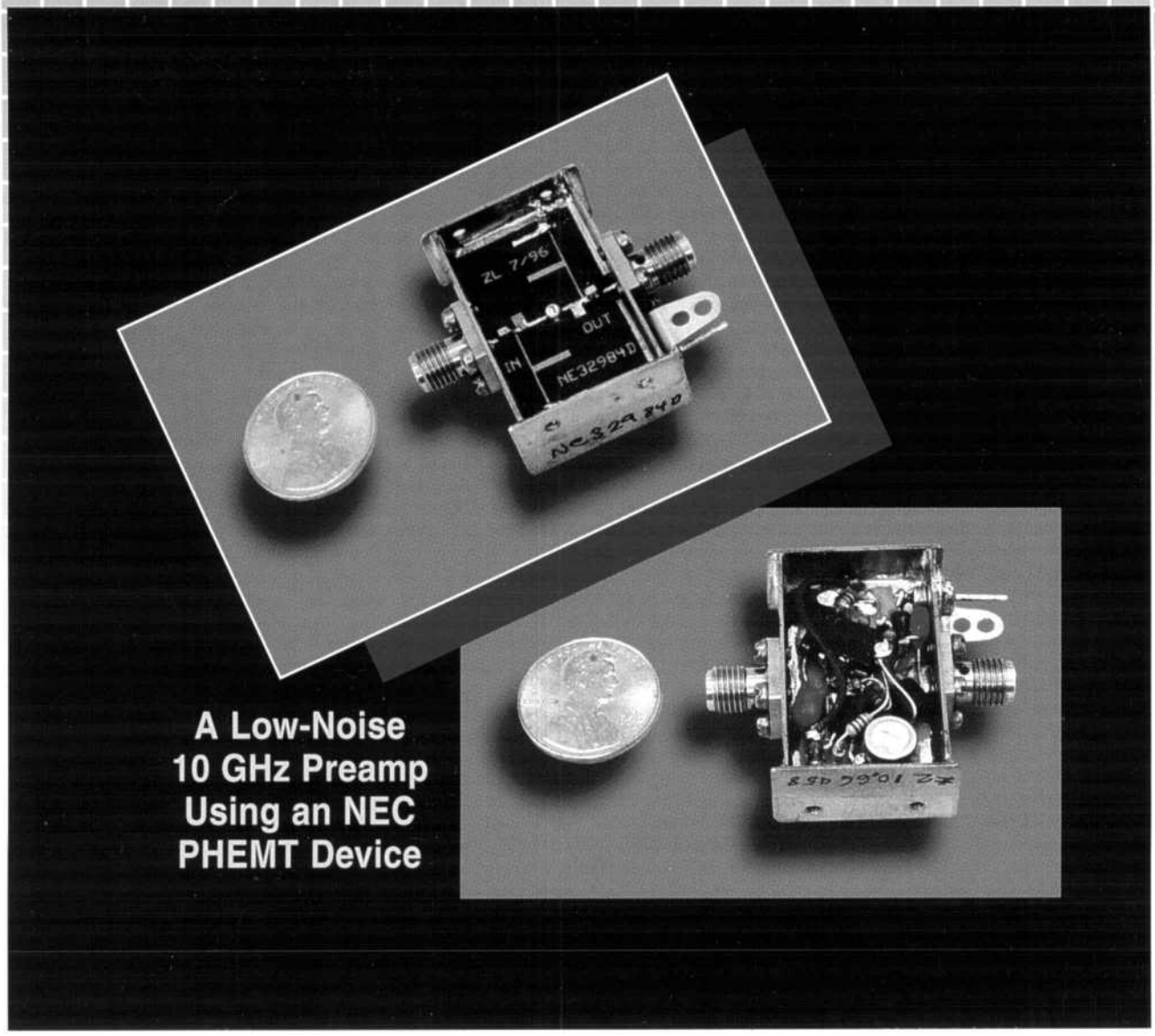
QEX

March/April 1998

Forum for Communications Experimenters

$4



A Low-Noise 10 GHz Preamp Using an NEC PHEMT Device

QEX: Forum for
Communications Experimenters
American Radio Relay League
225 Main Street
Newington, CT USA 06111-1494

QEX (ISSN: 0886-8093) is published bimonthly in January 98, March 98, May 98, July 98, September 98, and November 98 by the American Radio Relay League, 225 Main Street, Newington CT 06111-1494. Subscription rate for 6 issues to ARRL members is $18; nonmembers $30. Other rates are listed below. Periodicals postage paid at Hartford CT and at additional mailing offices.

POSTMASTER: Form 3579 requested.
Send address changes to: QEX, 225 Main St, Newington CT, 06111-1494

Issue No. 187

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)
Electronic Mail: MCIMAILID: 215-5052
Internet:qex@arrl.org

Subscription rate for 6 issues:

In the US: ARRL Member $18,
nonmember $30;

US, Canada and Mexico by First Class Mail:
ARRL Member $31, nonmember $43;

Elsewhere by Surface Mail (4-8 week delivery):
ARRL Member $23,
nonmember $35;

Elsewhere by Airmail: ARRL Member $51,
nonmember $63.

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

In order to insure prompt delivery, we ask that you periodically check the address information on your mailing label. If you find any inaccuracies, please contact the Circulation Department immediately. Thank you for your assistance.

About the Cover

In *RF*, Zack presents a 10 GHz preamp using a new, low-noise NEC PHEMT

Features

Columns

Mar/Apr 1998 QEX Advertising Index

THE AMERICAN RADIO RELAY LEAGUE

The American Radio Relay League, Inc, is a noncommercial association of radio amateurs, organized for the promotion of interests in Amateur Radio communication and experimentation, for the establishment of networks to provide communications in the event of disasters or other emergencies, for the advancement of radio art and of the public welfare, for the representation of the radio amateur in legislative matters, and for the maintenance of fraternalism and a high standard of conduct.

ARRL is an incorporated association without capital stock chartered under the laws of the state of Connecticut, and is an exempt organization under Section 501(c)(3) of the Internal Revenue Code of 1986. Its affairs are governed by a Board of Directors, whose voting members are elected every two years by the general membership. The officers are elected or appointed by the Directors. The League is noncommercial, and no one who could gain financially from the shaping of its affairs is eligible for membership on its Board.

"Of, by, and for the radio amateur, "ARRL numbers within its ranks the vast majority of active amateurs in the nation and has a proud history of achievement as the standard-bearer in amateur affairs.

A bona fide interest in Amateur Radio is the only essential qualification of membership; an Amateur Radio license is not a prerequisite, although full voting membership is granted only to licensed amateurs in the US.

Membership inquiries and general correspondence should be addressed to the administrative headquarters at 225 Main Street, Newington, CT 06111 USA.

Telephone: 860-594-0200
Telex: 650215-5052 MCI
MCIMAIL (electronic mail system) ID: 215-5052
FAX: 860-594-0259 (24-hour direct line)

Officers

President: RODNEY STAFFORD, W6ROD
5155 Shadow Estates, San Jose, CA 95135

Executive Vice President: DAVID SUMNER, K1ZZ

Purpose of *QEX*:

1) provide a medium for the exchange of ideas and information between Amateur Radio experimenters
2) document advanced technical work in the Amateur Radio field
3) support efforts to advance the state of the Amateur Radio art

All correspondence concerning *QEX* should be addressed to the American Radio Relay League, 225 Main Street, Newington, CT 06111 USA. Envelopes containing manuscripts and correspondence for publication in *QEX* should be marked: Editor, *QEX*.

Both theoretical and practical technical articles are welcomed. Manuscripts should be typed and doubled spaced. Please use the standard ARRL abbreviations found in recent editions of *The ARRL Handbook*. Photos should be glossy, black and white positive prints of good definition and contrast, and should be the same size or larger than the size that is to appear in *QEX*.

Any opinions expressed in *QEX* are those of the authors, not necessarily those of the editor or the League. While we attempt to ensure that all articles are technically valid, authors are expected to defend their own material. Products mentioned in the text are included for your information; no endorsement is implied. The information is believed to be correct, but readers are cautioned to verify availability of the product before sending money to the vendor.

Empirically Speaking

Our transition to bimonthly has gone well. I'm sure you've noticed the new cover graphics and some changes in the banner. Now we must keep the stuff inside as nice as the cover. In that regard, I am very pleased by how many of you have submitted articles. We are not overflowing but have a good content selection, as you can see in the last few issues. Every *QEX* issue gobbles up material. We publish about twice as much technical material as *QST*, on a monthly basis. We can't rest for a moment. Keep the stuff coming and remember that *QEX* is what *you* make it. Let's look at the workings of *QEX*: how we handle articles and what we expect from authors.

QEX Editorial Approach

As a purely technical forum, *QEX* is much less formal than *QST*. Authors exchange ideas here through their articles. They need not explain and support every detail. Although experts review submitted articles, our readers can fill in some gaps and work out unclear areas. The *QEX* editorial process is much quicker than that of *QST*, with much less author contact. *QST* articles get several days of editorial time, *QEX* articles get less than a day. We consider articles nearly camera ready, and contact the author only if there's a problem. Most manuscripts (about 60%) are complete, and require no communication except to work out the final review schedule.

How to Submit Articles

Send your articles to the *QEX* Editor at ARRL Headquarters. You can download an author's guide at **http://www.arrl.org/qst/aguide/.** If possible, send manuscripts and drawings in electronic form as well as on paper. You can e-mail them to **qex@arrl.org**. Hand drawn figures are okay, but better drawings make it easier for us to process the article. (If your drawings are correct and legible—even if hand drawn—we'll scan them for publication.) Once logged into the HQ database, articles are sent to me for review.

Given our great need for material, we return very few manuscripts these days, so don't be bashful. We need material at all skill levels. After the review, we either send an acceptance letter and Publication Release form or return the manuscript. We can't proceed without a signed release. Then I edit the article lightly for technical content, and send it to Bob Schetgen, who prepares it for layout. The author gets one final review before it goes to the printer.

We don't plan *QEX* far in advance, and content may change until it's sent to the printer. Therefore, we can't predict when your article will appear until well after it's edited.

This Issue in *QEX*

In this issue, we begin a three-part series by Doug Smith, KF6DX/7, about the workings of DSP receivers. Doug is the designer of the new Kachina receiver, and he gives the inside dope on this new technology. He begins with basic principles and goes on to concrete examples.

Ever wonder what really goes on with the waves travelling back and forth on your transmission lines? Walt Maxwell, W2DU, gives us some insight, in the context of impedance matching. This gives us a good feel for what's happening on transmission lines.

The voltage multiplier is an old standby circuit for making higher voltages at low power, but we tend to forget about it. Parker Cope, W2GOM, reminds us of this useful circuit family and puts them in a more modern context.

Sunspots are on the rise, and it's time to think about 6 meters again. Grant Bingeman, KM5KG, makes some interesting comparisons between $\lambda/4$ groundplane antennas and vertical dipoles. The result is a slick 6 meter vertical that can be scaled to other frequencies, too.

Talk about strange happenings, the ionosphere is full of surprises. Eric Nichols, KL7AJ, gives us some insight to the more mysterious happenings, all of which are very real.

Need a little more selectivity? Maestro Wes Hayward, W7ZOI, tells us how to go from double-tuned circuits to triple with very little pain.

Antenna modeling and the associated software are hot topics these days. L.B. Cebik, W4RNL, gives an overview of what's available and how to find out a lot more.

If you're into microwaves, you may have used waveguide filters. John Reed, W6IOJ, shows how to build custom waveguide filters, with an example for 2.4 GHz.

This month in RF, Zack Lau, W1VT, presents an inexpensive homebrew feed for 3456 MHz dishes. He's also got a 10 GHz preamp using the low-noise NEC NE32984D PHEMT.—*73, Rudy Severns, N6LF,* **rseverns@arrl.org**

Signals, Samples and Stuff: A DSP Tutorial (Part 1)

"DSP" is a buzzword of the '90s. Have you wondered what it's all about? This article begins an ongoing QEX *walk through the forest of DSP.*

By Doug Smith, KF6DX/7

What *is* all this DSP business, anyway? How does it really work? Certain crucial concepts are used in digital radio design. In this first article of a series, I'll describe these concepts in some detail. All of them are important to understand in the execution of a DSP transceiver. In the second article, we'll investigate an actual design. We'll look at *architectural* issues that impact decisions made during development, and we'll review the final configuration. In the third article, we'll survey some advanced DSP techniques such as adaptive filtering and special demodulation methods.

The Mathematics of Complex Signals

DSP implementations of radio transceiver functions compel designers to reexamine the mathematics that describe them. Computers and microprocessors are good at crunching numbers, but one thing stands out about them: they do exactly what they are told! Therefore, if we expect a DSP system to generate an SSB signal, we'd better know those calculations to perform, and those to avoid.

PO Box 4074
Sedona, AZ 86340
e-mail **dsmith@sedona.net**

Real and Complex Signals

Let's start with the job of taking a real input signal, say the audio from a microphone, and converting it to an SSB signal that can be transmitted over the air. We have to translate its frequency upward by the carrier frequency's value, and in so doing, preserve the spectral content. If we wish to produce an upper sideband (USB) signal, we want the carrier and lower sideband to be suppressed by as much as possible. We'll explore how the mathematics of complex signals achieve this using the so-called "phasing method" of SSB generation.

Of course, we generate SSB signals in other ways, such as the filter method. Because DSP makes it easy to build broadband phase shifters, and because the use of complex signals gives rise to a great deal of flexibility and precision, the phasing method has dominated DSP SSB generation to date.

Complex signals are not generally well understood, and

they're an obstacle in the way of those who wish to grasp the concepts. The idea of negative frequency is especially troublesome. A real signal, such as a cosine wave, is normally thought of as a positive frequency. It can be seen on an oscilloscope or spectrum analyzer in the positive-frequency domain. It can be transmitted and detected normally. We shall see, however, that such a signal actually consists of positive *and* negative frequencies when examined in the complex domain.

Our real cosine wave embodies the relation:

$$X_t = \cos \omega t \qquad \text{(Eq 1)}$$

where $\omega = 2\pi f$, and t is time. In the complex domain, the cosine wave is really the sum of two complex signals:

$$X_t = \frac{1}{2}\left[(\cos \omega t + j \sin \omega t) + (\cos \omega t - j \sin \omega t)\right] \qquad \text{(Eq 2)}$$

This signal has both positive- and negative-frequency components! The left-hand term is positive, the right-hand negative—the imaginary terms cancel and the real terms reinforce to make the equation true. In the complex plane, where the real part is one axis and the imaginary part the other, this signal can be represented as two vectors rotating in opposite directions. See Fig 1.

While this depiction is beautiful and elegant to a mathematician, what does it really mean to you and me? Well, it means that signals represented in complex form can have a one-sided spectrum, ie, having only positive or only negative frequencies. This is useful as we mix our signal upward to its final RF position.

Frequency Translation and Complex Mixing

We see that if we were able to translate the spectrum of our cosine wave—with its symmetrical positive and negative parts—upward in frequency far enough, we'd have two positive frequencies separated by twice the original signal frequency. For a real signal, this is exactly what happens when it's applied to an analog mixer! Both the sum and difference frequencies are generated, and the amplitude of each is half the original amplitude. So it's no coincidence that a mixer's conversion loss is about 6 dB—precisely what physics predicts.

We now invoke an identity discovered by Euler, which states that:

$$e^{j\omega t} = \cos \omega t + j \sin \omega t \qquad \text{(Eq 3)}$$

This is a more convenient notation for complex sinusoidal signals. So now, our real cosine wave takes the form:

$$\cos \omega t = \frac{1}{2}\left(e^{j\omega t} + e^{-j\omega t}\right) \qquad \text{(Eq 4)}$$

When we mix this with a real carrier, say

$$y_t = \cos \omega_0 t \qquad \text{(Eq 5)}$$

we get the product of the two inputs:

$$(\cos \omega_0 t)(\cos \omega t) = \left[\frac{\left(e^{j\omega_0 t} + e^{-j\omega_0 t}\right)}{2}\right]\left[\frac{\left(e^{j\omega t} + e^{-j\omega t}\right)}{2}\right] \qquad \text{(Eq 6)}$$

$$= \frac{\left[e^{j(\omega_0+\omega)t} + e^{-j(\omega_0+\omega)t}\right] + \left[e^{j(\omega_0-\omega)t} + e^{-j(\omega_0-\omega)t}\right]}{4} \qquad \text{(Eq 7)}$$

$$= \frac{1}{2}\left[\cos(\omega_0 + \omega)t + \cos(\omega_0 - \omega)t\right] \qquad \text{(Eq 8)}$$

The multiplication of the two cosine waves results in a frequency translation of their two-sided spectra. Now, let's consider what happens when we mix or multiply two complex, one-sided signals—as opposed to real, two-sided signals—together.

SSB Generation

We now present a new function, I_t, to represent the amplitude of the microphone audio versus time. This is a real signal with a two-sided spectrum. It's possible to convert this real signal to a complex signal—with only positive-frequency components—by generating a quadrature signal, Q_t, wherein all frequencies are phase-shifted by 90° from I_t, and treating I_t and Q_t as a complex, or *analytic*, pair. The signal $I_t + jQ_t$ contains only positive frequencies. The negative frequencies cancel each other, while the positive frequencies reinforce. The function that phase shifts all the frequency components by 90° is called a *Hilbert transform*. This would be difficult to achieve in the analog world, but is easy in DSP.

We now multiply this analytic signal by a complex oscillator:

$$Y_t = e^{j\omega_0 t} \qquad \text{(Eq 9)}$$

and the translated signal takes the form:

$$e^{j\omega_0 t}(I_t + jQ_t) = (\cos \omega_0 t + j \sin \omega_0 t)(I_t + jQ_t) \qquad \text{(Eq 10)}$$

$$= (I_t \cos \omega_0 t - Q_t \sin \omega_0 t) + j(I_t \sin \omega_0 t + Q_t \cos \omega_0 t) \qquad \text{(Eq 11)}$$

Since we're interested in transmitting this SSB signal via a single path over the air, we only have to compute the real part. This "half-complex" mixer is implemented as shown in Fig 2. It produces a real USB signal. This is, in fact, the good old phasing method.

Properties of SSB Signals

Since the amplitude of the carrier is constant, the amplitude of the SSB signal can be specified as some function of the modulating signal. If we think of the converted microphone audio signal $I_t + jQ_t$ as a vector, it follows that its length is equal to the instantaneous amplitude:

$$A_t = \left(I_t^2 + Q_t^2\right)^{\frac{1}{2}} \qquad \text{(Eq 12)}$$

The phase of the signal is the instantaneous angle of this rotating vector:

$$\phi_t = \tan^{-1}\left(\frac{Q_t}{I_t}\right) \qquad \text{(Eq 13)}$$

Now we can rewrite the real part of Eq 11 as:

$$\Re\left[e^{j\omega_0 t}(I_t + jQ_t)\right] = A_t \cos(\omega_0 t + \phi_t) \qquad \text{(Eq 14)}$$

This shows that an SSB signal is a hybrid of both amplitude and phase modulation. Also, note that having defined the amplitude and phase of the baseband signal in Eq 12 and 13 above, we can write:

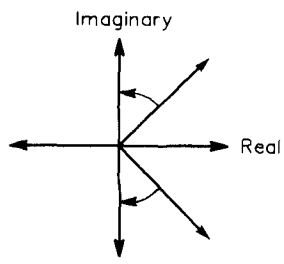


Fig 1—Vector representation of a real cosine wave.

$$(I_t + jQ_t) = A_t e^{j\phi_t} \quad \text{(Eq 15)}$$

directly relating the envelope and phase to the analytic baseband signal. The amplitude and phase of this analytic signal are identical to those of the SSB wave of Eq 11.

DSP Receiver/Exciter Architectures for SSB

With all this under our belts, we're ready to examine how these concepts apply to actual digital SSB receivers and exciters. While we won't explore all the ways DSP transceivers can be configured, and won't yet dive into all the minute detail, we will look at several designs to illustrate the principles.

Fig 3 is the block diagram of a digital exciter. The audio is normally low-pass filtered before sampling to remove components above half the sampling frequency of the analog-to-digital converter (ADC). A sampling frequency that is one quarter of the output sampling frequency is convenient. The audio passes through two band-pass filters, one of which incorporates a 90° phase shift. This converts the real signal to a complex signal $I_t + jQ_t$ having only positive frequencies. The filters are identical in frequency response and differ only in their phase responses.

The construction details of these filters will be treated later, under "Digital Filters." For now, it's sufficient to say that it's easy to build a frequency-independent DSP phase shifter—a fantasy in the analog world! The key concept here is that the real and imaginary parts of the analytic signal $I_t + jQ_t$ are handled separately in DSP.

The analytic signal is translated to the output frequency

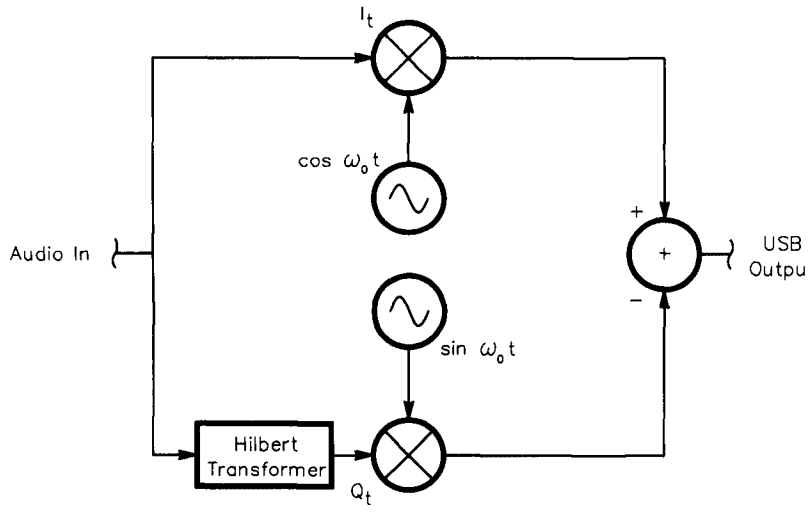


Fig 2—Block diagram of a half-complex mixer.

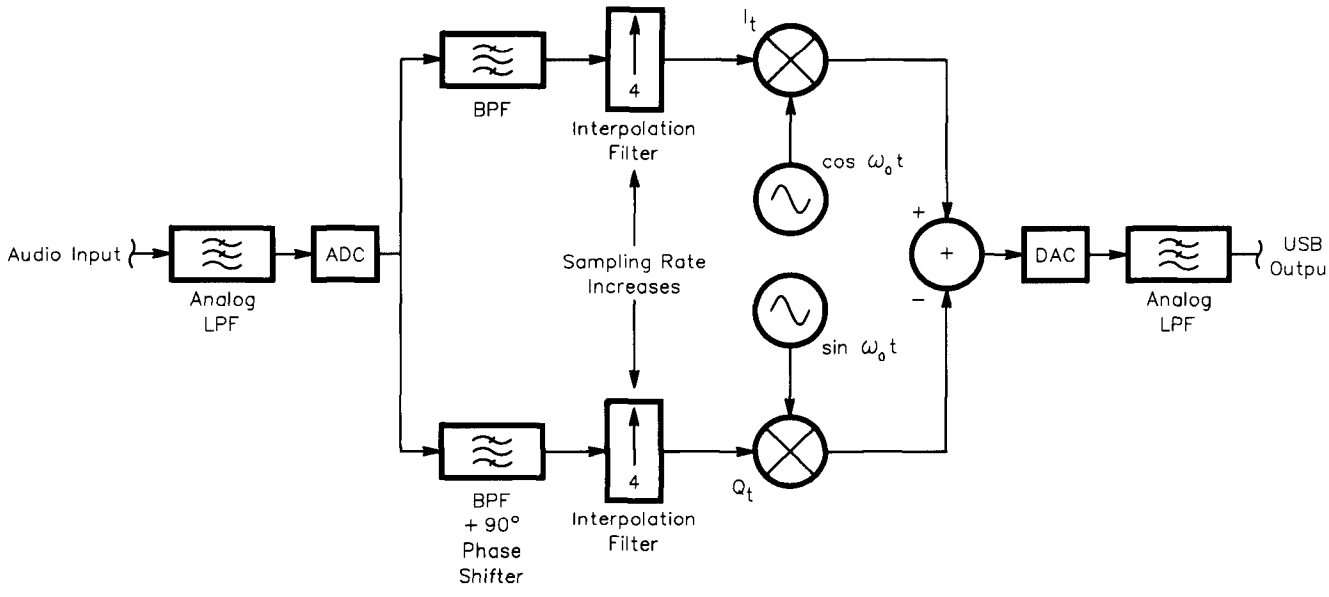


Fig 3—Block diagram of a digital SSB exciter.

through multiplication by a complex carrier:

$$e^{j\omega_0 t} = \cos \omega_0 t + j \sin \omega_0 t \qquad \text{(Eq 16)}$$

and the result is just as it was for Eq 11. Only the real part is computed, and that is at a sampling rate four times the input sample rate. For every sample from the filters, we use four samples of the complex oscillator.

This is beneficial, because for each full output cycle, the cosine oscillator produces values 1, 0, –1, and 0; the sine oscillator produces values 0, 1, 0, and –1. No actual multiplications take place, which saves time and accuracy. The sampling rate of the filter outputs must be artificially increased to make this procedure work—it's called *interpolation*.

The sampling process causes the output spectrum to repeat at harmonics of the sampling frequency. To remove these *aliases*, an analog *anti-alias* filter is required after the digital-to-analog converter (DAC). A digital interpolation filter eases the requirements of this anti-alias filter. It operates at the higher output sampling rate, taking the additional input samples to be zero. The filter shape is designed to attenuate the harmonic spectra in the original signal.

In the example, a USB signal is produced. Had the sum of the real and imaginary parts been taken instead of the difference, an LSB signal would have emerged.

An Independent Sideband (ISB) Exciter

We saw how easy it was to change sidebands in the above example by either subtracting or adding the real and imaginary parts. This makes it easy to create an ISB exciter that transmits separate information on each sideband.

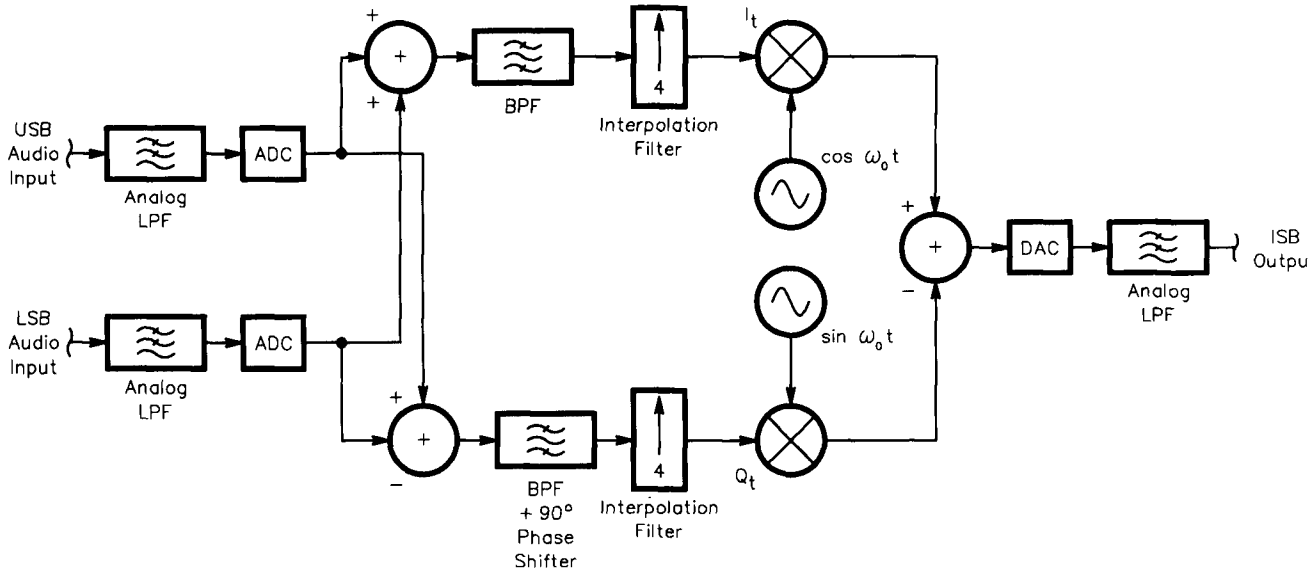


Fig 4—Block diagram of a digital ISB exciter.

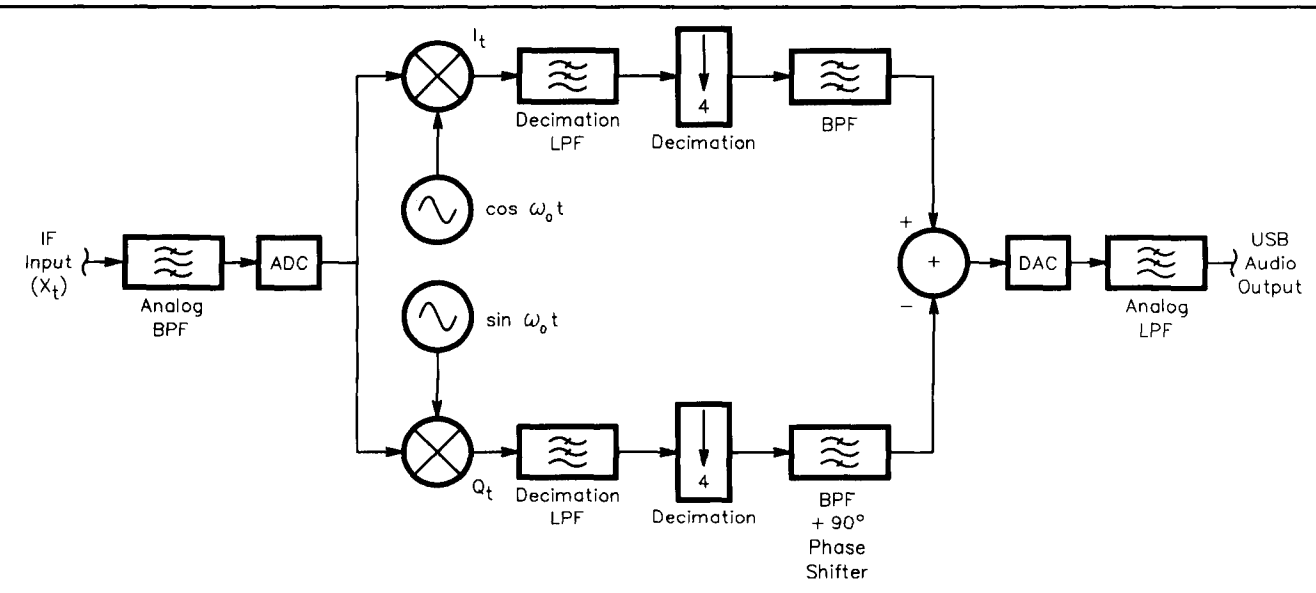


Fig 5—Block diagram of a digital SSB receiver.

The block diagram of such an exciter is shown in Fig 4. Working through the math for each sideband—and taking the input for the other sideband to be zero—shows that the system works as advertised.

While there are other ways to design sideband exciters in DSP, these examples illustrate the flexibility afforded by using the phasing method. The basic structures in the exciter architecture are directly applicable to a sideband receiver, too. Let's consider what happens on the receiving end of an SSB radio link, and see how the math works in reverse.

A Digital SSB Receiver

As in digital exciters, phasing methods prevail in receivers because it is easy to build frequency-independent phase shifters and because amplitude accuracy is preserved throughout. The first job is to convert the analog IF signal to digital form. Because of dynamic range considerations, the frequency at which the sampling is performed is limited to just above the audio range. Current technology dictates that a 16-bit ADC is used at a low-frequency IF.

As expertise in ADCs progresses, the digitization point will move closer to the antenna. For now, we must stick with traditional analog front ends, using very high-frequency first IFs to avoid image and spurious problems, and low-frequency second IFs to satisfy the ADCs. The sampled IF signal must be sharply band pass limited because of the restrictions imposed by current DSP processor performance.

Fig 5 presents a block diagram for a digital SSB receiver. After the IF signal is digitized, we wish to reduce the sampling rate—and the filtered bandwidth—as soon as possible. This is so because we need as much time as possible between input samples for the intense computations we must perform. Reduced sampling rates also ease the design of the digital filters that provide the final selectivity.

A technique known as *harmonic sampling* allows us to reduce the initial sampling rate to twice the input signal bandwidth. Let's say this bandwidth is about 15 kHz, enough for FM and other modulation formats, yet not so wide as to let in lots of QRM.

Therefore, the sampling rate must be at least 30 kHz, and we select something slightly higher to ease the analog IF filtering requirements. The digitized signal is then translated to baseband using the complex mixing algorithms outlined previously. Since the input signal, X_t is real, only two multiplications are necessary:

$$X_t e^{j\omega_0 t} = X_t \cos \omega_0 t + jX_t \sin \omega_0 t \qquad \text{(Eq 17)}$$

Now we have an analytic signal as before, and the value of the oscillator, ω_0, is chosen to beat the carrier frequency to zero. Again, the sampling rate is converted by choosing an IF that is four times the sample rate after translation, so that the oscillator values are only 1, 0, –1 or 0. This time though, we're *reducing* the sampling rate, and this is called *decimation*.

The sampling process again produces spectra at harmonics of the sampling frequency. To avoid mixing these into the passband at the reduced sampling rate, a decimation filter is required. This filter operates at the higher sampling rate, and limits the bandwidth to half the lower sampling rate.

Since the carrier frequency is at zero, the spectrum of our analytic signal contains both negative and positive components. The negative frequencies represent the lower sideband, and the positive frequencies the upper sideband. The I_t and Q_t signals are passed through two band-pass filters, one of which has a built-in 90° phase shift. These filters provide the final receiver selectivity, and are again identical in frequency response. The outputs of these filters are either subtracted or added to demodulate the USB or LSB audio. The digital audio signal is converted back to analog by the DAC, and the output is low-pass filtered to remove the sampling-frequency-caused harmonic spectra.

Obviously, we could both add and subtract the terms to produce an ISB receiver. Very little additional processing overhead would be involved, but we would need another DAC and low-pass filter.

We see that this receiver processes signals in a way that is the reverse of the digital exciter above. The mathematical relationships we described above define the transforms between real audio signals and real SSB signals, and we've seen how they work in both directions. Analytic signals can

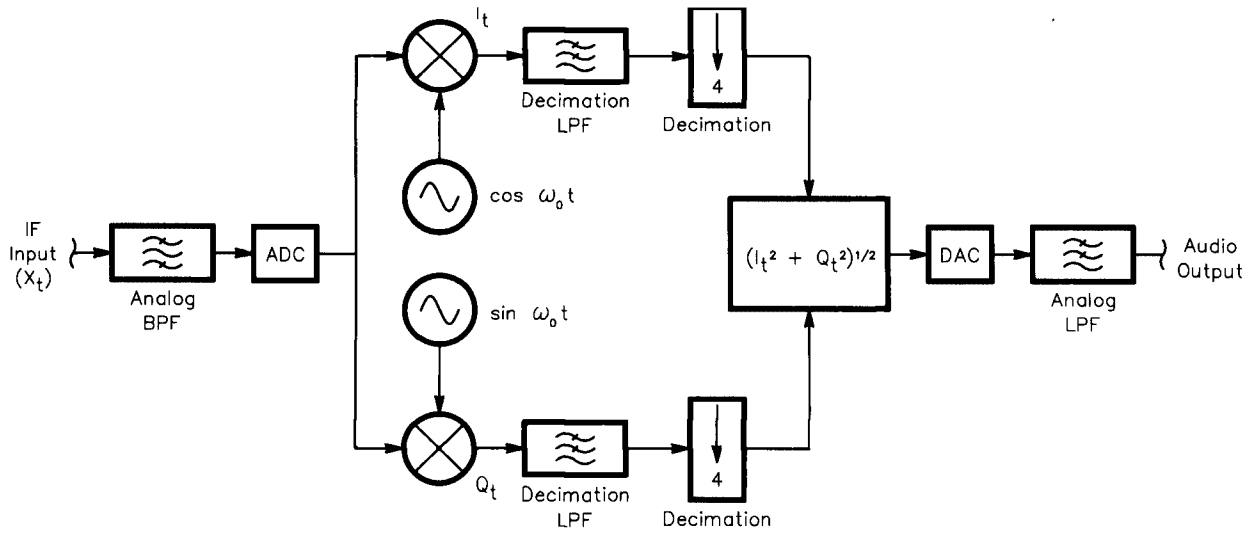


Fig 6—Block diagram of a digital AM receiver.

also be used to demodulate other types of signals, such as AM and FM. Let's look briefly at how these are handled in DSP.

