QEX

November/December 2006

A Forum for Communications Experimenters

Issue No. 239

25th Anniversary Edition

W1GHZ uses a surplus waveguide slotted line to design transition blocks for waveguide to coaxial cable transmission lines.

ARRL The national association for AMATEUR RADIO
225 Main Street
Newington, CT USA 06111-1494

QEX

QEX (ISSN: 0886-8093) is published bimonthly in January, March, May, July, September, and November by the American Radio Relay League, 225 Main Street, Newington, CT 06111-1494. Periodicals postage paid at Hartford, CT and at additional mailing offices.

POSTMASTER: Send address changes to: QEX, 225 Main St, Newington, CT 06111-1494
Issue No 239

Harold Kramer, WJ1B
Publisher

Doug Smith, KF6DX
Editor

Larry Wolfgang, WR1B
Managing Editor

Lori Weinberg, KB1EIB
Assistant Editor

L. B. Cebik, W4RNL
Zack Lau, W1VT
Ray Mack, WD5IFS
Contributing Editors

Production Department

Steve Ford, WB8IMY
Publications Manager

Michelle Bloom, WB1ENT
Production Supervisor

Sue Fagan
Graphic Design Supervisor

Devon Neal, KB1NSR
Technical Illustrator

Joe Shea
Production Assistant

Advertising Information Contact:

Janet L. Rocco, W1JLR
Business Services
860-594-0203 direct
860-594-0200 ARRL
860-594-0303 fax

Circulation Department

Cathy Stepina, *QEX Circulation*

Offices

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

Subscription rate for 6 issues:

In the US: ARRL Member $24, nonmember $36;

US by First Class Mail: ARRL member $37, nonmember $49;

Elsewhere by Surface Mail (4-8 week delivery): ARRL member $31, nonmember $43;

Canada by Airmail: ARRL member $40, nonmember $52;

Elsewhere by Airmail: ARRL member $59, nonmember $71.

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

In order to ensure 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

Paul Wade, W1GHZ, shows us how to design waveguide to coaxial cable transition blocks. The cover photos show a waveguide slotted line, which Paul uses to measure the complex impedance inside the waveguide, along with some resulting transitions.

Features

Columns

Nov/Dec 2006 *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 interest 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 the 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 three 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:

ARRL, 225 Main Street, Newington, CT 06111 USA.

Telephone: 860-594-0200

FAX: 860-594-0259 (24-hour direct line)

Officers

President: JOEL HARRISON, W5ZN
528 Miller Rd, Judsonia, AR 72081
Chief Executive Officer: DAVID SUMNER, K1ZZ

The purpose of *QEX* is to:
1) provide a medium for the exchange of ideas and information among Amateur Radio experimenters,
2) document advanced technical work in the Amateur Radio field, and
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 letters for publication in *QEX* should be marked Editor, *QEX*.

Both theoretical and practical technical articles are welcomed. Manuscripts should be submitted in word-processor format, if possible. We can redraw any figures as long as their content is clear. Photos should be glossy, color or black-and-white prints of at least the size they are to appear in *QEX* or high-resolution digital images (300 dots per inch or higher at the printed size). Further information for authors can be found on the Web at **www.arrl.org/qex/** or by e-mail to **qex@arrl.org**.

Empirical Outlook

Doug Smith, KF6DX
kf6dx@arrl.org

Twenty-Five Candles Burning Brightly

It wasn't hard to select an editorial topic this time. It's *QEX*'s 25th birthday! You'll find me near the center of the celebration, but it isn't about me. It's about us.

Editors often write in first-person plural, as if writing for the entire staff. For this occasion, that form shall embrace not only staff but also writers, readers, artists, printer — the whole lot of us. I'll unabashedly exercise my prerogatives, as head cheerleader, in the singular.


QEX
December
1981
Number 1
The ARRL Experimenters' Exchange

***QEX*, Number 1, December 1981.**

You're reading the best technical journal in its field. *QEX* rose to that status through hard and intelligent work of the same kind that sustains it now. Like many other magazines, we began somewhat modestly in an era when more emphasis was placed on technical know-how than is placed on it today. It's irrefutable what a difference 25 years can make. I don't equivocate because judging by the results in the publishing world, things have certainly changed.

Ham radio (1968-1990) had a 22-year run. Billed as *ham radio*'s logical successor, *Communications Quarterly* (1990-1999) ran for nine. *73* (1960-2003) survived for 43 years. *QST* (1915-) is still going strong, of course, and looks better than ever in my opinion; but its formerly heavy technical content has diminished by choice. *CQ* (1945-) is still percolating along after more than 60 years but its technical material has also declined. Its sister publication, *CQ VHF*, faltered in 1999 but came back in 2002. *QEX* had its little stumble, too, in 1997 when we missed six issues. One of the given reasons was lack of sufficient article submissions.

I'm happy to report that we're in good shape both financially and editorially now. No, we don't make a big monetary gain and our circulation hasn't grown significantly in recent years, as we'd hoped; but we're not slipping either. We're fortunate to get articles and readers from all corners of the globe. As opposed to some of those other periodicals, our technical content has steadily and dramatically improved. *QEX* has garnered six of the last eight ARRL Doug DeMaw, W1FB, Technical Excellence Awards. We are justifiably proud of all the great articles you've written! The last few issues have been great and I think you'll find this Silver Anniversary issue particularly outstanding.

Sometimes it's not enough to produce a good magazine, though. Folks can't be expected to read it if they don't know it exists. We need to promote ourselves because just about the only advertising we get is in ARRL publications. Ask me for complimentary issues and anything else that may entice new readers to subscribe. I'll fulfill all reasonable requests.

Since I mentioned advertising, let me ask whether any of you intrepid entrepreneurs could use a four-color, full-page cover ad. It's a lot less expensive than you might think and you'll be reaching a readership that includes leaders in communications from academia, industry and Amateur Radio. Check out our rate sheet at **www.arrl.org/ads/#QEX**.

Aside from being head cheerleader, a magazine editor must play other roles from time to time, such as referee, business manager, grammar cop, jury and chief cook and bottle washer. From this point in the middle of my ninth year, I have to write that it's been a blast! A highlight for me personally is getting acquainted with so many of you astute technical types whose thoughts and projects light the way for the rest of us. Only one thing is missing: I can't recall having received even a single article submission or letter from a female writer. Where are you gals hiding?

Casting a gaze to the future, we must ask ourselves what else we need in *QEX*. Our stated policy has been one of many small improvements but big improvements don't hurt. How about more microwave? More simple construction articles? We've decided to press again for more practical information about new components and software of interest to communications experimenters. Please send us your announcements.

I extend hearty thanks to you who've supported this forum through the years. We continue to gobble up the material that helps us meet our stated purpose (in column 1 at the left). If your last submission wasn't accepted, please refer to our Author's Guide at **www.arrl.org/qex/#aguide** and try again. Drop me a line anytime. My door is always open. QEX

Measurement of Soil Electrical Parameters at HF

The author describes the technique he uses to measure soil conductivity and relative dielectric constant over a range of frequencies on the HF bands.

Rudy Severns, N6LF

Introduction

Modeling of antennas over real ground requires at least a reasonable guess of the values for the soil conductivity (σ) and relative permittivity (ε_r), also referred to as "relative dielectric constant." Unfortunately these numbers are usually not readily available. From broadcast (BC) work we have charts of ground conductivity covering large areas, but these numbers give only σ, not ε_r. In part, the absence of ε_r data is because, for sites where you would want to build a BC station, the soil characteristics are usually dominated by σ, and ε_r has only a second-order effect. This is often true at BC frequencies but is usually not the case in all but the most conductive soils at HF. In addition, the values for σ will be different between BC frequencies and HF. Another problem is that the BC ground conductivity charts cover much too large an area to take into account the details of local ground variation, which can deviate greatly from local averages.

It would appear that the best approach is to simply measure your local soil characteristics at the frequencies of interest. Unfortunately, this is much easier said than done. None of the known methods is anywhere near perfect, and many are difficult to implement. In fact there is a school of thought that the problem is impossible and we should not waste our time worrying about it. I don't share that view as a general proposition, but it is not without some justification, given the difficulties involved.

There are common situations where the values for the soil constants (which are anything but constant!) are really not very important for modeling purposes. For example, for horizontal polarization with antenna heights above $^1/_4$-λ, the numbers are not very critical for determining feed-point impedances, near-field losses or the formation of the far-field radiation pattern. Another case would be for vertical antennas where one has the space, money and patience to lay down a large number of long radials. With this brute force approach, the near-field ground loss can be made arbitrarily small regardless of the soil, and you really don't care what the ground constants are, at least from a local loss point of view. The far-field pattern, however, is still just a guess without real data.

If your space and/or financial resources are more limited, then a modestly accurate estimate of your ground characteristics will allow you to design a ground system that minimizes the loss within the constraints of the space and resources you do have. There is also the situation that arises fairly often on 80 m. On that band a $^1/_2$-λ is about 130 feet, which is not an exceptionally tall tower for amateur use. Horizontal gain antennas are certainly practical but they're not easy, being large and relatively expensive. The option is to go to a vertical array, which may be easier. Accurately predicting the performance of a possible vertical array in comparison to a competing horizontal array requires at least a reasonable guess for the ground characteristics on that band. The decision as to which way to go may depend, at least in part, on having a reasonable estimate of the ground characteristics.

Another problem with present ground modeling practice is the assumption that soil parameters, whatever they may be in a given location, are constant over frequency. For example, for a given soil, the assumption is that ground constant values at 160 m are the same as the values at 20 m. That's not the case. As pointed out by Bob Haviland, W4MB, and in many professional papers, ground parameters at HF vary substantially with frequency.[1,2]

There is a need for a practical method to estimate soil parameters at HF for amateurs. By practical, I mean a mechanically simple test apparatus and measurement equipment

[1]Notes appear on page 8.

PO Box 589
Cottage Grove, OR 97424-0025
rudys@ordata.com
www.antennasbyn6lf.com

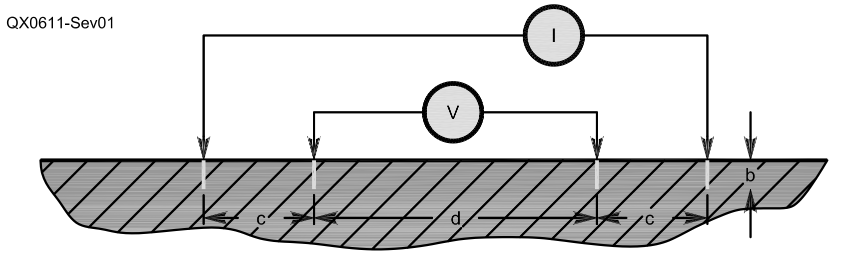


Figure 1 — This diagram shows the Wenner array, the traditional method used by amateurs to measure ground parameters.

no more advanced than an AEA or MFJ impedance analyzer. Fortunately, great accuracy is not required and it makes little difference in the modeling if the values are off by 25%. This article will discuss one particular approach, the use of ground probes, which at least approximate this ideal.

Soil Parameter Measurement

There are many ways to measure ground parameters. Each has advantages and limitations. None is the perfect answer. The traditional method used by amateurs is the Wenner array (or similar variations), which uses four probes in line as shown in Figure 1, excited with line-frequency ac.[3, 4, 5] This approach can give a good estimate of ground conductivity at 50 to 60 Hz, and by varying the spacing of pairs of probes, can be used to define subsoil layering. It gives no information on ε_r, however. A measurement of this type provides only a lower bound on soil conductivity, which will be higher at HF.

Another technique, frequently used in BC work, is to measure the rate of decrease of the E-field intensity as you go away from the antenna on a radial line. It is possible, by some judicious curve fitting to the measured data, to infer the average ground conductivity along the measured line. This is a reasonable approach at BC frequencies, where the soil characteristics are dominated by the conductivity. At HF, however, the soil is both resistive and capacitive. Typically, when trying this technique at HF, more than one pair of parameters (σ and ε_r) may generate curves that fit the data. This ambiguity is a problem. In addition, the measurements need to be made at some distance from the antenna where there are significant amplitude differences between measuring points and so do not give a very good idea of the ground characteristics within a $^1/_2$-λ of the base. Information on ground parameters close to the antenna is needed for ground system design, especially in the initial design stages for a new vertical.

A technique that would seem to fit our requirements is to insert a probe into the soil and measure its impedance.[6] In the simplest case the probe is basically just a capacitor, and the ground parameters are inferred from the change in impedance of this capacitor from when the probe is in air and to when it's in soil. This approach can yield a detailed characterization of the soil in the immediate area of the antenna and at distance also. A basic limitation of this procedure is that it is usually not possible to use a probe that reaches very far down into the soil. The result is characterization mainly of the top few feet of soil, which is usually substantially less than the skin depth. By making measurements at many spots over the area of interest the probe method can give a very good picture of the lateral variation of soil parameters. We know that the properties will also vary vertically (variations in moisture content, stratification, and so on) and we would like to know the variations down to a skin depth. It is possible to take a surface measurement, then dig down three feet or so in the same spot and reinsert the probe in the undisturbed soil at that level and make another measurement. This can be repeated until sufficient depth is achieved. That, however, defeats our goal of keeping the process simple, and is not practical for large-area surveys.

Is a fairly accurate characterization of only the top layer of soil of any real use? Certainly that's debatable but I think it is worth doing. There will be cases where the soil characteristics change slowly and the probe measurements are pretty close. It is also possible to have an entirely different strata a few feet down, with completely different characteristics. It probably is a good idea to dig one test hole as suggested above, to get a feeling for the local stratification and then do a survey with surface probes in the near area. In any case, I think probe measurements are a vast improvement over nothing but we should not be fooled into thinking the results are exact. Like everything in modeling, the information has to be used cautiously.

Monoprobe Technique

This method uses an impedance analyzer to measure the impedance of a single ground probe with a ground screen, as shown in Figure 2. The ground screen can be either square or circular, with a radius greater than the length of the longest probe. Rupar used a copper sheet for the ground screen.[7] I initially used a sheet of $^1/_8$-inch-thick aluminum, but a large metal sheet is awkward to work with and I found that a piece of ½ inch galvanized hardware cloth (as shown in Figure 3) worked just as well. Note the weights on the screen. The hardware cloth is flexible, and the weights are used to keep it in contact with the soil. This is an advantage if the ground is a little uneven in that the flexible screen may fit it better, minimizing any air gap between the screen and soil. More on this later. Anything will do for weights; bricks or rocks are fine. The flexibility of the hardware cloth means you can roll up the wire to make an easier package for

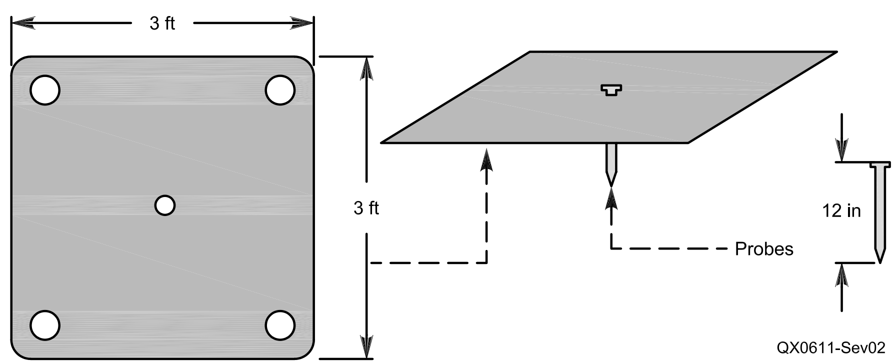


Figure 2 — This illustration shows the construction of the monoprobe. A three-foot square of half-inch hardware cloth forms the base. A hole cut in the center provides space for the probe to go through the screening without contacting the wire.

Figure 3 — This photo shows the monoprobe in use. Note the circular blocks used to hold the hardware cloth in contact with the ground, and the AEA impedance analyzer connected to the probe in the center.

carrying. The example in Figure 3 shows an AEA complex impedance analyzer being used. An MFJ-259B or other impedance analyzer will work just as well.

Figure 4 shows examples of typical probes (12 inches and 18 inches long). The crossbars in this example are phenolic, but any reasonable insulating material will work fine, even wood. The crossbar is there to help push the rod into the ground and pull it back out. The rod is inserted through a small hole (1 inch or so) in the screen and pushed down until the crossbar is pressed firmly onto the screen. The rods shown are brass but that's not essential. I just happened to have some $^3/_8$ inch brass rod stock on hand. For later experiments I found some inexpensive $^7/_{16}$ inch aluminum rod at a scrap yard. Anything from $^1/_4$ inch to $^1/_2$ inch should work fine. In fact, you can even use square rods if you wish. You can find suitable rod stock at most hardware stores. The larger diameters make for more sturdy probes, with a little more capacitance, but they may be harder to push into the ground. The presence of the thin insulating layer of oxide on aluminum rods has essentially no effect on the measurements.

Initially I threaded the top of the rod and the crossbar, then screwed them together, adding a nut on the top for a connection. You don't need to be so fancy. Later on I just drilled a tight fitting hole in the crossbar, drove the rod into it and added a cross 6-32 machine screw to hold it and to provide an electrical contact.

The impedance is measured between the top of the rod and the ground screen as shown in Figure 5. Note that I have used a lead from the top of the rod to the analyzer and a ground clip on the analyzer to connect to the screen. You could also mount a coaxial connector on the screen with a lead going to the top of the rod. The choice of which way to go affects the stray inductance and capacitance and is discussed in the appendix in the context of probe calibration.

OWL probes

The OWL (Open Wire Line) probes are simply two parallel rods and a crossbar without a ground screen, as shown in Figure 6.[6] The impedance is measured between the tops of the two rods. For a battery powered impedance analyzer like an MFJ, the measurement is floating (once you take your hands off the instrument!) and no balun is needed. If you want to use a more advanced analyzer, such as the Ten-Tec TAPR or N2PK vector network analyzers, with a cable, then a balun would be a good idea. I made up a test balun, which is included in the Figure 6 photo. I used a Fair-Rite FT240-43 ferrite core. This is standard core available from Amidon. The winding is a 3 foot length of RG-58 with BNC connectors at the ends. This length results in about 12 turns, and should give adequate isolation down to 1 MHz.

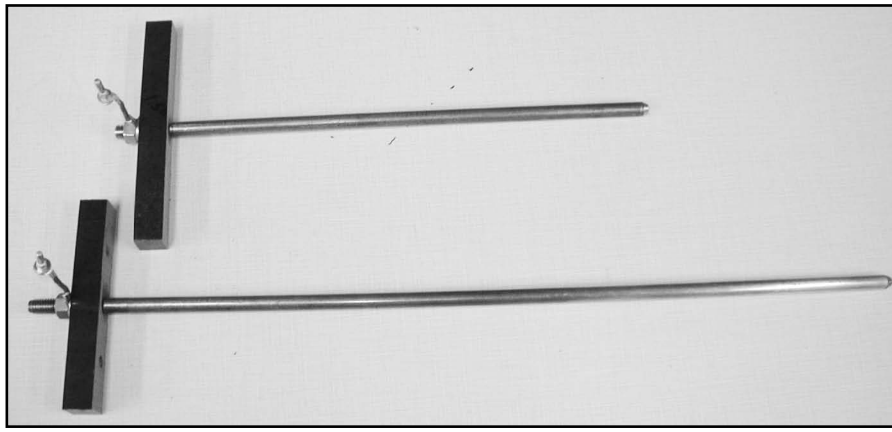


Figure 4 — Examples of 12-inch-long and 18-inch-long probes for use with the monoprobe technique.

Figure 5 — This photo shows the details of how the impedance analyzer is connected to the probe and ground screen.

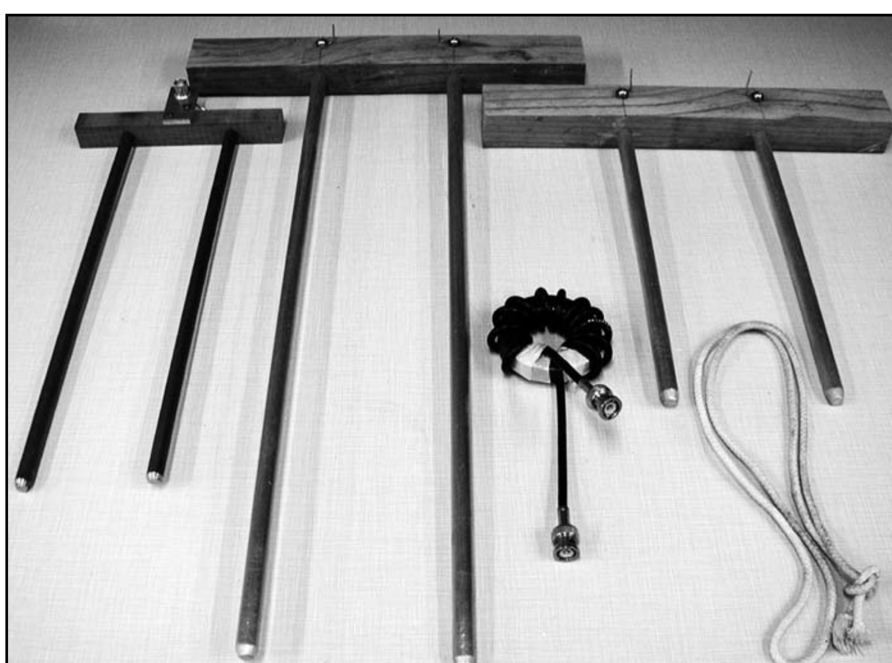


Figure 6 — Here are three open wire line (OWL) probes, along with a balun and a loop of rope used to pull the probes out of the ground after the measurements have been made.

The probes with 4 inch spacing use clip leads like those shown in Figure 5 but the 3 inch spacing probe (the probe on the left in Figure 6) has a BNC connector on the crossbar, to which the balun is connected. There is nothing magical about either arrangement.

Figure 6 also shows a vital piece of equipment: the cord! Before pushing a probe into the soil it's a really good idea to loop the cord around the crossbar. You will use it to pull the probe out of the ground. In hard soils, with the bar pressed against the ground, getting the probe out without the cord can be a chore. It also helps to put a handle on the cord.

I'd like to emphasize that the diameter, spacing and length of the probe rods is not critical. The only thing you must do is to measure or calculate the probe capacitance (as shown in the appendix) for your particular probe. Larger diameter rods and closer spacing result in lower measured impedances. With the AEA and MFJ analyzers, measurements of impedances above 200 Ω or so become less reliable and it is better to work with lower impedances. In soils with poor conductivity the measured impedance, with the same probe, will be higher than in more conductive soils, and it may be better to use closer rod spacing.

Choosing Between a Monoprobe or an OWL Probe

Both types of probes will work just fine, but each has advantages and disadvantages. The single probe is much easier to insert than a double probe. There is also the issue of keeping the rods parallel with the OWL. If the rod spacing varies between air and in the soil then the calibration of C_0 will be off. Given the modest accuracy required for ham applications this is usually not a problem. The OWL is much more compact to carry around because you don't need the large ground screen and weights to hold it down. That's a very practical advantage!

The monoprobe is influenced by a much larger volume of soil and provides an average over that volume whereas the OWL pretty much characterizes a small cylinder of soil. The monoprobe measurement is intrinsically unbalanced whereas the OWL may require a balun or other isolation for measurement with non-isolated instruments.

In the end, either will work. You just have to decide what suits you.

Taking and Reducing the Impedance Data

The procedure is very straightforward. You simply lay the screen on the ground if using a monoprobe, insert the probe into the soil and record the impedance reading on the analyzer at each frequency of interest. In the case of an OWL probe, you simply insert the probe into the ground up to the bar and make an impedance measurement. This should only take a few minutes and then you move on to the next points, recording the impedance measurements as you go. You can cover a lot of ground in an hour or so.

The next step is to convert the impedance readings to σ and ε_r using the equations given below. Putting these equations into a Microsoft *Excel* spreadsheet makes the whole process very painless.

The impedances can be in either of two formats: R + *j*X or magnitude (|Z|) and phase angle (θ). The equations for converting the measured impedances using R and X are:

$$\sigma = \frac{8.84}{C_0}\left[\frac{R}{R^2 + X^2}\right] \quad \text{(Eq 1)}$$

$$\varepsilon_r = \frac{10^6}{2\pi f_{\mathrm{MHz}} C_0}\left[\frac{X}{R^2 + X^2}\right] \quad \text{(Eq 2)}$$

If you prefer to work with |Z| and θ, the equations take the form:

$$\sigma = \frac{8.84}{C_0}\left[\frac{1}{|Z|\sqrt{1+\tan^2\theta}}\right] \quad \text{(Eq 3)}$$

$$\varepsilon_r = \frac{10^6}{2\pi f_{\mathrm{MHz}} C_0}\left[\frac{\tan\theta}{|Z|\sqrt{1+\tan^2\theta}}\right] \quad \text{(Eq 4)}$$

where C_0 = capacitance in pF of the probe in air. This can be either measured or calculated quite closely, as shown in the appendix. Frequency is in MHz and impedances are in ohms.

These equations assume the probe is simply a capacitor. If you make the probe longer at a given frequency, or push the measurement frequency up for a given probe length, there comes a point where the probe is no longer simply a capacitor, it becomes an antenna, buried in the soil. It can still be used but data reduction is more complex.

Some Actual Measurements

Now it's time to look at actual measurements taken on my property. Tables 1 and 2 show typical impedance measurements taken at two different locations, and their reduction to σ and ε_r. The data in Tables 1 and 2 is graphed in Figures 7 and 8.

Because of the relatively poor accuracy of the AEA analyzer, the graphs are a bit "lumpy."

Table 1

18 Inch Monoprobe, C_0 = 7.41 pF. On my antenna hill with an AEA-CIA analyzer.

Frequency (MHz)	*Resistance (Ω)*	*Reactance (Ω)*	*Conductivity, σ (S/m)*	*Relative Permittivity, ε_r*
1	129	–134	0.0044	83
2	83.3	–95.8	0.0062	64
3	66.3	–76.2	0.0078	53
4	56.7	–65.2	0.0091	47
5	51.5	–57.2	0.010	41
6	46.2	–47.8	0.012	39
7	40.2	–46.2	0.013	38
8	35.1	–40.4	0.015	38

Table 2

Four Inch × Nine Inch OWL, C_0 = 2.71 pF. In my backyard with AEA-CIA, no balun.

Frequency (MHz)	*Resistance (Ω)*	*Reactance (Ω)*	*Conductivity, σ (S/m)*	*Relative Permittivity, ε_r*
1	176	–137	0.0042	59
2	123	–119	0.0050	44
3	95.2	–98.5	0.0061	38
4	83.4	–86.3	0.0069	32
5	77.7	–75.0	0.0079	28
6	73.0	–63.4	0.0093	24
7	60.9	–60.9	0.0098	25
8	54.7	–52.8	0.011	25

In Figure 7 I have smoothed things out by inserting a linear trend line, which fits quite well. The lumpiness is typical using this class of instrument for measurement, but the lumps are still small enough not to matter. The data in Tables 1 and 2 is a bit sparse, but taking a large number of closely spaced data points and then smoothing with a trend line works even better.

You may notice that in Figure 3, the grass has been dug away so that the screen is in direct contact with the soil. I made measurements with and without the grass to see what the effect the grass would have. The results are shown in Figures 9 and 10.

As you can see from the graphs, the presence of the grass doesn't have much effect on the conductivity measurements, but does substantially affect the ε_r measurements. What the grass does is to insert a layer of air under the screen, which reduces the effective capacitance. That, in turn, reduces the value for ε_r. This is not a big issue but you should at least take a string trimmer and cut the grass as low as possible. If you are using an OWL probe then the effect of the grass is very small if the probe is pushed firmly down against the ground.

Another concern is the effect of using different probe lengths. The moisture in the very uppermost layer of soil responds rather quickly to weather conditions. When it rains, σ and ε_r go up and, when things dry out, σ and ε_r fall. This rate of variation with time and depth depends on the soil itself but for the most part the soil characteristics respond much more slowly at depths beyond 12 inches or so. This effect on measurements is illustrated in Figure 11. The soil at W6XX is quite sandy and the top layer dries out fairly rapidly. We can see this in the graph. σ is substantially lower in the upper layer which is being measured by the short probe. The longer probes reach down into soil that dries much more slowly, and as you can see the two longer probes give essentially the same data. A close look at the 24 inch probe data line illustrates a limitation mentioned earlier on probe length. As the probe is made longer the current distribution along the probe is no longer essentially constant. Instead of behaving like a simple capacitor (which Equations 1 and 2 assume) it is starting to act like an antenna. Notice how the 24 inch probe curve starts to bend over at the higher end. This can be corrected by using more complex equations for the data reduction, but for most users that may be more trouble than it's worth. The usable range is still above 40 m. Very high conductivity soils may require shorter probes.

Comments on Ground Data

The conductivity graph (Figure 7) has an important feature: the ground "constants" are not constant at all with frequency. It is very typical in the HF region for σ to increase with

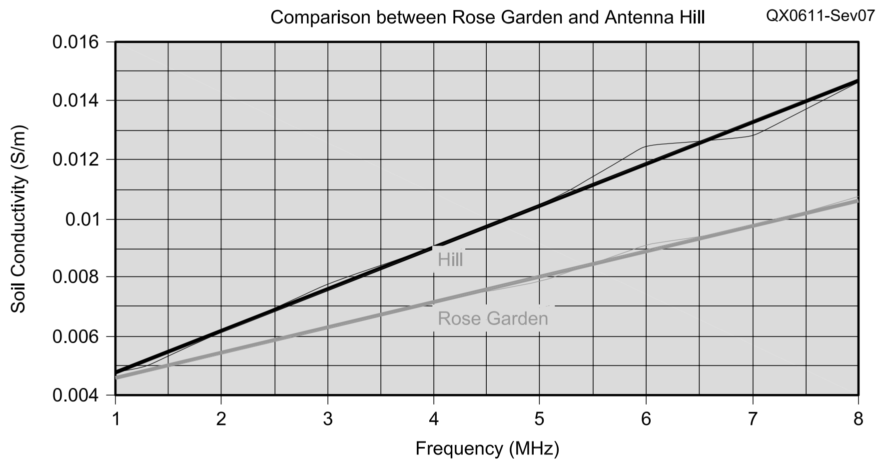


Figure 7 — This graph compares the conductivity I measured in two areas of my property.

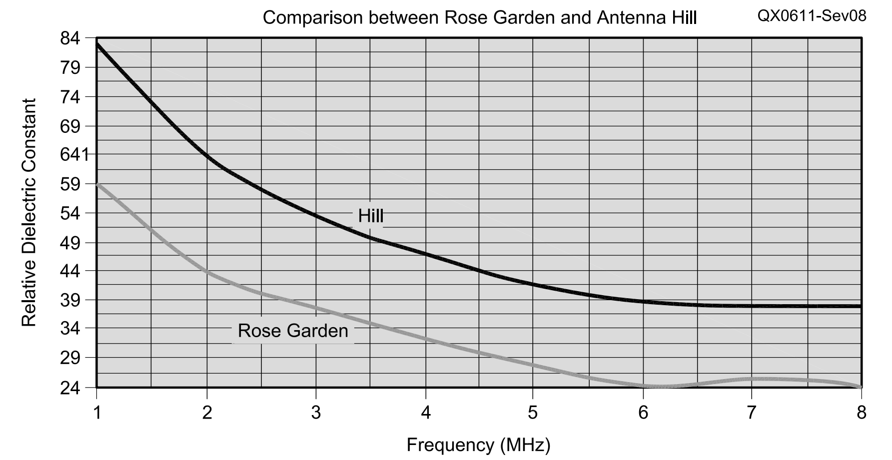


Figure 8 — The relative dielectric constant values on the antenna hill and in the rose garden, based on the measurements taken from 1 to 8 MHz.

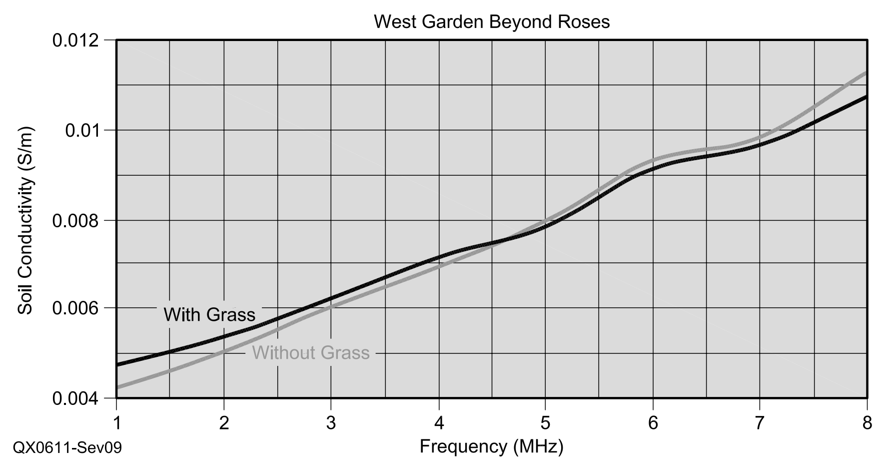


Figure 9 — I measured soil conductivity in an area of grass in my west garden, beyond the roses. Then I dug up a patch of grass and repeated the measurements to estimate the effects of grass.

frequency. In addition, as shown in Figure 8, ε_r is not constant either and tends to decrease as you go up in frequency to about 5 MHz, but then stabilize above that. The general shape and trends displayed in Figures 7 to 11 agree very well with those seen in the large body of professional work on ground parameter values.

At both sites in Figure 7, σ corresponds to what is generally called "average ground." Average ground is usually defined as σ = 0.005 S/m and ε_r = 13. In Figure 8, however, ε_r is much larger than 13, especially below 5 MHz. This is particularly characteristic of soils with a lot of clay particles. For many years there was a great deal of controversy over the large values of ε_r measured at low frequencies. The consensus is that it is very real. The following quote is from the King and Smith book, *Antennas In Matter*, which is considered a definitive work:[8]

"For some time, the high values of permittivity and the dispersion at these lower frequencies were thought to be artifacts of the measuring procedure; that is, it was thought that they were caused by electrochemical effects at the interface between the metallic electrodes and the sample of rock or soil. Measurements made using several different materials for the electrodes, however, indicate that there is a high permittivity associated with the geological material apart from any electrode effects."

Summary

How should we use the numbers we get? First, I try to take my readings at the end of the driest part of the year. Because both σ and ε_r are strong functions of soil moisture content, measuring near the end of the dry season will give you a conservative estimate. One exception I make is for my 80 and 160 m antennas, which I normally only use during the winter, which is definitely the wet season in Oregon. I use the winter ground parameters for these bands. Second, I average the readings found at different places over the site.

These are the values I use when designing a new antenna. Am I kidding myself? Well, perhaps, but I find it hard to believe that I am worse off than if I simply guessed and took the traditional value for mountains of σ = 0.001 S/m and ε_r = 5, which would appear to apply in my location even though these values are much lower than what I measure at my QTH.

I think that ground probe measurements are worth doing and I use them, but with care. Of course we would like even better methods, and in fact a number of workers are trying to find better ways.

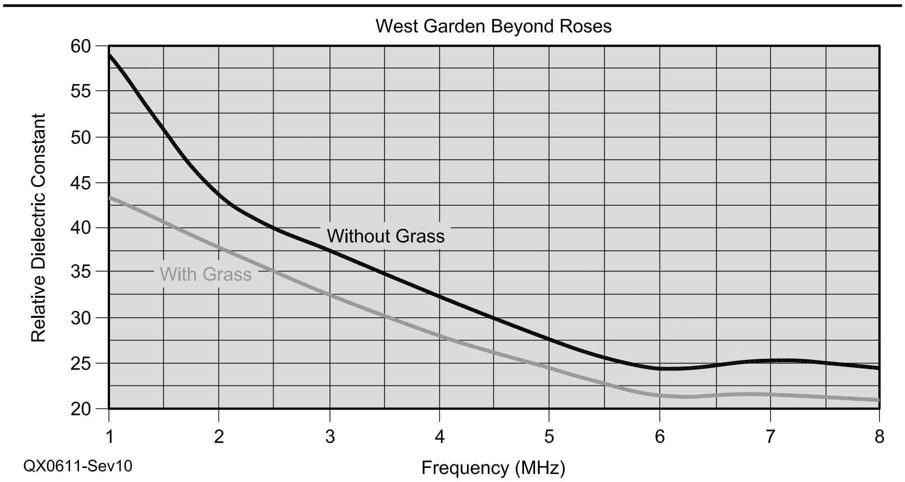


Figure 10 — This graph compares the relative dielectric constant as measured with and without grass.

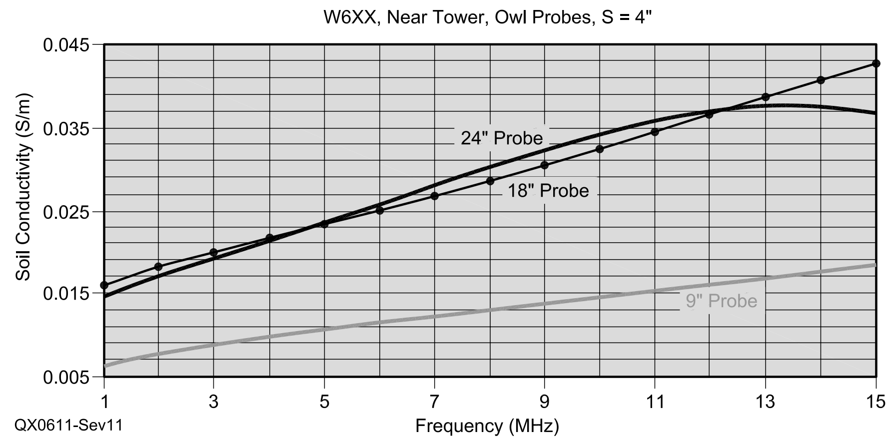


Figure 11 — This graph shows the data collected by Pete Gaddie, W6XX, near his tower, using three different-length OWL probes.

Acknowledgement

Pete Gaddie, W6XX, and I have been building and testing a variety of ground probes and much of what is in this article comes from our results. I would like to acknowledge the great assistance Pete has given me in supplying technical papers, extended discussion on technical details, measurement advice and a great deal of parameter measurement on his own. Thanks, Pete.

Notes

[1]R.P. Haviland, W4MB, "Ground Parameters for Antenna Analysis," *ARRL Antenna Compendium 5*, 1996, pp 96-100.

[2]Dean Straw, N6BV, *ARRL Antenna Book*, 20th ed, 2003, pp 27-36 through 27-41.

[3]Jerry Sevick, W2FMI, "Measuring Soil Conductivity," *QST*, Mar 1981, pp 38-39.

[4]*IEEE Guide for Measuring Earth Resistivity, Ground Impedance and Earth Surface Potentials of a Ground System*, IEEE Std 81-1983, Mar 11, 1993.

[5]*IEEE Guide for Measurements of Electromagnetic Properties of Earth Media*, IEEE Antennas & Propagation Society, IEEE Std 356-2001.

[6]George Hagn, "Ground Constants at High Frequencies (HF)," *Proceedings of the 3rd Annual Meeting of the Applied Computational Electromagnetics Society*, Mar 24-26, 1987.

[7]Michael A. Rupar, "Theoretical and Experimental Investigation of the Impedance of a Vertical Monopole Over Perfect, Imperfect and Enhanced Ground Planes," *Naval Research Laboratory Publication NRL/MR/5550-97-7941*, 30 April 1997, pp B1-B8.

[8]R. W. P. King and Glenn S. Smith, *Antennas in Matter*, MIT Press, 1981, p 427.

[9]I. L. McNally, K6WX, "Notes on Ground Systems," *ham radio*, May 1980, pp 27-29.

[10]Glenn S. Smith, "Measurement of the Electrical Constitutive Parameters of Materials Using Antennas, Part 2," *IEEE Transactions on Antennas and Propagation*, Vol AP-35, No. 8, Aug 1997, pp 962-967.

[11]R. L. Smith-Rose, "Electrical Measurements on Soil With Alternating Currents," *Proceedings of the IEE*, Vol. 75, Jul-Dec, 1934, pp 221-237.

[12]Kenneth P. Spies and James R. Wait, "Determining Electrical Ground Constants From the Mutual Impedance of Small Coplanar Loops," *IEEE Transactions on Antennas and Propagation*, Jul 1972, pp 501-502.

[13]F. E. Terman, *Radio Engineer's Handbook*, McGraw Hill, 1943, New York, NY, Equation 140, p 118.

Rudy Severns, N6LF, is a retired engineering consultant in the design of power electronics, magnetic components and static power conversion equipment. He was first licensed in 1953 as WN7WAG. He has an Extra class license, a BS Engineering from UCLA and is the author of two engineering texts and over 80 technical papers. He is a fellow of the IEEE and a life member of the ARRL. Rudy was QEX *Editor from August 1997 to Sep/Oct 1998. His current Amateur Radio interests are antennas, particularly vertical phased arrays, and 160 m CW DXing. In earlier days he held first radiotelephone and second radiotelegraph licenses and served in the US Army as a special forces A-team radio operator. His other interests are sailing, skydiving and scuba. He is 69 years old and currently a graduate student at the University of Oregon, studying physics and mathematics.*

Appendix — Determining C_0

The value for C_0, which appears in both of the equations for σ and ε_r, is the capacitance of the probe in air. It has to be determined before the impedance data can be reduced to σ and ε_r. There are two ways to go about finding C_0: direct measurement and calculation. For the OWL probes C_0 can be determined very closely from the following equation taken from Terman:[13]

$$C_0 = \frac{3.677}{\log_{10}\left\{\frac{D}{d}\left(1+\sqrt{1-\frac{1}{\left(D/d\right)^2}}\right)\right\}} \ \frac{\text{pF}}{\text{ft}} \qquad \text{(Eq A1)}$$

where:

D = center-to-center distance between rods

d = diameter of the rods.

