was worth plotting. I downloaded some Smith-chart routines from the Internet, picked a good one and added some code to plot my data in smooth curves on the Smith chart.

The confusing SWR data in Fig 5 is plotted on the Smith chart in Fig 6, with the addition of phase to make each data point a complex impedance. The data shown is for a probe diameter of 0.040 inches. Each solid curve plots a constant probe length with different backshort distances. The points are still well scattered, but one curve, for the shortest probe length of 0.109 λ, circles around the center of the Smith chart while the other curves are all on the left side of center. Thus, we might suspect that the optimum probe length may be somewhere between the shortest length and the next longest.

Fig 7 removes the curves for some of the longer probe lengths to concentrate on the ones that bracket the bull's-eye, the center of the chart. The dashed curves intersecting them are plots of constant backshort distance with varying probe length—what you would see if you soldered together a transition and could only trim the probe length. It's clear that neither of these adjustments alone will produce a good transition unless the other is

just right, but one of the dashed lines goes right through the center. The optimum backshort distance should be very close to 0.25 λg, or 0.250 guide wavelengths. The points I measured bracket the bull's-eye, so based on those values I tried to estimate a probe length that would land dead center.

Fig 8 is a Smith chart showing the three curves from Fig 7 plus one additional curve, the dashed line, for a probe length that was my estimate of optimum. The best point on this curve has a SWR of 1.05, a decent match. There isn't much point in trying to do better, since neither the slotted line nor the 50-Ω termination on the SMA connector is perfect—each of the measurements has some error and uncertainty.

The horizontal centerline on the Smith chart is the resistive axis; points on the line are purely resistive. Impedances greater than Z_0, the wave impedance, are to the right of the center point, while lesser impedances are to the left. Points off the centerline have reactive components—inductance is above the centerline, capacitance be-
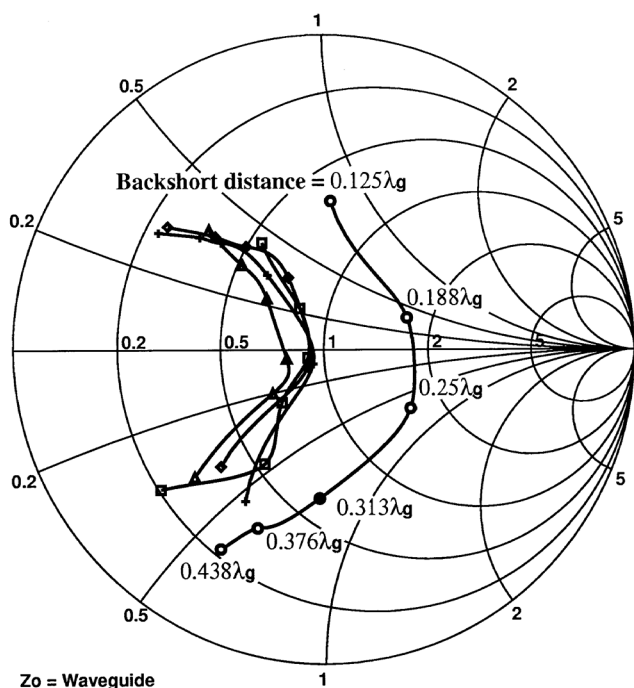
Fig 6—The SWR data of Fig 5 plotted on a Smith chart.
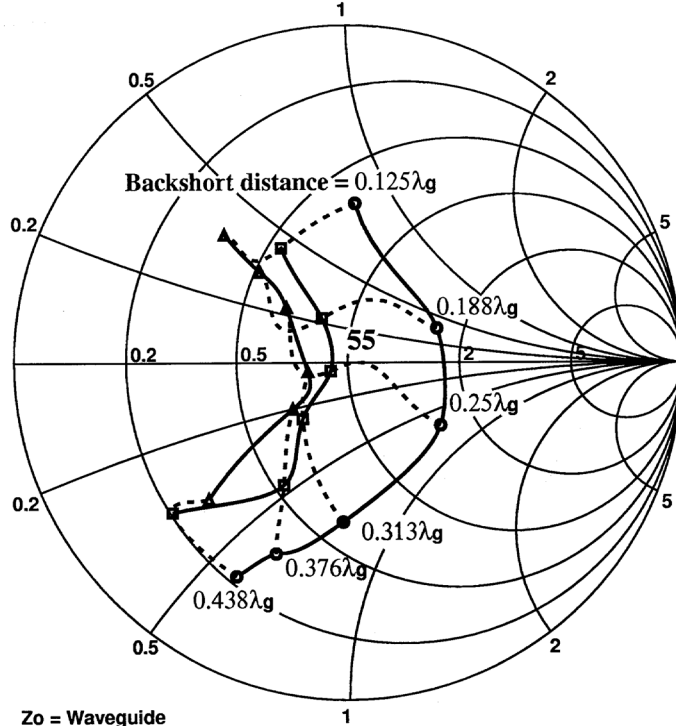
Fig 7—The SWR data of Fig 5 with some of the worst candidates removed.

low. So we can see some trends on these Smith charts: Longer probes produce lower resistances, while longer backshort distances move the reactance from inductive to capacitive. While it is clear from Fig 7 that these aren't straight lines, we can use these trends to "zero in" once we are close.

Returning to Fig 5, we can start to understand the curves. The three shortest lengths are the curves in Fig 7 that bracket the bull's-eye, but the SWR plot doesn't really make this clear. The other curves, for longer probes, all have a sharp minimum somewhere near a backshort distance of 0.375 λ_g; this looks suspiciously like a resonance. Since it is much easier to shorten a probe than to make it longer, I start with an overly long one and trim until it seems to be too short. Therefore, more data is taken with long probes than short ones.

On a Smith chart, we can see combinations of probe length and backshort distance that bracket the desired match at the center of the chart, then estimate the ideal combination from

these data points. I did this for a range of probe diameters from 0.010 to 0.062 inches. Fig 9 is the Smith chart for a probe diameter of 0.050 inches, the diameter of the center pin of an SMA jack. For this diameter, I again found three curves that bracket the bull's-eye, then estimated the best length and measured it for the dashed curve. The lighter dotted lines are plots for constant backshort distance. Fig 10 shows the same set of curves for the largest probe diameter, 0.062 inches, but without the lighter dotted lines.

Turning to smaller probe diameters, Fig 11 shows the curves for a probe diameter of 0.032 inches. In addition to the three that bracket the bull's-eye and the dashed line for my best-estimated length, there is an additional curve for a much shorter probe of 0.107 inches in length. This curve illustrates the trend toward a higher resistive component with shorter probe length —the curve crosses the horizontal axis further to the right.

For a probe diameter of 0.020 inches, I got lucky and cut the probe to

a length that produced a very low SWR, so I stopped there. Fig 12 shows the curves for this diameter, and Fig 13 shows the curves for the smallest diameter, 0.010 inches. At the latter diameter, the best length is significantly longer than with larger diameters, so the second length I tried was already too short and bracketed the bull's-eye. The dashed line is the final curve, for the estimated best length.

For each diameter, there is a good combination of length and backshort distance that provides a well-matched transition from circular waveguide to coax. Our other goal is to find a set of dimensions that is not critical, so that it may be readily reproduced and scaled to other frequencies. If the transition is broadband, with impedance that does not vary rapidly with frequency, it is probably more forgiving than one with rapidly varying impedance. Using the best dimensions for each probe diameter, I made measurements over the frequency range from 9 to 11.2 GHz. This frequency range corresponds to waveguide diameters of 0.617 λ to
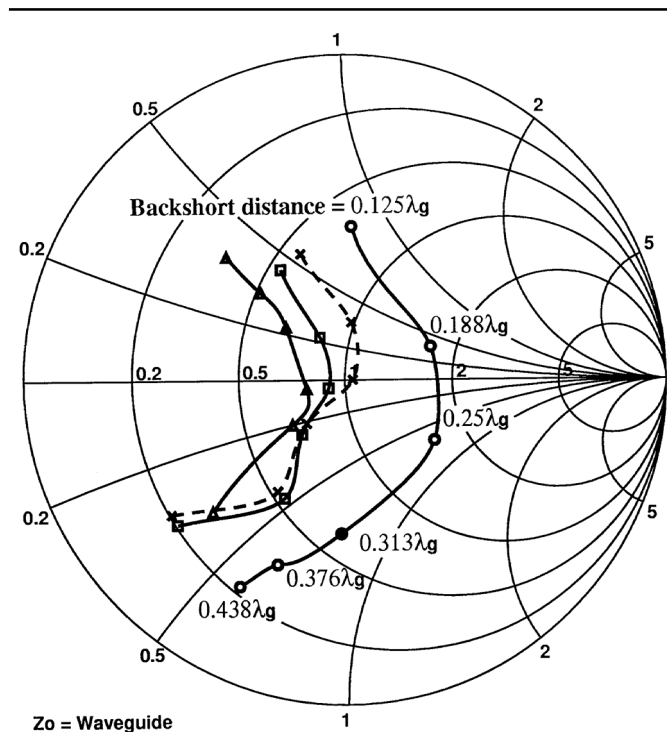
Probe Diameter = 0.040"
- ⊙ Probe length = 0.215" = 0.109 λ_g
- ▲ Probe length = 0.260" = 0.132 λ_g
- ▢ Probe length = 0.305" = 0.155 λ_g
- ✕ Probe length = 0.252" = 0.128 λ_g

Fig 8—The SWR curves of Fig 7 with a new dashed line representing the estimated optimum solution.
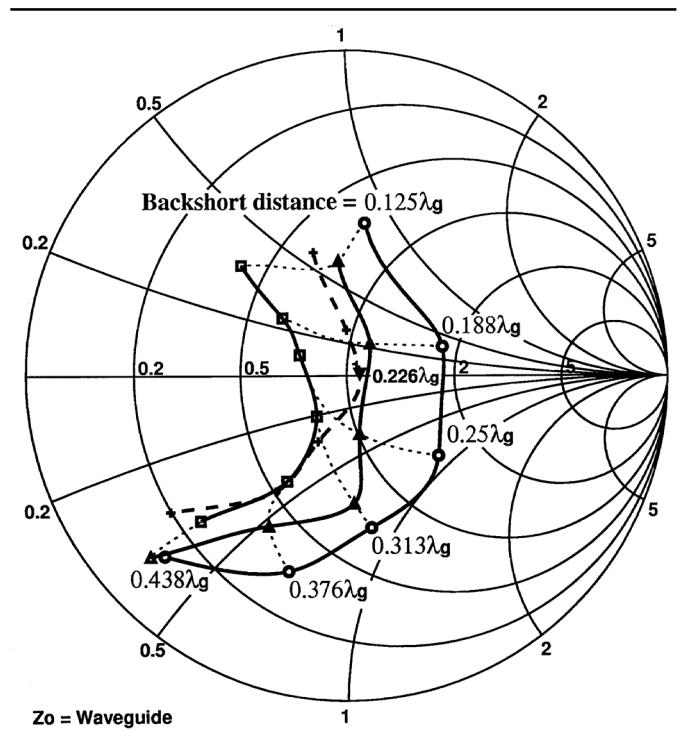
Probe diameter = 0.050"
- ⊙ Probe length = 0.205" = 0.104 λ_g
- ▲ Probe length = 0.235" = 0.119 λ_g
- ▢ Probe length = 0.290" = 0.147 λ_g
- ✚ Probe length = 0.248" = 0.126 λ_g

Fig 9—SWR curves for a ³/₄-inch water pipe circular waveguide-to-coax transition. Probe diameter=0.050 inches.
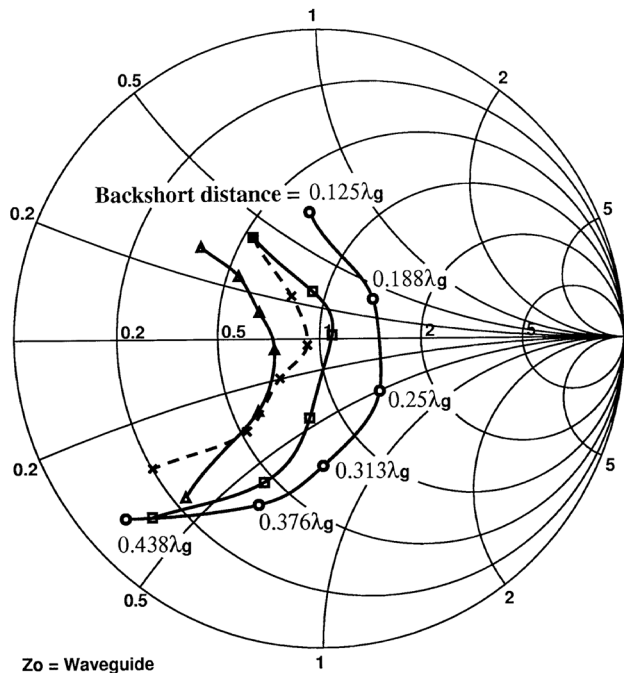
Probe diameter = 0.062"
- ○ Probe length = 0.215" = 0.109 λ_g
- ▲ Probe length = 0.250" = 0.127 λ_g
- ▢ Probe length = 0.305" = 0.155 λ_g
- ✕ Probe length = 0.260" = 0.132 λ_g