AM Demodulation

In conventional AM, the envelope of the received signal is equal to the amplitude of the baseband audio. So to implement an AM detector, we can compute the magnitude of the analytic signal $I_t + jQ_t$ as in Eq 12. The resulting audio signal is free of the distortion encountered in detectors using rectification methods. Note that in the AM demodulator shown in Fig 6, the second pair of filters isn't necessary; and the decimation filters provide the final selectivity.

Now that we must use Eq12, we're stuck with computing the square root of a number. We could use the relation

$$X^{\frac{1}{2}} = \log^{-1}\left(\frac{\log X}{2}\right) \qquad \text{(Eq 18)}$$

but this involves computing a logarithm, or looking it up from a very large table. Fortunately, Sir Isaac Newton comes to the rescue with his root-finder algorithm!

In the 17th century these calculations were quite a burden, and anything that sped them up was a major blessing. Logarithms had only just been invented, and large tables of them were both expensive and scarce. Newton found a quick, simple iterative method for finding the roots of certain equations. For square roots, it goes like this: Take a crude guess at the square root of the number in question. Divide the number by the crude guess. Add the crude guess to this result, and divide it all by 2. Then, use this result as the new crude guess and repeat the process to obtain the desired accuracy.

$$let\ GUESS_{new} = \left(\frac{\frac{Number}{GUESS_{old}} + GUESS_{old}}{2}\right) \qquad \text{(Eq 19)}$$

$$let\ GUESS_{old} = GUESS_{new}$$

REPEAT

In practice, the accuracy of the result reaches the limit of 16-bit representation in five or six iterations. This method is much quicker and more accurate, for assembly-language implementations, than any alternative.

FM Demodulation

In FM, the instantaneous frequency of the carrier is equal to the amplitude of the baseband signal. We discovered how to compute the phase of our analytic signal above, so we can build a PM demodulator right away using Eq 13. The arctangents can be looked up from a table, or we can compute them using power series.

Once the phase at each sample time is found, the frequency is calculated as the rate of change of the phase:

$$f_t = \frac{d\phi_t}{dt} \qquad \text{(Eq 20)}$$

Differentiating the string of phase samples is accomplished using the technique of *first differencing*. We simply take the difference between the adjacent samples:

$$f_n = \phi_n - \phi_{n-1} \qquad \text{(Eq 21)}$$

This is the FM demodulator output.

Alternatively, FM can be directly calculated by evaluating the vector relationships and using:

$$f_t = \frac{Q_t\left(\frac{dI_t}{dt}\right) - I_t\left(\frac{dQ_t}{dt}\right)}{I^2 + Q^2} \qquad \text{(Eq 22)}$$

then low-pass filtering.

Finally, a digital discriminator can be used to translate the FM to AM, which is then demodulated as described above.

While the mathematics of the various modulation formats can be complicated, they dictate exactly what the computational block diagrams must look like. In a DSP implementation, the equations hold true and the results are precisely as predicted by theory. Next, we'll develop the theories of sampling and data-rate conversion to see what limitations they impose on us, and what advantages apply to digital radio design.

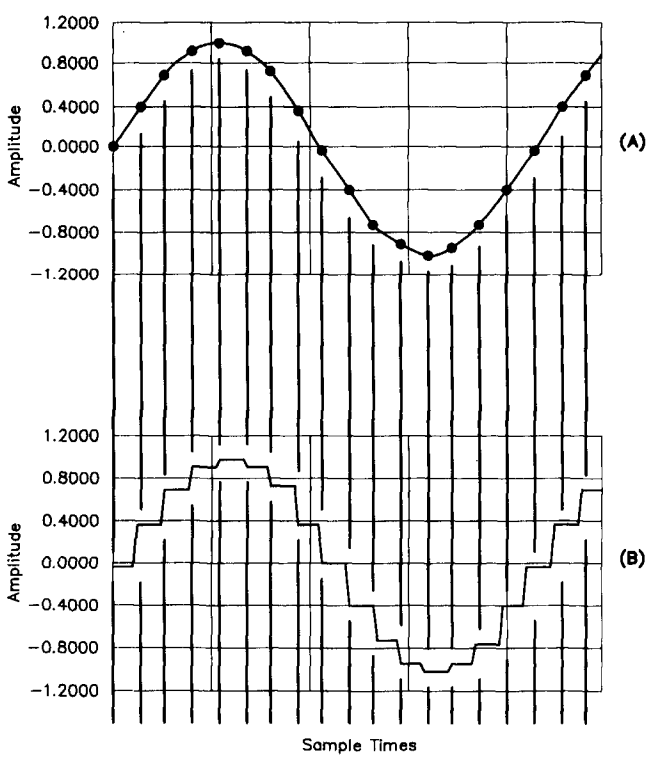


Fig 7a—Sine wave of frequency much less than the sampling frequency.
Fig 7b—Sampled sine wave.

Sampling Theory and Multirate Processing

In the field of radio communication, we deal with analog signals. In order to perform our DSP magic on them, therefore, the first and last steps in the processing chain involve conversion to and from digital form. Once a signal is digitized, we fight to preserve its integrity through the numerical precision of our calculations. At the interfaces to the analog world, the enemies are manifold: noise, distortion and lack of dynamic range and resolution. We'll explore the characteristics of sampled signals, and the reasons for degradation in conversion techniques.

Sampling Theory: Digitizing the Analog World

Sampling is the periodic measurement of the signal voltage and the conversion of this voltage to a number. If we take f_s measurements per second, we call f_s the sampling

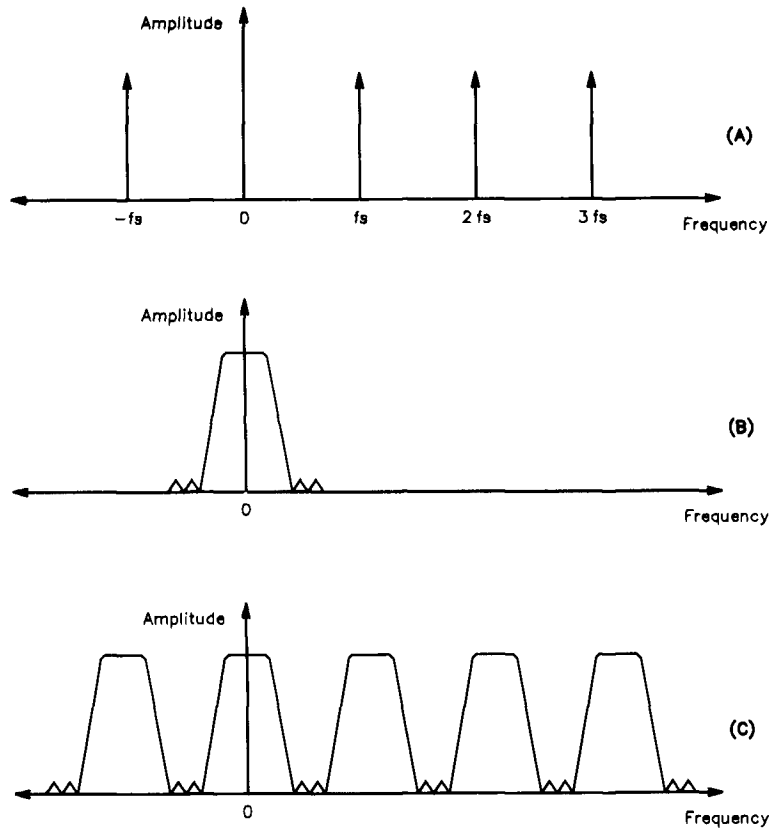


Fig 8a—Spectrum of sampling impulses.
Fig 8b—Spectrum of a band of real input signals.
Fig 8c—Spectrum of the sampled band of signals.

frequency. We can represent the sampled signal as a string of amplitude measurements over time, as shown for the case of a sine wave in Fig 7.

To completely understand the nature of the sampled signal, first consider a function δ_t, consisting of unit impulses spaced at intervals of the sampling time, f_s^{-1}. This is the time-domain representation of the sampling function. The sampled signal can be thought of as the sampling function multiplied by the continuous input signal, x_t:

$$y_t = x_t \delta_t \qquad \text{(Eq 23)}$$

The result is the *convolution* of the two signals—equivalent to their mixing in the analog world. A look at the spectrum of this result shows that it repeats at intervals of the sampling frequency, as shown in Fig 8 for a band of signals.

These repetitions are called *aliases*, and are as real as the fundamental in the sampled signal. They each contain all the information sufficient to describe the original signal. In our example, the sampling frequency is much higher than the signal frequency, so the output still roughly resembles the original sine wave.

Sine Wave, Alias Sine Wave: Harmonic Sampling

Imagine that the input frequency is greater than the sampling frequency. (See Fig 9.) Now, the output no longer matches the input. Note that the sampled signal is still in the shape of a sine wave, though, and that its frequency is lower than that of the input. Ordinarily, this wouldn't be a happy situation.

Nevertheless, a downward frequency translation is useful in the design of an IF-DSP receiver. In addition, lower sam-

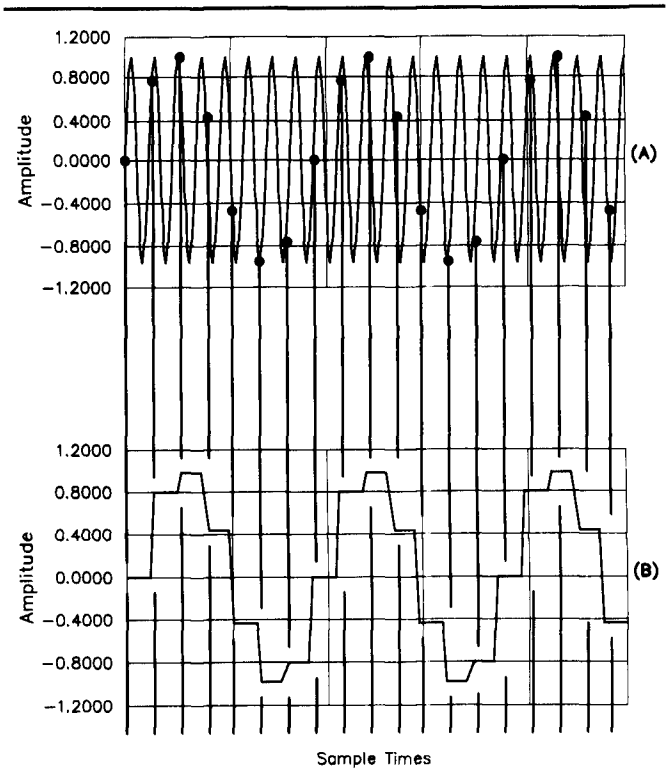


Fig 9a—Sine wave of frequency greater than sampling frequency.
Fig 9b—Harmonically sampled sine wave.

pling frequencies are desirable because they provide more time between samples for signal modification algorithms. Caution is required, though. An input signal near twice the frequency of that in Fig 9 would produce the same output. To use this technique, therefore, we must first band pass limit the input. This trick is known as *harmonic sampling.*

The input signals must fall between the fundamental, or some harmonic, of the sampling frequency and the point halfway to the next harmonic. A frequency translation will take place, but no information about the shape of the input signal will be lost because of this sampling technique. A frequency-domain representation of harmonic sampling is illustrated in Fig 10.

Information *can* be lost, however, because of inaccuracies and noise introduced by the ADC. After digital processing, additional distortion can be introduced by the DAC as we convert back to analog. Let's look at these errors.

Noise and Distortion in Signal Conversion

An ADC is a device that measures an analog signal voltage and outputs a proportional number. A DAC is another device that performs the reverse operation, outputting some analog voltage proportional to the numerical input. Both of these devices inject noise and distortion, at levels that are predicted by sampling theory. Their performance is so critical to any DSP transceiver that it's worth our effort to examine causes of degradation and learn how to combat them.

Aside from stray noise picked up in the physical circuits, the deleterious effects occurring in signal converters can be grouped as follows, and discussed:

- Nonlinearities in quantization step sizes
- Quantization noise
- Aperture jitter
- Noise figure and distortion in analog stages
- Zeroth-order sample-and-hold distortion

Nonlinearities

Nonlinearity means distortion, and—in this case—noise. The quantization steps of any real converter are not perfectly spaced, and conversion results are contaminated by the inaccuracy. In general, two types of nonlinearities can be characterized: differential nonlinearity (DNL), and integral nonlinearity (INL).

DNL is the measure of the output nonuniformity from one input step to the next. It is expressed as the maximum error in the output between adjacent input steps, measured over the entire input range of the device. In an 8-bit converter, for example, the worst errors typically occur when the output changes from 01111111 to 10000000, using base-2 notation.

Since we're talking about the accuracy of the smallest steps the converter can resolve, noisy low-order distortion products caused by this effect limit the dynamic range of the device. Current technology uses correction systems to compensate for temperature variations that would otherwise degrade performance beyond acceptable bounds.

A converter is considered *monotonic* if a steady increase

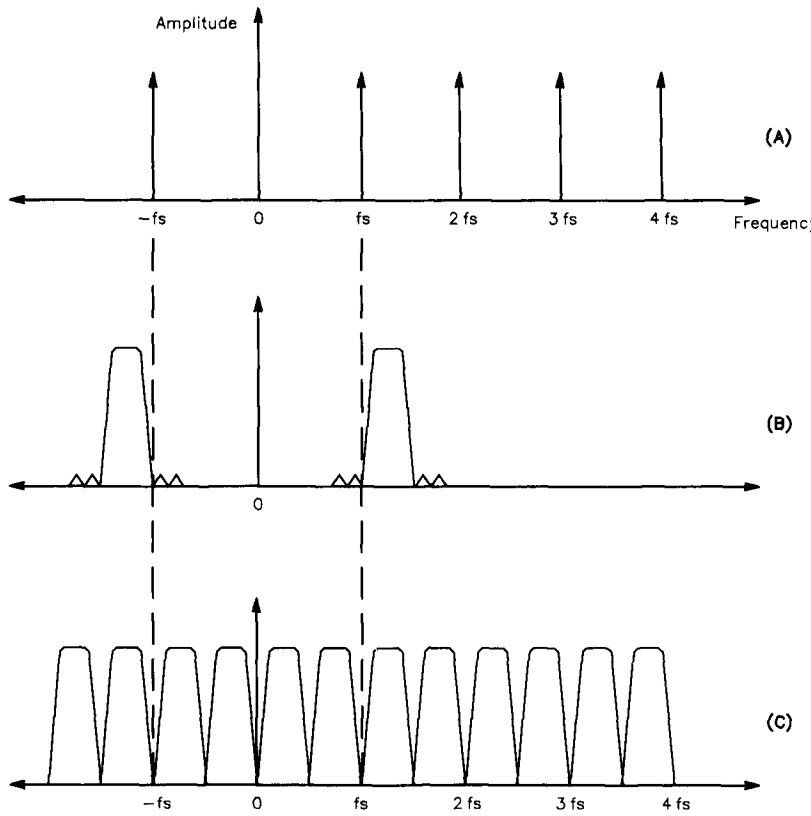


Fig 10a—Spectrum of sampling impulses.
Fig 10b—Spectrum of a band of real signals.
Fig 10c—Spectrum of harmonically sampled band of signals.

in the input signal always results in an increase in the output. Device manufacturers attempt to hold DNL to ±0.5 bits so that monotonicity is maintained.

The second measure of nonlinearity, INL, is a measure of the device's large-signal-handling ability. If we inject a signal of amplitude A and measure the output, then inject a signal of amplitude 100 A, we expect the output to increase in exact proportion. The INL is a measure of the maximum error in the output between any two input levels. Another way of depicting this measurement is to plot the input against the output, and see how straight the line is.

INL produces harmonic distortion and IMD that are obviously undesirable. Typical values are from ±1 to ±2 bits over the entire range.

Quantization Noise

Quantization noise, caused by the inability to resolve signals near the amplitude of the smallest quantization step, is basic to all converters and spread uniformly over the entire input bandwidth of $f_s/2$. The noise power is:

$$P_{qn} = \frac{V_{peak}^2}{3R2^{2b}} \qquad \text{(Eq 24)}$$

in watts, and so the noise density is:

$$ND_{qn} = \frac{P_{qn}}{\left(\frac{f_s}{2}\right)} = \frac{2V_{peak}^2}{3\,f_s R 2^{2b}} \qquad \text{(Eq 25)}$$

in watts/Hertz, where V_{peak} is half the maximum peak-to-peak input signal, R is the ADC input resistance, and b is the number of bits of resolution. Note that the quantization noise density decreases by 6 dB for every bit of resolution in the converter and by 3 dB every time we double the sampling frequency.

Since the maximum sine wave the converter can handle produces a power of:

$$P_{sine} = \frac{\left(\frac{V_{peak}}{\sqrt{2}}\right)^2}{R} = \frac{V_{peak}^2}{2R} \qquad \text{(Eq 26)}$$

the maximum signal-to-noise ratio (SNR) is:

$$\frac{P_{sine}}{P_{qn}} = \frac{3\left(2^{2b}\right)}{2} = (6.02b + 1.76)dB \qquad \text{(Eq 27)}$$

For a 16-bit converter, this is about 98 dB. If we could increase the sampling rate by some factor N, then digitally filter the output back down to the lower rate, we could

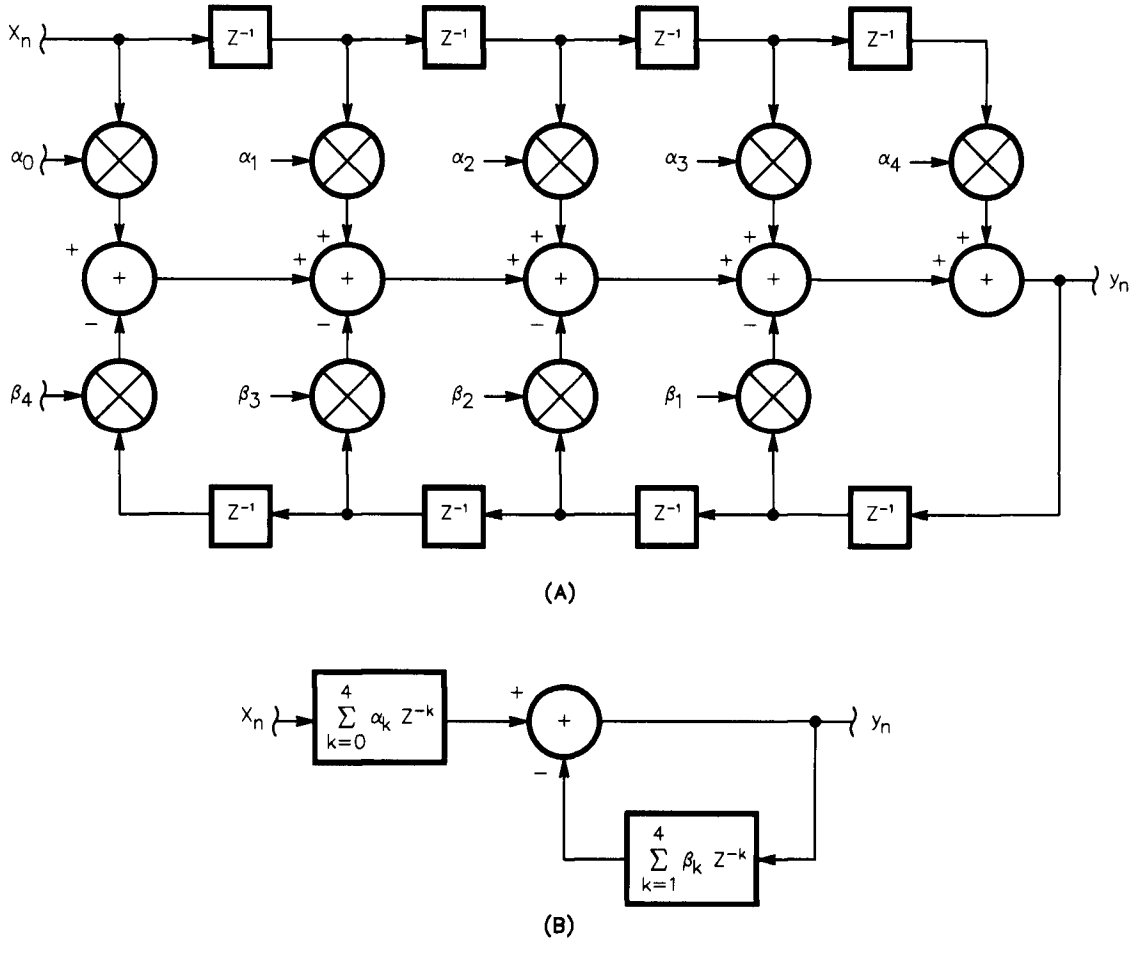


Fig 11a—Block diagram of an IIR filter for L = 5.
Fig 11b—Equivalent block

increase the SNR by almost the factor N. The quantization noise would be spread over a much larger bandwidth, and we'd use a decimation filter to eliminate the higher-frequency noise in bringing the sampling rate back down. This technique is known as *over-sampling*.

So-called "sigma-delta" ADCs use this method to achieve the best possible dynamic range and noise performance. They represent the state of the art in ADC technology.

Aperture Jitter

In addition to the above sources of noise and distortion, noise is introduced by slight variations in the exact times of sampling. Phase noise in the clock source used to derive the sample times, as well as other inaccuracies in the sampling mechanisms, produce undesired phase modulation of the sampled signal. If we assume this phase noise is not correlated with the input signal, a sine wave input generates a white noise component in the output having a power density relative to the input power:

$$\frac{N_{aj}}{P_{sine}} \approx \frac{8\pi^2 f^2 \sigma_a^2}{f_s} \qquad \text{(Eq 28)}$$

where f is the signal frequency, and σ_a is the RMS time jitter in the sampling rate. Note that the result is expressed in dB/Hz, and is proportional to the squares of both the signal frequency (f) and the time jitter (σ_a).

Noise Figure Issues

Quite often, the smallest signal the converter can discern is so small that noise generated in the analog stages of the converter becomes a problem. Consider a converter with 16 bits of resolution and a peak-to-peak input limit of 5 V. Say its input impedance is 10 kΩ. The noise power contributed by quantization is:

$$P_{qn} = \frac{2.5^2}{(3)(10{,}000)2^{(2)(16)}} \approx 48.5 \times 10^{-15}\,W \approx -103\,dBm \qquad \text{(Eq 29)}$$

It's likely, if the ADC comparators have a wide input bandwidth and a noise figure of more than 6 dB, that their noise contribution exceeds this amount. It's imperative, therefore, to evaluate converters based on all their properties before deciding to incorporate them in a DSP transceiver.

Zeroth-Order Sample-and-Hold Distortion

Typical converters are sample-and-hold devices. That is, they continue to output the last sampled value throughout the sample period. This effect acts as a filter having a frequency response:

$$H_\omega = \frac{\sin\left(\frac{\pi\omega}{\omega_s}\right)}{\left(\frac{\pi\omega}{\omega_s}\right)} \qquad \text{(Eq 30)}$$

This results in a high-frequency rolloff that is quite undesirable in many circumstances. For example, if the output frequency is one quarter of the sample frequency, an attenuation of about 1 dB will occur. This is mainly a problem in the final DAC stage, where signals are converted back to analog. We see that increasing the output sampling frequency reduces the attenuation, and that may call for interpolation of the output.

I've referred to changes in sampling rate called decimation and interpolation and alluded to the advantages in performing these rate changes. Now, let's look at them in more detail. This will lead us naturally into a discussion of digital filtering.

Sampling Rate Reduction: Decimation

Sampling at higher rates can be quite beneficial because it eases the design of the analog filters we must use to avoid the aliasing phenomenon. It also helps reduce noise introduced by quantization and aperture jitter.

At some stage, however, we want to reduce the sampling rate in order to provide as much time as possible in between samples for other calculations. When it's time to digitally filter some signals, making the filter bandwidth a large fraction of the sampling frequency makes it easier to build a sharp-skirted filter. Reduction of the sampling rate is usually referred to as decimation.

Decimation is normally done by integer values—although it doesn't have to be—and is the same as resampling the signal at the lower rate. The resampled signal has a spectrum repeating at intervals of the lower sampling frequency, and so we have to reduce the bandwidth to less than half this value to avoid the aliasing that would destroy information.

The decimation filter, operating at the higher sampling rate f_{in}, eliminates components above $f_{out}/2$ so that aliasing won't occur after the rate reduction. The output signal can be processed at the f_{out} rate, which requires less time. We see, though, that in sampling our signal at the lower rate, we're going to have to either average or discard some of the input samples.

When we filter the signal at the higher rate, then resample at the lower, it turns out it's legitimate to just discard the unneeded samples during decimation. No information about the input signal will be lost since aliasing is avoided. So, why compute these output samples when we're only going to throw them away? We'll calculate only the ones we're going to keep, and this is equivalent to running the decimation filter at the lower sampling rate. This *boxcar* technique is typical of those used by DSP designers to save time and effort.

Increasing the Sampling Rate: Interpolation

While a low sampling rate is pleasant for the reasons outlined above, it may cause difficulties when it's time to convert signals back to analog. The alias products aren't far above the desired signals, and are therefore difficult to filter out. The solution is to increase the sampling rate, and hence the frequencies of the alias products, using the process of interpolation.

We usually do this by an integer factor, although again, we don't have to. While there may be advantages to changing the sampling rate by other rational factors, such as 3/2, we're after an arrangement that solves the aliasing problem with a minimum of processing. In most cases, a doubling or tripling of the rate is sufficient to allow a reasonable anti-aliasing filter to be constructed.

To double the sampling rate, we'll insert additional samples with a value of zero between the existing samples. An interpolation filter is required, using the zero-inserted data as its input. It is a low-pass, operating at the higher sampling rate, which removes the alias components due to the lower sampling rate.

We've seen some properties of sampled signals and learned how the sampling theorem can be used to our advantage in designing a DSP transceiver. We've also seen the potential trade-offs that selection and conversion of sampling rates present. Next, on our path toward acquaintance with the important DSP methods, a thorough understanding of digital filtering is in order.

Digital Filters

The ability to construct high-performance filters is probably the most important reason for using DSP in radio transceivers. An expensive crystal or mechanical filter with a single bandwidth can be replaced by a set of superior digital filters, which offer as many bandwidths as the associated memory can support.

We can build digital filters that have linear phase response, which is very difficult in the analog world. This can be important for certain data modulation modes. Once the filter is designed, each unit is identical to the next—no alignment is necessary! Finally, filter shapes and responses that are impractical in the analog world can be easily implemented in DSP because there are no production variations.

Infinite-Impulse-Response (IIR) Filters

IIR filters are notable for the presence of feedback, which finite impulse response (FIR) filters do not have. For this reason, IIR filters are usually designed by converting traditional analog filter responses, such as Chebyshev and elliptical. IIR filters can have much sharper transition regions than FIR filters (for the same number of multiplications), but they bring with them the nonlinear phase responses of their analog brethren. They are much more susceptible to overflow problems, and are not necessarily unconditionally stable. They are also prone to *limit cycles*, low-level oscillations sustained by inaccuracies in numerical representation.

Designers can attempt to compensate these unwanted traits, but they may find that the resulting computational load isn't worth it. Nevertheless, we shall describe the synthesis and use of IIRs, as they have their places in modern radio development.

The transfer function of an analog Chebyshev low-pass filter can be written as the ratio of a constant to an *n*th-order polynomial:

$$H_s = \frac{K}{s^n + a_1 s^{n-1} + a_1 s^{n-2} + \ldots + a_n} \qquad \text{(Eq 31)}$$

Tables in the literature, such as Zverev, list the values of the coefficients, a, which are related to the cutoff frequency and used to derive actual component values for the filter. The low-pass design can be transformed to band-pass or bandstop response. Two popular methods exist for deriving the digital transfer function from the analog transfer function. These are known as the *impulse-invariant* method and the *bilinear transform* method.

The impulse-invariant method assures that the digital filter will have an impulse response equivalent to its analog counterpart, and therefore the same phase response. Problems arise, though, if the bands of interest are near half the sampling frequency. The digital filter's response can develop serious errors in this case. Because of this, the impulse-invariant method isn't as good as the bilinear transform method.

The bilinear transform method makes a convenient substitution for *s* in Eq 31 above, and the filter output comes out looking like:

$$y_n = \sum_{k=0}^{L-1} \alpha_k x_{n-k} - \sum_{k=1}^{L-1} \beta_k y_{n-k} \qquad \text{(Eq 32)}$$

This filter has L zeros and L – 1 poles.

The block diagram of such a filter for L = 5 is shown in Fig 11. Each box marked "Z^{-1}" is a one-sample-time delay. Feedback is evident in the diagram: The paths involving coefficients labeled ß loop back and are added to the signal path.

This *direct form* equation can be factored into 2-pole sections, and implemented in cascaded form. For each section:

$$y'_n = \left(\sum_{k=0}^{2} \alpha'_k x'_{n-k} - \sum_{k=1}^{2} \beta'_k y'_{n-k} \right) \qquad \text{(Eq 33)}$$

The output of each section serves as the input to the next. This configuration requires a few more multiplications than the direct form, but is less prone to instability and limit-cycle problems when proper pole-zero pairing is used. Additional information about IIR filters can be found in the literature.

Finite-Impulse-Response (FIR) Filters

This digital filter is by far the most popular for SSB use because of its linear phase response. The transfer function of an FIR filter has only zeros, so to implement one, we eliminate the poles and compute only the left half of Eq 32. The output then takes the form:

$$y_n = \sum_{k=0}^{L-1} h_k x_{n-k} \qquad \text{(Eq 34)}$$

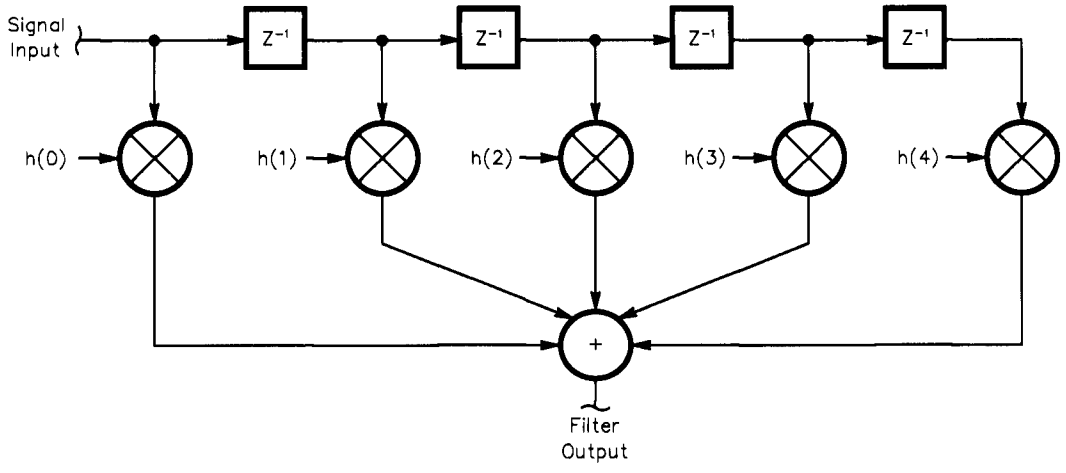


Fig 12—Block diagram of an FIR filter for L = 5.

where h is the set of L coefficients. The coefficients represent the impulse response of the filter, and the filter is said to have *length* L.

The block diagram of such a filter for L = 5 is shown as Fig 12. The set of input registers holding delayed samples of x is just a tapped digital delay line, and for this reason, this filter can be said to have 5 *taps*. Since the output depends only on past input values, the filter can be said to be a *causal process*.

Normally, the impulse response has a symmetry about center, and it turns out this is enough to ensure that no differential or group delay distortion is produced. Since there is no feedback, the filter is unconditionally stable. Almost any frequency response can be generated if enough taps can be used. Many excellent CAD programs are available to design digital filters, relieving us of the burden of generating coefficient sets.

Numerical Accuracy Effects in FIR Filters

In general, adding more taps sharpens the transition regions in the frequency response. In addition, accurate frequency and phase responses require a large number of taps. Because it is truncated at both ends, the finite coefficient set is only an approximation of that required for the exact response we want. Some error in the result must be tolerated.

After Rabiner and Gold, the number of taps required can be estimated using:

$$L = 1 - \frac{10\log(\delta_1\delta_2) - 15}{14\left(\frac{f_T}{f_s}\right)} \qquad \text{(Eq 35)}$$

where δ_1 is the allowable passband ripple, δ_2 is the stopband attenuation, f_T is the transition bandwidth, and f_s the sampling frequency. This assumes, of course, enough bits of resolution are used to achieve the required accuracy. In actual practice, filters of over 100 taps are used to realize shape factors of less than 1.14.

When computers are used to design FIR filters, coefficients can be represented to the full accuracy of the program—usually in floating-point format with 12 or more decimal significant figures in the mantissa. Real implementations ordinarily don't achieve this accuracy, as we're typically limited to 16 bits in an embedded DSP design. The truncation of coefficients and data degrades the response and, of course, sets the dynamic range.

We know that the product of two 16-bit numbers is a 32-bit number, so we need at least that many bits in the final accumulator to avoid losing accuracy. Another point of interest: While the delay through the filter isn't dependent on frequency, the absolute value of that delay can be quite large. In fact, it is equal to:

$$T_{FIR} = \frac{L \bullet t_s}{2} \qquad \text{(Eq 36)}$$

When we get around to using these filters in our design, we'll want to consider the effects of these delays.

Fixed-Point Mathematics and Scaling Problems

The typical DSP microprocessor is a 16-bit machine using fixed-point math. This means numbers are represented internally as signed fractions between –1 and 1 in *2s complement* format. The most significant bit (MSB) represents the sign of the number, and the 15 least significant bits (LSBs) the magnitude—or the complement of the magnitude for a negative number. In hexadecimal format, \$7fff is the largest positive number and represents $1-2^{-15}$, while \$8000 is the most-negative number, and represents (–1). This is a convenient format, since the product of two fractions is always another fraction! When we start adding the fractions together, however, (as in our FIR filter above) we may generate a number whose magnitude is greater than unity—a result known as *overflow*.

Most DSPs using 16-bit data and coefficients have final accumulators with at least 32 bits. A trade-off exists between the possibility of overflow—which is catastrophic—and the loss of accuracy in the LSBs, via truncation during operations. Especially in FIR filters with sharp transition regions, and under certain input conditions, the output can exceed ±1. The worst-case output can grow as large as the sum of the absolute value of all the coefficients:

$$y_{\max} = \sum_{k=0}^{L-1} |h_k| \qquad \text{(Eq 37)}$$

We must scale either the data or the coefficients by the reciprocal of this factor to ensure against overflow.

At the small-signal end of things, the bit-resolution of the system determines the dynamic range because of the presence of quantization noise, just as in the case of ADCs or DACs. It is computed almost the same way. Its normalized amplitude is:

$$V_{qn} = \frac{2^{-(b+1)}}{\sqrt{3}} \qquad \text{(Eq 38)}$$

where b is the number of bits used, internally, to represent numbers.

Truncation of numerical results is another form of quantization noise, also computed in the same way. If each of the L multiplications is truncated, then the noise amplitude from this source is:

$$V_{trunc} = \frac{2^{-(b+1)}L}{\sqrt{3}} \qquad \text{(Eq 39)}$$

It's interesting to note that while truncation of the coefficients affects the response characteristics of the filter, it doesn't contribute to the noise in the output.

Hilbert Transforms

We know that to build a DSP radio, it's convenient to use phasing methods as discussed previously. We need a way to shift the phase of signals by 90°. For a single frequency, this is easy: We just insert a delay of one-quarter cycle. Over a range of frequencies, though, an FIR structure is required to obtain a frequency-independent phase shifter. We'll call this a Hilbert transformer.

Coefficients for a Hilbert transformer can be generated

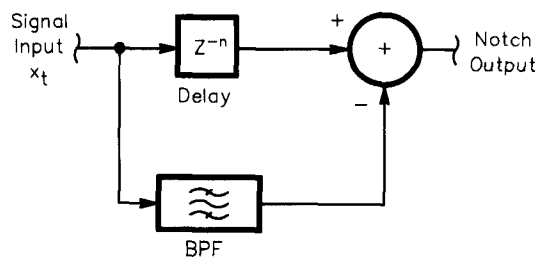


Fig 13—Block diagram of digital notch filter.

by CAD programs, and optimized for the bandwidth of interest. As noted above, the coefficient set will be symmetrical about center. The phase response is, of course, a straight line. A nice property of Hilbert transform impulse responses is that if the total number of taps is odd, the even-numbered coefficients are zero. This cuts our number of calculations in half.

