KD6OZH uses an Analog Devices ADF4157 fractional N PLL IC and an Atmel ATtiny2313 microcontroller to build a VFO that produces a clean, stable 75 to 150 MHz output signal. He also describes a similar 390 to 425 MHz VFO.

QEX

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

John Stephensen, KD6OZH, built two versions of a clean stable VFO circuit that uses an Analog Devices ADF4157 fractional N PLL IC and an Atmel ATtiny2313 microcontroller. One version is a 390 to 425 MHz VHF/UHF VFO, and the other is a 75 to 150 MHz oscillator he built for a 3 to 16 MHz HF transceiver.

In This Issue

Features

Index of Advertisers

Larry Wolfgang, WR1B

Empirical Outlook

Old *QEX* Article Files

Have you ever carefully examined a spider web? Perhaps you have come across an elaborate web suspended between trees across a trail through the woods, as I have. Or perhaps you have watched a spider spinning its web among plants in a garden. The patterns can be truly amazing and beautiful, especially if you happen to spot the web early in the morning, with droplets of dew clinging to the threads and the early light hitting it in just that perfect way.

Sometimes there will be several threads stretched for 6 or 10 feet across a space, with the thread making an almost perfectly horizontal line. I always wonder how the spider managed to cross that space, and how it could stretch the thread so tight. From a few supporting threads, the spider is able to then weave an intricate pattern, often with concentric circles and carefully drawn radial lines all designed to catch its next meal of some unsuspecting insect that may fly into the trap and be caught.

After all that work, some careless animal (you or me?) may walk along that trail and not notice the web until it is a tangled mess of silk that we usually can't wait to rip off our face or clothes, all the while imagining the monster spider in attack mode, ready to bite whatever just destroyed all that hard work. Or maybe it was a sudden heavy rain that washed away the web. Spider webs are rather tenuous creations, in any case. They are built as a labor of love and as a matter of survival, perhaps in a few hours or a day, and can be gone in an instant. They may be rebuilt in the same location, or perhaps the spider will move on to a less vulnerable spot.

For many years I have known that the *QEX* files section of the ARRL website is very popular. Readers enjoy being able to download files related to the articles they have read in the pages of *QEX*. Those files include computer program files, source code for microprocessor projects, *EZNEC* antenna model files, Microsoft *Excel* spreadsheets and more.

Many authors maintain their own websites, where they post various articles and files related to those articles. There they can update a program file when some new feature is added or when a bug is discovered and fixed. They can post pictures of new versions of the project, and even of new projects. Some authors maintain websites that become reader favorites. Have you looked at **w7zoi.net** lately? Wes Hayward, W7ZOI, is an ever popular author (*Experimental Methods of Radio Frequency Design* along with many *QST* and *QEX* articles) with lots of interesting information on his website. How about Jack Smith, K8ZOA, with **www.cliftonlaboratories.com**? These are just two of many examples of websites that many *QEX* readers probably have bookmarked for regular visits.

Much like those spider webs, the World Wide Web can be a rather transient place. Sometimes new domain names are obtained and old names are allowed to fade away. Sometimes a sudden "storm" may cause a once popular site to lie dormant, no longer updated or just no longer available. Whatever the reason, individual websites can be the most subject to sudden closure, with the loss of the files and information formerly available there. That is one of the reasons we have encouraged authors to share their files with us to be posted to the *QEX* files section of the ARRL website. ARRL is in it for the long haul, and although authors may have updated files or fixed software bugs on their websites, we have a way for readers to obtain the information (at least as it may have existed when the article was published) for many years to come. After all, we know that someone may pick up an old issue of *QEX* and discover an article that interests them now, years after that issue was published.

During recent months I have had a significant number of requests for old files that were mentioned in *QEX* articles and posted to our QEXfiles website. This reminds me that even a website intended to be "secure" and available for the long haul can be lost as quickly as those delicate spider webs. You see, when the ARRL unveiled a new website as part of a completely new design, for some reason older files were not transferred to the new structure. I had been assured they would be, because we knew how popular that section of our website was, but suddenly the files were gone. Some authors indicated that readers had contacted them, and the author was able to supply the requested file, but that was not always the case.