Fig 10—SWR curves for a ³/₄-inch water pipe circular waveguide-to-coax transition. Probe diameter=0.062 inches.
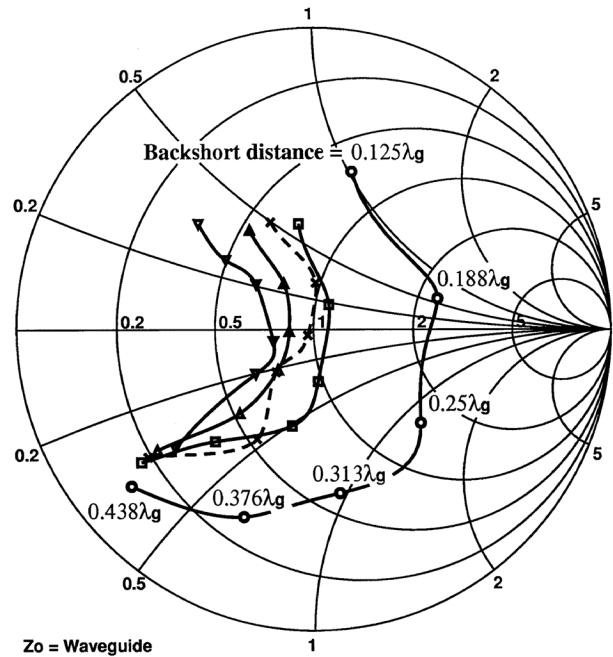
Probe diameter = 0.032"
- ○ Probe length = 0.210" = 0.107 λ_g
- ▲ Probe length = 0.243" = 0.123 λ_g
- ▢ Probe length = 0.268" = 0.136 λ_g
- ◆ Probe length = 0.305" = 0.155 λ_g
- ✕ Probe length = 0.252" = 0.128 λ_g

Fig 11—SWR curves for a ³/₄-inch water pipe circular waveguide-to-coax transition. Probe diameter=0.032 inches.
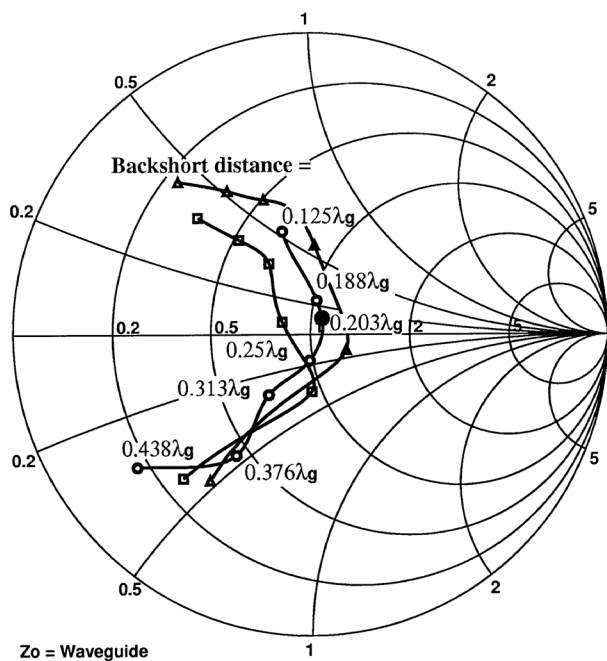
Probe diameter = 0.020"
- ○ Probe length = 0.255" = 0.129 λ_g
- ▲ Probe length = 0.295" = 0.150 λ_g
- ▢ Probe length = 0.355" = 0.180 λ_g

Fig 12—SWR curves for a ³/₄-inch water pipe circular waveguide-to-coax transition. Probe diameter=0.020 inches.
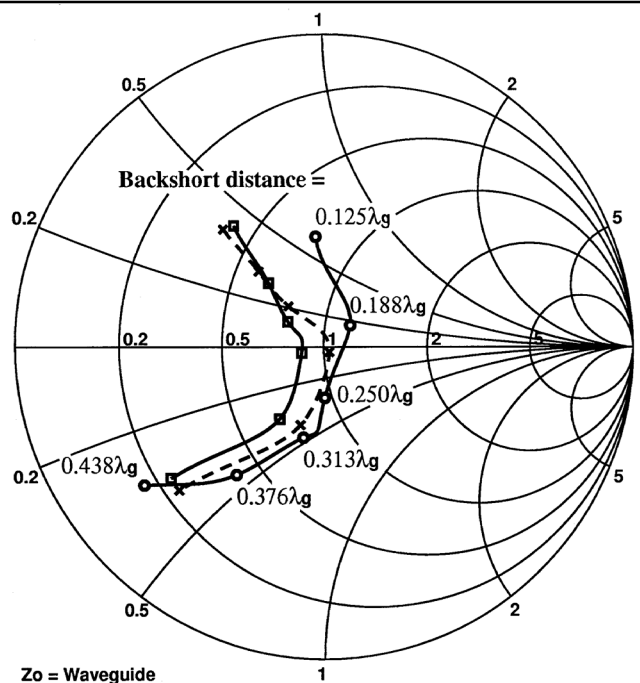
Probe diameter = 0.010"
- ○ Probe length = 0.245" = 0.124 λ_g
- ▲ Probe length = 0.285" = 0.145 λ_g
- ✕ Probe length = 0.265" = 0.135 λ_g

Fig 13—SWR curves for a ³/₄-inch water pipe circular waveguide-to-coax transition. Probe diameter=0.010 inches.

0.77 λ, the full useful range of circular waveguide.

Impedance is plotted versus frequency in Fig 14 for a probe diameter of 0.020 inches, with the impedance curve forming a tight circle around the bull's-eye over the full frequency range. Notice that these impedances are with respect to the waveguide wave impedance, Z_0, which changes with frequency; what we are plotting is SWR and phase at each frequency. Even though Z_0 changes by roughly a factor of two across the frequency range, this empirically-designed transition is able to match it to the coaxial 50-Ω characteristic impedance.

Larger probe diameters (in Fig 15) also produce reasonably tight groupings near the center of the Smith chart, indicating a broadband match. Very small probe diameters are less forgiving, with narrower bandwidth as shown in Fig 16, so the #14 wire at 1296 MHz probably also has a narrowband
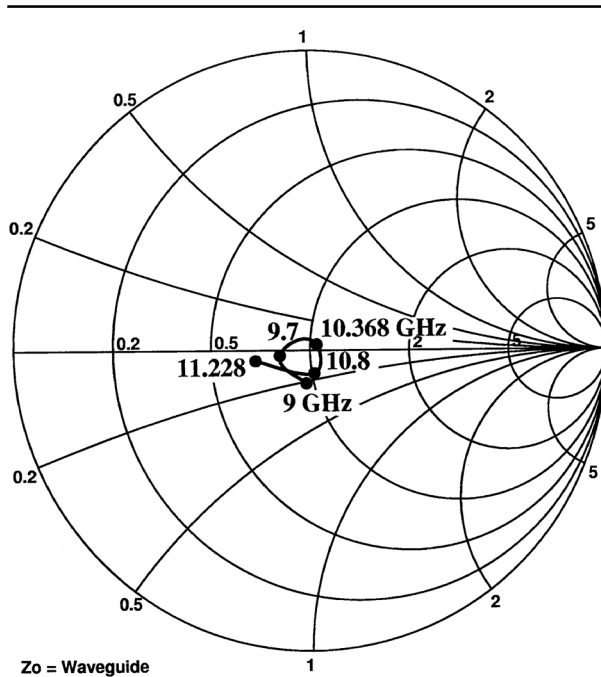
Fig 14—Waveguide impedance variation with frequency for probe diameter=0.020 inches; length=0.255 inches; backshort=0.430 inches.

Smith Chart for Dummies

The Smith chart intimidates many people, including many electrical engineers. For the examples in this article, it isn't necessary to understand the Smith chart—simply think of it as a target, with a bull's-eye in the center where the circle labeled "1" crosses the horizontal (resistive) line. We are trying to hit that bull's-eye, and anything that gets us closer to it is a move in the right direction. Our bull's-eye, in the center of the chart, is the resistive part of the waveguide characteristic impedance (not necessarily 50 Ω) with no reactive component. Any mismatch moves us away from the center of the chart. The shaded bull's-eye in the center of the Smith chart here is the area where SWR is less than 1.5:1, representing a reasonably well matched transmission line.—*W1GHZ*
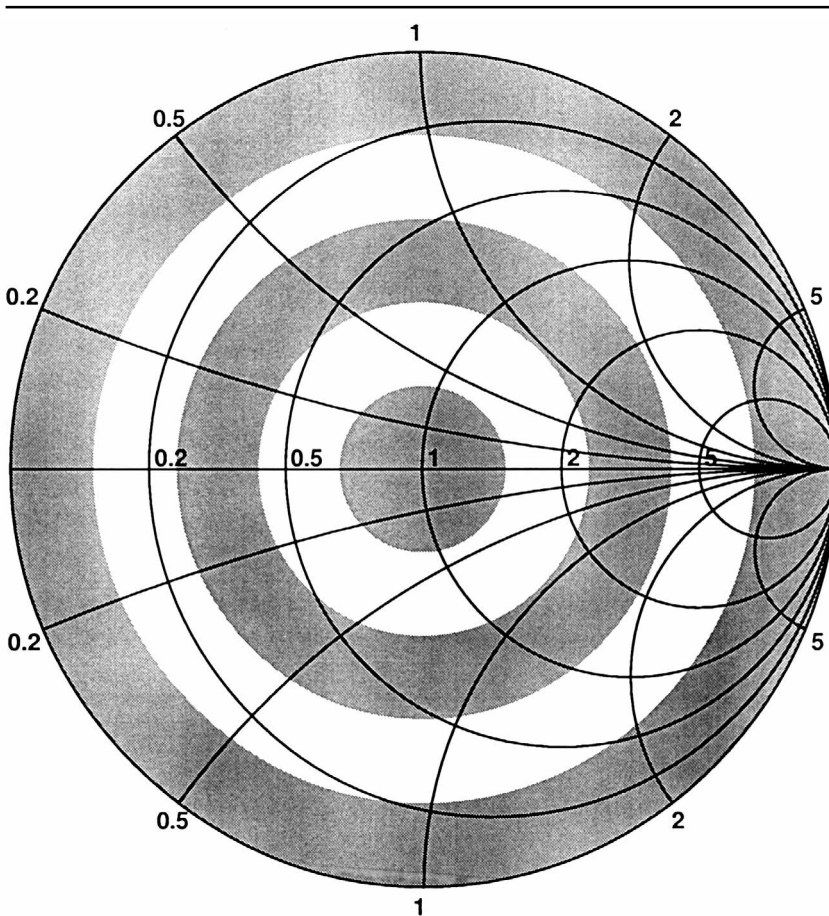
Fig A—For this article, we can imagine a Smith chart as a target.

character. Fig 17 summarizes the best dimensions that I found; the probe length and backshort distance are fairly constant for all probe diameters except for the smallest diameter, 0.010 inches.

The final question is reproducibility. The most convenient probe diameter was 0.050 inches, since that is the diameter of common SMA-connector center pins. So fabrication is simply a matter of cutting and filing the pin to
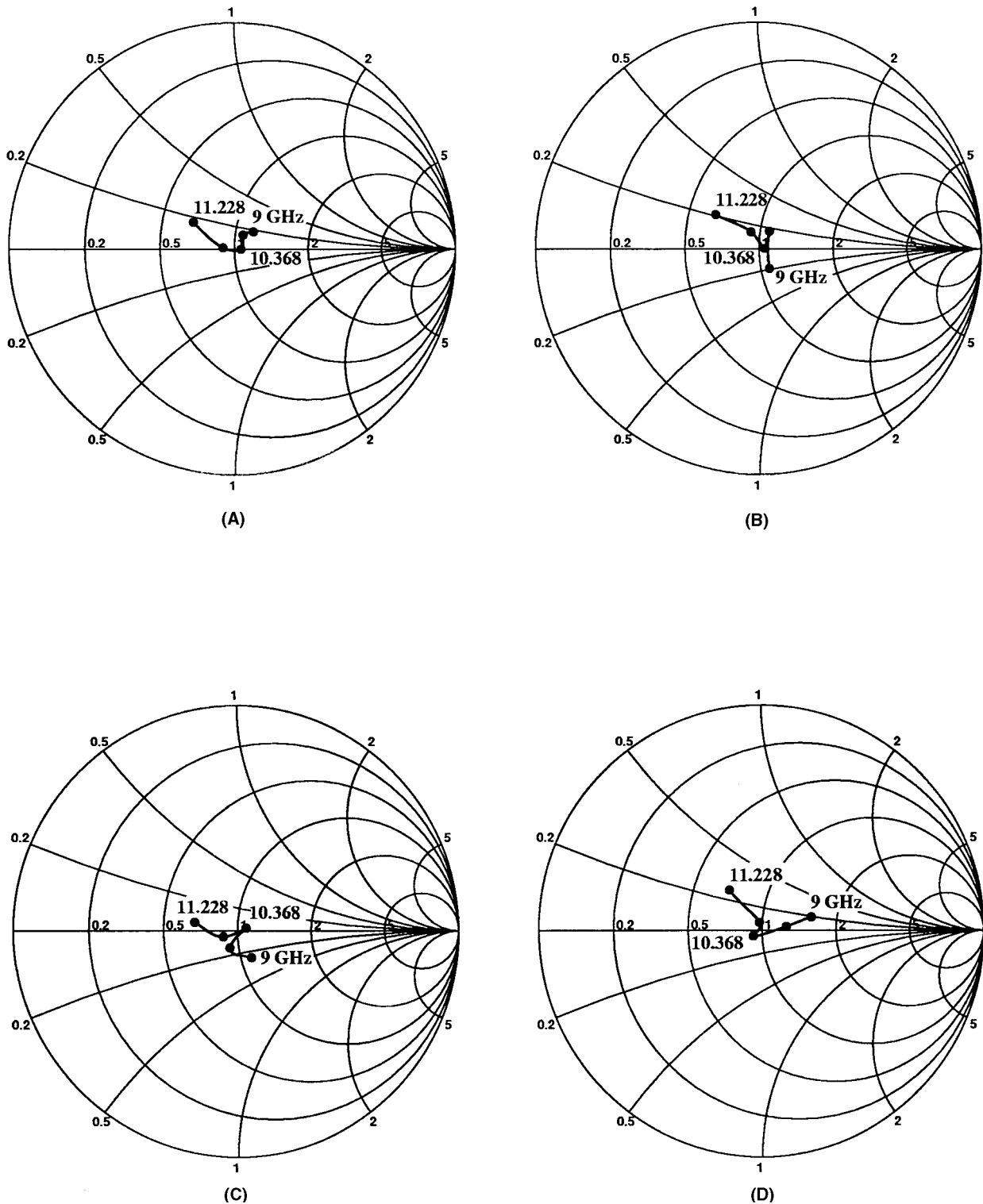
Fig 15—SWR curves for larger probe diameters. (A) Probe diameter=0.032 inches; length=0.252 inches; backshort=0.485 inches. (B) Probe diameter=0.040 inches; length=0.252 inches; backshort=0.493 inches. (C) Probe diameter=0.050 inches; length=0.248 inches; backshort=0.440 inches. (D) Probe diameter=0.062 inches; length=0.260 inches; backshort=0.500 inches.