Also, to minimize the computational load, let's see if we can somehow avoid having a separate Hilbert transformer in our phasing-method implementation. We'll attempt to build a pair of band-pass filters, with the frequency response we need and whose phase responses are 90° different from each other. These can then be used directly in our receiver/exciter to create and operate on analytic signals.

Analytic Filter-Pair Synthesis

The frequency translation theorems we explored above can be used to advantage in creating our pair of filters. If we start with a low-pass filter having impulse response, h_t, and frequency response, H_ω, multiplying the impulse response by a complex sinusoid

$$e^{j\omega_0 t} = \cos\omega_0 t + j\sin\omega_0 t \qquad \text{(Eq 40)}$$

results in two sets of coefficients, one for the real part, and one for the imaginary part:

$$h_{I_t} = h_t \cos\omega_0 t$$
$$h_{Q_t} = h_t \sin\omega_0 t \qquad \text{(Eq 41)}$$

The frequency response of either one of these filters is given by:

$$H_\omega = \frac{H_{(\omega-\omega_0)} + H_{(\omega+\omega_0)}}{2} \qquad \text{(Eq 42)}$$

which is a band-pass filter centered at ω_0. The first filter in Eq 41 has a phase response 90° different from the second. The frequency translation theorem works on the responses of filters just as well as it does on real signals!

To perform this transformation on the L coefficients of the initial low-pass filter, we calculate new coefficients:

$$\text{For } 0 \le k \le L-1,$$
$$h_{I_k} = h_k \cos\omega_0\left(k - \frac{L}{2} + \frac{1}{2}\right)t_s \text{ OR}$$
$$h_{Q_k} = h_k \sin\omega_0\left(k - \frac{L}{2} + \frac{1}{2}\right)t_s \qquad \text{(Eq 43)}$$

We can implement an IF shift in our receiver design simply by altering the value of ω_0, and computing the new coefficients. We can alter the transmitter's frequency response by convolving the impulse response of our analytic filter pair with that of a filter having the desired characteristic. It's evident that FIR filters yield flexibility beyond that of any analog technique.

Digital Notch Filters

The other type of filter of interest to us is the notch, designed to remove a single frequency. Note that such a filter can be constructed by subtracting the output of a narrow band-pass filter from the broadband input, as shown in Fig 13. We include a delay of

$$Z^{-n} = \frac{L \bullet t_s}{2} \qquad \text{(Eq 44)}$$

in the broadband input to compensate for the delay through the band-pass filter, whose length is L.

An unusual type of notch filter has been described by

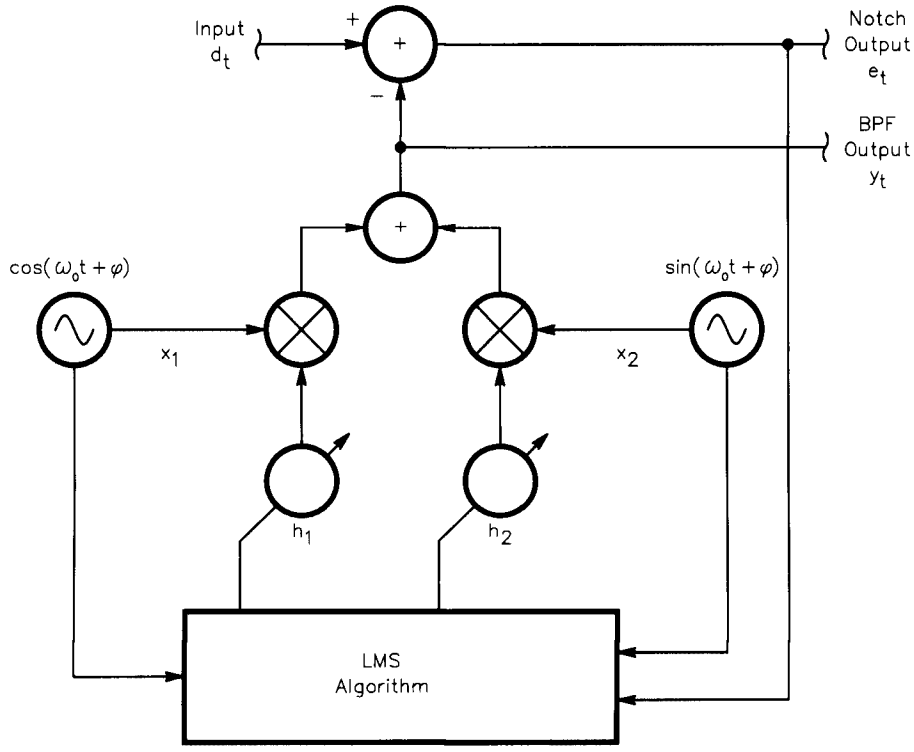


Fig 14—Block diagram of adaptive, manually tuned notch filter.

Widrow and Stearns, in which the number of taps in the band-pass filter is minimized. In fact, they were able to show that only two taps were needed for each frequency to be notched! DSP designers love this, since the amount of computation is almost nil. In describing this notch filter, we'll introduce the concept of the *adaptive interference canceler*, and we'll touch on some of the theory involved in adaptive signal processing.

The Adaptive, Manually Tuned Notch Filter

The situation is this: We want to copy a broadband signal, such as an SSB voice signal, and suddenly, a dreadful carrier appears in the passband! Our notch filter will remove it, and we'll have complete control over the notch width, along with a depth limited only by the bit resolution of our system.

Dr. Widrow discovered that one can build a filtering system to minimize repetitive signal energy by altering the filter coefficients "on the fly" using a certain algorithm. Known as the *least-mean-squares* (LMS) method, it describes a way to adjust FIR filter coefficients over time to remove an undesired tone in the input! A reference signal is used, which is of the exact frequency of the interfering tone. The algorithm then forms a band-pass filter that is subtracted from the broadband input to create the notch.

The block diagram of this system is shown in Fig 14. The broadband input is called d_t, and the reference input is a pure cosine wave:

$$x_t = A\cos(\omega_0 t + \phi) \qquad \text{(Eq 45)}$$

The cosine wave is sampled and fed to the input of one multiplier. It is also phase-shifted by 90° to produce a sine wave, which is fed to the second multiplier. The multiplier outputs are then added, as in a regular FIR filter, to form the band pass output. This output y_t is then subtracted from the broadband input signal to produce the notch output, e_t. Note that the band pass output is also available at no additional overhead.

While the initial values of the coefficients h_1 and h_2 are unimportant, the procedure for updating them is defined by the LMS algorithm as:

$$h_{1(t+1)} = h_{1t} + 2\mu e_t x_{1t}$$
$$h_{2(t+1)} = h_{2t} + 2\mu e_t x_{2t} \qquad \text{(Eq 46)}$$

where $0 < \mu < 1$, and the sampled reference inputs are:

$$x_{1t} = A\cos(\omega_0 t + \phi)$$
$$x_{2t} = A\sin(\omega_0 t + \phi) \qquad \text{(Eq 47)}$$

In the final analysis, it can be shown that as the reference inputs are sinusoidal, the system is linear and time-invariant for the output e_t. Several interesting points follow about the characteristics of this notch.

Adaptive Notch Filter Properties

First, the 3 dB bandwidth of the notch can be shown to be:

$$BW = \frac{2\mu A^2}{t_s} \qquad \text{(Eq 48)}$$

radians/second. The Q of the filter is simply the center frequency divided by the bandwidth:

$$Q = \frac{\omega_0}{2\mu A^2}(t_s) \qquad \text{(Eq 49)}$$

Therefore, we have control over the bandwidth by varying the factor, μ, and the amplitude of the reference signal, A, in the equation. The depth of the null is, in general, superior to that of a fixed filter because the algorithm will maintain the correct phase relationship for ideal cancellation, even if the reference frequency is changing slowly.

Each additional tone to be notched creates the need for two more taps in the adaptive filter. Any noise in the input causes us to add more taps to achieve sufficient accuracy. More detail of adaptive processing will be provided in future articles, and still more details may be found in the references listed.

Summary

DSP microprocessors are optimized for the multiply-and-accumulate (MAC) operation that forms the backbone of digital filtering and other DSP algorithms. We've discovered that FIR filter structures are definitely the way to go when designing DSP subsystems, and that quite a few nuances of theory make life easier.

In Part 2 of our review of DSP techniques, we'll look at an actual transceiver scheme, and examine how the strategies we've learned result in an efficient, high-performance design.

REFERENCES

1. Oppenheim, A. V. and Schafer, R. W., *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
2. Sabin, W .E. and Schoenike, E. O., editors, *Single Sideband Systems and Circuits*, McGraw-Hill, New York, NY, 1987.
3. Alkin, O., *PC-DSP*, Prentice-Hall, Englewood Cliffs, NJ, 1990.
4. Widrow, B. and Stearns, S. D., *Adaptive Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1985.
5. Frerking, M. E., *Digital Signal Processing in Communications Systems*, Van Nostrand-Reinhold, New York, NY, 1993.
6. Rabiner, L. R. and Gold, B., *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1975.
7. Zverev, A. I., *Handbook of Filter Synthesis*, John Wiley & Sons, New York, NY, 1967.
8. Panter, P. F., *Modulation, Noise, and Spectral Analysis*, McGraw-Hill, New York, NY, 1965.

Doug Smith, KF6DX/7, is an electrical engineer with 17 years experience designing HF transceivers, control systems, DSP hardware and software. He joined the amateur ranks in 1982, and placed on the air one of the first HF scanning bulletin-board systems. At Kachina Communications, Inc. in central Arizona, he is currently exploring the state of the art in digital transceiver design. □□

Examining the Mechanics of Wave Interference in Impedance Matching

Impedance matching is a mystery to many of us. Here's a short explanation of the subject from a well known engineer.

By Walter Maxwell, W2DU
ARRL TA for Antennas and Transmission Lines

Preface

The purpose of this article is to examine a powerful, but relatively unused treatment of wave-reflection mechanics as an aid in understanding how the matching of impedances is achieved by means of networks and transmission lines. This treatment focuses directly on the fundamental actions of four reflected waves, the voltage and current reflected waves, which result from two mismatches, the mismatched load and the mismatch introduced by the matching device. As we proceed, it will become evident that all impedance-matching operations are achieved through the superposition of these four reflected waves. By wave interference resulting from the superposition, the reflections from the mismatched load are canceled by the reflections generated by the matching device.[1] The cancellation of reflections at the matching point simultaneously creates a virtual, totally reflecting open or short circuit to all the reflected waves at the matching point *in the network or transmission line.* Consequently, no reflected waves travel rearward past the matching device, and a matched condition exists between two previously mismatched impedances.

243 Cranor Ave
DeLand, FL 32720-3914
e-mail **w2du@iag.net**

Background

Before presenting the discussion however, it seems wise to acquaint the reader with some background on the subject. During the 1970s, I published a series of *QST* articles,[2] and my ARRL book[3] in 1990, both aimed at dispelling prevalent misconceptions concerning wave propagation on transmission lines. One of the most serious misconceptions concerned reflected power reaching the tubes in the RF amplifier of the transmitter. The prevalent, but erroneous thinking was that the reflected power enters the amplifier, causing tube overheating and destruction. However, I dispelled this misconception in the above-mentioned publications, using the wave-mechanics treatment, discussed here in greater detail, by showing that when the pi-network tank is tuned to resonance, a virtual short circuit *to rearward*

[1]Notes appear on page 24.

traveling waves is created at the input of the network. Consequently, instead of the reflected power reaching the tubes of the amplifier, it is totally re-reflected toward the load by the virtual short circuit appearing at the network input.

However, please note that, in his *QEX* article, Bloom[4] has disputed my *QST* and ARRL book explanations (and that of MIT electrical engineering professor Slater[1]) of the wave action that results in total re-reflection of the reflected waves. Bloom maintains that such total re-reflection is impossible, because he claims the creation of a virtual short circuit at the pi-network input is "unrealistic."

Discussion

Whenever there is a mismatch of impedances that generates reflected waves—which must be eliminated by a matching device—it is fundamental that the elimination of the undesired reflections is achieved through wave cancellation by introducing new reflections that are complementary to the undesired reflections. The new reflections are generated by introducing a new mismatch tailored to produce reflections whose complex reflection coefficients are complemen-

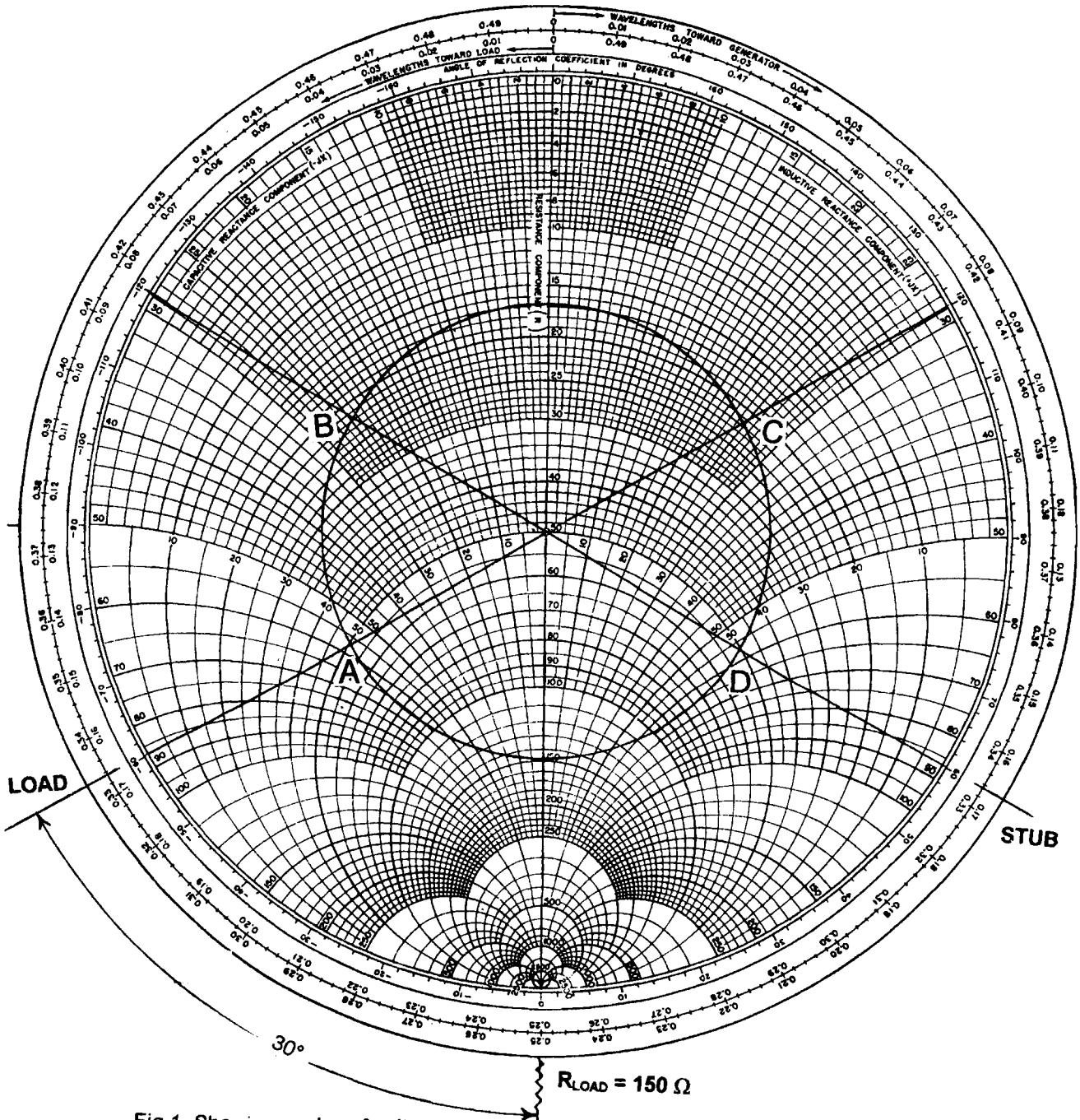


Fig 1. Showing angles of voltage and current reflection coefficients of load and stub.

tary to those of the original reflections.[1] By complementary, I mean coefficients that have identical magnitudes, but angles of the opposite sign. The new mismatch may consist of any appropriate network, or a transmission-line configuration.

The technique of stub matching on a transmission line provides a simple, but elegant model for an illustration of the wave-interference treatment, with the stub establishing the new mismatch. However, before discussing this technique using wave-interference treatment, it will be a study in contrast to review briefly the traditional treatment of matching with transmission-line stubs. To solve a matching problem in the traditional manner, engineering handbooks and texts provide equations and graphs to determine the length of the stub, and its position on the line, based on the characteristic impedance of the line (Z_0), the impedance of the load terminating the line and the resulting SWR. It is understood that the matching is achieved because the reactance (or susceptance) of the stub cancels the line reactance (or susceptance) at the stub point. However, the handbooks and texts do not explain the wave actions that produce the impedance matching or *why* the reflected waves produced by the mismatched load do not travel rearward beyond the stub. So let's now examine the treatment that *does* reveal what the handbooks and texts do not.

Matching via a Transmission-Line Stub

To illustrate stub-matching technique using a wave-interference treatment, consider first a 50 Ω line terminated with a pure 150 Ω resistance that yields a reflection coefficient of magnitude 0.5, for a 3:1 SWR on the line, as shown in Figs 1 and 2. By placing a series stub of the correct length at the correct position on the line, a 50 Ω impedance *match is obtained at the stub point; this results in a 1:1 SWR on the line between the source and the stub*, but the 3:1 SWR between the stub and the load remains (see Fig 3).

This we already knew, so now let's make a change that provides a fresh perspective. Leaving the stub in place, we replace the mismatched 150 Ω load with a 50 Ω matched resistance, as shown in Fig 4. Now there is no reflection from the load, so the SWR between the stub and the load is 1:1, but there is a reflection from the stub. This stub that initially reduced the 3:1 SWR to 1:1 also generates a reflection of magnitude 0.5, presenting a reactive 3:1 mismatch to the 50 Ω line. Thus, with the 50 Ω matched load terminating the line, the stub alone now produces a 3:1 SWR on the line between the source and the stub. So how does the stub achieve an impedance match with the 150 Ω resistance at the load? Because the reflection coefficients of magnitude 0.5 for both voltage and current produced by the stub reflections at the stub point are complementary to those arriving at the stub point from the 150 Ω mismatch at the load. So you now ask, "Why are there no reflected waves on the line between the source and stub with the 150 Ω load terminating the line?" *This is the crucial point of the discussion.* It is because the superposition of the voltage and current components of the reflected waves, generated by both the *load* mismatch and the *stub* mismatch, creates a *totally reflecting open-circuit to all reflected waves arriving at the stub point.* These reflected waves are then totally *re*-reflected by the open–circuit condition at the stub point, to travel only in the *forward* direction toward the load—they cannot travel rearward past the open circuit provided by the stub. This phenomenon is explained below and illustrated in Fig 5.

In the illustration above, the action of the stub created an open circuit at the stub point to rearward-traveling reflected waves (but *not to forward* traveling waves). However, the stub can create either an *open* circuit or a *short* circuit, depending on the relation between the terminating load impedance, Z_L, and the characteristic impedance, Z_0, of the line. When Z_L is *greater* than Z_0—as in the illustration above—the reflecting condition at the match point is an infinite impedance, a virtual *open* circuit. When Z_L is *less* than Z_0 the reflecting condition is a zero impedance, a virtual *short* circuit. To avert any misunderstanding concerning the conditions just described, I repeat for emphasis: *These open-or short-circuit conditions apply only to the* ***rearward****-traveling waves, not to the forward-traveling waves,* which see *only* the characteristic impedance of the line, Z_0.

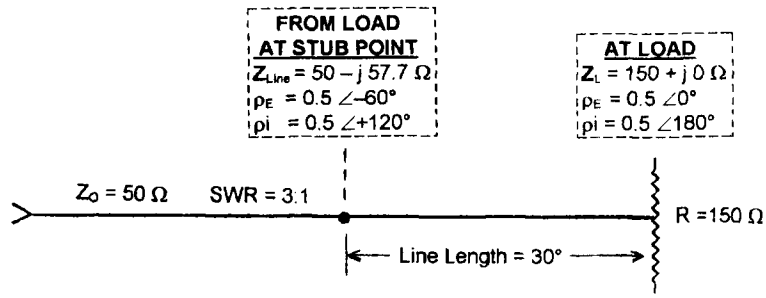


Fig 2. A 50-ohm transmission line terminated with a 150-ohm load, creating a 3:1 SWR along the entire line.

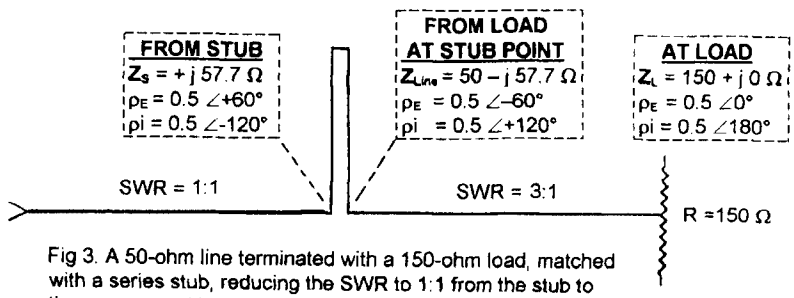


Fig 3. A 50-ohm line terminated with a 150-ohm load, matched with a series stub, reducing the SWR to 1:1 from the stub to the source, and leaving a 3:1 SWR from load to stub.

Let us now examine the wave-interference phenomenon that produces the totally-reflecting condition at the stub, or match point, *which is the basis for all impedance-matching operations.* Then following, we'll discuss how the same phenomenon occurs in λ/4 transformers and in pi networks comprising lumped components, which, when resonant, behave as λ/4 impedance transformers. After this discussion, the reason why the short circuit to rearward traveling waves appears at the input of a pi network in RF amplifiers will become clear.

By continuing the use of the 3:1 mismatch with the stub treatment, as in the illustration above, we obtain nice, even numbers that are easy to work with. With a 3:1 mismatch, the correct position on the line for the stub is 30 electrical degrees from the 150 Ω resistive terminating load, because at 30° from the load, the real part of the line impedance is 50 Ω (see point A in Fig 1). The reactive portion of the line impedance at this point is $-j57.7$ Ω. Thus the line impedance at this point is 50 $-j57.7$ Ω, normalized to the 50 Ω characteristic line impedance as 1 $-j1.15$. In this illustration, we will use impedance treatment, not admittance. Therefore (as we already know), placing an inductive stub having $+j57.7$ Ω of reactance ($+j1.15$ normalized) in *series* with the line produces an impedance match. This is so because the stub's $+j57.5$ Ω of inductive reactance cancels the $-j57.7$ Ω capacitive line reactance, which leaves 50 + $j0$ Ω (1 + $j0$ normalized) as the input impedance looking into the line toward the load. This is true when looking from anywhere between the source and the stub point. This is also old hat, but it doesn't explain *why* there are no reflected waves on the line between the source and the stub.

However, as we will see, more action occurs at the stub point—in preventing reflected waves from traveling rearward past the stub—than simply canceling reactances. To discover how the open circuit to rearward-traveling waves at the stub point is established by wave interference, we must examine the reflection coefficients produced by both the load and stub mismatches. Keep in mind that for every electrical degree of motion along a mismatched transmission line the angle θ of the reflection coefficient changes *two* degrees. Thus the angle of the voltage reflection coefficient at the stub point (30° from the resistive load) resulting from the 3:1 load mismatch is −60°, as indicated on the reflection-angle scale at its intersection with the radial line extending from point A in Fig 1. Because reflected current is always 180° out of phase with its corresponding reflected voltage, the angle (θ) of the current reflection coefficient at the stub point for the load mismatch is +120° as indicated by the radial extending from point C.

Let us now examine the reflection characteristics of the stub mismatch. Keep in mind that, to examine the stub mismatch separately, we replaced the 150 Ω mismatched load with a 50 Ω matched load, as shown in Fig 4. Thus the $+j57.7$ Ω (point D in Fig 1) of inductive stub reactance in series with the 50 Ω matched line results in an impedance of 50 + $j57.7$ Ω and a voltage reflection-coefficient angle θ of +60° at the stub point, as indicated by the radial extending from point D. And again, since reflected current is always 180° out of phase with the reflected voltage, angle θ of the current reflection coefficient of the stub reflection at the stub point is −120° (see radial through B). Next, it is important to note that—because the load and stub both yield 3:1 mismatches—the *magnitudes* of all four reflection coefficients are identically 0.5. This is one of two requirements for the total cancellation of the load-reflected wave by wave interference and creation of the virtual open circuit at the stub.

We will now discover that the second requirement for creating the open circuit is the relationship of the angles (θ) of the two sets of reflection coefficients. First we consider the two *voltage* angles at the stub point: −60° for the load reflection and +60° for the

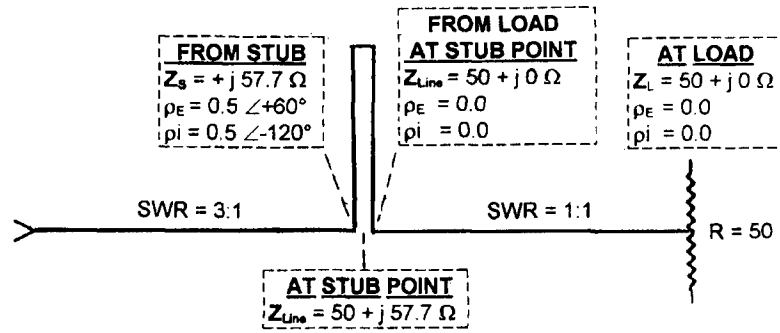


Fig 4. Series stub on **matched** line, generating a mismatch complementary to load mismatch generated by 150-ohm load in Figs 1, 2, and 3.

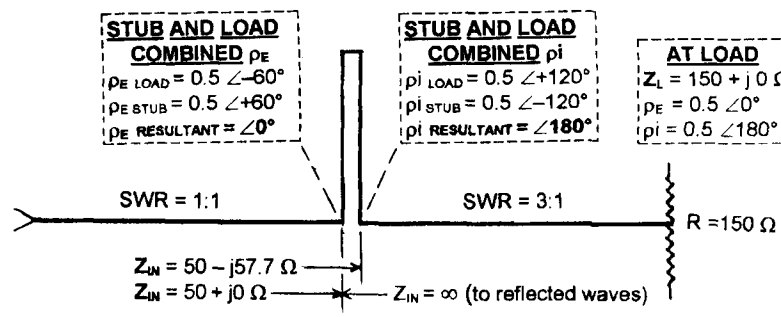
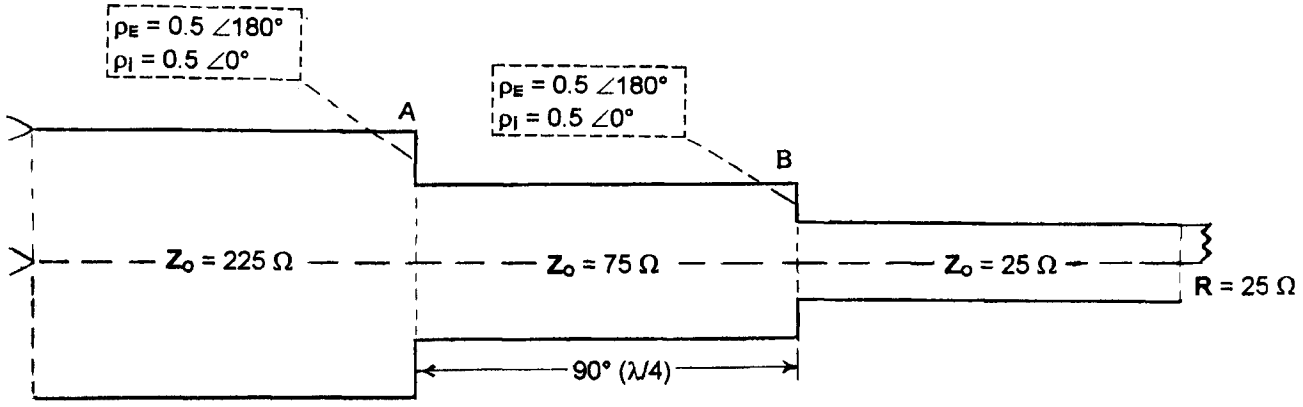


Fig 5. Showing how resultant angles of voltage and current reflection coefficients $\rho_E\theta$ = 0° and ρiθ =180° at the match point define a virtual open circuit at the match point to waves reflected by the load.

stub reflection makes their resultant 0°. Next we consider the two *current* angles at the stub point: +120° for the load reflection, and –120° for the stub reflection makes their resultant 180°. In these two resultant angles, voltage 0° and current 180°, lies the secret of the open circuit to reflected waves being created at the stub point, which I'll now explain with the aid of Fig 5.

From universally–known transmission-line theory, when the terminating load on a line is an open circuit, the angles of the reflection coefficients are 0° for voltage and 180° for current. Conversely, when the terminating load is a short circuit, the angles are reversed, 180° for voltage and 0° for current. Now for the crucial factor in establishing the virtual short or open circuit at the stub. In both cases, these conditions are reciprocally related. An open-circuit termination yields 0° and 180° for voltage and current, respectively, but conversely, if by some means we *establish* reflection coefficients of equal magnitude, with angles of 0° for voltage and 180° for current at any given point on a line, that combination of reflection–coefficients *creates* an open circuit at that point to waves traveling rearward. Means for establishing these reflection coefficients of equal magnitude, angles 0° for voltage and 180° for current, is illustrated in Fig 5. Continuing, if we *establish* reflection-coefficient angles of 180° for voltage and 0° for current at a given point, that combination *creates* a short circuit at that point to waves traveling rearward. So having already established the resultant voltage and current θ angles of 0° and 180°, respectively, with equal magnitudes of 0.5 at the stub point in the present example, we have demonstrated that an *open* circuit to rearward traveling waves has been created there, causing them to be totally re-reflected toward the load at the stub point. Consequently, if there are no reflected waves on the line between the stub and the source, the line must be matched, with an SWR of 1:1 between the source and the stub.

Before proceeding to discuss the corresponding conditions of matching with λ/4 transmission lines and pi-networks, let us first examine what happens to the reflected waves after re-reflection at the stub point. On re-reflection by the open circuit appearing at the stub point, the resultants of the voltage and current components of the load- and stub-re-reflected waves emerge in the forward direction *in phase* with the corresponding components of the source waves. Why do the re-reflected resultant waves emerge *in phase* with the source waves? Keep in mind that just prior to the re-reflection the reflection angles θ of the resultants were 0° and 180° for voltage and current, respectively. Also, keep in mind that on reflection from an open circuit, angle θ for voltage remains unchanged, while angle θ for current changes by 180°. Therefore, on re-reflection the voltage angle remains at 0°, and the current angle reverses from 180° to 0°, bringing both in phase with the forward voltage and current from the source. Consequently, the power in the reflected waves adds directly to that of the source waves at the stub point. This accounts for power measurements in the transmission line between the stub and the load that are greater than that supplied by the source, by the amount of the power reflected. Of course, the power reflected at the load then subtracts from the sum of the source and reflected powers, leaving the source power to be totally absorbed by the load. (We're considering only loss-less lines here.) This same total-reflection phenomenon occurs with a properly tuned antenna tuner, which is why no reflected power is observed on the line between the transceiver and the input of the tuner. (With an SWR of 12:1 on your open-wire feed line, has it puzzled you why there is *no reflected power* on the line from the transceiver to the tuner? Why is the forward power in the open-wire line *greater*



Due to the 180° phase delay explained in the text (two-way 90° from point A to B to A), forward waves reflected at B arrive back at A after complete reversal of reflection angles from $\rho_E\theta$ = 180° and $\rho i\theta$ =0° to $\rho_E\theta$ = 0° and $\rho i\theta$ =180°, 180° out of phase with their corresponding reflections of forward waves reflected at A. Result: Creation of a totally-reflecting virtual short circuit at A to waves reflected at B.

Fig 6. Impedance Matching with 90-degree (λ/4) Transmission-Line Transformer

than the *true* power, the power delivered by the source and absorbed by the load? Here we're treating the open-wire line as *loss-less* line.)

Matching via a λ/4 Transformer

Before we discuss impedance matching with λ/4 transmission-line impedance transformers, you should be aware of an important refinement concerning the definition of reflection-coefficient angles: When a traveling wave in a transmission line encounters either an open-circuit termination or a change in line impedance from low to high at a junction, reflection angle θ is 0° for voltage and 180° for current, when the higher-impedance line is terminated in a matched load. Conversely, when the wave encounters a short-circuit termination, or a change in line impedance from high to low at a junction, angle θ is 180° for voltage, and 0° for current when the lower-impedance line is terminated with a matched load.

In explaining the wave action in the λ/4 line-matching transformer as shown in Fig 6, let's assume that two transmission lines having impedances of 225 and 25 Ω, respectively, are to be matched by a λ/4, or 90°, matching transformer. The source feeds the input of the 225-Ω line, and the 25 Ω line is terminated in a pure resistance of 25 Ω.

To match the 225 Ω and 25 Ω impedances of the two lines, the characteristic impedance (Z_0) of the matching line must be equal to the geometric mean of the two impedances being matched, which in this case is 75 Ω. The 225 Ω line from the source is connected to the input of the 75 Ω matching line, point A, the 25 Ω line is connected to the output of the matching line, point B, and terminated with a load of 25 $+j0$ Ω. In the absence of the matching line, the mismatch between the two lines is 9:1 for a reflection-coefficient magnitude of 0.8. However, with the matching line in place there will be no wave reflections on either transmission line, but there will be reflections on the matching line resulting from the 3:1 mismatches appearing at both ends of the matching line. The matching is performed by the behavior of two sets of reflected waves traveling within the matching line, which are produced by the 3:1 mismatches appearing at both the input and output junctions of the matching line and the 225 Ω and 25 Ω lines, respectively. The 3:1 mismatch to forward-traveling waves occurring at the input of the matching line produces two wave reflections, one of voltage and one of current. Because the impedance change is from high to low, the reflection coefficients at the input are ρ_E 0.5 at θ = 180° for voltage, and ρ_I0.5 at θ = 0° for current. These same reflection coefficients appear again as the continuing forward-traveling waves arrive at the output of the matching line, because the impedance change for the continuing forward waves is again from high to low with the identical mismatch as that at the input. Initially the waves reflected at the input of the matching line travel rearward into the 225 Ω line, and the waves reflected at the output (of the matching line) travel rearward into the 75 Ω matching line. However, the matching action occurs at the input of the matching line. So to determine how the matching is accomplished, we are concerned with the reflection coefficients of the four reflected waves as they appear on simultaneous arrival at the input of the matching line. As stated above, the magnitudes of all four reflected waves are identical, ρ = 0.5, and the angles of the coefficients generated at both the input (point A) and output (point B) of the matching line are 180° for voltage (E), and 0° for current (i). On arriving rearward at the input, however, the waves reflected from the output junction have traveled 180° farther; that is a half-wavelength more than the forward waves arriving at the input from the 225 Ω line. Therefore, while angles at the input of the matching line for waves reflected at the input mismatch are 180° for voltage, and 0° for current, the angles at the input for the waves reflected from the output mismatch are reversed: 0° for voltage and 180° for current, due to the extra 180° of travel. Consequently, all corresponding components of voltage and current are 180° out of phase at the matching point. This condition is equivalent to two waves a half-wavelength apart. With equal magnitudes and opposite phase at the same point (point A, the matching point), the sum of the two waves is zero. Result—zero impedance—a short circuit to all rearward-traveling waves, and total re-reflection of the waves reflected from both the input and output mismatches in a manner identical to that described in stub matching.

λ/4 Transformer versus Pi Network

Let us now examine similarities among of the impedance-transforming characteristics of the pi-network and those of the λ/4 transmission line to explain why a virtual short circuit to reflected waves appears at the input of a pi-network fed by an RF power amplifier. Because the pi-network at resonance in a RF power amplifier is electrically equivalent to a λ/4 (90°) transmission-line impedance transformer, the network performs in the same manner as described above. The only difference is in the values of the impedances.