Dimension units for D and d must be the same but can be anything.

For an 18.5-inch OWL probe with D = 4 inches and d = 0.44 inches, this gives C_0 = 4.51 pF. Of course there will also be a small additional capacitance due to end effect. A later measurement gave C_0 = 4.83 pF, which indicates that the end effect adds about 10% to the calculated capacitance. Unfortunately, there isn't a similar simple expression for the monoprobe.

Measuring C_0 poses a problem because it is so small, typically less than 10 pF. I use an inexpensive L/C meter made by Almost All Digital Electronics, model L/C meter IIB, shown in Figure A1.

This meter operates at about 1 MHz. By being very careful to zero the instrument just before a measurement and taking great care not to change the layout between zeroing and measuring, I have found that this instrument does measure small values of capacitance very well. In the case of the OWL probes I always measured a value which was just a little bit higher than calculated, which is what you would expect taking end effect into account.

A direct measurement of a probe will give a capacitance that is the sum of C_0 and the part of the probe that sticks out of the ground and is connected to the impedance analyzer. This is a parasitic capacitance (C_p), which has to be subtracted from the total measurement. I determined C_p by building a dummy probe that is identical in all respects to the actual probe except that the portion of the rod or rods that would normally be in the soil is cut off. The mechanical layout for the part sticking out of the ground is carefully replicated and a direct measurement of C_p made. This is then subtracted from the total capacitance measurement for the probe. In principle C_p is in parallel with the impedance you want to measure to determine σ and ε_r and causes a small error. In practice, C_p will be roughly the same magnitude as C_0. When the probe is inserted into soil, however, C_0 is multiplied by ε_r and the effective capacitance is much larger than C_p. You can modify the equations to take C_p into account but except for soils with very low ε_r I don't think it matters much.

Figure A1 — I used this Almost All Digital Electronics L/C meter to measure the capacitance of the probes in air. By measuring the total capacitance of the probe and leads, and then measuring the parasitic capacitance, C_p, of the part of the probe above ground along with the connecting leads, I am able to calculate the capacitance of the probe alone, C_0.

Again, it is important to realize how small the measured capacitances are. You have to keep your body and any other conductors well away from the probe and the L/C meter. I place the probe and meter on top of a large plastic garbage can, well away from benches and other objects. I zero the meter by holding it with one stick and pushing the zero button with another, so the effect of my body is minimized.

Even with this simple and inexpensive instrument I believe I get quite accurate values for C_0. I confirmed the measurements using an HP 3577A vector network analyzer. Table A1 shows the parameters for my measurements.

C_p is in shunt with the measured impedance and might cause some error. You can, of course, modify the equations to remove this effect when C_p is known but I found that for most soils the values for the measured impedances were much lower than the shunt impedance presented by C_p and adding a correction factor was unnecessary.

Table A1

Measurement Parameters

Probe Type	*Rod Diameter (Inches)*	*Spacing (Inches)*	*Length (Inches)*	*Ground Screen (feet)*	C_o
Monoprobe	0.375	—	18	3 × 3	7.41 pF
OWL	0.44	4	18.5	—	4.83 pF
OWL	0.44	4	9.5	—	2.7 pF
OWL	0.44	3	12	—	3.4 pF

Rectangular Waveguide to Coax Transition Design

Learn how to find the optimum dimensions for a waveguide to coax transition using an empirical approach that relies on a set of impedance measurements and a few calculations.

Paul Wade, W1GHZ

A question I am frequently asked is, "Why do the antenna dimensions in the *W1GHZ Microwave Antenna Book — Online* not include the probe dimensions (for the transition from waveguide to coaxial transmission line)?"[1] The answer is that the transition is part of the transmission line, not the antenna, and does not directly affect the performance of an antenna. The transition may be right at the antenna, seemingly part of it, or at the other end of a run of waveguide transmission line, many meters away.

The transition is an important part of most microwave systems, however, since solid-state components are usually constructed on microstrip transmission lines and interconnected with coax, while microwave antennas normally use waveguide techniques.

A typical transition consists of a coaxial connector on the broad side of a rectangular waveguide with the center conductor extended as a probe into the waveguide, with one end of the guide ending in a short circuit, like Figure 1. Since the structure is well-known, *design* may be too ambitious a term, but the correct dimensions are far from obvious and are difficult to calculate. A number of sets of dimensions have been published, but there is little agreement between them, so it is difficult to tell which are right. Also, many of the published transitions are part of an antenna, so the dimensions may have been chosen to compensate for a poorly matched antenna impedance.

Therefore, I chose an empirical approach: making a comprehensive set of measurements from which the optimum dimensions may be reached. I had previously used this technique to determine probe dimensions for circular waveguide made from copper water pipe.[2] It is easier to determine the dimensions for rectangular waveguide, since commercial guide and components are readily available from surplus sources.

[1]Notes appear on page 16.

161 Center Rd
Shirley, MA 01464
w1ghz@arrl.net

Characteristics of a transition are best viewed by measuring the complex impedance (magnitude and phase) in the waveguide, using a waveguide slotted line. Figure 2 shows a typical X-band slotted line — the precision impedance measurement instrument of a few years ago, and, for waveguide, still more accurate than most network analyzer measurements. Since all professional microwave work today uses automatic network analyzers and computers, and few remember how to use a slotted line, slotted lines are almost given away today. I have paid as little as $2 for one at a hamfest.

To find the optimum dimensions for a transition, I needed to make measurements over a range of transition dimensions, so an adjustable transition was desirable. The probe dimensions are readily varied by unscrewing the coax connector and trimming the probe, but the distance to the shorted end of the waveguide, or backshort, must also be varied. I machined a sliding plug to fit inside the guide, with alternating quarter-wave sections of high and low impedance to form an electrical short circuit, so that the performance of the short does not depend on intimate contact with the waveguide walls. Figure 3 is a photograph of adjustable transitions for two common sizes of X-band waveguide, WR-90 and WR-75, as well as one for circular waveguide.

Measurement Technique

The first measurement is with a short circuit (flat metal plate) closing the end of the slotted line. The short provides a clear standing-wave pattern with sharp nulls at half-wavelength intervals, so we can measure the guide wavelength, and make any adjustments to the slotted line measuring probe.

The next measurement is of the sliding tran-

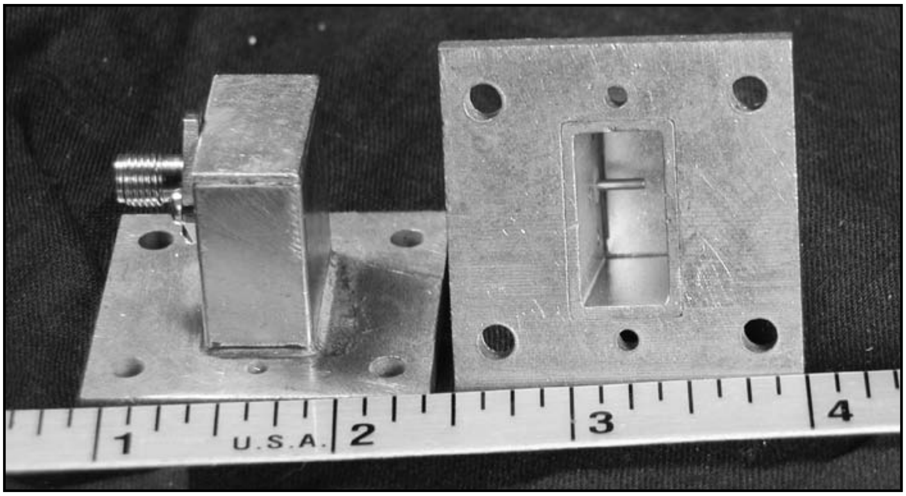


Figure 1 — WR-75 waveguide to coax transition for 10 GHz.

sition with no transition probe installed, with the face of the sliding short at the centerline of the coax connector. The location of the standing-wave nulls is recorded as the phase reference for the impedance measurements.

Finally, we are ready to measure the transition. A probe is installed in the transition, and a precision 50-Ω termination installed on the coax connector. Then the sliding short is moved to a series of locations and the impedance recorded at each location. (Actually, the SWR and the location of a null are recorded, and the complex impedance calculated later, using either a Smith Chart or a computer.) Then the probe is removed and replaced with a different one, and the measurement sequence repeated.

In my previous experiments with circular waveguide, I had found that the diameter of an SMA inner conductor, 1.27 mm, was in the optimum diameter range for X-band transitions, so I chose to limit the number of measurements by using only this diameter for rectangular waveguide at 10 GHz. With a single probe diameter, only the probe length is varied, so changes are easy, starting with a long probe and cutting off various increments.

Many commercial transitions have the probe surrounded by a cylinder of dielectric, so that the Teflon dielectric of the SMA connector continues for the full height of the waveguide. Since it is difficult to shorten the probe length inside the dielectric, my preferred transitions have the Teflon dielectric of the SMA connector ending at the inner wall of the waveguide and only a bare probe extending into the guide. I did, however, want to compare the performance of bare-probe transitions with full-height-dielectric ones, so I carefully sliced the protruding dielectric from a connector, and then fit it back together for the measurements after each adjustment of the probe length.

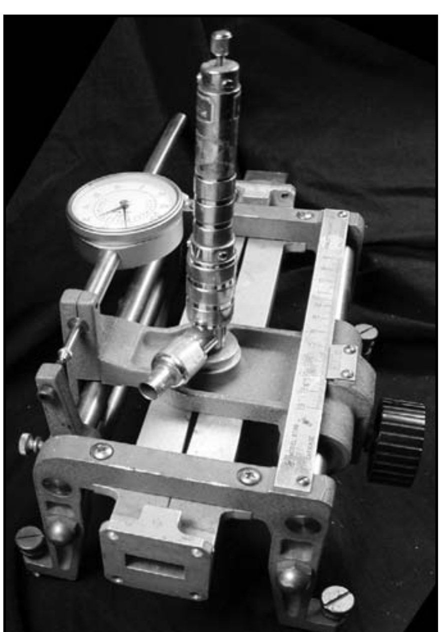


Figure 2 — A surplus waveguide slotted line.

Measurement Results

Without the slotted line or an expensive network analyzer, only SWR could be easily measured. Figure 4 shows a plot of the SWR versus probe position and diameter in WR-90 waveguide at 10.368 GHz. We see that there is a set of dimensions that produce an excellent SWR, and a number of other combinations that produce a usable SWR < 2:1. What isn't clear is how we reached the excellent dimensions, or whether there are other excellent combinations that we did not find.

A plot of complex impedance on a Smith Chart is much more meaningful. Each curve in Figure 5 is for a constant probe length, with varying backshort distance. The backshort forms a section of shorted transmission line behind the probe, creating a reactance at the probe that varies with the length of the line to the backshort. At some distance, this reactance is equal to the reactance of the probe but with conjugate phase, canceling the reactance so that only the resistive component of the impedance remains. This is where the curves cross the horizontal centerline, or real axis, of the Smith Chart.

The resistive component where each curve crosses the real axis varies with probe length — we can see that the curve for a short probe crosses the real axis to the right of center, at a higher resistive impedance than

Figure 3 — Adjustable transitions for WR-90, WR-75 and circular waveguide.

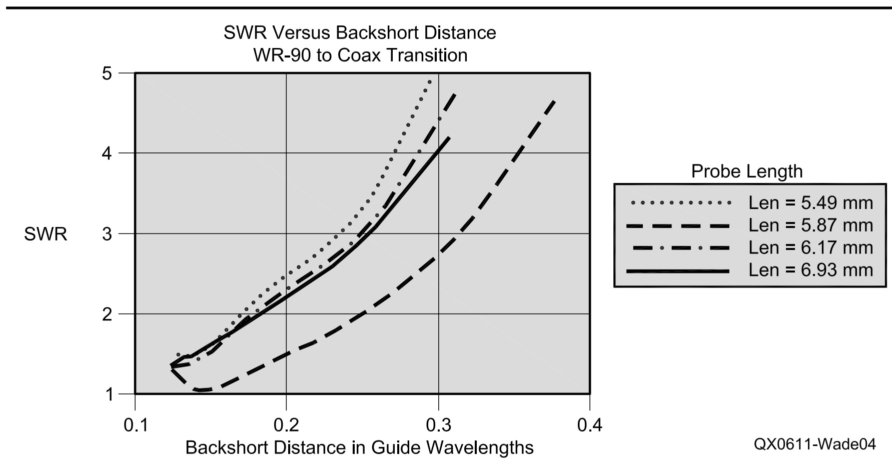


Figure 4 — WR-90 to coax transition, SWR versus backshort distance at 10.368 GHz.

Z_0. Conversely, a longer probe crosses to the left of center, so the resistive impedance is lower than Z_0. The relationship is clear and monotonic, so we can be confident that there is a single optimum set of dimensions that produces the best impedance match, which is in the center of the Smith Chart. Once we have a set of curves that bracket the center, we can interpolate to estimate the optimum probe length, then make more measurements to locate the backshort position for a perfect match — see the data point in Figure 5 marked 0.146 λ.

Discussion

The probe in the waveguide acts as an impedance transformer from the low impedance of the coaxial line, typically 50 Ω, to the high impedance of the waveguide, typically greater than the impedance of free space = 377 Ω. Since there is no loop for current to flow in and create a magnetic field, the probe must be an electric field transformer. The maximum electric field is at the center of the wide dimension of the waveguide, where we have placed the probe. Intuitively, we suspect that a longer probe intersects more of the electric field and thus couples more tightly. We saw this in Figure 5 — the longer probe produces a lower impedance in the waveguide, because it is coupled more tightly to the low impedance in the coax.

Adding a dielectric around the probe would tend to concentrate the electric field near the probe and increase coupling, so a shorter probe would be required. We can see this in Figure 6 — the same WR-90 waveguide as Figure 5, but with the Teflon dielectric of the SMA connector extended to the far wall of the guide. We see the same family of curves, but the impedance is matched with a shorter probe length.

There seems to be a persistent myth, found in older microwave books, that the backshort should be exactly λ/4 from the probe. We know that a shorted quarter-wavelength of transmission line acts an open circuit, so this distance would add no reactance at the probe; this distance would only work for an ideal probe, with no inductance or capacitance of its own. The probe is a wire, however, so it must have inductance, in addition to capacitance to the waveguide walls. We can see in the Smith Charts that the backshort distance required to achieve a pure resistive impedance varies slightly with probe length, as we would expect if it were tuning out the reactance of the probe. More to the point, Figure 4 shows that with the backshort λ/4 from the probe, the best SWR achieved is 2:1.

Since the backshort distance is a reactive phenomenon, it must be frequency dependent. We can see how critical by measuring the SWR versus frequency. I fabricated fixed transitions to the best dimensions, one with a bare probe and the other with extended Teflon dielectric. In Figure 7, we can see that the WR-90 transitions have good bandwidth, with excellent SWR from 10 to 11 GHz, and pretty good from 9.7 to 13 GHz. The transition with the Teflon dielectric shows slightly better bandwidth than the one with a bare probe, but not a significant improvement — both are more than adequate to cover the whole amateur band. Figure 8 shows the performance of the two WR-90 transitions connected together, with extremely low insertion loss from 10 to 13 GHz.

Other Waveguides

Scaling the WR-90 transition dimensions to other frequencies and waveguide sizes is more difficult in rectangular waveguide than it was with circular waveguide, since other standard waveguides have different rectangular aspect ratios. In circular waveguide, the characteristic impedance is a function of the guide wavelength:

$$Z_0 = 377 \bullet \left(\frac{\lambda_g}{\lambda_0} \right) \quad \text{(Eq 1)}$$

while in rectangular waveguide, the characteristic impedance is modified by the aspect ratio:

$$Z_0 = 377 \bullet \left(\frac{\lambda_g}{\lambda_0} \right) \bullet \frac{2b}{a} \quad \text{(Eq 2)}$$

where *a* and *b* are the large and small inner dimensions, respectively. The guide wavelength, λ_g, is easily measured with the slotted line, but it can also be calculated:

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{\lambda_c} \right)^2}} \quad \text{(Eq 3)}$$

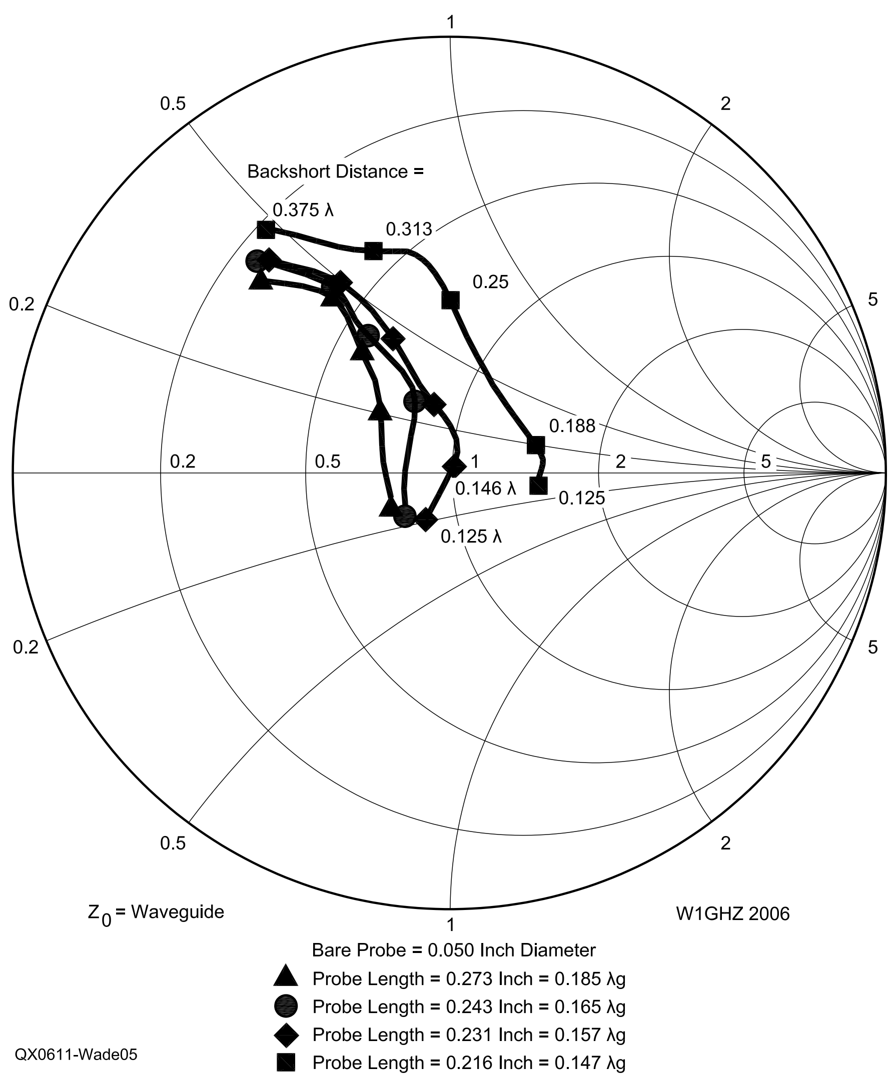


Figure 5 — Smith Chart plot of impedance as dimensions are adjusted for WR-90 to coax transition.

where the cutoff wavelength, $\lambda_c = 2a$, or twice the large inner dimension of the waveguide.

Thus, in different size waveguides, we are transforming to different waveguide characteristic impedances. Another good X-band waveguide is WR-75; the aspect ratio *b*/*a* is 0.5, while for WR-90 the aspect ratio is 0.444.

Since λ_g is larger in the smaller guide as well, the characteristic impedance of WR-75 at 10.368 GHz is significantly higher than the Z_0 of WR-90.

Thus, we would expect the transition dimensions for WR-75 to also be different. In Figure 9, the impedance curves measured for an adjustable transition are shown; we see the same family of curves as previously, but a shorter probe is required to match the higher impedance. I did not take a full set of measurements with the extended Teflon dielectric in WR-75, just enough to find approximate matched dimensions. In Figure 10, we can see that this shortcut was a mistake — the best probe length would be slightly shorter than the one measured. This is a problem with making a lot of measurements before doing the calculations necessary to review the results.

I fabricated two WR-75 fixed transitions, one with a bare probe and the other with extended Teflon dielectric. Figure 11 shows that there is no appreciable difference in bandwidth — both have good SWR from 10 to 10.5 GHz — but narrower bandwidth than the WR-90 transitions. In Figure 12, the insertion loss is very small from 10 to 11 GHz, and low from 9 to 16 GHz.

The recommended operating range for WR-75 is 10 to 15 GHz, while the range for WR-90 is 8 to 12.4 GHz. It is probable that if we were to find WR-75 transition dimensions at a frequency near the center of the recommended range, it would have a wider bandwidth. It would probably not be as good in the 10 GHz amateur band, however. Many surplus WR-75 transitions suffer from this deficiency — they work best around 12 to 14 GHz — so we are probably better off fabricating our own transitions optimized for 10 GHz.

Other Frequencies

We have determined optimum waveguide to coax transition dimensions for 10 GHz empirically. These dimensions are listed in Table 1, and a sketch of a transition is shown in Figure 13. The 10 GHz dimensions are not directly scalable to other waveguide sizes and frequencies, but we have demonstrated a technique applicable to other waveguide sizes and frequencies.

What can we do if we do not have a waveguide slotted line suitable for the target

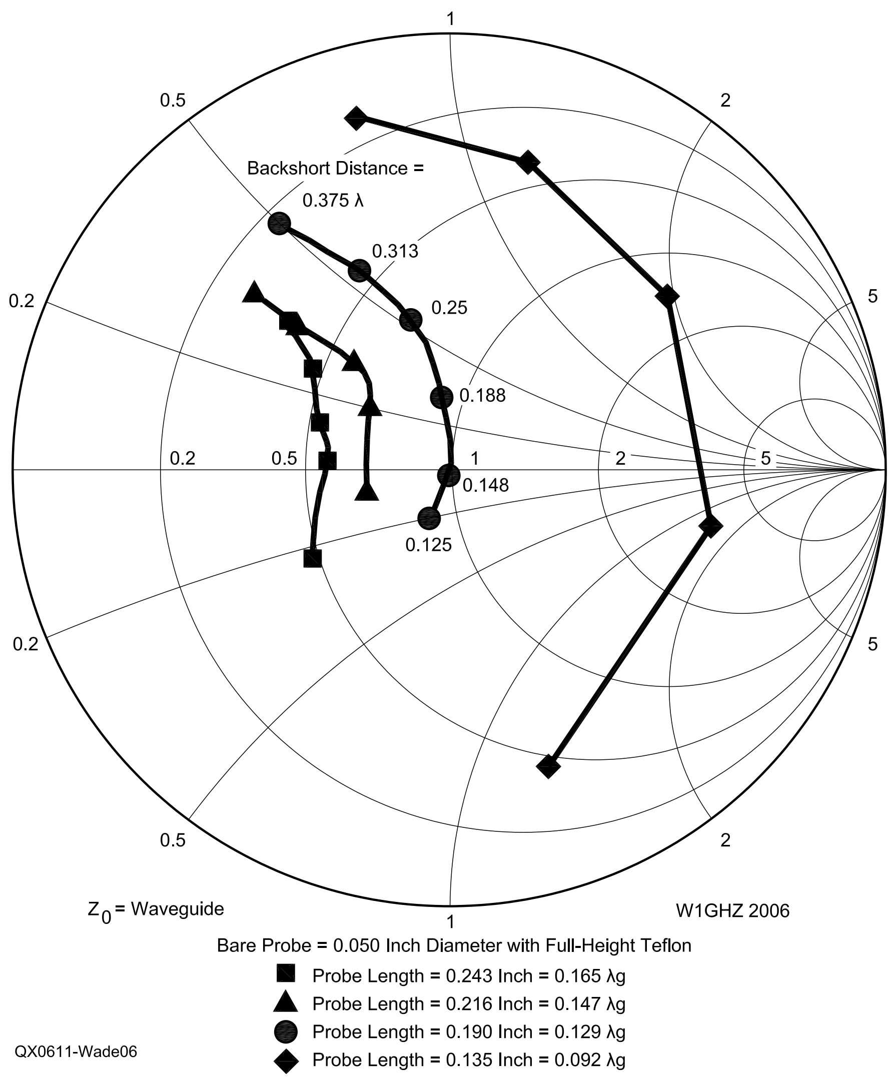


Figure 6 — Smith Chart plot of impedance as dimensions are adjusted for WR-90 to coax transition with full height Teflon.

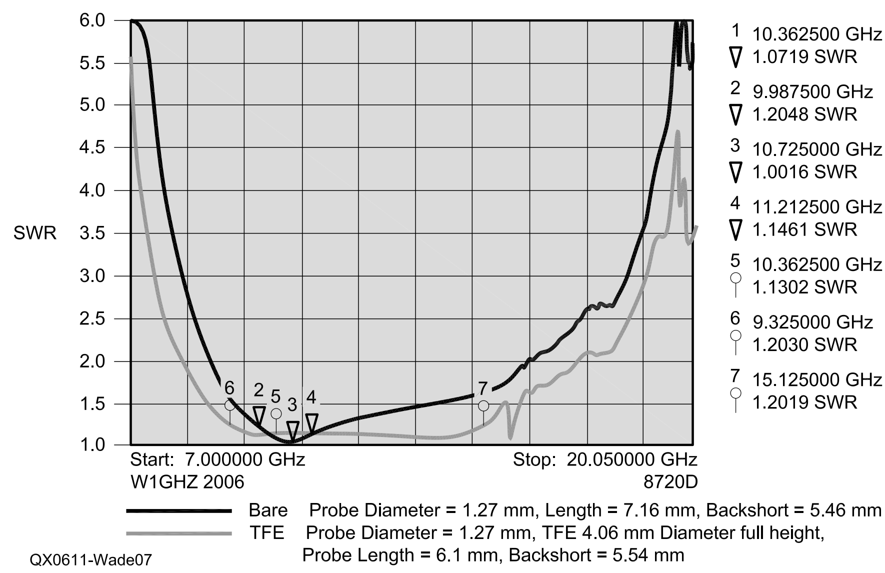


Figure 7 — SWR versus frequency of WR-90 to coax transition.

frequency? Can we make a transition if all we can measure is SWR or return loss? I think we can come pretty close, even without a sliding short. Since we have to drill a hole for the coax connector, we must first estimate the backshort distance, from the hole to the shorted end. For the WR-90 transitions, the backshort distance is 0.146 λ_g, while for the WR-75 transitions, the distance is 0.118 λ_g. Looking at Figures 5, 6, 9, and 10, we can see that a compromise distance of 0.125 λ_g, splitting the difference, will give us an SWR of 1.3 or better with the right probe length. So, if we put the coax connector about one-eighth of a guide wavelength from the shorted end and trim the probe length for best SWR, we can probably get a pretty decent transition in any size rectangular waveguide.

Figure 14 is a plot of SWR versus probe length for different backshort spacings, in $^1/_{16}$ λ_g increments, for the WR-90 transition with the Teflon dielectric extending the full height of the waveguide. The 0.125 λ_g spacing gives the best SWR. If we start with a long probe and trim, the SWR will decrease slowly, then increase rapidly when we pass the optimum length. Therefore, we should stop shortening as soon as we see an increase in SWR. One final note: the SWR for the $\lambda/4$ spacing recommended in the old books never gets below about 2 — still usable, but not very good.

Figure 14 shows that a fairly wide range of dimensions will result in a reasonable SWR, less than 2 or 3, which can be easily tuned out with some screws in the waveguide. Perhaps this is why some published transitions include tuning screws. Without good test equipment, though, it is just as easy to make it worse as it is to make it better. Many microwave operators are able to get on the air successfully with minimal test equipment, so any necessity for tuning creates an additional hurdle for beginners.

Software

Sometime after I finished the WR-90 and WR-75 transitions, I started a new job. Part of the new job involves electromagnetic simulations, and I had to learn to use some new software. One of the simulation programs, Ansoft *HFSS*, is well suited to waveguide calculations.[3] Since I needed some simple problems as learning exercises, I tried simulating some transitions, using the same empirical technique, but calculating the impedance with the software rather than measuring it.

After simulating my WR-90 and WR-75 dimensions to make sure the two techniques yield the same results, I moved to WR-42 waveguide at 24.192 GHz. The impedances were plotted on a Smith Chart until they

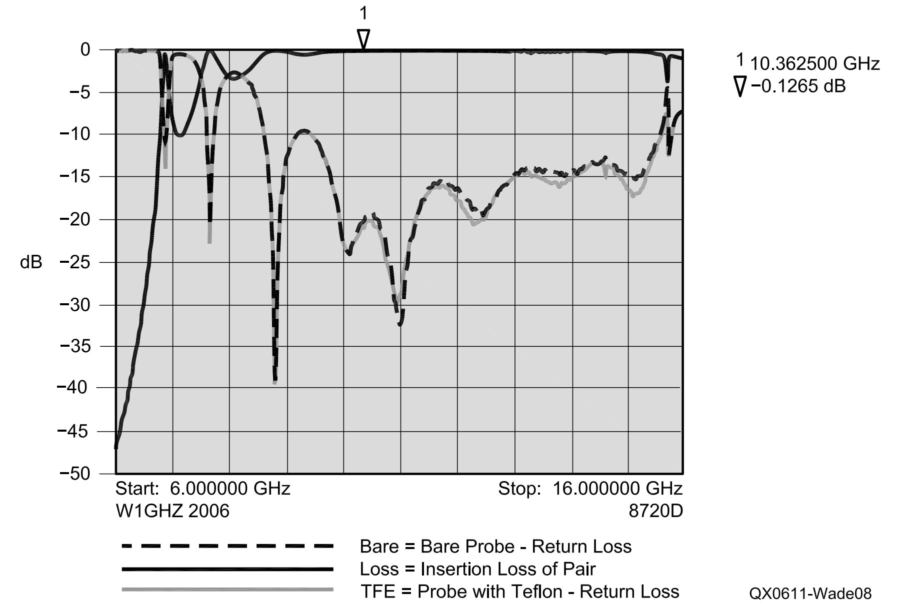


Figure 8 — Insertion loss and SWR of two WR-90 to coax transitions back-to-back.

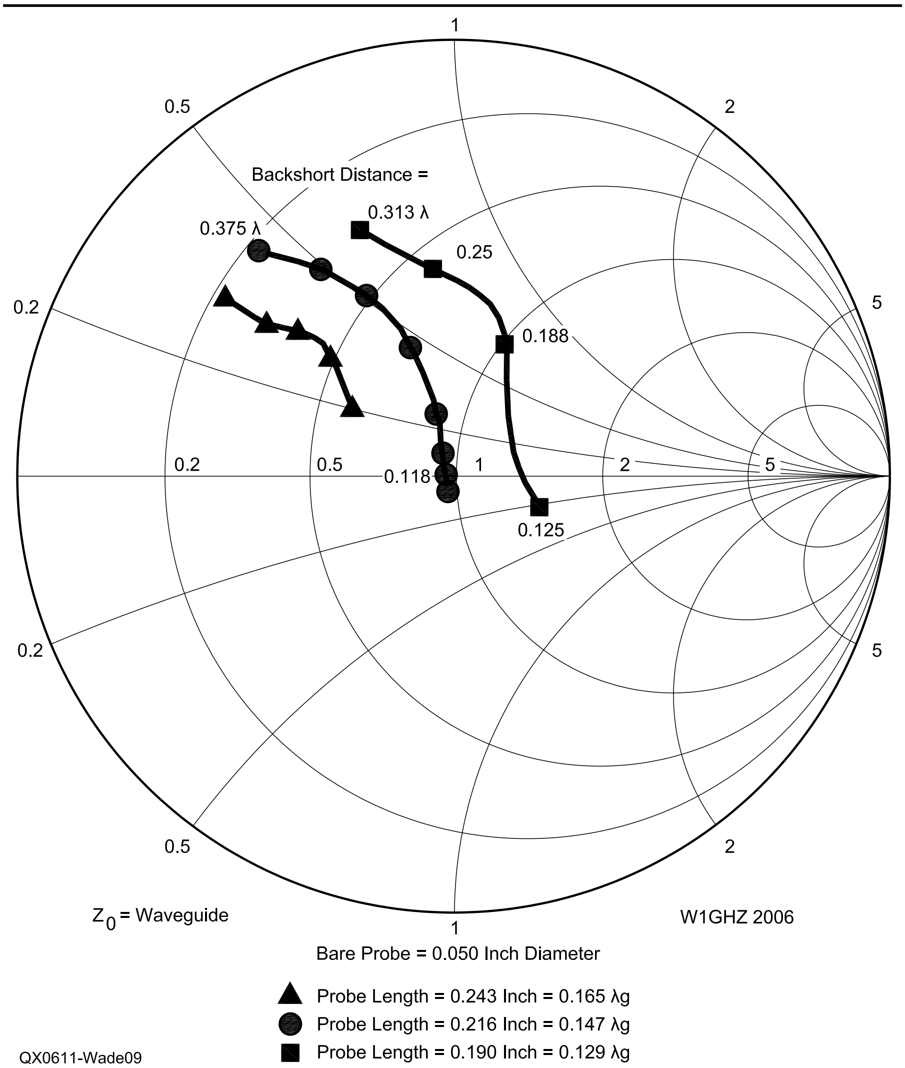


Figure 9 — Smith Chart plot of impedance as dimensions are adjusted for WR-75 to coax transition.

converged on optimum dimensions. Since the software calculated the impedance directly, it was easy to plot each point immediately and move quickly toward the optimum with fewer trials.

Once I had the optimum dimensions, I built a WR-42 transition, shown in Figure 15, and measured it with another surplus slotted line. The results were encouraging, with an SWR of 1.02. I quickly made several more, with similar results, so that I could finish my 24 GHz transverter.

Other Frequencies and Waveguides

To round out the results, in off-hours when the software was not being used, I calculated optimum dimensions for some of the lower microwave ham bands in various sizes of waveguides. These dimensions are listed in Table 1. Not all the dimensions have been verified experimentally, but I have built and tested a number of them in different sizes — many of them are shown in Figure 16. The one at the far right, in WR-159 waveguide for 5.76 GHz, has not had the backshort installed yet.

The bandwidths shown in Table 1 are for a Return Loss > 20 dB, equivalent to an SWR < 1.2:1. Larger diameter probes generally provide a wider bandwidth. For most amateur work, wide bandwidth is not important, but a very narrow bandwidth often makes the dimensions more critical. I found that the required tolerance on the probe and backshort dimensions is roughly 0.1 mm (0.005 inch) at 10 GHz, 0.05 mm (0.002 inch) at 24 GHz, and proportionally larger at the lower frequencies.

Some of the probe diameters were chosen to ease fabrication. The higher frequency probes are 1.27 mm diameter — the center pin diameter of an SMA connector — making fabrication a simple matter of trimming the Teflon, then cutting and filing the probe to length. At lower frequencies, the SMA pin diameter is too small to make an effective probe, but AWG no. 12 copper wire fits neatly into the solder end of a female N connector with little discontinuity. The alternative is to slip a larger diameter probe over the SMA pin — I tried this with WR-187 transitions at 3.456 and 5.76 GHz, and found that the discontinuity from the step in diameters required a change in probe dimensions to compensate. Table 1 shows dimensions calculated for a 2.36 mm diameter probe and the modified dimensions for the same probe diameter extending from an SMA pin.

Integrated Transitions

In general, it is a good idea to keep the transition separate from the antenna, so that each may be tested individually and tuned

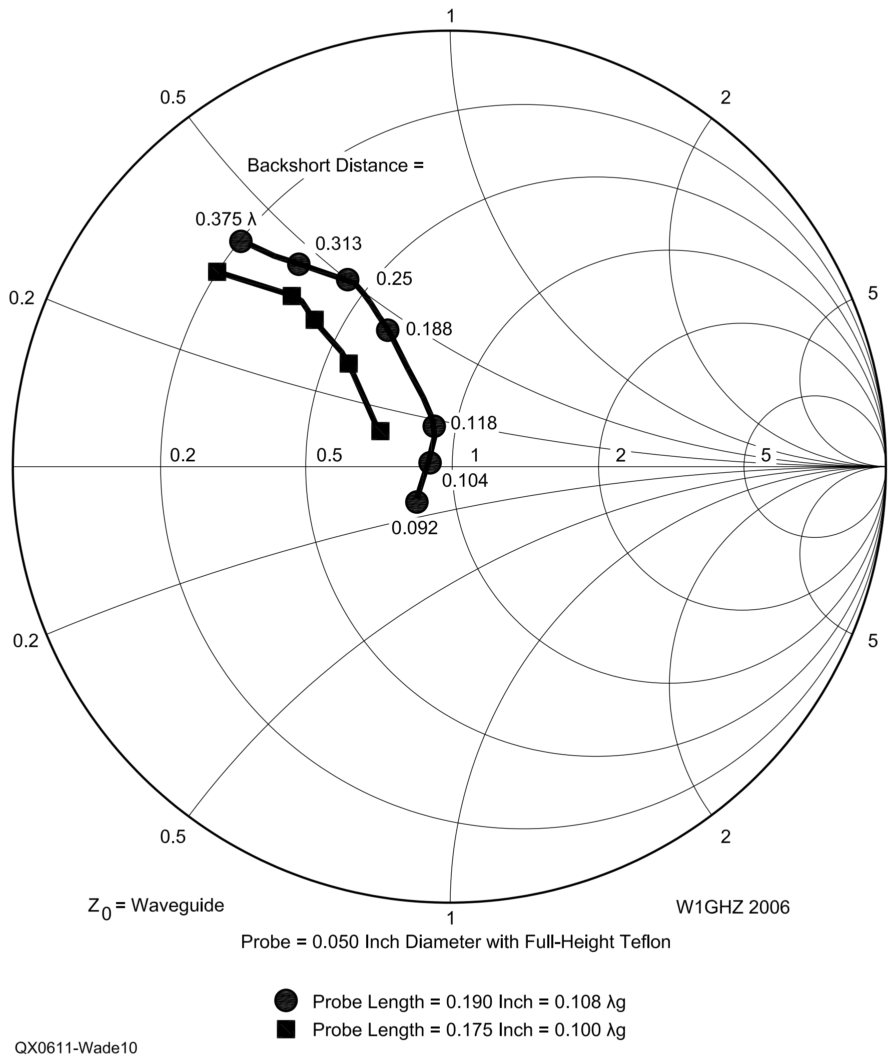


Figure 10 — Smith Chart plot of impedance as dimensions are adjusted for WR-75 to coax transition with full height Teflon.

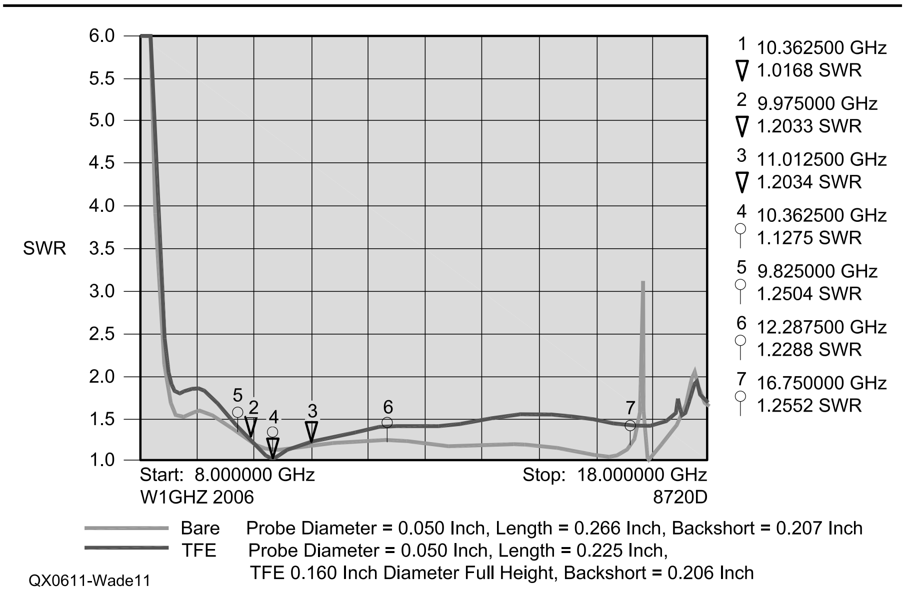


Figure 11 — SWR versus frequency of WR-75 to coax transition.

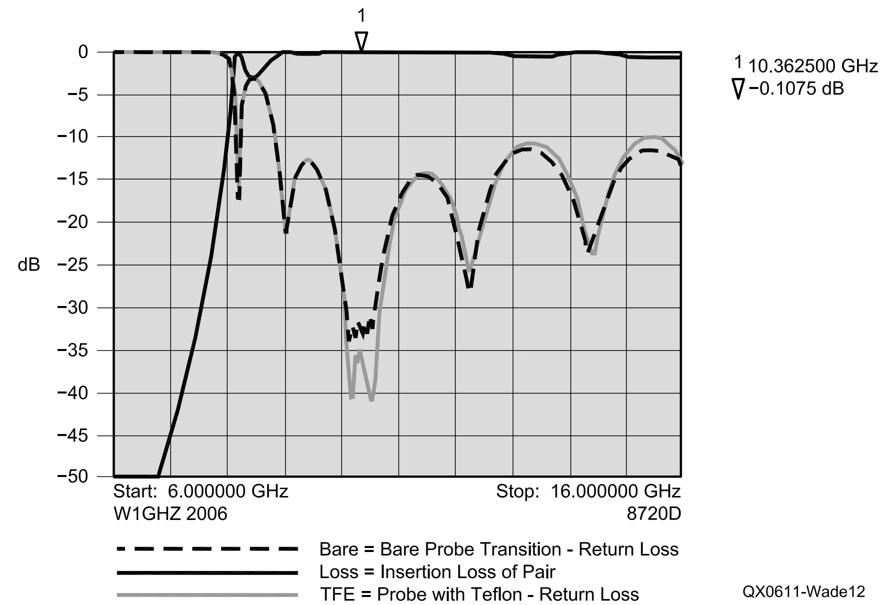


Figure 12 — Insertion loss and SWR of two WR-75 to coax transitions back-to-back.