the desired length read from Fig 17. I cut up another hexagonal plumbing fitting to provide several flats for connector mounting, then soldered one to the side of a piece of pipe as well as a small brass sheet for a backshort. I drilled and tapped holes in the flat for the connector and screwed it in place. The transition worked well, providing an SWR of about 1.05 feeding a diagonal horn, so I assembled three more, using flats from the hexagonal fitting. An alternative construction technique is to use an SMA connector with a threaded body, which allows for some adjustment. Screwing the connector flange to a flat surface is more robust and repeatable, however, and adjustment isn't necessary if we have good dimensions.

I made the final measurements of the four transitions from the coax input using an automatic network analyzer. Fig 19 shows the SWR of all four transitions, each feeding a diagonal horn[4] with SWR of about 1.06. The worst transition has an SWR of about 1.11 at 10.368 GHz, while the others are 1.05 or lower. The SWR is under 1.5 from 9.8 to 11+ GHz, a reasonable bandwidth. To measure the loss, I connected two transitions together with a simple plumbing joint slipped over them. The loss for a pair, shown in Fig 20, is 0.23 dB at 10.368 GHz and under 0.3 dB from 9 to 11 GHz. At 8.5 GHz, the cutoff frequency is obvious; waveguide makes an excellent high-pass filter.

Scaling to Other Frequencies

I believe the transition should scale well to other frequencies. The hardest part is finding a pipe of suitable diameter, between 0.6 λ and 0.76 λ. Then the other dimensions can be scaled directly to the ratio of pipe diameters, starting with a convenient probe diameter and taking the dimensions from Fig 17. All dimensions must be scaled by the same ratio! If the probe length is made slightly long, it can be trimmed for best SWR.

Antenna Applications

The real purpose of the circular waveguide was to investigate antennas with circular cross sections, particularly feed horns. I built several feed horns for offset dishes, including the ones shown in Fig 21: a conical horn, two large W2IMU dual-mode feed horns and a rectangular horn feed for DSS offset dishes. I also built a diagonal horn,[5] a feed horn with a square cross section rotated so that the probe is parallel with the diagonal. I made SWR and sun-noise measurements on these
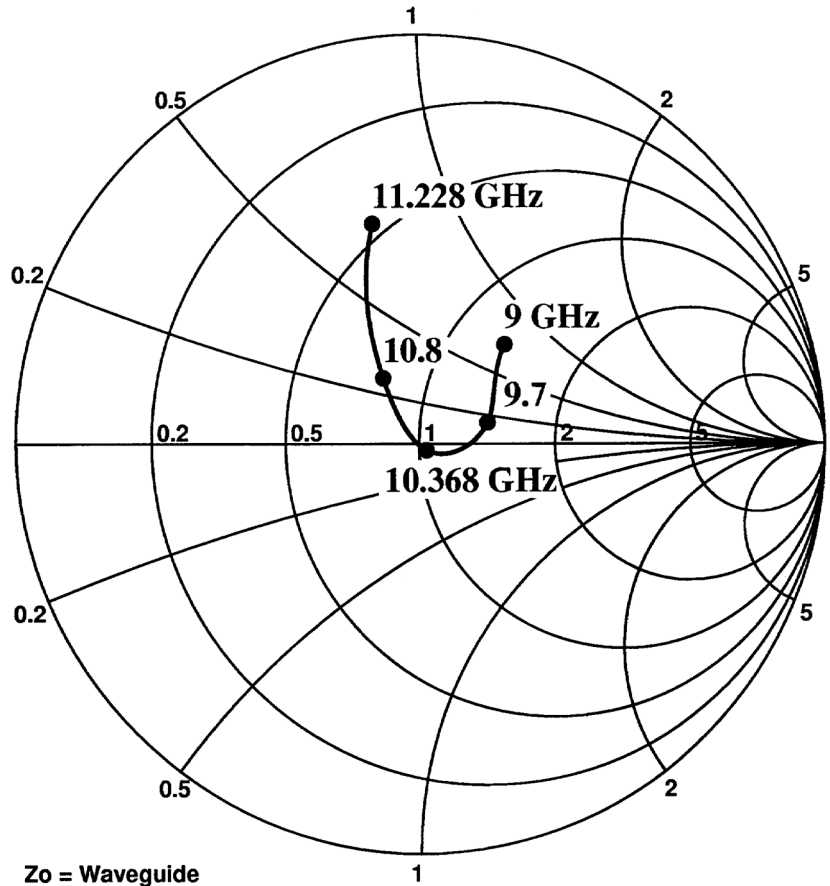
Fig 16—SWR curves for a very small probe diameter. Probe diameter=0.010 inches; length=0.265 inches; backshort=0.616 inches.
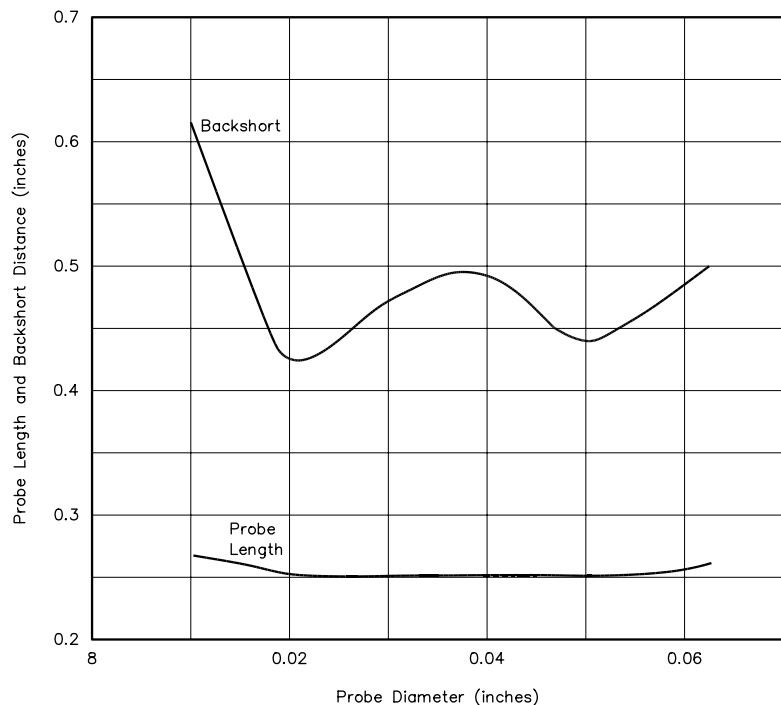
Fig 17—Circular waveguide-to-coax transition—best dimensions.

horns as well as several surplus corrugated horns with circular-waveguide inputs, which I machined to fit to $3/4$-inch pipe. All the homemade horns had SWRs better than 1.1. The diagonal horn had particularly good SWR but disappointing feed performance.

The two W2IMU dual-mode feed horns are large versions, optimized for offset dishes as described in the W1GHZ Microwave Antenna Book—Online.[6] The original W2IMU feed[7] provides best performance for an f/D around 0.5. The two dual-mode horns in Figure 21 are designed for larger f/D required for offset dishes: the larger version in Figure 21 is dimensioned for $f/D = 0.8$ and the smaller for $f/D = 0.7$. For a dual-mode horn to properly illuminate a larger f/D, not only must the aperture diameter increase, but also the length of the output section must increase and the flare half-angle must decrease. The dashed lines in Figure 22 illustrate these changes from the original version depicted by the solid lines. The two large dual-mode feed horns provided the highest efficiency I've measured to date, slightly higher than the rectangular feedhorn[8] I designed for offset dishes. One interesting result was that the efficiency was slightly higher with circular polarization than with horizontal polarization. The middle horn in the table is currently working well feeding a one-meter offset dish as part of my 10-GHz periscope[9, 10] antenna system.

The dimensions for these two larger dual-mode feeds as well as the original W2IMU feed are shown in the following Table 1. The dimensions are shown in wavelengths and may be scaled to any frequency. Dimension A is not shown in the table; it is the diameter of the input circular waveguide feeding the horn. The diameter of the input waveguide does not affect horn performance as long as only the TE11 mode is propagated. As we saw previously, a range of waveguide diameters is usable.
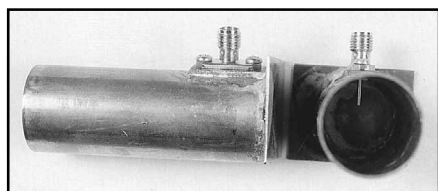
Fig 18—A completed circular waveguide-to-coax transitions for 10 GHz.

Table 1—Dimensions two dual-mode feeds and the original W2IMU feed

Dimensions shown in wavelengths may be scaled to any frequency.

f/D	Flare	B= Aperture Diam.	C= Output Length
0.55	30°	1.31 λ	1.31 λ
0.7	27.4°	1.63 λ	2.8 λ
0.8	24.9°	1.79 λ	3.52 λ

1	10.375000 GHz 1.1130 VSWR
2	9.800000 GHz 1.4428 VSWR
3	11.450000 GHz 1.3716 VSWR

Start: 8.000000 GHz
8/10/99 11:58:20 AM   W1GHZ   Stop: 12.000000 GHz   8720D
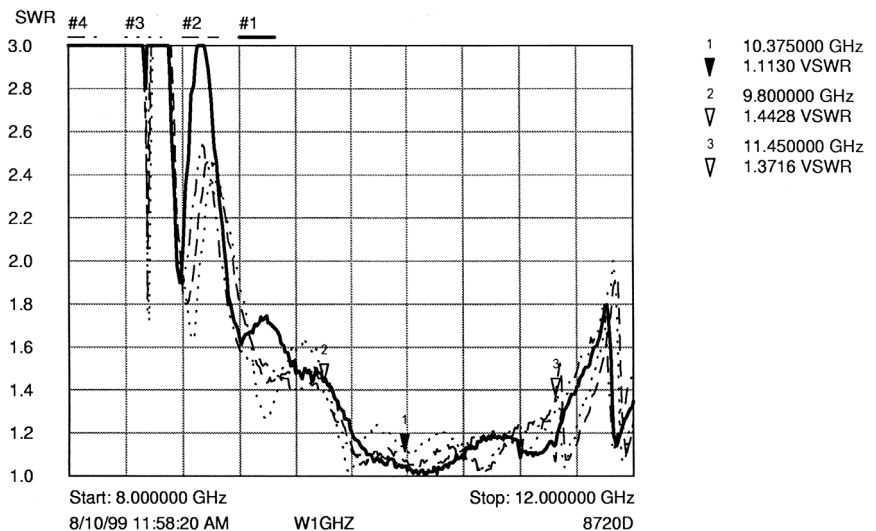
Fig 19—Measured SWR for the completed $3/4$-inch circular waveguide transitions with a diagonal-horn load.