For example, let's assume 1800 Ω as a realistic optimum load impedance of the amplifier, and 50 Ω as the output impedance. The geometric mean of these impedances is 300 Ω, establishing the characteristic impedance (Z_0) of the pi-network performing as a 90° matching transformer. When the network is tuned to resonance with a 50 + $j0$ Ω termination at the output, the 50 Ω resistive load produces a mismatch of 6:1 at the output of the 300 Ω network line. The amplifier source impedance of 1800 Ω also sees a 6:1 mismatch at the input of the 300 Ω pi network. As in the λ/4 matching-line operation described above, the electrical λ/4 of the network performs the same half-wave time delay for the reflection from the output mismatch to reach the input. Here too, the two sets of reflected waves are 180° out of phase at the input terminals to create a zero impedance there to all reflected waves.

Now we come to the crucial point of this entire discussion. So far, we have assumed the load at the output of the network is 50 Ω resistive. As we know, however, if the load is a mismatched 50 Ω transmission line, the input impedance of the line loading the network will not be 50 Ω. If the network has been loaded and resonated with a 50 Ω termination, and then connected to an unmatched line without retuning and reloading to derive the same output power as before, the waves reflected from the input and output mismatches will no longer be equal and opposite, and the zero impedance previously established at the network input will no longer be zero. In this condition, wave reflections resulting from the different load impedance will travel past the input of the network to combine vectorially with the source waves. This creates a reactive load impedance at the input of the network, which changes a straight load line into an elliptical load line. However, the short circuit will be re-established when the network is retuned and reloaded to

compensate for the difference in the load impedance, because then the network is again equivalent to a λ/4 transmission-line transformer. In other words (contrary to Bloom)[4], when the pi–network is properly tuned and loaded with any complex impedance within its compensation range, the impedance to rearward-traveling reflected waves at the input is zero. This prevents reflected energy from seeing the internal impedance of the source as a terminating load.

On the Input Impedance of Open- and Short-Circuit Transmission-Line Stubs

After examining the wave mechanics that establish the impedance-matching phenomenon in λ/4 transmission–line transformers, we would be remiss if we do not examine the wave mechanics that establish the input impedances of open- and short-circuited transmission-line stubs. In this discussion, all lines are considered loss less.

A λ/4 Open Stub

Let us first consider a λ/4 open-circuit stub. On arrival at the input of the stub the source voltage and current are in phase, at which point we assign a reference phase of 0°. On arriving rearward at the input, the voltage and current waves reflected from the open–circuit termination of the stub have traveled 180° relative to those arriving at the input. (They travel 90° from the input to the open-circuit termination and another 90° in returning to the input.) However, keep in mind that, on reflection from an open circuit, angle θ of the voltage reflection coefficient experiences no change, while angle θ for current changes 180°. Thus, on return to the input, angle θ of voltage waves reflected from the open-circuit termination is now 180°, a complete phase reversal relative to the source. Consequently, the reflected voltage appearing at the input of the stub is of the *opposite polarity* to that arriving from the source. Continuing, with the 180° angle change in current on reflection at the open-circuit termination, plus 180° travel on the stub line, on return to the input, angle θ for the reflected current is 0°. It is *in phase* with the source current. So we have two equal voltages of opposite polarity at the input, the source voltage and the stub voltage, with their currents in phase. What's the result? Maximum current—zero impedance—*a short circuit!* A perfect dc analogy is two batteries connected in series, plus to minus, and plus to minus—short circuit—maximum current.

A λ/4 Shorted Stub

Let us now consider a λ/4 short-circuit stub. The only differences relative to the conditions prevailing in the open-circuit stub are the θ angles of the voltage and current reflection coefficients of the waves reflected from the short-circuit termination that appear on return to the input. On reflection from a short circuit, angles θ of the reflection coefficients for both voltage and current are opposite to those reflected from an open circuit. That is, on reflection from a short circuit the voltage angle θ changes by 180°, while the current angle θ does not change. Consequently, on return to the input, the θ angles of the voltage and current waves reflected from the short-circuit termination of the stub are 0° and 180° respectively, opposite to those returning at the input of the open-circuit stub. As a result, the source and stub voltages are equal and of the *same polarity at the input*, while the currents are *of opposite phase*. What's the result? Zero current—an infinite impedance—an *open circuit*. A perfect dc analogy is two batteries (of equal voltage) connected in parallel, plus to plus, and minus to minus—open circuit—zero current.

Nonresonant Stub Lengths

So far we have considered only stubs of λ/4, the length that yields zero reactance, because the stub-reflected voltage is either exactly in phase, or 180° out of phase with the source voltage. However, as we know, *capacitive* reactance is present in a circuit when current leads the voltage, and *inductive* reactance is present when voltage leads the current. So, let us examine the wave mechanics of stubs that do yield reactance by causing the voltage and current to be out of phase. Let us begin with stubs that provide *capacitive* reactance by causing current to lead the voltage. These are *open*-circuit stubs having lengths *less* than λ/4. When the stub length is less than a λ/4, the stub *current leads the voltage* at the input of the stub because the current reflected from the open–circuit termination has traveled *less* than 180° relative to the waves arriving from the source. As a result (while including the 180° phase reversal on reflection from the open circuit), the reflected current arrives back at the input of the stub *ahead* of the source voltage. Because of this *current-leading-voltage* phase relationship between the source voltage and the reflected current arriving at the input of the stub, a *capacitive* reactance $-jX$ is developed at the stub input.

For this knowledge to be useful, we need to establish the relationship between stub length and the resulting reactance obtained from the superposition of the source and reflected waves at the input of the stub. First, the phase relationship between the source and reflected waves at the stub input is determined by angle θ of the reflection coefficient of the reflected waves. Angle θ is directly related to the stub length by the expression $L_S = \theta/2$, where L_S is the electrical stub length. Both L_S and θ are in degrees. The stub length in wavelengths may be found using the expression $L_{S\lambda} = L_S°/360°$. Finally, the capacitive reactance at the stub input may be found using the expression $-jX = Z_0 \cot L_S$.

As an example, let us find the capacitive reactance obtained with an open-circuit stub having a measured current reflection coefficient angle θ of 90°, and where the characteristic impedance Z_0 of the stub line is 50 Ω. From angle θ, the stub length is $L_S = \theta/2 = 90°/2 = 45°$. The length $L_{S\lambda}$ is 45°/360° = 0.125 λ, or λ/8. The capacitive reactance $-jX = Z_0 \cot L_S = 50 \cot 45° = 50 \times -1.0 = -j50\ \Omega$. The stub length required to obtain a given capacitive reactance may be found by reversing the above procedure.

Having examined the open-circuit stub for its capability of obtaining *capacitive* reactance, let us now examine the short-circuit stub for its capability of obtaining *inductive* reactance. Keep in mind that in the open–circuit stub, angle θ of the voltage reflection coefficient remains unchanged on reflection, and angle θ for current reverses 180°. The reverse is true with the short-circuit stub; angle θ for voltage reverses 180°, and angle θ for current remains unchanged.

When the short-circuit stub length is less than a λ/4, the stub *voltage leads the current* at the input, because the voltage wave reflected from the short-circuit termination has traveled *less* than 180° relative to the waves arriving from the source. As a result (while including the 180° phase reversal on reflection from the short circuit), the reflected voltage wave arrives back at the input of the stub *ahead* of the source current wave. Because of this *voltage-leading-current* phase relationship between the source current and the reflected

voltage arriving at the input of the stub, an *inductive* reactance $+jX$ is developed at the stub input. Borrowing from the open-ircuit expressions above, the stub length is obtained in the same manner, and the inductive reactance at the stub input may be found from the expression $+jX = Z_0 \tan L_S$. (Note that the tangent function is used for short-circuit stubs, yielding $+jX$, and the cotangent is used for open-circuit stubs, yielding $-jX$.)

As an example, let us find the inductive reactance obtained with a short-circuit stub having a measured voltage reflection-coefficient angle θ of 90° and a characteristic impedance (Z_0) of 50 Ω. From angle θ, the stub length is $L_S = \theta/2 = 90°/2 = 45°$. The stub length L_{S_λ} is $45°/360° = 0.125\lambda$, or $\lambda/8$. Inductive reactance $+jX = Z_0 \tan L_S = 50 \tan 45° = 50 \times 1 = +j50$ Ω.

It is interesting that, because tan 45° = 1.0, and cot 45° = −1.0, the $\pm jX$ reactance obtained with short- or open-circuit stubs of length 45° ($\lambda/8$) always equals the value of the characteristic impedance Z_0 of the stub line.

Summary

This article has revealed five fundamentals concerning the phenomenon of impedance matching:

1. The cancellation of mismatch-engendered reflected waves is achieved through wave interference.

2. The mechanism of wave interference is fundamental to impedance matching with transmission-line stubs, $\lambda/4$ transmission-line transformers and impedance-matching networks comprising lumped elements.

3. Reflected waves cannot travel rearward beyond the matching point in the matching device because they are stopped by the virtual open or short circuit to rearward–traveling waves, which appears at the matching point.

4. The virtual open or short circuit results from wave interference between two sets of reflected waves. One set is produced by the load mismatch and the other by a matching device tailored to produce reflected waves that are complementary to those produced by the load mismatch.

5. Because a resonant pi network is equivalent to a $\lambda/4$ transmission-line transformer, a virtual short circuit to rearward-traveling waves appears at the inputs of pi networks used in the output circuitry of RF amplifiers, when the network is at resonance and correctly matched to any complex load within the range of the matching capability of the network.

In addition, the article has also examined the capabilities of open- and short-circuit transmission-line stubs to obtain capacitive and inductive reactances from the distributed constants of the line.

Acknowledgement

My thanks to Roy Lewallen, W7EL, for taking the time to review the manuscript and for his helpful suggestions, which have contributed to the clarity of my message.

Notes

[1]Slater, J. C., *Microwave Transmission*, (New York: McGraw–Hill Book Company, 1942), pp 7–69.

[2]Maxwell, Walter, W2DU, "Another Look at Reflections," *QST*, a series of articles, 1973: April, June, August, October*; 1974: April, December; 1976: August*. *(Specific attention to matching by wave interference.)

[3]Maxwell, Walter, W2DU, *Reflections—Transmission Lines and Antennas*, Newington, CT, ARRL, 1990, pp 4-3 to 4-10* , 9-3.*

[4]Bloom, Jon, KE3X, "Where Does the Power Go?" *QEX*, Dec 1994, p 17. □□

Voltage Multipliers

Need a little more voltage or a bipolar supply? Here are some basic circuits to meet many needs.

By Parker R. Cope, W2GOM/7

When your project requires ±15 V for op-amps and you have only a single +9 V supply, a voltage multiplier can be a relatively simple solution to the problem. The voltage multiplier is the predecessor of the more modern dc/dc converter or switch mode power supply (SMPS). A voltage multiplier is unregulated and not as versatile as an SMPS, but it is simpler. It's quite appropriate for developing the relatively high dc voltage for light single-polarity loads, such as plasma displays, as well as split-voltage supplies for op-amps or other analog circuits. For many home projects, the voltage multiplier is less expensive and more convenient than winding or locating suitable magnetic parts for use in a SMPS.

8040 E Tranquil Blvd
Prescott Valley, AZ 86314

A voltage multiplier can produce a dc voltage, of either polarity, that is (approximately) a whole number multiple of the peak-to-peak input voltage. To illustrate the theory and procedures for designing any voltage multiplier, the design of +15 V and −15 V, 5 mA supplies operating from a 9 V battery are described. The voltage multiplier requires an input power level slightly greater than the output power. Of course, an efficiency of less than 100% should not come as a surprise, for to paraphrase the second law of thermodynamics, "You can't get something for nothing, and you're lucky to break even." The ac-input power can be provided by a high-power oscillator or a low-power oscillator followed by an amplifier. The ac voltage is rectified to produce the output dc.

The voltage doubler, sometimes called a peak-to-peak detector, is named for its dc output, which is approximately double its peak input. The average (dc) value of a full-wave rectified sine wave is $0.637 \times V_{pk}$, while the RMS value is $0.707 \times V_{pk}$. The peak, average and RMS values of a full-wave rectified square wave are all equal. In passing, note that a square wave can be generated more efficiently than a sine wave.

A CMOS square-wave oscillator, shown as U1A and U1B in Fig 1, has an approximate operating period of 1.2 R1 C1. R2 merely limits the current in the input protection circuit of U1A and can be any large value. When

U1 is a CD4009 hex inverter, each section can swing 1.6 mA to within 50 mV of the dc supply rails. Four inverters in parallel could supply about 6.4 mA (peak). To obtain higher currents, use an amplifier such as the one comprised of Q1 and Q2.

The pair of complimentary BJTs, Q1 and Q2 in Fig 1, amplify the oscillator's output current. The maximum output voltage of the amplifier is determined by the collector-emitter saturation voltage VCE_{sat} of the transistors. VCE_{sat} occurs when the base-collector junction is forward biased. For example, a 2N3904 or '3906 bipolar transistor has a VCE_{sat} on the order of 0.1 V, when the base current is 1 mA and the collector current is 10 mA (ie, the "forced beta" is 10). The VCE_{sat} increases to 0.3 V when the base current is 0.1 mA and collector current is 10 mA (the forced beta is 100). In the example, the peak collector current is 40 mA, and base current is limited to about 0.8 mA. This produces a VCE_{sat} of 0.3 V. With a 9 V dc source, the output of the amplifier swings from 0.3 V to 8.7 V; the output is 8.4 V_{P-P}.

When the output is taken from the transistors' collectors as shown in Fig 1, "dead time" must be introduced in the base drive to prevent both transistors from conducting simultaneously during the on-to-off transition. The turn-off delay is lengthened and the turn-on delay shortened as the base drive increases to reduce VCE_{sat}. Prudence requires delaying the turn-on of the nonconducting transistor until after its counterpart switches off. For the 2N3904 and 2N3906, the storage time is given as about 0.2 µs and the turn-on time is about 0.035 µs. The networks at the input of the inverters (U1C, D, E and F) control the turn-on drive delays of the transistors. The R3-C2 combination slows the rise of the voltage on the input of inverters U1C and U1D and delays the turn-on of Q1 by about 0.2 µs. D1 allows the voltage on C2 to fall with minimum delay and turn-off Q1 as quickly as possible. In a similar fashion, R4, C3 and D2 control the conduction delay of Q2. D1 and D2 can be any fast switching diode such as the 1N4148 or 1N914.

Coupling the 8.4 V_{P-P} output of the amplifier to the rectifiers through C11 blocks the 4.5 V dc that is coincidental to the amplifier output. The result is an ac voltage with a positive peak of 4.2 V and a negative peak of 4.2 V. Connecting the diodes as shown in Fig 2 produces a positive voltage equal to the peak-to-peak ac input minus two diode drops. D11 clamps the input to one diode drop below ground on the negative half of the input cycle. On the positive transition of the input, the voltage at the cathode of D11 and the anode of D12 rises 8.4 V. Diode D12 charges C12 to the peak of its input less its forward drop. The output of the voltage doubler is the peak-to-peak input voltage less two diode drops. Schottky diodes similar to the 1N5817 have a forward drop of about 0.3 V, while silicon rectifiers similar to the 1N4001 have a forward drop of about 0.6 V. With Schottky diodes the

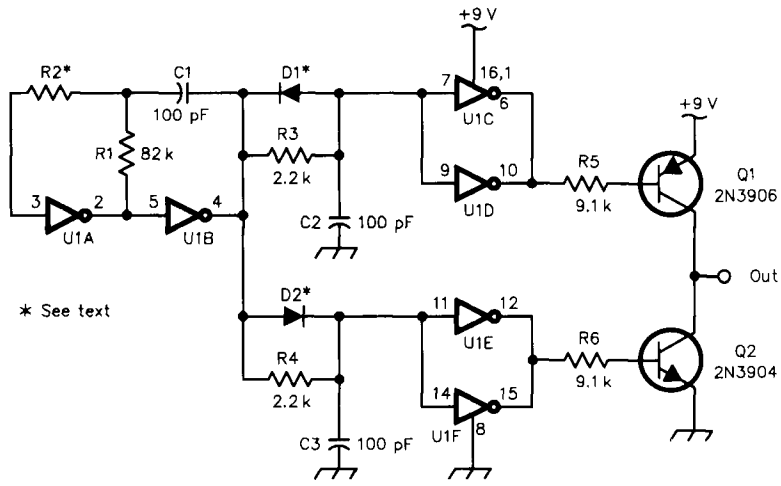


Fig 1—A simple RC oscillator and amplifier generate an ac voltage.

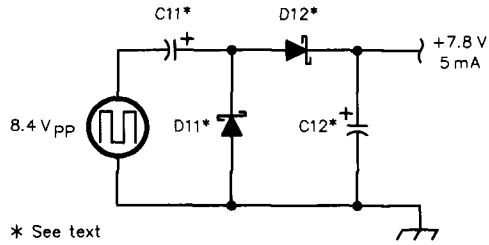


Fig 2—A peak-to-peak rectifier or voltage doubler.

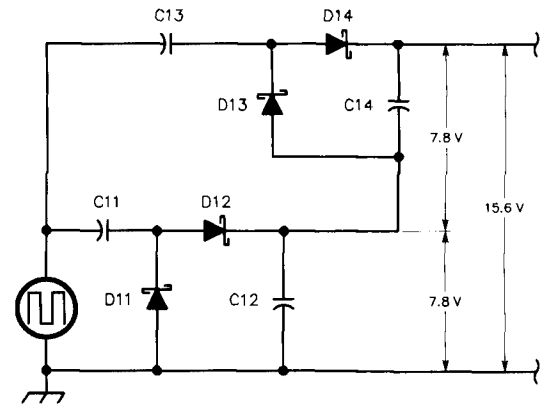


Fig 3—Two voltage doublers can be stacked to produce twice the voltage.

output is 7.8 V, but with silicon diodes the output would be 7.2 V. Therefore, Schottky diodes are recommended, and they are used in the example.

A second doubler can be stacked on the first (as shown in Fig 3) to produce an output of 15.6 V. A third can be stacked on the second to produce 23.4 V and so on. [Note: as you stack doublers beyond 2 or 3, the impedance of the supply increases, and the output voltage becomes somewhat less than predicted.—*Ed.*] The polarity of the output voltage is changed by reversing all the diodes. Balanced 15.6 V supplies can be produced with two positive doublers and two negative doublers as shown in Fig 4. The doublers in Fig 4 are redrawn in the conventional form.

Ideally, the dc power output of the multiplier is equal to the ac power input less the circuit losses. The diode's forward voltage drop is the major loss in the rectifier, but the diodes have another loss mechanism that comes into play when the input frequency is high: *reverse recovery time*.

Reverse recovery time is the time it takes for the minority carriers in the diode junction to be swept out when the voltage reverses. During the reverse recovery time, a diode conducts in the reverse direction. Schottky diodes have reverse recovery times of a few nanoseconds while those for silicon rectifier diodes may be in the range of 0.05 to 1 μs. Therefore, Schottky diodes should be used with ac input frequencies greater than 1 MHz. Reverse recovery time of slower silicon rectifiers may become a factor for frequencies as low as 20 kHz. Above 10 MHz, we must consider equivalent series inductance and equivalent series resistance of the capacitors. A 0.01 μF capacitor with two $^1/_8$-inch leads resonates near 10 MHz. Above the resonant frequency, the capacitor acts like an inductor, and its impedance increases with frequency. For frequencies below a few hundred kHz, the capacitor's inductance is not significant. The shunt capacitance of the diodes also reduces the output voltage at high frequencies. Forward voltage drop remains a major loss at any frequency. Absent some special requirements, a frequency of about 100 kHz is a reasonable compromise.

The ripple on the output dc voltage is determined by the load current and the capacitance of the filters. The voltage across any capacitor changes as it is charged or discharged. The initial rate of change of voltage on a capacitor is related to the current into or out of it and can be expressed as:

$dV/dt = I/C$, or $dV = (I/C)dt$

where,

dV = change in voltage across the capacitor.

dt is the time duration of charging or discharging, in seconds.

I is the charging or discharging current, in amperes.

C is the capacitance (in Farads) being charged or discharged.

For the circuit of Fig 2, the discharge

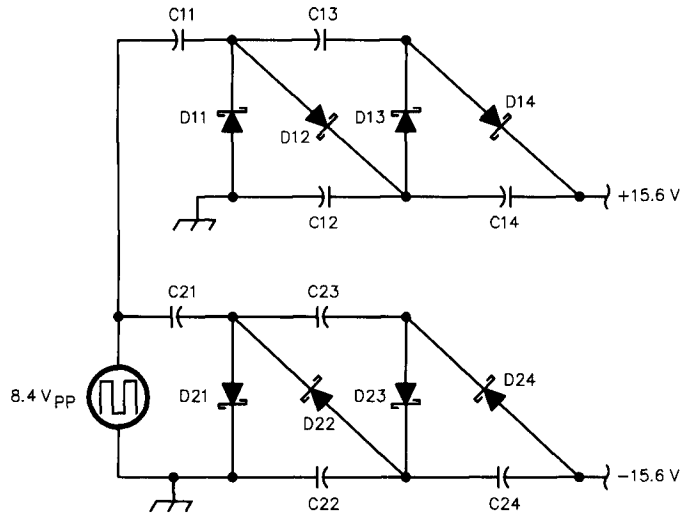


Fig 4—A voltage multiplier can produce either positive or negative voltages.

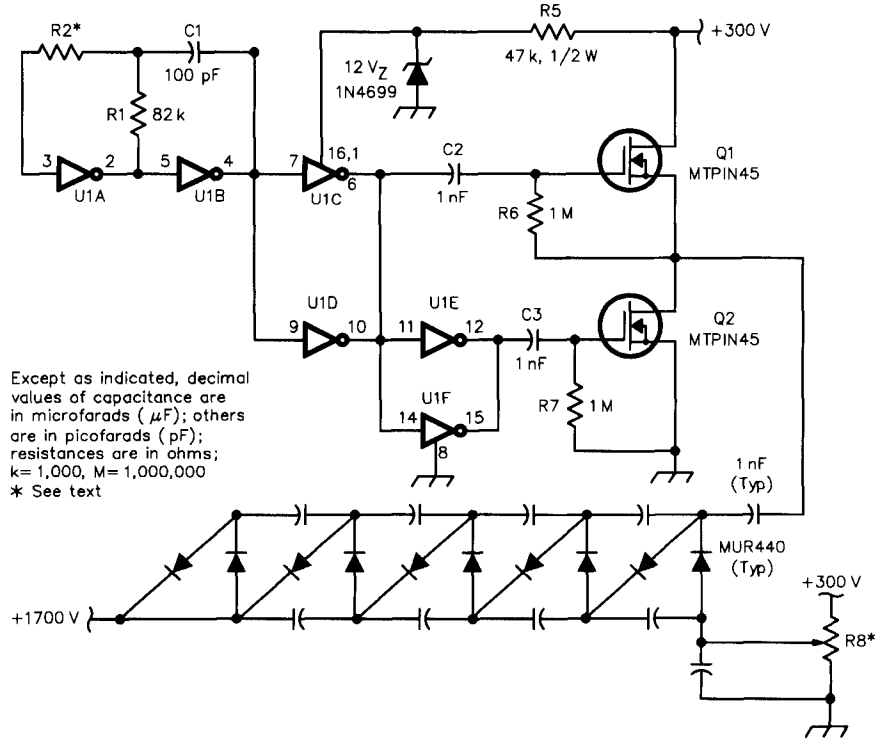


Fig 5—A voltage multiplier can develop scope anode supplies.

time for C11 is the duration of the positive half-cycle of the ac-input voltage. The discharge time of C12 is the duration of the negative half cycle. From this, we can conclude that a higher input frequency needs smaller capacitors for a given ripple voltage and load current. Choosing the operating frequency is a trade-off: high frequencies require consideration of parasitics in the capacitors, diodes and the ac input generation, as well as circuit layout. Low frequencies require high capacitance, but ac generation and layout are less of a problem. When the frequency is 100 kHz or so, and the current/ripple voltage demands it, solid tantalum capacitors such as Sprague 196, Kemet T350 or Panasonic ECS are satisfactory. The dc working voltage (DCWV) of the capacitors and the peak inverse voltage (PIV) of the diodes must be greater than the peak-to-peak input voltage. Operating below the DCWV and PIV ratings of the components leads to longer component life.

In the circuit shown in Fig 1, the oscillator has a frequency of about 100 kHz, (half cycle period of 5 μs). When the load current on the voltage doubler of Fig 2 is 5 mA and the capacitors are 2.2 μF, the voltage on the capacitors changes 2.3 mV per microsecond, or 11.5 mV_{p-p} ripple for a 5 μs discharge time. The voltage doubler requires 10 mA_{pk} from the amplifier when the doubler's load current averages 5 mA. When the output voltage is 7.8 V, 39 mW is dissipated in the load. The current from the 9 V source is 10 mA_{pk}, 5 mA_{ave}, and power drawn from the 9 V source is 45 mW, for an efficiency of about 87%.

Fig 3 shows two doublers stacked to produce a quadrupler, with an output that is approximately four times the peak-to-peak ac input. The ripple voltage is also twice the ripple voltage of the doubler. The current required from the amplifier for a quadrupler is twice the current required for a doubler or 20 mA_{pk} (10 mA_{ave}). When two quadruplers are used—one for +15.6 V at 5 mA and the other for −15.6 V at 5 mA as shown in Fig 4—the amplifier draws 40 mA_{pk} (20 mA_{dc}). The quadruplers shown in Fig 4 are redrawn in the more usual form.

The voltage multipliers shown are not regulated, and the output varies with dc source or load changes. For many applications, however, supply voltage regulation is not critical. For example, most op-amps can operate with a wide range of supply voltage, although the maximum output swing of any op-amp is ultimately determined by the supply voltages. Plasma displays are relatively tolerant of power-supply voltage variation. Supply changes of 10% or even 30% affect brightness but are otherwise quite acceptable.

If regulated or variable supplies are necessary, the output can be varied by varying the dc source voltage. It is conceivable to vary the source voltage by means of a resistance in the collector of Q1 or to control the dc voltage supplied to Q1 with a pass transistor. When regulation is required, a sample of the output is fed back to the pass transistor. Choose the time constants and gains in the feedback loop carefully to avoid oscillation.

Unregulated voltage multipliers are adequate for many applications and offer a simple solution to obtaining higher voltages or polarity changes when only a single dc source is available. It is practical to generate the anode voltage for a scope with voltage multipliers.

For example, if a scope has a 300 V supply, a high voltage amplifier and rectifier can produce 300 V blocks, which are added to produce the desired voltage. Such an arrangement may not be as simple as generating supplies for an op-amp, but it's still easier and cheaper than procuring a high-voltage transformer.

Fig 5 shows a +1700 V, 200 μA supply. The amplifier uses MOSFETs (MTP1N45) instead of BJTs. The turn-off and turn-on delays are much shorter for MOSFETs; they are determined primarily by the current that we can apply to charge the internal gate capacitance. External delays are not needed. Two N-channel MOSFETs are used instead of complimentary bipolar transistors. Therefore, the gate drives for the MOSFETs are inverted. The diodes used in the rectifier must have a PIV greater than 300 V. Schottky diodes aren't available at these voltages, but ultrafast silicon diodes such as Motorola MUR440 (400 PIV) or MUR450 (500 PIV) have acceptable reverse recovery times. Filter-capacitor values of 1 nF (0.001 μF) will produce a peak-to-peak ripple of 1 V. Disc ceramics such as Sprague's 5GAD10 (0.001 μF, 1000 V) are an economical choice. Five doublers develop +1500 V.

The voltage multiplier is a versatile circuit to have in your bag of tricks. You can "roll your own" voltage multiplier for a special project without spending a lot of time and money on special parts. A word of caution: If you have grown up with low-voltage solid-state electronics, you are probably casual about the touching the supply voltages, but voltage multiplier can develop voltages that can "bite." When the voltages are 48 V or more, keep one hand in your pocket. The multiplier may not be able to provide lethal currents, but it can get your attention. □□

Editor's Sidebar

The voltage multiplier circuits shown in this article are just a small sample of the wide variety used. These circuits were originally used for high-voltage X-ray tubes early in this century and in the first particle accelerators in the 30s, so there is a lot of information around. The *ARRL Handbook* has a good discussion of voltage multipliers in the Power Supplies chapter. Some older references like the *Radiotron Designers Handbook* and the *Electronic Designers Handbook* (Landee, Davis and Albrecht) are out of print, but they can be found in used book stores and college libraries. They have useful discussions of voltage multipliers. Another great source of information is Jack Althouse's (K6NY) *QST* article (Oct 1971, pp 29-33).

In modern IC versions, some diodes have been replaced with switches to provide voltage division as well as multiplication. These ICs are often referred to as "charge pumps." Such ICs are available from many semiconductor manufacturers and are usually inexpensive.—*QEX Editor, Rudy Severns, N6LF*

Elevated Monopoles Vs Active-Mast Antennas

That whip antenna may not work as well as you think it does! Let's address some myths, look at some alternative designs and some 6 meter examples that you can easily scale for other bands.

By Grant Bingeman, KM5KG, PE

A class of vertically polarized antennas, commonly called *monopoles,* consists of a single vertical element erected over an image plane, ground plane or counterpoise. This ground plane usually consists of several horizontal, elevated wires arranged radially around the base of the vertical element. Typical medium-wave AM broadcast radio towers are monopoles operating over buried radial ground systems consisting of about 120 $\lambda/4$ wires. Because propagation above 30 MHz is often line-of-sight only, a monopole antenna is usually mounted high above the earth, so it takes its ground system up into the air with it. We can call this configuration an elevated monopole (Figure 1).

4212 S. Buckner Blvd
Dallas, TX 75227
e-mail **DrBingo@compuserve.com**

Some of the monopole questions we will address in this article include:

1. Do the transmission line and mast affect antenna performance? Yes!
2. Is the monopole as good as a dipole? No!
3. Must we give up gain to obtain a good impedance match?
4. Is there always a trade-off between gain and bandwidth?

Contrary to popular belief, the gain of an elevated dipole is often better than that of an elevated monopole. A dipole also weighs less, has less wind resistance, less parts, etc. So why bother with elevated monopoles at all? This is a good question when we consider that there is a way to obtain dipole performance without an artificial ground or a balun. This is accomplished by integrating the mast into the electrical design of the antenna. The rest of this paper will describe the development of such an antenna for the 6-meter amateur band.

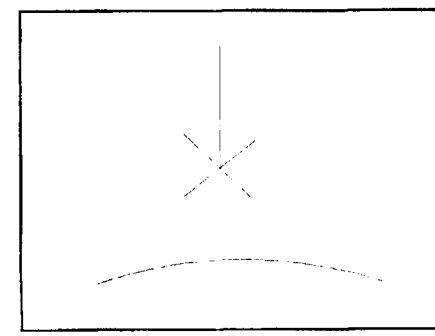


Figure 1 elevated monopole

First, let's consider a 5/4 λ dipole about 10 meters above the ground at its center feed point (when each arm is 5/8 wave long we obtain maximum

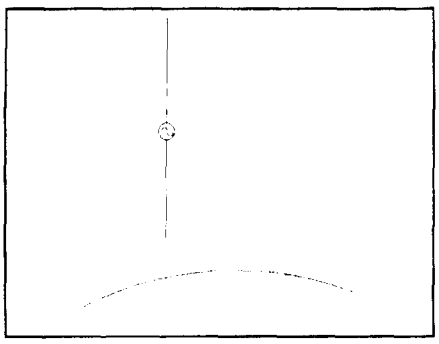


Figure 2 5/4 wave dipole

gain from a dipole, Figure 2). Assume we are using 1 inch, type M copper pipe, which has a radius of 0.56 inches, and we are operating at 52 MHz. The free-space wavelength is 5.77 meters, so a good length for each dipole arm is about 3.5 meters. When we suspend the center of this vertical dipole 10 meters above a perfect earth and radiate 1000 W, our vertically polarized, loss-less electric field at one kilometer is about 608 mV/m RMS. (This is much higher than for the same antenna in free-space, by the way.) The 52 MHz input impedance of this dipole is 139 $-j418$ Ω (see Table 1). The current distribution on the dipole is described by Figure 3 and Table 2. Note that the area under the current magnitude curve is directly related to the radiated field intensity. However, the areas marked with a (–) symbol are out of phase with those marked with a (+) symbol, so the negative area must be subtracted from the positive area.

I used *NEC3* moment-method analysis to obtain the impedances, currents and radiated field data for this article. All wires were modeled with four current segments per meter at 52 MHz, thereby keeping each segment less than 0.1 λ long. When modeling transmission line stubs where two parallel wires are electrically close together, more segments (current elements) are required, especially if wires of different radii are used. The earth was modeled as a perfect loss-less reflector, just to keep things simple. Actual field intensities that occur in practice will be quite a bit lower than the numbers quoted in this article, but the loss-less numbers are okay for comparing various antenna designs. Keep in mind that the power delivered to the antenna throughout this article is 1000 W, regardless of the antenna feed-point impedance. In practice, this would require a matching network of some kind.

Now let's consider the same 5/8 λ (3.5 meter) vertical element over an

Figure 3, 5/4 Wave (7.0 meter) Vertical Dipole Current Distribution
52 MHz, 1000 watts

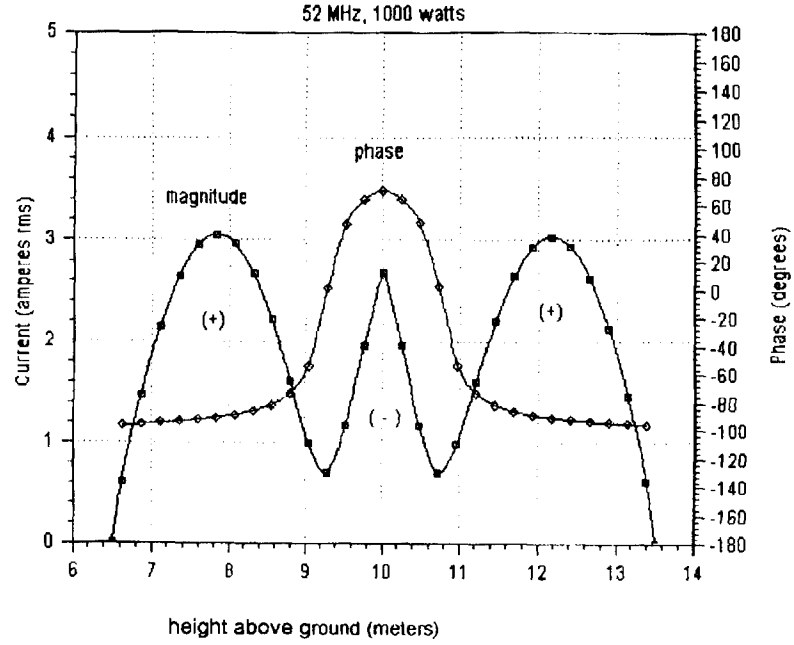


Figure 4, 3.5 Meter Monopole Input Impedance and Field Intensity
52 MHz, 1000 watts, vertically polarized electric field at 1 km

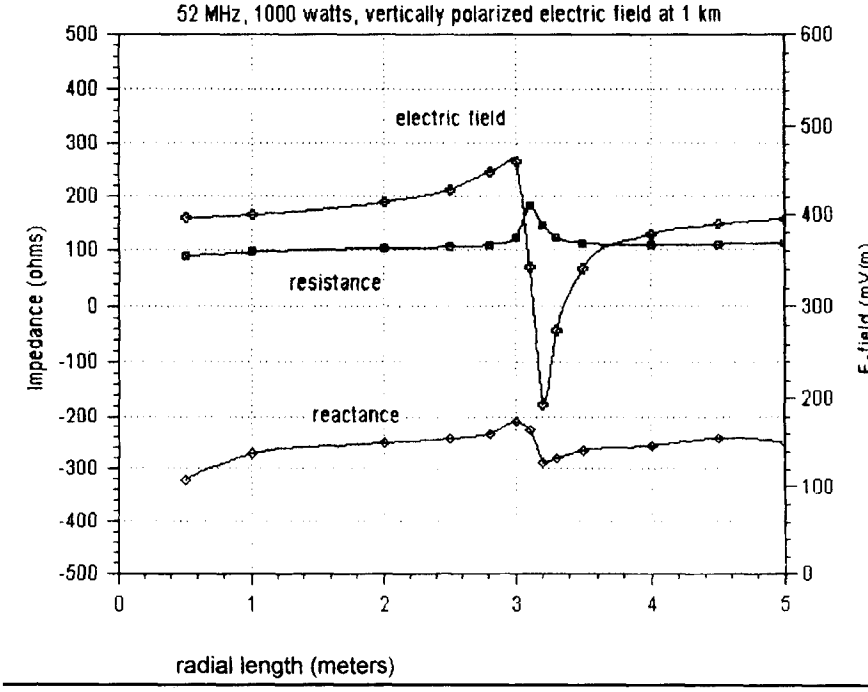


Table 1
5/4 λ Dipole Bandwidth

Freq (MHz)	*Input Impedance (Ω)*	*1 kW E Field (mV/m at 1 km)*	*Resonated† (Ω)*	*SWR*
50	191 –*j* 518	603	191 –*j* 116	2.14
51	162 –*j* 467	606	162 –*j* 57	1.50
52	139 –*j* 418	608	139 +*j* 0	1.00
53	122 –*j* 373	606	122 +*j* 53	1.53
54	108 –*j* 329	602	108 +*j* 105	2.38

†Series reactance has been tuned out by a series component.

image plane consisting of four horizontal radial elements, 10 meters above the ground. How long should these horizontal elements be? As you can see in Table 3 and Figure 4, the input impedance and gain of this monopole are very sensitive to the lengths of the ground radials. What, exactly, is going on? Apparently, when the radial length is near $\lambda/2$ (about 3.2 meters), the current distribution on the vertical element is disturbed (Table 4,Figure 5). Nonetheless, even the best performance of the monopole-and-ground-plane configuration falls 2.4 dB short of the gain we had with the dipole arrangement (20 log [608/460] = 2.4 dB).