Over the past few weeks I found (or perhaps re-discovered) what I had forgotten. At least some of the old files had been archived for ARRL HQ Staff access. I have begun what may be a rather long process of recreating the tables of file names and descriptions, author names and links to those old files. As I write this, I have added the 2006 files to our new website at **www.arrl.org/qexfiles**. In the coming weeks and months, I plan to continue restoring older files, as time permits. It won't happen immediately, but I will recreate as much of that broken web as I can.

If you are looking for an older file, please check the ARRL website to see if it has been restored. If you try to download a file and find that it is not the file you expected, or if you find other problems, please let me know. Since I don't see an automated way to do this, I will be copying files and text, and creating hyperlinks as those files are put into the new web structure. Our new web may not be exactly like the old one, but hopefully it will fulfill all our needs.

John B. Stephensen, KD6OZH

3064 E. Brown Ave, Fresno, CA 93703; kd6ozh@arrl.net

A Stable Low-Noise VFO for VHF-UHF Transceivers and Transverters

Most new VFO designs seem to use direct digital synthesis, but DDS designs suffer from increased spurs as the output extends into the VHF and UHF range. This modern PLL design eliminates that problem.

When I first became interested in Amateur Radio, variable frequency oscillators (VFOs) used large air-core coils and mechanically variable capacitors with 100:1 gear-reduction drives. These were easy to design and build but not very stable beyond 10 MHz. For SSB operation on the upper HF bands and all VHF and UHF bands, crystal-controlled converters were necessary. Later on, frequency synthesizers using multiple phase-locked loops (PLLs) allowed the construction of stable electrically tuned VFOs for SSB radios operating on HF, VHF or UHF bands. These were complicated devices, however. A 2 meter radio that I built in 1993 used a $3 \times 12 \times 17$ inch chassis to contain such a synthesizer. By 1999, I could combine a direct digital synthesizer (DDS) chip and a PLL chip, and achieve a similar result for an HF radio.[1] The DDS spurs increase as the PLL output increases into the VHF and UHF range, though.

There is now a much simpler option — the sigma-delta fractional-N PLL. The principles of its operation have been described previously in *QEX*.[2] In that case, a CPLD-based integer-N PLL and a microcontroller were combined to make a low frequency fractional-N PLL. A complete fractional-N PLL is now available in a 16-pin TSSOP package and operates at up to 6 GHz! The lead spacing is 0.65 mm (25.6 mils) so it can be soldered by hand. When combined with commercial off-the-shelf voltage controlled oscillators (VCOs), simple and stable synthesized VHF and UHF VFOs can be created.