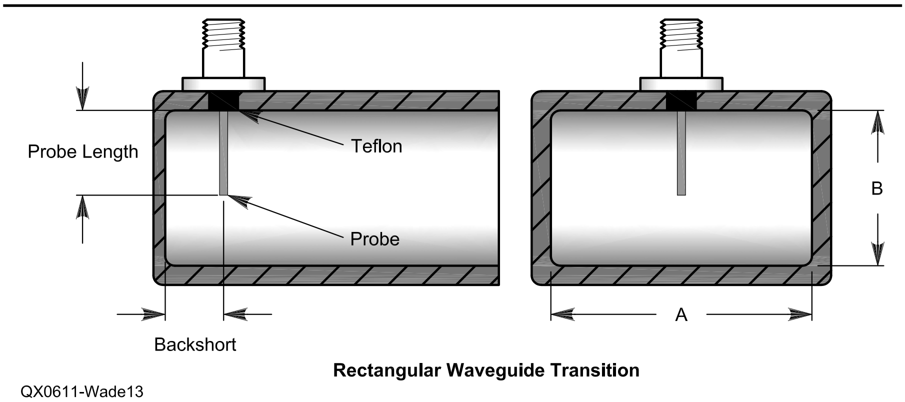


Figure 13 — Sketch of rectangular waveguide to coax transition, showing dimensions.

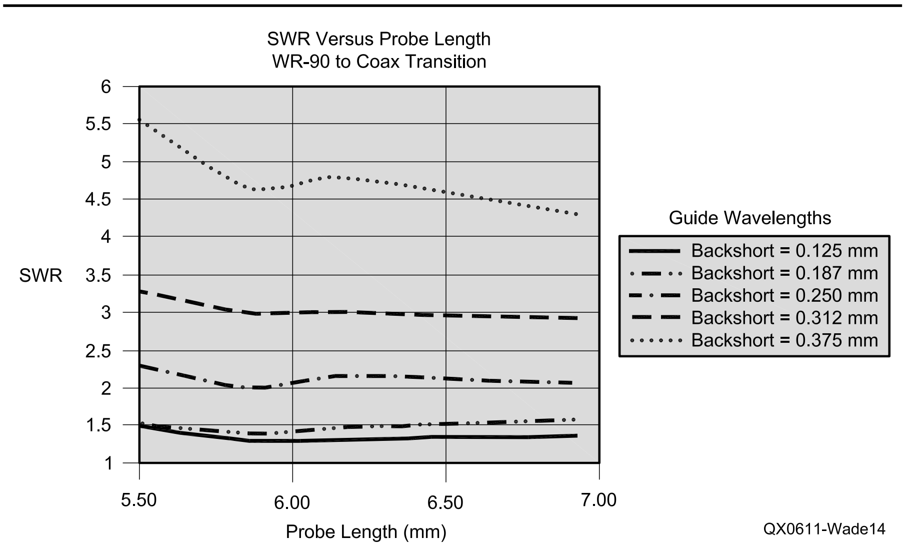


Figure 14 — WR-90 to coax transition, SWR versus probe length at 10.368 GHz.

if necessary. The length of waveguide between the transition and the antenna does not matter if it is simply a matched transmission line. For simple, foolproof antennas like a rectangular horn, however, it may be convenient to integrate them into one piece, with the horn soldered to one end of a scrap of waveguide and the transition at the other. Figure 17 is a photograph of two 10 GHz integrated rectangular feed horns for a common offset dish.[4]

Summary

By measuring impedance from the waveguide side of a transition, we are able to predictably adjust and optimize dimensions. The impedance may be found using either cheap surplus hardware or expensive software.

For common waveguide sizes, the optimum dimensions are listed in Table 1 for most of the amateur microwave bands. Armed with this data, only some simple metalwork is required to turn a bit of surplus waveguide into a working transition with no further tuning required.

Notes

[1]P. Wade, W1GHZ, *The W1GHZ Microwave Antenna Book — Online* at **www.w1ghz.org**.

[2]P. Wade, W1GHZ, "Understanding Circular Waveguide — Experimentally," *QEX*, Jan/Feb 2001, pp 37-48. (Reprinted with ARRL permission at **www.w1ghz.org/QEX/circular_wg.pdf**).

[3]**www.ansoft.co**.

[4]*The ARRL UHF/Microwave Projects Manual, Volume 2*, ARRL, 1997, p 1-34. Also "Microwavelengths," *QST*, Aug 2005, pp 75-76.

Paul Wade, W1GHZ, previously N1BWT and WA2ZZF, has been licensed since 1962 and

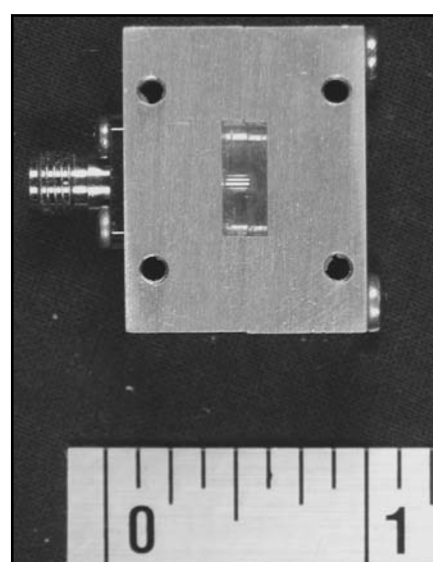


Figure 15 — WR-42 to coax transition for 24 GHz in milled aluminum block.

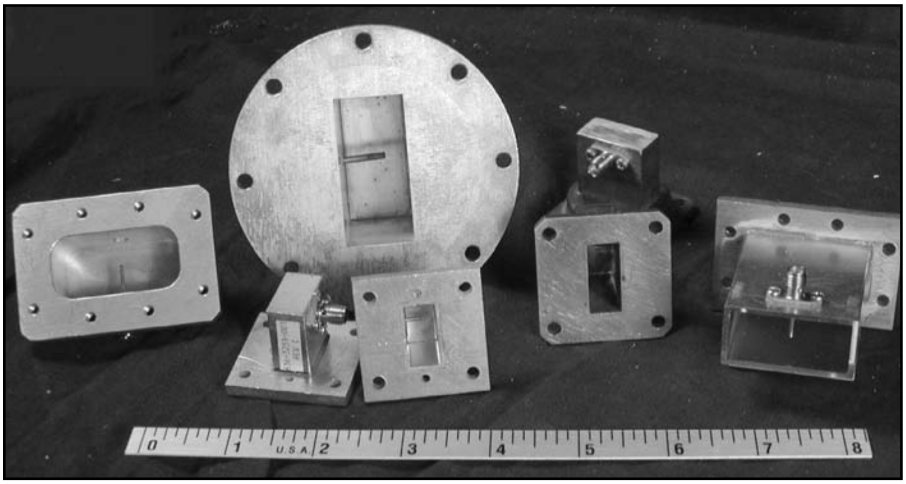


Figure 16 — Transitions in several sizes of waveguide for different microwave bands.

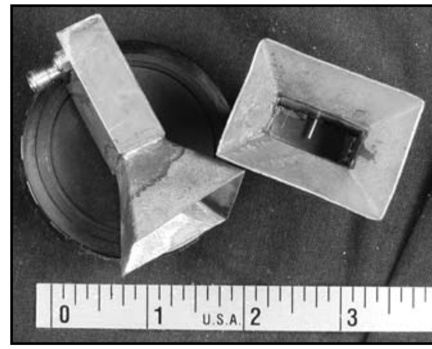


Figure 17 — WR-90 transition integrated into 10 GHz feed horn for DSS offset dish.

Table 1
Rectangular Waveguide to Coax Transitions W1GHZ 2006

Waveguide	*Frequency (GHz)*	*Probe Diameter (mm)*	*Probe Length (mm)*	*Backshort Length (mm)*	*Bandwidth*	*Number Tested*
WR-42	24.192	1.27	2.41	2.49	>17%	4
WR-75	10.368	1.27	5.49	5.26	14%	1
WR-90	10.368	1.27	5.89	5.46	7%	5
WR-112	10.368	1.27	6.5	6.6	15%	1
WR-112	5.76	1.27	9.8	5.8	7%	
WR-137	5.76	1.27	10.5	8.5	10%	1
WR-159	5.76	1.27	11.17	10.0	11%	1
WR-159	5.76	AWG no. 12	10.9	10.0	14%	
WR-187	5.76	2.36	11.3	11.0	16%	
WR-187	5.76	SMA to 2.36	11.6	9.7	16%	1
WR-187	5.76	AWG no. 12	11.3	11.2	14%	
WR-187	3.456	2.36	14.5	18.0	5%	
WR-187	3.456	SMA to 2.36	15	16.5	5%	1
WR-187	3.456	AWG no. 12	14.9	17.4	7%	
WR-229	3.456	1.27	18.2	15.0	8%	
WR-229	3.456	AWG no. 12	17.7	15.1	10%	
WR-229	3.456	2.36	17.4	15.06	11%	
WR-229	3.456	3.175	17	15.6	11%	
WR-229	3.456	4.76	16.2	16.2	14%	
WR-229	3.456	6.35	15.5	16.75	17%	
WR-284	3.456	6.35	17.8	17.5	23%	
WR-284	3.456	AWG no. 12	19	17.5	11%	
WR-284	2.304	6.35	20	28	8%	
WR-284	2.304	AWG no. 12	24	24	7%	
WR-284	2.4	AWG no. 12	23.7	22	8%	
WR-340	2.304	6.35	25	23	11%	
WR-340	2.304	AWG no. 12	27	23	9%	
WR-340	2.4	AWG no. 12	26	23	9%	

has never made a contact below 50 MHz. He has been a microwave experimenter for years and published numerous articles, is active in the Vermont 10 GHz group and a past President of the North East Weak Signal Group. An ongoing project is the W1GHZ Microwave Antenna Book — Online *at* **www.w1ghz.org**. *In 1997 Paul was honored to be named by the Central States VHF Society as the recipient of the Chambers Award. More recently, he was honored by the ARRL with the 2000 Microwave Development Award, and in 2001 with the Thomas Kirby Eastern VHF/UHF Society Award.*

A former microwave engineer and retired ski instructor, he is currently employed by Mercury Computer Systems designing computer hardware. A frightening experience at Microwave Update 2001 involving a giant meatball resulted in an angioplasty with two stents, and swearing off meatballs in favor of a low-fat diet.

QEX

IMD in Digital Receivers

Performance limitations of receivers with the A/D converter at the antenna – how to measure and work around them.

Leif Åsbrink, SM5BSZ

Today, most receivers have an A/D converter at some point in the signal path because digital technology can provide better filters at lower cost compared to analog technologies. The A/D converters gradually move nearer the antenna and digital technology takes over a larger fraction of all the filtering work. Someday receivers will be completely digital and sample directly at the signal frequency — at least for HF bands. Sampling directly at the signal frequency is already possible today with amateur equipment. The first commercially available equipment is the SDR-14 from RFSPACE. This article highlights the problems of characterizing this new class of radio receivers.

Standard methods of characterizing performance in terms of third-order intercept and intermodulation-free dynamic range fail badly and cannot be used at all to give an adequate representation of how well these new receivers perform as compared to conventional radios when used on the same antenna with real signals. New ways of doing measurements are called for, and this article is intended to shed some light on the problems of measurement and on the methods we can use to improve performance by adding preselectors and reducing out-of-band signals. The SDR-14 is used as the example, but all the problems are inherent in the technology. Digital technology is developing rapidly; better A/D converters become available and performance will improve drastically in the future but still the problems of properly characterizing a radio receiver will remain.

Practical Consequences of Dynamic Range Limitations in the SDR-14

Look at Figure 1. This computer screen capture shows two waterfall graphs taken from the 40-m band. The lower one is made with the SDR-14 hardware while the upper one is made with a Delta 44 soundcard connected to WSE converters that bring the signals down to baseband I and Q. The waterfalls are produced with two instances of Linrad running simultaneously in the same computer under *Gnome*, a graphical user interface for Linux. The two systems are connected to the same antenna and the noise figure is made equal by having a 1-dB attenuator at the input of the SDR-14. See Figure 2. A preselector with adequate dynamic range and a gain figure of 26 dB followed by a stepped attenuator is placed between a simple wire antenna and a –3-dB hybrid. The Delta 44 sampling speed is 96 kHz and therefore its waterfall only covers 96 kHz of the 40-m band. The SDR-14 always samples at 66.667 MHz but its decimation chip is set to deliver the baseband I and Q at a sampling rate of 189.4 kHz, and therefore it shows about twice as wide a spectrum. A quick glance at the lower part of Figure 1 shows that there are many false signals in the SDR-14 spectrum. They occur at every 5 kHz and it is pretty obvious that they originate in some kind of nonlinearity that mixes the AM broadcast carriers of the 40-m band with each other, producing high-order intermodulation products — or at least spurious signals that occur at frequencies where intermodulation products would be expected.

Now look at Figure 3. That image was made a few minutes later than Figure 1. The only difference is that the stepped attenuator is set for much stronger signals and that the color scale of the waterfalls is changed accordingly. Both receivers see 40 dB more signal. Note that the "intermodulation" products have disappeared completely. There is not the slightest trace of the nonlinearities when the SDR-14 is fed with strong enough signals! It is obvious that the mechanism producing the false signals is not what we normally call intermodulation and characterize with an IP3 number. The SDR-14 is an excellent receiver provided it is preceded by a suitable preselector that removes most of the HF signals outside the band of interest and amplifies what passes the filter to a level close to the maximum level the unit can handle. There is an LED on the

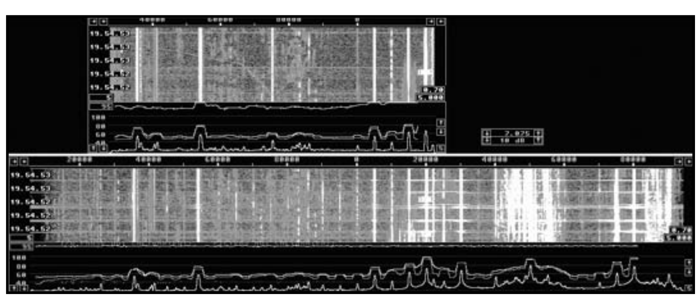


Figure 1 — The top screen capture is made with a Delta 44 soundcard connected to WSE converters that bring the signals down to baseband I and Q. The lower screen capture is made with the SDR-14 hardware.

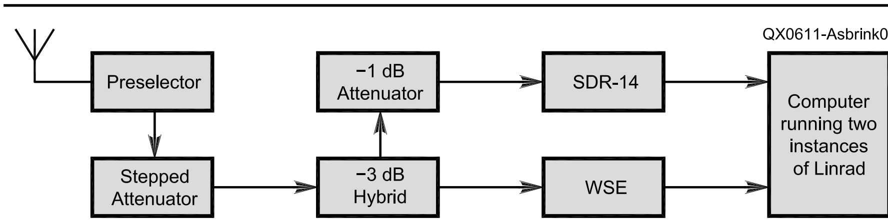


Figure 2 — Test setup for the simultaneous test of two hardware systems on a single computer running two instances of *Linrad*.

Jaders Prastgard
63505 Eskilstuna
Sweden
leif@sm5bsz.com

SDR-14 showing if too much gain is applied. More color images and detailed information about this real life test can be found on the Internet.[1]

[1]www.sm5bsz.com/digdynam/practical.htm

The Two-Tone Test Applied to the SDR-14

The level of the third-order intermodulation products do not follow the third-order law at all. It is absolutely unacceptable to attribute a third-order intercept point (IP3) to this receiver.

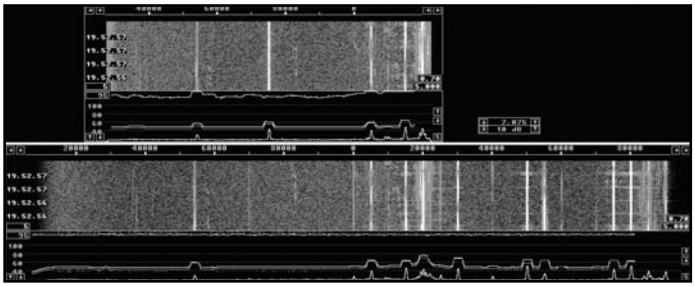


Figure 3 — The screen captures here are made the same way as those of Figure 1. The only difference is that the stepped attenuator is set for much stronger signals and that the color scale of the waterfalls is changed accordingly. Both receivers see 40 dB more signal. Note that the "intermodulation" products have disappeared completely.

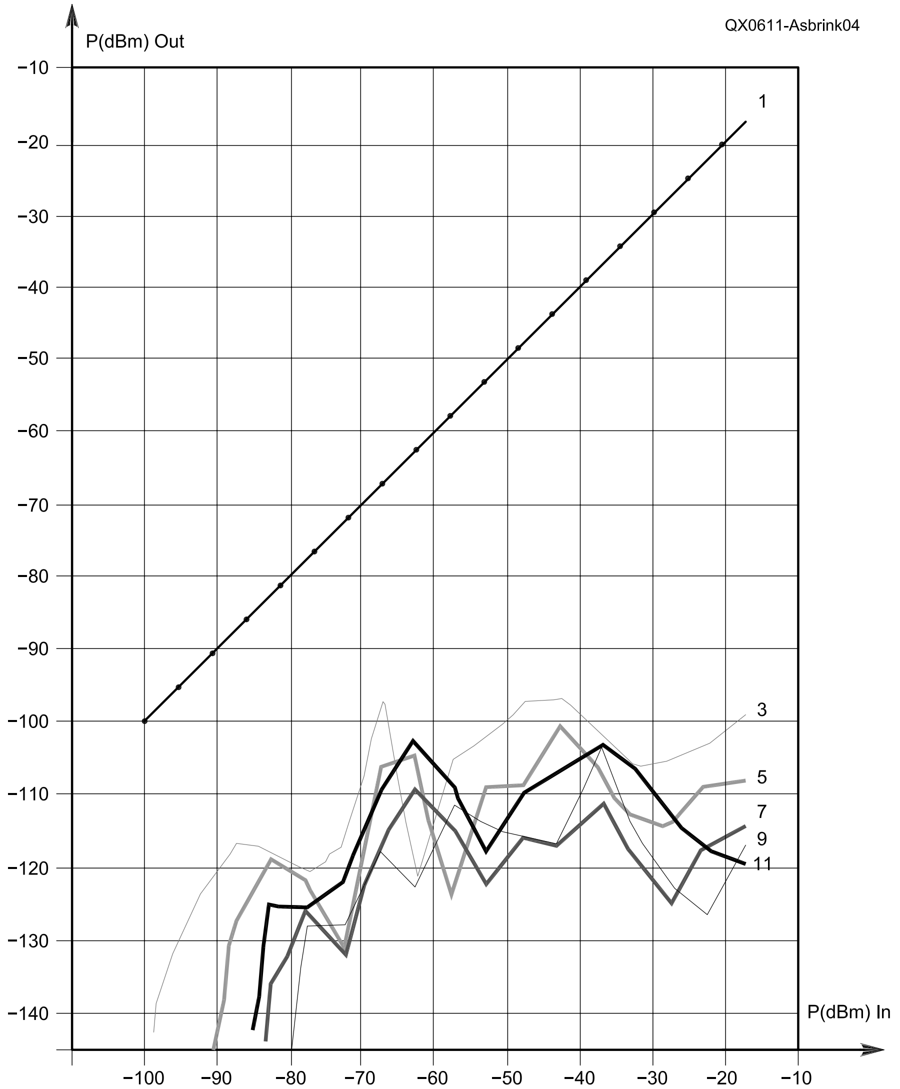


Figure 4 — Two-tone test for the SDR-14 using the 0.1 to 30 MHz input. Third-order "intermodulation" is at about –100 dBm, independent of the test signal level. Higher order interference, IM5, IM7, IM9 and IM11 are at similar power levels.

The performance has to be characterized with some other figure of merit. The measurement, however, is the same. Feed two equally strong signals simultaneously into the unit and measure the level of the signals at the intermodulation frequencies. The measurements have to be done at many signal levels in order to produce the function intermodulation versus input level. Such functions in steps of 5 dB are plotted in Figure 4 for several orders of the "intermodulation product." The mechanism is not ordinary intermodulation and the signals at frequencies corresponding to different orders of intermodulation do not have amplitude relationships that bear the slightest resemblance to what one would see in an analog receiver.

The curves in Figure 4 do not allow an IP3 to be determined, but they do allow a determination of the third-order intermodulation-free dynamic range. Two input signals at –17.5 dBm produce IM3 at –99.1 dBm. The best two-tone intermodulation-free dynamic range is thus about 82 dB, and the noise floor would have to be put at about –100 dBm at 500 Hz bandwidth. This is 35 dB above the internal noise floor of the SDR-14 itself.

–174 dBm (room temperature) + 12dB (SDR-14 noise figure) + 27 dB (500 Hz bandwidth) = –135 dBm (noise floor power of SDR-14) This result is well in line with the observations made on real life signals (see Note 1). Table 1 shows the measured data.

Dithering

The interference observed in a two-tone test is produced by the addition of an error signal that depends on the digital value of the A/D output. It does not matter whether the reason is limited accuracy of the A/D process or whether the reason is pickup of a voltage from one of the lines of the digital data bus. The error will occur periodically and therefore it will generate periodical signals. In a two-tone test, such signals will look like intermodulation. Dithering is often used to destroy the periodicity of the error signal. By adding noise, one can force the errors to occur at random times and thereby smear out the signal energy so no false signal is created. As a consequence there will of course be some degradation of the noise floor.

When adding wideband noise for dithering, it is a good idea to use a notch filter that will not allow any noise at the signal frequency to reach the A/D converter. When adding wideband noise at a power level 6 dB below the level where the saturation indicator flashes occasionally, the noise floor only increases by 3 dB in the selected passband. The false signals are drastically reduced, however, as can be seen in Figure 5. Table 2 shows the measured data.

With the test tones at –32 dBm, the only observable interference signal is at –130 dBm, which is only 2 dB above the intrinsic noise floor power of the dithered SDR-14 unit. The best two-tone, third-order, intermodulation-free dynamic range is thus improved from 82 dB to 98 dB by dithering, with notched white noise.

Dithering does not require random signals, however. Any signal outside the passband can be used. If the dithering signal is a sine wave, the spurs will move to other frequencies that correspond to mixing products between the dither signal and its overtones, and the test signals. To smear out such signals one can frequency modulate a sine wave with a low frequency sine wave. Figure 6 shows the performance of an SDR-14 on 10.7 MHz, when dithered by a frequency-modulated sine wave at 7.22 MHz. The power of the dithering signal was –20 dBm, about 9dB below the point of saturation. The modulation frequency was 1 kHz and the frequency swing about 200 kHz. The higher-order intermodulation products are smeared out over many MHz and no low-order spur is visible within about 190 kHz on the computer screen. The frequency modulated sine wave makes a better dithering signal than random noise. It sweeps nearly the full range of the A/D converter continuously while the white noise only occasionally reaches full amplitude and sometimes has a very low amplitude for quite some time. The noise floor is degraded by 4 dB, because there is more noise produced on the digital data bus, causing the analog input to pick up a little more. The false signals of higher order have disappeared completely and the third-order signal behaves as if the SDR-14 were an analog radio with an IP3 of +21 dBm and a noise figure of 15 dB. See Figure 6. Table 3 shows the measured data.

The third-order intermodulation seen here originates in the preamplifier. By sending the test signal into the direct input, one can avoid the preamplifier and the low-pass filter inside the SDR-14. With a dithering signal of –3 dBm at 7.22 MHz and two test tones at 10.7 MHz, each at –5 dBm, one finds the third-order signals at –95 dBm. Higher-order signals are visible at a similar level. The noise figure of this input is about 30 dB without any signals and about 36 dB with all three signals present. An RF amplifier that could drive the A/D converter without degrading the dynamic range much would have to have an output IP3 well above

Table 1

Power in dBm for signals of order N

N: 1	3	5	7	9	11
–17.5	–99.1	–107.7	–114.5	–117.4	–118.9
–22.5	–103.0	–108.7	–117.1	–126.4	–117.2
–27.4	–104.8	–114.2	–113.2	–121.8	–124.5
–32.4	–106.4	–113.0	–106.2	–108.8	–118.7
–37.3	–102.4	–106.1	–103.2	–103.3	–111.0
–42.4	–96.2	–100.7	–106.5	–117.1	–116.6
–47.2	–97.3	–108.7	–110.1	–116.7	–115.5
–52.0	–101.7	–108.7	–118.0	–122.2	–114.0
–57.0	–104.8	–123.8	–109.6	–114.4	–111.4
–61.9	–121.1	–104.5	–102.9	–109.0	–122.3
–66.7	–96.8	–106.0	–109.8	–116.2	–117.7
–72.4	–117.1	–130.5	–122.4	–131.9	–128.7
–77.3	–120.9	–122.0	–125.7	–125.3	–127.2
–82.1	–117.8	–119.0	–125.6	–137.8	—
–87.3	–116.1	–129.7	—	—	—
–92.3	–124.2	—	—	—	—
–96.6	–134.5	—	—	—	—
–100.1	—	—	—	—	—

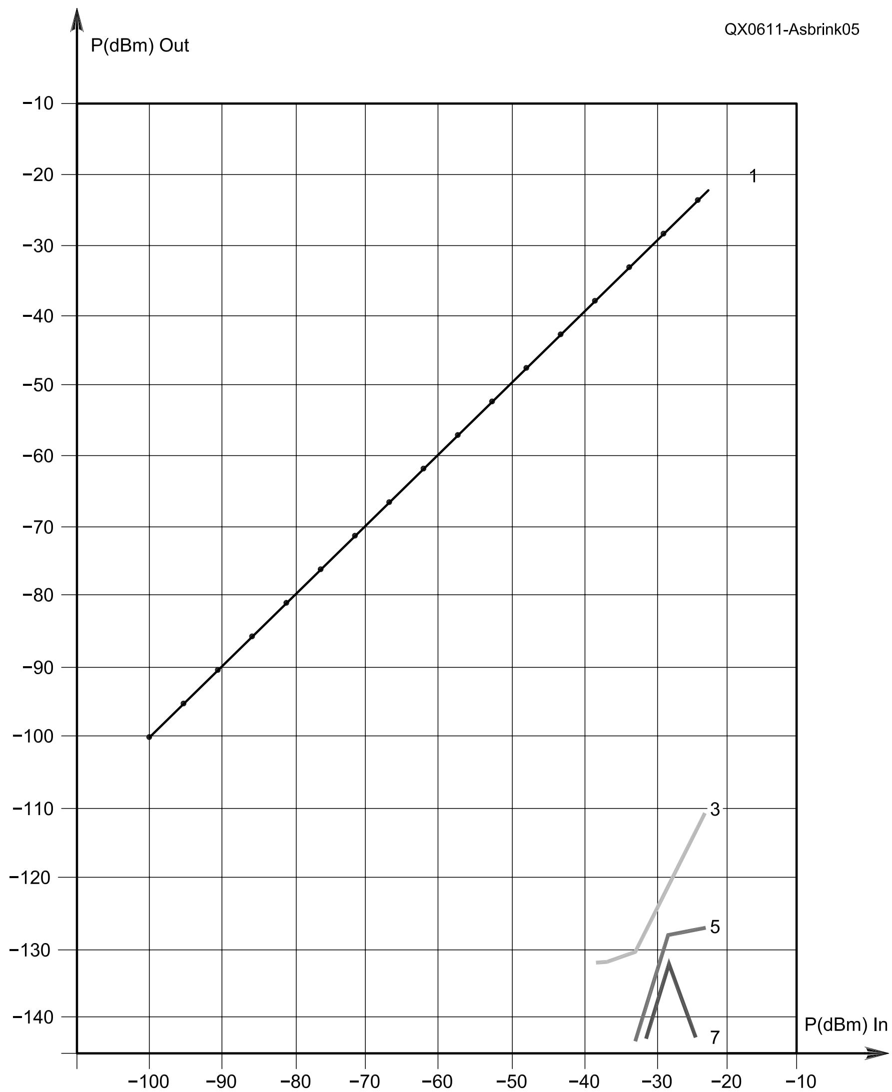


Figure 5 — With white noise used for dithering, the false signals are reduced significantly.

Table 2

Power in dBm for signals of order N

N: 1	3	5	7
–22.5	–110.1	–127	—
–27.6	–119.7	–128	–132
–32.6	–130	—	—
–37.5	–132	—	—

+40 dBm, which would call for a class A amplifier capable of delivering several watts. Such an amplifier with enough gain to place the antenna noise floor at –95 dBm in 500 Hz bandwidth would give a spur-free dynamic range of 90 dB at the same time as the noise floor is allowed to be lifted 15 dB by the preamplifier chain. For a 144 MHz weak signal operator this is very hard to accomplish in any other way, since the increase in the noise floor of about 15 dB that is necessary for a really low system noise figure is associated with a loss of dynamic range by about 15 dB in a conventional receiver.

Test Procedures

The standard procedure for measuring the two-tone, intermodulation-free dynamic range, used for example at the ARRL Lab, is to gradually increase the test signal until the third-order intermodulation signal equals the noise floor. While adequate for conventional analog receivers, although sometimes technically extremely difficult (on really good receivers), this procedure does not give a valid result for a digital radio like the SDR-14. Therefore, a different procedure should be adopted. It would give identical results as the old procedure on analog receivers but it would give a true figure of merit for digital radios. Rather than measuring what level is required for getting IM3 equal to the noise floor, one should measure the largest difference (in dB) between the test

Table 3

Power in dBm for signals of order N

N:	*1*	*3*
	–21.7	–105.8
	–26.7	–121.8
	–31.7	—

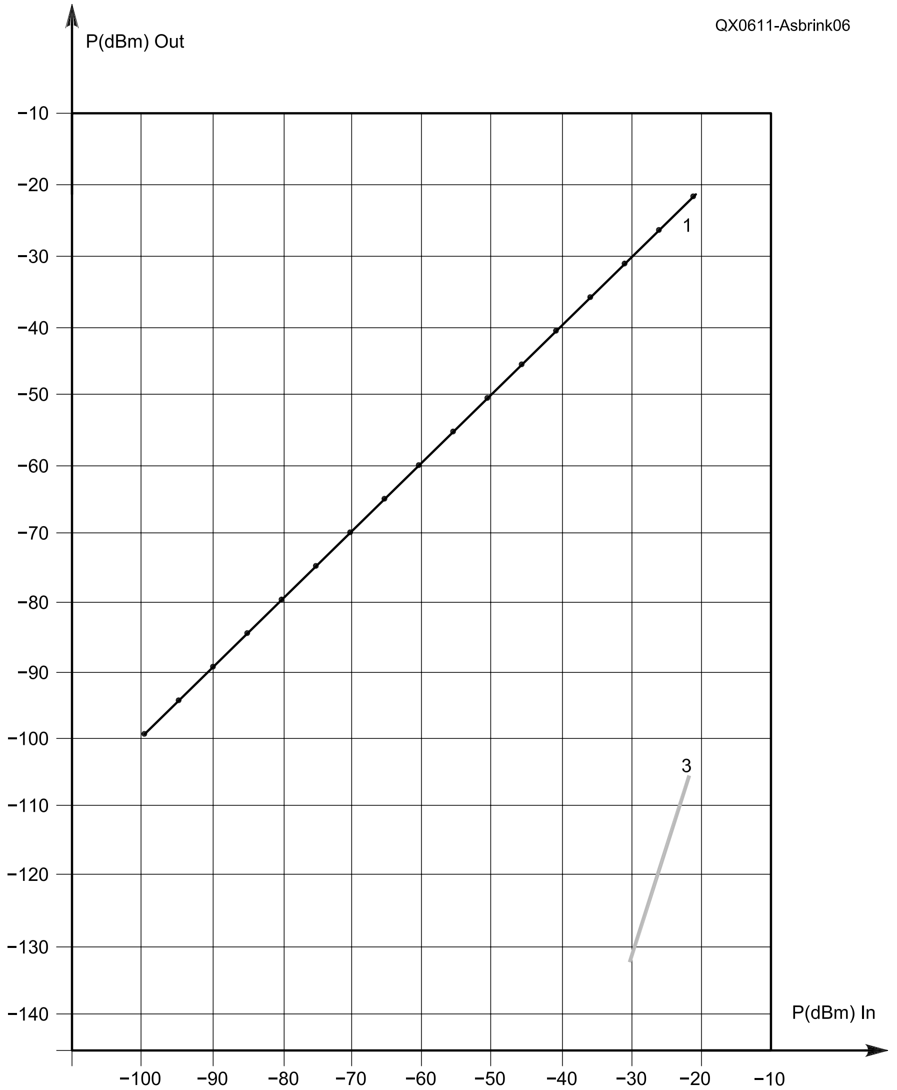


Figure 6 — With a frequency modulated carrier used for dithering, the false signals created by the A/D converter are weaker than the third-order intermodulation of the built-in preamplifier.

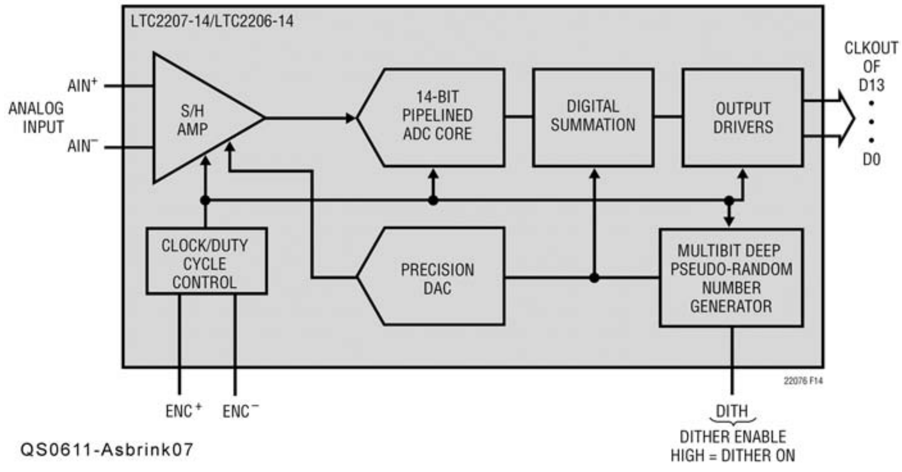


Figure 7 — The dithering used in LTC2206/2207. When dithering with noise, one normally has to incorporate filters that prevent any noise within the desired passband from reaching the A/D converter. This chip adds the same noise with opposite sign to the input and to the digital output. The cancellation means that the dithering noise will not be present in the output regardless of frequency, without any filters. (Copyright Linear Technologies, Inc, 2006, used with permission.)

tones and the intermodulation product. It will be close to saturation of the A/D converter on the SDR-14, while it will be at the noise floor for an analog receiver. Obviously the digital radio might need an RF amplifier to lift the external noise floor to the IM3 level observed in the test.

Both analog and digital radios need attenuators or amplifiers to place the external noise at the optimum level in order to get the maximum performance. Using an analog receiver with a very low noise figure on 40 meters in evening times without an attenuator will cause similar interferences as using a digital radio without an amplifier during the same circumstances!

The Future of Digital Receivers

The two causes of interference on intermodulation frequencies, A/D converter non-linearities and interference from the digital data lines, can be largely decreased by smart chip design. The LTC2207 from Linear Technology uses a random number generator that drives a D/A converter that adds noise to the input signal. The same noise is then subtracted at the digital side, as illustrated in Figure 7, taken from the data sheet.

The following excerpt from the LTC2207 data sheet illustrates the problems with the data bus:

"Interference from the ADC digital outputs is sometimes unavoidable. Interference from the digital outputs may be from capacitive or inductive coupling or coupling through the ground plane. Even a tiny coupling factor can result in discernible unwanted tones in the ADC output spectrum. By randomizing the digital output before it is transmitted off chip, these unwanted tones can be randomized, trading a slight increase in the noise floor for a large reduction in unwanted tone amplitude. The digital output is "Randomized" by applying an exclusive-OR logic operation between the LSB and all other data output bits. To decode, the reverse operation is applied; that is, an exclusive-OR operation is applied between the LSB and all other bits.

Leif was born in 1944 and licensed in 1961. He holds a PhD in physics and worked with research on molecular physics for 15 years at the Royal Institute of Technology. Since 1981, he has been running his own company developing various electronic products. He is the inventor of the magnetic intermodulation EAS system now owned by Checkpoint. Leif is essentially retired and works mainly with Amateur Radio related technical projects. QEX

In Search of New Receiver-Performance Paradigms, Part 1

Advances in receiver design force new approaches to certain tests. Some areas of receiver performance aren't getting nearly enough attention.

Doug Smith, KF6DX

Many well-established methods of modeling and measuring receiver performance fail to keep pace with the achievements of modern designs. Those methods must be revised or discarded in favor of something better. Other standardized methods are perfectly adequate as they are, but they're not seeing much use in Amateur Radio. Those tests deserve more of our attention.

We say that the primary goal of testing is to discover a unit's limitations. Underlying that goal is the necessity to relate those limitations to real conditions the unit is likely to encounter. Thus we strive to measure how well the goals of the designer and the user converge. Does the unit do what the designer intended? Was it designed to the realistic expectations of the user?

The aim of product evaluation should be to convey everything that matters to a user. Those things include not only spot performance data but also performance over temperature, time and vibration, ease and cost of repair, and warranty, among other things.

225 Main St
Newington, CT 06111-1494
kf6dx@arrl.org

Test Philosophy

What is the perfect receiver? It's a device that:

1) receives a desired signal, and
2) rejects all undesired signals.

From that definition it should be self-evident that receiver design heavily depends on the nature of both desired and undesired signals. Most receivers on the market today are general-purpose, but even some of those contain specialized circuits and subsystems. Our testing therefore depends entirely on intended usage, and tests must be designed with specific operating conditions in mind.

Receiver performance must be judged by its ability to achieve goal (1) — to receive an undistorted copy of what was transmitted — in the face of all possibilities for goal (2) — to reject interference. That judgment must be made over the range of environments in which a receiver is expected to operate.

Individual tests must be carefully de-

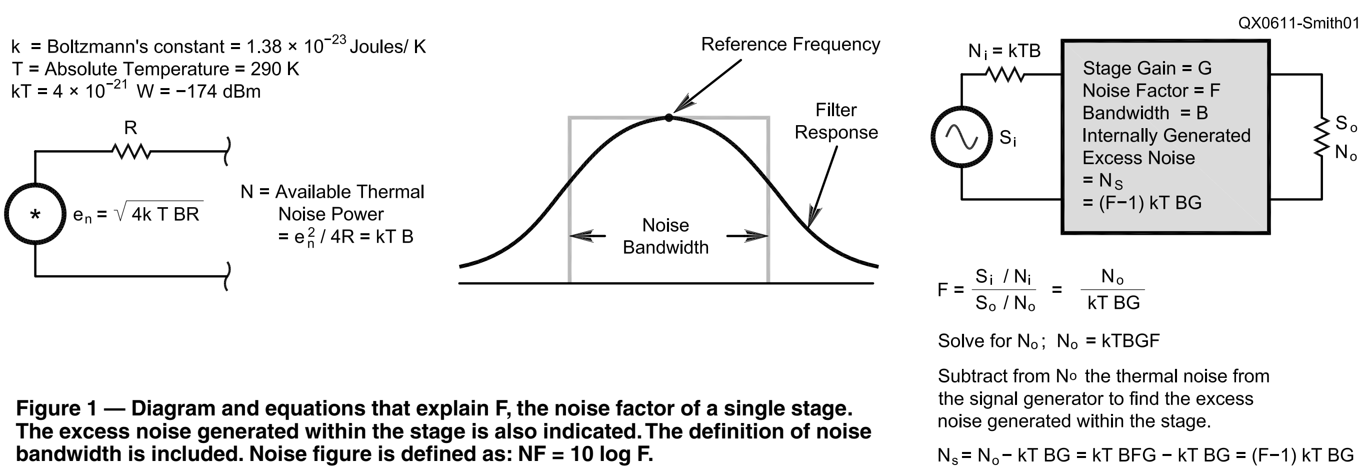


Figure 1 — Diagram and equations that explain F, the noise factor of a single stage. The excess noise generated within the stage is also indicated. The definition of noise bandwidth is included. Noise figure is defined as: NF = 10 log F.

signed to measure only a particular parameter of interest, to the exclusion of other effects. Instrumentation and calibration must be traceable to standards of known precision. Procedures must be documented well enough to allow others to exactly duplicate experiments and thereby support or refute the data obtained. Margins of uncertainty must be declared for all measurements.

In no way should a test procedure influence a definition of what is measured. The first step, then, is to define what we want to measure and why. Thus we start with goal (1) and go from there. The first two parts of this article series discuss certain critical receiver performance areas as a prelude to suggesting revisions in test procedures. We begin inside an isolated receiver.

The Enemies From Within

Receivers face at least two limitations even in the absence of external interference: noise and internal spurious signals or "birdies." To those limitations we can add issues of signal distortion, electromagnetic compatibility, ergonomics and both short- and long-term reliability.