Because of weight and cost considerations, elevated monopole ground radials do not usually extend far enough to encounter this half-wave resonance disturbance. However, there is a gain advantage if we extend the radials just short of this condition. The cantilevered horizontal radials can be supported with plastic guy lines from the top of the whip if necessary.

Table 2—5/4 λ Dipole Current Distribution

Height (m)	Current Magnitude (A RMS)	Current Phase (°)
6.5, 13.5	0.0	–96.8
6.7, 13.3	0.61	–95.7
6.9, 13.1	1.48	–94.8
7.1, 12.9	2.15	–93.9
7.3, 12.7	2.65	–92.8
7.6, 12.4	2.96	–91.6
7.8, 12.2	3.06	–90.1
8.1, 11.9	2.96	–88.1
8.3, 11.7	2.67	–85.3
8.6, 11.4	2.20	–80.9
8.8, 11.2	1.61	–72.8
9.0, 11.0	0.99	–52.9
9.3, 10.7	0.70	2.8
9.5, 10.5	1.19	48.9
9.8, 10.2	1.97	65.1
10.0 (center)	2.70	71.8

Table 3—5/8 λ Elevated Monopole Performance vs Ground-Plane Radial Length

Radial Length (m)	Input Impedance (Ω)	1 kW E Field (mV/m at 1 km)
0.5	89 –*j* 322	396
1.0	97 –*j* 273	400
2.0	104 –*j* 252	414
2.5	106 –*j* 244	428
2.8	110 –*j* 234	448
3.0	125 –*j* 210	460
3.1	184 –*j* 227	342
3.2	148 –*j* 289	194
3.3	123 –*j* 280	274
3.5	113 –*j* 266	341
4.0	111 –*j* 257	378
4.5	111 –*j* 243	389
5.0	113 –*j* 251	395

So why do we see so many whip antennas above a skirt of horizontal radials and so few vertical dipoles when the dipole maximum gain is better? Probably this evolved because of mounting considerations. Certainly, it is a bigger deal at lower frequencies where the physical dimensions of the antenna are bigger. A dipole requires twice as much vertical space as a monopole. Feed-line routing can also be more of a problem with a dipole, and often a horizontal boom is required to space the dipole away from the main support structure. If this boom or its support is metal, it

Figure 5, 3.5 Meter Monopole Current Distribution

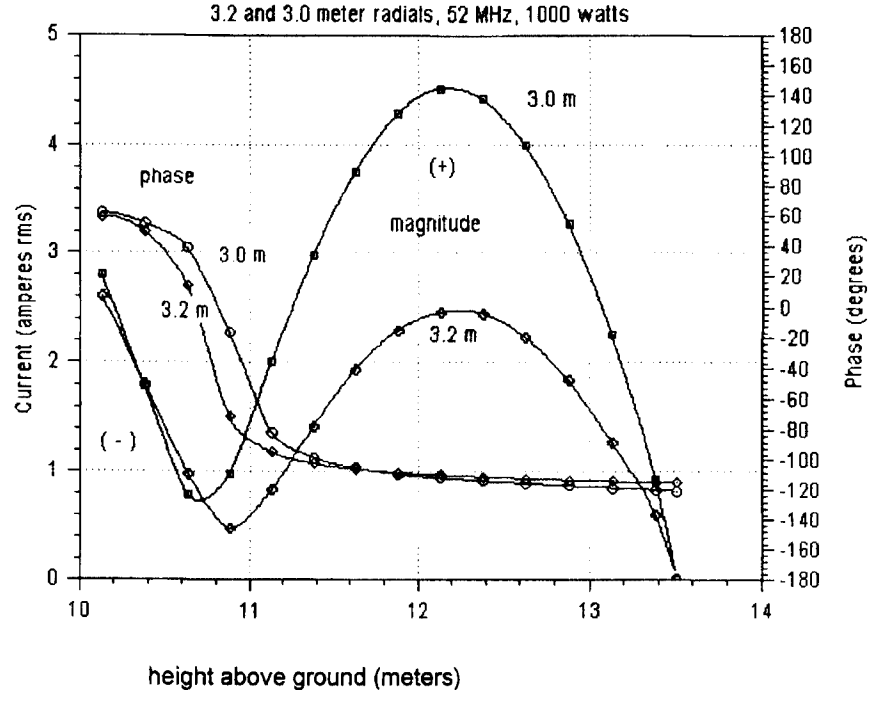
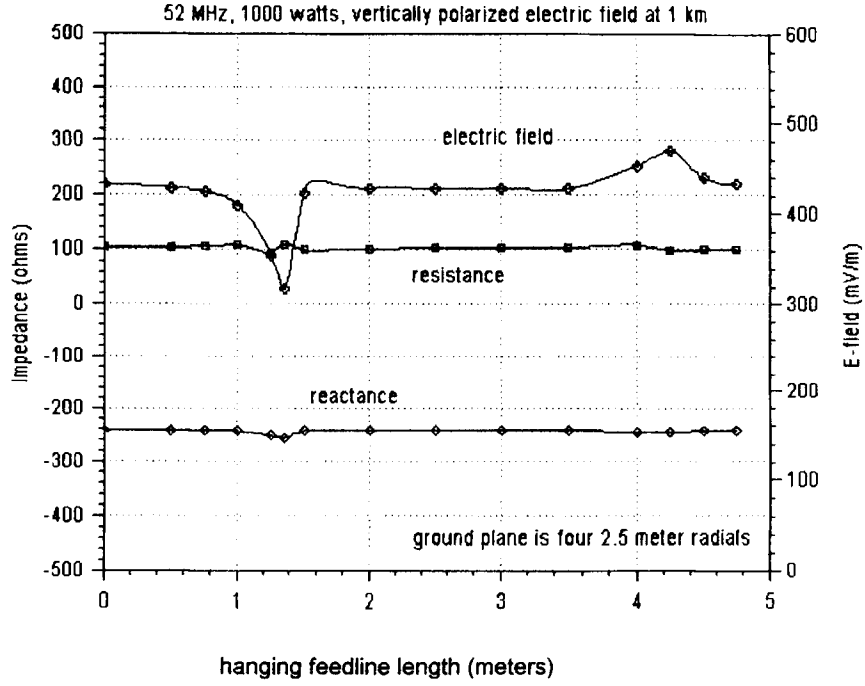


Figure 6, 3.5 Meter Monopole Input Impedance and Field Intensity



throws another twist into the equation.

However, a dipole requires a balun or choke of some sort in order to isolate the transmission line from the antenna. It is a mistake, however, to assume that you don't need to worry about this when feeding a monopole. You cannot simply connect a coaxial transmission line—shield and center conductor to the radials and whip, respectively—to an elevated monopole and expect the antenna to perform as if the transmission line were invisible. If in fact the feed-line or support pole becomes a radiating part of the antenna, we need to know; it can affect impedance, gain and radiated pattern shape.

Consider the case of a 3.5 meter whip above four 2.5 meter radials. If we attach a piece of coaxial transmission line (radius 5.0 mm) to our artificial ground plane, and allow this coax to hang vertically directly below the whip, conventional wisdom says that the length of this transmission line should have no affect on gain or impedance because it is shielded from the whip by the ground plane. In point of fact, the following variations occur (Table 5, Figure 6). This suggests that it might be a good idea to wind a few turns of the coax around the mast in order to form a choke, which would help to isolate the antenna from its feedline:

There is an obvious resonance condition that disturbs the monopole performance when the length of the coaxial line hanging below the antenna is near a quarter wavelength multiple. It appears, however, that we don't need to decouple the support pole or transmission line from the elevated monopole, so long as it's situated directly beneath the whip and it is not an odd multiple of a quarter wavelengths long. Notice that we are only modeling the outer conductor of the coax and still feeding this antenna at the base of the whip. We're not considering mismatch and subsequent standing waves inside the

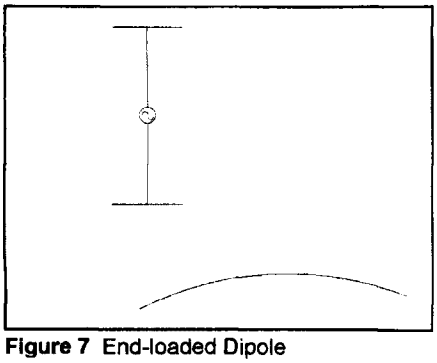


Figure 7 End-loaded Dipole

Table 4—Current Distribution on a 5/8 λ Monopole Over Four Radials

	3.0 m Radials		*3.2 m Radials*	
Height (m)	*Magnitude (A RMS)*	*Phase (∞)*	*Magnitude (A RMS)*	*Phase (∞)*
13.5	0.00	–115.0	0.00	–121.0
13.4	0.94	–115.2	0.60	–119.9
13.1	2.25	–114.4	1.27	–118.7
12.9	3.26	–113.5	1.83	–117.4
12.6	3.99	–112.6	2.22	–115.9
12.4	4.41	–111.5	2.44	–114.2
12.1	4.51	–110.2	2.45	–112.2
11.9	4.28	–108.6	2.28	–109.5
11.6	3.75	–106.2	1.93	–105.5
11.4	2.97	–102.6	1.42	–98.7
11.1	1.99	–95.3	0.83	–81.9
10.9	0.98	–7.2	0.47	–17.1
10.6	0.78	14.2	0.98	39.2
10.4	1.78	49.4	1.78	55.2
10.1	2.80	59.1	2.60	62.8

Table 4A—3.5 Meter Whip Over Four 2.8 Meter Radials, Bandwidth

Freq (MHz)	*Input Impedance (Ω)*	*1 kW E field (mV/m at 1 km)*	*Resonated† (Ω)*	*SWR*
50	151 –*j* 323	426	151 –*j* 98	2.23
51	128 –*j* 278	435	128 –*j* 49	1.55
52	110 –*j* 234	448	110 +*j* 0	1.00
53	97 –*j* 188	465	97 +*j* 51	1.65
54	91 –*j* 135	480	91 +*j* 108	2.85

†Series reactance has been tuned out by a series component.

Table 5—3.5 Meter Whip Over Four 2.5 Meter Radials and Transmission Line

Line Length (m)	*Input Impedance (Ω)*	*1 kW E Field Intensity (mV/m at 1 km)*
0.0	105 –*j* 243	430
0.5	105 –*j* 243	428
0.75	106 –*j* 244	424
1.0	109 –*j* 245	409
1.25	96 –*j* 253	354
1.35	110 –*j* 257	316
1.5	100 –*j* 245	423
2.0	102 –*j* 244	428
2.5	103 –*j* 243	428
3.0	103 –*j* 243	427
3.5	104 –*j* 243	427
4.0	108 –*j* 247	452
4.25	99 –*j* 248	469
4.5	100 –*j* 244	439
4.75	102 –*j* 244	433

Table 6—5.0 Meter End-Loaded Dipole Performance

Loading Length (m)	*Input Impedance (Ω)*	*1 kW E Field Intensity (mV/m at 1 km)*
0.0	1310 +*j* 217	507
0.5	1470 –*j* 769	536
1.0	343 –*j* 771	560
1.5	138 –*j* 577	578
2.0	65 –*j* 446	587
2.5	33 –*j* 353	578
3.0	18 –*j* 277	516
3.5	13 –*j* 208	328

Table 6A—5.0 Meter Dipole with 2.0 Meter End-Loading, Bandwidth

Freq (MHz)	*Input Impedance (Ω)*	*1 kW E Field (mV/m at 1 km)*	*Resonated† (Ω)*	*SWR*
50	98 –*j*546	573	98 –*j* 117	4.08
51	80 –*j*494	580	80 –*j* 57	2.21
52	65 –*j*446	587	65 +*j* 0	1.00
53	53 –*j*402	592	53 +*j* 53	2.43
54	44 –*j*361	595	44 +*j* 102	5.63

†Series reactance has been tuned out by a series component.

coax at this time. Don't forget that the whip must be insulated from the ground plane radials.

We can reduce the height of our antenna if we top load it. That is, we can use a shorter vertical element, and add some horizontal elements at the end to increase the current on the vertical element. The easiest way to do this is to add a tee at each end of the dipole (Figure 7). There may also be an impedance advantage with top loading. We can sometimes get 50 Ω of resistance simply by selecting the right amount of top loading, and still obtain almost the same maximum gain compared to a straight vertical configuration. Table 6 and Figure 8 predict that about 2.2 meters of end loading on a 5.0 meter dipole will produce maximum gain and about 50 Ω of input resistance. The price we pay, however, is decreased impedance bandwidth (Table 6A). The current distribution in Table 7 and Figure 9 shows an increase in current compared to the 7.0 meter dipole. This increased current compensates for the shorter length of the end loaded dipole, so the area under the curves (hence the radiated fields) are about the same from either configuration.

Now let's take the best top-loaded

Table 7—Current Distribution on 5.0 Meter Dipole End Loaded with 2.0 Meter Tees

Height (m)	Current Magnitude (A RMS)	Current Phase (°)
7.5, 12.5	0.0	–87.1
7.6, 12.4	4.78	–86.90
7.8, 12.1	4.78	–86.6
8.0, 11.9	4.53	–85.9
8.3, 11.7	4.01	–84.8
8.6, 11.4	3.24	–82.9
8.8, 11.2	2.28	–78.8
9.0, 11.0	1.21	–66.3
9.3, 10.7	0.53	13.7
9.5, 10.5	1.51	68.8
9.8, 10.2	2.79	78.4
10.0	3.92	81.7

Table 8—2.5 Meter Monopole Top-Loaded with 2.0 Meter Tee

Radial Length (m)	Input Impedance (Ω)	1 kW E Field Intensity (mV/m at 1 km)
0.5	89 –*j* 322	396
1.0	97 –*j* 273	400
2.0	104 –*j* 252	414
2.5	106 –*j* 244	428
2.8	110 –*j* 234	448
3.0	125 –*j* 210	460
3.1	184 –*j* 227	342
3.2	148 –*j* 289	194
3.3	123 –*j* 280	274
3.5	113 –*j* 266	341
4.0	111 –*j* 257	378
4.5	111 –*j* 243	389
5.0	113 –*j* 251	395

Table 9—2.5 Meter Monopole with 2.0 Meter Top Load Tee and Active Mast

Mast Length (m)	Input Impedance (Ω)	1 kW E Field Intensity (mV/m at 1 km)
0.17	44 –*j* 1440	365
0.50	33 –*j* 543	319
0.83	35 –*j* 384	263
1.17	38 –*j* 285	185
1.5	48 –*j* 202	100
1.83	73 –*j* 106	103
2.17	155 +*j* 42	225
2.5	534 +*j* 208	355
2.83	939 – *j* 656	461
3.17	234 –*j* 598	544
3.5	112 –*j* 450	587
3.83	75 –*j* 349	594
4.17	64 –*j* 263	555
4.5	71 –*j* 177	468
4.83	75 –*j* 40	330

Figure 8, 5.0 Meter End-Loaded Dipole Input Impedance and Field Intensity
52 MHz, 1000 watts, vertically polarized electric field at 1 km

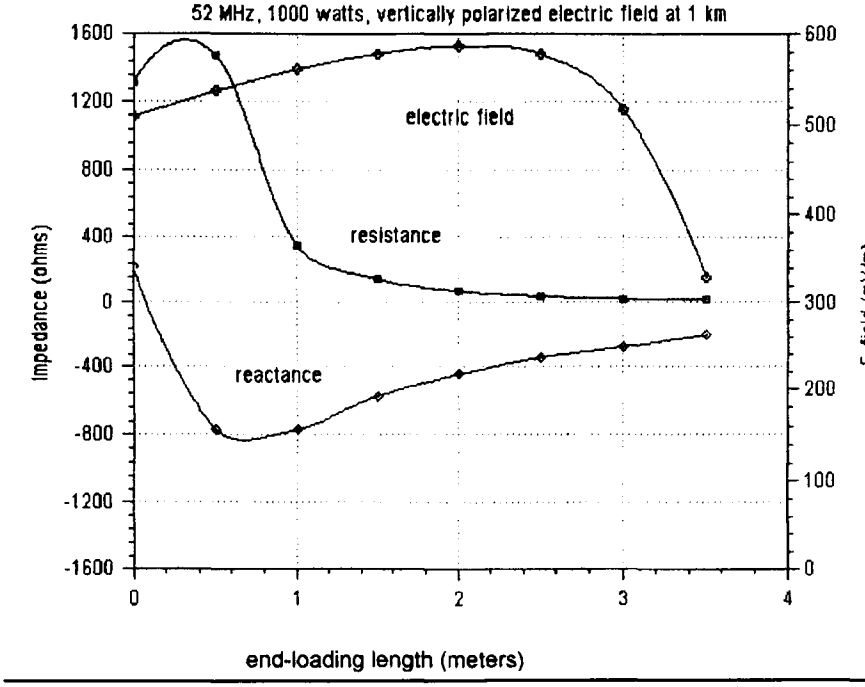
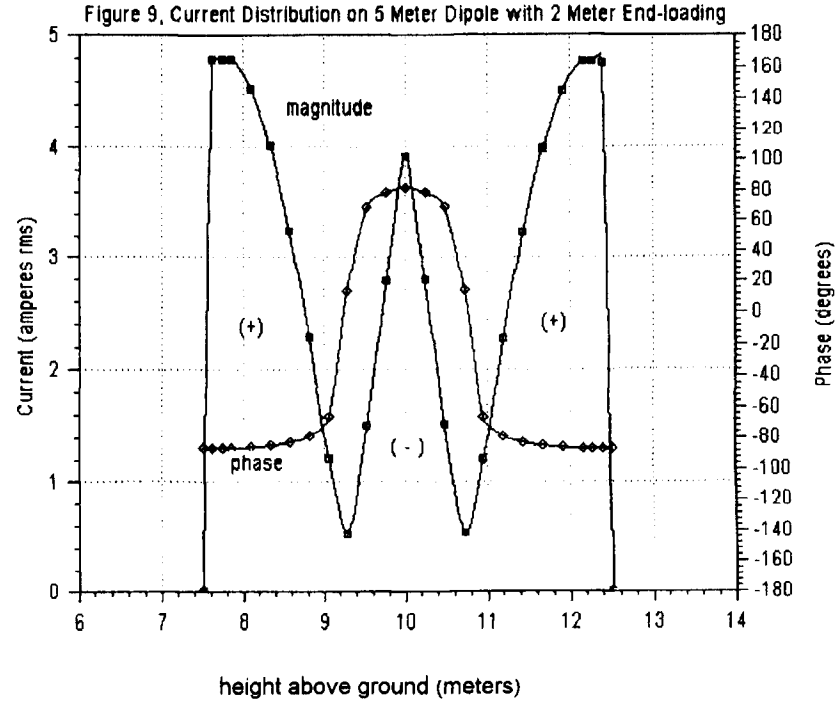


Figure 9, Current Distribution on 5 Meter Dipole with 2 Meter End-loading



configuration and apply it to our monopole configuration. Using a 2.5 meter vertical element and a single 2 meter tee, Table 8 and Figure 10 tell the story. Again, there is a region we need to avoid near the 3.2 meter ground radial length where performance is very poor.

Therefore, we have found that our best monopole just doesn't have very good gain compared to our best dipole. Let's look at what happens when we replace the ground plane with a conducting mast. Maybe we can create a practical dipole configuration if we choose the correct active-mast length. By eliminating the four horizontal radial arms, we also significantly reduce wind loading, weight, volume, etc. If we can make the mast behave as if it were the bottom half of a dipole, maybe we can retain the gain advantage that the dipole has over the monopole.

A mast length of 3.5 meters seems to produce the best gain (Table 9, Figures 11 and 12), which makes sense when we consider that this is about 5/8 λ. We could get similar gain with a shorter mast and some end loading at its base (say 2.5 meters loaded with 2 meters, same as at the top of the mast). However, this is probably an unnecessary complication. We also have to be careful about how we run the feed-line to this antenna, because unlike the monopole over a ground plane, our active mast is not isolated from the antenna. The mast is, indeed, a radiating part of the antenna, and any transmission line that hangs down from this mast will make the mast look longer. Therefore, we need to wind a few turns of the coaxial feed-line around the base of the mast to make a high-impedance RF choke that will isolate the line from the antenna. Of course, the mast must be insulated from ground also. A few feet of plastic pipe works well.

We can match the impedance of this antenna to the transmission line (50 Ω) by way of a stub: about one meter of ½ inch copper pipe (see Figure 13). We run the coax inside the mast to a point 2.5 meters from the top, and pull it out through a hole drilled in the side of the mast. The outer conductor is then attached to the mast, and the inner conductor run to the end of the stub where we have a coaxial coupling capacitor. This capacitor is simply a ¼×20 screw threaded into a piece of drilled and tapped plastic rod inserted in the end of the ½ inch pipe. You adjust the length of the stub, via a shorting strap, to get 50 Ω of resistance, then you adjust the capacitor to tune out about 500 Ω of reactance. It's good to start the shorting strap at 80 cm (31½ inches) from the end of the stub, with the screw started about two inches into the pipe.

In addition, stubs can have a lot of circulating current, hence power losses. Stubs tend to narrow the impedance bandwidth as well. Therefore, if you can obtain the desired 50 Ω simply by adjusting the amount of top loading, you may be better off to do so. In that case, you would tune out the reactance with a coil. Of course then you must worry about an insulator to separate the mast from the whip and also the I^2R power losses in the coil! Plus, the whip does not have a dc path to ground as with the stub configuration. That can lead to some large static charges on the center conductor of the coax.

I found the impedance bandwidth of

Figure 10, 2.5 Meter Top-loaded Monopole Input Impedance and Field Intensity
52 MHz, 1000 watts, vertically polarized electric field at 1 km

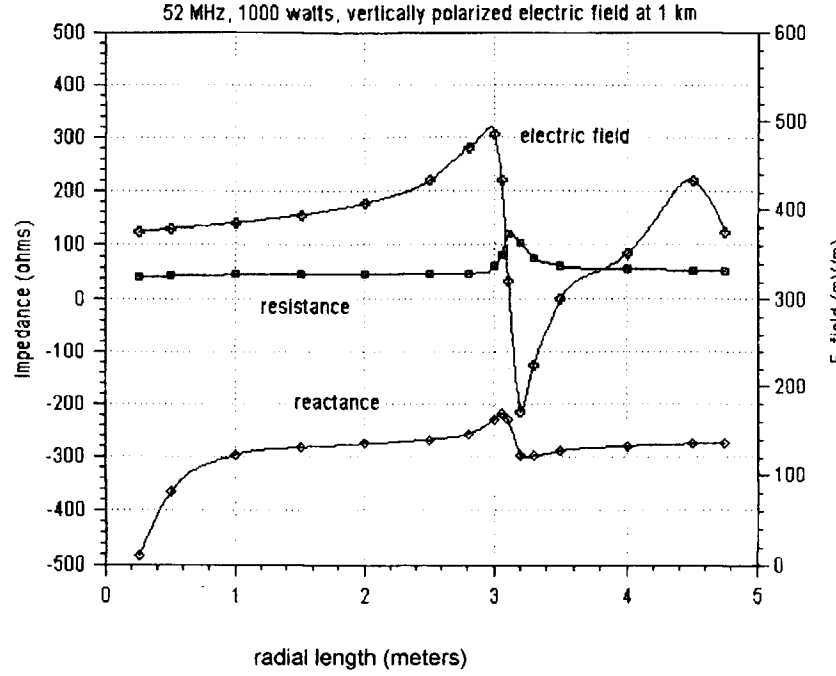
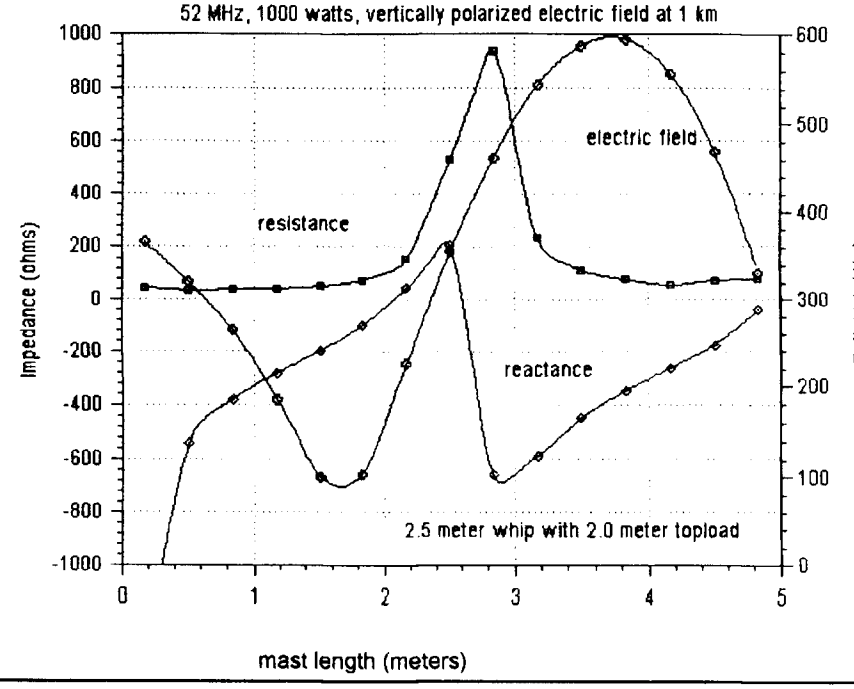


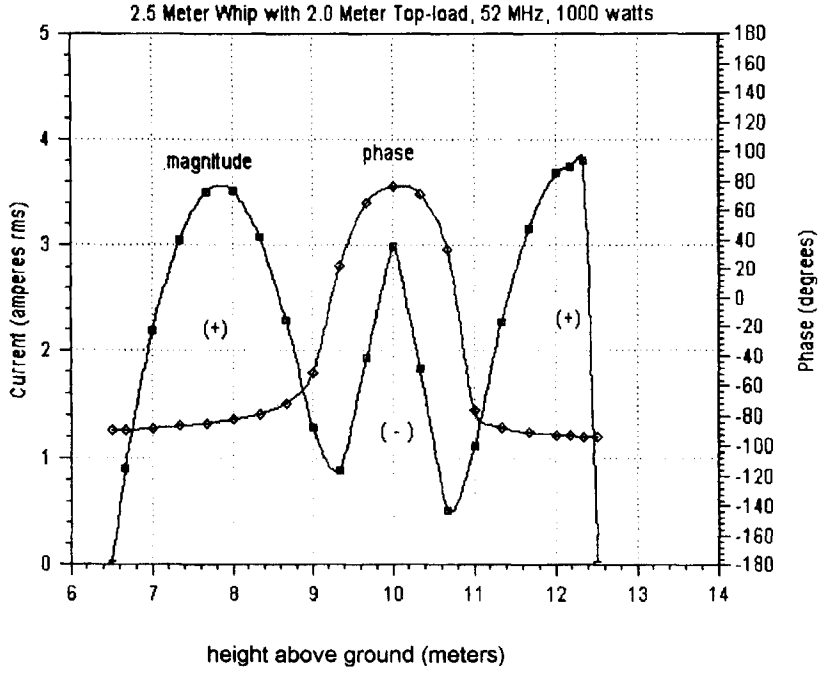
Figure 11, Top-loaded Active Mast Input Impedance and Field Intensity
52 MHz, 1000 watts, vertically polarized electric field at 1 km



this stub configuration to be rather narrow. It's okay for CW and FM, not so hot for SSB or AM. You will need an adjustable impedance matching network to operate over the full 6 meter band, unless you select a different feed arrangement. Keep in mind that it is dangerously easy to obtain a good impedance match at the expense of gain. For example, in Table 9 one might be tempted to use a series coil to tune out the reactance for the case of the 1.5 meter active mast length. That would yield an input impedance of 48 Ω, but the gain would be the absolute worst—about 17% of the optimum value. It might make a good dummy load, but it would be a lousy antenna! However, we can get a decent gain and match using a series coil at the feed point when we have about four meters of active mast.

The vertical radiation patterns of these antennas over real, lossy earth also tell an interesting story. I used Roy Lewallen's *EZNEC* 2.0, a ground conductivity of 5 mS/m and a ground dielectric constant of 15 to produce the following vertical pattern plots. Comparing the 5/4 λ dipole, 10 meters above ground at its center (Figure 15) to the 5/8 λ monopole (four 3 meter radials), we can see that there is still a big difference in gain (1.5 dB).

Figure 12, 3.5 Meter Active Mast Current Distribution
2.5 Meter Whip with 2.0 Meter Top-load, 52 MHz, 1000 watts



Summary

We have found that:

1. The maximum gain of an elevated monopole does not appear to be as good as that of a dipole.
2. The gain and input impedance of a monopole can be quite sensitive to the length of the ground radials.
3. The proximity of the mast and coaxial feedline beneath the ground radials influence monopole performance.
4. Top loading can be used as an additional handle on impedance matching without unduly compromising antenna gain, although there is some trade-off.

The ground plane can be eliminated

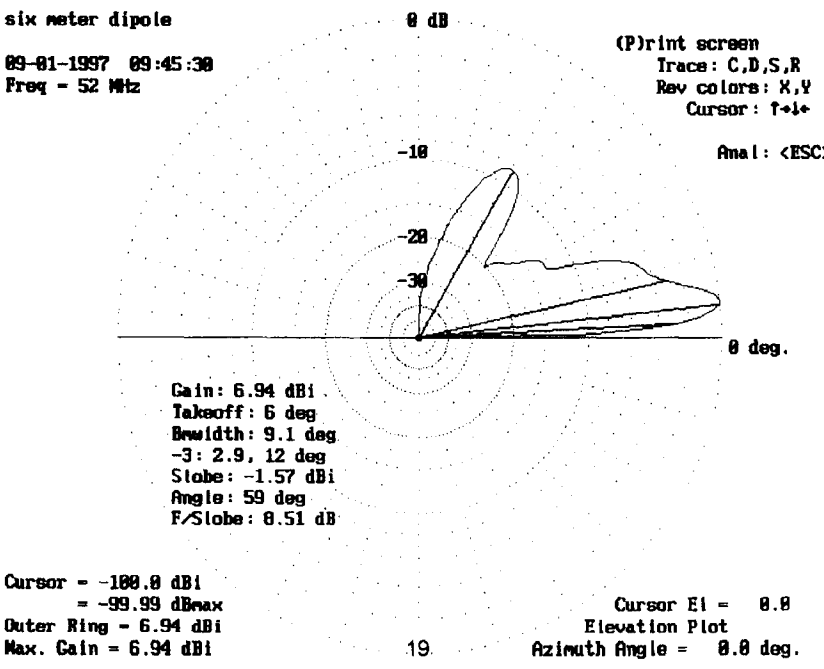


Figure 15 7.0 meter dipole 10 meters above ground

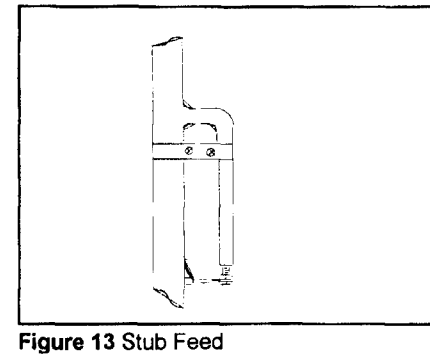


Figure 13 Stub Feed

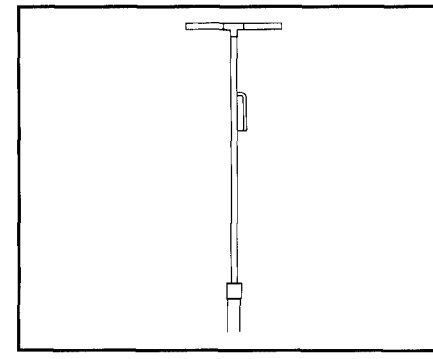


Figure 14 Overall antenna view

if the mast acts as the lower half of a dipole and a choke and insulator are installed at the base of the mast.

References

The following articles relate to these antennas and may be of interest.

1. Les Moxon, G6XN, "Ground Planes, Radial Systems and Asymmetric Dipoles," *ARRL Antenna Compendium* Vol 3, 1992.
2. Rudy Sevens, N6LF, "The Lazy-H Vertical," *Communications Quarterly*, Spring 1997, Vol 7, No. 2, pp 31-40.
3. Dick Weber, K5IU, "Optimal Elevated Radial Vertical Antennas," *Communications Quarterly*, Spring 1997, Vol 7, No. 2, pp 9-27.

Grant Bingeman has been a Senior Engineer at Continental Electronics in Dallas, Texas, for 17 years, where he designs high-power antennas and transmitters for frequencies from dc to light. His e-mail address is ***DrBingo@ compuserve.com****. His call signs from the 1960s were WN6AIW, WB6MBX and WB4AOI. After a 30 year hiatus from Amateur Radio, he has returned to the fold as KM5KG.*

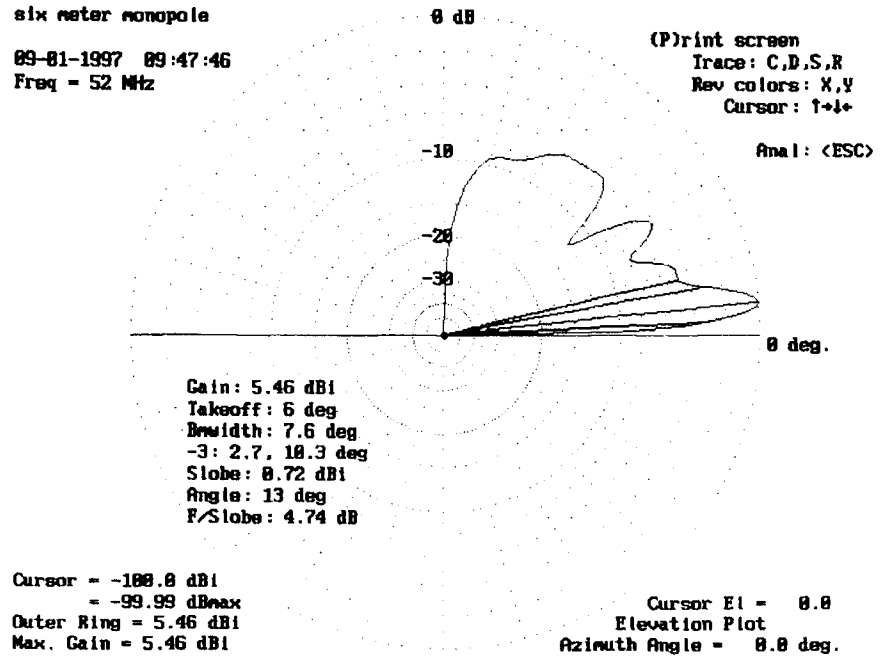


Figure 16 3.5 meter monopole with four 3-meter radials 10 meters above ground

How the Ionosphere REALLY Works

Current research approaches explanations of some unusual propagation behaviors.