[1]Notes appear on page 13.
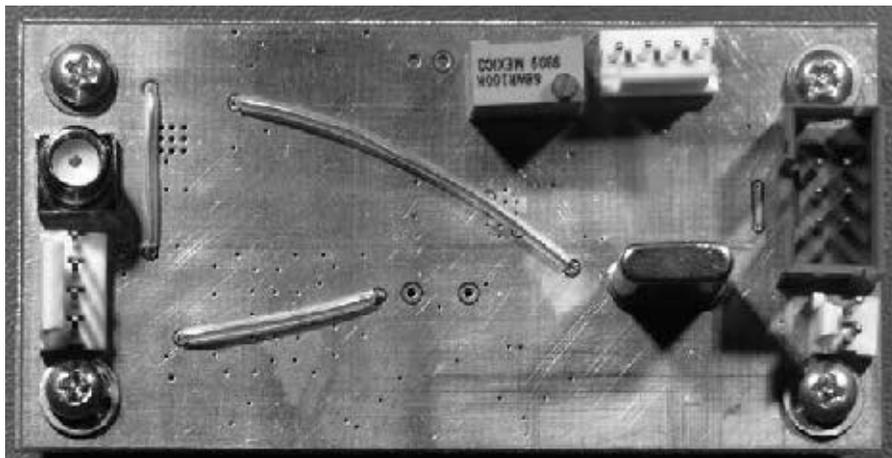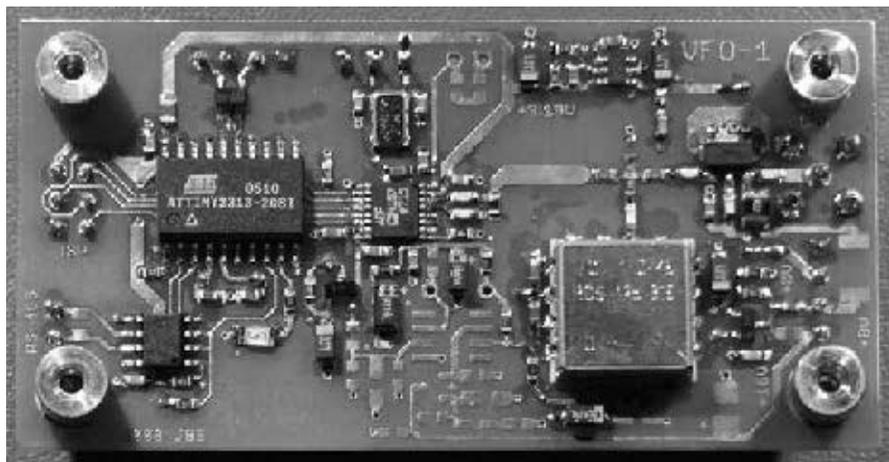
Figure 1 — Top and bottom views of the 390 to 425 MHz VFO (1.5 × 3 inch circuit board).

Figure 1 shows a synthesizer constructed for a 70 cm transceiver. The circuit board was designed so it can be configured for multiple types of VCOs, for use in different projects. It uses an inexpensive 8-bit microcontroller to convert ASCII frequency inputs arriving on the serial port into chip-specific commands sent to the PLL. The frequency reference is a 5×3.2 mm temperature compensated crystal oscillator (TCXO) that is accurate to 2.5 PPM over a –20 to +70°C range. The

Figure 2 — This photo shows a MiniCircuits Surface Mount VCO Module. Photo courtesy of Mini-Circuits (minicircuits.com).

UHF oscillator is an ROS-425-319 that tunes from 390 to 425 MHz with a 0 to 5 V input. An MMIC amplifier raises the output level to +16 dBm to drive an FET resistive mixer. This synthesizer is used in a single-conversion 70 cm radio with a 21.4 MHz IF, but could also be used in transverters with a 28 MHz IF.