1	10.362500 GHz -0.2297 dB
2	9.000000 GHz -0.3041 dB
3	11.000000 GHz -0.2294 dB
4	10.362500 GHz -29.6220 dB
5	8.500000 GHz -6.1196 dB

Start: 6.000000 GHz
8/10/99 12:05:12 PM   W1GHZ   Stop: 16.000000 GHz   8720D
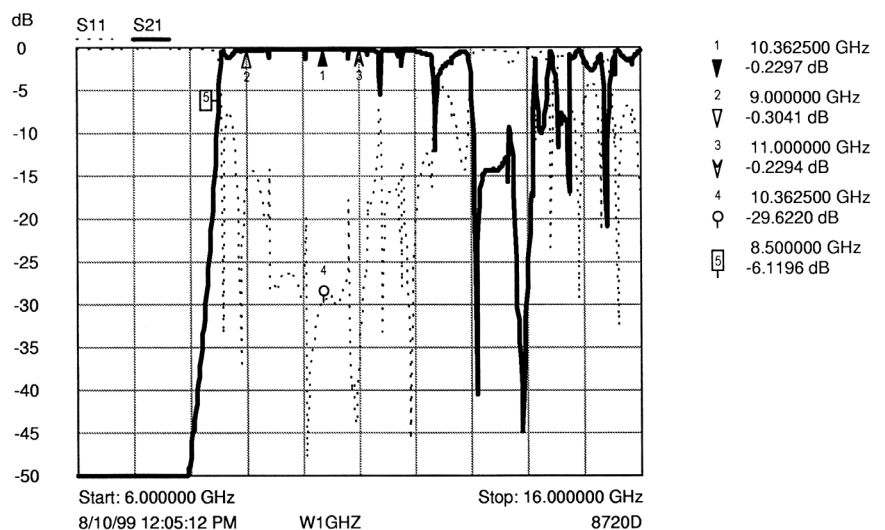
Fig 20—Losses for a mated pair of $3/4$-inch circular waveguide transitions.

Fig 21—10-GHz feed horns for DSS offset dishes.
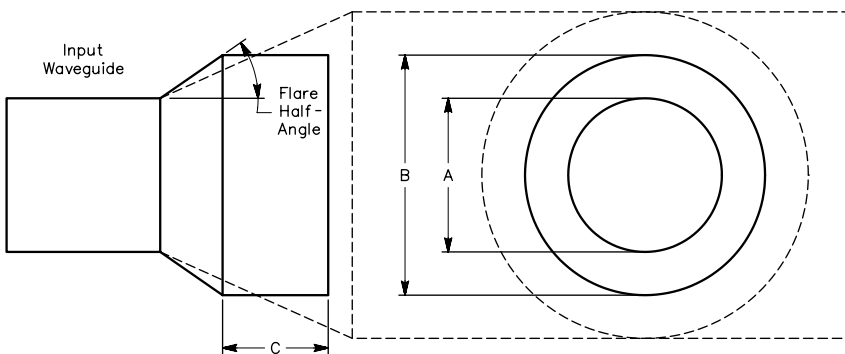
Fig 22—The W2IMU dual-mode feed horn. See Table 1 for dimensions.

Conclusions

Circular waveguide for 10 GHz use made from ordinary copper pipe is both useful and inexpensive. Using simple test equipment and old-fashioned experimental microwave engineering, I have found some good working dimensions for quality circular-waveguide components. In the process, I learned more about circular waveguide. I hope this demonstrates that fancy test equipment is not always necessary for microwave work.

Notes

[1] G. C. Southworth, *Principles and Applications of Waveguide Transmission*, (Princeton, NJ: Van Nostrand, 1950).

[2] D. M. Pozar, *Microwave Engineering*, second edition, (New York: Wiley, 1998) pp 79-82.

[3] G. C. Southworth and A. P. King, "Metal Horns as Directive Receivers of Ultra-Short Waves," *Proceedings of the IRE*, February 1939, pp 95-102. (reprinted in A. W. Love, *Electromagnetic Horn Antennas*, IEEE, 1976, pp 19-26.)

[4] R. Chatterjee, *Elements of Microwave Engineering*, (Ellis Horwood Limited, 1986; New York: Halsted Press, 1986) pp 168-169.

[5] A. W. Love,"The Diagonal Horn Antenna," *Microwave Journal*, March 1962, pp 117-122 (reprinted in A. W. Love, *Electromagnetic Horn Antennas*, IEEE, 1976, pp 189-194.)

[6] You can view this at **www.qsl.net/n1bwt/ preface.htm**; go to Chapter 6.5.

[7] R. H. Turrin, W21MU, "Dual Mode Small-Aperture Antennas," *IEEE Transactions on Antennas and Propagation*, AP-15, March 1967, pp 307-308. (reprinted in A. W. Love, *Electromagnetic Horn Antennas*, IEEE, 1976, pp 214-215.)

[8] P. C. Wade, N1BWT, "More on Parabolic Dish Antennas," *QEX*, Dec 1995, pp 14-22.

[9] P. Wade, W1GHZ, "10 GHz without Feed-Line Loss," *Proceedings of the 24th Eastern VHF/UHF Conference*, (Newington, Connecticut: ARRL, 1998), pp 227-237.

[10] P. Wade, W1GHZ, "Periscope Antenna Systems," *Proceedings of Microwave Update 2000*, Trevose, Pennsylvania, 2000, pp 182-202.

RF

By Zack Lau, W1VT

A Directional-Feedback Amplifier

Many RF amplifiers use feedback circuitry to improve performance. Simple resistive feedback networks commonly used have significant disadvantages. They provide an output-to-input path, decreasing isolation. They also increase the noise figure. The increase in noise figure can be overcome through the clever use of transformers—a design by

225 Main St
Newington, CT 06111-1494
zlau@arrl.org

Norton combines excellent matching, low noise figure and high intercept points.[1] However, the gain is low, and

[1]Notes appear on page 51.

the design suffers from even worse output-to-input isolation. This may make a Norton amplifier difficult to use. One solution is to use directional couplers to generate the feedback path, so that re-

Table 1—Measured performance data for the 20-dB gain preamplifier

Frequency (MHz)	MS11 (dB)	MS12 (dB)	MS21 (dB)	MS22 (dB)	Noise Figure (dB)
2	−19	−54	21.5	−34	
5	−35	−51	21.2	−28	
10	−28	−47	20.3	−24	3.2
12	−26	−44	19.8	−23	3.2
20	−22	−40	18.0	−20	3.2
30	−17	−36	15.7	−19	3.3
50	−17	−31	12.2	−17	3.4
100	−21	−25	7.0	−14	5.1
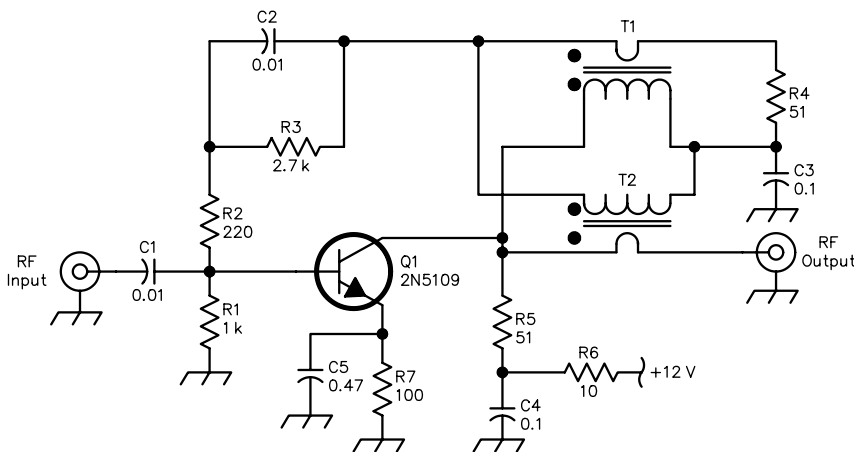
Except as indicated, decimal values of capacitance are in microfarads (μF); others are in picofarads (pF); resistances are in ohms; k = 1,000.

Fig 1—Schematic of the 20-dB gain preamplifier using directional feedback.
Q1—2N5109 with heat sink
T1, T2—10:1 transformers. Primary is 10 turns of # 28 AWG enameled wire. Secondary is a component lead stuck through the hole of FT-37-43 toroid core. Phasing is important.
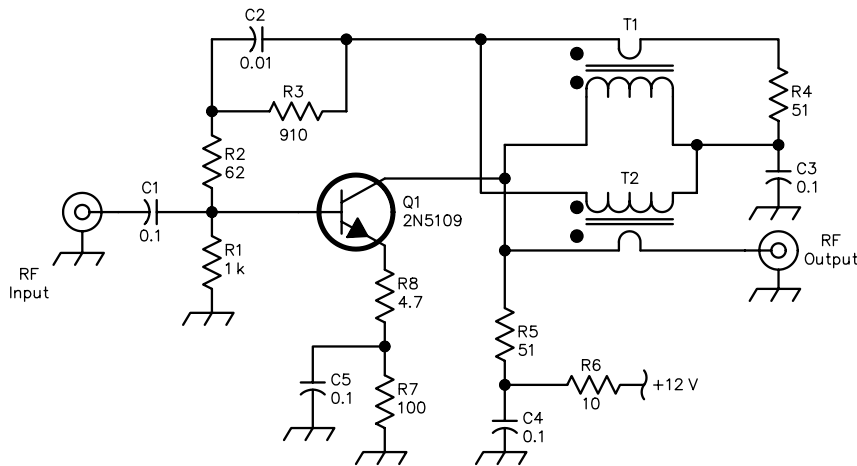
Fig 2—Schematic of the 12-dB gain preamplifier using directional feedback.

Q1—2N5109 with heat sink
T1, T2—10:1 transformers. Primary is 10

turns of # 28 AWG enameled wire.
Secondary is a component lead stuck

through the hole of FT-37-43 toroid
core. Phasing is important.

Except as indicated, decimal values of capacitance are in microfarads (μF); others are in picofarads (pF); resistances are in ohms; k = 1,000.
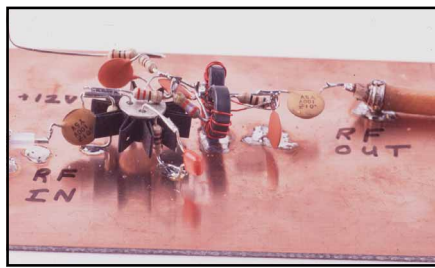
(A)
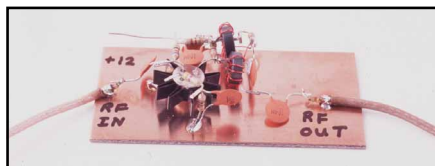
(B)

Fig 3—(A) Photograph of the 20-dB gain preamplifier using directional feedback. (B) Photograph of the 12-dB gain preamplifier using directional feedback.

Table 2—Measured performance data for the 12-dB gain preamplifier

Frequency (MHz)	MS11 (dB)	MS12 (dB)	MS21 (dB)	MS22 (dB)	Noise Figure (dB)
2	−20	−49	11.5	−29	
5	−27	−49	11.7	−37	
10	−28	−45	11.7	−34	6.0
12	−32	−44	11.7	−33	5.9
20	−33	−39	11.2	−28	6.0
30	−32	−35	10.5	−24	6.0
50	−29	−31	8.7	−21	6.3
100	−23	−25	5.0	−16	6.8

flections from a load have minimal effect on circuit performance.

The 20-dB preamplifier of Fig 1 is my first attempt at a design using directional feedback. It features high reverse isolation, measuring as high as 54 dB at 2 MHz. Performance data is shown in Table 1. In contrast, a Norton amplifier configured for 11 dB of gain has just 12 dB of isolation. The input and output return losses are also good at HF. The 3-dB NF is moderately low, though quite a bit higher than the 1 dB possible with a highly optimized Norton design.[2] The noise figure was measured with an HP-8970 with a HP-346A noise source. Due to the 4-MHz measurement bandwidth and 10 MHz lower-frequency limit of the HP-8970, noise figure was not measured below 10 MHz. The output intercept is +12 dBm, which means that the input intercept is just −8 dBm.

I used a directional coupler at the output to obtain a unidirectional feedback path. By only feeding forward power back to the amplifier input, the feedback path doesn't vary with load conditions. Just as important, reverse power primarily ends up in R4, the reverse-power load termination. With a typical broadband amplifier, reverse power is fed back to the input of the amplifier by the feedback network. The directional coupler works as an isolator, a one-way gate for feedback signals. Obviously, properly phasing of the windings is important for proper performance.

I used a pair of transformers T1 and T2, to make a broadband directional coupler. With a turn ratio of 10:1, the coupling is −20 dB. I chose mix-43 ferrite material, for a good compromise between loss and permeability. I adjusted the exact level of feedback by varying R2 in order to optimize the input return loss. R3 and bypass capacitor C2 allow the bias current to be adjusted without changing the feedback network significantly at RF. Similarly C3 provides a RF bypass so that R4 does not short out the dc bias circuit. R5 simultaneously provides dc power to the transistor and a 51-Ω termination to the directional coupler. The performance of the amplifier seems good, although the feedback port of the directional coupler isn't properly terminated in a 51-Ω load.

The transistor's intrinsic emitter resistance significantly affects the input match, since there is little feedback from the emitter to ground. This may actually be an advantage in some cases—the input impedance can be adjusted to match a preceding circuit by adjusting the dc bias. Increasing the current lowers the input resistance. Adding an emitter-feedback resistor can raise the input resistance and reduce the sensitivity of the input match to bias current—but this also decreases the gain of the amplifier. It also raises the noise figure.

The revised circuit using emitter

feedback is shown in Fig 2. The current is increased from 20 to 40-mA, which significantly reduces the input resistance. More importantly, it helps improve the output intercept. To further improve the output intercept, the gain is lowered by reducing the shunt feedback resistance from 220 to 62 Ω. This improves the input intermodulation-intercept point, as less gain means that stronger signals are required to reach the compression point of the amplifier. However, the input impedance is further reduced with shunt feedback. The input impedance is optimized to 50 Ω by increasing the emitter-feedback resistance with a 4.7-Ω resistor. This also has the effect of making the input impedance less sensitive to bias-current variations. The net effect is a decrease in gain to

12 dB and an increase in input intercept from –8 dBm to +11 dBm. This is a substantial improvement of 19 dBm. Photographs of the two amplifiers are shown in Fig 3.