By Eric P. Nichols, KL7AJ

Introduction

From the earliest beginnings of Amateur Radio, hams have been fascinated with the unknown, or at least the unpredictable. To plagiarize and paraphrase a well-known quotation about the weather, "Everyone talks about radio propagation, but nobody does anything about it."

Certain recent and not-so-recent developments in the expanding field of ionospheric research have given hope that we may, to some degree, have some say in how the ionosphere behaves. At the very least, even if we can't turn on the sky at will, we are learning more about the conditions that make it "do what it do."

PO Box 56235
North Pole, AK 99701
e-mail **enichols@gci.com**

A Catalogue of Aberrant Behaviors

Since the days of Heaviside, we have had a vague idea about this thing called an Ionosphere, and most of us have had to spew forth at least a semblance of understanding of propagation in order to pass our amateur exams. Most of the traditional models are, of necessity, oversimplified. Although they explain a great deal of our day-to-day and year-to-year operating conditions, they fail to shed much light on the weird phenomena that many us have experienced. (Or at least *claim* to have experienced.)

New ionospheric probing techniques have revealed a great deal in recent years. There are now plausible explanations for all the following behaviors that theoretically "can't be":

"One-way" skip: By far the easiest to explain. *Yes*, not only does it exist, but it is the *predominant* mode of behavior in places like KL7 land, where the ionosphere is disturbed.

Long delayed echoes: No, these are *not* reflections from the Van Allen Belts or undiscovered planets in our solar system. Nor are they long-running hoaxes. They are purely ionospheric phenomena that can be consistently reproduced in plasma chambers, (with shorter time scales, of course). Yes, 20 second delays *are* possible in the Earth's ionosphere. These are elegant examples of *mode conversions.*

Frequency shifting: Not just tempo-

rary frequency pulling due to Doppler shift, this is a well-documented factor. We're talking about actual long-term frequency shifting of a pure carrier. Yes, you *can* receive WWV a few hertz off from where it should be, under the right conditions. This is a special case of the *Luxemberg effect*, requiring only one transmitter. It is caused by *parametric decay*, a fairly well understood plasma phenomenon.

"60 Hz" buzz coming from "nowhere": Actually this auroral buzz can vary from a few tens of hertz to a few hundred, but it still has a very "man-made" characteristic. This is still largely misunderstood, but it *is real* and it *is ionospheric* in origin.

Extreme selective fading: The primary reason why HF packet doesn't work. New ionosonde techniques (Doppler ionosondes, in particular) have shown us just *how* abrupt the critical frequency/MUF contour can be. Most of the time the ionosphere behaves like Jell-O, but sometimes it's more like an optical prism.

Just about anything else that seems "funny": More than likely, it's caused by the ionosphere—enough said.

Plasma Soup

What is a plasma? To start out with, let me say that I have yet to find a word that even *rhymes* with plasma. Perhaps this is fitting, because there isn't really anything that *acts* like one, either. (One of my colleagues has suggested *hasma*, as in "Who has ma plasma, I know I done had it around here somewhere!"—but that's a bit of a stretch.) One standard working definition of a plasma is: "A gas that has been ionized to the state where the individual species exhibit collective behavior."

Okay, how do we describe this collective behavior? I will begin with the assumption that everyone is familiar with the standard atomic model of matter: a heavy nucleus (protons, and perhaps neutrons) orbited by a lighter electron or two, etc. That is standard grade-school stuff. I also assume that you have a basic understanding of the nature of gases, how the density, pressure, temperature, etc, all relate. Most of this is pretty intuitive; it's this "collective behavior" that can really throw us for a loop. I will try to convey this principle with an analogy that was given to me by a not-to-be-named Elmer.

Let us imagine a large metropolitan area populated with nothing but dog lovers. Each darling little yapper is attached to its owner by a leash of a different length. Now, imagine every one of these residents is out on the sidewalk giving FiFi her (his) daily walk. Regardless of the length of the various leashes, each canine is associated with one and only one owner, even though, as viewed from above, the movement down Fifth Avenue might seem chaotic, if not totally random. Now, depending on the particular length of its particular leash, each poodle is allowed a different *degree of freedom*. Remember that term; it will be important. Up to now, the movement of each poodle has been more or less independent of the other poodles; they are each content to do their own thing, within the limits of their leashes.

Now, what happens when the local hot dog vendor pushes his cart by? Suddenly, each and every poodle within a block takes note, and tugs in the same direction. In other words, the poodles are exhibiting a collective behavior, which is far more compelling than their previous, semi-random wanderings.

What does this all have to do with ionospheric propagation? Let's start by saying that dog owners are ions, poodles are electrons, and the lengths of the leashes are (sort of) electron mobility, which is closely related to "electron temperature." Let's tie this in with something that should be familiar to most of us who have been hamming for any length of time. Most of you have seen that funny little curve called the *electron density profile*. It's been published in every radio text since father Abraham first clamped a Texas Bugcatcher onto the rear end of a roving camel on his way to the Promised Land.

Why is electron density so important? Because the electrons *alone* interact directly with radio waves. Now, this is not to say that ions are unimportant in radio propagation; indeed, they account for some of the more interesting phenomena encountered from time to time. However, ion (and neutral) interaction is always a secondary, or *derived* effect. It is electron density that tells us where the action is. To put it succinctly, if you've got a lot of electrons, you've got a lot of radio.

Lest we oversimplify things, the electron density profile should not be thought of strictly as a weather map. The electrons don't just sit up there in a cloud, like they do in a vacuum tube. The electrons are part of a complex soup of ions, neutrals and electrons of varying degrees of freedom (temperature).

Keep our owners and poodles in mind. Although our electron cloud exhibits a strong collective behavior, it is still restrained by the presence of its associated ions, which are slow moving, massive, and more "gas like" in behavior. Even if our dog owners should let go of the leashes, the stray dogs would still have to wend their ways through a lot of legs to get to the hot dog vendor, or any other collectively interesting target.

Now, perhaps the obvious question at this time is: "Just how many electrons do we *need* to collectively reflect a radio wave? Is it possible to reflect a radio wave off a single electron? The answer, surprisingly, is *yes*. As a matter of fact there is a type of diagnostic radar called Incoherent Scatter Radar (ISR), which relies on the fact that individual electrons can reflect (scatter) radio waves. This is rather esoteric technology, however, and of little relevance to Amateur Radio for the time being. For any reasonable effect to be imparted to a radio signal in the HF region, we need something that resembles a large, relatively flat surface, a collective electron plasma, to be precise.

Fortunately, the natural ionospheric plasma cooperates with this requirement by kindly arranging itself in regions somewhat resembling *layers*. How nice of it. Lest we be tempted to oversimplify this situation, we must recognize that even when the electron density is at its greatest, there is an appallingly small number of the little critters up there. A typical electron density maximum occurs at an altitude of around 300 kilometers (during a really sweet F2 day). We're talking about *maybe* a million electrons per cubic centimeter—almost nothing. To put a perspective on this situation, sea level atmospheric air contains about a trillion times that number of particles. So, basically, we're bouncing our radio signals off next to nothing. This largely explains the fickle nature of the ionosphere. (*Bouncing* may be a misnomer; actually, our waves are being absorbed and reradiated.)

Now, if our electrons merely sat up there by themselves, like a cloud in a vacuum (they *don't*, as we've established), they would pretty much respond to radio waves of any frequency in much the same manner. As long as there was a layer of collective electrons significantly larger than a half wavelength or so in any dimension, they would be expected to act like any reflector, sort of like an inverted ground radial system. Well, we all know that such is not the case. The

ionosphere is *highly* frequency dependent.

This is where the *ions* come in to play. Remember that I said that "reflection" from the ionosphere is really a process of absorption and reradiation? What do we really mean when we say a radio wave is *absorbed?* As in any physical system, when we absorb energy, we usually think of converting energy from one form to another. In this particular case, our radio waves (electromagnetic energy) are converted into *kinetic energy* (electron motion). I hear a collective, "Say *what?*" Am I saying that when you send a signal skyward, you're actually moving electrons around 200 miles away? That's exactly what I'm saying. You can all feel very proud of yourselves for having an actual mechanical effect on the heavenlies when you send "CQ" on your QRP rig.

Now, if we simply let our electrons go their collectively merry way after kicking them with a little RF, that would be the end of it. Much as we would like to see otherwise, dog owners are still in the picture. Our stimulated poodles are going to reach the ends of their leashes. When they do, they're going to rebound to their original position. Upon rebounding to their owners, they release all that energy as *electromagnetic radiation. Voila!*

How Do We Know All This Stuff, Anyway?

We don't, but we're learning more all the time. There are numerous methods for exploring the ionosphere; and for some reason, Alaska seems to have more of the stuff than anywhere else. This is good job security for this particular KL7. However, as a rule, high-latitude ionospheric research is in a class by itself, conducted by HIPAS Observatory (High Power Aurora Stimulation Experiment), HAARP (High Frequency Active Aurora Research Project) and a few other similar outdoor laboratories scattered around the arctic. Arecibo in Puerto Rico is the exception, the only equatorial-region ionospheric heating facility.

Although I want to concentrate our attention on things that *every* ham can experience, we should be aware of some of the more interesting fringes of radio as well. Radio Amateurs have contributed a lot to ionospheric research and will continue to do so. Let's peek at some of the hardware for doing ionospheric exploration.

Ionosonde: This is the classic ionospheric measurement tool. It is nothing more than swept radar with a vertical-incident beam. You just send short pulses skyward and look at the reflection as you increase the frequency. Eventually, you reach the critical frequency or frequencies, above which you get no reflections, but you can learn much more than that. You can see, with varying degrees of clarity, all those layers the literature has taught you exist all along. *QST* and other publications have published ionograms over the years, and the Internet is full of locations from where you can download them—fascinating stuff for all you DXers.

Doppler ionosonde: This is a variation of the Ionosonde. First proposed by a Korean gentleman by the name of Valery Kim, a respected name in ionospheric studies. I consider myself privileged to have had a part in the development of this technique. As any ham who has had the merest taste of HF knows, radio propagation changes quickly. Ionograms are great at telling you where the layers are, but they take a minute or so to develop the data, so you can miss some interesting short-term stuff. The Doppler ionosonde not only tells us how high a certain layer is, but it can tell us exactly how fast it's moving upward or downward, and believe me, it can move *fast*. Arctic flutter is one common manifestation of this.

Skymapping (HF holography): This is a relatively new, undeveloped technique for giving us detailed pictures of small regions of the ionosphere in the north-south/east-west dimensions. Very interesting, because it reveals how sharply the ionosphere can tilt. It is anything but flat; it can have slopes on the order of 35° to 40° over certain regions. I hear a collective *aha!* out there as you all realize how nicely this explains one-way propagation. A radio signal going one direction might be above the critical angle in an east-to-west direction, at the same launch angle and below the critical angle going west-to-east.

Imaging Riometer: The Japanese have dumped huge amounts of money into this technology. Riometer is an acronym for Relative Ionospheric Opacity meter. It is a passive system (receive only) that does skymapping of the D and lower E regions, in the 30 to 40 MHz regions. It gives us a lot of information about the actual electron activity, and a bit about the chemistry up there.

Lidar: Vertical incident laser radar: Believe it or not, there are layers of metallic ions up there, eg Iron, Calcium and Sodium. All you meteor-burst folks already know this, but its effects might have more to do with HF propagation than previously thought. At any rate, we can use a high powered laser tuned to the atomic species of our choice, which will resonate (fluoresce) when excited. We can then look at the re-emitted photons from the atoms of choice. Excited atoms are rather high-Q circuits, so they can be very precisely selected. It's sort of like shooting a gun into a dark room and listening for whatever screams.

Magnetometers and Electric-Field Sensors: These are used to determine nonelectromagnetic effects in the ionosphere. There are all sorts of pseudo radio-like happenings in a plasma. *Weird* things. Things like Ion Acoustic Waves, Plasma Surface Acoustic Waves, Alfven Waves, Whistlers, and conversions from these to electromagnetic waves and vice versa. This is the most probable explanation for long-delayed echoes.

In Situ *Measurements*: Sometimes, measurements from a distance just aren't good enough. Any or all of the aforementioned equipment and techniques can be stuck on a rocket and launched through the ionosphere. Obviously, this is expensive; but again, the Japanese think it's a very worthwhile endeavor. (So do we, to some extent.) The first of the really neat space-based ionospheric research was the downward looking ionosonde (Topside Sounder), which gave us the first look at the electron density profile *above* the peak density.

Try This at Home

Can an average ham on a shoestring budget contribute to this fascinating research? You can contribute more than you might think. Ionospheric research, like any research, benefits from vast quantities of data, from as many sources as possible. Next time you're on the air, see if the conditions you're seeing match the propagation forecasts. If you see or hear something strange or new, write it down. Keep records of *everything*, and compare notes with other hams. Build your own Riometer. All it takes is a dipole, a 10 meter receiver and a chart recorder or a PC data acquisition program. Measure the Doppler shift of WWV. All it takes is a receiver, an XY oscilloscope and a crystal time base. Set up an HF holography club. With GPS receivers and imaging software, it's a snap.

A Few Loose Ends

This field is vast and deep. I have only been able to present a tantalizing overview of the subject. If there is sufficient interest, I would like to follow a few of the more interesting tangents in greater detail. Among other thing—if there appears to be enough demand—I would like to present the latest, most reproducible theory on long-delayed echoes. That, alone, would require several articles.

In closing, I would like to publicly thank all the fine people at HIPAS Observatory, especially Ralph Weurker, a scientist's scientist; Helio Zwi, now with HAARP, a gentleman and a scholar, and finally; UCLA grad student Jacqueline "Wacky-Jackie" Pau, who tirelessly followed me around the tundra in snow up to her armpits (or my waist, depending on one's point of view) as we set up countless antennas in the field. □□

Extending the Double-Tuned Circuit to Three Resonators

Here's an empirical design method for triple-tuned filters.

By Wes Hayward, W7ZOI

Although the most common LC band-pass filters used by radio amateurs are single and double-tuned circuits, there are situations where a higher-order filter is desired. The triple-tuned circuit, a band-pass filter with three resonators, is an especially useful example. It is easily designed and offers improved stop-band attenuation with little increase in insertion loss. This paper presents two simple methods that allow the VHF triple-tuned circuit to be adjusted and tuned with simple instruments and little more complication that a double-tuned filter.

7700 SW Danielle Ave
Beaverton, OR 97008
e-mail **w7zoi@teleport.com**

The Double-Tuned Circuit as a Step toward the Triple-Resonator Filter

Two filter schematics are shown in Fig 1. Both were designed for a Butterworth response and a center frequency of 110 MHz with a 3-dB bandwidth of 2 MHz. The N = 3 (three pole) filter has slightly higher insertion loss than the double-tuned circuit, but is more selective at high attenuation. The reader can calculate the responses with any of dozens of available computer programs.[1]

It is interesting to compare the two schematics. The coupling capacitors are identical for the same bandwidth. However, the end-section matching is slightly different, with the end Q being higher for the double-tuned circuit (DTC) than for the triple-tuned circuit (TTC). This leads us to ask what the performance would be for a DTC that has the same end designs as a TTC. This is shown schematically in Fig 2 with responses shown in Fig 3.

[1]Notes appear on page 46.

Experimental Methods

The introductory example presented above used computer generated and analyzed designs. This is all the preparation that is needed for lower-frequency filters where discrete reactances are large. This works well for filters up through 30 to 50 MHz with capacitor values of 1 pF or more and inductors of a couple hundred nanohenries or more. As we move into the VHF area, and higher, the components are not described well by

"printed" values. Stray reactances begin to dominate, encouraging us to use experimental methods.

I described the DTC in detail in a previous *QST* paper.[2] Experimental methods were emphasized in that tutorial, which described a method for building a double-tuned circuit without ever going through the numbers. The method entails adjusting the end loading on both resonators while also adjusting coupling. The resonators are tuned for maximum response after K and Q adjustment; the insertion loss is then measured. The reader is urged to study that paper if he or she is not very familiar with the procedure.

Perhaps the most important detail presented in the DTC tutorial was the need to perform a wide-band sweep to locate any double-humped response that might be present. A common and potentially disastrous error that the experimenter can make with the DTC is severe over coupling that produces two widely separated response peaks. It is easy to miss one of the two peaks if a wide sweep is not done.

With this background in mind, we can use the earlier observations to implement a triple-tuned circuit. The three-element filter is built with the middle resonator eliminated. The loading is kept identical at the two ends and is adjusted for a desired filter bandwidth. Coupling is set up between the first and *third* resonator and the combination is adjusted as a DTC. A wide-band sweep is done to find the double humped response, if present. It is then eliminated through further adjustment, if found. This is

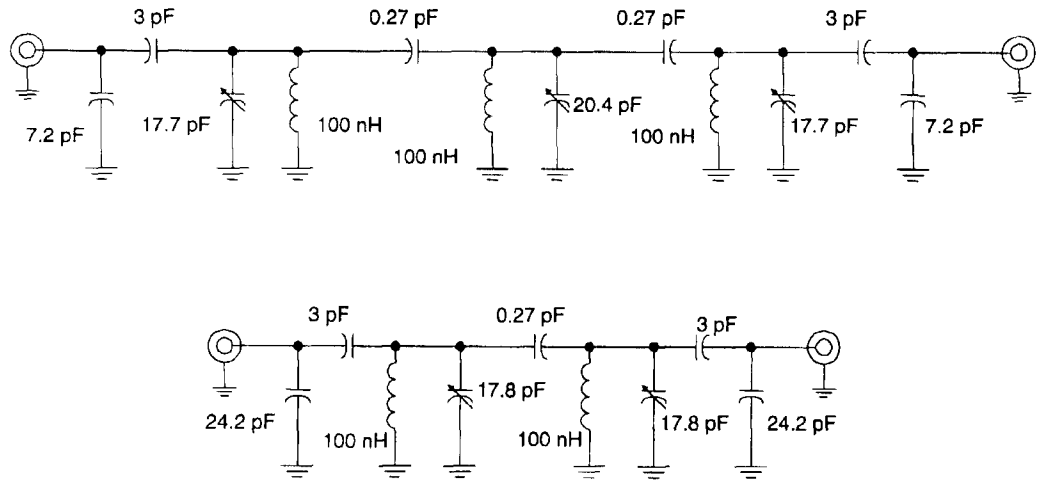


Fig 1—Butterworth LC Band-pass Filters with a 2 MHz bandwidth, centered at 110 MHz. The upper schematic is a triple-tuned circuit (N = 3), while the lower schematic is a double-tuned circuit.

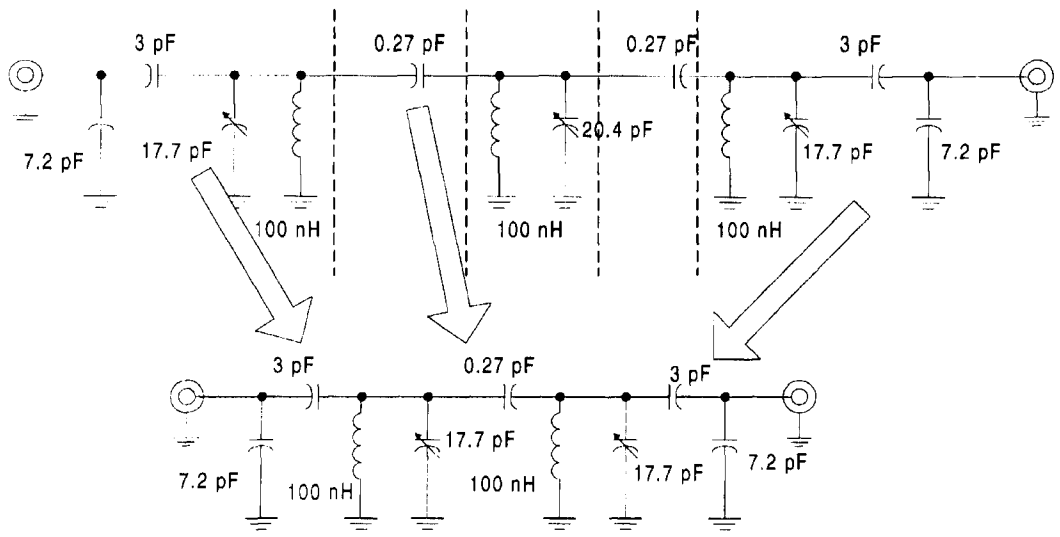


Fig 2—A TTC is designed "on paper" and is mentally segmented into end resonators, a middle resonator, and coupling elements. A DTC is then fabricated consisting of the two end resonators and a coupling element. The response of the DTC, and the parent TTC are shown in Fig 3.

repeated until the desired bandwidth is achieved in a DTC.

With a working DTC now in hand, the coupling is examined and duplicated as the third, middle resonator is added. The coupling from resonator 1 to 2 should be the same as that from 2 to 3. The TTC is then tuned and measured, performing a wide-band sweep to confirm the absence of extra peaks. The final bandwidth should be close to that of the intermediate DTC with a slightly higher insertion loss.

The filter shown in Photos A and B was built for use as the first IF of a spectrum analyzer. Filter design began with selection of an inductor. Values around 100 nH are practical at 110 MHz. The inductor was built with a 6.0 inch piece of #18 enameled wire. The ends were stripped and five turns of the wire were wound on the shank of a 1/4 inch drill bit.[3] The unloaded Q was assumed to be around 200, a value later confirmed with measurements. This inductor resonates at 110 MHz with a capacitor of about 21 pF, realized in the filter with combinations of fixed and glass trimmers. The predicted coupling capacitor value is 0.27 pF for each of the two components. A computer design is useful to provide guidance during construction, even if some components are less than practical.

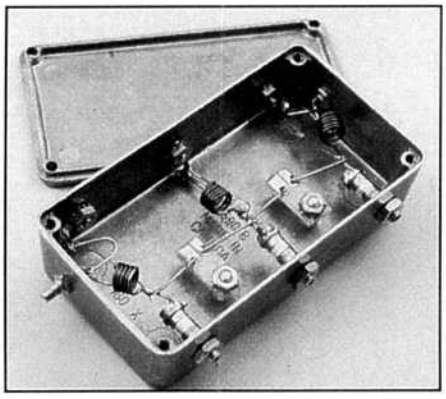


Photo A—A TTC filter for 110 MHz.

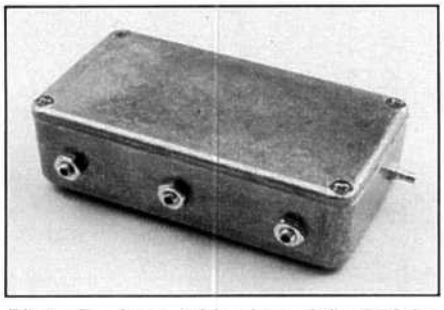


Photo B—An outside view of the TTC in its cast box.

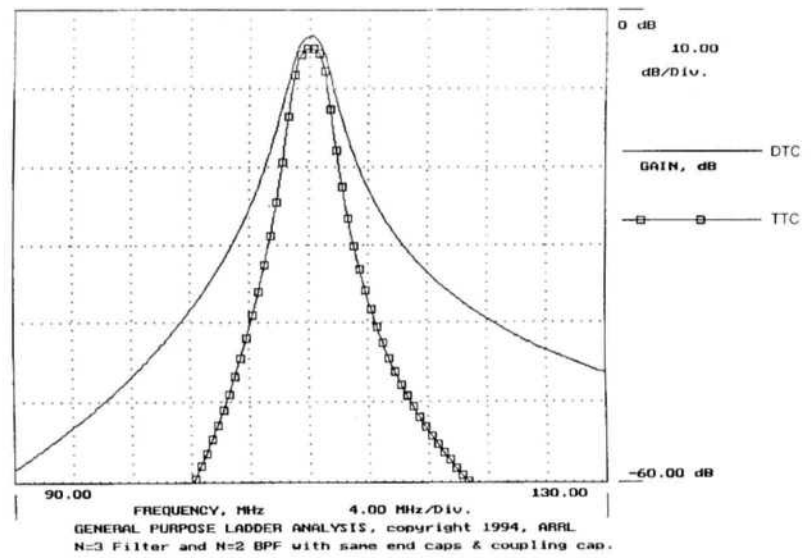


Fig 3—Response of a TTC and a DTC derived from it. Both have approximately the same bandwidth.

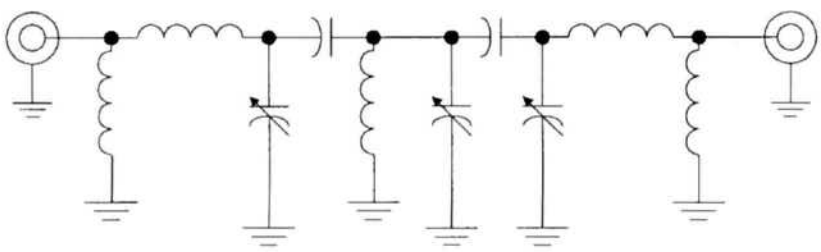


Fig 4—Filter form used with a band-pass circuit shown in photograph A.

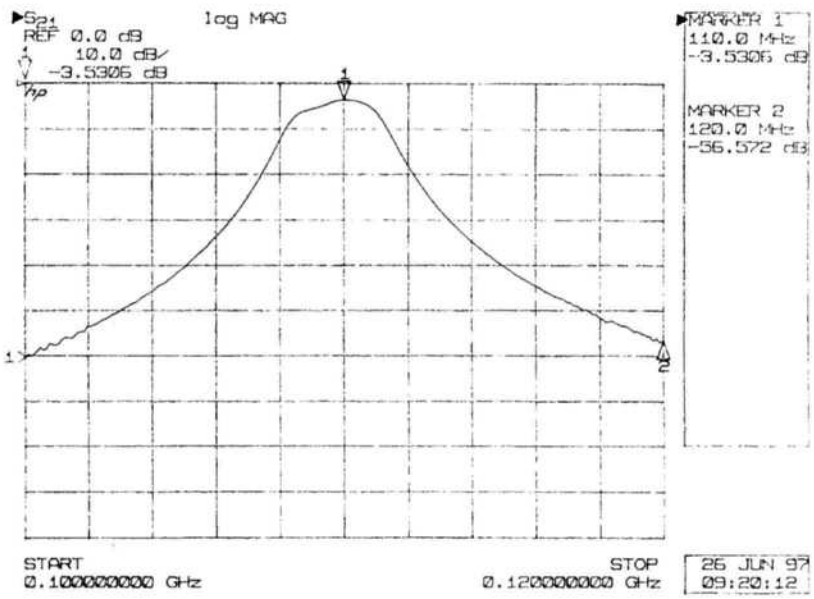


Fig 5—Experimental results with an example triple-tuned filter.

The generality of these methods allows great flexibility. Any of several coupling methods can be used for end-section loading. Even though the schematic shows capacitive taps for attachment between end resonators and the 50 Ω loads, other methods could just as well be applied. For example, the filter shown in the photos used end loading realized with the circuit of Fig 4, where the grounded end of the inductor is lifted from ground and attached to a coaxial connector. Then, a small inductor is attached from the connector to ground. The size of this inductor is varied to establish the end Q. Even though the physical details are different, the measurement schemes and results are not.

Figure 5 shows experimental results, a plot for the experimental filter. The evaluation measurements were performed with an HP-8510B network analyzer. However, the sophisticated instrument was not used for any of the adjustments.

Filter Adjustment with a Return-Loss Bridge

A second useful method is available to the experimenter with a return-loss bridge in his or her laboratory. This instrument, which has often been described in the literature[4], is an impedance bridge where the error, or unbalance signal serves as a measure of reflection coefficient (or SWR) looking into a load. Fig 6 is the schematic for a filter where the input reflection coefficient is to be measured. This analysis was performed with the evaluation version of the popular *PSpice*.[5] The use of a pair of voltage sources at the input generates a voltage, shown as "gam," which *is* the voltage reflection coefficient.[6] Return loss relates to Gamma through

$$Return\ Loss = -20 \cdot Log(\Gamma) \qquad \text{(Eq 1)}$$

Figure 7 shows the gain and magnitude of S11 versus frequency for the complete filter. The match plot dips down to a return loss of 20 dB at the filter center frequency.

Figure 8 shows a modification where the match is calculated for an end resonator that is no longer coupled to the rest of the filter. The result is shown in Fig 9. This match, now only 6.5 dB at resonance, is not nearly as good as the complete filter. If the unloaded resonator Q was much higher, the return loss would be even less. Hence, it is important that the unloaded Q value be reasonably accurate during this simulation.

We eventually wish to use this as an aid to measurements. This was to be a measure of impedance at resonance, but we have not obtained enough information to establish an impedance. Two different resistive loads can provide a 6 dB return loss, just as two different pure resistances can provide a 3:1 SWR. This dilemma is solved with the computer simulation shown in Fig 10, where the single resonator is swept several times with different values of matching inductor with each sweep. A return loss of 6 dB occurs with matching inductor values of 4 nH and 10 nH. Yet only the higher, 10 nH, value produces the desired filter response.

The information from the simulations can now be used to tune our circuit. The TTC filter is built, but the center resonator is either short circuited, or removed. One end is driven at 110 MHz by a return-loss bridge and the resonator is tuned for best return loss, occurring with the largest dip. The matching (grounded) inductor is then varied and retuned until a return loss of about 6 dB is measured. The matching inductor should have a total wire length of about an inch, corresponding to an inductance of about 10 nH using the rule-of-thumb that a

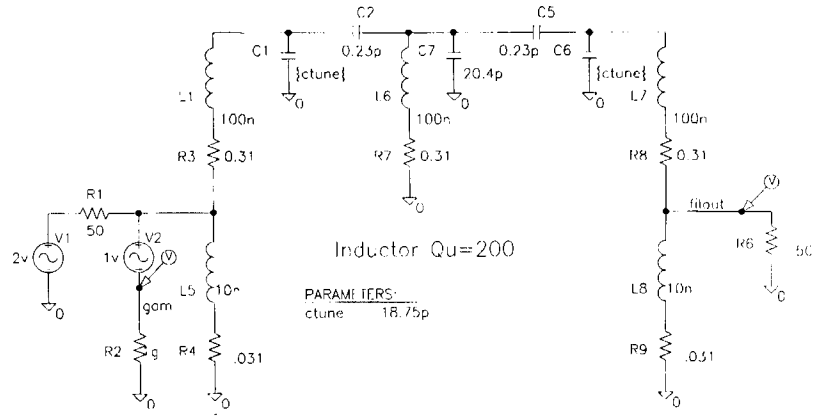


Fig 6—Schematic for practical version of the filter, set up for analysis in *PSpice*.

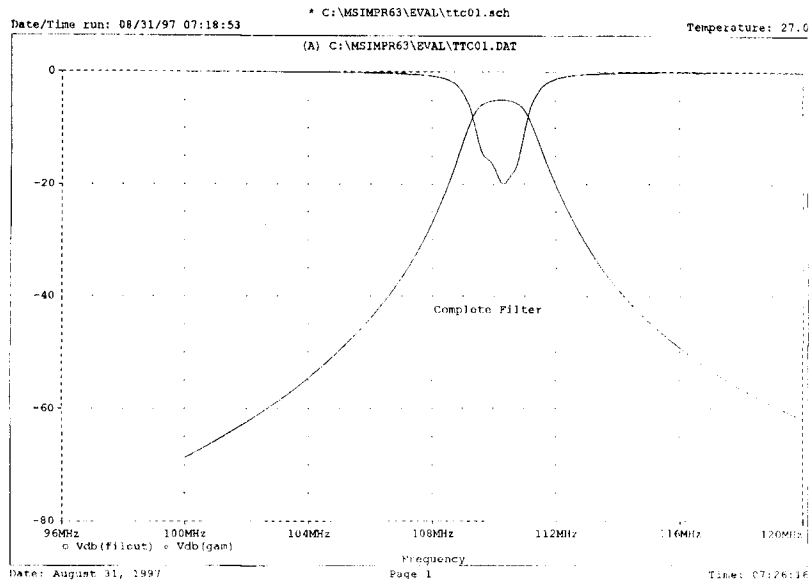


Fig 7—Transfer function and impedance match for complete filter. The return loss looking into the filter is about 20 dB at band center.

wire has L = 0.5 nH per millimeter of wire length. This procedure is repeated for the other end of the filter.

Having established the end Q, the center resonator is added to the circuit. Equal valued coupling capacitors are added between resonators, always keeping the values as small as "seems" reasonable. The three resonators are tuned at 110 MHz as the coupling caps are adjusted. A wide-band sweep is performed during the process to guarantee that overcoupling is not producing extra response peaks. If extra peaks appear, the coupling capacitors are reduced until the desired band-pass shape and bandwidth are obtained.

Some Practical Considerations

Band-pass filters are critical elements in most RF systems and should be built with care. Shielding is often needed; not only is it important to shield the resonators from the outside

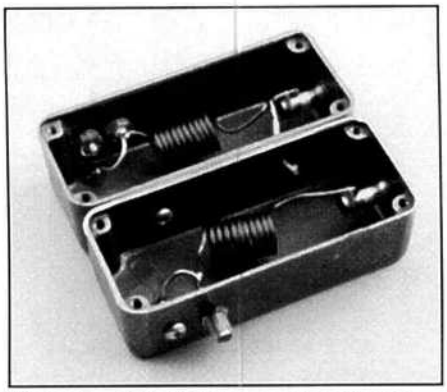


Photo C—This DTC is mounted in *two* boxes to improve isolation.

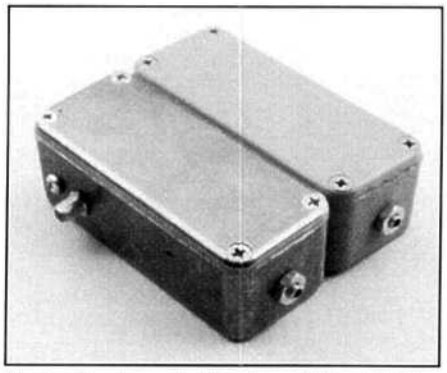


Photo D—An outside view of the DTC in two boxes.

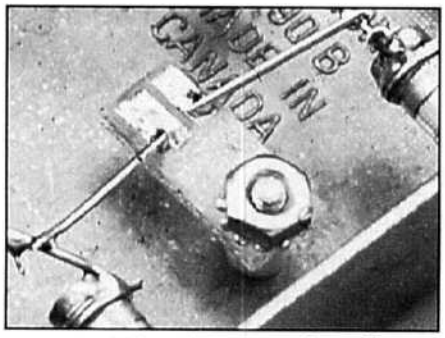


Photo E—This enlargement from Photo A shows one method to achieve tiny coupling capacitances. This method of capacitively coupling two conductors is sometimes called a "gimmick" capacitor. See text.

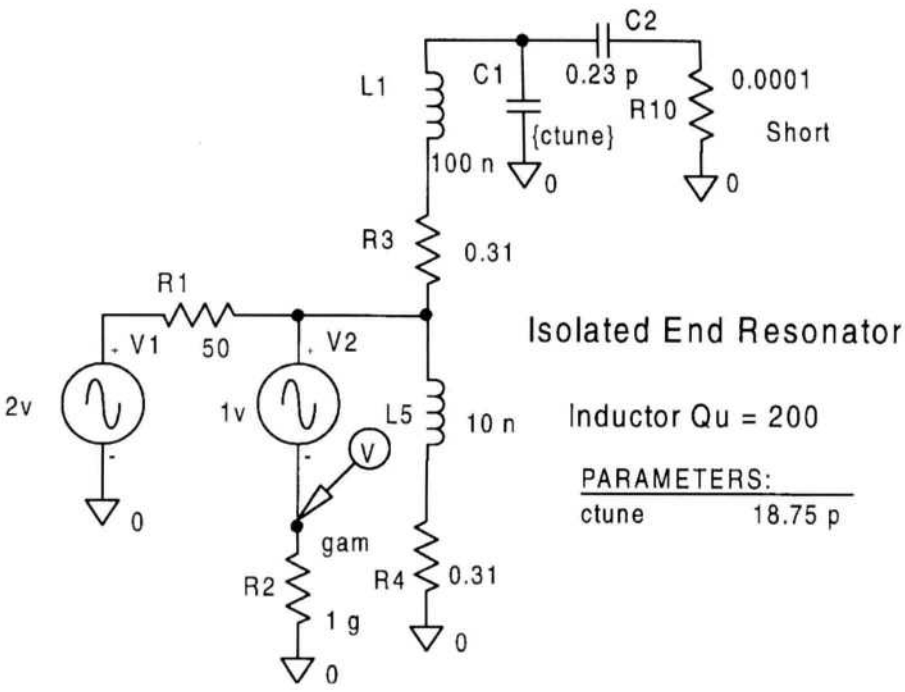


Fig 8—Calculation of reflection when looking into an end resonator without other coupled resonators.