Noise

Even in the absence of externally applied signals, receivers generate their own noise because of the thermally induced motion of electrons and atoms in their circuits. Accelerating electric charges produce changing electric and magnetic fields and thus, in a circuit, electromotive forces: voltages. In a conductor, very many free electrons are available: about 10^{23} per cubic centimeter. We say the electrons are free because it is easy for them to move amongst the atoms in the conductor while remaining bound to the conductor itself. That is a property of metals and certain other substances.

A common misconception is that electrons forming a current in a conductor move at the speed of light. Imagine a conductor whose cross-sectional area is about 1 cm^2: the equivalent of a 0000 AWG wire. If a current of 1000 A dc were flowing, the net electron velocity would be only about 1 cm/s — little more than a snail's pace! The effects of an applied voltage in a conductor, however, can travel as a wave at almost the speed of light.

Free electrons in a conductor also have a random thermal motion. They bop around in every direction at high velocities because of the effects of heat — about 10^7 cm/s at room temperature, or 20°C. So, thermal motion is much more significant at normal temperatures than that of net current.

Why? Because free electrons don't go far between collisions with neighboring particles. They don't necessarily collide in the sense of coming in actual physical contact, but particle physicists tend to use that term.

The resistance of a conductor can be shown to be proportional to the number of collisions. Resistance decreases with decreasing temperature because of lessening thermal motion and consequently, a decreasing number of collisions. For example, as the temperature of copper is reduced from room temperature to liquid-helium temperature (–269°C), its conductivity increases 1000 times.[1]

So, what we have is a churning electron stew at the front end and in every other part of a receiver, that, along with distributed gain, determines receiver noise performance. The available thermal noise power from a simple resistor is k*TB*, where k is Boltzmann's constant, *T* is the temperature in kelvins and *B* is the bandwidth in hertz.[2] At 20°C, that's about –139 dBm in a 3-kHz bandwidth, –147 dBm in a 500-Hz bandwidth, and –174 dBm in a 1-Hz bandwidth. See Figure 1. Such performance is what you'd expect to measure in a perfect receiver that adds no noise of its own. Of course, components in a receiver do add noise. The above numbers are theoretical limits for *noise-floor power*.

Chilling receiver components with liquid nitrogen (–196°C at atmospheric pressure) or even liquid helium improves their noise performance dramatically. That technique is often used for radio astronomy and off-Earth communications where sensitivity is critical.

A receiver's noise-floor power influences its ability to demodulate weak signals. It's defined as the lower limit of all dynamic-range measurements. It has traditionally been measured under pristine, single-signal conditions. Engineers accept a receiver's *noise figure* (NF in dB) as the ratio of output signal-to-noise ratio (SNR) to input SNR. NF is therefore bandwidth-independent.

In the old paradigm, noise-floor power was measured in a certain bandwidth, say, nominally 500 Hz. The trouble is that one manufacturer's nominal 500-Hz bandwidth might be closer to 350 Hz; another's, closer to 700 Hz. That difference would introduce significant bias in the comparison of noise-floor powers and make one receiver appear better than the other when in fact, they were equal. Noise-figure measurement eliminates

[1]Notes appear on page 30.

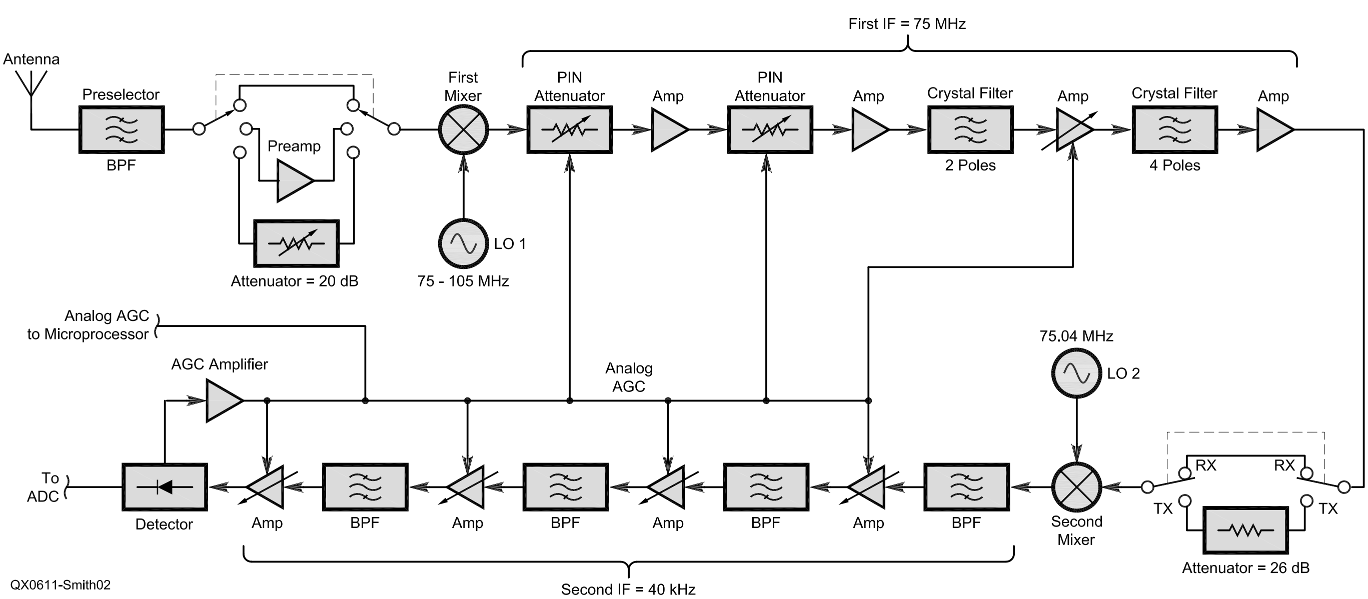


Figure 2 — Block diagram of an up-converting receiver with a 75-MHz first IF and a 40-kHz second IF.

the problem. Many well-documented NF measurement methods exist.[3, 4, 5]

Because of the nature of modern receiver architectures, NF should be investigated over the full range of input signals and gain settings. Users are interested in how SNR increases with increasing input signal strength. See the section on "Audio Quality" below for more on this topic.

In any case, noise-floor power is used in dynamic-range calculations and so cannot be discarded solely in favor of a NF declaration. To relate NF to noise power density, *normalization* is called for. In it, noise power density is stated relative to a 1-Hz bandwidth.[6] When bandwidth and the noise power in it can be accurately measured, you simply divide the power by the bandwidth. When logarithmic units of dBm are used for power and bandwidth, the operation becomes one of subtraction. Example: (–139 dBm) – (10 log 3000 Hz) = –174 dBm/Hz. Ultimately, that removes the bandwidth dependence of dynamic range declarations as described in Part 2 of this article.

Birdies

A birdie is an unwanted tone or other response in a receiver that is present even when external signals are not. Birdies are either:

1) Discrete internal signals that find their ways into the receiver's antenna input or intermediate frequencies (IFs) directly, or

2) Signals that result from two or more internal signals that find their ways into mixers in the receiver.

Designers must constrain signals of both types by proper shielding, grounding and design. Signals of the first type and their harmonics have to be identified by inspection. Signals of the second type are not always so easy to find since they are generally harmonics of one internal signal mixing with harmonics of another.

Predictions of birdies may readily be made with a computer program that iteratively computes $f_3 = mf_1 \pm nf_2$ and compares that frequency with both f_{IF} and f_{RF} for all possible values of f_1 and f_2, which are two internal oscillator frequencies, and where m and n are natural numbers. The *order* of a birdie is simply the sum of those two natural numbers, or $m + n$.

Start by using two of the receiver local oscillators (LOs) because they are already present at the mixers and they're usually high-level signals. Even if the LO signals themselves are clean, the mixers will generate significant harmonics. Up-converting rigs use LOs with harmonics that fall well into the microwave range. Containing all those harmonics between mixers and IF stages can be a formidable task, especially where square-wave LOs are used.

It's usual to step through the entire tuning range of a receiver in, say, 200-Hz or other small steps during a birdie search. High-order responses may be missed if the steps are not small enough. Note that f_1, f_2 or both may change with receiver tuning, and not necessarily at the same rate as the receiver's center frequency. With computer control of receivers, birdie searches may be automated advantageously.

A prediction program must look for any birdies that fall within the widest IF bandwidth of a receiver prior to the demodulator. Take as an example an up-converting superheterodyne design with a 75-MHz first IF and a 40-kHz second IF. See Figure 2. It employs a 75-MHz crystal roofing filter with a nominal bandwidth of 20 kHz. The birdie search, therefore, should report any birdie appearing within ±10 kHz of either 75 MHz or the tuned RF.

The first LO, a DDS-driven PLL, moves in approximately 1-Hz steps in the range 75.04 to 105.0 MHz and is responsible for all receiver tuning. Its relationship to the tuned receiver frequency, f_{RF}, is simply $f_{LO} = f_{RF} + 75.00$ MHz. The second LO is fixed at 75.04 MHz.

Figure 3 shows a compact QuickBasic program that predicts birdies up to the 30th order in that design by considering up to the 15th harmonic of each LO. Even-order harmonics aren't normally a problem when double-balanced mixers are used, but they're worth checking anyway. Odd-order harmonics in a square wave, such as are produced in a commutating mixer, tend to decrease in amplitude with increasing harmonic number in proportion to the inverse of the harmonic number, as in the sequence: 1/3, 1/5, 1/7, 1/9… So, the 15th harmonic is approximately 20 log 15 ≈ 24 dB down from the fundamental. Further, in a 75-MHz system, that harmonic lies at 15 × 75 MHz = 1.125 GHz and that's where the mixers ought to start to become "numb" — but you never know!

The program may readily be expanded to examine higher-order birdies by changing the two loop limits for *M* and *N*. Figure 4 is a sample of program output. The polarity sign appearing under the "TYPE" column indicates the "sense" of the birdie, as explained below.

The program artificially increments the first LO frequency when it finds a birdie of a particular type and order to eliminate a lot of

```
DEFDBL A-Z: REM ALL VARIABLES DOUBLE PRECISION
F2 = 75040000: REM LO2
F3 = 75000000: REM THE IF
F4 = 10000: REM MAXIMUM DELTA F FROM CENTER
STEPSIZE = 200: REM TUNING STEP IN HZ
COUNT = 1: REM TOTAL NUMBER OF RESPONSES

CLS : PRINT "SEARCHING..."
PRINT "LO1 ORDER", "LO2 ORDER", "TUNED FREQ", "TYPE", "BIRDIE FREQ"

FOR F1 = F3 + 100000# TO F3 + 30000000# STEP STEPSIZE: REM LO1 RANGE
 FOR M = 1 TO 15: REM LO1 HARMONIC
  FOR N = 1 TO 15: REM LO2 HARMONIC
   RF = F1 - F3: REM TUNED RF
   BIRD = M * F1 - N * F2: REM BIRDIE FREQUENCY
    IF ABS(ABS(BIRD) - F3) <= F4 THEN TYP = 1: GOSUB OUTBIRD: F1 = F1 + M * STEPSIZE
    IF ABS(ABS(BIRD) - RF) <= F4 THEN TYP = 0: GOSUB OUTBIRD: F1 = F1 + M * STEPSIZE

 NEXT N, M, F1

PRINT "COUNT="; COUNT: REM TOTAL RESPONSES FOUND

RESET: SYSTEM: REM QUIT

OUTBIRD: REM FORMATTING TO GET THE DECIMAL PLACES TO LINE UP
  PRINT USING "###"; M;
   PRINT " ",
    PRINT USING "###"; N;
     PRINT " ",
      PRINT USING "###.####"; RF * .000001;
       IF TYP = 1 THEN PRINT , "IF";  ELSE PRINT , "RF";
         IF SGN(BIRD) >= 0 THEN PRINT "+",  ELSE PRINT "-", :REM THE BIRDIE "SENSE"
      PRINT USING "###.####"; ABS(BIRD) * .000001: REM ROUND UPWARD
  COUNT = COUNT + 1
 RETURN
```

Figure 3 — QuickBasic 4.5 program listing for birdie search of 75-MHz up-converting LF/HF receiver from 100 kHz to 30 MHz.

the duplicate entries that would otherwise appear. It computes the actual frequency and the sense of each birdie, and reports it. A positive sense means that the birdie moves upward in frequency (either at IF or RF, depending on birdie type) with increasing receiver center frequency. A negative sense means the birdie moves downward in frequency with increasing receiver center frequency. Note that in the design of Figure 2, an IF birdie with a positive sense would have a negative sense with respect to RF, since the first mixer inverts the sense. Searching in that way, the program produces a little over 100 results, each of which must be checked starting with the low-order responses. A few duplicates still appear because of the small f_{RF} step size used (200 Hz). SSB mode is recommended for confirmation tests, and an audio spectrum analyzer (or one on the IF) can be used to detect birdies that are inaudible or slightly outside the final passband.

The tester must search for birdies at the predicted frequencies but also he must run a general check across bands of interest for unpredicted birds. The test might take all night; but if it's automated, who cares? You can often identify high-order signals by tuning the receiver by amount Δf and taking the ratio of how much the birdie moves to Δf. The ratio definitely indicates something about the source of the birdie by inspection of the receiver architecture. It's sometimes amazing how high-order harmonics of internal oscillators, lying in the microwave region, can find their way into and out of mixers.

Digital control systems and synthesizers tend to generate signals that are not integrally related to their clock frequencies. Prediction of those spurious signals can be problematic and that's why a general birdie search must be conducted, regardless of other predictions.

These days, a reasonable design goal is to have *no birdies greater in power than the noise floor of a receiver*. I want to add more emphasis there than simple italics can convey! When designers first began producing receivers with VHF first IFs several decades ago, some atrociously birdie-laden equipment hit the streets. Ham-radio manufacturers of HF equipment had a bit of an advantage in that they were primarily concerned with only six rather narrow bands. They have the disadvantage, of course, of having to cover almost five octaves of spectrum. Now, general-coverage receivers are the norm in LF/MF/HF equipment and there are nine HF ham bands; many rigs also cover VHF and even UHF bands.

DC-to-daylight receivers that continuously cover spectrum from LF to UHF and beyond are common in commercial and military surveillance applications. Such units generally up-convert to microwave IFs and so contain UHF or even SHF LOs. Tracking and eliminating birdies is therefore as difficult as ever but it deserves more attention in receiver testing than it is currently getting.

In one form or another, direct digital synthesis (DDS) is almost ubiquitous in modern receivers. DDS is a brute-force way of generating a sine-wave LO that uses a numerically controlled oscillator (NCO) coupled to a digital-to-analog converter (DAC). See Figure 5. Its main drawback to date has been spectral impurity. Both discrete and broadband AM and PM impurities may be characterized and analyzed.

Some of the spurs at the output of a DDS are harmonics of the output signal, and the fundamental and harmonics of the clock. Harmonics of the output signal are chiefly caused by nonlinearities in the DAC. Harmonics that fall more than half the clock frequency away may cause unexpected birdie problems because they can produce aliases that fold back to frequencies at less than half the clock frequency. Those aliases may appear at the IF or at the tuned RF of a receiver.

To predict the frequencies of those spurs,

SEARCHING...

LO1 ORDER	LO2 ORDER	TUNED FREQ	TYPE	BIRDIE FREQ
14	15	5.0394	RF–	5.0484
13	14	5.3966	RF–	5.4042
12	13	5.8086	RF–	5.8168
13	15	5.8148	IF–	75.0076
11	12	6.2892	RF–	6.2988
12	14	6.2960	IF–	75.0080
10	11	6.8574	RF–	6.8660
11	13	6.8646	IF–	75.0094
9	10	7.5390	RF–	7.5490
10	10	7.5408	IF+	75.0080
10	12	7.5470	IF–	75.0100
11	12	7.5490	RF+	7.5590
8	9	8.3724	RF–	8.3808
9	9	8.3740	IF+	75.0060
9	11	8.3812	IF–	75.0092
10	11	8.3830	RF+	8.3900
7	8	9.4138	RF–	9.4234
8	8	9.4152	IF+	75.0016
8	10	9.4238	IF–	75.0096
9	10	9.4254	RF+	9.4286
6	7	10.7530	RF–	10.7620
7	7	10.7542	IF+	74.9994
13	15	10.7566	RF–	10.7642
7	9	10.7644	IF–	75.0092
8	9	10.7658	RF+	10.7664
12	14	11.5808	RF–	11.5904
5	6	12.5384	RF–	12.5480
6	6	12.5394	IF+	74.9964
5	6	12.5408	RF–	12.5360
11	13	12.5426	RF–	12.5514
12	15	12.5492	IF–	75.0096
6	8	12.5518	IF–	75.0092
7	8	12.5530	RF+	12.5510
6	8	12.5546	IF–	74.9924
10	12	13.6792	RF–	13.6880
11	14	13.6864	IF–	75.0096

Figure 4 — Sample output from birdie search.

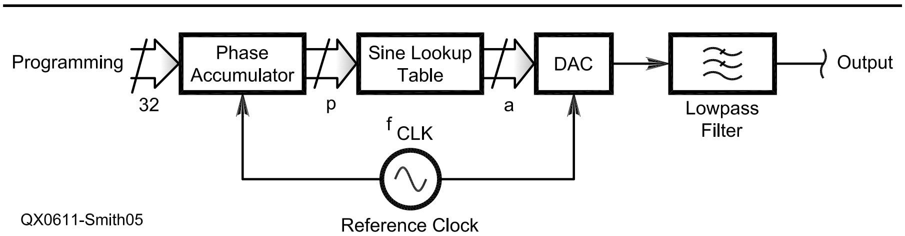


Figure 5 — Block diagram of a direct digital synthesizer (DDS).

start by computing the ratio, k, of a particular integral harmonic, n, of the LO frequency, f_{LO}, to half the sampling frequency, f_s: $k = (nf_{LO}) / (f_s / 2) = (2\ n\ f_{LO}) / f_s$. Define $a = \mathrm{I}(k)$ as the integer part of k. Define $b = \mathrm{F}(k)$ as the fractional part of k. For all odd a, spurs appear at $(0.5 - b) f_s$; for all even a, spurs appear at $b\, f_s$.

It's also possible to get discrete spurs and broadband noise at both integral and fractional subharmonics of the output frequency and clock. Those are traceable to nonlinearities in the DAC, to the limitations of the phase accumulator, or to jitter in the clock. With today's large phase accumulators in NCOs, significant spurs are almost always AM and not PM spurs.

AM spurs can virtually be eliminated by squaring the output of a traditional sine-wave DDS. That's almost what happens in a double-balanced mixer, but not quite. A separate squaring circuit with hysteresis (like a high-speed logic gate or the input to a PLL chip) will do the job. However, such a squaring circuit may generate lots of odd harmonics of the LO and spray them throughout the receiver — more possible birdies. Placing an amplifier and the squaring circuit immediately adjacent to the mixer mitigates the birdie problem, since only a low-level, relatively clean LO signal need be supplied from synthesizer to mixer. I shall have more on the hazards of DDS spurs in Part 2.

Electromagnetic Compatibility (EMI/RFI)

The enemy from within here is the unwanted emission of signals from the unit under test. FCC, CE and other standards require that radiated and conducted emissions from equipment be held under certain limits. In the US, *Title 47 Code of Federal Regulations (CFR) Part 15* defines those limits for two classes of equipment: class A and class B. Class-A equipment is intended for commercial use; class-B equipment is intended for domestic use.

In general, the CE (European) requirements are much more stringent than those of Part 15. In addition, for CE certification, an accredited laboratory must test units. The tests include spark discharge and external electromagnetic field susceptibility. In the US, Amateur Radio equipment can be simply verified by the manufacturer. For every market, however, *equipment must be tested and must carry a sticker or label attesting to compliance.*

Any equipment sold in the US that uses digital signals of significant frequency must carry a label with the following, or similar, wording:

This device complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions: (1) this device may not cause harmful interference, and (2) this device must accept any interference received, including interference that may cause undesired operation.

Furthermore, a statement for class-A devices must contain the following or a similar statement in the instructions provided to the user:

This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

For a class-B device, the following or a similar statement must be included:

This equipment has been tested and found to comply with the limits for a Class B digital device, pursuant to part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference in a residential installation. This equipment generates, uses and can radiate radio frequency energy and, if not installed and used in accordance with the instructions, may cause harmful interference to radio communications. However, there is no guarantee that interference will not occur in a particular installation. If this equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, the user is encouraged to try to correct the interference by one or more of the following measures:

- *Reorient or relocate the receiving antenna.*
- *Increase the separation between the equipment and receiver.*
- *Connect the equipment into an outlet on a circuit different from that to which the receiver is connected.*
- *Consult the dealer or an experienced radio/TV technician for help.*

Conducted emissions from receivers include LO and other signal leakage to the antenna. Conducted emissions also include those induced by equipment onto the ac mains. Power supplies not incorporating power-factor correction can generate surprising amounts of RFI because they are conducting only during small portions of the ac mains waveform. Those types have virtually been banned from Europe. Switch-mode power supplies employ LF, MF and even HF frequencies and must be checked for conducted and radiated emissions. European requirements *are* much tougher than those here in North America.

Although some argue on the grounds of ambiguous regulatory wording that Amateur Radio equipment is exempt from Part-15 requirements, some ham radio manufacturers disagree. You will find Part-15 and CE stickers on their gear. They believe that the FCC intended to protect consumers from interference from *all* sources. Unfortunately, that belief is not yet universal, and it remains a buyer responsibility to look for those stickers. The fact remains: Virtually all amateur receivers produced these days qualify as Class-B unintentional radiators. Those operating above 30 MHz may be subject to the FCC's verification or certification rules, as well.

Control Systems and Operating Manuals

Control system theory deals with dynamic systems that are under either manual or automatic control to achieve a desired response. Modern receivers have both manual and automatic control systems that need testing.

In our general terms of reference, a control system:

1) elicits input from a source, whether it be from an operator or other means;

2) provides information about system response to inputs; and

3) allows changes to inputs based on feedback.

When a human being is in the feedback loop, we say that control is manual; when feedback is incorporated without an operator's intervention, we say it's automatic. That distinction is the difference between manual and automatic notch filters, for example.

In an erstwhile era, things were simpler than they are today. Manual controls were dedicated to single functions because they were hard-wired switches, potentiometers, variable capacitors and inductors. Today, we have sophisticated microprocessor-based control systems that are defined by software or firmware. Many of those systems have grown to such complexity that the myriad of control functions available must be carefully organized into hierarchical structures. More than a few "knob-twisters" have found the assignment of multiple functions to single controls and the depth of elaborate menus difficult to navigate. So we're faced not only with the basic issue of efficient control but also those of ergonomics.

In fairness, we should admit that modern receivers incorporate advanced features that were never dreamed of in yesteryear. The challenge, then, is to present controls to the user in a logical manner while minimizing the number of them on the control panel. Actually, designers have always grappled with that challenge and they always will.

Most modern rigs incorporate extended control through serial, USB or Ethernet ports. They propagate feedback and other information through those interfaces, which are often used in logbook programs, antenna-heading controls and the like. Full remote control of receivers — complete with audio, command and control — has been achieved in several designs of late. Those features deserve testing. What is a receiver's response time or *latency*? What is the likelihood of malfunction because of command corruption in the control system?

An engineer's being "cooped up" in a laboratory for months on end doesn't necessarily produce the best ergonomic design. From the beginning, and at regular intervals continually, designers must expose their control schemes to outsiders who ideally know little or nothing about a particular receiver. Designers thereby get feedback about what's intuitive and what's not. That process is a control system in itself because the loop is closed between input and output around a feedback path.

Admittedly, a control system that is very new and different, yet superior to antecedent technologies, may have a steep learning curve; but every new piece of equipment should improve on its predecessors and intuition often equals the total of a user's previous experience. Spending a day or two with a new receiver may be a good gauge of intuition but not necessarily a good evaluation of the system. The talents of an evaluator are therefore likely a large part of subjective test results — a situation to be avoided.

So what's the right way to engage in objective control-system testing? Well, one way to start is to develop a test plan. Manufacturers should have control-system test plans because qualifying a receiver for release to production requires them. Whenever there is a design change, the plan identifies unintended affects on the system. Testers should have test plans because they cannot always anticipate combinations of inputs that produce unexpected responses.

First, simply verify that the receiver does what the manufacturer says it does. Follow the operating manual and rate its accuracy, noting any anomalies. Measure system response against the information provided by the control system. Second, assemble matrices of control inputs versus expected responses and identify mutually inclusive or exclusive responses. For example, automatic notch and noise reduction may fight one another to achieve an intended goal: The notch tries to eliminate continuous tones and the other tries to enhance them.

Third, measure the effectiveness of automatic control systems within a receiver. What is the depth of the automatic notch filter? What is its response time? What is the signal-to-noise-ratio improvement afforded by a receiver's noise-reduction algorithm? What are the capture and hold-in ranges of a receiver's synchronous AM mode? What is the automatic gain control (AGC) overshoot versus signal level?

Fourth, place emphasis on the logical organization of physical controls. Are knobs, buttons and switches used for multiple functions? What is the interaction among them? Are the controls adequately accessible? Is positive auditory or visual feedback provided for each control? Could a blind or deaf person operate the unit? Can the system be reset to a known configuration should the operator get lost?

Finally, can you crash the control system? Try the notorious "dead-man" test, in which controls are operated in simultaneous, sequential and random fashion in an attempt to find anomalous behaviors. If you do find a fatal error, does cycling the power bring the unit back to life?

An installation and operation manual should impart everything a user needs to install a product and get it up and running. It should also include information about how to investigate and solve common problems during setup that cause the product not to perform as intended. Manuals get only cursory examination these days in receiver testing, but they are critical to the correct use of products and must be evaluated right along with the equipment itself.

Prominent warnings are appropriate where the product or procedure presents any danger to the installer. That is especially true for ac-powered equipment. Product liability lawsuits are a bane of manufacturers, particularly where the installer or user was not reasonably made aware of hazards. In some instances, the first thing the unpacker should see is a one-page warning sheet on fluorescent-colored paper atop the package contents. It has been shown that those steps go a long way toward defraying liability for personal injury or death.

It may be that a defect in, or misuse of, the product would present a danger to the user. Manufacturers of ladders, for example, have had the "cheese" sued out of them for not warning of inappropriate usage, of which a reasonable person would not be aware. A chain-saw manufacturer, believe it or not, needs to warn against cutting anything but the type of wood for which it was designed. We see some silly cases, of course, such as the desiccant manufacturer who found it necessary to place on his product the warning, "Do not eat. Not for internal use." You can bet, though, that the warning arose from a lawsuit in which someone ate the moisture-absorbing chemical.

Most states have laws that address merchantability and fitness of purpose. Those laws go to what is called an implied warranty — that the product is safe and fit for its designed purpose. Design flaws and production defects may cause a breach of that implied warranty. Admittedly, receivers don't generally pose much of a health threat. Wherever you're dealing with the ac mains, though, due diligence is always in order.

A good step-by-step installation procedure indicates at the start what tools or other user-supplied equipment is necessary to complete installation and check-out. Pay careful attention to the temporal order of the steps. It may be bad, for instance, to plug a cable into a computer when the power is on. The installer should be informed of such hazards and what damage they can cause if he does not follow the procedure.

There has been an unfortunate tendency to organize operational documents according to the product controls. While that method is all right to help an operator identify controls, detailed operating information is better organized according to what the user is trying to do. In conjunction with a good index, an operating section is much easier for a user to navigate and understand that way.

That's not to state that exhibits showing all controls, connectors and displays aren't good ways to begin — they are. Some simple drawings or photographs, with reference designators indexed to tables and accompanied by brief verbal descriptions, should be sufficient.

A key ingredient of a good operating manual is a set of explanations of why each feature is useful in doing what the user is trying to do. Once a reader understands why he needs a feature and some of the theory behind "what makes it tick," you will have a more adept user. Again, the talents of the operator may come into play when evaluating an operating manual but details of its organization are worth reporting, including whether theory of operation, schematics, parts lists and detailed troubleshooting and maintenance information are provided.

Audio Quality

How a receiver sounds depends on the performance of several key subsystems: AGC; RF, IF and AF in-band intermodulation distortion (IMD); total harmonic distortion (THD); audio response and acoustic design, including loudspeaker size, quality and placement; enclosure design and peak audio power output capacity. Traditionally, measurements of THD and in-band IMD are made with analyzers connected electrically to the loudspeaker terminals; but that technique ignores the effects of acoustic design. Better would be to add measurements taken from a microphone closely coupled to the loudspeaker.

A receiver's nominal bandwidth settings are almost always stated in its technical specifications but what are often not stated are frequency response, passband ripple and shape factor. All four parameters are easy to measure with an audio spectrum analyzer's peak-hold function when only noise is present in the receiver. Most current test regimens measure only the –3-dB or –6-dB points of the passband. Tests should be conducted both electrically at the loudspeaker or ancillary audio output terminals, and acoustically. A receiver's frequency response might be perfect electrically, but its acoustic response is almost certainly not. IF shift and other passband controls should be exercised over their entire ranges to ascertain their effects on frequency response limits.

In-band distortion includes THD and IMD. Those parameters need to be measured under realistic conditions. An unfortunate tendency has been to measure in-band distortion with AGC set to fast or off. It makes no sense to defeat the subsystem that's designed to prevent the very thing you're trying to measure![7] Fast AGC produces envelope distortion on two-tone signals that wrecks IMD performance. Instead, measure IMD with the AGC set to its slowest decay time. There is no reason that tone spacing should be any different than in transmitter tests, although both wide (1.8 kHz) and narrow (100 Hz) spacings would be desirable. THD should be measured at rated output power and also at normal listening levels.

Multi-mode receivers often suffer more from THD in AM and FM modes than in SSB. Modern units using DSP and analytic signal processing (the "I/Q" method) may be particularly prone to THD in their demodulators. THD should be reported for each operating mode.

AGC transient response to on-channel signals may have a decisive impact on receiver audio quality. That is discussed in Part 2 as being dependent on conditions outside the receiver: large and rapid excursions in received signal amplitude.

A proper noise-floor measurement discovers signal-to-noise ratios (SNRs) of receivers near their noise floors but says nothing about how SNR increases with increasing signal levels. We're interested in how quiet a receiver gets when the received signal is large. Only by careful distribution of gain and gain control among a receiver's stages will SNR increase to large values with large signals. It may be argued that SNRs greater than about 50 dB aren't worth pursuing, since the human hearing system has trouble detecting anything that far below a desired signal at normal listening levels. The ultimate SNR of a receiver, however, is a sure figure of merit of its design quality. Background hiss and noise, especially when the receiver is muted, squelched or when the audio volume is turned down, are also of interest.

Reliability, Serviceability, Warranty

How long will a unit last? If it breaks, will the fault be covered under warranty? For how long? If out of warranty, how much time and money will it cost to get it fixed?

Run-in and life testing should be parts of any manufacturer's test program. Run-in is generally conducted after a unit is assembled and its basic functionality confirmed. It's usually performed under a limited set of realistic but extreme operating conditions, often at elevated temperatures or under other environmental stresses. The idea is to weed out infant mortalities in components. Some manufacturers elect to run units for a day or two under those conditions and some may go as long as several months, depending on the reliability level desired. High operating temperatures tend to accelerate the path down the "bathtub curve" of failures versus time.

Certain commercial and military units require a specification of mean time before failure (MTBF). The US military provides guidelines for calculating MTBF for nomenclatured electronic equipment.[8] Each component part contributes to the MTBF. Relays and complex integrated circuits, like microprocessors, generally hurt reliability the most, but passive components like electrolytic capacitors also stand out.

Some would say that life testing of Amateur Radio gear is impractical but I disagree. Many receiver models have sold thousands of units and so the data are there from actual users. Web sites like **www.eham.net** compile user comments but it is hardly a scientific evaluation. We could do better by formally standing units against our yardsticks of expected reliability.

Mean time to repair (MTTR) is also an issue. Stories of both horror and joy can be found in Amateur Radio cyberspace, but they do not form a homogenous picture of service speed and quality. To remove bias, what's needed is a holistic approach. Warranty information needs to be stated in product evaluations and it wouldn't be too hard to purposefully break a rig or two, contact factory service and see what happens.

Having written that, however, I must add that reliability and serviceability information needs to come mainly from manufacturers, since it's beyond the abilities of most testers to gather consistently. Asking manufacturers for that information at evaluation time would give them the chance to "toot their own horns" and to show what separates them from others. You might think that every manufacturer would say, "Ours is the best," but pressing them for specifics about quality assurance and customer service is not out of line. On a rolling production line, serious equipment makers do documented inspections, spot checks, life testing and record keeping. Witness the ISO-9000 series of specifications pertaining to such things.[9]

A scientific survey or two wouldn't hurt. We could benefit ourselves by holding manufacturers to account for their product quality.

Product quality assurance extends all the way through the entire production process, from purchasing to shipping. Good internal documentation allows anyone in the production chain to have enough information to do a high-quality job. Without a specification document, for example, I might not know whether to buy a 16-V tantalum capacitor or a 25-V capacitor. Equipment buyers are interested in that and other procedures used by manufacturers because they give insight into general quality. It doesn't all have to be published; interested parties could be referred to the manufacturer's Web site.

Summary

In this part, we've covered the enemies of receiver performance from within. In Part 2, we'll take a look at the enemies from without: temperature, vibration, power supply variations, electromagnetic susceptibility, and those limitations placed on receiver performance by both wanted and unwanted signals at the antenna.

Acknowledgment

This paper was graciously edited by *QEX* Contributing Editor L.B. Cebik, W4RNL. His insight and aplomb are greatly appreciated.

For Further Reading

IEC 60315-6, *Methods of measurement on radio receivers for various classes of emission — Part 6: General purpose communication receivers,* International Electrotechnical Commission, ed. 1.0 b: 1991.

ISO/IEC 17025, *General requirements for the competence of testing and calibration laboratories*, International Organization for Standardization/International Electrotechnical Commission, 2005.

ITU-R Recommendation SM.331-4, *Noise and sensitivity of receivers*, International Telecommunication Union, 1978.

U. Rohde, N1UL, "Receiver Measurements, How to Evaluate Receivers," *QEX*, Jul/Aug 2005, pp 3-11.

D. Smith, KF6DX, "Tech Notes: Quantifying Measurement Uncertainty," *QEX*, Jan/Feb 2006, pp 54-56.

D. Millar, K6JEY, "Test Procedures," *The ARRL Handbook of Radio Communications*, 2006 edition, p 25.43.

M. Tracy, KC1SX and M. Gruber, W1MG, *Test Procedures Manual*, ARRL, Rev H, Jun 2004, **www.arrl.org/members-only/prodrev/testproc.pdf**.

W. Sabin, WØIYH, "Measuring SSB/CW Receiver Sensitivity," *QST*, Oct 1992, pp 30-34.

M. Tracy, KC1SX, "Product Reviews — In Depth, In English," *QST*, Aug 2004, pp 32-36.

R. Sherwood, WBØJGP, and G. Heidelman, K8RRH, "Present-Day Receivers — Some Problems and Cures," *ham radio*, Dec 1977, p 10-18.

Notes

[1]E. Purcell, *Electricity and Magnetism*, Berkeley Physics Course, Vol. 2, McGraw-Hill, 1965, ISBN 0-07004-908-4.

[2]A. Einstein, "On the Motion Required by the Molecular Kinetic Theory of Heat of Small Particles Suspended in a Stationary Liquid," *Annalen der Physik*, 1905.

[3]A. Helfrick, K2BLA, "Measuring Noise Figure," *ham radio*, Jan 1984, pp 26-30.

[4]B. Olson, W3HQT, "Noise Figure Measurement," *QEX*, Nov 1986, p 12.

[5]F. Brown, W6HPH, "A Do-It-Yourself Noise Figure Meter," *QEX*, Jul/Aug 2005, pp 52-53.

[6]W. Sabin, WØIYH, "Receivers and Transmitters," *The ARRL Handbook of radio Communications*, 2006 edition, D. Straw, N6BV, ed, ISBN 0-87259-948-5, p 14.8. The ARRL Handbook is available from your local ARRL dealer, or from the ARRL Bookstore, ARRL order no. 9485. Telephone toll-free in the US 888-277-5289, 860-594-0355, or fax 860-594-0303; **www.arrl.org/shop/**; **pubsales@arrl.org**.

[7]Although first pointed out over eight years ago, it's still being done. See D. Smith, "Signals, Samples and Stuff, Part 2," *QEX*, May/Jun 1998, pp 22-37.

[8]MIL-HDBK-217, US Government Printing Office, Washington, DC.

[9]International Organization for Standardization, **www.iso.org**.

Doug Smith, KF6DX, has edited QEX *since 1998.*

Seventh-Order Unequal-Ripple Low-Pass Filter Design

The author discusses the advantages of unequal-ripple low-pass filters.

Dave Gordon-Smith, G3UUR

Equal-ripple Chebyshev low-pass filters offer increasing harmonic attenuation as the ripple in the passband is allowed to increase, and a design with a ripple of 1 dB gives significantly more harmonic suppression than a 0.01 dB design. However, the latter normally have to be used to present broadband RF power amplifiers with a close match to 50 Ω over a reasonably wide frequency range. In power amplifier applications, though, only about one-third of the upper part of the passband is utilized because the second-harmonic attenuation would be insufficient for lower fundamental frequencies. It doesn't matter, therefore, what sort of response the filter provides at frequencies lower than about half of the cutoff frequency in this application.

A band-pass response would be sufficient if it were not a component-inefficient way of achieving the required attenuation in the stopband. A better solution is to use an unequal-ripple design, which has high ripple in the lower half and low ripple in the upper half of the passband, as shown in Figure 1. The response in the upper portion of the passband can be made to offer a low VSWR (less than 1.1:1) over up to 45% of the –3 dB bandwidth, depending on the depth of the lower valley and the ripple chosen for the equal-ripple upper part. Such designs can offer considerably more harmonic suppression than equal-ripple low-pass filters with a similar low VSWR in the used part of the passband, as well as more convenient capacitor values.

This article describes the approach used by the author to design unequal-ripple filters, and also discusses the network constraints required to achieve a low-ripple region at the upper end of the response of a 7th-order low-pass design with a deep valley in the lower half. In addition, it presents a table of normalized coefficients (See Table 1), and describes how this table can be used to provide pi- or T-configuration low-pass filter designs that use standard preferred-value capacitors, with little or no compromise in the cutoff frequency. Table 2, a table of 3-dB-down k and q values for unequal-ripple low-pass filters, is also presented as a starting point for those who require filter specifications that are not covered by the designs presented in this article.

Whitehall Lodge, Salhouse Rd
Rackheath, Norwich,
Norfolk, NR13 6LB, UK
g3uur@dg-s.fsnet.co.uk

Theory

Filters that have equal ripple in the upper part of the passband and a deep valley in the lower part are known as "acromorphic" filters. Their name is derived from Greek, and reflects the special shaping (*morphe)* of the high end (*akros*) of the passband to tailor the response for broadband amplifier harmonic filtering applications. All the poles of acromorphic low-pass filters lie on an ellipse, apart from the one at ω = 0. This lies inside the ellipse by an amount that increases as the lower valley ripple (LVR) increases. The other poles also move round the ellipse away from the negative real axis as the LVR increases, and their angular separation also decreases as they do so, to maintain equal ripple in the upper part of the passband. Calculating the positions of the poles for a particular response is a tedious and time-consuming process. A quicker way of designing a low-pass filter with an unequal-ripple response is through curve fitting. The amount required to be added to 1 in the denominator of the power transfer function (Equation 1) at various radian frequencies between zero and ω = 1 can be calculated from the desired passband. It is convenient to make the additional amount required to produce the deepest valley in the passband equal to 1, and also to make this the level at ω = 1 as well. The additional amount required at each peak in the response is zero. Then the polynomial can be scaled by a factor ε^2 to provide the right variations with respect to 1 for the desired passband ripple, as is done for Chebyshev designs. The starting point for designing a low-pass filter with an arbitrary response using curve fitting is the power transfer function,

$$|H(j\omega)|^2 = \{1 + \varepsilon^2 P_n^2(\omega)\}^{-1} \qquad \text{(Eq 1)}$$

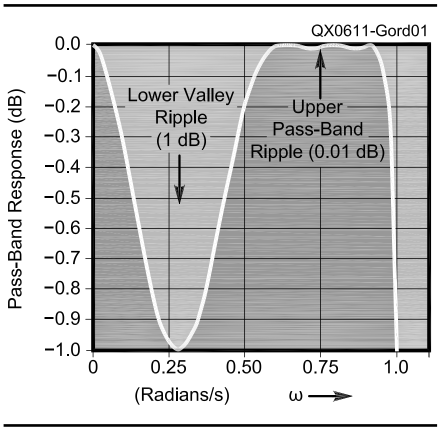


Figure 1 — Predicted passband response for normalized 1 dB/0.01 dB unequal-ripple (acromorphic) low-pass filter (normalized to ω = 1 at 1 dB down).

$P_n^2(\omega)$ is a polynomial in even powers of 2n, where n is the order of the filter being synthesized. It is the square of the polynomial on which the filter response is based. The squared polynomial that gave the best fit for the author's non-standard 7th-order design with a 1 dB dip at 0.28 ω and three equal-ripple peaks and two valleys of 0.01 dB at the upper end of the response was

$$P_n^2(\omega) = 754.15\ \omega^{14} - 2800.97\ \omega^{12} + 4257.98\ \omega^{10} - 3387.15\ \omega^8 + 1485.39\ \omega^6 - 340.18\ \omega^4 + 31.77\ \omega^2 \qquad \text{(Eq 2)}$$

This was obtained by using a scientific curve fitting routine on a mainframe computer, and only 11 data points were needed, including the zero at ω = 0. Today, this can be done on a PC at home. The positions of the peaks and valleys on the radian frequency axis had to be shuffled to reduce the RMS error of the fit to a low level. The same polynomial can be used for a variety of designs with the same shape of curve, but different levels of upper passband ripple and lower passband dip, just by changing the value of ε. The corresponding filter polynomial, on which Equation 2 is based, is

$$27.462\,\omega^7 - 50.998\,\omega^5 + 30.173\,\omega^3 - 5.637\,\omega \quad \text{(Eq 3)}$$

It can be seen that Equation 3 is similar in form to a Chebyshev polynomial, but with quite different non-integer coefficients.