The reverse isolation is degraded only a few decibels, and it still represents a substantial improvement over conventional amplifiers. A typical post amplifier has perhaps 20 dB of isolation. In contrast, the directional feedback amplifier has over 40 dB of isolation, an improvement of 20 dB. Not surprisingly, the increase in feedback has resulted in a noise figure rise from 3 to 6 dB.

The directional feedback concept should be useful in designs that require lots of flexibility, such transceivers where a lot of switching is used to improve functionality. These ampli-

fiers can be used to insure that output ports don't become unwanted "input" ports. This is particularly important when "tapping off" low-level signals for features like panadapters and subreceivers. I hope that a clever designer can further improve upon these designs. The noise figure and input intercept, while good, might be further improved. Who will be the first to design an amplifier that has it all—low noise figure, high input and output intercepts, low return loss and high reverse isolation?

Notes

[1]W. Hayward, W7ZOI, *Introduction to Radio Frequency Design*, pp 218-219.
[2]Z. Lau, KH6CP/1 now W1VT, "6-Meter Transverter Design," *QEX*, Sep 1995, pp 24-30.

Tech Notes

Build a Data Acquisition System for Your Computer

By Steven C. Hageman

[*Gathering, storing and interpreting data can be a tedious and time-consuming task! Steve Hagman shows how computers can be put to work measuring and storing information garnered from external sources*—Peter Bertini, K1ZJH, *QEX* Contributing Editor, k1zjh@arrl.org]

I first began experimenting with A/D converters using the parallel port of a personal computer back in 1994. Since this was done under the DOS operating system, my software needed to *peek* and *poke* the parallel port for data exchanges. I built a half dozen or so designs ranging from 8 to 12 bits of resolution; and while these were moderately useful, the programming for one particular project was not generally adaptable to the next. Today, Windows is firmly entrenched as the operating system of choice; although direct parallel-port access is blocked under the Windows environment.

By 1996, I was experimenting with Microchip Technology microcontrollers, and I was using the computer's serial port (sometimes called RS-232, although that standard has been superceded; Windows identifies serial ports as COM1, etc) for communications instead of the parallel port. This offers much more flexibility; I can now communicate at reasonable data rates—in both directions—between the PIC device and the personal computer. Windows supports serial communication, and nearly every PC has at least one available serial port. A serial port may be the only means of communicating with a notebook computer.

Enter the PIC-DAS

The PIC-DAS (Programmable Interface Controller—Data Acquisition System) features a 12-bit, eight-channel A/D converter, a 12-bit, four-channel D/A converter and includes an 8-bit bidirectional digital I/O port. The PIC-DAS is nearly universal and—unlike my earlier projects—is programmable under either DOS or Windows. Most programming languages support serial communication. The PIC-DAS can be run from Microsoft Office applications, such as Excel and Word, via the VBA (Visual Basic for Applications) language that these applications offer. The PIC-DAS can even communicate with a terminal program. The heart of the system is Microchip Technology's PIC16C63A micro-controller. The PIC controls the bidirectional serial communication with the PC, implements the 8-bit digital I/O port and controls the A/D and D/A converters.

Circuit Description

Fig 1 is the PIC-DAS schematic drawing. The PIC16C63A contains an internal RS-232 UART for communicating with a PC's serial port. Transistors Q3 and Q4 perform the necessary logic inversion between the PIC and the PC port. RS-232 standards specify minimum logic levels between plus and minus 3-25 V. To minimize parts count, the PIC-DAS simplistic design limits logic swings between zero and +5 V. This is not a true implementation of RS-232 standards, but it works fine with a relatively slow 9600 baud rate and cables limited to 15 feet or less. A programmed PIC is available.[1]

A Maxim MAX186 (U1) is a fully self-contained 12-bit A/D converter featuring a built-in 8-channel multiplexer and voltage reference, which is typically 4.096 V. Since the 12-bit converter provides 4096 conversion codes ($2^{12} = 4096$) the resolution is exactly 1 mV/bit. The A/D is controlled by a four-wire serial interface. When the PIC receives an A/D conversion command it responds by sending the proper serial commands to the A/D, which performs the conversion and then returns the conversion results to the PIC. The PIC then completes the command by sending the A/D value back to the PC serial port.

The D/A (U3, a Maxim MAX525) also features 12-bit resolution and provides a single-chip solution with four output amplifiers. The reference for the D/A is taken from the A/D, permitting the same 1-mV-per-bit resolution. Both the A/D and D/A are programmed from the PIC based on commands from the

[1]Notes appear on page 57.

PC serial port using a simple three-wire serial interface

The PIC (U2) implements the control program and the bidirectional 8-bit I/O port. PICs have a unique, relatively high 25-mA per pin current rating, with a package limitation of 200 mA maximum. This, combined with 20-mA clamp diodes, give the PIC excellent immunity to static or misapplied voltages.

The digital I/O port can be setup as any combination of inputs or outputs, and each I/O pin can be read or written to individually or byte-wise, as a single 8-bit wide byte. Port B on the PIC also has built in 'weak' pull-ups that can be switched on or off. These are small current sources (on the order of 250 µA) from the 5-V rail to each output. While the pull-ups are not particularly accurate, they are useful for contact-closure detection and can be set to an "on" or "off" state by a serial command.

The PIC-DAS requires a scant 7 mA while active; this permits about 70 hours of continuous operation using a common 9-V alkaline battery as a power source. Battery power can be

---

Fig 1—The schematic of the PIC-DAS is simplicity itself. Only four ICs and some protection resistors are all it takes to build a 12-bit-resolution universal digital/Analog data-acquisition and control subsystem. The PIC16C63A from Microchip Technology is the heart of the system. This microprocessor contains enough code space to program a reasonable user interface that responds to RS-232 control via ASCII commands.

C1-C3—10 mF, 16 V Tantalum, Kemet, Mouser #80-TC322C106K02

C4-C8—0.1mF, 50 V ceramic, Panasonic (ECU-S1H104KBB), DigiKey #P4923-ND

D1-D3—1N4148, 1N914, etc. silicon diode, small signal

U1 —MAX186BCPP, Maxim, DigiKey #MAX186BCPP-ND

U2—PIC16C63A, Microchip Technology, see text and Note 1.

U3—MAX525BCPP, Maxim, DigiKey #MAX525BCPP-ND

U4—LM2936Z-5.0, National, DigiKey #LM2936Z-5.0-ND

J1-J3—Terminal Block, 2 position 5.08 mm, (13 blocks are needed for this project), On-Shore, DigiKey #ED1643-ND

Y1—8-MHz resonator, Panasonic (PX800MC), DigiKey #PX800MC-ND

Q1—NDF0610, PMOS transistor, Fairchild, Digikey #NDF0610-ND

Q2, Q4—2N7000, NMOS transistor, Fairchild, DigiKey #2N7000-ND

Q3—2N3906 PNP transistor, Mouser #625-2N3906

Misc—Plastic enclosure, RadioShack #270-1808, 8×3×1 inches. DB9—female, RadioShack #276-1538. PCB—Printed Circuit Board—FAR Circuits, see Note 2.
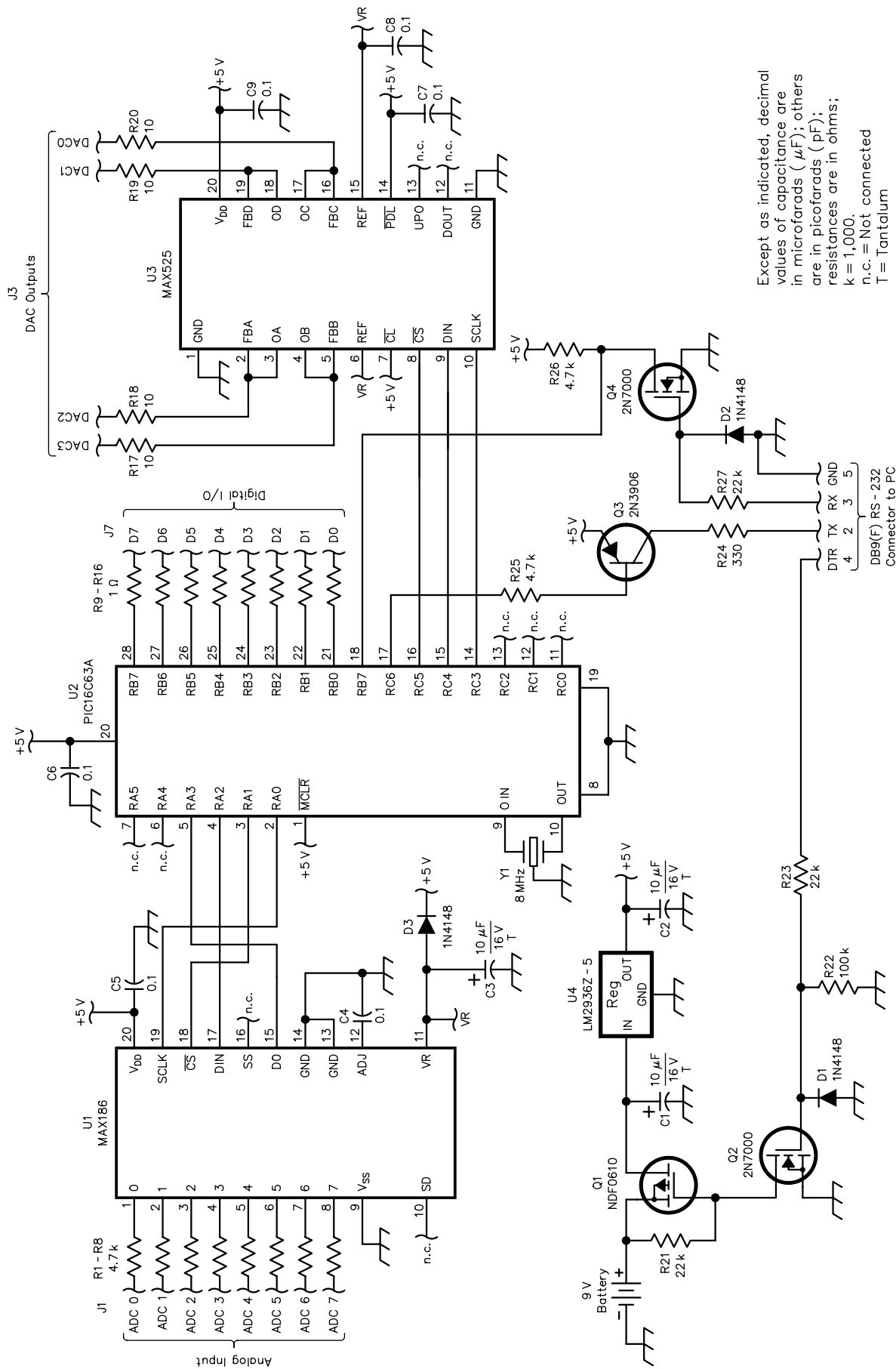
conserved by switching the PIC-DAS on and off via the RS-232 DTR line from the PC serial port. Transistors Q1 and Q2 form an electronic switch to connect the battery to the 5-V regulator (U4) whenever the DTR line is high. U4 is a low-dropout, low-quiescent-current 5-V regulator.

When a terminal program (or other application) opens the COM port connected to the PIC-DAS, the PIC-DAS powers up and is ready for commands in less than 0.5 second. (Note: some programming languages allow independent, programmatic control of the DTR state.) For illustration, let's use an example where we use the PIC-DAS to take hourly temperature readings over a span of several days. We could conserve battery power by opening the COM port only when the data is being recorded and closing the port between readings.

Construction

As shown in Fig 2, the PIC-DAS is built on a PC board that fits nicely into a 8×3×1-inch RadioShack enclosure. A $3/8$-inch notch, cut along one side of the enclosure's top lip, lets the screw-terminal connections protrude from enclosure for easy access. These connectors are modular and supplied in multiples of two. They are snapped together to produce the needed terminal count before soldering to the PC board. Two $1/2$-inch standoffs secure the PIC-DAS board to the enclosure. The construction is not difficult, and sockets

may be used for all of the ICs.

The prototype board may look somewhat strange because of its topside copper layer. The circuit board was not chemically etched, but rather milled on a computer-controlled prototype maker. This system reduces the lead-time for a prototype board from several days to an hour or so. Fig 3 is a close-up of the milled traces on a prototype board. Whole sections can be rubbed out, but this takes some time when the milling tool is only 0.010 inches wide. The advantage of this system is obvious. I started the board layout one morning and by the next morning I had all the parts soldered in place and was testing code. A low-cost circuit board for this project is available from FAR Circuits.[2] Fig 4 shows the completed board, mounted in the enclosure and ready for use.

The PIC-DAS can be built in stages. The heart of the PIC-DAS is the 16C63A micro-controller, and it is always required for operation. You might choose to omit the D/A or both the A/D and D/A converters, but the D/A requires the A/D reference voltage for proper operation.

PIC-DAS Software Support

The software support for the PIC-DAS ranges from native RS-232 Prints-and-Inputs, as would be used with a DOS based *BASIC*, to an ActiveX Control[3] I developed to make Windows programming more automated. In all, there are 10 commands

that the PIC-DAS understands. Seven of these deal with the digital I/O, two with the analog I/O and one is an identification command.