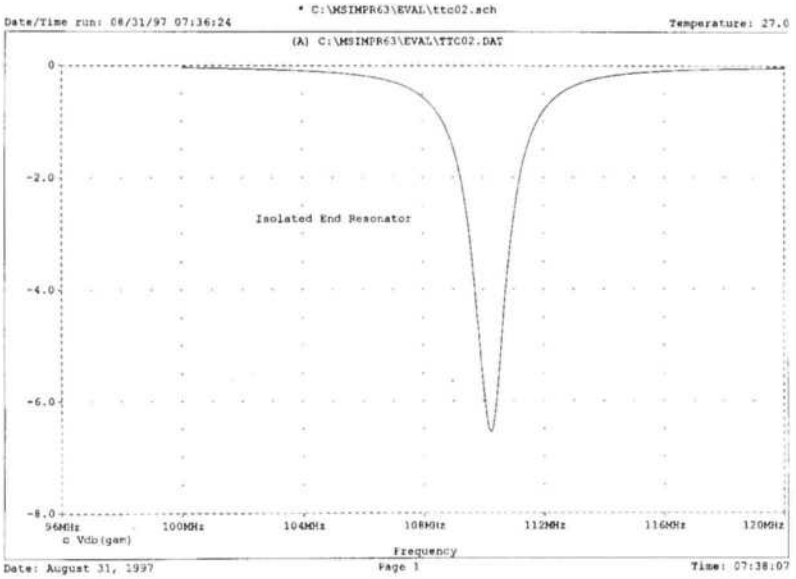


Fig 9—Reflection looking into the isolated resonator.

world, but shielding between resonators is often required. The 110 MHz example TTC is built in a single cast-aluminum box. (See Photos A and B.) Small-diameter coils were picked specifically to minimize interaction between resonators, eliminating the need for internal shields. Photos C and D show a double-tuned circuit built in two cast-aluminum boxes that have been bolted together. Large, higher Q inductors are used. This scheme can be expanded to numerous filter elements. It is often worthwhile to flip adjacent boxes so that lids alternate, side to side to accommodate cast boxes with nonparallel sides.

The small valued coupling capacitors needed in band-pass LC filters are often difficult to realize. One scheme that I have applied is sometimes called a gimmick capacitor (shown in Photo E). Two small, isolated pads are fabricated on a scrap of single sided circuit board material. (Double-sided board has excessive capacitance related to the board material.) Wires are then run from the "hot" ends of the resonators to the pads. The wires are kept a bit longer than required to reach the pads. The excess ends can then be bent close to each other, as needed, to adjust the capacitance. In the example filter, the excess wire lengths were completely trimmed away, for the stray pad-to-pad capacitance provided the needed coupling.

The measurements outlined are best done with the best instrumentation available. This may be elegant commercial gear, or simple home built tools. It is quite possible to build and adjust a TTC with a signal generator and a home-brew power meter, such as the unit recently described by K7OWJ.[7] Once the filter is finished, it can be integrated into a spectrum analyzer, an instrument that will then simplify the adjustments the next time a TTC is needed.

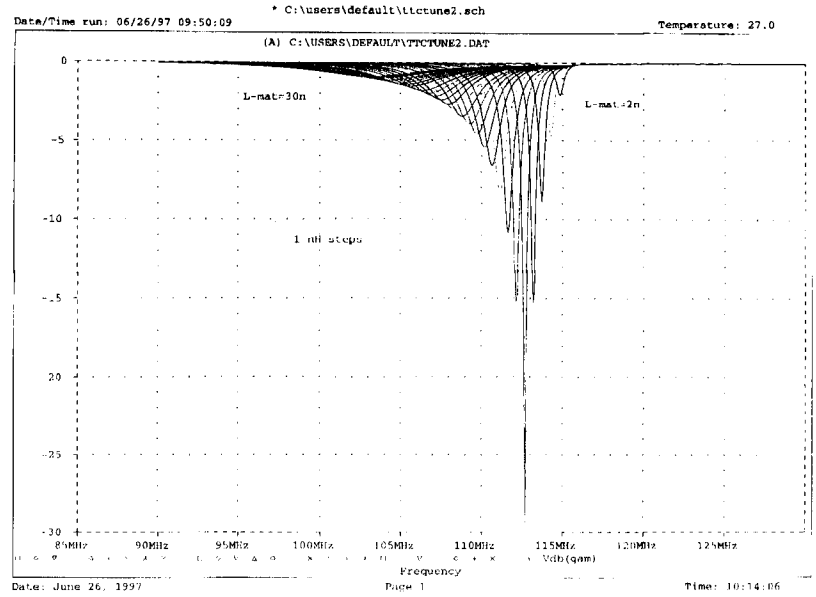


Fig 10—Reflection looking into an isolated resonator with the matching inductor changed in 1 nH steps.

References:

[1]The filters were designed with *DTTC*, a program offered with the ARRL version of "*Introduction to Radio Frequency Design*," ARRL Order No. 4920. The analysis of Fig 3 was generated with *GPLA*, also offered with the text. *ARRL Radio Designer* is also suitable for analysis of filters of this sort. *PSpice* is a program offered by MicroSim, 20 Fairbanks, Irvine, CA 92718. Also, see **www.microsim.com**. The evaluation version of *PSpice* is a very effective, yet affordable tool. ARRL publications are available from your local ARRL dealer or directly from ARRL. Mail orders to Pub Sales Dept, ARRL, 225 Main St, Newington, CT 06111-1494. You can call us toll-free at tel 888-277-5289; fax your order to 860-594-0303; or send e-mail to **pubsales@arrl.org**. Check out the full ARRL publications line on the World Wide Web at **http://www.arrl.org/catalog**.

[2]Hayward, "The Double-Tuned Circuit," *QST*, Dec. 91, pp 29-34.

[3]See IRFD (See Note 1). The coils are designed with a program on the disk, *COILS.EXE*.

[4]The return-loss bridge is described in IRFD, in *Solid-State Design for the Radio Amateur*, (Newington: ARRL, 1977; ARRL Order No. 0402) and in several editions of the *ARRL Handbook (ARRL Order No. 1786)*, as well as numerous articles.

[5]See information on MicroSim presented in Note 1.

[6]Hayward, "Reflections on the Reflection Coefficient," *QEX*, Jan 1993, p 10-12.

[7]Bramwell, "The Microwatter," *QST*, June 1997, pp 33-35. □□

NEC *and* MININEC
Antenna Modeling Programs: A Guide to Further Information

Thinking about getting some antenna modeling software? First, read this overview of offerings—and pointers to more information.

By L. B. Cebik, W4RNL
ARRL TA for Antennas and Antenna Modeling

Much antenna research and design work is based on systematic antenna modeling in a version of the *Numerical Electromagnetics Code* (*NEC*). The following brief notes are not a review of existing antenna modeling programs, but a basic guide to getting further information on these programs from the software providers.

NEC and *MININEC* are method-of-moments calculation programs for antenna modeling. *NEC-2* and *MININEC* are public-domain programs. *NEC-4* is proprietary. References at the end of this article provide some background information on the history, techniques and limitations of these methods.

1434 High Mesa Dr
Knoxville, TN 37938-4443
e-mail **cebik@utk.edu**

The basic calculating programs are not user-ready. Most individual and small-business users purchase a version of *NEC* or *MININEC* from a vendor who has added input and output interfaces (and often added correction factors) to the basic calculating module(s) to make them user-ready. This guide provides ways to contact commercial sources, starting with the most advanced versions.

NEC-4

The latest version of *NEC* is *NEC-4*, which overcomes most shortcomings of earlier codes. It permits modeling of underground radial systems, elements of varying diameter sections and carefully constructed close-spaced parallel wires, as well as all the modeling capabilities of earlier versions of the code. (Some of these capabilities, however, while superior to those of *NEC-2*, have limitations that are just now becoming well documented.) *NEC-4* is a proprietary code of the Lawrence Livermore National Laboratory, University of California, from whom a user-license must be obtained. Export restrictions apply.

To obtain a user-license, contact Gerald J. Burke, L-156, Lawrence Livermore National Laboratory, PO Box 5504, Livermore, CA 94550; e-mail **burke2@llnl.gov**. The price of the license is $850 ($150 for an approved educational site).

In addition to a license, users must also create or purchase the required interface programming. At present, I

know of only one source of commercial software for *NEC-4*: Roy Lewallen, W7EL. *EZNEC Pro* has an option for *NEC-4* (*EZNEC/4*), if the purchaser has a confirmed license for *NEC-4* ($600). *EZNEC Pro* is also available for *NEC-2* (*EZNEC-M*—$425). Both versions offer the same features, except for the calculating engine options, and can handle more than 3000 segments and 1000 wires—very large antenna models. (See the next section for notes on *NEC-2*.) W7EL also makes available *EZNEC*, a segment-restricted (500) version of *NEC-2* ($89, now in version 2 with enhanced pattern-plot exploration features, SWR plots and other upgrades) and *ELNEC*, a version of *MININEC 3* ($49). All W7EL software packages are written for DOS. They employ similar user interfaces that have earned praise for user friendliness in many circles. Contact Roy Lewallen, W7EL, PO Box 6658, Beaverton, OR 97007; e-mail **w7el@teleport.com**. URL **http://www.teleport.com/~w7el/.**

On the horizon, and possibly available by the time these notes appear in print, is *GNEC*, a Windows-based version of *NEC-4* from Nittany Scientific. A license for *NEC-4* will be required to purchase this program. *GNEC* attempts to implement the full *NEC-4* command set and provides a large array of plotting and other graphical outputs, including rectangular plots of many performance figures and color-coded analytical antenna views. See the entry under *NEC-2* for contact information.

NEC-2

NEC-2 is a highly capable version of the code, which is in the public domain. It is restricted to antenna elements of a single diameter (although some software providers have introduced corrections for linear elements of varying diameters). It cannot handle buried radial systems, although above-ground systems close to the earth can be handled. It is equipped with the Sommerfeld-Norton high-accuracy ground model, for accurate modeling of horizontal wires close to the earth.

Nittany Scientific produces a *Windows* version of *NEC-2* at two levels: *NECWin Pro*, V 1.1 (NWP; $425) and *NECWin Basic* (NWB; $75: restricted to 500 segments). Both employ a spreadsheet geometry construction page, pull-down boxes for other antenna parameters, and *Windows*-style graphical outputs, including three-dimensional pattern views. In addition, NWP provides direct entry or importation of *NEC* model input "cards" and provides a large assortment of rectangular output graphics, along with other advanced *NEC* capabilities. These include the color-coded analytical antenna views (Necvu 2). Contact Nittany Scientific, 1700 Airline Hwy, Ste 361, Hollister, CA 95023-5621; e-mail **sale@nittany-scientific.com;** URL **http://www.nittany-scientific.com.**

In addition to *EZNEC* and *NECWin*, *NEC-2* is also offered by Brian Beezley, K6STI, in a package called *NEC/Wires* 2.0. This package is part of a suite of programs offered by K6STI, the most well known of which are *AO* and *YO*. Contact information appears in the section on *MININEC*.

MININEC "Professional"

Before recent advances in speed and memory, it was not feasible to run *NEC* on a PC. Rockway and Logan developed *MININEC*, a Basic language adaptation of *NEC* for PCs. More recently, they have advanced the *MININEC* algorithms and code to overcome many of its initial limitations. The "new" *MININEC* can handle sharp angles in antenna geometry directly (without segment length tapering) and antennas close to ground with much better accuracy. However, the *MININEC Professional* code is a proprietary product. EM Scientific offers several levels of *MININEC Professional*, including the basic level *MININEC* for *Windows* (400 wires; $125), *MININEC Professional* (1000 wires; $390), and *MININEC Broadcast Professional* (2000 wires; $790). These are all *Windows* products. Contact EM Scientific, Inc, 2533 N Carson St, Ste 2107, Carson City, NV 89706; e-mail **76111.3171@compuserve.com**. URL **http://www.emsci.com**.

MININEC

The public domain *MININEC* code (version 3.13) is available with several commercial user interfaces, as indicated in these notes. For general antenna analysis that does not press its limitations, *MININEC* is a highly competent code. It handles elements of changing diameter directly and, with segment-length tapering, can accurately model a wide range of antenna geometries. However, horizontal antennas must be at least 0.2 wavelengths above ground for accurate results. Moreover, specification of ground conditions affects only antenna far-field results, not feed point conditions. Antenna size is limited to 256 segments. Despite limitations, *MININEC* has shown results superior even to *NEC-4* on certain types of antenna structures involving close-spaced wires, large sudden changes of element diameter and angular junctions of wires with dissimilar diameters.

Orion of Canada (Madjid Boukri, VE2GMI) offers a *Windows*-based version of *MININEC*, *NEC4WIN*, using a spreadsheet geometry input page, pull-down boxes for other antenna parameters and a pattern-plotting output that includes lobe identification and bandwidth. In addition, the user can vary the height of the antenna without invoking a complete recalculation of the matrix for faster results. The current version is 1.9L ($35). A *Windows 95*-only version with enhanced capabilities may become available soon. Contact Madjid Boukri, VE3GMI, Orion Microsystems, 197 Cr Joncaire, Ile Bizard, Quebec, Canada H9C 2P7; e-mail **mboukri@cam.org**. URL **http://www.cam.org/~mboukri**

Brian Beezley, K6STI, also offers a wide range of *NEC*-related DOS-based software, as well as a Yagi-optimization program and a terrain-analysis program. Perhaps the best known program is *YO*, the Yagi-optimizing program that uses special algorithms calibrated to *NEC-2* (offered as *NEC-Wires*). *AO* (Antenna Optimizer) is a *MININEC*-based program with an optimizing function included. *AO* includes a frequency-correction factor to bring *MININEC* results in line with *NEC* results as antenna frequency increases. The input section of the program includes the ability to use symbolic expressions for antenna dimensions, thus permitting alteration of multiple parts of the structure with a single entry revision. *AO*, *YO* and *NEC/Wires* are each $70 (or any three of the K6STI antenna programs for $120). Contact Brian Beezley, K6STI, 3532 Linda Vista, San Marcos, CA 92069; e-mail **k6sti@n2.net**.

ELNEC is W7EL's DOS-based version of *MININEC*. It uses an interface very similar to those of his programs for *NEC-2* and *NEC-4*. The program contains a parallel-wire correction. See the section on *NEC-4* for contact details.

Other Information Sources

The Applied Computational Electromagnetics Society (ACES) is perhaps the professional focal point of advances in all forms of electromagnetics codes and related mathematical models. It holds an annual meeting on the West Coast in March with a very

full and varied program. Further information can be found at their Web site **http://www.emclab.umr.edu/aces/.**

The Unofficial *NEC* Archives are maintained by Ray Anderson, WB6TPU. Formerly, this collection of *NEC*-related software was available only via FTP. However, the entire contents are now accessible via the Web. They include many source codes for *NEC* and for pre- and post-processing of *NEC*, along with some sample input files. Look for them at **http://www.qsl.net/wb6tpu.**

The Unofficial *NEC* Home Page is supported by Peter D. Richeson, who maintains on-line copies of the *NEC-2* manual (Part III). URL **http://www.dec.tis.net/~richesop/nec/.**

There is also a mailing list for those interested in a forum for questions and information. The address is **neclist@ee.ubc.ca.**

This listing is necessarily limited, even within the scope of *NEC*-related software and information sources. However, the indicated Web pages lead to other information on details, specifications, related developments and a more complete understanding of the rapidly expanding field of electromagnetic modeling.

In addition to *NEC*-related software, the Web and related FTP sites provide a variety of antenna modeling, optimizing and calculating programs. See, for example, the collection at the following URL **http://www.qsl.net/k7on/shareware/1997.** Most of these programs are either freeware or shareware, and they can ease a large number of antenna design problems—especially if the resulting designs are then further evaluated on one of the *NEC*-based programs.

For Further Reading

NEC-2 and NEC-4: Basic References

J. Burke, A. J. Pogio, *Numerical Electromagnetic Code (NEC) Method of Moments, a User Oriented Code*, Vol 2 (Part III: User's Guide), Tech Doc 116, Naval Systems Center, San Diego, 1982.

Gerald J. Burke, *Numerical Electromagnetic Code—NEC-4: Method of Moments*, (Parts I and II), Lawrence Livermore National Laboratory UCRL-MA-109338, 1992.

MININEC: Basic References

A. J. Julian, J. C. Logan, J. W. Rockway, "*MININEC: a Mini-Numerical Electromagnetic Code*," Tech Doc 516, Naval Systems Center, San Diego, 1982.

J. C. Logan, J. W. Rockway, "*The New MININEC (Version 3): A Mini-Numerical Electromagnetic Code*," Tech Doc 938, Naval Systems Center, San Diego, 1986.

Method of Moments Techniques

John D. Krauss, *Antennas*, 2nd ed. (New York: McGraw-Hill, 1988), pp 384-408.

Antenna Modeling Articles

Roy Lewallen, W7EL, "*MININEC*: The Other Side of the Sword," *QST* (Feb 1991), pp 18-22.

L. B. Cebik, W4RNL, "A Beginner's Guide to Using Computer Antenna Modeling Programs," *ARRL Antenna Compendium*, Vol 3 (1992), pp 148-155.

John S. Belrose, VE2CV, "Modeling HF Antennas with *MININEC*—Guidelines and Tips from a Code-User's Notebook," *ARRL Antenna Compendium*, Vol 3 (1992), pp 155-164.

R. P. Haviland, W4MB, "Programs for Antenna Analysis by the Method of Moments," *ARRL Antenna Compendium*, Vol 4 (1995), pp 69-73.

In addition to these references, the lengthy manuals accompanying such programs as *EL/EZNEC*, *AO/NEC/Wires*, and *NECWin Pro/Basic* can also make significant contributions to your understanding of antenna modeling. A detailed background document on antenna modeling is also available at the EM Scientific (*MININEC* Pro) Web site **http://www.emsci.com.** □□

A Homebrew 2.4 GHz Waveguide Filter

Here's a filter with a 20 dB low-frequency cut-off at –22 MHz and an insertion loss of less than 0.5 dB.

By John Reed, W6IOJ

Waveguides are natural filters. As an example, scanning an input through an S-band waveguide will result in an insertion loss of less than 0.01 dB/ft from near cut-off to twice cut-off frequency. However, when passing through the waveguide cut-off frequency there is a sudden, near-vertical signal strength drop of over 10^3 dB. What are the disadvantages? For one thing, it's simply a high-pass filter with the instrumentation probes and baffles controlling the high-frequency response (plus multimode responses at frequencies higher than twice the cut-off frequency). The primary problem is that the cut-off frequency is determined by the waveguide width, which means a standard size waveguide works as a filter with only on one cut-off frequency. The following describes a relatively simple home-brew waveguide that does not seriously compromise the conventional waveguide characteristic, and the size is the builder's choice. It is made from available materials using ordinary hand tools. The size of the described filter suits it for use as a 2.4 GHz receiver's front-end filter. It results in excellent image and other low-frequency signal rejection, while using a HF receiver as the IF. As an example, the receiver image for 30 MHz IF will be –37 dB down.

770 La Buena Tierra
Santa Barbara, CA 93111
e-mail **reedioj@aol.com**

Description

The filter schematic is shown in Fig 1, and Photos A and B show the completed filter. [Notice that the sketch in Fig 1 shows only one half of the filter. The filter is symmetrical about the frequency trimmer.—*Ed.*] There are five adjustments: Two are the sliding end baffles, which match the input/output probes. The center adjustable probe fine tunes the cut-off frequency, and two adjustable probes mounted above the input/output probes optimize the coupling coefficients. Aligning the filter for optimum performance is a fairly involved process because all five adjustments are interrelated. Begin alignment by tuning the end baffles and frequency trim for maximum signal with minimum coupling-probe protrusion into the waveguide. The output will indicate a normal bandpass character, but the insertion loss will be 1 or 2 dB. Next, gradually turn in the coupling probes while repeaking the output with the other adjustments, until there is minimum insertion loss. Minimum insertion loss indicates peak performance. Continued coupling-probe increases will gradually degrade the cut-off characteristic with no more insertion-loss improvement.

Fig 2 shows the frequency response, following a careful adjustment sequence. The irregular high-frequency response is due to input/output coupling variables. Insertion-loss measurements were made by first minimizing source/detector matching variables (by inserting 20 foot RG-58 cables at the filter's input and output ports,

about –7 dB per cable). Comparison measurements were made with the filter and then substituting a 1 foot cable for it. This substitution method indicates a 0.25 dB insertion loss, which represents the filter loss without cables and connectors. A second measurement was made by simply installing the filter in front of a 2.4 GHz receiver. That measurement indicates an insertion loss of 0.4 dB. The actual cut-off response depth is unknown due to test equipment limitations. This low-level floor response is mainly related to the distance between the input/output probes. The response decays exponentially as probe spacing increases.

The filter is very stable and it will stand moderate physical abuse, as long as there is no deformation of the aluminum shield.

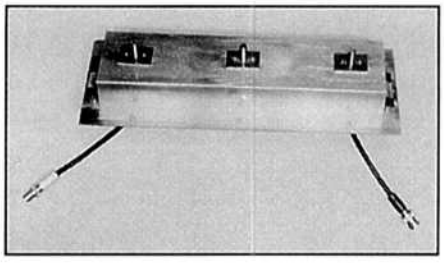


Photo A—The finished 2.4 GHz filter.

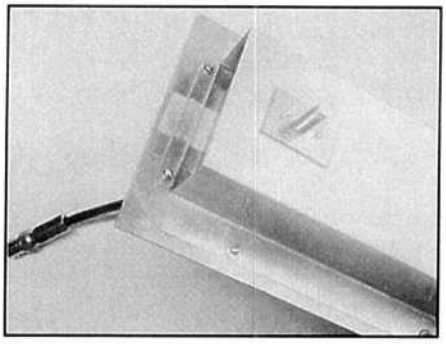


Photo B—An end view of the completed 2.4 GHz filter.

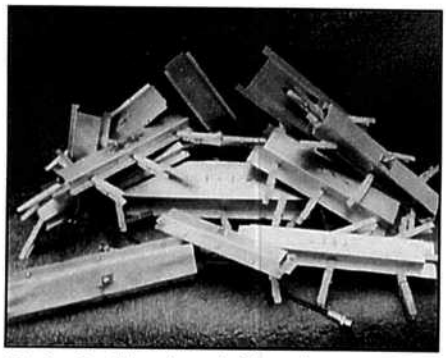


Photo C—Wooden clothespins secure the parts of experimental filters.

Fabrication

Forming the aluminum cover is the most difficult process. (Figs 3 and 4 show details of the probes and adjustment assemblies.) A width error of $^1/_{64}$ inch changes the cut-off frequency by 15 MHz. A consistent error of that magnitude is of little consequence because it can be compensated for by the center-frequency trim probe. However, a random irregular error along the waveguide length, or sides that are not parallel, will contribute to a

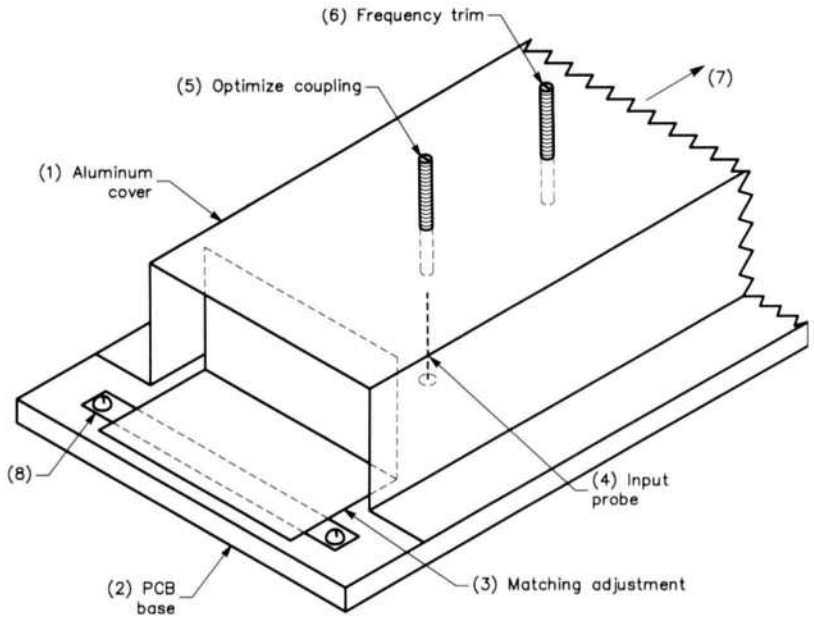


Fig 1—Schematic of the high-pass waveguide filter. (1) Aluminum cover, 0.020 inch thick. (2) Double-sided glass-epoxy PC board base, $^1/_{16}$×4×11$^1/_4$ inches. (3) Sliding end baffle, 0.025 inch thick brass, 1$^5/_8$ inch wide, $^7/_8$ inch sliding base, baffle tailored to clear cover by $^1/_{32}$ inch. (4) 0.5 inch cable inner conductor (detail in Fig 3). (5, 6) Adjustable probes are $^1/_4$-28 brass screws (detail in Fig 4). (7) Input/output assemblies are identical. (8) Plastic bars clamp the sliding baffles in their final positions.

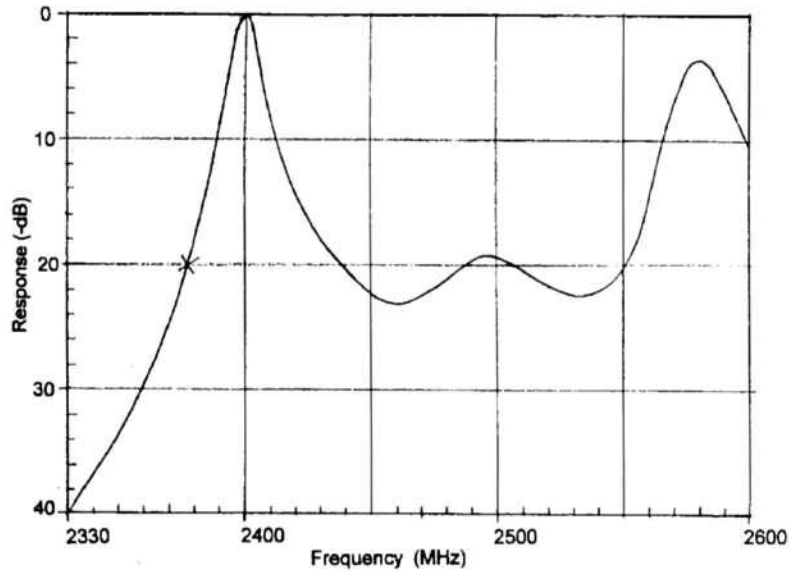


Fig 2—Frequency response of the 2.4 GHz high pass filter. Notice the –20 dB response at –22 MHz from the cut-off frequency (11 MHz at the IF).

mushy cut-off. In addition, insertion loss increases if the mounting surfaces between the cover and the mounting PCB are not flat.

The bends are made using a jig made of two extruded aluminum pieces: 1 inch square, by 13 inches long, with 1/16-inch-thick walls (available in most hardware stores). The pieces are bolted together at each end with 1/4-28 bolts, to make an assembly 1×2×13 inches. By clamping the aluminum sheet between the two pieces, you can form a sharp 90° bend using a hammer together with a 3/4×1 1/2×11 wood interface ("beater bar") to prevent hammer distortions in the aluminum.

I suggest that you first start with a 6 1/8×10 inch piece of 0.020 inch aluminum sheet. Scribe a centerline and parallel lines, 1-15/64 inches away, on each side of the centerline. In addition, drill all holes required for the finished cover. See Fig 5 for the detailed layout. After drilling, remove any burrs and confirm the flatness of the aluminum. Then bend at the scribe lines to form a U-shape cover. Next, scribe a line on each sidewall, 1 1/4 inch down from the U cover top. Finally, again using the jig, bend the mounting flanges. At this point, you should have a cover with the proper dimensions. I recommend a dry run of the bending procedure with a narrow piece of aluminum before making the final cover.

A second critical area is the RF joint between the PC board base and aluminum cover, and to a lesser degree, between the base and sliding brass baffle. Both are RF butt joints, which means they must be flat and smooth. To facilitate processing of the aluminum cover, I used a jig consisting of two 3/4×1 1/2×12 inch wood pieces. The 1 1/2 inch sides are firmly spaced apart—about 1/8 inch wider than the cover—by screwing them to two cross members. The exposed, parallel 3/4 inch surfaces are flattened by sanding the assembly with sandpaper backed by a glass surface. The same procedure is used to finish the cover mounting surfaces, placing the cover in the jig and starting with 150 grit (fine) sandpaper, then 220 grit (very fine) and finishing with 320 grit. As the assembly is moved back and forth on the sandpaper, the cover will probably slip in the jig. This can be avoided by drilling a 5/16 inch diameter hole in one of the jig's cross members to accommodate one of the cover's adjustment screws. Assemble the adjustment assembly and lock the 1/4-28 screw in an extended position, such that it fits in the jig hole.

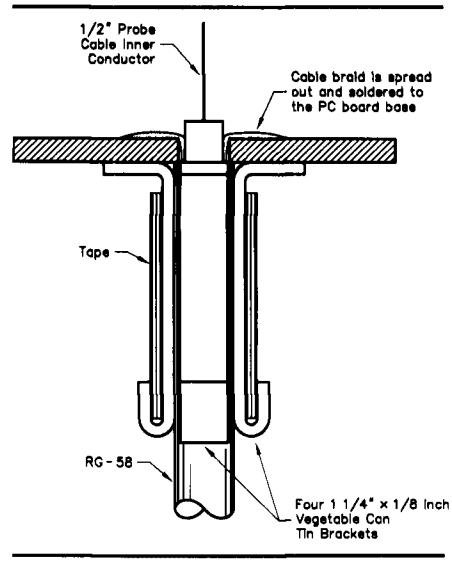


Fig 3—Detail of the input/output probes. The tin L brackets are held against the cable and soldered to the base. Firmly wrap 3/4 inch wide tape around the brackets (about six layers), to clamp the cable in a vertical position.

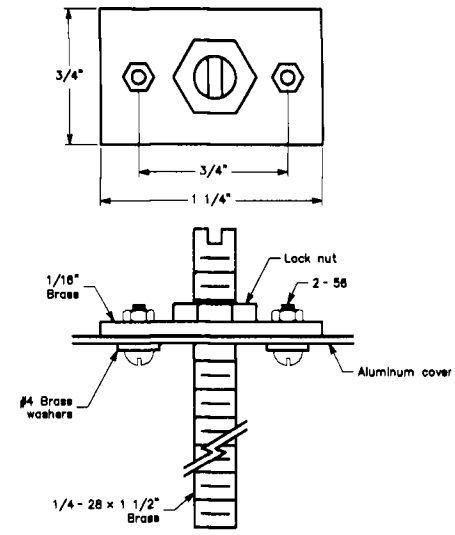


Fig 4—Detail of the adjustment assemblies. Polish the interface surfaces between the cover and the brass plate.

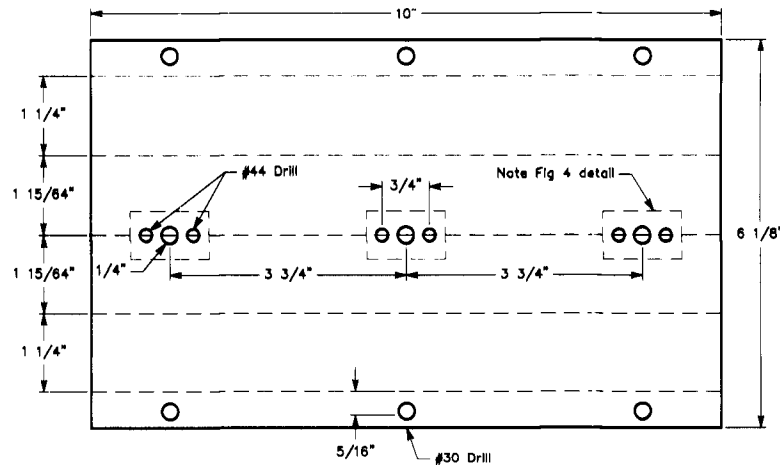


Fig 5—Layout of the aluminum cover.

Following the polishing procedure, wash the cover assembly in soap and water; aluminum oxide makes a poor butt joint. At this point, and when not mounted on the base PC board, the polished cover mounting surfaces are fairly fragile; handle them with care. Use the same polishing procedure to finish the sliding brass baffles.

During initial tests, I use wooden, spring-loaded clothespins to hold the assembly together (see Photo C). Use three on each side of the cover and one for each sliding baffle. You can estimate whether the cover is leaking RF by squeezing the joints at various positions with your fingers. If there is no change in the output, it is good evidence that the joints are good. In the final assembly, 3/16×1/2×10 inch wood molding is placed on top of the cover mounting surfaces when fastening the cover to the base with #4-40 hardware. The wood distributes the force of the mounting hardware to prevent possible stressing of the aluminum.

This home-brew waveguide fabrication method has been used to fabricate waveguide filters up to 5.9 GHz. The test data was limited to a 20 dB range due to inadequate test equipment. However, this data indicated similar performance as compared to that of the 2.4 GHz filter. □□

RF

By Zack Lau, W1VT

An Inexpensive 3456 MHz Dish Feed

After looking at Paul Wade's excellent set of articles and seeing what the TVRO industry uses, it appears that a scalar ring or Chaparral feed is an excellent choice. I'll describe how to build one without busting the bank.

The first step is choosing the appropriate material for the circular waveguide. I decided on 5 oz macadamia-nut cans. While a bit shorter than normally recommended, they do have a diameter in the proper range. More importantly, they are nicely tinned on both the inside and outside, so you can easily solder to them. Unfortunately, the removable lid and attachment ring are made out of aluminum, so you can't easily solder to it to extend the waveguide. Copper or brass tubing would work even better, if you don't mind the cost or can find a cheap source. According to Doug Sharp, W2SZ/1 uses a racquet-ball can for 3456 MHz.[1] I suggest looking around with a ruler—you might already have something suitable in your closet!

The cylindrical dish feed material by Paul Wilson, W4HHK, suggests a feed-diameter range of 2.25 to 2.60 inches, assuming a design frequency of 3456.1 MHz.[2] Most of the narrowband DX activity occurs close to this frequency. Margarete Ralston suggests a diameter range of 0.70 to 0.75 λ, which corresponds to a tighter range of 2.39 to 2.56 inches.[3] I estimate the inside diameter of the macadamia nut can as being about 2.57 inches, although it's not easy to measure accurately. The inside diameter of the aluminum attachment ring that remains after you remove the lid is about 2.41 inches.

The attachment rings posed another problem—how do I mount a scalar ring on the can? If I make a big enough hole to clear the attachment rings, the scalar assembly won't fit snugly against the can. One option is to split the scalar assembly, and then reassemble it around the can. I opted for a simpler solution, using brass sheet as shim stock to form a tight fit. (See Fig 2.) This reduced the need for precision

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

[1]Notes appear on page 56.

metal work, since one can compensate for a too-large hole by using thicker brass sheet.

I make my scalar rings out of brass sheet—both 25 and 32 mil stock seems to work adequately. First mark the sheet with a compass; then make the big hole for the cylindrical horn. Don't bother trying a Greenlee punch to make a big hole in a piece of brass sheet—I found it distorts the metal too much to be useful. Instead, I put on my safety goggles and carefully use a hole or "fly cutter." Not only is it necessary to use a drill press, but one has to carefully clamp the metal it place. It would probably help [and be much safer!—*Ed.*] to clamp the metal between thin sheets of wood and cut through the "sandwich," but I've gotten by without that effort.

One-inch-wide 25 mil brass strip stock works well for making the rings. Just as when building loop Yagis, you bend the strips around a metal can to form their shape. It looks a little nicer if you bend the strips first, and then cut them to size, but cutting them to the exact size can be a challenge. Cutting the strips to length first makes it more difficult to bend the ends of the strips to the precise shape. It probably doesn't make a lot of difference which approach you choose, in terms of electrical performance.