To find the normalized values of the components of the filter being designed, the coefficients from Equation 2 had to be scaled by $\varepsilon^2 = 0.2589254$ for a 1 dB dip in the response and added to 1, and then compared with those of the product of the transfer function derived from circuit analysis and its complex conjugate. The latter calculation produced a constant of 4, rather than 1, because these normalized filters are terminated at both ends by 1 Ω resistors and the a_0 term is 2 in the circuit analysis polynomial. Therefore, the coefficients of $P_n^2(\omega)$ had to be multiplied by 4 after scaling by ε^2 to allow direct comparison, which provided the following results:

$$a_7^2 = 781.07,\quad a_6^2 - 2a_5a_7 = -2900.97,$$
$$a_5^2 - 2a_4a_6 + a_3a_7 = 4410.00,$$

$$a_4^2 + 2a_2a_6 - 2a_3a_5 - 2a_1a_7 = -3508.08,$$
$$a_3^2 + 2a_1a_5 - 2a_2a_4 - 2a_0a_6 = 1538.42$$

$$a_2^2 + 2a_0a_4 - 2a_1a_4 = -352.32,$$
$$a_1^2 - 2a_0a_2 = 32.90 \quad \text{(Eq 4)}$$

The terms a_7 through a_0 are the coefficients of the voltage transfer function. Solving the above equations gave numerical values for these coefficients,

$$27.95(j\omega)^7 + 39.62(j\omega)^6 + 79.97(j\omega)^5 + 69.79(j\omega)^4 + 63.30(j\omega)^3 + 30.84(j\omega)^2 + 12.50(j\omega) + 2 \quad \text{(Eq 5)}$$

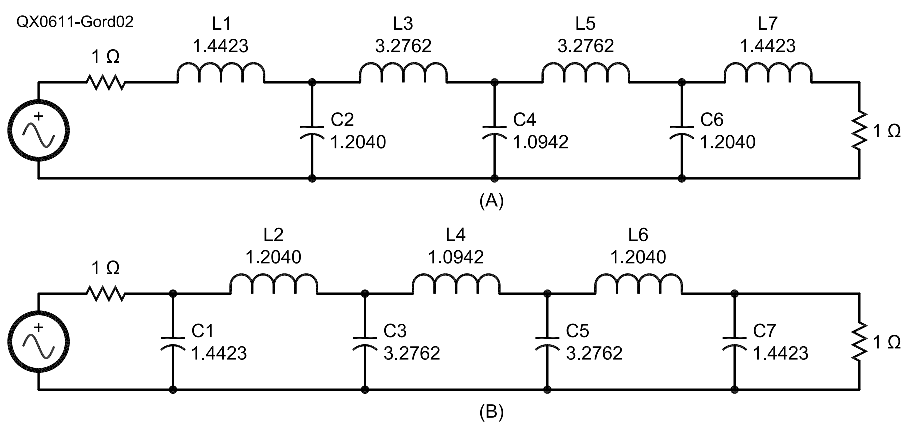


Figure 2 — Normalized (3-dB-down) component values for a 1 dB/0.01 dB unequal-ripple low-pass filter in (A) T-configuration and (B) pi-configuration circuits.

Table 1
Normalized (3-dB-Down) Coefficients for 7th-Order Unequal-Ripple (Acromorphic) Low-Pass Filters

LVR	*L1/C1*	*L2/C2*	*L3/C3*	*L4/C4*	*L5/C5*	*L6/C6*	*L7/C7*	*FL*	*FH*	*FCR*
0.10 dB	1.0191	1.4819	2.2619	1.6415	2.2619	1.4819	1.0191	0.4597	0.8677	1.1525
0.12 dB	1.0355	1.4678	2.2982	1.6128	2.2982	1.4678	1.0355	0.4724	0.8683	1.1517
0.14 dB	1.0516	1.4568	2.3340	1.5863	2.3340	1.4568	1.0516	0.4833	0.8693	1.1504
0.16 dB	1.0663	1.4459	2.3666	1.5627	2.3666	1.4459	1.0663	0.4919	0.8701	1.1493
0.18 dB	1.0802	1.4358	2.3974	1.5407	2.3974	1.4358	1.0802	0.4992	0.8712	1.1478
0.20 dB	1.0936	1.4264	2.4272	1.5198	2.4272	1.4264	1.0936	0.5056	0.8723	1.1464
0.22 dB	1.1068	1.4173	2.4566	1.4999	2.4566	1.4173	1.1068	0.5112	0.8735	1.1448
0.24 dB	1.1193	1.4087	2.4843	1.4819	2.4843	1.4087	1.1193	0.5158	0.8743	1.1438
0.26 dB	1.1313	1.4006	2.5109	1.4646	2.5109	1.4006	1.1313	0.5199	0.8756	1.1421
0.28 dB	1.1428	1.3931	2.5364	1.4485	2.5364	1.3931	1.1428	0.5236	0.8766	1.1408
0.30 dB	1.1545	1.3854	2.5624	1.4325	2.5624	1.3854	1.1545	0.5269	0.8777	1.1393
0.32 dB	1.1657	1.3784	2.5873	1.4176	2.5873	1.3784	1.1657	0.5297	0.8787	1.1380
0.34 dB	1.1762	1.3713	2.6107	1.4035	2.6107	1.3713	1.1762	0.5324	0.8798	1.1366
0.36 dB	1.1872	1.3644	2.6350	1.3893	2.6350	1.3644	1.1872	0.5347	0.8809	1.1352
0.38 dB	1.1980	1.3576	2.6590	1.3757	2.6590	1.3576	1.1980	0.5359	0.8819	1.1339
0.40 dB	1.2090	1.3511	2.6833	1.3623	2.6833	1.3511	1.2090	0.5389	0.8830	1.1325
0.42 dB	1.2185	1.3433	2.7045	1.3475	2.7045	1.3433	1.2185	0.5406	0.8840	1.1312
0.44 dB	1.2289	1.3386	2.7276	1.3380	2.7276	1.3386	1.2289	0.5423	0.8849	1.1301
0.46 dB	1.2391	1.3328	2.7502	1.3261	2.7502	1.3328	1.2391	0.5438	0.8861	1.1285
0.48 dB	1.2484	1.3269	2.7708	1.3154	2.7708	1.3269	1.2484	0.5450	0.8869	1.1275
0.50 dB	1.2577	1.3221	2.7914	1.3046	2.7914	1.3221	1.2577	0.5462	0.8880	1.1261
0.52 dB	1.2682	1.3152	2.8148	1.2930	2.8148	1.3152	1.2682	0.5472	0.8889	1.1250
0.54 dB	1.2772	1.3106	2.8348	1.2828	2.8348	1.3106	1.2772	0.5486	0.8902	1.1233
0.56 dB	1.2869	1.3045	2.8563	1.2725	2.8563	1.3045	1.2869	0.5492	0.8910	1.1223
0.58 dB	1.2964	1.2990	2.8774	1.2623	2.8774	1.2990	1.2964	0.5501	0.8919	1.1212
0.60 dB	1.3053	1.2939	2.8972	1.2529	2.8972	1.2939	1.3053	0.5508	0.8928	1.1201
0.62 dB	1.3145	1.2888	2.9175	1.2436	2.9175	1.2888	1.3145	0.5514	0.8937	1.1189
0.64 dB	1.3236	1.2836	2.9378	1.2342	2.9378	1.2836	1.3236	0.5521	0.8946	1.1178
0.66 dB	1.3324	1.2787	2.9572	1.2254	2.9572	1.2787	1.3324	0.5526	0.8954	1.1168
0.68 dB	1.3411	1.2737	2.9766	1.2168	2.9766	1.2737	1.3411	0.5531	0.8963	1.1157
0.70 dB	1.3500	1.2689	2.9964	1.2081	2.9964	1.2689	1.3500	0.5536	0.8971	1.1147

The coefficients of the various terms in Equation 5 were already known in terms of the products and the sums of products of the inductor and capacitor designations from circuit analysis. This gave 7 simultaneous equations containing the 7 unknown values of the filter components. From these the normalized component values could be established in terms of the ripple bandwidth. It turned out that the filter was actually symmetrical, and there were really only 4 unknown values, but this was not known until the initial design was completed. Modeling the circuit showed that the calculated values were slightly out compared with the exact values required for the desired response, but they were close enough for the final adjustment of the values to be done easily. The corrected, normalized values for an unequal-ripple low-pass filter, with a 1-dB dip in the lower part and an equal-ripple 0.01-dB passband in the upper part of the response, are given in Figure 2 parts (A) and (B) for the T-configuration and pi-configuration circuits, respectively. These values are in terms of the 3dB-down bandwidth, and not the ripple bandwidth used in the original calculations. The passband response curve predicted for this design of low-pass filter is shown in Figure 1. Low-pass filters based on this design have given good service on the HF bands for some years now. They can be tailored to fit in with available capacitor values, and also offer improved harmonic suppression over low-ripple Chebyshev designs.

Designing with Normalized Tables

The use of normalized tables of component values for designing low-pass filters greatly simplifies the procedure for determining the values required for particular cutoff frequencies and termination impedances. Table 1 presents the normalized coefficients for 7th-order acromorphic low-pass filters with LVR values of 0.1 to 0.7 dB. The normalized values, which assume equal terminations of 1 Ω at the input and output and a cutoff frequency of 0.159 Hz ($\omega = 1$), are shown as L1, C2, and so on, to distinguish them from the de-normalized values, which are designated L_1, C_2, etc. The ratio of L3(C3) to L1(C1) has been fixed at 2.22:1 in this table because it provides both low upper-region ripple (variable, but below 0.01 dB) and, for the pi-configuration, the convenience of using either two parallel capacitors, from adjacent positions in the standard preferred-value range, or, in some circumstances, a single standard capacitor with close to the right ratio to C_1 and C_7 for positions C_3 and C_5 in the filter.

The effect of slight variations in this ratio on the upper part of the passband can be corrected by minor alterations to the inductor values. This can be done quite conveniently while checking the VSWR at the input of the filter when the output is terminated by the correct load, or by simulating the circuit and making alterations to suit available capacitor values. The input VSWR measurement is a very sensitive way of checking that the filter is performing as required, despite component variations. The input VSWR should be 1.1:1, or less, between FL, the lowest useable frequency and FH, the highest useable frequency in the passband. Outside this range the VSWR increases considerably. FL and FH are specified in terms of the 3-dB-down cutoff frequency, FC, which in Table 1 is normalized to 1.000. The ratio of FC to FH is given in the FCR column because generally FC has to be calculated from a known FH.

The de-normalization process provides the component values for any cutoff frequency and termination impedance. To work out the actual value of L_1 for a particular cutoff frequency, F_c, and termination resistance, R_o,

$$L_1 = (\mathbf{L1} \times R_o)/(2 \times \pi \times F_c) \qquad \text{(Eq 6)}$$

For C_2 this is

$$C_2 = \mathbf{C2}/(2 \times \pi \times F_c \times R_o) \qquad \text{(Eq 7)}$$

Note that inductors and capacitors are treated differently with regard to termination de-normalization. Capacitors need to be smaller and inductors larger for a higher termination impedance at any cutoff frequency. Table 1 is ranked in order of the LVR of the filter passband response. The higher the value of LVR, the better the stop-band performance. Designs at the top of the table, where **LVR = 0.1 dB**, are marginally better than 0.01-dB Chebyshev designs, and the stop-band attenuation performance improves as the value of LVR increases. At **LVR = 0.7 dB**, the harmonic suppression is about 9 dB better than a 7th-order 0.01-dB Chebyshev filter. The passband ripple is less than 0.01 dB for all designs presented in this table. Normalized coefficients can be used to predict the component values for either pi- or T-configuration low-pass filters. It's just a matter of which type of component is de-normalized first. If the T-configuration is required, the first coefficient should be de-normalized as an inductor. After that the components just alternate between inductors and capacitors until all the coefficients have been de-normalized. In fact, due to symmetry, only just over half of the coefficients have to be de-normalized.

To design a filter that uses single standard-value capacitors in each, or most, of the capacitor positions, it is useful to de-normalize the first capacitor coefficients in both the **LVR = 0.1 dB** and **LVR = 0.7 dB** rows to discover the range of values that are possible for a given cutoff frequency. Then, the most convenient value can be selected from this range. The tolerance of the components used will cause some variation in the actual cutoff frequency, and it is wise to make this frequency a few percent (equal to the worst component tolerance) higher than the band limit to allow for the worst case where it might be low of the design figure. To calculate the desired cutoff frequency for a given application, multiply the band-edge frequency by a tolerance multiplier (1.04 for 4%), and then by the value given in the **FCR** column for the particular row used. This cutoff frequency can then be used in the de-normalization process to work out the actual component values for the filter.

Notice that there is a T-configuration low-pass design near **LVR = 0.44 dB** that has equal capacitor values. If the design values for C2, C4 and C6 do not coincide with preferred values for a particular cutoff frequency,

Table 2
3-dB-Down k and q Values for Unequal-Ripple (Acromorphic) Low-Pass Filter Design

LVR	q_1	$\pm\Delta q_1$	k_{12}	$\pm\Delta k_{12}$	k_{23}	k_{34}
0.1 dB	1.0370	0.0391	0.8059	0.0181	0.5446	0.5187
0.2 dB	1.1081	0.0396	0.7948	0.0174	0.5369	0.5205
0.3 dB	1.1649	0.0400	0.7923	0.0166	0.5305	0.5218
0.4 dB	1.2090	0.0405	0.7824	0.0159	0.5252	0.5230
0.5 dB	1.2458	0.0409	0.7794	0.0151	0.5208	0.5242
0.6 dB	1.2800	0.0414	0.7768	0.0144	0.5167	0.5252
0.7 dB	1.3117	0.0418	0.7751	0.0136	0.5129	0.5263
0.8 dB	1.3373	0.0423	0.7740	0.0128	0.5097	0.5272
0.9 dB	1.3662	0.0428	0.7727	0.0121	0.5062	0.5282
1.0 dB	1.3991	0.0432	0.7700	0.0113	0.5034	0.5288

three equal values of capacitor can still be used for LVR values less than 0.44 dB if a small value of capacitance is added across C4. For LVR values larger than 0.44 dB, a lower value of capacitor must be used for C4 than either C2 or C6, and additional capacitance added across that to make up the value. It should still be possible to limit the total number of preferred-value capacitors to 4 for most T-configuration low-pass designs without any compromise in the cutoff frequency.

Designing with 3 dB-Down K and Q Values

Dishal[1] and Green[2] have shown that low-pass filters can be designed using k and q values through the relationship.

$$k_{n,n+1}^2 = 1/(L_n C_{n+1}) \text{ or } 1/(C_n L_{n+1}) \quad \text{(Eq 8)}$$

where n is 1 through 6 for a 7th-order filter, and L and C are the normalized 3-dB-down filter component values. The first normalized component in a T-configuration low-pass filter is $L1 = q_1$ and subsequent components are

$C2 = 1/(k_{12}^2 \times L1)$, $L3 = 1/(k_{23}^2 \times C2)$, and $C4 = 1/(k_{34}^2 \times L3)$.

The symmetry of the network about C4 for a 7th-order T-configuration filter means that L1 = L7, L3 = L5, and C2 = C6. For pi-configuration filters the network is symmetrical about L4, and $C1 = q_1$.

Specifying filter designs in these terms can provide a starting point for a wide variety of designs without the need for an enormous number of tables of normalized values. Using Equation 8 the normalized values of 7th-order acromorphic low-pass filter components can be calculated using the figures given in Table 2. These 3-dB-down k and q figures are for an upper passband ripple of 0.003 dB at each specified value of LVR. This gives a median value for every parameter at each value of LVR that is roughly halfway between the limits (**$\pm\Delta q_1$** and **$\pm\Delta k_{12}$** for **q_1** and **k_{12}**) that correspond to the passband ripples of 0.001 dB to 0.01 dB. The values of **k_{23}** and **k_{34}** vary very little for different passband ripples at a fixed value of LVR. The values given for them in Table 2 can be used to predict the starting values for L3 (C3) and C4 (L4) for any upper passband ripple, and minor adjustments to the component values made later during circuit simulation to perfect the design. Estimates of the normalized component values of L1(C1) and C2(L2) can be obtained by interpolation between the limits of **$k_{12} \pm\Delta k_{12}$** and **$q_1 \pm\Delta q_1$** given in Table 2 for other ripple values between 0.001 and 0.01 dB, and then the design can be perfected using a circuit simulation package to determine the exact component values after de-normalization. Note that the changes in **q_1** and **k_{12}** must be in the opposite sense, so to raise the upper passband ripple the value of **q_1** must increase and **k_{12}** must decrease by a lesser amount so that the overall kq product increases very slightly. To help prospective experimenters perfect their designs during circuit simulation the following section on design constraints and component relationships has been included.

Design Constraints

The values of **L1**(**C1**) and **C2**(**L2**) determine the LVR, and the ratio of **L3**(**C3**) to **L1**(**C1**) determines the ripple produced in the upper portion of the passband once **L1** and **C2** have been set. This ratio changes slightly with the depth of the LVR for a particular upper passband ripple. To keep the ripple in the upper part of the response below 0.01 dB, the ratio of **L3** to **L1** must be greater than 2.13:1 for values of LVR greater than 0.04 dB. This ratio increases gradually from 2.13 at LVR=0.04 dB to 2.27 at LVR=1 dB for an upper passband ripple of 0.01 dB, and from 2.26 to 2.41 for 0.001 dB passband ripple. The relative values of **C2** and **C4** have to change to maintain equal ripple in the upper part of the passband as they both vary relative to **L1** and **L3**. **C4**, or **L4** in the pi-configuration, can be altered to modify the shape of the upper passband with very little effect on the value of LVR. To increase the depth of LVR the value of **L1** must increase, and that of **C2** must decrease by slightly less, so that the kq product increases overall. The ratio of **L3** to **L1** must increase very slightly as LVR increases to maintain the same upper passband ripple. In addition, the value of **C4** required to maintain equal-ripple in the upper passband must decrease as LVR increases.

Conclusions

Unequal-ripple low-pass filters that have a deep valley in the lower part and an equal-ripple passband in the upper part of the response can be achieved using symmetrical LC ladder networks. The harmonic attenuation of these designs improves as the lower valley ripple is allowed to increase. The sensitivity of these designs to components variations is greater than that of low-ripple Chebyshev designs, but the change in shape of the passband due to this effect can very often lead to a perfectly acceptable and usable alternative passband response.

Sometimes compromises in the cutoff frequency or ripple level have to be made to accommodate preferred-value components in equal-ripple Chebyshev designs, and the harmonic attenuation suffers. Shifting the cutoff frequency of a 7th-order 0.01 dB Chebyshev low-pass filter up in frequency by 10% to get the design to coincide with the next nearest preferred value of capacitor can easily lose 6 dB of harmonic suppression. Changing from a 0.01 dB Chebyshev to a 0.001 dB Chebyshev equal-ripple design can also lose 6 dB of harmonic attenuation. The harmonic suppression that Acromorphic low-pass filters offer can be as much as 15 dB better than 0.001 dB ripple and 9 dB better than 0.01 dB ripple Chebyshev designs of the same order, and they require no compromise in the choice of cutoff frequency to make use of the preferred values of capacitance that are available.

Many solutions exist for low-ripple upper regions with a higher ripple in the lower valley, but solutions where equal ripple is produced in the upper part of the response tend to offer the widest usable bandwidth in this region. One great advantage of the numerous solutions that exist for this type of filter is the range of designs that can be produced, all with the same cutoff frequency but considerably different component values. This allows the selection of the most convenient components for a particular cutoff frequency, rather than being tied to using parallel combinations of several capacitors to achieve the right value, as is often required for low-ripple Chebyshev designs. There is tremendous scope for experimentation in the design of unequal-ripple low-pass filters, and excellent performance can be obtained with the convenience of standard component values if the information presented in this article is applied with understanding and a bit of ingenuity.

Notes

[1]M. Dishal, "Two New Equations for the Design of Filters," *Electrical Communications*, Vol 30, pp 324-337; Dec 1953.

[2]E. Green, "Exact Amplitude/Frequency Characteristics of Ladder Networks," *Marconi Review*, Vol 16, Number 108, pp 25-68; 1953.

First licensed in 1965, Dave concentrated mainly on home construction and 160 meter DXing during his first few years on the air. Constructing his own equipment was a necessity in those early days because he was an impoverished schoolboy, but later it became a great source of fun and satisfaction. He holds a PhD in materials science from the University of Bath, and was a tenured member of the academic staff at Warwick University for many years, specializing in the characterization and study of defects in crystalline solids. During this time he was also a Visiting Professor at the State University of New York (Stony Brook) and a regular Guest Scientist at Brookhaven National Laboratory on Long Island. While at BNL, in the late '80s, he discovered a ham shack that was full of old tube radio gear from the '50s and '60s, and this revived his flagging interest in Amateur Radio. He's now semi-retired and a confirmed vintage radio enthusiast, spending much of his hobby time restoring and operating old tube equipment, as well as writing articles on that and his Amateur Radio experimental work from the past.

A High-Efficiency Filament Regulator For Power Tubes

A simple design that provides voltage step-down, true power regulation and ramp-up start with minimal hardware.

Eric von Valtier, K8LV

Advantages of Filament Regulation

Historically, filament regulation has not been widely used in radio transmitters, and the few examples that I have seen were in very high-end commercial products. The necessary hardware for a filament regulator was quite extravagant prior to the modern solid-state era, and it is understandable that transmitter designers regarded it as overkill. But that has changed and as this article clearly demonstrates, it is now possible to do a first-rate job of filament regulation with very little hardware. It offers three principal advantages and enhancements over traditional designs:

- Precise control of filament emission
- Accurate ramp-up start of filament heating
- Ability to modify filament transformer voltage

6793 Big Trail
Holly, MI 48442
evonvaltie@aol.com

The first two items are obviously significant but unfortunately, I have seen very little detailed information on the subject. Tube manufacturers typically specify a filament voltage to be maintained within ±5% and a ramped start-up, but no actual specification for the ramp-up parameters. Based upon my experience in observing the current inrushes that accompany virtually all filaments, I believe that the best general rule is to profile the start-up current such that ramp-up occurs at nearly a constant current, which is equal to the normal running current. The ideal ramp-up would then be a current source set at the target filament current. In practice, I have found that this scenario is well-approximated by a simple voltage ramp with a duration of 5-10 seconds and in the final controller design, we use a linear ramp-up of RMS voltage from 0 to the final value. This value is then maintained within ±5% (worst case) of the calibrated value.

The third item is an interesting problem that has presented itself to many amplifier builders when faced with filament transformers of higher than necessary voltage. Typical solutions use variacs and power resistors for voltage reduction. The present design is a step-down-only regulator that will regulate the RMS output voltage value, with a lower limit that is within about 250 mV (minimum) of the transformer voltage. For instance, it will allow proper operation of a 5 V filament from any transformer having a minimum loaded voltage output of about 5.25 V. The upper limit of voltage that can be safely stepped-down is determined by several factors as shown below. Generally speaking it works best with transformers having voltages of up to about 60% over the desired voltage, and possibly higher if the current demand is not very great (only a few amperes).

The original regulator design was implemented for an amplifier using a Russian triode with a 12-V/4-A filament, which contained a 24 V dc control circuit and relay supply using a transformer supplying 18 V

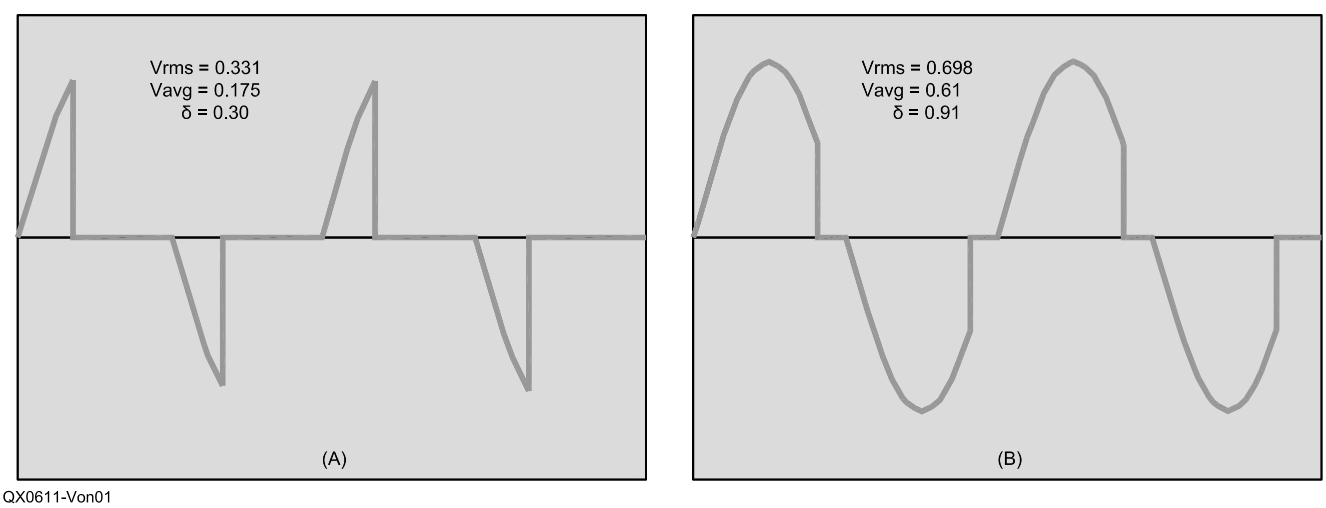


Figure 1 — Filament current versus time for two values of duty cycle.

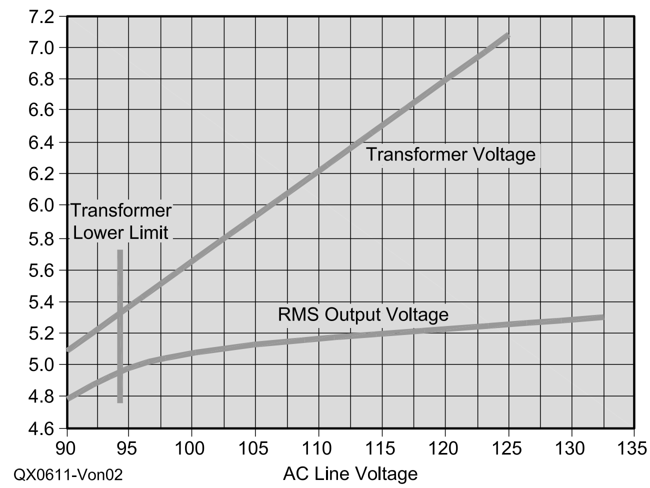


Figure 2 — Measured output of a regulator for an 8877 utilizing a 6.3 V filament transformer.

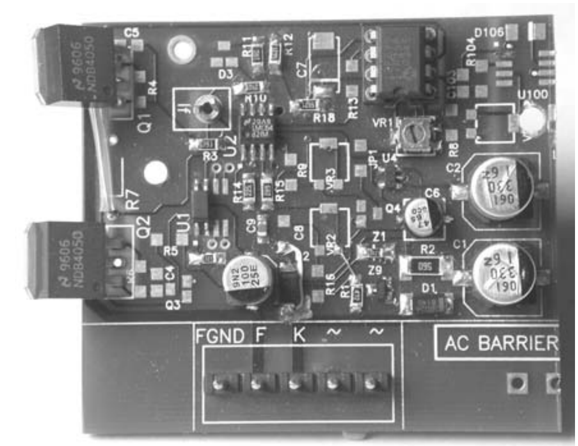


Figure 3 — PCB for 10 A regulator using onboard 2 mΩ MOSFETs, using schematic of Figure 4.

ac that had an extra 4 A for filament supply. The regulator produces a 12.6 V_{RMS} filament supply with excellent regulation and controlled ramp-up. It has performed perfectly for several years to date, and it coexists peacefully within the RF deck (inside of a shielded area for the control and switching circuitry). The following sections describe the theory and circuit design of the regulator.

Regulator Topology and Circuitry

The topology for this regulator is based upon standard synchronous PWM (pulse width modulation) of the 60-Hz filament current, which is readily achieved with modern MOSFETs. The tube filament is uniquely suited for this type of control because of its extremely long thermal time constant. As a result, the PWM rate can be very low and the rise times for turn-on and turn-off can be slowed down to 50 µs or more. This completely eliminates any noise or RFI problems that accompany traditional switching regulators operating at high-kHz frequencies. By combining low PWM speed and low MOSFET on-resistance, a highly efficient regulator results. The circuit contains several unique features, which will next be described.

Figure 1 shows the filament current waveforms for two values of duty cycle, one low and one high. The PWM is applied identically to both positive and negative half-cycles, with turn-on always applied at zero phase. The lower δ would typically occur when the line voltage is quite high, and the higher value when line voltage drops. The regulator is initially "calibrated" to produce the desired RMS output with δ = 60% to allow a good margin in either direction for line voltage variation. It responds to changes in line voltage by varying δ, which varies both the average and RMS output voltages.

Table 1

Full-Cycle Parameters for a 1 V pk Sine Wave with Synchronous PWM at Duty Cycle δ

δ	*Avg*	*RMS*	*RMS/Avg*	*η*
0.10	0.016	0.056	3.65	3.28
0.20	0.073	0.178	2.44	2.20
0.30	0.131	0.272	2.08	1.87
0.40	0.220	0.391	1.78	1.60
0.50	0.318	0.500	1.57	1.41
0.60	0.417	0.588	1.41	1.27
0.70	0.521	0.661	1.27	1.14
0.80	0.587	0.694	1.18	1.06
0.90	0.624	0.703	1.16	1.05
1.00	0.636	0.707	1.11	1.00

η is the correction factor for average-reading voltmeters and is defined as η = RMS/(AVG*1.11).

Regulation is controlled by measuring the output voltage and comparing it to an internal setpoint. The error between the two is used as a control signal to ramp up/down the value of δ until the error nulls. Since the output waveform is a moderately complex waveform, its accurate measurement is not trivial. In this design, the measured output quantity is the RMS voltage squared, which is equal to the actual power times the load resistance. In effect, we are regulating more than just the voltage: We are regulating the power, which is a desirable result. This is a somewhat more advanced form of regulation than normal, and is fortuitous here because the simple relationships among the peak, average, and RMS values of a sine wave are not valid when that sine wave undergoes synchronous PWM.

That is illustrated in Table 1. Note that a regulator based upon measurement and feedback of the average output voltage would not maintain a constant RMS value as the duty cycle changes in response to input voltage variations. This difficulty is entirely circumvented by measuring the RMS voltage and using it to complete the control feedback loop. It provides the added advantage of regulating the output power under variations of both input voltage and waveform.

RMS measurements are normally much more complex than peak or average values; but with high-speed analog-to-digital (A/D) conversion and the time-measurement capabilities found in modern microprocessors, these measurements become quite practical. In our earlier designs the controller constantly used the instantaneous value of δ, which it always "knows," to apply the correction factor η to the measured *average* output before applying it to the control algorithm.[1] But even this method was a compromise because it only produced accurate regulation for an input that was free of harmonics, and the typical filament transformer secondary voltage contains significant distortion.

A practical solution is to measure the true RMS voltage, which becomes feasible with intelligent controllers that have high-speed A/D capability. These considerations all strongly support the concept of RMS regulation, which is the most unique aspect of this regulator. Table 1, which will be required later for calibration purposes, contains the precisely calculated average and RMS values of a 1-V-peak sine wave, subject to synchronous PWM in increments of 10%. The last column of data contains values of the parameter η, the ratio of RMS to average value, which is required in the software design for the controller and for calibration of the final hardware.

The circuit is just a series pass transistor in series with the tube filament and the filament transformer. The control strategy is all implemented in the microprocessor, which derives its fixed power supply directly from the filament transformer. The output circuit is quite unique and requires some explanation. It relies on the increasing availability of MOSFETs with just a few milliohms of on-resistance. These two transistors are in series with the load, so the full load current produces an I^2R loss in them. By use of the best possible devices it is possible to build regulators for all but the largest power tubes that require no external heatsinking of the pass transistors. This allows assembling the entire regulator on a small PCB with no special thermal provisions.

The MOSFET pair is designed to emulate the typical triac, which would typically be used here. However, the forward voltage drop of such devices (1-2 V) is far too great to operate as a high-efficiency pass element in low-voltage circuits carrying high currents. The resulting circuit is a novel bipolar switch using "bootstrapped" drive. Both devices are always conducting simultaneously, and on alternate half-cycles one transistor or the other is operating with a small reverse V_{DSS} caused by the drop across its low on-resistance, because of the load current. The pass transistors Q1 and Q2 can be selected as needed for the desired filament current, based primarily on their on-resistance. The total power loss to be dissipated is calculated from the on-resistance and the RMS filament current as: $P = 2I^2R_{DSS}$.

If this value is kept well below one watt, then PCB-mounted MOSFETs can be used without heatsinks, which is desirable for construction purposes. For a 10 A filament, this would require MOSFETS with a maximum on-resistance of 2-3 mΩ, which are now readily available. As the current increases into the 20 A range and above, it becomes necessary to either parallel several transistors (for no heatsinks) or mount the transistors on the chassis or some type of heatsink. I prefer to avoid the heatsinks when possible, paying for it, if necessary, with additional silicon. Another option is use heatsinks on the PCB-mounted transistors, which is useful up to about 5 W dissipation for normal ambient temperature.

Diode D1 and associated components de-

[1]Notes appear on page 39.

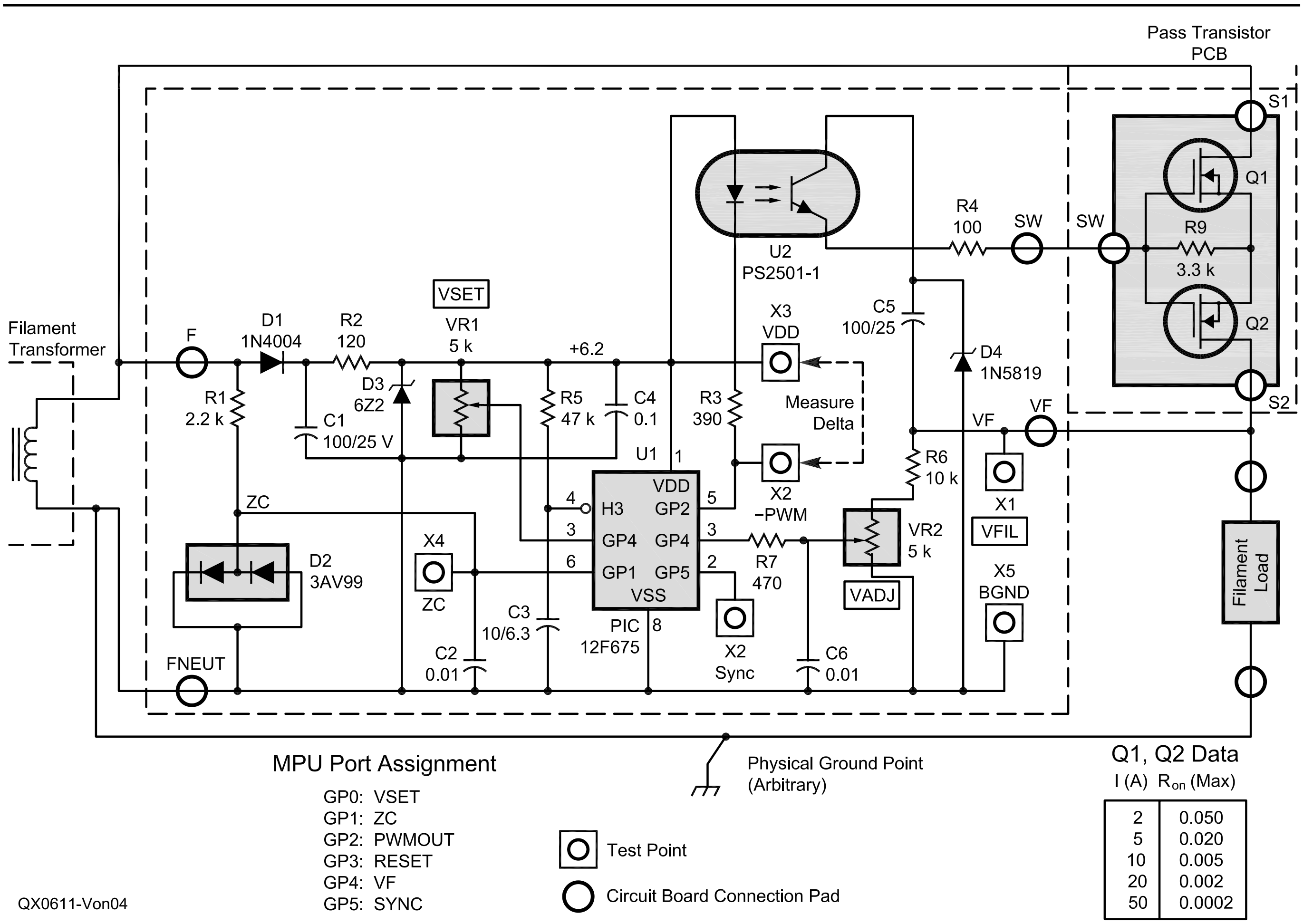


Figure 4 — Schematic diagram of the final regulator design as described in the text.

rive a zero-crossing trigger signal directly from the filament transformer secondary, which is sampled by the MPU at input GP1. This signal is used to establish an internal time base synchronous with the ac mains, and to actually measure the period of the ac cycle. The PWM generator uses the zero-crossing as a time reference and is identical on both half-cycles. In each half-cycle, the current is turned on at zero degrees and then turned off at a time that corresponds to the value of δ, which is continuously calculated and updated by the control algorithm. Hence, control updates occur at the ac line rate of 60 Hz, which is more than adequate here because of the long time-constants inherent in the load.

For best performance, the value of δ should be in the range of 60% at nominal input voltage, which can be seen by inspection of Table 1. This will allow the regulator to accommodate excursions of the input voltage of ±25% and still produce correct RMS output. If even wider input range is required, then δ should be reduced by an additional 10-15%. In practice, the nominal value of δ is established by a constant in the regulator control algorithm and is a function of the output set point divided by the internal voltage reference value. This internal scale factor is input to the software by means of VR1, in the form of a fixed analog voltage in the range of 0-6.2 V. This voltage is measured by the ADC and converted to a number used by the control algorithm, as will be shown in the Setup and Calibration section. Additional onboard test points provide for measurement of the active duty cycle δ with a dc voltmeter.

Software Design

The software currently being used in regulators at K8LV is based upon a few algorithms common to regulator design. I urge those who are adequately experienced to do their own software design, as this is the type of project that is good for developing experience with small, embedded systems. For others not so inclined or equipped, I will supply programmed MPU chips at a very reasonable price along with additional design notes and answers to common questions. Send a request via the e-mail address shown. The general guidelines for algorithm coding are as follows.

The MPU is operated in the internal RC oscillator mode, which is moderately frequency-stable, but not enough so for precise ac phase control. Hence, the time between successive zero-crossings is constantly measured and updated, which measures and defines the period of the PWM half-cycle in units of 64 microseconds. The A/D converter reads 128 samples of the filament voltage during this interval and accumulates the square of these values to form a running average of V-squared times 128, which represents the mean-squared voltage and power. The control loop nulls this value against the SET value, by varying δ, which essentially regulates the power. The actual regulation takes place by varying the turn-off time, with a resolution of 1 bit out of 128 (for the full cycle). This, along with the finite number of samples per cycle, limits the overall accuracy to a little better than 2% of the power output (V^2).

The control loop is somewhat difficult to frequency compensate because of the great degree of numerical approximation used in the control loop, and we chose to leave it slightly underdamped. This produces very mild overshoot to large steps in input voltage or load resistance; but considering the extremely long thermal time constants of the load, this is totally harmless and the regulator achieves remarkable performance for such a small amount of hardware.

Setup and Calibration

The regulator design permits a wide range of input and output voltages, with some limitations. The output voltage set point can range from a few hundred millivolts below the loaded transformer rating, down to zero. The peak input voltage range is fixed by the gate voltage ratings of Q1 and Q2, which are in the 20-30 V range. Actual input voltage will usually be limited by consideration of transformer losses.

The initial calibration process ensures that the desired output voltage will be produced with a corresponding value of δ that allows for proper regulation. This value is established by the settings of both VR1 and VR2, which interact considerably. Hence, the initial setup requires simultaneous adjustment of both values, after which the SET control VR2 acts as the fine-adjustment control for V_{out}.

It is preferable to first test a new regulator without connecting to an expensive power tube, but rather with a 50-100 Ω resistor in place of the tube filament. An RMS voltmeter should be used if available, but an ordinary ac voltmeter will suffice if necessary. Note that the ordinary meter actually measures the average value of the unknown voltage, and then displays it on a meter that has the factor 1.11 built into the meter scale to display RMS volts. This RMS reading is only accurate for a perfect sine wave, so will be used here only to measure the average value of V_{out}, which is simply $V_{meter}/1.11$. This average value is then multiplied by the factor η from Table 1, after measuring the duty cycle to obtain the actual value of δ. All output measurements are to be made in this way if an RMS meter is not available. The duty cycle δ is measured with a dc voltmeter connected from X3 to X1, with a scaling factor: 6.2 V dc (the supply voltage) equals 100%.