VCO

A wide variety of VCOs covering the 50
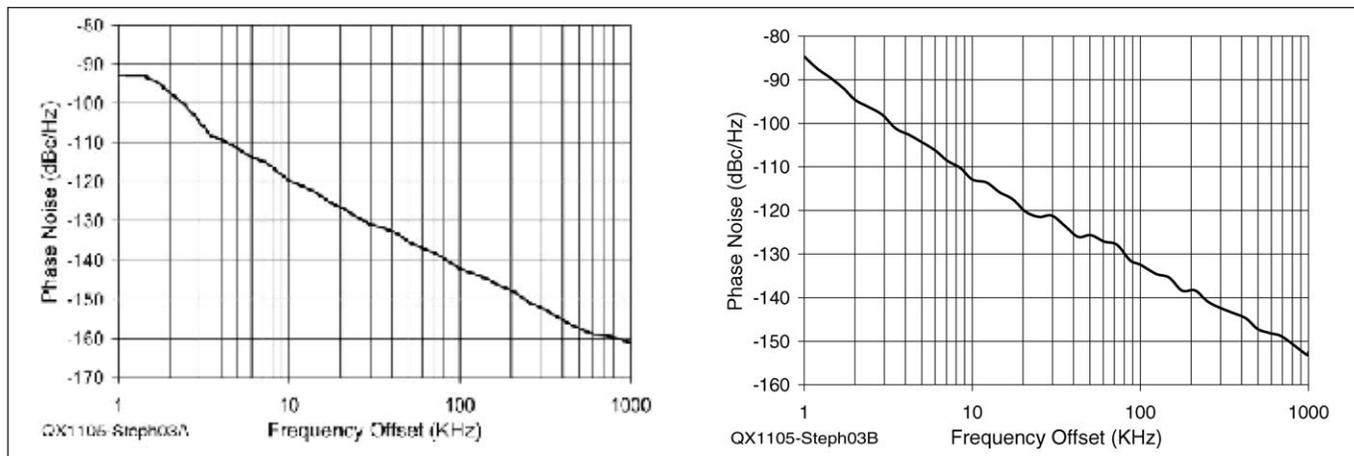
Figure 3 — Part A is a graph of the 390 to 425 MHz VCO phase noise and Part B is the 3060 to 3200 MHz VCO.