I've used the scalar feed dimensions on page 9-31 of *The ARRL UHF / Microwave Experimenter's Manual* with good results.[4] It suggests different scalar-ring mounting positions for dish f/D ratios varying between 0.30 and 0.45. It may not be necessary to use all three rings—I've heard of EME stations claiming good results with just one ring. Obviously, if someone has good measurements that compare the trade-offs between the number of rings and antenna efficiency or G/T, I'd certainly like to know. With small dishes, such as a 2 foot dish on 2304 or 3456 MHz, it is quite possible that better results may be obtained with one or two rings, since the blockage from the extra rings typically becomes considerable.

The tricky part of using the tinned-

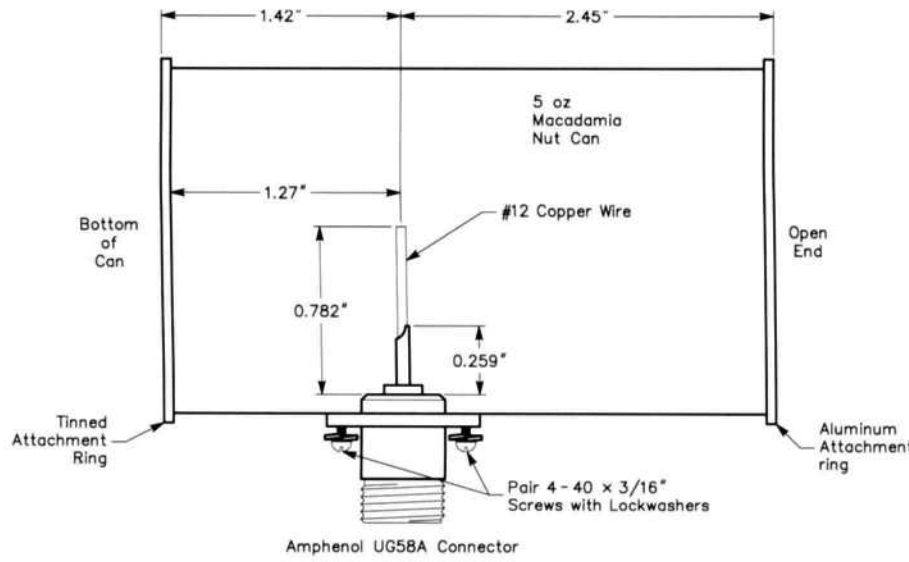


Fig 1—Dimensions of the 3456 MHz coax to cylindrical waveguide transition.

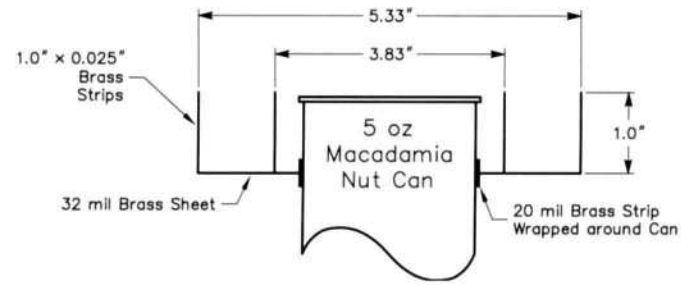
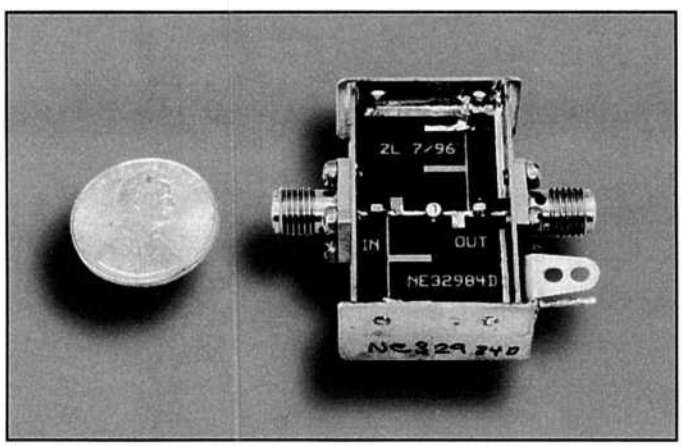
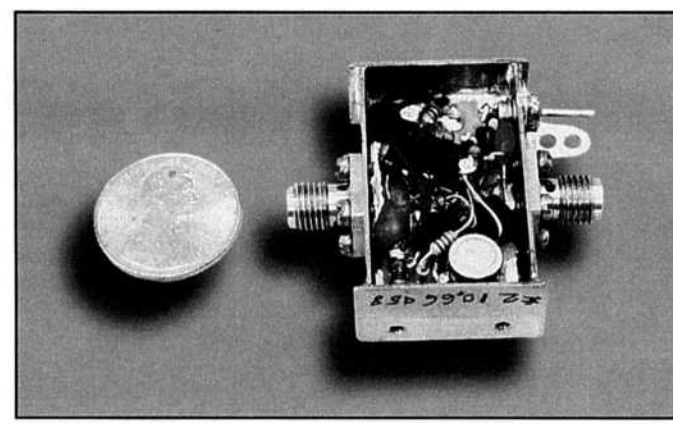


Fig 2—Suggested dimensions for a scalar feed suitable for feeding a small 0.45 f/D parabolic dish.



(A)



(B)

Fig 3—The NE32984D preamplifier.

steel can is making the hole for the coax-to-waveguide transition. (See Fig 1.) The best method I've found is to use a Uni-bit drill to make the hole. Ordinary drill bits don't work well, because they tend to tear ugly holes in the thin metal. While I have soldered connectors to the cans, a better approach is to make a 0.8×1.5×0.032-inch brass mounting plate with tapped #4-40 holes for the connector and solder the plate to the can. This allows you to easily remove the connector and adjust the probe.

The probe for the transition is made out of bare #12 copper wire. It extends the center pin of the N connector from the original 258 mils to 782 mils, measured from the surface of the Teflon to the tip of the probe. The spacing from the back wall is 1.27 inches. Consider these as starting dimensions that may be tweaked for better performance. The unit I built has a return loss of roughly 24 dB.

Measuring return loss or SWR may be a problem at this frequency. Paul Wade published a simple return-loss bridge in the Feb 1995 *QEX*, but its performance is marginal. (It might help to use even smaller chip components and thinner circuit board.) I used a surplus directional coupler, a spectrum analyzer, signal generator and a transverter. While a power meter will work as the signal detector, the spectrum analyzer can sort out unwanted spurs, so your signal source need not be well filtered for you to get meaningful results. The transverter converts the signal generator over to 9 cm, since I don't have a high-quality signal generator that covers this band.

NEC NE32984D 10 GHz Preamp

The latest NEC low noise PHEMT offers a very low noise figure—only 0.37 dB at 10 GHz according to the data sheet. The challenge is to preserve the low noise figure while adding matching networks for use in a 50 Ω coaxial system. This design shown here has a 0.7 dB noise figure and a gain of 11 dB. Figures 3 through 6 show details of the preamp.[5]

While I had good performance with the NE32684A in a *QEX* (Dec 1992) design that was relatively easy to copy, I decided that performance was more important in this design. Thus, I chose to use 10 mil 5880 Rogers Duroid PC board. The thinner board reduces board radiation losses to the point where a cover over the circuit has only a small effect on performance. The disadvantage is that I don't know of a small quantity supplier of this board, although it appears to be popular in European 24 GHz designs. If you don't mind the $250 minimum order, it can be obtained directly from Rogers Corporation.[6]

When matching with microstrip, it generally helps to make the input and output lines as short as possible, since the lines can be pretty lossy at high frequencies. The trade-off is that shorter and shorter lines become more difficult to tune properly with small pieces of copper foil. Thus, while lines only a quarter wavelength can match just about anything in theory, longer lines may be required in practice. The input line is about one-half wavelength. At this frequency, a little tuning is often required to get the best performance possible out of a device.

The manufacturer recommends a bias current of 10 mA with a V_{ds} of 2 V. This is easily achieved with a PNP active bias circuit. PHEMTs are low voltage devices, however, the maximum V_{sg} is 3 V. Thus, I use an LM317L wired as a 3 V regulator to generate the drain voltage, and invert the regulated 3 V with an ICL 7660 regulator to get the negative gate voltage. This is sim-

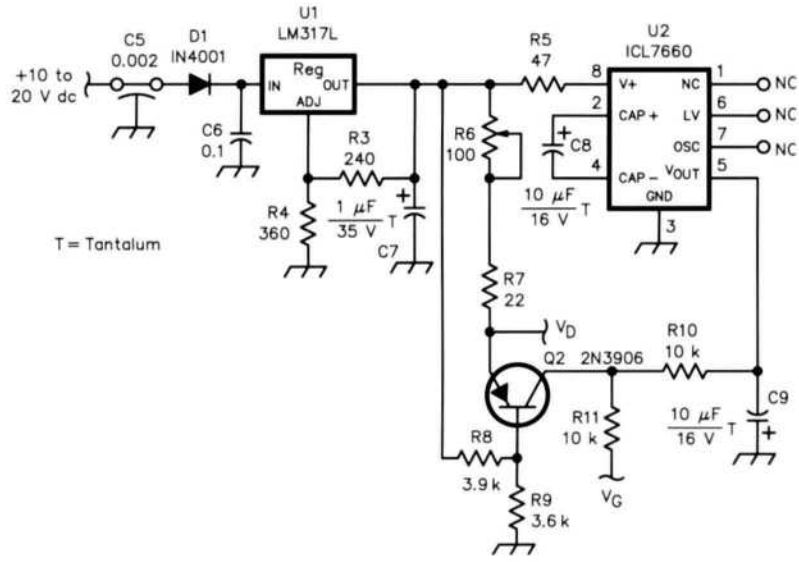


Fig 4—Schematic diagram of the preamplifier power supply.
C1, C3—1 pF ATC 100A, 55 mil chip capacitors
C2, C4—0.01 µF chip capacitors
J1, J2—SMA Panel jacks with center conductor that matches the 30 mil wide microstriplines.

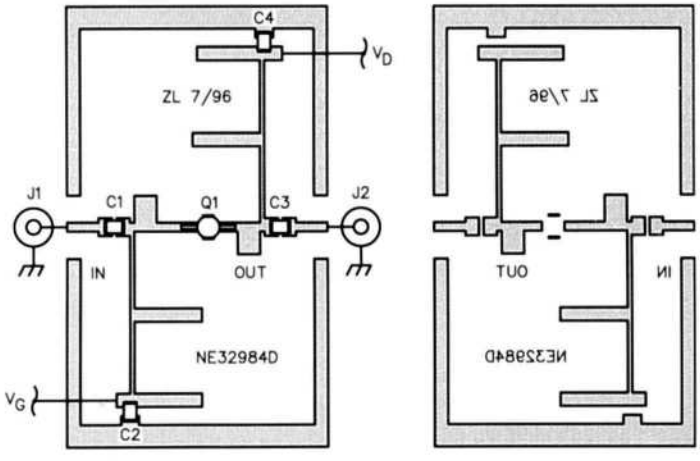


Fig 5—Parts layout of the NE32984D preamplifier.

pler than using two separate regulators for the drain and gate voltages.

Notes

[1]"W2SZ/1 Equipment List as of the September 1992 VHF QSO Party" *Proceedings of the 38th Annual West Coast VHF/UHF Conference*, pp 142-145. The Proceedings are ARRL Order #4327. ARRL publications are available from your local ARRL dealer or directly from ARRL. Mail orders to Pub Sales Dept, the ARRL, 225 Main St, Newington, CT 06111-1494. You can call us toll-free at tel 888-277-5289; fax your order to 860-594-0303; or send e-mail to **pubsales@arrl.org**. Check out the full ARRL publications line on the World Wide Web at **http://www.arrl.org/catalog**.

[2]D. Straw, N6BV, Ed., *The ARRL Antenna Book*, 18th ed. (ARRL Order #6133) p 18-14.

[3]"Design Considerations for Amateur Microwave Antennas," Margarete Ralston, KI4VE, *Proceedings of the Microwave Update 1988*, pp 57-64 (available from ARRL, see Note 1).

[4]*The ARRL UHF/Microwave Experimenter's Manual*, ARRL Order #3126.

[5]You can download a Postscript file of the etching pattern from the ARRL "Hiram" BBS (tel 860-594-0306), or the ARRL Internet ftp site: **oak.oakland.edu** (in the **pub/hamradio/arrl/qex** directory). In either case, look for the file 98RF03.ZIP.

[6]Rogers Corporation, Microwave Materials Unit, 100 S Roosevelt Ave, Chandler, AZ 85226; tel 602-961-1382, fax 602 961-4533; URL **http://www.rogers-corp.com/mwu/**.

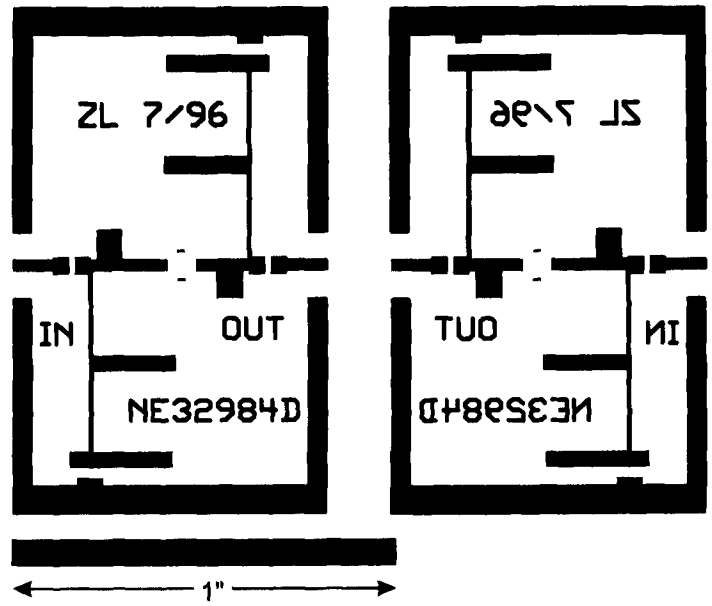


Fig 6—Twice-scale PC board template. The board material is 10 mil 5880 Duroid (ϵ_r = 2.2).

Letters to the Editor

Clarifications of "Exploring the 1:1 Current (Choke) Balun"

◊ As a result of inputs from several readers and further efforts by myself, I would like to offer the following notes and clarifications of this article in *QEX* for July 1997.

Coax Braid Return Current

Text in the middle column of page 13 refers to Fig 1 and a ground connection somewhere on R_L. There is a discussion of the current that returns from the coax center wire to the ground connection and from there through the bottom part of R_L to the coax braid and back to the generator. It states:

"...without the balun the current to ground will not return through the transmission line back to the generator, but will return directly to the generator via the ground path."

This needs some clarification. The current that returns through the coax braid is precisely the current that flows through the bottom part of R_L, which obviously depends on the voltage drop across that portion of R_L. The value of that voltage drop depends on the voltage with respect to ground at point B (Fig 1). To be more exact, a portion of the coax output voltage (V_{IN} in Fig 1) appears across the impedance that is formed by the bottom part of R_L in parallel with the coax braid impedance between point B and the ground at the input of the coax.

Assume for the moment that the balun is not installed. If the impedance between point B and the ground at the input of the coax is zero, then no current will flow in the bottom segment of R_L, and therefore the return current through the braid is zero. It all returns through the ground path, not the coax. This zero-impedance condition exists if the coax braid length is zero wavelengths, 1/2 wavelength, etc. As the length of the coax changes from these values, the impedance of point B with respect to ground becomes greater than zero. At 1/4, 3/4, etc wavelengths the longitudinal impedance of the coax braid is very large, so virtually all of the return current is forced through the coax braid, not the ground path. In other words, the coax performs the balun function. In this special case, the balun may not be needed.

The point is that the value of the return current via the coax braid is wavelength sensitive. On the other hand, this wavelength sensitivity disappears if we use the balun, and the current is "forced" to return through the coax at all wavelengths. It is an important reason for using the balun. This idea is discussed in the section "The Ungrounded Balun," but it should have been mentioned in this part of the article, too.

The zero-wavelength condition mentioned above is often found in circuit-design situations involving short lengths of coax. The balun assures that the return current flows in the coax, not via some poorly defined ground path that may be troublesome. I have experienced this problem.

The Conventional Transformer

In the middle column of page 16, the discussion of Fig 5 concerns the use of a short length of coax as a "conventional" transformer, with the center wire as primary and the braid as secondary. The center wire (primary) current causes an EMF-generating E-field to be generated, which then induces a voltage onto the braid (secondary). This is a very routine conventional-transformer behavior. The value of this voltage, divided by the length of the braid is, by definition, an E-field of some kind.

The problem is that the E-field in a perfect conductor is supposed to be precisely zero. So there is some confusion on this point. The answer is as follows, as described by Skilling, *Fundamentals of Electric Waves*, 1948, page 102 and Shadowitz, *The Electric Field*, 1975, page 384. An inducing E-field generated by the current in the primary (center wire) creates an exactly equal and opposite E-field due to electric charges (electrons) at every point in the secondary (braid). The result is that the *net* E-field in the braid is zero, just as it should be. These charges create the voltage distribution (charge gradient) that is measured by the voltmeter. That is, the voltmeter responds to the charges, but it ignores the inducing E-field. Ordinary voltmeters respond only to charges. This voltage increases uniformly along the length of the braid.

The statement in the article is correct. There is an E-field in the braid that is due specifically to the electronic charge distribution. If the braid were a perfect insulator (no free electrons) the induced E-field would also exist inside the material. The introduction of free electrons reduces the net E-field to zero, so the electrons are responsible for the zero E-field.

There is another way that the E field in a conductor can be made equal to zero. It is described in Note 9 on page 20 of the article. In this case, the voltage measured between any two points is virtually zero. This note is correct.

Coax Center-Wire Coupling

It is incorrectly believed by some that the center wire of a coax is relatively immune to any event that takes place on the outside of the braid. This is not true at all. A perfect example is to measure the impedance of the center wire of a balun, using a vector impedance meter. It is found that this impedance is almost identical to the impedance of the braid. That is, the center wire is interacting with the ferrite cores *through* the braid. In other words, the center wire, the braid and the ferrite constitute a very good conventional transformer with a 1:1 turns ratio. Also, the "Experiments" section gives a further demonstration that is unambiguous. In addition, the transformer works equally well from center-wire-to-braid and from braid-to-center-wire.

A longitudinal EMF-producing E-field that is induced by some external cause—such as the near field of an antenna onto the outer surface of the braid—is transformer-coupled onto the center wire. This is true even though any braid current that might result flows only on the *outside* of the braid. The experiment in Fig 9 is one illustration of this. This easily demonstrated fact is perhaps not widely appreciated.

Because of the interaction of the center wire with the external environment, the claim in the article that an ungrounded coax line, *suspended in mid-air*, is basically no different from a two-wire transmission line, is quite correct. This runs contrary to the traditional beliefs of some, but the proof (Fig 9) is easy, and it is convincing.—*William E. Sabin, WØIYH,* —e-mail **sabinw@mwci.net**

A Solid-State R-390?

◊ Pete,[1] it is the R392 that runs on 24 V. It uses 6AJ5s in the gain and mixing stages—I think they are selected 6AK5s. Their g_m under those operating conditions is about 1500 µmhos, which is about one tenth that of the passing 40673. (As far as I can tell, no American maker of dual-gate silicon MOSFETs has made any in about five years. The tale they told me is that they were all using 2 inch silicon wafers with special processing, and the only supplier of those quit. There are some still available from Europe but the g_m tends to be even higher, and supply is often illusive.)

The R-392 has five or six IF stages to get its gain with little gain per stage. It would oscillate if devices with *real* gain were installed. A ham who lived near here 15 or 20 years ago played with that idea using low-gain junction FETs. Even then, I think he used coupling capacitors to replace some stages. His last name was Keyes and he went to Virginia from Iowa as I recall. I don't remember his first name or call sign. The tube R-392 works better if the plates are run from 32 V with the heaters at 24 V.—*Dr. Gerald N. Johnson, P E, KØCQ, RR1 Box 62, Gilbert, Iowa 50105; e-mail* **geraldj@ames.net**

[1]Reference Pete Traneus Anderson's Letter to the Editor in Oct 1997 *QEX*, which responded to "Synthesizing Vacuum Tubes," by Parker R. Cope, W2GOM/7 in Aug 1997 *QEX*.

October *QEX* Editorial

◊ My October *QEX* arrived today with the "Empirically Speaking" comment about the lack of Letters to the Editor from the general readership. E-mail makes it much easier to respond! I have been a *QEX* subscriber since Issue 1.

I usually scan each issue when it arrives, and if there is an article I find especially interesting, I read it. Most of the time however, I am so busy, and the articles are technical enough, that I usually give it a quick scan and put it on the shelf for later reading. Someday I hope to get back to a lot of what is in *QEX*. I wish there were more hours in each day, so I could spend a bit more time with some of the articles! I have found several times that *QEX* has been a good reference on something that I had no interest in when it was published. Keep up the good work—someday I will have time to go back and do a few things based on what has been published in *QEX!—Bill Kelsey, N8ET, Kanga US; tel 419-423-4604; e-mail* **kanga@mail.bright.net**; **http://qrp.cc.nd.edu/kanga/**

◊ I subscribe to *QEX* particularly because it is a technical magazine. Even though there is not something that strikes my fancy in every issue, I keep every issue so I can refer to it as my interests change. I agree with the idea of adding more "straightforward practical articles," but I do not want to see it degrade to the technical content I see in many ham publications (ie, how to connect coax to a dipole). While there is a need for these articles, I believe the intent of *QEX* is to cater to the interests of technically minded hams with articles that are just beyond the audience of *QST* and would be in *QST*. Some thoughts would be to reprint technical articles from nonUS ham publications as you previously have. Also, I always liked the surveys of foreign articles presented in the old *AMSAT Newsletter* by Kaz Deskur, K2ZRO. (Kaz, being multilingual, brought some good things that us single-language persons would have otherwise missed). Also, I believe the radio portion of the hobby has been overtaken, in some cases, by the digital side. I would like to see more RF oriented articles, as there is plenty of digital stuff in the other technical magazines. For example, your recent series on microwave antennas was great. In fact an untapped source of material maybe some of the fine local newsletters put out by local radio clubs and organizations. I always pick these up at ham stores when I get the opportunity to travel around the country. Some of these articles are written by "experts" who want to support their local newsletters, but do not want to be "bothered" to write for national magazines. I know none of this is earth shattering, but I hope it will help you form or modify your editorial policy. Best of luck in your new editing duties.—*Domenic M. Mallozzi, N1DM, e-mail* **DMallozzi@aol.com**

Valve Amplifier with Three or More Anode Lines!

◊ I am currently designing and building a 145 MHz power amplifier to use either an Eimac 3CX5000A7 or YC156 triode. The anode resonator was intended to be a λ/2 line, with the tube at the centre and the two ends of the line shorted to ground for RF—but not dc! The dimensions of the λ/2 line were designed by computer to have an optimal characteristic impedance[2] and to have the correct length for resonance at 145 MHz.

When the anode resonator was tested with a GDO, the resonant frequency was only 132 MHz, without any anode tuning or loading capacitors, which will reduce it further. This error is believed to be due to the inductance of the 92 mm (3.6 inch) long shorting bars at each end of the line. Attempts to reduce the anode line length to increase the resonant frequency were not practical, as the line length would have been less than 226 mm (8.9 inch)—the diameter of the Eimac SK-306 chimney. The method used to solve the problem may be of interest to others.

Instead of using a λ/2 line (which is in effect two identical λ/4 lines placed 180° apart), three λ/4 lines were tried, each placed 120° from the other, with the tube again in the centre. The resonate frequency increased from 132 MHz to 162 MHz, as each line then resonates with a third, rather than half, of the tube's output capacitance. Increasing the number of anode lines (N) to four, spaced 90° apart, increased the resonant frequency further to 190 MHz. A test with just one λ/4 line (N = 1), resulted in resonance at 95 MHz.

Given that frequency is proportional to

$$\frac{1}{\sqrt{C}}$$

where C is capacitance, and a fraction $1 / N$ of the tube's capacitance resonates with the anode lines, theoretically, the expected resonant frequency for multiple lines is proportional to

$$\sqrt{N}$$

If a single line resonates at f, we would expect two identical lines to resonate at 1.414 f, three lines at 1.732 f, four lines at 2 f and N lines at $\sqrt{N}$ f. There is good agreement between this theory and the frequencies measured with a GDO. As the number of lines is increased, the lines will eventually interact, and this theory would break down. When N is large, a cavity is formed. Increasing the number of anode lines improves the uniformity of the RF current from the tube anode, so there is an added benefit of three or more lines.

The amplifier has not been completed, so these results must be con-

sidered preliminary. The complete amplifier design will be offered for publication at a later date.—*Dave Kirkby, G8WRB, (QTHR) e-mail* **davek@medphys.ucl.ac.uk.**

[2]D. R. Kirkby, G8WRB, "Finding the Characteristics of Arbitrary Transmission Lines," *QEX*, Dec 1996, pp 3-10.

Printed Circuit Boards Available for DDS VFO

◊ FAR Circuits now sells PC boards for Curtis Preuss, W2BV's Direct Digital Synthesis VFO (*QEX*, July 1997, pp 3-7. The boards cost $7.50 each, plus $1.50 for shipping and handling of up to four boards. For more information, contact Fred at FAR Circuits, 18N640 Field Ct, Dundee, IL 60118-9269; tel 847-836-9148 (also voice mail and fax); URL **http://www.cl.ais.net/farcir/**).—*Bob Schetgen, KU7G, QEX Managing Editor*

Another Exchange About Screen Bias Supplies

From Hal Jones to Ian White:

◊ Congratulations on your article "Power and Protection for Modern Tetrodes" in *QEX* for Oct 1997. I enjoyed reading it.

Submitted for your consideration are the following comments regarding the caption and text for the block diagram on page 17:

"Fig 3—The Collins 30S-1 used two separate high-current supplies for the cathode and screen, with choke-input filters, but no other voltage regulation."

The text correctly refers to "...both the anode and the cathode/screen supplies..." According to Collins Radio, both the plate (anode) and the screen power supplies in the 30S-1 used tuned filter-chokes.[3]

Pappenfus, et al, states on page 227:

"Single-phase full-wave rectifiers are usually used for transmitters up to about 3 kW PEP output. A single-section filter with a tuned choke is generally best for these medium-power transmitters."[4]

This chapter examines using a tuned choke to reduce the value of input choke inductance needed for regulation. The authors conclude (on p 225) that the most effective means of reducing the transient response of the output voltage to sudden changes in the load is to use the lowest possible value of filter-choke inductance. *"Second in importance is increasing the value of filter capacitance as much as necessary."* (From p 226.)

Perhaps this is another example of those later designers missing selected features from the 30S-1, ie, the point that both power supplies in the 30S-1 used tuned filter chokes. (The "all-or-nothing deal" mentioned on page 17 of the article.)—*Hal Jones, W6ZVV, 2059 Kings Ln, San Mateo, CA 94402*

[3]Collins Radio Co, *Fundamentals of Single Side Band*, (Part No. 597 0331 00, 2nd ed.; Cedar Rapids: Collins Radio Co, 1 Sep 1959) p 13-29 and 30, Fig 13-18, 30S-1 Schematic Diagram.

[4]E. W. Pappenfus, Warren B. Bruene, E. O. Schoenike, *Single Sideband Principles and Circuits*, (New York: McGraw-Hill, 1964), Ch 15, "Power Supplies for SSB Amplifiers," pp 222-229.

Ian White replies:

◊ The choice between choke-input and capacitor-input supplies is usually not relevant to the screen-grid supplies that were the subject of this article. The Collins 30S-1 is an exception because its combined screen/cathode supply has to deliver the full plate current. Hal is correct that the Collins 30S-1 uses tuned chokes. However, all kinds of choke-input filters perform best with a steady load. With the fluctuating loads of SSB and keyed CW, the logic of the statements on pages 225 and 226 (quoted by Hal) is that the lowest possible value of choke inductance is zero—in other words, to use a capacitor-input filter with enough capacitance to prevent the voltage from falling significantly under full load ...but that would have been a different article.—*Ian White, G3SEK,* —**http://www.ifwtech.demon.co.uk/g3sek.** □□

Upcoming Technical Conferences

Southeastern VHF Society Technical Conference

The second annual Southeastern VHF Society Technical Conference will be held Friday and Saturday, April 3 and 4, 1998 in Atlanta, Georgia!

The dates and location of our inaugural conference proved so popular that we are staying with the same weekend and location in 1998—the Atlanta Marriott Northwest located between Atlanta and Marietta. A special conference room rate of $69 per night (plus tax) is available to attendees. For reservations, please call the Marriott at 1-800-228-9290.

SVHFS Award

The Southeastern VHF Society will present the SVHFS Award annually in recognition of exceptional contributions and service to the Society or the VHF/UHF community in general. Presentation of the first SVHFS Award will take place at the Saturday evening banquet. For more information on the SVHFS Award, please contact Steve Adams, K4RF (ex-WS4F), SVHFS President at **stevews4f@aol.com**

Call for Papers

Program Chairman Bob Lear, K4SZ, has issued a call for papers. If you are interested in making a technical presentation or having a paper published in the *Proceedings*, please contact Bob at: Bob Lear, K4SZ, PO Box 1269, Dahlonega, GA 30533; tel 706-864-6229; e-mail **k4sz@stc.net**

Antenna Measurements

Antennas will be measured Friday, April 3, starting with 144 MHz and working upward in frequency—amateur bands only, please! A maximum of two antennas per band, per individual may be tested. Please supply a female N connector or SO-239. Please indicate on your registration form the antennas you will bring. Walk-ups will be tested as time and space permits, so please preregister! For more information, or if you would like to test a commercial antenna, please contact Antenna Measurements Chairman Dale Baldwin, WBØQGH, at **wbØqgh@mindspring.com**

Noise Figure Testing

Noise figure testing will be conducted on Saturday, April 4. Attendees will be assigned a time for testing based on receipt of their conference registration form. For more information, contact Noise Figure Measurements Co-Chairmen Charles Osborne, WD4MBK, at **cosborne@pipeline.com** or Fred Runkle, K4KAZ, at **engineer@rightmove.com.**

In addition to technical presentations, antenna measurements and noise figure testing, activities will include: Friday evening flea market, Saturday evening banquet, SVHFS auction and a family program.

Please make plans now to attend this fun event! We are determined to make our second conference even better than our first—with your help! Register before March 14, 1998, and receive a discount!

For more information on the Southeastern VHF Society or our 1998 Conference, please visit our Web site at **http://www.akorn.net/~ae6e/svhfs.**—Tad Danley, K3TD; e-mail **K3TD@contesting.com**; TRA 4501, TARA 18, NAR 14020.

For Amateur Radio contesting, try **http://www.contesting.com.** For high power rocketry try **http://www.rocketryonline.com** and **http://tara.xyzzy.net.**

"Four Days In May" '98 QRP-ARCI Conference

QRP Amateur Radio Club, International (QRP-ARCI) proudly announces the third annual "Four Days In May"

(FDIM) QRP Conference commencing Thursday, May 14, 1998—the first of four festive days of 1998 Dayton Hamvention activities. Mark your calendar for this extra bonus day and register early for this not-to-be-missed QRP event of 1998. Amateur Radio QRP presentations, workshops and demonstrations will be the focus of the full day Thursday QRP Symposium to be held at QRP ARCI headquarters—the Days Inn Dayton South. Last year, this sold-out event had a "standing room only" crowd of 125 enthusiastic attendees. FDIM QRP Symposium attendees will start their day with a "wake-up" coffee social and then plunge into a full day of multimedia QRP presentations by renowned QRP authors and designers. Here is a sampling of the topics already on the slate:

- Antenna Feeders
- PCB Alternatives
- G3RJV "Six-Pack"
- Coherent CW
- Transistor Modeling
- Beyond the NE-602
- QRP PIC Designs

This first day will culminate with an evening QRP ARCI author social for folks to meet the QRP presenters. The QRP ARCI '98 QRP Symposium will be the talk of the Dayton Hamvention.

The QRP extravaganza continues with the annual Friday night QRP ARCI Awards Banquet, honoring QRP dignitaries for their service to the Amateur Radio community. QRP ARCI is proud to announce that this year's featured Banquet speaker will be Ade Weiss, WØRSP, one of the pioneers of the QRP movement, as we know it today. Following the Awards Banquet, will be the FDIM QRP vendor social, with prizes.

FDIM Saturday will be special this year, with an evening social for QRPers to meet the many regional North American and International QRP Club members—bring your banners! Saturday culminates with the annual QRP building contest sponsored by the NorCal QRP club. This year the contest will highlight awards for QRP rigs based on the venerable 2N2222 transistor—in honor of the invention of the transistor.

"Four Days in May"
'98 QRP Conference
Frequently Asked Questions:

QRP Symposium Presenters—Please submit your QRP technical manuscripts to FDIM '98 Technical Paper Chairperson Ken Evans, W4DU (848 Valbrook Ct, Lilburn, GA 30047; e-mail **w4du@bellsouth.net**). FDIM '98 QRP Symposium *Proceedings* will be available for sale during and after the conference for folks who are not able to attend the Symposium.

FDIM QRP Symposium Registration—Registration for the Thursday, May 14, 1998 FDIM QRP Symposium is $10 if prepaid by May 1, 1998, $12 after that date and at the door. "At the door" registration will be limited if we sell out. Please register early to guarantee a seat. Registration covers a full day of QRP Symposium activities, which include the QRP technical presentations, a "Six-Pack" printed circuit board kit, those famous "special" Symposium bag stuffers and an endless QRO coffee pot. The $10 registration fee also includes a complimentary copy of the FDIM '98 QRP Symposium *Proceedings*.

You can pay your $10 registration fee by US check, money order or international money order made out to "QRP ARCI." Send the fee and a SASE to Cam Bailey, KT3A, (FDIM Symposium Registration, PO Box 173, Mt Wolf, PA 17347) by May 1, 1998. E-mail **kt3a@juno.com** for information.

QRP-ARCI Awards Banquet—This not-to-be-missed Friday May 15, 1998 event will be hosted by FDIM Banquet Chairperson—Scott Rosenfeld, NF3I. Please send your $22 banquet ticket fee (US check, money order, international money order payable to "QRP ARCI") and a SASE to Scott (QRP ARCI Banquet Tickets, 4015 Sparrow House Ln, Burtonsville, MD 20866-1333). E-mail **ham@w3eax.umd.edu** for information.

FDIM QRP Vendor Social—This tradition started at FDIM '96. It's a special evening set aside to officially introduce our QRP vendors from around the world. All may attend this wonderful Friday, May 15, 1997 evening social. Jim Stafford, W4QO, QRP ARCI Vice President, will be the gracious host this year. QRP Vendors who desire registration information please contact Jim (QRP Vendor Evening Chairperson, 11395 West Rd, Rosewell, GA 30075; or e-mail **w4qo@amsat.org**).

QRP ARCI FDIM Headquarters—The Days Inn Dayton South will be the 1998 FDIM QRP headquarters. QRP ARCI has secured a special block of rooms for our guests. Please contact Hank Kohl, K8DD (e-mail **k8dd@tir.com**), regarding availability of rooms. Hank's address is 1640 Henry St, Port Huron, MI 48060.

On behalf of the QRP ARCI team I invite you all to join us for the QRP Event of 1998: "Four Days In May" '98 QRP Conference at the 1998 Dayton Hamvention. See you all there!—Bob Gobrick, NØEB, FDIM '98 Publicity Chairperson (e-mail **rgobrick@worldnet.att.net** or **rgob@tengizchevroil.com**). □□

ARRL Educational Activities Department Announces New Videos

Two new ARRL Continuing Education Workshop videos now are available from the ARRL Video Library:

• *Digital Signal Processing* (Volume 10; 1993; 4.5 hours) is Order #6605; price $21. This in-depth workshop tells how to get started in Amateur Radio projects involving digital signal processing (DSP). It's taught by ARRL Senior Engineer Jon Bloom, KE3Z.

• *Computer-Aided Antenna Design* (Volume 11, 1996; 4 hours) is Order #6613; price $15. This video introduces antenna system design with computers. It's taught by ARRL Senior Assistant Technical Editor Dean Straw, N6BV.

Both videos are best used in small study groups; and they may be shown in four installments. To order by credit card, call Jean Wolfgang, WB3IOS, 860-594-0219 (9 AM until 4:30 PM Eastern Time). Credit-card orders received by 2 PM Eastern will be shipped the next day. Send orders paid by check to ARRL EAD, 225 Main St, Newington CT 06111. (For a complete list of EAD videos, send a request and an SASE to ARRL EAD.)

□□