The duty cycle is gradually increased or decreased to obtain a δ of about 60% by alternating between VR1 and VR2. This adjusts the circuit to work optimally with the selected transformer and output voltage. Once set, the output voltage can be finely adjusted using VR2 alone (the VR1 setting is not critical). Regulation can be checked using a variac on the transformer primary if available.

Figure 2 shows the measured output of a regulator for an 8877 utilizing a 6.3 V filament transformer. The low-voltage dropout point, which is clearly visible, occurs when the transformer overvoltage drops below a few hundred millivolts and the PWM saturates at 100%. The non-zero slope of the regulation curve is due primarily to numerical inaccuracy in the control algorithm and can be further improved with some improvement in the software. This processor is somewhat taxed to perform all of the necessary code at 4 MHz, and it does an outstanding job of implementing this system for just a few dollars. Enhancements are obviously possible but this minimal design effectively solves all of the problems related to filament control for radio transmitters.

Additional Considerations

I would like to add a few notes about the application of this concept overall. The circuit is similar to, but not quite identical to, typical switching regulators, which generally use inductance to convert power at one voltage level to power at another. These inductors also enhance the overall efficiency of the conversion process, one effect which is missing here. As a result this regulator has slightly degraded transformer efficiency.

This regulation scheme produces copper loss in the filament transformer which is proportional to the percentage of overvoltage of the transformer. When this percentage is zero (transformer voltage equals load voltage) the copper loss is the same as that of the bare transformer, typically about 5%. As the percentage increases, so does the percentage of copper loss, so it is not efficient to use this scheme for a very large percentage of voltage reduction. The RMS current rating of the transformer should be observed since unlike a switching regulator, this circuit will still draw the full RMS load current from the transformer even if it is selected to produce a much higher voltage than the load. This factor only becomes important with higher current filaments, those that require over 100 W. For example, this regulator works well with 6.3 V ac transformers powering 5 V filaments, 12 V ac to 10 V, etc. But any transformer with the proper full-load current rating is permissible if operated within its current rating.

Regulator Construction

The purpose of this article has been to introduce the basic theory and application of a new concept, rather than details of construction. It is hoped that others will be motivated to utilize this information to design and build their own regulators based upon these principles. I have built several working models to verify the design and study its operation with various loads. Most of the data presented here is based upon the final design implemented on a small PCB (4.5 in^2) that can be installed almost anywhere. The finished board, using two TO220 pass transistors for a 10 A filament, is shown in Figure 3. The circuit schematic is in Figure 4. The board was laid out in such a way that the pass transistor section could be removed from the main PCB and chassis mounted for heat sinking. In the worst case Q1 and Q2 will dissipate about 10 W, so direct mounting on the amplifier chassis is a viable option.[2]

The entire circuit can be routed on a few square inches of PCB as shown, even if the power transistors are PCB mounted instead of on an external heat sink. Copies of this circuit board, along with a programmed MPU chip and some additional instructions, will be made available to readers of this article at low cost. Contact me at the e-mail address shown on the first page for details.

Notes

[1]η is the factor by which an averaging voltmeter reading must be multiplied to obtain the correct RMS value when PWM duty cycle is less than 100%.

[2]A forward voltage drop of 1 V across a triac carrying 10 A of current would produce 10 W of loss, about 20% of the power supplied to a 5 V filament.

Eric von Valtier, K8LV, is an engineer and physicist with advanced degrees in electrical engineering, nuclear engineering, and physics from the University of Michigan. He was first licensed in 1951 as WN8IRO, and spent his entire teen years obsessed with ham radio. He went QRT from 1960 to 1998, returning to ham radio as a "retirement" activity. His professional career was spent as an independent design and R&D engineer, during which he produced over 100 product designs that were successfully marketed, with several patents issued. He is equally experienced in both analog and digital electronics and has written over a half-million lines of code covering numerous applications from commercial database access to scientific computation. He enjoys building equipment for his station using new and experimental concepts, and pursuing LF DX using antennas that perform beyond the expectations of a QTH which cannot "be taken very seriously for DX purposes!"

Command and Control: Talk to Your Radio and Your Radio May Talk Back

Tell your radio where to go, and it can, while using PC voice-activated software control.

Steve Gradijan, WB5KIA

Simple software uses voice commands to control a Kenwood TS-2000, TS-570 and perhaps TS-480 or any of eleven ICOM transceiver models. The techniques involved in controlling station equipment using voice control, called *Command and Control* involve the free Microsoft Speech Engine and are not difficult to master. Some versions of the *Windows XP* operating system have the engine pre-installed, and it is available for *Windows 98* and higher systems, too.

Command and Control is the use of specific voice commands through your PC sound card to control your station equipment. Language compilers like Microsoft *Visual Basic*, Borland *Delphi*, *C++* and others are used to homebrew speech recognition software. *Visual Basic* or *Delphi* programming of the required software is easily within the capabilities of even novice programmers by using the examples provided and mentioned here. For those experimenters without a language compiler, the example software available on the *QEX* Web page and shown in Figure 1, can be used with over a dozen different transceivers to change frequency and mode using voice commands.[1] The software demonstrates the principles involved in *Command and Control* even without a radio, should you not own any of the two dozen radios that can be controlled.

What Can You Do with *Command and Control*?

Use voice commands to tune your radio, change mode and settings, or change your antenna heading. The physically handicapped may be able to control a radio by voice commands. Use the technology to control your radio during contests, especially during CW contests or when you could use an extra pair of hands; add voice commands to your homebrew logging software, and so on. Electronic devices, robots or room lights can be controlled too.

How It Works

Speak a command into a microphone connected to a PC sound card microphone jack as shown in Figure 2. The sound card converts the audio to digital form, and that is read by Microsoft's Speech Recognition Engine (part of Microsoft's speech *SAPI 5.1*, discussed below). The Speech Recognition Engine uses a grammar to "filter" the spo-

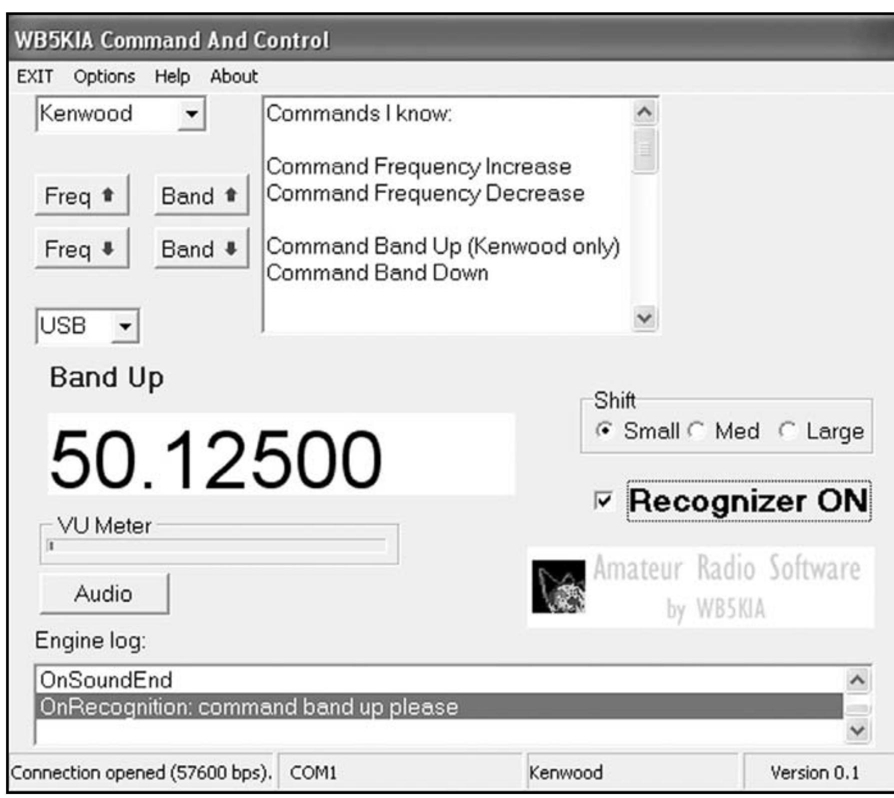


Figure 1 — Screenshot of *Command and Control* controlling a TS-2000. At the bottom of the screen, a log records the speech recognizer's activities.

[1]Notes appear on page 48.

1902 Middle Glen Dr
Carrollton, Texas 75007
sjg47@lycos.com

ken word. The grammar limits the words the software can "recognize." Homebrew software transforms the words or phrases that are identified as valid, by using the grammar. A valid recognition creates a response that sends a digital command to your station equipment through the PC serial COM port. Depending on the equipment you are controlling, the command is sent through an RS-232 cable either directly to the equipment or through a level converter. The level converter transforms the RS-232 signals from the PC to the TTL signal levels that may be required by some equipment manufacturers.

A grammar is a file that contains the words and phrases the program understands. The initial part of the grammar defines constants used to communicate with the software. Rules are chosen to suit the application. Figure 3 shows the top level rule called **start**, which is defined as the optional word command, a value from a different rule also called **command** and the

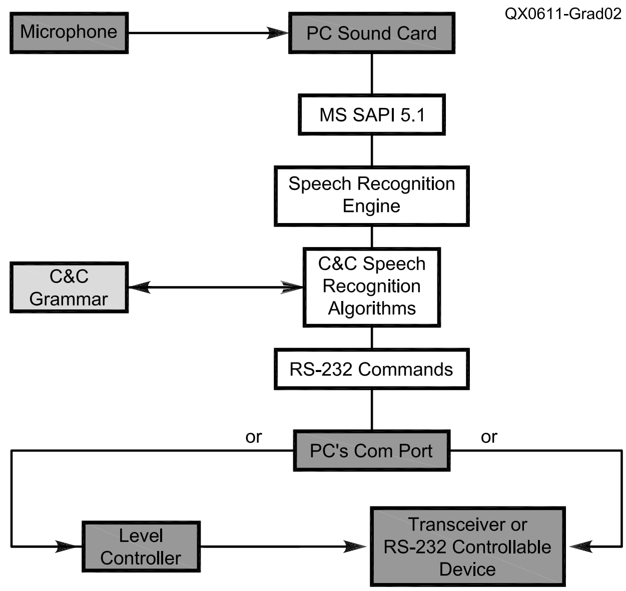


Figure 2 — Voice commands go to the PC, processed by the speech recognizer and are conveyed to the radio over a serial cable.

```
- <GRAMMAR LANGID="409">
    <!--  "Constant" definitions   -->
  - <DEFINE>
      <ID NAME="RID_start" VAL="1" />
      <ID NAME="PID_chosencommand" VAL="2" />
      <ID NAME="PID_commandvalue" VAL="3" />
    </DEFINE>
    <!--  Rule definitions   -->
  - <RULE NAME="start" ID="RID_start" TOPLEVEL="ACTIVE">
      <O>command</O>
      <RULEREF NAME="command" PROPNAME="chosencommand" PROPID="PID_chosencommand" />
      <O>please</O>
    </RULE>
  - <RULE NAME="command">
    - <L PROPNAME="commandvalue" PROPID="PID_commandvalue">
        <P VAL="1">color red</P>
        <P VAL="2">color blue</P>
        <P VAL="3">color green</P>
        <P VAL="4">color yellow</P>
        <P VAL="5">color black</P>
        <P VAL="6">color gray</P>
        <P VAL="7">color navy</P>
        <P VAL="8">mode Lower sideband</P>
        <P VAL="9">mode Upper Sideband</P>
        <P VAL="10">mode CW</P>
        <P VAL="11">mode FM</P>
        <P VAL="12">mode AM</P>
        <P VAL="13">mode RTTY</P>
        <P VAL="14">frequency increase</P>
        <P VAL="15">frequency decrease</P>
        <P VAL="16">Windows shut down</P>
        <P VAL="17">Windows sleep</P>
      </L>
    </RULE>
  </GRAMMAR>
```

Figure 3 — The grammar is an XML file that determines what control words will be recognized by the Microsoft Speech Engine.

optional word **please**. The command rule used in the software described here defines one of two **modes**, two **frequency** operations, six **colors** or two ***Windows*** operations that the example program recognizes. The program can access these with code, too, since they are also defined as properties (**chosenmode**, **chosencolor** and so on) in the software code. The grammar is stored in an XML file and is loaded with the software's **OnCreate** event handler when the program's main form is loaded.

The **OnRecognition** event handler locates the **chosencommand** property and then finds the nested **commandvalue**. The **commandvalue** determines how the program responds to the user's spoken request. Phrases like "command frequency decrease" or "red please" will lower the rig frequency or change the color of the program's form in this example.

The **TfrmCommandAndControl.SpSharedRecoContextRecognition** event handler does all the hard work in the main *Delphi* example. The **case** statements invoke the desired response for each phrase that is defined in the grammar and recognized by the Speech Engine. The action selected by the **case** statement corresponds to the **commandvalue** identified by the Speech Recognizer. The code snippet in Table 1 is annotated and abbreviated to describe the key elements of the *Delphi* coding. The **comport1.PutString** invokes the actual command to the radio using the **AsyncPro** COM port control. You can use other port programming controls, design your own or download this free control for *Delphi* as indicated in the Resources sidebar. Lines starting with double forward slashes (//) in Table 1 and source code are program comments for the *Delphi* examples; lines starting with an apostrophe (') in Table 2 serve the same purpose in the *Visual Basic* example.

What You Need to Operate

Developing your own voice control program for your radio is not very difficult to implement if you have any of several, easy to use language compilers and if the radio or equipment you want to control can be accessed by RS-232 (serial) protocols. To implement *Command and Control* at your station using the techniques described here, you need a PC with a *Windows 98* or higher operating system.

Microsoft *SAPI 5.1*, the speech software package that is part of some *Windows XP* installations, but which can work with earlier operating systems, must be installed on your PC; a microphone must be connected to your sound card's microphone jack; and you need some control software. The voice control software will give you the most problems; software to control equipment with voice commands is not generally available. You

Table 1
The *Delphi* Event Handler that Does the Work.

```
procedure TfrmCommandAndControl.SpSharedRecoContextRecognition(
                   Sender: TObject; StreamNumber: Integer; StreamPosition: OleVariant;
                   RecognitionType: TOleEnum; var Result: OleVariant);
var
SRResult: ISpeechRecoResult;

begin
  SRResult := IDispatch(Result) as ISpeechRecoResult;
  with SRResult.PhraseInfo do
         begin
            Log('OnRecognition: %s', [GetText(0, -1, True)]); //0 is the place it
            //starts recognizing; -1 tells it to recognize all words - a 1 would
            //recognize one word, a 2 would recognize two words etc. The more "words"
            //used, the greater the reliability of invoking the desired response.
            //The grammar commandvalue of the recogized phrase triggers the specific case
            //statements that follow. The commandvalues are assigned when you build
            //the xml grammar.
            if (checkbox2.checked) and (interference1 = false) then
             case GetPropValue(SRResult, ['chosencommand', 'commandvalue']) of
               1: begin //used to change the program form's color
                   Color := clRed;
                   label1.caption:='Red';
                    end;
               8: begin //used to change the rig's mode; to lower sideband in this case
                   Color := clBtnFace;
                   Try
                      Case Combobox1.itemindex of
                             1: begin //for Kenwood mode
                                    comport1.PutString('MD1;'); //LSB mode
                                  end;
                             2..11: begin //for ICOM
                             // ICOM code goes here
                                end;
                             else begin
                               // ignore, do nothing
                                        end;
                              end; //end case
                                except
                                  // do nothing
                               end;
                              //LSB
                              label1.caption:='LSB';
                      end;
         14: begin //used to increase frequency
                   Color := clBtnFace;
                    //increase freq
                   button4.click; //routine to invoke a frequency shift
                   label1.caption:='Increase frequency';
               end;
         16: begin //routine to shut-down Windows (turn your PC off)
                   Color := clBtnFace;
                   //windows shut down
                   button2.click;
                   label1.caption:='Windows shut-down';
               end;
         17: begin // used to put Windows in "sleep" mode
                   Color := clBtnFace;
                   //windows sleep
                   label1.caption:='Windows Sleep';
                   SendMessage(Handle, WM_SYSCOMMAND, SC_SCREENSAVE, 0);
               end;
         else
               begin
                   Color := clBtnFace;
                   end;
               end ;
         interference1 := false;
         end;
end;
```

will have to write your own program or use the simple example software described below which controls the frequency and mode of several radios. It will introduce you to the unlimited possibilities of *Command and Control*.

The example program controls the frequency of several Kenwood models including the TS-2000 and almost a dozen ICOM radios. You need a serial cable to connect your PC to the Kenwood radios. ICOM's CI-V level converter or a homebrew alternative and the serial cable are needed to control most ICOM radios. Some Yaesu and Ten-Tec radios require similar level converters but the software described here will not control them as written. You will use code similar to that required by the ICOM radios to write code appropriate to provide the command signals those radios require. The *SAPI 5.1* is pre-installed on some *Windows XP* systems. *Windows 98* users and higher may need to install the full SDK if the *SAPI 5.1* is not already installed. The example program Help file describes how to determine if *SAPI 5.1* is installed. While either *SAPI 4* or *5* can be used to implement *Command and Control*, the examples provided here work only with *SAPI 5.1*. *SAPI 5* is compliant with *Windows 98* and higher while *SAPI 4* works with *Windows 95* and above.

Depending on your voice characteristics, you may need to train your PC to recognize your voice and to calibrate voice levels. The process involves reading some provided text into the microphone connected to your sound card to compensate for differences in pronunciation and dialect that may be a result of where you were raised. Voice training is automated, very simple to use and the tools are part of the example software. Once calibrated, settings are valid for future *Command and Control* sessions and other programs that provide speech recognition.

What You Need to Design Your Own Software

I used Borland *Delphi* to program the main example for this article (see Figure 4). Microsoft *Visual Basic* works fine and it is possible to do some control work by writing in HTML. Any *Windows* platform capable compiler can be used. The Microsoft *SAPI SDK* contains examples of how to use the technology to communicate with the PC for Microsoft language flavors. The examples in this article show how to link the PC to the radio to actually do something useful.

Figure 2 outlines the basics of *Command and Control* speech recognition. It shows the hardware pieces in dark gray and the software with a white background. Connect an ordinary PC microphone to the sound card mic jack; turn on the software and communication is established by the *SAPI* algorithms that link to the speech recognition engine software. A "grammar" file is used as a "filter" to make it easier for the recognizer engine to determine what is said. Algorithms identify what words the recognition engine determines are spoken and, based on that result, directs the software to send appropriate commands to the PC COM ports. Command is achieved with

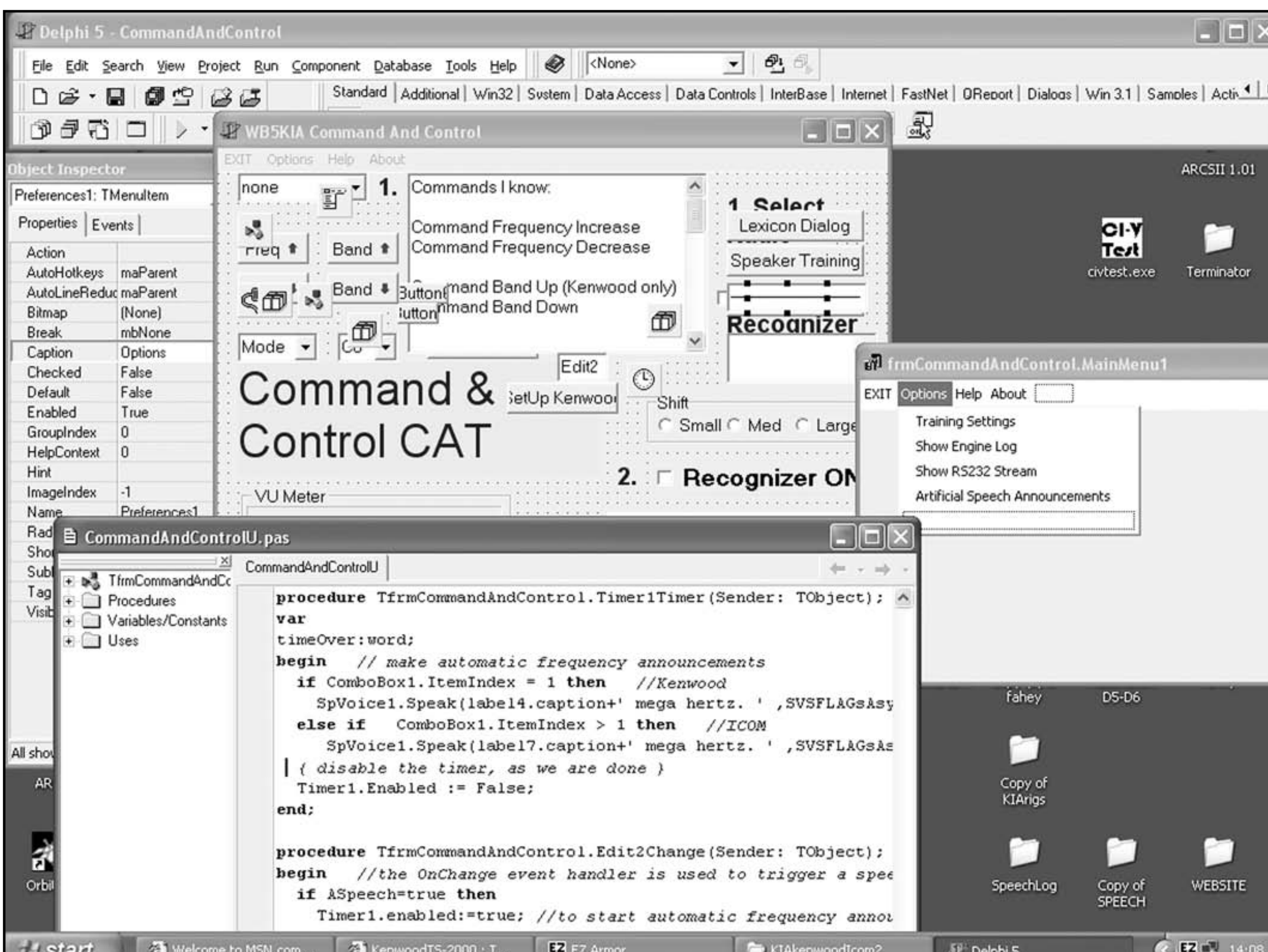


Figure 4 — A view of the *Delphi* 5 Professional language compiler windows.

some radios through a direct RS-232 serial connection (Kenwood and Elecraft are two examples); others require a level control to convert RS-232 levels into the signals that are recognized by ICOM, Yaesu and other radios.

Grammars (see Figure 3) are needed to “train” the speech recognizer and tailor it to the language being spoken. The grammar used in the examples shown is for American English (**<GRAMMAR LANGID =“409”>**); British English is “809.” The grammar provides the speech recognition software a way to decide what words have already been spoken. *SAPI 5.1* uses an XML grammar to help the speech recognizer determine what has been said. XML files “look” like ordinary text or html files, but need to be created and modified with an XML editor. An excellent tool is provided as part of the Microsoft *Speech Development Kit* (*SDK*). A copy of the free *SDK* is also required if you do not have a stand-alone XML editor. The *SDK* is a large file download (68 MB). If you have a XML editor, you can develop without the *SDK* but note the Microsoft *SAPI* references and example code included in the download are invaluable aides if you intend to code something elaborate.

A grammar implemented in XML is required to educate the speech recognizer as to the allowed commands that can be given to the radio or other instrument you are attempting to control. A “shared” recognizer (**TSpsharedRecognizer**) is described in the example code that can be used with multiple applications. A control mode is used to pass commands to your radio. The words that are understood by the *SAPI* with *Command and Control* are limited to the supported commands defined in the grammar. The grammar is established to define rules to dictate what will be understood by the PC. Since the vocabulary is limited, the recognizer’s job is relatively easy compared to when the *SAPI* is used to allow continuous diction (a technique to write spoken words to a word processor, for example).

Rather than trying to understand anything spoken, the software only needs to recognize speech that follows the supplied rules. An example grammar used with *Command and Control* is shown in Figure 3.

A second way of providing voice control involves continuously speaking, as in normal conversation, to transfer a large amount of information. This might involve dictating to have a word processor type a document or speaking individual ordinal numbers to change the frequency setting on a radio. Continuous diction might allow sending of Morse code or teletype while avoiding a key, keyer or keyboard. Continuous diction is a

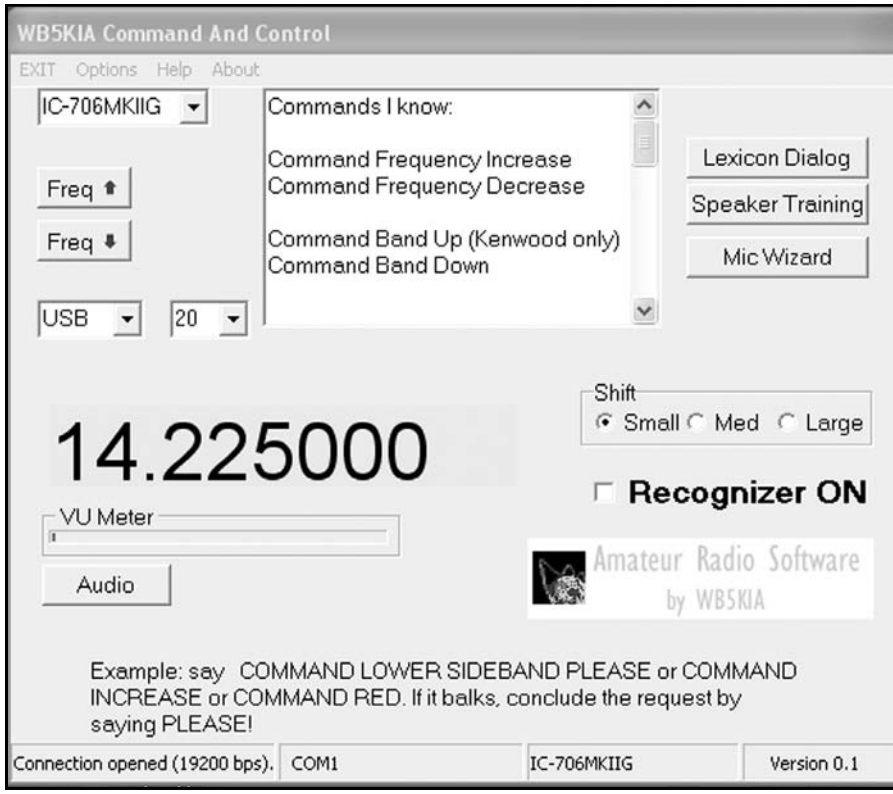


Figure 5 — The control program has facilities to ‘train’ speech to be recognized as shown on the right of this panel demonstrating an ICOM IC-706MKIIG being controlled. The program has manual ‘back up’ controls.

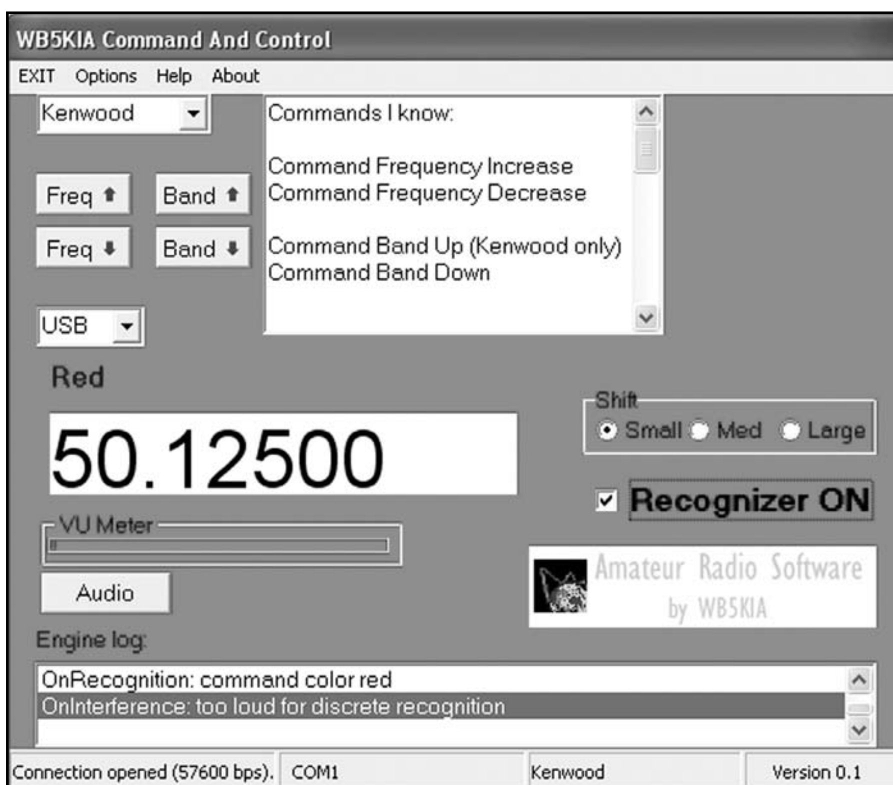


Figure 6 — This screen has a brilliant red background, which can be changed to other hues by voice commands. Various *Windows* operating system functions can also be triggered.

method that is simpler to implement in some ways than *Command and Control*, but it is not as reliable with simple software.

Brian Long (*Speech Synthesis & Speech Recognition Using SAPI 5.1*) documents methods for communicating with Microsoft's artificial speech engine and provides *Delphi* example code.[2] His articles provided the framework for this article. *Visual Basic* code examples and examples in several other programming languages are included in the *SDK* development kit. The example programs discussed here use code snippets from these materials and are available on the *QEX* Web page. See Note 1.

Visual Basic works fine too. A *Visual Basic* code example (Figure 7) demonstrates one way to use spoken commands to control the frequency on a Kenwood radio. It also shows how to select activities within a program for execution using one of the recognition routines. The difficult part in writing this type of software is hooking up the code through a COM port. The *Visual Basic* example uses the Mscomm port control. Depending on your version of *Visual Basic*, you may have to reload the port control into the project file or use a third-party control. Only the VB5/6 source code is provided.

Examples

Two examples demonstrate simple techniques you can use to develop your own speech recognition control software. Download the examples, which contain sample code and a compiled program from the *QEX* Web page. (See Note 1.) Non-programmers can use a compiled version of the *Delphi* code to experience using simple software to "control" a radio using only voice commands. Speak simple phrases like "Command frequency decrease please" or "Command mode upper sideband" to change frequency and mode. It is a bit eerie.

Developed with *Delphi*, *AsyncPro* (a free control for the COM port), a class library from Brian Long's article and Microsoft's *SAPI 5.1*, the main example controls frequency and mode for a Kenwood TS-2000 or any of eleven ICOM models including the IC-706MKIIG. The second example demonstrates *Visual Basic* code for implementing basic frequency control of a TS-2000. The examples consist of a grammar and a piece of software. The software source code is annotated. For those who do not have any of these radios, the main example program provides a few voice controlled activities that can be experienced without a radio.

Brian Long introduced the programming class SpeechLib_TLB in the article *Speech Synthesis & Speech Recognition Using SAPI 5.1* and his code is a necessary tool for *Delphi* programmers. The free class file needs to be added to the *Delphi* Project (see Resources).

Voice recognition software can be programmed in any modern programming language. A second example was coded using *Visual Basic*. The identical grammar is used with both examples. The COM1 port was coded to communicate with your radio and is "hard wired" to keep the code simple. You can modify the port and other program features to suit your preferences if you have the software compiler. The default communication speed for the COM port is set to 57600 for the TS-2000 and 19200 for the ICOM radios. The corresponding communication speed must also be set on the transceivers; consult the appropriate documentation for your radio to find out how to do this.

Voice Commands

A *Command and Control* grammar in the main example allows 22 different voice commands. When developing a grammar, choose words or phrases that are as different from one another as possible to avoid confusing the software voice recognizer. The longer the length of a vocal command, the less likely the chance for a miscued control. Three types of commands are used in the *Delphi* example. Two command groups were established exclusively for radio control, the other groups demonstrate related activities. Table 3 shows the voice commands the software "understands" and their purposes.

The *Delphi* example code results in a simple, usable transceiver voice controller that has several options. It controls the frequency and mode of a TS-2000 and

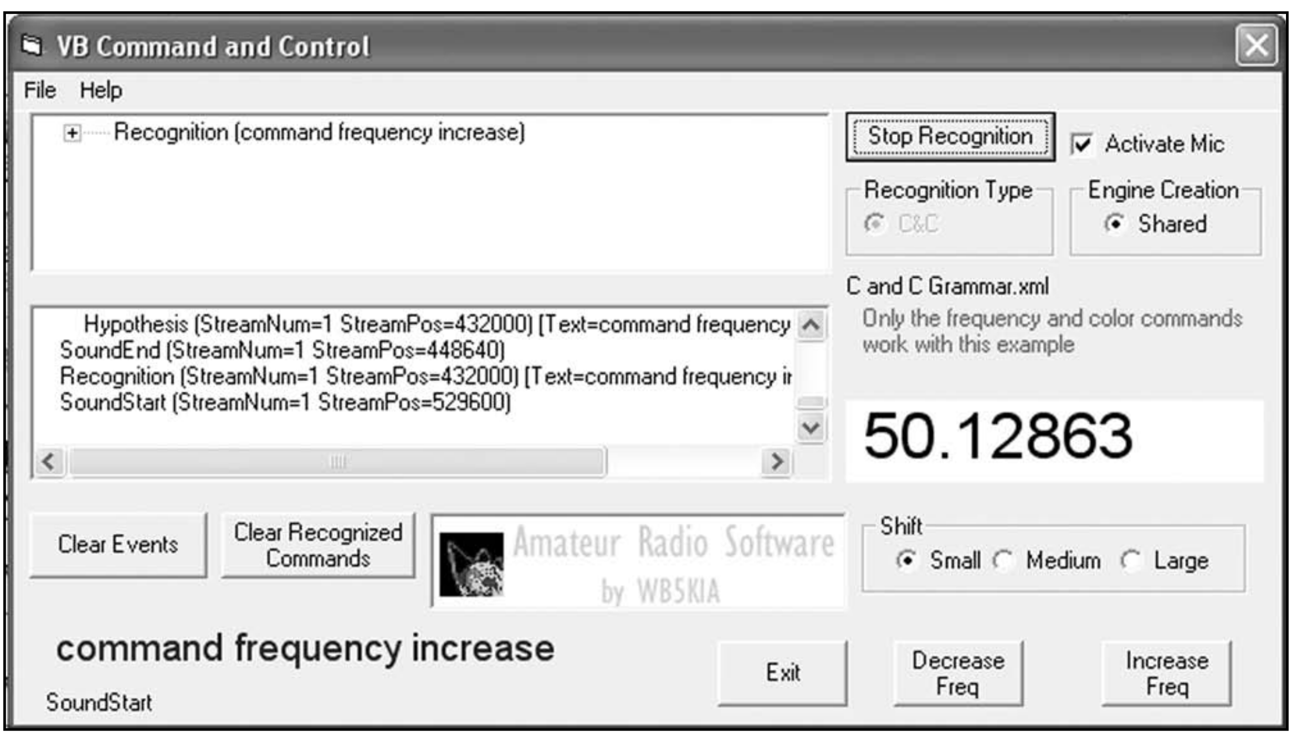


Figure 7 — The VB *Command and Control* example demonstrates how *Visual Basic* can be used to control a Kenwood transceiver.

Table 2
A Visual Basic Example Controlling the Frequency of a TS-2000
The Two Routines Work Together to Change Frequency and Evoke Screen Events

```
' Recognition event handler
Private Sub RC_Recognition(ByVal StreamNumber As Long, ByVal StreamPosition As Variant, ByVal
       RecognitionType As SpeechLib.SpeechRecognitionType, ByVal Result As SpeechLib.ISpeechRecoResult)
  Dim RecoNode As Node
  Static i As Integer
  ' Update Event List window first
  UpdateEventList StreamNumber, StreamPosition, "Recognition", " [Text=" &
       Result.PhraseInfo.GetText() & ", RecoType=" & RecognitionType & "]"
  ' Increment unique value for RecoNode's key name.
  i = i + 1
  ' Add top level node
  Set RecoNode = TreeView1.Nodes.Add(, , "Reco" & i, "Recognition (" & Result.PhraseInfo.GetText() &
       ")")
  ' Call the BuildResultTree subroutine to build up the Result tree
  BuildResultTree Result.PhraseInfo, Result.Alternates(5), RecoNode
  ' Save the recognition Result to the global RecoResult
  Set RecoResult = Result
  ' This is where you can make the recognition results do something useful.
   Label1.Caption = " " & Result.PhraseInfo.GetText() & ""
  Text1.Text = Trim(Label1.Caption) 'the change event handler of Text1 is used to evoke a response
End Sub

Private Sub Text1_Change() 'this is one way to evoke responses
  ' it should also be possible to capture the grammar
  ' P VAL data ... property value i.e. red is 1, blue 2 etc.
  'having that data as in the Delphi example would make coding
  ' recognition of identified commands to desired response easier
  If Text1.Text = Trim("command color red") Then    'to change screen colors
       Form1.Label1.BackColor = RGB(255, 0, 0)
  ElseIf Text1.Text = "command color blue" Then
       Form1.Label1.BackColor = RGB(0, 0, 255)
  ElseIf Text1.Text = "command color black" Then
       Form1.Label1.BackColor = RGB(0, 0, 0)
  ElseIf Text1.Text = "command color green" Then
       Form1.Label1.BackColor = RGB(0, 255, 0)
  ElseIf Text1.Text = "command color yellow" Then
       Form1.Label1.BackColor = RGB(255, 255, 0)
  ElseIf Text1.Text = "command color navy" Then     'actually this is cyan; this is navy: &H00800000&
       Form1.Label1.BackColor = RGB(0, 255, 255)
   Else
       Form1.Label1.BackColor = &H8000000F 'button face color
  End If

  If (Text1.Text = Trim("command frequency increase")) Or (Text1.Text = Trim("command frequency
       increase please")) Then    'to go up frequency

  'routine to increase frequency is involved and goes here (see source code)

  ElseIf (Text1.Text = Trim("command frequency decrease")) Or (Text1.Text = Trim("command frequency
       decrease please ")) Then      'to go down frequency

  ' routine to decrease frequency is involved and goes here

 End If
End Sub
```

Table 3
Voice Commands Used in the Example and the Rig or PC Operations Evoked.

Voice Command Example	*Purpose*
"Command **frequency** increase"	Increase the frequency 1 kHz.
"Command **frequency** decrease"	Decrease the frequency 1 kHz.
"Command **band** up" / "Command **band** down"	Change the radio operating band. (Kenwood radios only, ICOM radios use manual band change.)
"Command **mode** upper sideband please"	Change the mode.
"Command **shift** large"	Change the frequency increment/decrement amount of the frequency command (options are small, medium, large).
The following are non-rig "control" functions:	
"Command **color** red please"	It also "knows" blue, green, red, navy, yellow, and black. The color command group demonstrates voice control of functions in a program analogous to using a logging programs button or scroll bar.
"Command ***Windows*** sleep"	Invoke the *Windows* screen saver. The *Windows* group demonstrates the capability of even controlling the PCs operating system activities.
"Command ***Windows*** shut-down"	Shut down the *Windows* operating system.

Resources

Articles Regarding Command and Control:
Speech Synthesis & Speech Recognition Using SAPI 5.1 by Brian Long
www.blong.com/Conferences/DCon2002/Speech/Speech.htm

Software:
Microsoft SAPI 5.1 SDK. The SpeechSDK51MSM.exe is available from **www.microsoft.com/speech/download/SDK51**.
Some versions of the *Windows XP* operating system come with *SAPI 5.1* installed but most people will need the free 68 MB download. A cd is available from Microsoft for a nominal fee.
Files associated with Brian Long's article regarding *SAPI 5*, including SpeechLib_TLB and similar methods to implement *SAPI 4*: **www.blong.com/Conferences/DCon2002/Speech/SAPI51/SAPI51.zip**; **www.blong.com/Conferences/DCon2002/Speech/SAPI4HighLevel/SAPI4.htm**
AsyncPro is a free *Delphi* add-in control used with a Borland *Delphi* or C++ compiler to provide control of the COM port. It is available from: **sourceforge.net/projects/tpapro/**.

Programming Amateur Radio Software:
B. Wood, WØDZ, "The Return of the Slide Rule Dial," *QST*, Feb 2002, pp 33-35.
M. E. Erbaugh, N8ME, "Customize the Ten-Tec Pegasus — Without Soldering," *QEX*, Sept/Oct 2002, pp 3-9.
S. Gradijan, WB5KIA, "Build a Super Transceiver — Software for Software Controllable Radios," *QEX*, Sept/Oct 2004, pp 30-34.

Language Compilers:
Suitable compilers were described in: S. Gradijan, WB5KIA, "Beginners' Computer Programming for Ham Radio — Part 1," *QST*, Feb 2003, pp 32-37.

possibly other Kenwood radios and, with slightly less elegance, the same for eleven modern ICOM transceivers. The resulting software also controls features of the software, commands *Windows* into sleep mode or shuts *Windows* down. The *Visual Basic* example demonstrates how to control the frequency of a TS-2000.

The difficult part of providing *Command and Control* for Amateur Radio activities is not the creation of the speech software links and grammar, but is the coding to invoke the commands and transmit them to the radio. This is relatively easy with the Kenwood radio but fairly difficult for ICOM, Yaesu and some other radios.

Controlling the Kenwood is neat, but my algorithms for ICOM could be improved. I provide a simplified, semi-manual system for the ICOM radios. The process is graceful for Kenwood using voice commands.

Performance

Performance on my Athlon 2200 PC with 317 MB of memory is remarkable with my TS-2000 or IC-706. Results may not be as good with slower equipment. The example program's ability to recognize commands prior to setting the microphone level and before voice 'training' was fair. Recognition increased dramatically after using the example program's "Mic Wizard" and "Speaker Training" buttons and following the directions.