The command structure to the PIC-DAS is as follows:

COMMAND [Parameter 1] [Parameter 2] {CR}

The command is always required, and it may be abbreviated to just the first two characters. A space character follows the command if required for the command Parameter 1. Likewise, a space is required between parameter 1 and 2. The command executes when a carriage return is received by the PIC-DAS. All commands and states are reset when the power to the PIC-DAS is cycled. The command parser in the PIC-DAS only decodes the first two letters in the command string. For example: **DIR 255** works exactly like sending **DI 255** to the PIC-DAS. This is perhaps a less readable but faster form of communication, especially for longer commands. Table 1 lists the available commands.

In *QBasic*[4] the "Direction" command would be sent as:

PRINT #1 "DIR 255"

When the **PRINT** command finishes, it terminates the string with a carriage return. The PIC-DAS then parses the command and executes it.

In the *standard mode,* PIC-DAS executes a command then sends back data (if any) followed by a carriage return. Even if there is no data, a carriage return sent back, this can be
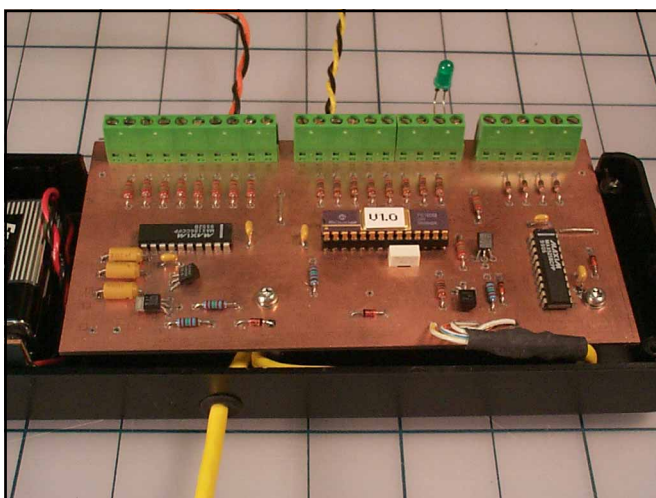
Fig 2—The prototype circuit board was milled from copper-clad fiberglass board with the help of a prototype-milling machine. The screw connectors allow quick and easy connection of the PIC-DAS to whatever circuit you want to control or monitor.
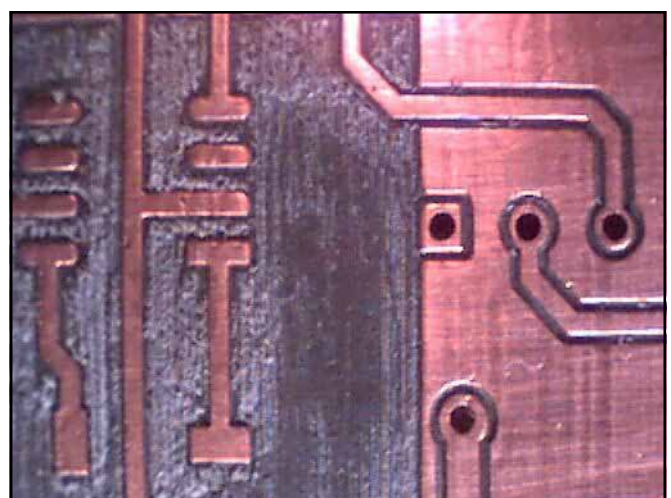
Fig 3—Here is a close-up view of a milled prototype board. The PC-board milling software accepts standard Gerber format data that is produced by all PC board-layout programs. The software then determines the path for the milling tool to isolate the traces on the board. The only differences between this and a production PC board are the lack of plated through holes and the somewhat more difficult task soldering the board.

used to keep a very fast PC from getting ahead of the PIC-DAS. When PIC-DAS encounters a unrecognizable command it responds with the string "UNKNOWN COMMAND." This is useful in debugging code. The sidebar contains a program written in Microsoft *QBasic* that exercises some PIC-DAS commands.

I have written an ActiveX Control for Windows programming. While it is possible to work directly with the serial port under Windows, an ActiveX Control reduces every PIC-DAS command to a simple single-function call. ActiveX Controls are plug-in-like components—they are libraries of commands that are useable across a variety of Windows applications and programming languages. Even the spreadsheet program Excel, or any modern 32-bit programming language, can use ActiveX technology to control the PIC-DAS. A download package, available from the ARRL Web site, has all the source code and many more programming examples for the PIC-DAS (see Note 3).

An Application Example

Temperature readings are arguably the most frequently made measurements in the world. It is easy to use the PIC-DAS for this purpose via a device such as the National Semiconductor LM35, a three-terminal semiconductor temperature sensor with a calibrated output that changes 10 mV/°C. Since the PIC-DAS has a A/D resolution of 1 mV/count, the resolution will be 0.1°C. LM35s require only a small current to operate. I connected them to digital outputs 0 and 1. This allows them to be turned on and off as required; they don't need to operate continuously.

Fig 5 shows how the LM35s were connected to the PIC-DAS. I decided on a simple system that would simultaneous measure and record both indoor and outdoor temperatures. To measure outdoor temperature, I ran 20 feet of twisted-pair wire through the shack wall, where a LM35 is at-

---

A Simple PIC-DAS Control Program

```
'Standard open com port command, you may need to change the COM1
'to match the COM port that your PIC-DAS is connected to
'The extra parameters here disable all the hardware timeouts
'When the COM port is opened the DTR line is enabled which turns on
'the PIC-DAS, to turn off the PIC-DAS, close the COM port again
'Opening the Com port clears the transmit and receive buffers
Open "COM1:9600,N,8,1,CD0,CS0,DS0,OP0,RS," For Random As #1
'===== Digital example
'Even though the DIR command returns no parameters, it does
  send back
'a carriage return when the command completes and we need to remove
'this from the RS232 input buffer.
Print #1, "DIR 0" 'Digital set to all outputs (0=output)
Line Input #1, ret$ 'Discard the CR sent from the PIC DAS
'Count from 0 to 255 on the digital output port
For i% = 0 To 255
Print #1, "OUT "; i% 'This sends "OUT 0", "OUT 1"...etc
Line Input #1, ret$ 'Discard the CR sent back
Next i%
'===== Analog example
'Read A/D channel 0, 100 times, calculate the average
sum% = 0
For i% = 1 To 100
Print #1, "AI 0" 'Read A/D channel 0
Line Input #1, ret$ 'Get result
sum% = sum% + VAL(ret$) 'Sum the reading
Print i%, ret$ 'Print result
Next i%
Print "Average = "; Sum% / 100 'Print the average
Print "End of test program..."
'===== Close the COM port
Close #1
End
```

This example shows a small segment of code, written in *QBasic,* that sets the digital port to all outputs, then counts from 0 to 255 on that port. The next section reads the A/D and prints the result. The "Line Input" function is a special input function that terminates the input command when a carriage return is read, this command is perfect for use with the PIC-DAS as it terminates any string sent back with a carriage return. The VAL( ) function is used to convert the string representation of the A/D value sent back by the PIC-DAS to a real number.

Table 1—PIC-DAS Command Summary

Command	Parameter 1	Parameter 2	Description
DIR	**num**	n/a	Sets the data direction on the digital I/O port. **num** = 0 to 255. 0 = Output, 1 = Input
BSET	**pin**	n/a	Sets **pin** to a logic 1
BCLEAR	**pin**	n/a	Clears **pin** to a logic 0
BIN	**pin**	n/a	Returns the state of **pin** 1 or 0
IN	n/a	n/a	Returns the entire port as a byte. Returns the port value (0-255)
OUT	**num**	n/a	Sets the entire port to byte **num**
PULLUP	**state**	n/a	**state** = 1 for pull-ups on **state** = 0 for pull-ups off
AIN	**ch**	n/a	Reads A/D channel **ch** (0-7) Returns the A/D code read (0-4096)
AOUT	**ch**	**val**	Sets D/A channel **ch** to value **val**
VER	n/a	n/a	Returns a string identifying the firmware version
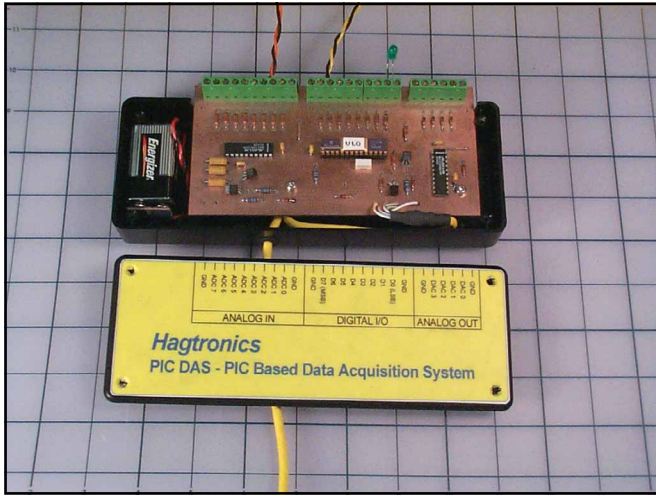
Fig 4—The author's completed PIC-DAS project, assembled and ready to go.
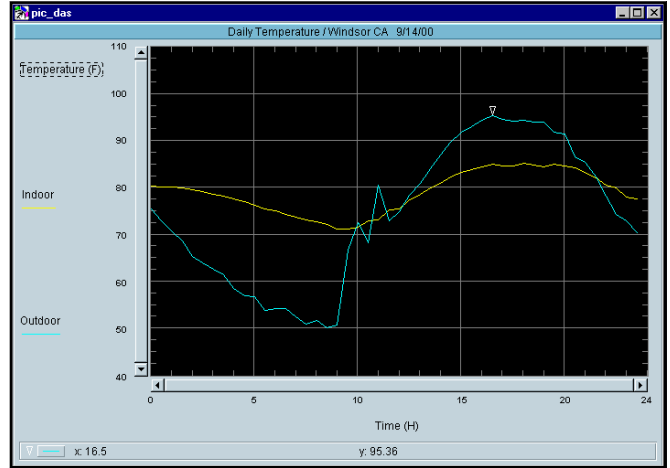
Fig 6—A plot of temperature versus time around my workshop on a warm summer day. The sharp rise in outside temperature corresponds to sunrise. The two peaks after sunrise are when the sun was briefly shining through a tree, directly illuminating the outside temperature sensor. The rest of the day the outside sensor was in the shade.
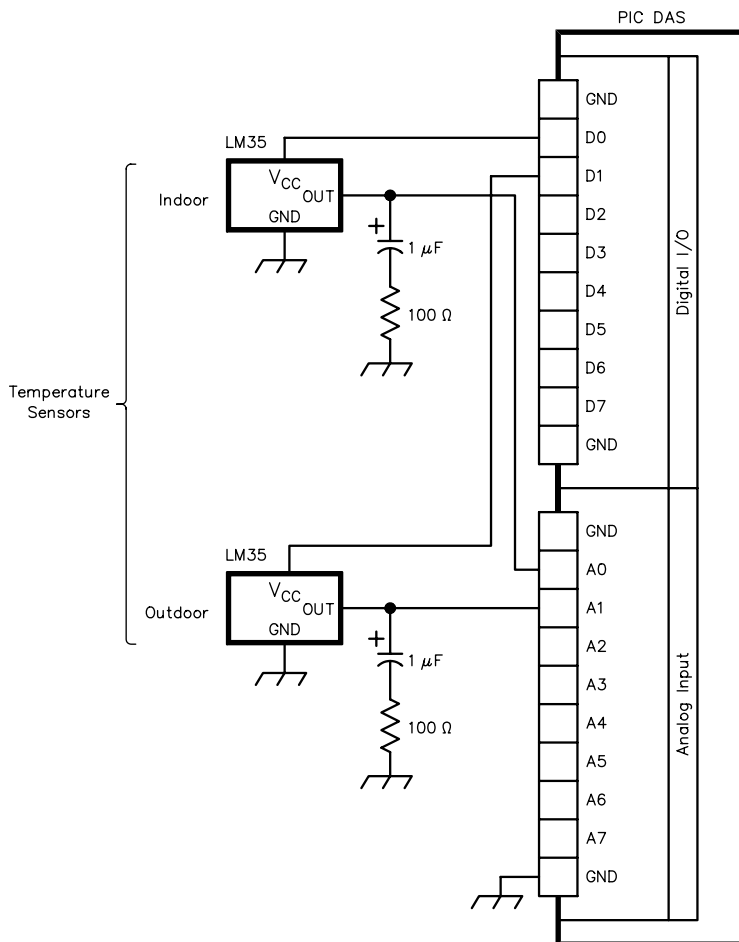
Fig 5—The PIC-DAS was designed to be easy to use; otherwise, you won't want to use it. This application is for an indoor/outdoor thermometer using LM35 temperature sensors. It took 10 minutes to wire the circuit and about 20 minutes to write the 16 programming lines that gather and plot temperature readings over a 24-hour period. Temperature sensing has never been easier thanks to these single-chip silicon temperature sensors from National Semiconductor.

tached to sample the outside temperature. The other sensor was placed near the PIC-DAS. A quick application was written for Agilent *VEE*[5] to read the sensor temperatures every half-hour for a 24 hour period. The entire *VEE* program needs only 16 lines of code to input the A/D values from the PIC-DAS, scale the A/D readings into Fahrenheit degrees and display the data! The data is plotted on a graph in Fig 6, and shows time versus temperature measurements taken around my workshop on a warm, sunny day.