The difficult part of providing *Command and Control* does not involve the creation of the speech software and grammar. The coding to invoke the commands and transmit them to the radio is relatively easy with the

Kenwood radio but is comparatively difficult to implement for ICOM, Yaesu and some other radios. The example software for Kenwood is efficient, but the algorithms to control the ICOM radios could be improved to provide more flexibility.

Your Radio May Talk Back

I almost forgot to mention, your radio can also talk back to you. The Microsoft *SAPI* has provisions for artificial speech. Turn on the artificial speech option in the program and frequency changes will be announced over your PC speakers. This process was described in the May/June 2006 issue of *QEX*, in the article "A Talking LogBook With Rig Control."[3]

The Future

Since this article was written, *Command and Control* was added to the free *TalkingLog Book* program (**http://mysite.verizon.net/cloudyr/talkinglogbook**) described above. It works with a large variety of radios. A free CAT program for the TS-2000 called *ARCS II* uses voice commands to control 27 functions of the radio (**www.qsl.net/wb5kia**).

Acknowledgments

Significant contributions involving code to connect to the Microsoft Speech Recognizer came from the work of Brian Long. Some code snippets used to decode BCD signals used with ICOM radios was borrowed from software by David Fahey. He wrote several algorithms for CAT control of an ICOM R8500 radio receiver using *Delphi 3* in 1997.

Notes

[1]The software version described in this article, and current at publication time, is available from the ARRL Web site at **www.arrl.org/qexfiles/**. Look for **11x06Gradijan.zip**. You may want to check the author's Web site periodically for other Command and Control software (**www.qsl.net/wb5kia**).

[2]B. Long, S*peech Synthesis & Speech Recognition Using SAPI 5.1*, **www.blong.com/Conferences/DCon2002/Speech/Speech.htm**.

[3]S. Gradijan, "A Talking Logbook with Rig Control," *QEX*, May/June 2006, pp 40-44.

Steve Gradijan is a Geoscience Consultant in the Dallas, Texas area. He holds an Extra Class license WB5KIA and has BS and MS degrees in geology. He has been involved in recreational programming in various programming languages for over 25 years.

Tech Notes

Linux for the Radio Amateur

Robert P. Haviland, W4MB

The following material assumes that you are a computer user, but that you are really a specialist in ham radio, not in computers. However, the material may be helpful to some computer geeks who have specialized in what is called Windows.

> **Where to Find Things to do With Linux**
>
> Looking for an up-to-date list of Linux projects? You can find them at **www.linux.org/projects**. Projects are divided into *General* (non-technical projects to promote Linux), *Hardware Port* (porting Linux to the PowerPC, Macintosh, Alpha and other systems), *Software* (developmental distributions, applications and drivers), and *Scientific* (high performance, parallel processing and research oriented).

What is Linux?

A working computer really includes three things: (1) a set of hardware, (2) an operating system to make the hardware functional and (3) a set of application programs to accomplish desired tasks.

An operating system is basically made of two parts: (1) the BIOS (Basic Input-Output System), plus (2) a set of programs to complete the hardware functionality. The BIOS is built into the computer by the manufacturer, but the remainder is entered separately, a step made possible by the BIOS. Common usage calls this additional set of programs the *operating system*.

Linux is such an operating system. It started in 1991 when a Finnish student, Linus Torvalds, became interested in a version of the Unix operating system called MINIX, and decided to write his own version. He published this as Linux under the General Public License (GPL) established by The Free Software Foundation. In the process, he used many of the routines of the foundation's GNU project, so you often see the terminology *GNU-Linux*.

Under the GPL, anyone can copy, use and modify the program code, create their own version and sell it. However, they must give full credit, and must make the source code available for others to use. A result was that programmers around the world wrote revisions and extensions to Linux code. Linus decided to act as coordinator and arbitrator to ensure that conformity to Unix continued, and that there was only one true version at any given time. This is the Linux kernel, the GNU code being kept separate, both making up the Linux operating system. The GNU project also completed their goal of a revised Unix, now called the HURD operating system. There is also another based on Unix, called FreeBSD.

A further result has been the creation of different "flavors" of Linux, called distributions. The Linux kernel and the needed GNU programs are always included, but there may be additions and extensions to those, designed to meet some goal, plus other original programs. Some of the common flavors are Debian, Mandriva, Red Hat, Slackware, SUSE and Turbo Linux, with Ubuntu being a late one. All provide one or more graphical interfaces as well as typed-in commands that give more comprehensive control of the hardware and programs. There is also Linspire, designed to look like Windows. And there are some special versions — for example, to run on an old computer with limited memory, or to fit on a single floppy disk. I have not checked all of the above flavors, but I believe each has a Web site.

In more detail, Linux is a multi-tasking, multi-user system. Subject to available memory, it can run many programs at the same apparent time, on a priority basis, by swapping from fast memory to disk. I once tried to find a limit, but gave up — it was more than I'll ever need. Users are assigned a virtual computer. In the graphics mode, it is necessary to log out the present user and log in the new one. If the GUI has not been started, switching is done by pressing F1 to F6, each independent of the others, and requiring separate log-in and password. Usually F1 is kept for a user called root, who has complete control of the hardware and software. Users other than root work under a permissions system, which can allow read, write and execute files (in Unix, programs are files), separated into root, group and all users. This is good for a family computer.

How Do I Get and Install Linux?

Getting Linux is free if you have an Internet connection; but remember that a full distribution can run to gigabytes, so a download can take a long time. Or, since there are no copy restrictions, perhaps get a distribution from a friend. More practical is to order a distribution on CD or DVD: Look on the Internet for vendors. Several issues of magazines include CDs or a DVD with a version, usually the same as the download version. There are books covering several of the popular versions that include CDs or a DVD of the distribution, probably the download version: They have the advantage of providing written instructions and reference material, and usually additional programs. The most complete offering — and also the most expensive — is a boxed set from a distribution vendor. That has the advantages of a full distribution, instruction manuals and support by telephone or e-mail, usually with a duration limit.

For a beginner, I suggest a download from a source listed later, or a magazine or book. Installation can be a matter of inserting a CD, and following instructions.

With hard disk prices as low as they are, I strongly recommend a separate drive for Linux. Even an old 10 MB drive will do; but if you are to do much work, go for 50-100 MB or more. When installing a Linux distribution, setting the partitions is one of the early actions. The recommended way is to partition the drive in three parts: to /(root) = say 10 MB, swap of 2-3 times the computer memory and home for the rest of the drive.

Another decision relates to the type of installation, such as Workstation, Desktop, Server, Custom or Upgrade. Except with a small drive, starting with Workstation is recommended. An important later decision relates to multi-booting; for example, selection of either Windows or the selected version of Linux with automatic startup of one. More complex arrangements are possible. This is recommended since it allows you to use a favorite program not yet available to run under Linux. Other than a few more decisions such as language and time zone, installation is automatic, although some distributions require entering the type of mouse, keyboard and monitor, instead of having an automatic determination.

Once Linux is loaded, I suggest spending some time in practice. Hint: Never use root unless you need to get full control of the computer. The other users cannot damage the operating system, nor can intruder programs get to it. Use the terminal program as a command line interface, and get familiar with the com-

1035 Green Acres Circle N
Daytona Beach, FL 32119
bobh@iag.net

mon commands. Most are abbreviations, cd = change directory, ls = list (contents of file), su = superuser (root), but beware of rm = remove.

Get familiar with the contents of directories, with moving among them and with the graphical interface programs, and choose the one you prefer as the default. Also learn to use the provided manuals, man1-6. I also suggest getting a copy of Open Office (**www.openoffice.org**), which includes a word processor, spreadsheet, database, graphics and other programs. Firefox is a better Web browser than the basic ones included with Linux (**www.mozilla.com**).

The other programs to obtain depend on your interests. Look in **www.linux.org** and **www.sourceforge.net**, as well as the links to other sites they provide.

Ham Radio Programs

You might find it interesting to enter "Linux ham" in a search engine such as Google. The last time I tried this, over 14,000 references returned but, of course, there are duplications and some were unrelated to radio. One reference was **www.radio.org/linux**. This included programs under the headings Satellite, Shack Automation, Packet, Morse, AMTOR, PACTOR, SlowScan, FAX and Design/Construction. Subheadings under the last included the *SPICE* analysis program and Audio Spectrum analyses.

Don't forget to look at **www.arrl.org**. Try **www.qsl.net**, and **www-w6yx.stanford.edu** and the call signs of hams you know are working on items you are interested in. The site **emlib.jpl.nasa.gov** has programs and links to sites of special interest to professionals in all aspects of electromagnetics, including antennas and propagation.

Many universities have specialized in particular fields. To check these, try **www."schoolname".edu**. A useful one is the Finnish University, **ftp.fu.net**. Look in the folder /pub for programs for Linux, under /linux/languages for compilers for Basic, Fortran and others, and under /ham for radio programs. There is an extensive list of mirrors, which duplicate other sites. An interesting site if you follow propagation is **www.solar.cfa.edu**.

While this material is intended specifically to be an introduction to Linux possibilities, most of the sites have material useful under other operating systems. For these, as for Linux, the possibilities are so vast that the best approach is to search the Internet under the name of the field you are interested in.

Should You Change to Linux?

That is a question you will have to answer yourself. Overall, Linux should be stable, more immune to viruses and intruders, and definitely more adaptable to special needs. If you have an old computer, you can increase the protection by loading it with a special Linux program, and using it as a firewall between your regular computer and the Internet or even a local net; see **www.smoothwall.org**. This works for all operating systems.

On the other hand, if you do not know Unix, Linux takes time to learn. And probably a decision to change will depend on the availability of programs to do things you want. If you have a pet commercial program, ask the source if there is a Linux version available. More and more, these are being provided. If not, check the above sites to see if there is a public-domain program that does the same thing (possibly better). There are public-domain programs to run Windows programs under Linux, such as *Wine*, and also commercial ones.

Don't forget the benefits of dual-booting: you can have what you have now, and much new. For example, I just completed an upgrade of an old box. It now has a 64 bit processor, 1 GB of RAM, 250 GB of hard drive, and Floppy, Zip and CD/DVD drives, all running under desktop and developer Linux, which, incidentally is the only current operating system to use the 64 bit mode of the processor. I am in the process of adding application programs, including *Wine*, *Grass* (a geographic information system) and others, and transferring data from files, spreadsheets, and so on to the new machine. I am enjoying the fast operation, even though the CPU is currently under-clocked by 40% in the interest of cooler operation. QEX

In the next issue of QEX

No radio transmission suffers more from selective fading than medium and high-speed digital types. John Stanley, K4ERO, explains how the DRM (Digital Radio Mondiale) international broadcasting standard copes with dispersive channels. He uses a PC with a sound card and one of the available software demodulation packages to plot the data. John offers insights about what to expect over certain paths, both long and short. The results may surprise you! QEX

Antenna Options

L. B. Cebik, W4RNL
1434 High Mesa Drive
Knoxville, TN 37938-4443
cebik@cebik.com

From Two to One

In the preceding two columns we examined clusters of options for vertical and horizontal bidirectional wire arrays based on the dipole as the root antenna. Each cluster only sampled the options available, and we might easily add to the number of clusters. Our goal was to present each cluster in an internally consistent environment to ease some of the difficulty in decision-making. Within each group, the individual can now use consistent performance data in combination with dimensional data, installation challenges, and budgetary cautions when deciding on the next antenna to build.

Very early on, a perennial question arose: how does one convert a bidirectional array into a directional beam? Once more, we have options. Most handbooks treat the options in

Table 1
Modeled Performance of a 40-Meter Half-Square Array and 2 Derivative Parasitic Beams. (See Figure 1.)

Version	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*	*Feed Point Z R ± jX Ω*
Single	3.28	18	—	86	53.2 + *j*1.7
Fixed Beam	6.53	17	25.14	76	59.7 + *j*1.0
Reversible	6.49	17	20.40	76	54.6 + *j*2.6

Notes
1. All arrays use AWG no. 12 copper wire and have a top height of 50 feet above average ground (conductivity = 0.005 S/m; relative permittivity = 13).
2. Reversible beam uses a 50-Ω transmission line with an electrical length of 17 feet at the reflector corner, corresponding to the driver feed point corner.

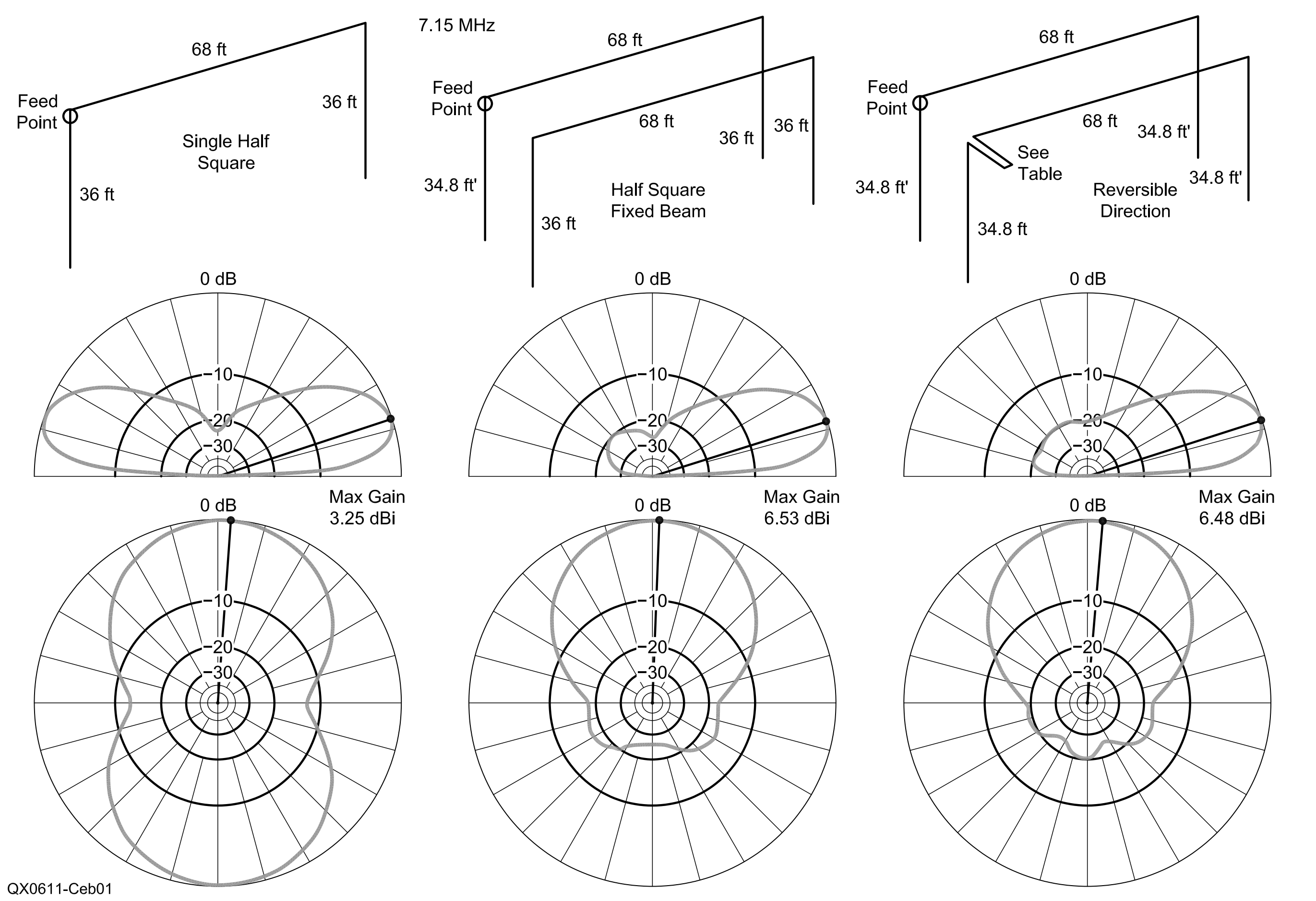


Figure 1 — Transforming the bidirectional half-square array into a directional beam with a parasitic reflector.

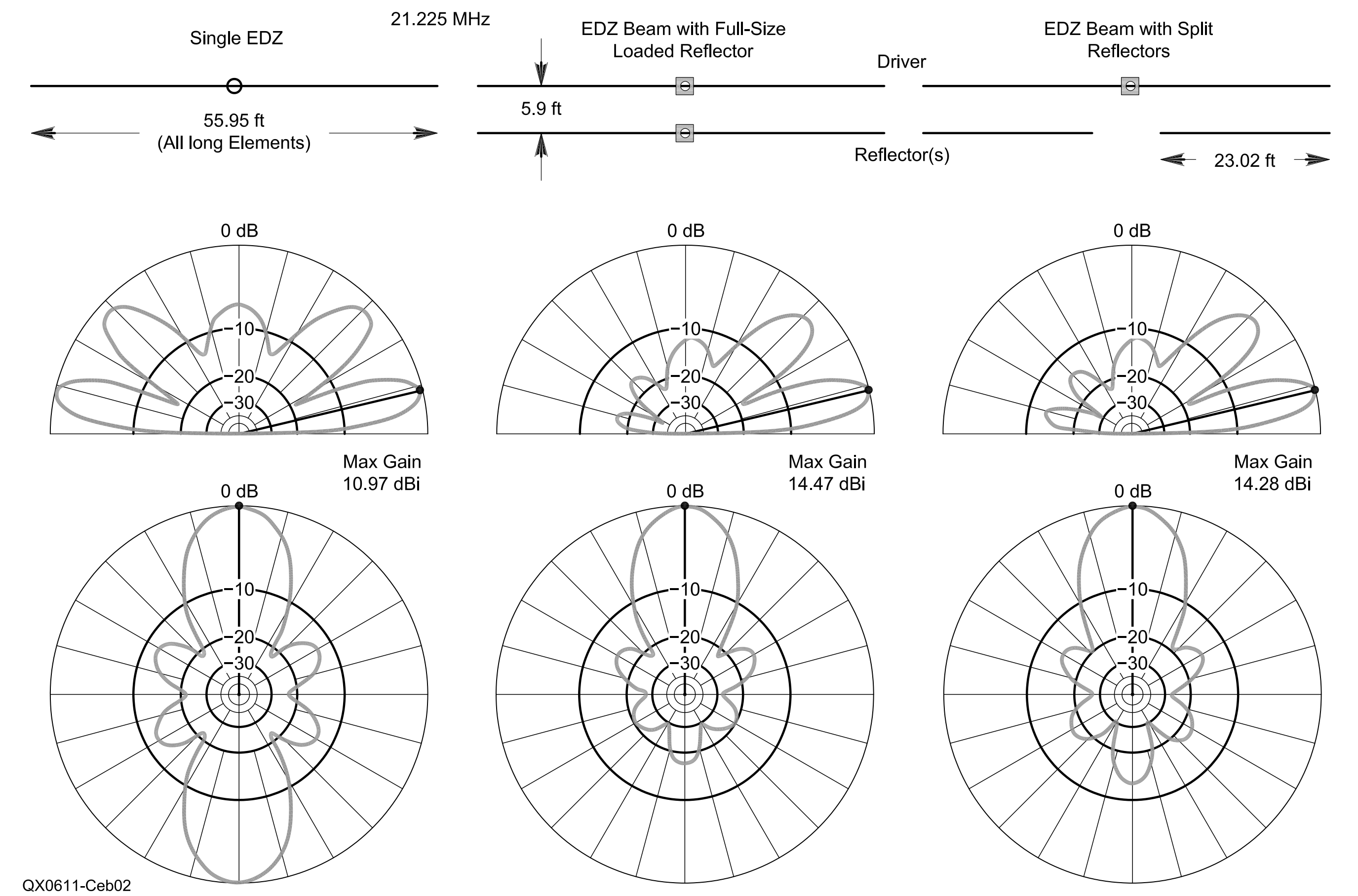


Figure 2 — Transforming a bidirectional extended double Zepp into a directional beam with parasitic elements.

different chapters, so we rarely receive any kind of overview of available techniques. Although our space is small and therefore our samples will be few, we can partly correct the oversight.

Consider two elements of similar size, at least one of which would serve as a bidirectional antenna. To create a directional end-fire beam from these two elements, we have to provide each element with the correct relative current magnitude and phase angle to yield the strongest possible forward lobe and the weakest possible rearward lobe. Similar considerations apply to both multielement arrays and to large collinear arrays. As a convenience in sorting out how we accomplish the goal, we can separate the techniques into four general groups: parasitic methods, phasing methods, screen reflection, and traveling-wave termination. Not all techniques are equally applicable to all antenna types and situations, but all have a place in the option set.

Parasitic Methods

Many hams associate parasitic elements almost solely with Yagi-Uda arrays based on the half-wavelength antenna element. Designers, however, have long realized that parasitic elements will work with almost any bidirectional antenna. A parasitic element approximates the conditions that we may achieve directly by complex phasing systems, but it does its work solely by virtue of its length, diameter, and spacing from the original element. For a given spacing and diameter, if we shorten an element relative to the original one with a feed point (now called the driver), the new element will show a negative relative phase angle for its current, and the main forward lobe will form in its direction. It has become a director. If we lengthen the element relative to the driver, then its current will show a positive phase angle relative to the current on the driver, and the main forward lobe will be in the direction of the driver. The new element has become — by convention — a reflector. Reflector-driver arrays tend to show a wider operating bandwidth but poorer gain and front-to-back characteristics than optimized driver-director arrays. Hence, reflector-driver arrays are easier to tame, that is, easier to field-adjust into acceptable performance.

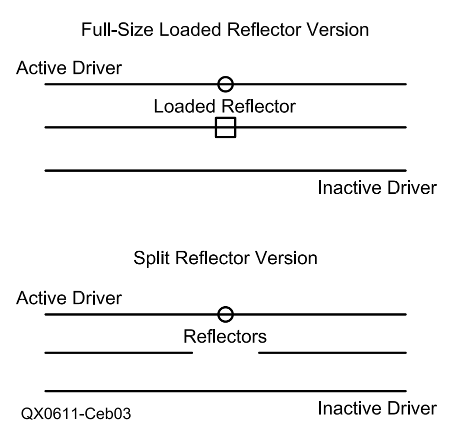


Figure 3 — Two possible ways to create reversible-direction EDZ beams.

We may fairly use these basic principles of parasitic element operation to create a large variety of directional beams.

Consider the half-square array, essentially a broadside array consisting of two vertical dipoles in an easily reproduced wire antenna. See Figure 1. The sketch shows the single half square along with two parasitic variations. The dimensions differ slightly from those used in our first array article (Jul/Aug 2006 *QEX*) to yield a near 50 Ω feed point impedance. We are still using AWG no. 12 copper wire, average ground, and a maximum height of 50 feet. The left-most plot shows the azimuth pattern with its slight misalignment due to the corner feed point. Table 1 supplies the modeled data.

We may add to the original array a reflector and optimize the vertical lengths for maximum front-to-back ratio with the selected spacing. The forward gain increases by about 3.2 dB with only a small narrowing of the beamwidth. The 180° front-to-back ratio is excellent. The rearward quartering lobes show a very slight imbalance, but not to a point of creating problems in using the array. The structure is very simple to replicate, and the feed point impedance remains "coax-ready."

Rather than lengthen the reflector physically, we may also lengthen it electrically. One convenient method is to use a shorted length of transmission line to add enough inductive reactance to a driver-length element to create an effective reflector element. The sample uses 50-Ω line at an electrical length of 17 feet. Suppose now that we add such a length (adjusted physically for the velocity factor of the selected coax) to the driver and bring it to a central point between the elements. We can install a remotely operated switch that converts one side into just an extension of the feed cable and the other side into the required shorted-stub. (Use a DPDT switch or relay to isolate the braids as well as the center conductors.) We now have a reversible half-square beam with virtually the same gain as the fixed version and a very useful rearward lobe set. The need to place the load and the feed point on the same side of the array creates a bit more imbalance in the pattern, but nothing fatal to the array's use.

We can also apply parasitic methods to collinear arrays, for example, the extended double Zepp (EDZ) shown in Figure 2. The sample EDZ uses a length just above 1.2 λ to reduce the side lobes. Like our earlier horizontal bidirectional arrays, the sample uses AWG no. 12 copper wire and is 1 λ above average ground. Table 2 provides the modeling data for both the single EDZ and two versions of a parasitic beam.

The first version uses a full-size reflector element. Since the element will be capacitively reactive, we can incorporate electrical lengthening into the center loading inductive reactance and tune the element for maximum front-to-back ratio. We also have the option of adding a coil to the center of the driven element to achieve resonance and simplify matching. For example, with the same parasitic array set for resonance, we might use a ¼ λ section of 75 Ω cable as a matching section for a 50 Ω main feed line. Providing the driver with inductive reactance at a normal Q range of 200 to 250 has its costs in terms of a slightly lower forward gain. The forward gain, however, shows well over a 3 dB increase relative to the single EDZ element.

We may also simplify the structure somewhat by using a split reflector element, with a section for each outer half-wavelength of the driver. The front-to-back ratio will not be as high as with a full-size element properly loaded, but we save the trouble of having to field adjust a reflector inductor for maximum front-to-back ratio.

Figure 3 shows two ways of obtaining reversible EDZ beams from the same reflector options. In both cases, we find decreases in the reported data from the model in one or another performance category, as shown in the bottom lines Table 2. Mutual coupling between the active and inert drivers is stronger than in the case of comparable reversible-direction wire Yagis. Nevertheless, some combination of elements may satisfy a given set of operating needs.

Phasing Methods

We often employ phasing methods to improve the rearward null performance of an array. Rarely do phasing methods show very marked gain increases over parasitic methods applied to essentially the same array. In some cases, such as the fixed half-square beam, resorting to phasing techniques would only add considerable complication for little, if any, detectable improvement. In other cases, however, phasing treatment may show marked improvement.

Phasing techniques essentially add an energy source to what we may achieve with parasitic methods. Parasitic techniques achieve the goal of setting the relative current magnitudes and phase angles to create a beam by virtue of mutual coupling between the elements. A driver-reflector Yagi, for example, rarely surpasses 12-dB front-to-back ratio. If we phase feed the two elements, we can create current magnitude and phase relationships that can easily raise the front-to-back ratio to 20 dB. Indeed, if we are willing to sacrifice some gain, we can create 180° front-to-back values as high as 50 dB — but only for the narrowest of bandwidths.

Many hams have a restricted view of

Table 2
Modeled Performance of a 15-Meter Extended Double Zepp and 2 Derivative Parasitic Beams. (See Figures 2 and 3.)

Version	*Driver Loaded?*	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*	*Feed Point Z R ± jX Ω*
Single	No	10.97	13	——	35	223.5 – *j*1042
Full Refl	No	14.64	13	16.99	34	118.9 – *j*959.4
	Yes	14.47	13	17.05	34	123.9 + *j*0.6
Split Refl	No	14.46	13	12.83	35	115.6 – *j*986.7
	Yes	14.28	13	12.83	35	120.6 – *j*1.7
Reversible Versions						
Full Refl	No	14.59	13	16.05	34	119.7 – *j*959.7
	Yes	14.58	13	9.42	34	122.8 – *j*3.8
Split Refl	No	14.43	13	12.36	35	114.8 – *j*987.2
	Yes	14.56	13	11.05	34	119.9 + *j*1.2

Notes
1. Full-length reflectors always center-loaded for maximum front-to-back ratio. Split reflectors sized for maximum front-to-back ratio.
2. Where applicable, drivers are center-loaded for resonance at 21.225 MHz.
3. All loads inductive with Qs of 200-250.
4. Elements are AWG no. 12 copper wire. Element lengths are about 1.2 λ to reduce sidelobes.
5. All antennas modeled are 1 λ above average ground.

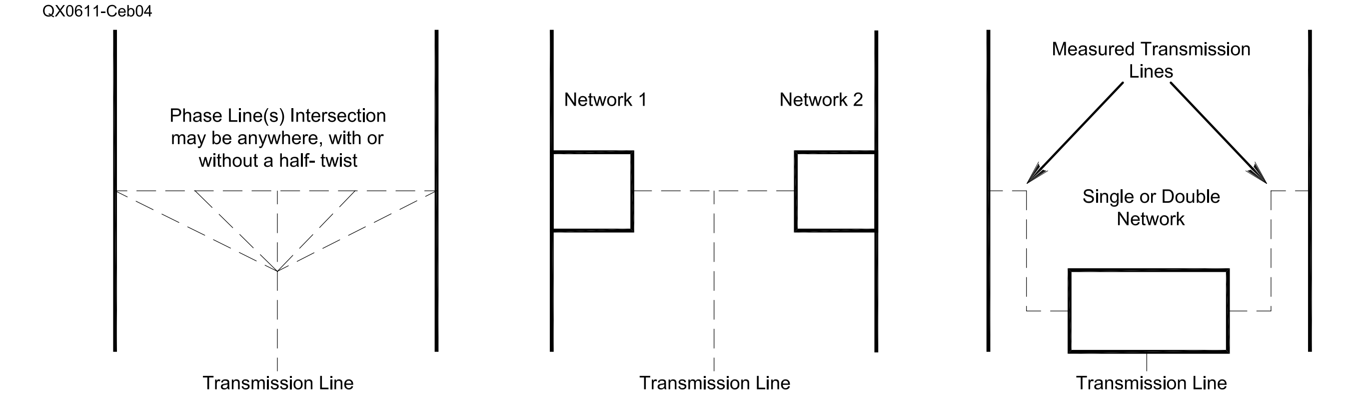


Figure 4 — General outlines of typical phasing setups for two-element end-fire arrays.

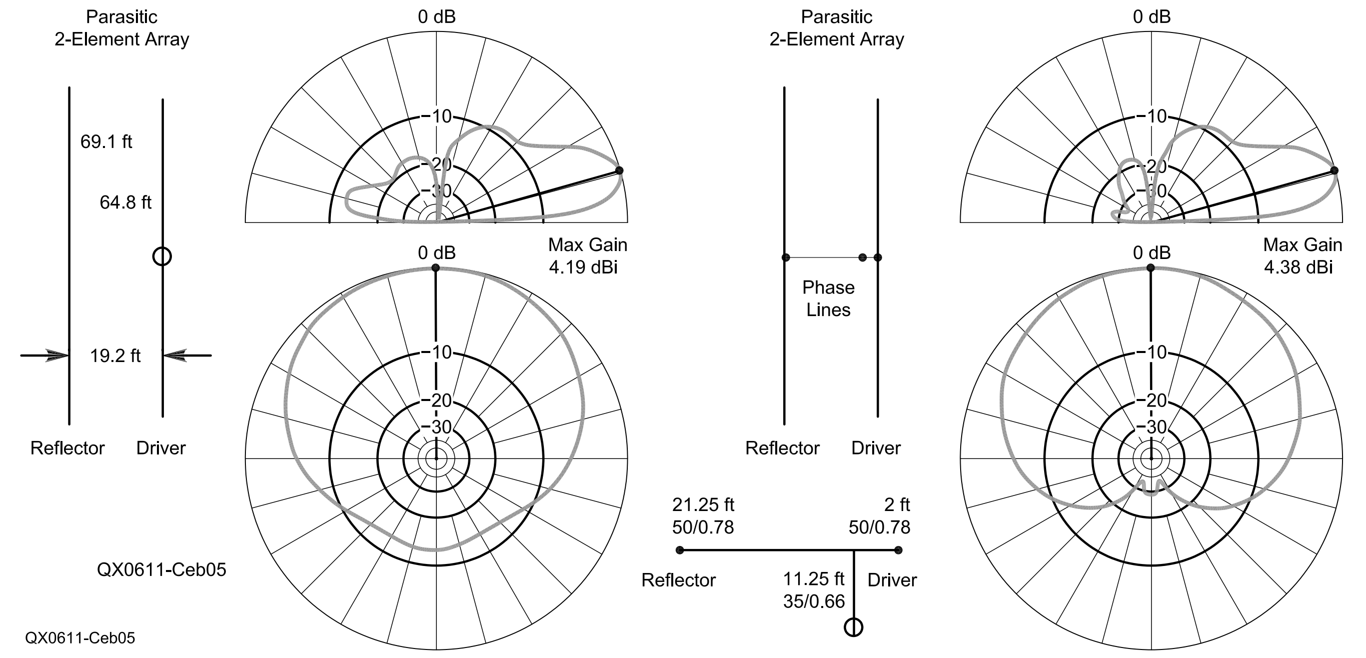


Figure 5 — Alternative 40-meter two-element vertical end-fire beams: parasitic and phased.

phasing techniques. *Low-Band DXing* by ON4UN has perhaps the largest collection of practical phasing techniques.[1] We may usefully, but incompletely, categorize the techniques into three groups, as shown in Figure 4. On the left are methods that use transmission lines to achieve the requisite splitting of current and transforming the current magnitudes and phase angles to optimal values. The ZL-Special technique of using a single $^1/_8$ λ line with a half-twist misled many builders into believing that line impedance transformation was the critical factor. However, the rate of current transformation along a line may not coincide with the impedance

[1]Notes appear on page 59.

Table 3
Modeled Performance of Parasitic and Phased 40-Meter 2-Element Vertical Dipole Array. (See Figure 5.)

Version	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*	*Feed Point Z R ± jX Ω*
Parasitic	4.19	15	12.19	139	43.8 + *j*6.6
Phased	4.39	15	27.95	139	54.4 + *j*2.4

Notes
1. Both versions use the same 1 inch diameter aluminum elements, sized for parasitic operation. Minimum height above ground is 15 feet. Element center points are the same height above ground.
2. Test frequency is 7.15 MHz. Ground is average.

transformation. Particular element setups may require a forward, a rear, or an intermediate feed point. One part of the line may require a half twist or not. Indeed, there are many situations in which available transmission lines or those we might construct simply will not provide the required current magnitude ratio and the phase angle difference that yields an effective beam.

The most direct route around this problem is to employ phase-changing networks at each element, as suggested by the middle sketch. Due to the weight of such networks, we ordinarily employ them with ground-mounted vertical monopole systems. An alternative is shown on the far right. We may employ ground-mounted networks and use measured transmission-line lengths to each element. Current undergoes a full cycle in 360° of line (not in 180°, as is the case for impedance). The lines may have different lengths, however, to add additional current and phase-angle transformations to the output of the network or networks.

Network calculations require more space than we have for this overview, but we can demonstrate the potential benefits of a transmission-line-based phasing system with a single example. Consider two vertical dipoles set up for parasitic operation, as shown in Figure 5. The element center points are at equal heights, and the lowest approach to average ground for this 7.15-MHz array is 15 feet. With the 1 inch diameter aluminum elements used for the array, we obtain the patterns shown in the figure and the modeled performance figures that appear in Table 3. Like virtually all good two-element driver-reflector parasitic systems, we obtain about 3.2 dB gain over a single vertical dipole. As we might expect from such an array, the maximum front-to-back ratio is just over 12 dB. With the element dimensions shown, the feed point impedance is close to 50 Ω.

Without changing any of the array dimensions, we may improve performance, especially with respect to the front-to-back ratio, as shown on the right side of Figure 5. The required phase line has two sections. The short section goes to the forward element. The longer section goes to the rear element and uses a half twist. The phase line sections consist of a foam dielectric coaxial cable with a 50 Ω impedance and a velocity factor of 0.78. (A 0.66 velocity factor would yield physical lengths too short to handle the spacing between elements.) The net feed point impedance is about 30 + j13 Ω, hence the need for a section of 35 Ω cable (or parallel sections of 70 Ω cable) to yield an impedance close to 50 Ω. The specified length entails a velocity factor of 0.66. For the effort of figuring or finding a usable phasing system, we obtain a very small increase in gain and a very large improvement in the front-to-back ratio. At the same time, we maintain a coax-ready array.

Planar or Screen Reflector Techniques

Screen or planar reflectors date back to the 1920s and the billboard array. Once considered ungainly at HF, they find extensive use in the UHF region. The reflector is largely untuned, although extensive modeling has shown that gain reaches maximum when the

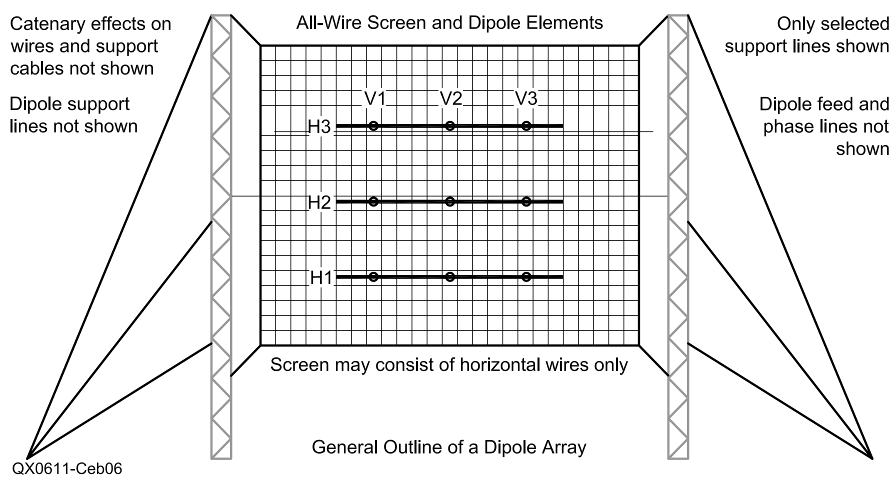


Figure 6 — General outline of a modern dipole array with a screen (planar) reflector.

Table 4
Modeled Free-Space Performance 10-MHz Dipole Arrays With a Planar (Screen) Reflector. (See Figure 6.)

Array Size	*1V-1H*	*1V-2H*	*1V-3H*
Maximum Gain (dBi)	7.45	8.68	10.11
Front-to-Back Ratio (dB)	19.24	21.18	21.78
Beamwidth (Degrees)	69.6	48.8	33.0
Array Size	*2V-1H*	*2V-2H*	*2V-3H*
Maximum Gain (dBi)	9.77	11.21	12.85
Front-to-Back Ratio (dB)	21.44	28.73	28.90
Beamwidth (Degrees)	72.4	48.8	32.8
Array Size	*3V-1H*	*3V-2H*	*3V-3H*
Maximum Gain (dBi)	11.72	13.22	14.87
Front-to-Back Ratio (dB)	21.58	29.33	29.34
Beamwidth (Degrees)	72.0	48.6	32.8

Notes
1. Test frequency 10.0 MHz.
2. Arrays use folded dipole drivers with AWG #10 wire throughout.
3. Gain above ground will be at least 4 dB higher, with the actual increase depending upon the height of the lowest dipole in the array.
4. Gain over bidirectional arrays (without the screen reflector) varies from 3.6 dB for the largest array to 5.3 dB for the smallest.

limits of the screen at about 0.5 λ beyond the limits of the driving array in all dimensions. The screen is unlike the parasitic reflector, which does not reflect but instead has a size to yield the correct current magnitude and phase angle to yield a maximized forward and a minimized rearward lobe. The screen is a reflecting surface that acts largely (but not exclusively) according to principles derived from optics.

In the HF region, the short-wave broadcast industry makes extensive use of screen reflector arrays with banks of dipoles fed in phase for a main beam that is broadside to both the screen and the driving array. Figure 6 shows the outlines of a very large array consisting of three rows and three columns of dipoles. Only a few of the support cables appear in the sketch. To have shown them all would have obscured the essential electrical elements of the array. It is possible to construct far smaller screen arrays to good effect. For example, a lazy-H is a bidirectional array that corresponds to a 1V-2H scheme in terms of the designations in Figure 6. Not only would a screen that is 2 λ high by 2 λ wide yield a directional signal, it would produce about 5 dB gain over the maximum gain of the basic lazy-H.

The models used to test the capabilities of the dipole array with a screen reflector use commercial standards. Commercial dipole arrays strive to be broad-banded. The wire is AWG no. 10 copper (although many arrays use aluminum-coated steel). A combination of folded dipoles and a spacing of 0.3 λ yields a wider operating bandwidth, although it produces less than absolutely maximum gain. Table 4 shows the modeled free-space performance of screened arrays from 1V-1H up to 3V-3H. The largest array of the group is among the models available for this Antenna

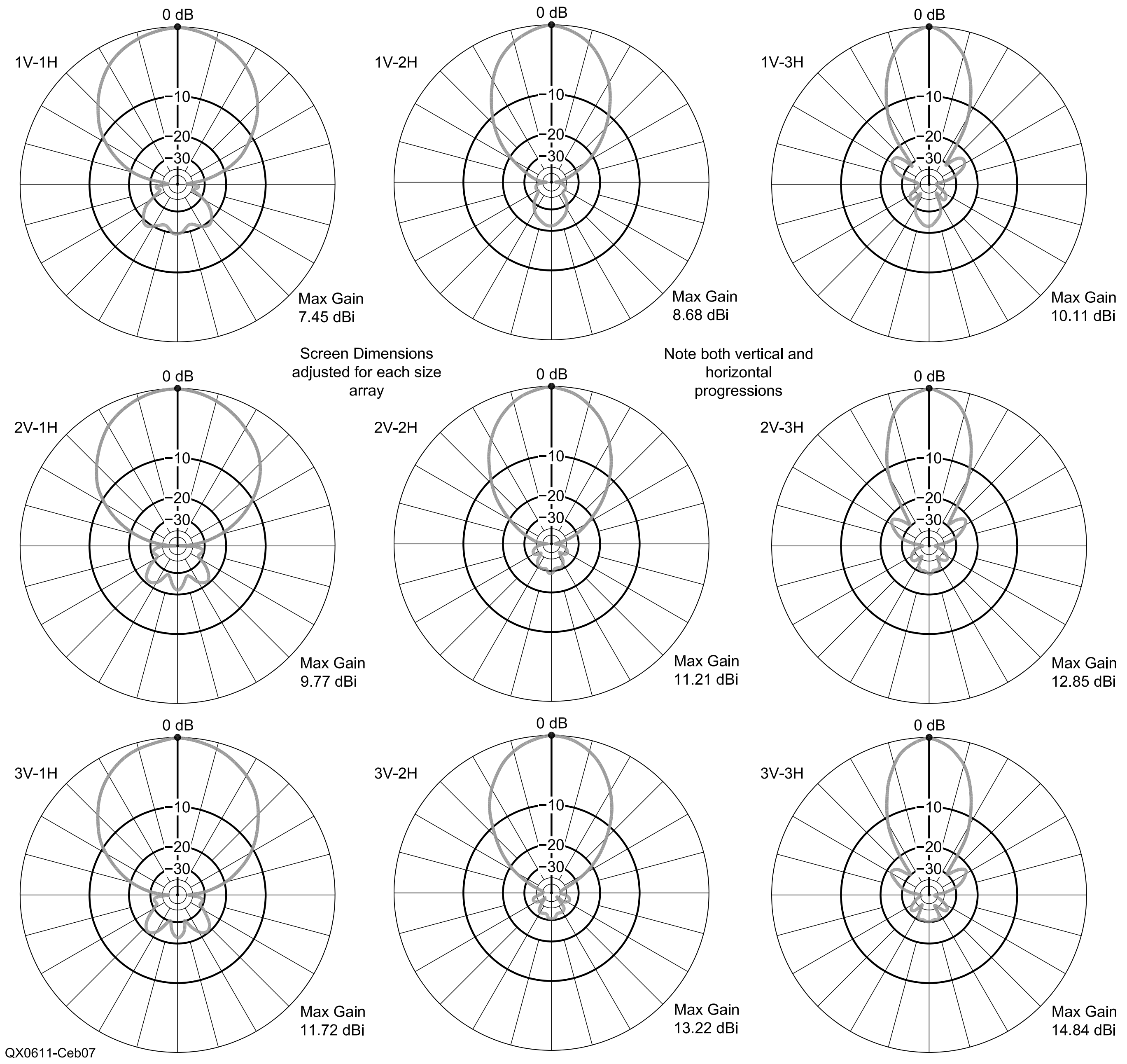


Figure 7 — Free-space 10-MHz E-plane patterns of typical dipole arrays with a screen (planar) reflector.

Options column.[2] Corresponding to the data entries are the azimuth patterns that appear in Figure 7. Each pattern rests on the use of a screen with dimensions optimized for the driving array size. (Commercial HF dipole arrays may skimp a bit on the screen size, and because the elements are horizontally polarized, the screen may use only horizontal wires. When a screen uses only horizontal wires, some reflector elements may exhibit parasitic properties, including higher-than-normal current magnitude excursions.)

The data and patterns are notable in several different ways. As we add elements vertically, the beamwidth does not change, although for each column, we find about the same gain increase with each added vertical bay. When we read the table and the patterns horizontally, we discover that each horizontal bay that we add reduces the beamwidth and increases the front-to-back ratio. One major advantage of the dipole array is that it allows the operator to select the best compromise between gain and beamwidth just by opting for an array size. In addition, by phasing each vertical bay in equal increments (for example, 30°, 60°, 90° for a three-bay array), it is possible to slew the signal angle with respect to the plane of the array without major distortion to the pattern shape.

In amateur use, a planar or screen reflector accepts a wide variety of drivers. The drivers may be either vertically or horizontally polarized and range from simple dipoles to more complex bidirectional arrays. Half-squares, rectangles, and side-fed quads are all applicable to a planar reflector to increase the forward gain and to yield a very good front-to-back ratio. In addition, for these single-feed-point drivers, we can often find a spacing between the reflector and the driver that yields a desired feed point impedance, such as 50 Ω, without sacrificing significant gain. One caution to observe is not to bother using an end-fire array in front of the screen. The screen will dominate so that the phased elements will not provide any significant further gain.

Traveling-Wave Termination

Long-wire technology arose in the late 1920s and more especially in the 1930s in answer to the need for very precise transoceanic point-to-point communications. Edmond Bruce stands at the head of the class of antenna engineers for his invention of the "king" of long-wire antennas. Although these collinear arrays have largely passed out of service as installations strive to save real estate, amateurs still dream of wire antenna farms. Indeed, one humorist from past decades once specified his ideal DX antenna as a very large rhombic installed on a rotatable island. The basic long-wire shapes are the single long wire, the V, and the diamond or rhombic. Each of these shapes is essentially bidirectional. We obtain a directional beam by adding a terminating impedance at the array end opposite the feed point. In doing so, we convert a standing-wave antenna into a traveling-wave antenna, although the conversion is imperfect.

We measure the basic single long-wire antenna in wavelengths. The sample shown in Figure 8 is 4 λ long. As the patterns show, the long-wire has a split lobe. As we make the antenna longer, the lobes in each direction grow closer to the axis of the wire itself, but they never meet for almost all practical

Table 5
Modeled Performance of 3.5-MHz Unterminated and Terminated Long-Wire Arrays.

Single Long Wire Type	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*
Unterminated	11.99	13	1.85	20.8/lobe
Terminated (800 Ω)	7.99	13	16.48	25.5/lobe
V Array Type	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*
Unterminated	15.48	12	1.86	14.2
Terminated (750 Ω)	12.35	12	21.97	14.8
Rhombic Type	*Gain (dBi)*	*TO Angle (Degrees)*	*Front-Back Ratio (dB)*	*Beamwidth (Degrees)*
Unterminated	18.65	13	2.90	13.6
Terminated (850 Ω)	17.39	13	45.68	13.4

Notes
1. Test frequency 3.5 MHz.
2. Arrays use 0.16 inch diameter (AWG no. 6) copper wire and are 1 λ above average ground.
3. Terminated single long-wire and V beams have vertical wires to effect a termination to ground. Rhombic termination is in series with the total wire structure.

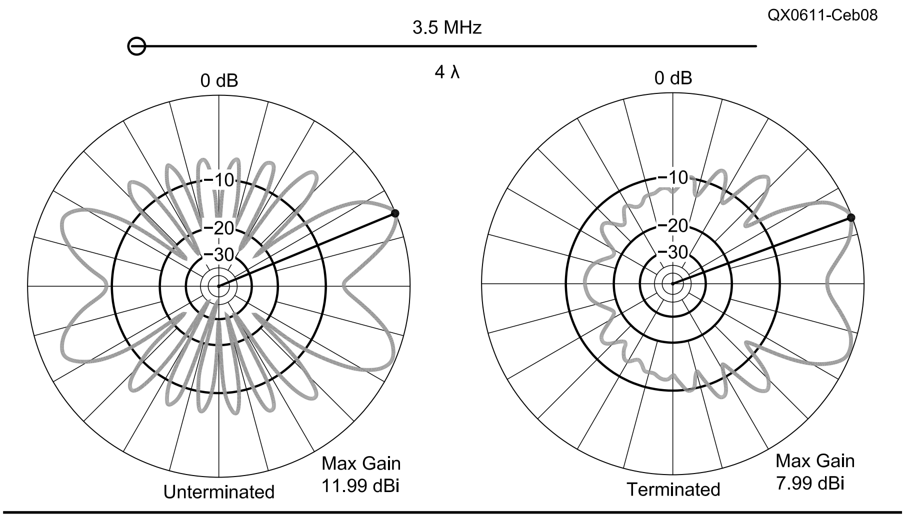


Figure 8 — General outline and azimuth patterns of unterminated and terminated 4 λ long-wire arrays 1 λ above average ground.

lengths. The sample long-wire antennas in these notes use 0.16 inch diameter (AWG no. 6) wire and are 1 λ above average ground at a test frequency of 3.5 MHz. The end-fed single wire has a lobe count that is twice the number we would expect of a center-fed wire of the same length. As the data in Table 5 show, an unterminated or standing-wave long-wire has a natural imbalance of nearly 2 dB away from the feed point.

When we add a terminating impedance (in this case 800 Ω), we obtain the directional pattern shown in Figure 8. The gain drops significantly, not only because of the terminating resistor, but as well because each end of the array must have a wire to a common ground. The two vertical lines tend partially to fill the pattern nulls and thereby prevent the array from obtaining full gain. Nevertheless, we also obtain a very wide operating bandwidth. In fact, the SWR bandwidth is significantly wider than the practical bandwidth for the pattern shape.

The V array in Figure 9 rests on the long wire. The angle between the two legs is a function of the lobe angle of each leg — as shown by the lobes of the single long wire. Two lobes coincide down the centerline between lobes, resulting in a bidirectional pattern that is close to equal strength in both directions. The feed point is in series with the legs at their junction. If we wish to add a termination, we must run lines to ground, where we place separate terminating impedances, or we must run a line between the far ends of the V with the termination centered. In either case, we discover a decrease in gain relative to the unterminated array. The pattern is very strongly directional, however, and the SWR bandwidth far outstrips the practical operating bandwidth. With a V array, changing operating frequency changes the leg length as measured in wavelengths. Hence, the angle for an acceptable pattern has a limited frequency range of about 2:1.

With the possible exception of VK3ATN's moonbounce 2-meter rhombic in the 1960s, I am unaware that anyone has ever bothered to use an unterminated rhombic, as shown in Figure 10. I have included data on it in Table 5, however, to compare with the far more common terminated version. A rhombic with 4 λ legs is, of course, twice as long overall as a V with 4 λ legs, but the width is about the same, since the same angle-based construction is involved. The terminating impedance of a rhombic (850 Ω in this sample) is in series with the collinear array wires. Therefore, we see far less difference between the gain of the unterminated and the terminated versions. We can achieve very high front-to-back ratios, however.

For commercial service, the major fail-

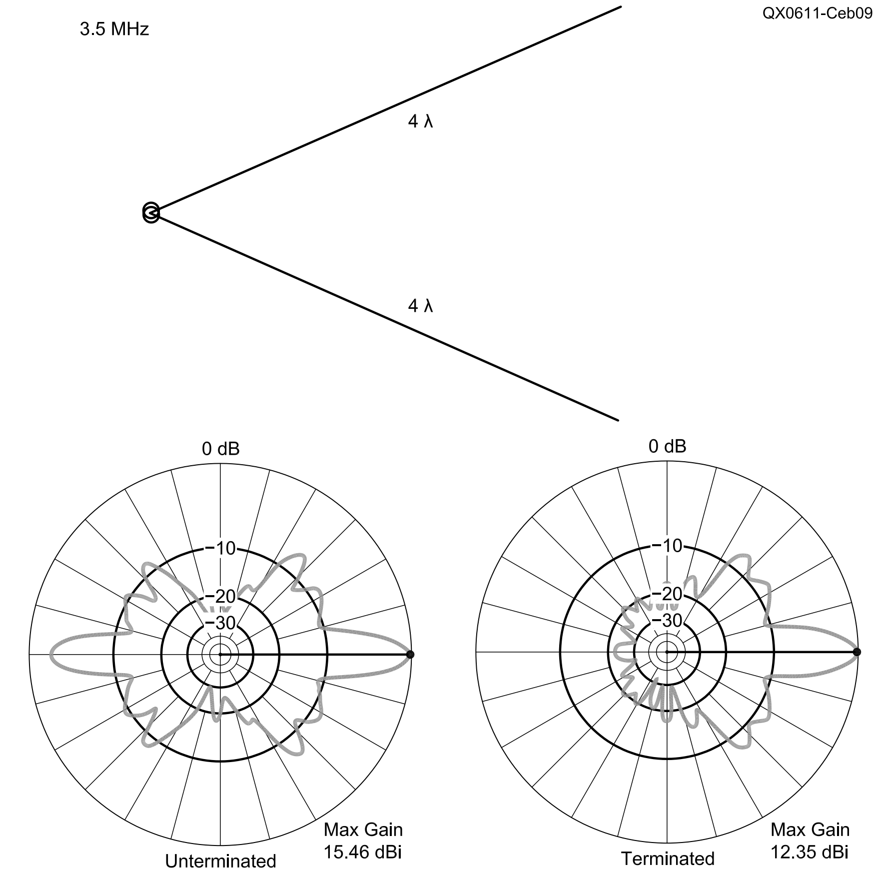


Figure 9 — General outline and azimuth patterns of unterminated and terminated V-arrays with 4 λ legs 1 λ above average ground.

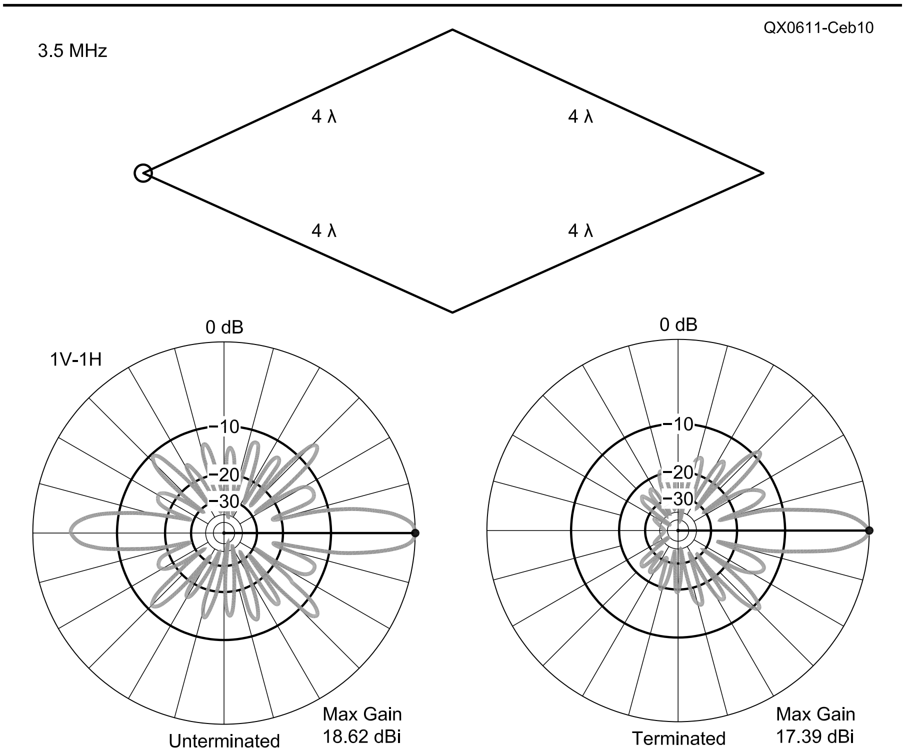


Figure 10 — General outline and azimuth patterns of unterminated and terminated rhombic arrays with 4 λ legs 1 λ above average ground.

ing of all long-wire technology was the high level of the sidelobes, clearly evident on all of the patterns. The correct V or rhombic angle might combine two or more long-wire lobes, but it did little to suppress the other lobes in the long-wire pattern. For amateur service, optimized V-beams and rhombics have two major drawbacks. Acreage is one of them. The other is the extremely narrow beamwidth. Nevertheless, the technique of using a termination to convert a bidirectional array into a directional beam deserves its place among the options available to us.

Conclusion

These notes have taken a somewhat different approach to converting a bidirectional array into a directional beam by providing an overview of the major techniques available to us. Not every technique is applicable to every bidirectional array that we might wish to convert into a directional beam. As well, in some cases, a technique might be applicable in principle but highly impractical to implement. In most cases, however, we likely will find more than one potential method of conversion and can weigh each against the other factors (for example, complexity, finickiness, budget, and so on) that go into our final building decisions.

Notes

[1] J. Devoldere, ON4UN, *Low Band DXing*, 4th Ed, ARRL, 2005, available from your ARRL dealer or the ARRL Bookstore, ARRL order no. 9140. Telephone 860-594-0355 or toll free in the US 888-277-5289; **www.arrl.org/shop/**; **pubsales@arrl.org**.

[2] Models for the antennas discussed in this "Antenna Options" column are available in *EZNEC* format at the ARRL Web site. Go to **www.arrl.org/qexfiles** and look for **11x06_ AO.zip**.

Out of the Box

Raymond Mack, W5IFS
17060 Conway Springs Ct
Austin, TX 78717-2989

Trimble GPS Timing Receiver

A November/December 2002 article in *QEX* described how to use a surplus GPS trained VCXO as a frequency standard.[1] The September/October 2006 issue of *QEX* has a similar article.[2] The GPS receiver in the 2002 article relied on GPS receivers that were pulled from service in CDMA cell phone sites, and those are no longer available. Fortunately, there is an alternative available for purchase.

Trimble has a line of Acutime GPS smart antenna receivers that are designed for accurate time keeping. The latest addition is the Acutime Gold GPS smart antenna, which is an RoHS (European term for lead free) version of their current model. This device is intended to be pole mounted outdoors with just the power and data cable going down to the rest of the system. It is completely sealed against the weather. The antenna is roughly the size of half a baseball, and is dome shaped to enhance runoff of the elements.

The RS-422 interface includes both a 1 pulse per second time reference signal as well as data interface to send a time tag for each pulse. The 1 pps signal is accurate to within 15 nanoseconds of UTC. This is adequate to provide an accuracy of 1 part in 10^{10} when averaged over a large time period. Trimble has several articles on their Web site (**www.trimble.com/tmg_cdmaclocks. shtml**) that will help you understand how GPS trained clocks work.

The base receiver has a list price of $445.00 in small quantities. They also have a starter kit available that includes 100 feet of cable, user guide, RS-422 to USB converter, and *Windows* software for monitoring and communication. It is available through distribution. Here in Texas it is available from Lelko Technologies (Embedded Products, Timing), 2060 North Collins Road, Suite 130, Richardson, TX 75080;
Phone: 972-680-9588
Fax: 972-680-9589
www.lelkotech.com

Acutime Gold GPS smart antenna

Notes

[1] B. Jones, K8CU, "Using the HP Z3801A GPS Frequency Standard," *QEX*, Nov/Dec 2002, pp 49-53.

[2] B. Zauhar, VE2ZAZ, "A Simplified GPS-Derived Frequency Standard," *QEX*, Sep/Oct 2006, pp 14-21.

Letters to the Editor

Dual Directional Wattmeters (May/Jun 2006)

Hi Doug,

In my article, I attempted to provide an accurate technical explanation while keeping the mathematics to a minimum. In conversations with readers, I have found that the simplifications left a small degree of uncertainty about subtle details. In particular, it is not perfectly clear how the calculations properly handled phase information. But as I noted in that text, all of the equations are accurate if one solves them using phasors (sometimes referred to as vectors in this context) instead of scalar values.

The only major revelation in my analysis is the proof that these circuits actually measure true power in arbitrary complex loads, in a way that really does not rely on dual-mode wave motion on transmission lines. All the discussion about transmission-line theory is merely an aid to understanding but as I have shown, the circuits will correctly measure power without a transmission line involved. The following is a complete proof of this using proper ac phasor notation. I recognize now that I should have included this in the original article, but better late than never!

Before jumping into a mathematical jungle, let me review the issue quickly. Any passive device, including an antenna, has an impedance $R + jX$. When connected to a sine-wave generator of frequency ω, there will be a transfer of power from the generator to the load. It will generally be accompanied by a "circulating" power whose time average in the load is zero. The goal of power measurement is to measure the non-zero true power even if it is accompanied by any amount of the latter.

Using the notation from the article, particularly from Figures A1, A2 and A3, we assume the wattmeter to be inserted between generator and load, Z_L, resulting in current as shown, with a voltage drop, V. These will now be represented as phasors **V** and **I** [here in boldface — Ed.], which are applied to Equations D1 and D2 as usual, resulting in the phasor "output" voltages v_1 and v_2:

$$\mathbf{v_1} = k_1 \boldsymbol{V} + k_2 \boldsymbol{I} \quad \text{(Eq 1A)}$$

$$\mathbf{v_2} = k_1 \boldsymbol{V} - k_2 \boldsymbol{I} \quad \text{(Eq 1B)}$$

where $k_1 = C_2 / (C_1 + C_2)$ and $k_2 = R_S / 2N$.

Those equations are easily manipulated to a useful form by taking the complex conjugate of both sides of each equation and noting that $XX^* = |X|^2$

$$|\boldsymbol{v_1}|^2 = k_1^2\, \boldsymbol{VV^*} + k_2^2\, \boldsymbol{II^*} + k_1 k_2(\boldsymbol{VI^*} + \boldsymbol{V^*I}) \quad \text{(Eq 2)}$$

$$|\boldsymbol{v_2}|^2 = k_1^2\, \boldsymbol{VV^*} + k_2^2\, \boldsymbol{II^*} - k_1 k_2(\boldsymbol{VI^*} + \boldsymbol{V^*I}) \quad \text{(Eq 3)}$$

Note that the voltages $\mathbf{v_1}$ and $\mathbf{v_2}$ are the actual measured voltages, which are then displayed in units of volts-squared on the meters, multiplied by a constant, with units of W/V^2. Such measurements are scalar and necessarily measure the *magnitude* of the voltages. Subtracting these two equations and replacing the amplitudes $\boldsymbol{V}$ and $\boldsymbol{I}$ by their RMS values, we get:

$$\boldsymbol{v_1}^2 - \boldsymbol{v_2}^2 = 2\, k_1 k_2\, (\boldsymbol{VI^*} + \boldsymbol{V^*I}) \quad \text{(Eq 4A)}$$

$$\boldsymbol{VI^*} + \boldsymbol{V^*I} = |\boldsymbol{V}|^*\, |\boldsymbol{I}|\, (e^{j\theta} + e^{-j\theta}) \quad \text{(Eq 4B)}$$

$$= 2VI \cos\theta$$

$$\boldsymbol{v_1}^2 - \boldsymbol{v_2}^2 = 4\, k_1 k_2\, VI \cos\theta \quad \text{(Eq 4C)}$$

where V and I are now RMS values, and θ is the phase angle between V and I.

That result now shows that the difference of the two meter readings is indeed proportional to the true load power, as expressed by its classical formula in terms of the phase angle θ. The final calibration formula is established by re-writing this using $P = VI \cos\theta$ and a meter-scale calibration factor C (in W/V^2):

$$P = C\, (v_1^2 - v_2^2) / 4\, k_1 k_2 \quad \text{(Eq 5)}$$

All of the circuit constants are absorbed into the constants k_1 and k_2, leaving a true-power value that is independent of the load. Hence, it is proven to be a true wattmeter.

The relevance of Z_0, the calibration impedance, is obviously secondary. It is buried in the circuit constants and its only relevance is that it can be selected to produce a null reading on the reflected power meter for one specific value of load resistance, usually chosen as 50 Ω. However, as the equations show, if this value were changed from 50, the resulting change in k_2 could be compensated by a re-adjustment of either k_1 or the meter calibration resistors (which change the constant C) to maintain overall power calibration.

Final conclusion: DDW operation using the circuits discussed does *not* depend upon transmission-line wave mechanics. These devices function perfectly well for arbitrary two-terminal loads. If there is a transmission line involved, however, those wave mechanics will force the V and I relationships on the line to assume values consistent with the load power and phase. That results in the apparent presence of the standard forward and reverse wave modes, and if the line impedance is Z_0, then the meters will correctly read the amount of power in each mode. Otherwise, the individual meter readings are of little use.

— 73, Eric von Valtier, K8LV; **evonvaltie@aol.com**

Uniform Current Loop Radiators (May/Jun 2006)

Dear Doug,

Bob, NP4B, has initiated a new era of uniform current loops! I have followed that discussion with a growing interest, because this kind of antenna would solve my problems in monitoring NCDXF beacons.

After the comments about the topic in *QEX*, and kind explanations of the theory from Bob, I decided to simulate a single wire loop with lumped capacitors at the corners. This concept, also using four short supporting spiders, would create a sturdy mechanical construction with minimal wind load. Raising the height to a half wavelength puts the maximum elevation gain to more or less the optimal ionospheric propagation angle and at the same time reduces the man-made ground-wave noise to the minimum.

To my positive astonishment, the simulation result was better than I ever could have imagined. Perhaps Bob, with a better theoretical backgound, may add some further comment. In honor of him, Robert Zimmerman, and of course not forgetting myself, OH2RZ, I named the antenna the 2RZ-LOOP.

I am very thankful to Bob for his great idea and help that inspired me to get a better monitoring antenna than the usual quarter-wave vertical. Now I'm waiting to find time to put the antenna up before the harsh Finland winter sets in.

— 73, Ahti Aintila, OH2RZ; **ahti@attocon.com**

CW Shaping in Software (May/Jun 2006)

Hi Doug,

I am giving some thought to a question raised by a good friend of mine. His proposition is that a high standard of close-in phase noise in a receiver is more hype than of practical significance. This train of thought returned me to the age-old question of how much bandwidth is needed for CW.

I have read many papers on this topic including yours [**www.doug-smith.net/cwbandwidth1.htm**] and the article in *QEX*. It is clear that only 100 Hz or so are needed and hence I want to claim that close-in phase noise is indeed a primary requirement — especially true for CW. I have not seen any discussion of close-in phase noise objectives and would appreciate your thoughts on the topic.

— 73, Ron Skelton; **ron-skelton@charter.net**

Hi Ron,

Well, I think we have two things to consider: 1) how much interference you create because of your TX phase noise, and 2) how much QRM you suffer because of your RX phase noise.

1. Although you didn't mention TX, I do because in a transceiver, there's generally a close correlation between TX and RX phase noise levels. Obviously, the more PM and AM noise you transmit, the more you could QRM someone. And, at some relatively low frequency offset from your carrier, phase-noise power density will surpass keying sideband power density — but that's at a very low level — when the keying is done correctly.

2. While you may have the lowest phase-noise RX in the universe, the QRM you get from others may depend on the phase noise they're transmitting. In other words, there's little point in having RX phase noise that's more than about 6 dB better than that of the transmitter of the guy whose signal is giving you grief. Make sense? So if you've got the lowest RX phase noise possible, then you at least know it's not your fault. *— 73, Doug Smith, KF6DX,* QEX *Editor;* **kf6dx@arrl.org**

Doug,

Thanks for the response. Does the same argument apply to TX/RX IMD performance? *— 73, Ron Skelton;* **ron-skelton@charter.net**

Hi Ron,

Good! Yes, I think it does. The low-order IMD performance of some receivers isn't as good as that of many transmitters. I'd be willing to bet that today's solid-state transmitters dominate in high-order IMD, so high-order receiver IMD might not be an issue, but who knows? *— 73, Doug, KF6DX;* **kf6dx@arrl.org**

An Innovative 2-kW Linear Tube Amplifier (Jul/Aug 2006)

Dear Doug,

There are some minor errors that I did not detect when reviewing my article. Thanks to the readers who pointed out some of them by e-mail. The errors are:

1) On page 21, Figure 4 — Power supply schematic:

a) Fuse F2 is not connected to the R11/R12 node as shown but only to the R14/C24 node.

b) Pins 2 and 6 of CN5 are tied to pin 3 of CN7, not to ground.

c) The ground connections in this page and in Figure 5 at page 22 are not the same as the rest of the circuit. This ground is not isolated from mains ac supply and must be floating referenced to amplifier ground.

2) On page 23, Figure 7 — Control schematic: Contact A4 of switch SW1:1 must remain unconnected.

3) On page 24, the first paragraph incorrectly states: "…presents an open circuit to odd harmonics, a short circuit to even harmonics." The correct statement is just the opposite: short circuit to odd harmonics and open circuit to even harmonics.

It is also informative to tell that components Q1-5, Q7, Q11, Q17, D1, D3, D5, U12 and U13 are all mounted to the main heat sink or bolted to the aluminum frame.

I liked very much the final presentation and so did the friends to whom I sent the magazine! I had received several e-mails from readers who also liked the article — many thanks to them.
— Regards, Saulo Quaggio, PY2KO; **saulo@auttran.com.br**

Effective Directivity for Shortwave Reception by DSP (July/August 2006)

Doug,

Carl Luetzelschwab, K9LA, e-mailed me about my comment on circular polarization: "I have found no information as to how this is for DX signals." Carl told me that Bob Brown, NM7M has written about this topic. He pointed me to Bob's article, "Power Coupling on 160 Meters," *Communications Quarterly*, Summer 1999, pp 95-101. I would like to forward this information because it is a very good and understandable article.
— Jan Simons, PAØSIM; **pa0sim@amsat.org**

***Octave* for System Modeling (Jul/Aug 2006)**

Dear Doug,

Chuck Hutton and Ben Bennett, N7IVM, were kind enough to point out a bug in my code that is included in my article. I've been editing the code using a text editor and then running the resulting file from the command line of either a DOS window or the Linux command line and it runs fine that way. When the code is executed from within *Octave*, though, it fails.

The reason for the failure was pointed out by Chuck and is also found on page 111 of the *GNU Octave Manual* by John W. Eaton. The manual points out that to differentiate between script and function files, a script file (which is what we are using here) must not begin with the keyword "function." I had thought that the presence of a line in the file that must be interpreted by the operating system to locate the executable (the first line in the file that begins with "#!") would meet that requirement, but such is not the case.

A fix recommended by Chuck and also covered on pages 111 to 112 of the *GNU Octave Manual* is to insert a line that *Octave* will execute but which will produce no effect on the cal-culations or output of the program. One such line is "1;" (a numeral one followed by a semicolon and an endline character). You can insert it immediately before the line beginning with "function" and it will take care of the problem.
— 73, Maynard Wright, W6PAP; **m-wright@eskimo.com**

Hi Maynard,

Thanks for sending the updated file. I have posted the corrected file as part of **7x06_Wright.zip** in the July/August section of the file listings at the **www.arrl.org/qexfiles** Web site.
— Larry Wolfgang, WR1B, QEX *Managing Editor*; **wr1b@arrl.org**

Magnetic Coupling in Transmission Lines and Transformers (Sep/Oct 2006)

Doug,

In the article by Gerrit Barrere, KJ7KV, the equation on page 30 does not yield approximately $53 + j\,0\ \Omega$. It yields something quite different.

Further, the equation is the high-frequency approximation for the characteristic impedance of a transmission line. 60 Hz is not a high frequency for RG-58. Assuming 262 nH, 93 pF and 0.046 Ω/m, the characteristic impedance of RG-58 at 60 Hz would be about $810 - j\,810\ \Omega$.
— Bert Weller, WD8KBW; **sodiumflame@sbcglobal.net**

Hello Bert,

Sorry, a "typo" sneaked into that equation. The text says 90 nF and 260 μH. It should say 93 pF and 261 nH, which is what the equation should be, too. With those values, Z_0 does come to 53 Ω at high frequencies.

You're right about the Z_0 equation I used not being valid for such low frequencies. I was trying to make the point that transmission line behavior reduces to circuit analysis for electrically short lines, but over-looked the fact that 60 Hz is well below the "break point" of RG-58, where it starts to behave like a fully coupled line. Ott has a good discussion of this (see the article's bibliography). I should have used something like 100 kHz and 50 m of RG-58 for the example. Then the point would have been made with the numbers also correct.
— Gerrit Barrere, KJ7KV; **gerrit@exality.com**

Dear Bert,

Thanks. For the benefit of readers:

You're right about the 93 pF/m and 261 nH/m. From the ITT book (*Reference Data for Radio Engineers*, 6th Ed, Howard Sams & Co, 1975):

$$Z_0 = \left(\frac{R + j\omega L}{G + j\omega C} \right)^{\frac{1}{2}}$$

where Z_0 is characteristic impedance, L is inductance, C is capacitance, R is resistance and G is the conductance of the dielectric material (polyethylene), all per unit length; and ω is angular frequency ($2\pi f$). Assuming ωL is very small compared with R and that G is small compared with $j\omega C$ allows the simplification:

$$Z_0 = \left(\frac{R}{j\omega C} \right)^{\frac{1}{2}}$$

$$Z_0 = \left(\frac{-jR}{\omega C} \right)^{\frac{1}{2}}$$

Plugging in the values, $Z_0 \approx 810 - j810\ \Omega$ at 60 Hz. The characteristic impedance is high, capacitive and very frequency-sensitive. The Z_0 of RG-58/U doesn't approach $50 + j0\ \Omega$ until you get near 1 MHz or so.
— 73 de Doug, KF6DX

Dear Doug,

I really enjoyed the article by Gerrit Barrere, KJ7KV. He did a great job explaining the relationship between transmission lines and transformers.

I just disagree with one detail: the assertion that the total magnetic flux external to the wire increases greatly as the wire size is reduced. In fact, I believe that total external flux of an infinitely-long straight wire in free space is infinite and so does not depend on wire size.

It's easy to prove. As stated in the article, the external field is proportional to $1/r$, where r is the distance to the center of the wire. If you integrate $1/r$ from R (wire radius) to infinity you get $\log(\infty) - \log(R)$, which is infinity. The consequence is that the inductance per unit length of an infinitely long straight wire in free space is infinite!

When I first realized this many years ago, it made me horribly confused for a while. As a young engineer I had been told that a useful crude rule-of-thumb is that a wire has an inductance of 20 nH/inch (8 nH/cm). That seems like a contradiction. Of course, the explanation is that in real life you never have a wire infinitely far away from all other conductors. A real wire usually connects to something on each end and the circuit is completed, either through a ground plane or a returning wire. Because the inductance is proportional to the logarithm of the distance to the return path, it changes only very slowly as the spacing is varied.

As an example, a piece of bare 22 AWG hook-up wire (0.0253 inch diameter) spaced 0.32 inches (0.8 cm) above a ground plane actually does have an inductance of 20 nH/inch. However, if you move the wire twice as far away, the inductance only increases to 23.5 nH/inch. If you switch to wire that is twice as thick (16 AWG), the inductance at 0.32 inch spacing drops to 16.4 nH/inch. So, as long as you realize that the "20-nH/inch rule" is only a very crude approximation, it can be useful when deciding if component lead length is likely to upset the operation of your circuit.
— Alan Bloom, N1AL, ARRL TA, 1578 Los Alamos Rd, Santa Rosa, CA, 95409; **n1al@arrl.net**

Hello Alan,

Thank you for the feedback! You are correct for the ideal case of the isolated wire. This is one demonstration of why this condition does not exist. For an actual case, however, the total field is proportional to log(big number) – log(R), which is finite and does increase rapidly as R approaches zero. The point is that the ***H*** field increases greatly near the wire as the radius is decreased. I should have put it this way instead of "…the total field…" It is helpful to know how the field near the wire varies with wire size to understand the mechanism of coupling.
— Gerrit, KJ7KV

Dear Gentlemen,

Thanks for an excellent article! I always appreciate a new viewpoint on what is often a confusing subject. I especially appreciated your adding mutual inductance to the models.

I must respectfully disagree with you that a common-mode choke is not a current balun. The common-mode choke, to the extent that it impedes common-mode currents and passes differential-mode currents does indeed force a current balance into the load.

Think of a load that is imbalanced — that is, grounded somewhere other than at the center of the load. The Ruthroff balanced impedance converter is a voltage balun because it attempts to force the voltages in the two parts of the load to be equal (with respect to ground) at the expense of a current imbalance. That is a voltage balun.

Your common-mode choke, when presented with an unbalanced load, will force the currents in the two parts to be equal at the expense of an unbalanced voltage (with respect to ground). That is a current balun.

It should be fairly obvious that you can have one or the other, but that with an unbalanced load you cannot force both the voltage and current to balance simultaneously. We really don't care in the current balun that no load gives an output node at ground potential — it's the current balance that counts, and that is still balanced: zero in both legs!
— Respectfully, Glenn Dixon, AC7ZN; **dixong@ieee.org**

Hello Glenn,

Thank you for the note. You make a good point about the symmetrical nature of the two baluns, that voltage balance is not to be expected from a current balun and vice-versa. I agree that the article doesn't make this distinction very well. The point is that currents are not forced to be equal in this circuit; the degree of current balance depends on the relative impedances between the windings and the rest of the circuit. In that sense, the common-mode choke is only part of an overall filter, and dependent on the rest of the circuit. The voltage balun forces its balance regardless of the rest of the circuit, however.

Thanks for the clarification.
— Gerrit, KJ7KV

An L-Q Meter (Mar/Apr 2006)

Hi Doug,

A sharp-eyed reader, Paul Kiciak, N2PK, has discovered an error in the source code which is in the ARRL files as part of my article. There is a line in the code which says:

Za.r = 1.0 / Rinstrument;

It should read

Za.r = 1.0 /Ri;

I cannot explain how this error appeared since my own current copy of the source code doesn't have this error. I do apologize.
— 73, Jim Koehler, VA7DIJ, VE5FP; **jark@shaw.ca**

Hi Jim,

Thanks for the note, and for the corrected program file, which we have posted at the Web site. Readers can find the corrected file in the March/April section of the file listings at **www.arrl.org/qexfiles.** Look for **3x06Koehler.zip**.
— Larry, WR1B; **wr1b@arrl.org**

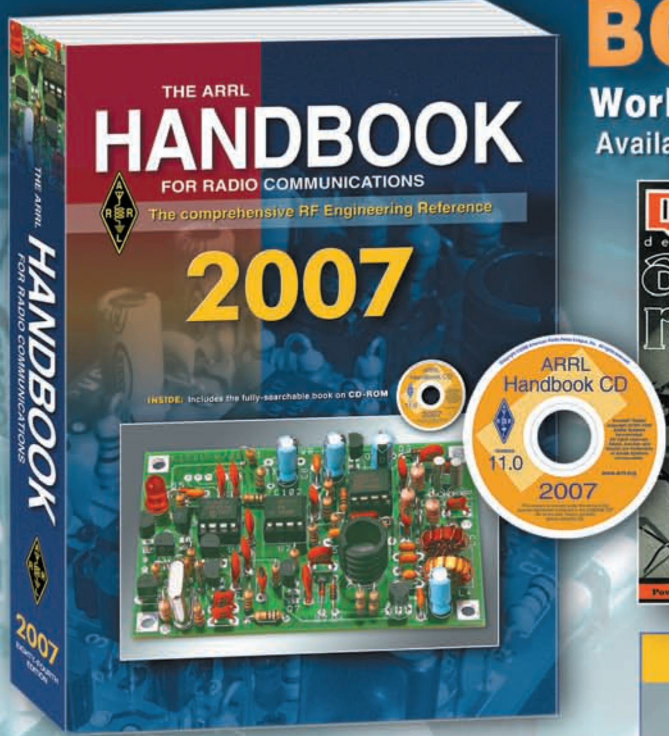
QEX Index 2006 — Issues 234 through 239

The ARRL Handbook — 2007 Edition

BONUS GIFT! While Supplies Last!

World War II Commemorative *QST* Offer!

Available from ARRL and Select ARRL Publication Dealers



You get all this when you order:

- **The ARRL Handbook**—2007 edition. Up-to-date and always revised!
- **The ARRL Handbook on CD-ROM**—version 11.0. included with every book.
- **BONUS "January 1942 issue of *QST*" reissue**—Limited edition.

65 Year Commemorative Reissue

Facts about the January 1942 issue of *QST*!

- First issue following Pearl Harbor attack, Volume XXVI, Number 1
- Contains last-minute press pages: "WAR COMES!"
- Includes the FCC Order suspending Amateur Radio operation in the US, issued December 8, 1941
- 100 pages of nostalgic editorial, feature articles, advertisements, and cartoons

Sixty-five years later, this commemorative reproduction of *QST* is a time capsule—a tribute to the perseverance and patriotism of radio amateurs!

The ARRL Handbook—2007
Eighty-Fourth Edition

The greatest applied electronics and communications reference of all time! Information on fundamental electronics concepts, components and building blocks, analog and digital radio design, troubleshooting techniques, antennas and much more. Make **The Handbook** your one-stop source for practical construction projects and weekend builds: radios, antennas, amplifiers, test equipment, DSP, accessories and more. Includes many new and updated projects:

- **High Power 50 MHz Low Pass Filter**
- **HiMite Low Power Transceiver**—build it for 20, 17, 15, 12 or 10 meters.
- **12-V dc Boost Regulator for Mobile Operation**—maintains 13.8 V when battery voltage decreases.
- **The TiCK-4**—a tiny CMOS keyer.

Plus, an improved digital communication interface for HF digital modes, low-cost remote antenna switch, a PIC-based HF/VHF power meter, top-loaded low-band antenna, and more.

CD-ROM INCLUDED! This edition is bundled with **The ARRL Handbook CD (version 11.0)**—the complete and fully searchable book on CD-ROM, including many color images, PC board templates, additional software and reference material.

Order Today www.arrl.org/shop or Toll-Free 1-888-277-5289 (US)

Softcover. Includes book, CD-ROM and Bonus "**January 1942 *QST***" reissue. ARRL Order No. 9760 .... **$44.95** plus s&h

Hardcover. Includes book, CD-ROM and Bonus "**January 1942 *QST***" reissue. ARRL Order No. 9779 .... **$59.95** plus s&h

Order now and get the special Bonus "**January 1942 *QST***" **reissue!** While supplies last.

Bonus "**January 1942 *QST***" reproduction offer while supplies last. A limited supply will be printed before it is returned to the "ARRL vault" for at least 10 years. Bonus "**January 1942 *QST***" premium cannot be redeemed for cash. No returns. Exchanges must be accompanied by premium. Valid only on pre-orders direct from ARRL or from Select ARRL Publications Dealers. This offer may be canceled or modified at any time due to system error, fraud or other unforeseen problem. Void where prohibited.

Shipping and Handling charges apply. Sales Tax is required for orders shipped to CA, CT, VA, and Canada. Prices and product availability are subject to change without notice.

ARRL *The national association for* ***AMATEUR RADIO***
225 Main Street, Newington, CT 06111-1494 USA

SHOP DIRECT *or call for a dealer near you.*
ONLINE WWW.ARRL.ORG/SHOP
ORDER TOLL-FREE 888/277-5289 (US)