These versatile A/D and D/A converters permit many varieties of closed-loop measurements and control-system operations. I've used similar systems to measure sophisticated switching power supplies by using the DACs to control the input-voltage source and the output constant-current loads. Another application is in the development of battery chargers—characteristic battery curves can be developed by using the DACs to control the charging and discharging current, while the ADC can measure battery voltage and cell temperature. The digital outputs could control relays to switch the charge/discharge path. I've used the digital outputs to quickly program new chips using either a serial or a parallel interface. The use of existing and proven interface designs allows my development time to be better used in evaluating new parts rather than debugging yet another interface.[6]

Notes

[1]A programmed PIC16C63A for the PIC-DAS is available from Steve Hageman (9532 Camelot Dr, Windsor, CA 95492) for $20 postage-paid in the USA and Canada. Foreign orders please add $5 additional for US Postal Service Global Priority Mail shipment. Check or money orders only, no credit cards accepted.

[2]A drilled and etched PC board is available from FAR Circuits, 18N640 Field Ct, Dundee, IL 60118-9269; tel 847-836-9148 (voice and fax); **farcir@ais.net**; **www.cl.ais.net/farcir/**.

[3]The *ActiveX* control for *Windows* (*9x* and *NT*), the DOS *QBasic* examples, the PIC firmware source code and HEX programming file are available for download from the ARRL Web **www.arrl.org/qexfiles/**. Look for PICDAS.ZIP.

[4]*QBasic* is an enhanced BASIC that has been in MSDOS since version 5 was released. *QBasic* is also supplied with *Windows 9x* and it will run on NT as well.

[5]Agilent *VEE* (Formerly HP *VEE*) is a Rapid Development Test Language that allows programs to be built with graphical blocks rather than text lines. It is perhaps the quickest way to program an instrument and produce a professional looking graph of the measurement results available. See **www.agilent.com/find/vee** for more information.

[6]The Author maintains a FAQ page for all of his published projects at **www.sonic.net/~shageman**. Please be sure to check for the latest updates and any other questions that may arise.

Steve has been interested in electronics since the fifth grade. A confirmed analogaholic, Steve has always been fascinated by the magic of receiving signals from the ether. Steve received a BSEE degree from the University of Santa Clara in 1978 and since then has spent his time designing everything from switching power supplies, to embedded systems, to most recently RF circuitry for test and measurement applications. Steve may be reached at **shageman@sonic.net**. □□

New Book

Radio-Electric Transmission Fundamentals, 2nd Edition, by B. Whitfield Griffith, Jr, Noble Publishing Corporation, Norcross, Georgia, 2000, ISBN 1-884932-13-4.

It has been said that a good teacher can take the most complex of subjects and boil it down to something even a simpleton can understand. That is a potentially dangerous statement, because what knowledge is left at the end of the boiling may be sublime, yet insufficient for true understanding. Perhaps it is better to say that a good teacher is one who knows what to include in his or her lessons and what to leave out: That is why being a good teacher is so difficult.

Radio-Electronic Transmission Fundamentals is not really a new title, since the first edition was published in 1962. We are glad to see that Noble have brought it back into print, though, because Whit Griffith found the above-mentioned elusive balance between theory and common-sense reasoning in his explanations of the basic workings of antennas, transmission lines and RF networks. He begins with a history of great discoveries in electromagnetism. Continuing on to fundamental electronic-network theory, he assumes very little expertise on the part of the reader as he asks and answers most of the right questions about electricity. ("What is this thing called 'juice'?" "Why attach all this importance to electric and magnetic fields?")

The bulk of the work concerns itself with showing how electromagnetic field theory neatly predicts many aspects of network, transmission-line and antenna behavior. It should prove interesting reading for those experimenters and engineers who want a clearer picture of what makes things tick. Graduate students and working designers may find it insightful. It is perhaps especially useful to technicians in other fields who need a concise introduction to electromagnetism.

Some of the material treats subjects that are now outmoded. For example, very few will gain from the discussion of computation using slide rules; however, an entire generation of mathematicians has grown up without knowing much about them, and even such quaint stuff may prove useful. The section on vacuum-tube transmitters remains a good introduction to the topic.

This book may be more valuable than many undergraduate texts to the electronics experimenter. It leaves out most of the mathematics, thereby avoiding the fog that is often created by rigorous derivations; but it includes just enough math to start working with RF networks and transmission systems. There is even a short chapter on calculus. It is highly recommended for those who want to take their RF knowledge beyond just a rudimentary understanding of simple circuits.

Whit Griffith, N5SU, went to MIT, then worked in the FCC's Boston field office before joining E. F. Johnson. There, he developed RF components and phasing/coupling equipment for AM directional antenna arrays. He also worked at Continental Electronics in Dallas, designing high-power transmitters and serving as their resident expert on antenna systems. He retired in the late 1980s and currently resides in the Dallas area, where he is still active on the amateur bands. —*Doug Smith, KF6DX* □□

ARRL Technical Awards

The ARRL Board of Directors has created a program of technical awards to recognize service, innovation and microwave development. In taking this action, the Board answered international, federal and organization motivations: International regulations recognize the "technical investigations" conducted by the Amateur Service. The FCC recognizes the amateur's "proven ability to contribute to the advancement of the radio art," the need to promote technical skills and expand the "existing reservoir . . . of trained operators, technicians and electronics experts." The League's purposes include promoting experimentation and education in the field of electronic communication, research, development and dissemination of technical, educational and scientific information.

All of these purposes are fundamental to the amateur service and they deserve recognition. Hence, the Board's initiative to encourage and reward achievement in technical areas. The ARRL technical awards encourage and tangibly reward amateurs who have given outstanding technical service. The awards provide an opportunity to publicly recognize amateur's service, Amateur Radio and its many benefits for society.

ARRL Technical Service Award

The annual Technical Service Award is given to a licensed radio amateur whose service to the amateur community and/or society is exemplary in the area of Amateur Radio technical activities. These include, but are not limited to:

• Leadership or participation in technically-oriented organizational affairs at the local or national level.

• Service as an official ARRL technical volunteer: Technical Advisor, Technical Coordinator, Technical Specialist.

• Communication of technical education and achievements with others through articles in Amateur Radio literature or presentations at club meetings, hamfests and conventions—this includes encouraging others to do the same.

• Promotion of technical advances and experimentation at VHF/UHF, with specialized modes and work with enthusiasts in these fields.

• Service as a technical advisor to clubs sponsoring classes to obtain or upgrade amateur licenses.

• Service as a technical advisor to Amateur Radio service providers, government and relief agencies establishing emergency-communication networks.

• Aid to amateurs needing specialized technical advice by referral to appropriate sources.

• Aid local clubs to develop RFI/TVI committees and render technical assistance to them as needed.

• Aid to local technical-program committees to arrange suitable programs for ARRL hamfests and conventions.

Each Technical Service Award winner receives an engraved plaque and travel expenses to attend an ARRL convention for the formal award presentation.

ARRL Technical Innovation Award

The amateur community has witnessed great technological changes over the past 75 years. Amateurs have been at the heart of many advances in the radio art. Enduring ARRL policy encourages amateurs to remain at the forefront of technological advancement. The ARRL's annual Technical Innovation Award is granted to a licensed radio amateur whose accomplishments and contributions are most exemplary in the areas of technical research, development and application of new ideas and future systems of Amateur Radio. These include, but are not limited to:

• Promotion and development of higher-speed modems and improved packet radio protocols.

• Promotion of personal computers in Amateur Radio applications.

• Activities to increase efficient use of the amateur spectrum.

• Work to alleviate long-standing technical problems, such as antenna restriction, competition for spectrum and EMC.

• Digital voice experimentation.

• Amateur satellite development work to improve portable and mobile communication.

• Improvements in solar, natural and alternative power sources for field applications.

• Practical compressed video transmission systems.

• Spread-spectrum technology and applications.

• Technology development to assist disabled amateurs.

Each Technical Innovation Award winner receives a cash award of $500, an engraved plaque and travel expenses to attend an ARRL convention for the formal presentation.

ARRL Microwave Development Award

The microwave and millimeter bands are great frontiers for amateurs. They present amateurs with a vast test area for development of both new and traditional modes. The annual ARRL Microwave Development Award is given to an amateur individual or group whose accomplishments and contributions are exemplary in the area of microwave development. That is, research and application of new or refined activity in the amateur microwave bands at 1 GHz and above. This includes adaptation of new modes in both terrestrial and satellite techniques.

Each Microwave Development Award winner receives an engraved plaque and travel expenses to attend an ARRL convention for the formal presentation.

Nominate Now!

Count yourself among those who know that technical advancement is not a lost ideal in the amateur community. Now is the time to nominate your colleagues for one or all of the awards described above!

Formal nominations may be made by any ARRL member. Submit support information along with the nomination document, including endorsements of ARRL affiliated clubs and League officials. Nominations should thoroughly document the nominee's record of technical service during the previous calendar year. Information concerning the character of the nominee should be as complete as possible. The ARRL's Amateur Radio Technology Working Group will serve as the award panel, review the nominations and select the winners.

The nominations form for these awards can be found at **www.arrl.org/ead/award/application.html**. Send nominations to: ARRL Technical Awards, 225 Main St, Newington, CT 06111. Nominations must be received at Headquarters by March 31. Send any questions to Jean Wolfgang, WB3IOS, Educational Programs Coordinator, at ARRL Headquarters or e-mail **jwolfgang@arrl.org**. □□

Impedance-Transformation Properties of a π-Output Network

Dear Editor,

One purpose of a π-output network in a linear power amplifier is to transform the impedance at the input of the coax (connected to the antenna-coupling system) to the correct plate load resistance, R_L. For this discussion, we will just examine the case where the π network is tuned to "match" (transform the impedance of) a 50-Ω load resistance to a plate-load resistance of 2000 Ω. The network chosen is the same as that chosen by Bill Sabin for analysis in his letters to the Editor that appear in the May/Jun 2000 and Sep/Oct 2000 issues of *QEX*. It has a *Q* of 12 (as defined in *The ARRL Handbook*). *The Handbook* gives equations for computing *Q1* and *Q2* of the equivalent **L** section, the sum of whose *Q*s is 12. *Q1* and *Q2* are computed to be 10.64 and 1.36, respectively.

The π network has a phase delay from the input RF plate voltage to the RF output voltage across the 50-Ω load resistor. This phase delay is the sum of the phase delays across the two **L** networks:

$$D = \left(\tan^{-1}(Q1) + \tan^{-1}(Q2)\right) \text{ degrees} \qquad \text{(Eq A)}$$

This is 138.34° in the example.

The impedance transformation properties of the π network can be represented exactly (neglecting losses) by a 138.4° length of 50-Ω coax and an ideal transformer at its input, with an impedance step-up of 50:2000 Ω (40:1). See Fig 1. Therefore, we can use a Smith chart, Fig 2, to find the impedance Z_L at the tube plate when the load resistance *R* is varied. We can multiply each of the values of *r* and *jx* at the coax input by 40, or we can add labeling for a Smith chart normalized to 2000 Ω (as in Fig 2).

We start with the π network correctly tuned for *R* = 50 Ω. To find Z_L when *R* = 100 Ω, find 100 + *j*0 on the Smith chart. It is on the 2:1 SWR circle. Move around the circle clockwise for 138.34 electrical degrees (276.68 compass degrees). We find *Z* = 43 + *j*32. Multiplying by 40 gives Z_L = 1720 + *j*1281 Ω. Now change *R* to 25 Ω. Using the same procedure, we find *Z* = 37.4 –*j*27.8 Ω and Z_L = 1496 + *j*1114 Ω. A line drawn from one Z_L to the other is a straight line passing through 2000 + *j*0 Ω at the center of the chart.
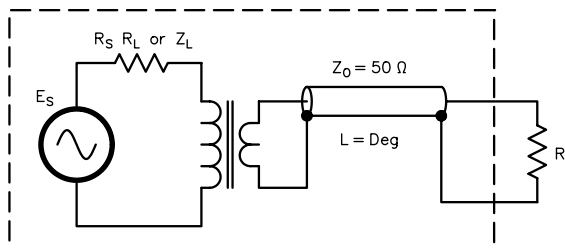
Now do the same for the two load resistances of 55 and 45 Ω [used in the experiment in question]. Their SWRs are 1.1:1 and 1.111:1, respectively. These values of Z_L are identical to the values Mr Sabin plotted in his Fig 1 of the May/June letter. This proves that my simple equivalent *Z*-transformation circuit is valid. It also clarifies why simple, amplitude-only "load-pull" methods of measuring R_S are invalid: It is the network phase delay.

Likewise, in his Sep/Oct letter (p 63), if he had plotted the output impedance (at the 50-Ω end) for different values of R_S on a Smith chart, all points would have fallen on the same straight line. This again verifies his calculations, but plotting the results on a Smith chart makes it much more understandable.

It also clarifies why the SWR is the same at both ends of the π network. If R_S is 6000 Ω (also 150 Ω on the right edge of the chart), the SWR is 3:1. The impedance looking back into the π network output terminal also has an SWR of 3:1, only it is not resistive in this case (the dot on the top edge of the chart).

By now, readers have probably recognized that adding a 41.66° length of coax between the π network and the load resistance would rotate the nearly vertical straight line clockwise so that both the input and output impedances fall on the purely resistive line. This makes the load-pull method of measuring R_S accurate enough to establish that a conjugate match does not exist. It would be good only up to 80% of PEP to keep the RF plate voltage out of the nonlinear region of the tube when the higher value of *R* is used.—*Warren Bruene, W5OLY, 7805 Chattington Dr, Dallas, TX, 75248-5307*
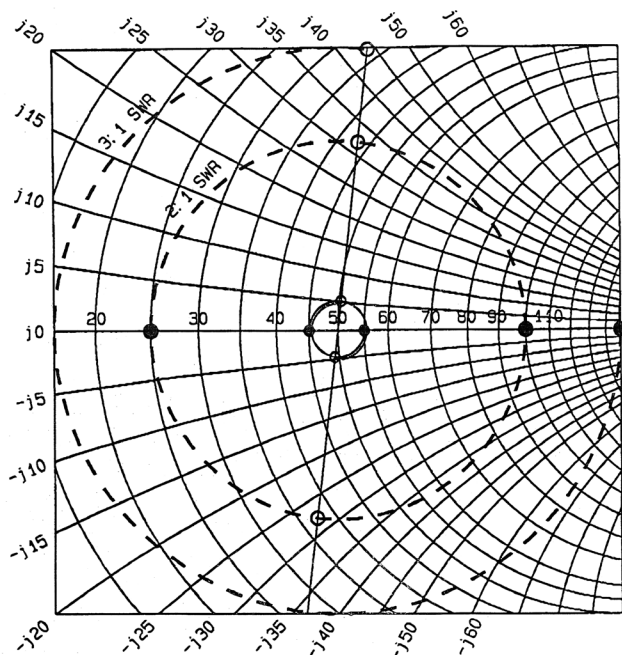
Fig 2—Expanded Smith chart illustrating the effect of varying the load resistance on the input impedance of a π network with 138.4° of phase delay. Consider the chart to be normalized to 50 Ω for the antenna side of the π network (heavy dots on *j*0 line) and normalized to 2000 Ω for the tube side of the π network (diagonal line).

Fig 1—Equivalent circuit of an RF amplifier and output network.

The Conjugate-Match Argument is Dead

I make this declaration now because the proponents of the claim that a conjugate match exists between an RF power amplifier tube plate and its output network have not been able to prove their case. The fact that RF feedback reduces the effective value of R_S without changing a tube's operating conditions is further evidence that no conjugate match exists. Except for Mr Sabin (to his credit), none [of the parties in this discussion] has admitted even their most obvious errors. Let's give [those arguments] the deep six. R I P—*Warren Bruene, W5OLY, 7805 Chattington Dr., Dallas, TX, 75248-5307*

Doug,

This letter is a supplement to Warren's that describes the coax add-on that makes the phase shift of the π network 180°. The resulting circuit is equivalent, under certain conditions, to a conventional transformer, which then makes the load-pull method for determining the r_p dynamic plate/drain/collector resistance a valid procedure. The method used here employs exact design equations. The values of $θ$ (phase delay), L and C are varied to see their sensitivity with respect to R_p, which is the estimate of the r_p value.

Assume a dynamic plate resistance r_p = 4000 Ω. From Chapter 13 of *The 2000 ARRL Handbook*, p 13.6, the design equations for the π network provide the following almost-exact values:

f = 3.75 MHz
$Q0$ = 12
Plate load, Z_{input} = 2000 + j0 Ω
R_L = 50 Ω average, switched between $RR1$ = 45 and $RR2$ = 55 Ω.
$C1$ = 225.751 pF
L = 8.9215 μH
$C2$ = 1155.88 pF
Delay line = 50-Ω coax at the π-network output, $θ$ = 41.662°.

The following procedures are used:

For RR = 45 or 55 Ω, the corresponding complex impedance Z_{in} at the input of the delay line is

$$Z_{in} = 50 \frac{RR + j50\tan(θ)}{50 + j\,RR\tan(θ)} \qquad \text{(Eq 1)}$$

Assume an input current of 1.0 A into the π network. Using nodal circuit analysis, the voltage V_{in} at the output of the π network (same as the input of the delay line) is

$$V_{in} = \frac{\dfrac{1}{jωL}}{\left(\dfrac{1}{r_p} + jωC1 + \dfrac{1}{jωL}\right)\left(\dfrac{1}{Z_{in}} + jωC2 + \dfrac{1}{jωL}\right) - \left(\dfrac{1}{jωL}\right)^2} \qquad \text{(Eq 2)}$$

The output voltage of the delay line is

$$V_o = V_{in}\left(\cos(θ) - j\frac{50}{Z_{in}}\sin(θ)\right) \qquad \text{(Eq 3)}$$

Use the two load-pull values $RR1$ and $RR2$ and the corresponding values of V_o1 and V_o2 in the load-pull formula to find R_r, the calculated estimate of r_p

$$R_r = \frac{V_o1 - V_o2}{\dfrac{V_o2}{RR2} - \dfrac{V_o1}{RR1}} \qquad \text{(Eq 4)}$$

Using these values, R_r = 4000 + j0.028 Ω, an extremely close correspondence to the exact r_p = 4000 Ω. Using the *magnitudes* of V_o1 and V_o2, as measured by an ordinary RF voltmeter, the result is 4000 Ω. This corroborates

mathematically the load-pull procedure for the π network, including the extra length of coax.

Of interest also is the impedance Z_{input} looking into the input end of the π network

$$Z_{input} = \frac{1}{\dfrac{1}{\dfrac{1}{\dfrac{1}{Z_{in}} + jωC2} + jωL} + jωC1} \qquad \text{(Eq 5)}$$

For the two values of RR, 45 and 55 Ω, the two values of Z_{input} are 1800 + j0.003 and 2200 + j0.006 Ω. The average of these two is 2000 + j0.005 Ω. Smaller and larger values of RR will cause Z_{input} to become significantly de-tuned, so it is best to use values close to 50 Ω, such as 45 and 55 Ω.

These design equations were set up in a *Mathcad* worksheet so that variations of parameters could be examined. The sensitivity to the length of the delay line is an especially important quantity. The results are as follows:

Delay line = 41.662° ±10%. R_r = 3938 –j428 Ω, and R_r = 3948 + j391 Ω. Using an ordinary voltmeter, the values are R_r = 3968 and 3974 Ω.

Delay line = 41.662° ±20%. R_r = 3763 –j809 Ω, and R_r = 3832 + j690 Ω. Using voltage magnitudes, R_r = 3877 Ω and 3914 Ω.

These values are close enough for practical purposes, so extremely close delay-line phase values are not necessary but they should be within 20% of the correct value. Note that errors in the delay line produce *reduced* values of R_r. As the delay approaches zero, the magnitude of R_r becomes approximately 2000 Ω.

For a positive error of 3% in capacitance values and a negative 3% error in inductance value, with delay equal to the correct value, R_r calculated to 4100 + j242 Ω, magnitude 4110. This is acceptable. The reader should set up a *Mathcad* (or similar) worksheet to explore error analysis in detail.

The values of inductance and capacitance need to be known within a few percent, and this leads to a problem. In a particular manufactured amplifier that has been tuned and loaded to the output power and linearity (intermodulation distortion) levels that are specified by the manufacturer, the values of $Q0$, L, $C1$ and $C2$ are seldom accurately known. Therefore it is necessary to dig into the amplifier, open up the connections to the components and measure the component values. The accuracy of these measurements *can* be within a few percent, which leads to modest errors in the load-pull calculations. Then we put everything together, calculate the length of delay line needed to achieve 180°, and make the load-pull measurements, with no changes in settings. Of particular interest is the danger of high-voltage entanglement; extreme caution by qualified personnel is *essential*. This is a tedious process, especially if repeated on several frequency bands, so we must ponder the question: "Just how important is this r_p determination for me?" The consensus these days is: "not very important in nearly all Amateur Radio applications."—*William E. Sabin, W0IYH, 1400 Harold Dr SE, Cedar Rapids, IA 52403;* **sabinw@mwci.net**

Dear Bill,

It is not necessary to take the output network apart to determine the extra length of coax needed. You don't even need to know R_L or the Q of the π network.

The straight-line locus of impedance versus load resistance variation extends all the way across the Smith chart with resistance changes from zero to infinity.

Therefore, the following procedure can be used:
1. Tune and load into 50 Ω. It is assumed only a few feet of coax will be used to connect the dummy load to the transmitter.
2. Remove transmitter power.
3. Disconnect the coax from the dummy load and measure reactance looking back into the coax.
4. Find this value of X on the outside edge of a Smith chart normalized to 50 Ω. It will be on the upper right side for a π network.
5. Read the degrees for the reflection coefficient at that point and divide by two. That is the extra electrical length, in degrees, needed to make the total delay 180°.

No soldering iron is required. Sincerely and 73—*Warren*

Introduction to Adaptive Beamforming (Nov/Dec 2000)

Doug,

Your summary of work on adaptive antennas and beamforming was well done. I would like to call attention to some work in the field that I participated in and co-authored.

These studies centered on multiple receiving antennas for GPS to form nulls on jammers. In this case, the GPS signal is noise-like, while the jammers may be noise or de-terministic signals. It may be applicable to the amateur case of reception in conditions of interference by strong unwanted signals, as well as several smaller undesired signals. It is noted in the papers that great lengths have been taken by several investigators to match gain and phase per channel or to have advance knowledge of gain and phase through calibration and look-up tables.

These papers make the claim that given an adaptive sys-tem that is stationary over short intervals (a few tens of samples) and [by] applying weights based on some crite-rion—whether it be maximum unwanted-signal rejection, best S/N or minimum total power, as examples—the adap-tation process will effectively find an optimum without knowledge of channel-to-channel errors, mutual coupling or drift. This comes directly from the math involved provided that the errors are linear and independent. The first paper deals with errors that are nearly independent of frequency, while the second paper deals with errors that change dras-tically across the passband under consideration. In both cases, the added degradation in the settled value of S/N is found minimal compared to cases where complete or partial advance knowledge was available. This may have useful implications for amateur work as it shoots down the need to have perfectly matched stages in each channel. Please refer to the following for more details:

"Commercial Off-the-Shelf (COTS) GPS Interference Canceller and Test Results," *Proceedings of the ION National Technical Meeting*, Jan 1998.

"Impacts of Frequency Dependent Mutual Coupling and Channel Errors on Digital Beam Forming Antenna Performance," ION Inter-national Technical Meeting of the Satellite Division, Sept 98, Nashville, Tennessee.

Photocopies of the articles ($10 each) are available from the Institute of Navigation at **http://www.ion.org/publicationorder.html**. I hope these are of interest.—*David M. Upton, WB1CMG, 25 Harwood Rd RR#1, Mont Vernon, NH 03057-1517;* **david@wb1cmg.mv.com**

Who Invented SSB?

Doug,

In their ad (Nov/Dec 2000, *QEX*), Rockwell-Collins claim they invented SSB. What utter nonsense!

In 1910, Major G. O. Squier, a United States Signal Corps officer, developed an experimental multiplex system. Shortly afterwards, the Bell System (Western Electric, later Bell Telephone Laboratories) realized that one sideband contained all the information and that the carrier could be eliminated if it could be reinserted at the receiving end. Thus, SSB was born about World War I time. The first Bell System SSBSC link went into service in 1914 between South Bend, Indiana and Toledo, Ohio. The first commer-cial SSBSC system went into service in 1918 between Balti-more, Maryland and Pittsburgh, Pennsylvania. By 1928, the Bell System had full-blown SSB transcontinental sys-tems. Transoceanic SSB radio systems were in operation about that time—the Bell System played a large part.

Where did Art Collins fit into this development? Their claim sounds a bit like Al Gore inventing the Internet.—*Ken Uthus, KT7E, 1720 Horseman Ct, Santa Maria, CA 93454-2671;* **kennyuth@pronet.net** ☐☐

Next Issue in QEX/Communications Quarterly

Here at *QEX*, *Next Issue In....* gets written about 13 or 14 weeks before the cover date of that *next* issue. It isn't always easy to predict which articles will appear in each issue until many weeks later. We find that we are batting 0.936—fair, but not good enough, especially for those whose articles are mentioned and then aren't published on time. We can and will do better in the new decade by planning farther ahead. Now, as we step up to the plate once again…

The remaining segments of two series we began quite some time ago will appear in the Mar/Apr 2001 issue. John Stephensen, KD6OZH, has forwarded to us the third part of his ATR-2000 homebrew transceiver art-icle. In case you forgot, John's first two parts appeared in the Mar/Apr and May/Jun 2000 issues. Those covered the vast majority of his receiver design, along with the local oscillator, mixer and product detector. Part 3 covers the power amplifier, noise blanker, control circuitry and power supplies.

Editor Doug Smith, KF6DX, concludes his two-part series on ana-log bandwidth compression of voice signals. Part 1 appeared in the May/Jun 2000 *QEX*. Part 2 goes into the details of his invention and explains how perceptual audio coding packs four or more voice signals into the spectrum of one. WAV-file samples of original, coded and decoded audio will be available from our Web page by the time that issue hits the street.

Cornell Drentea, KW7CD, presents his views on modern synthesizer technology in "Beyond Fractional-N." Cornell was involved in DDS-driven PLL synthesizers before most of us had even heard of them. He shares valuable insights he has gained about them over the years and conducts a careful analysis of the variables and performance in real systems. ☐☐