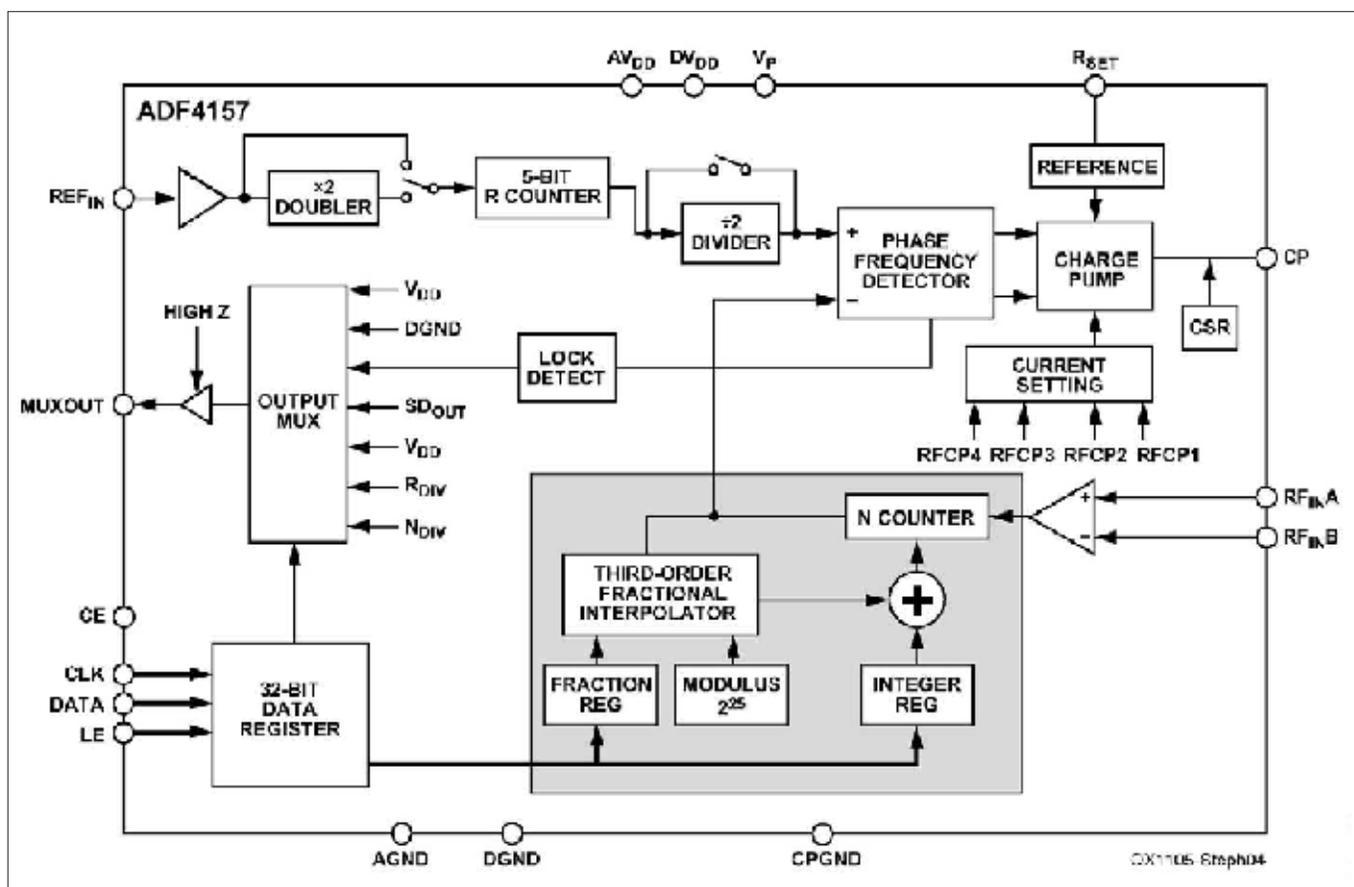
Figure 4 — This Sigma-Delta PLL Block Diagram is courtesy of Analog Devices (www.analog.com).
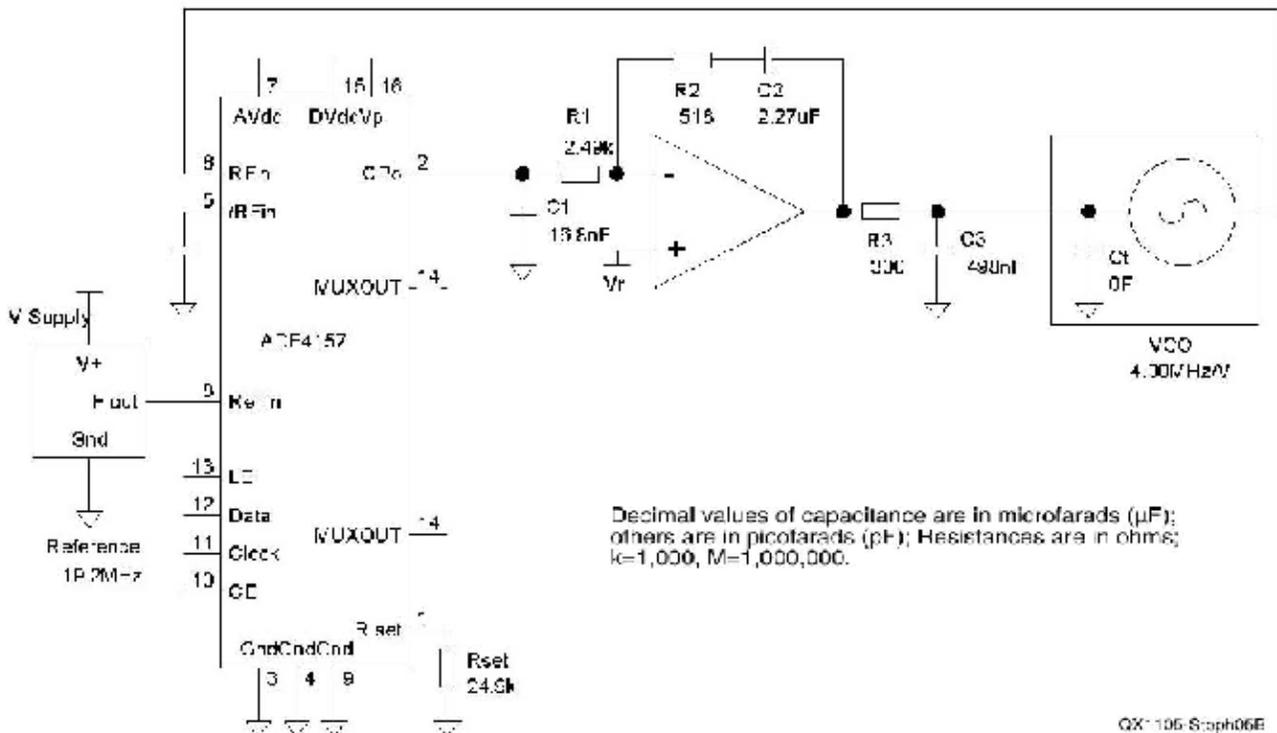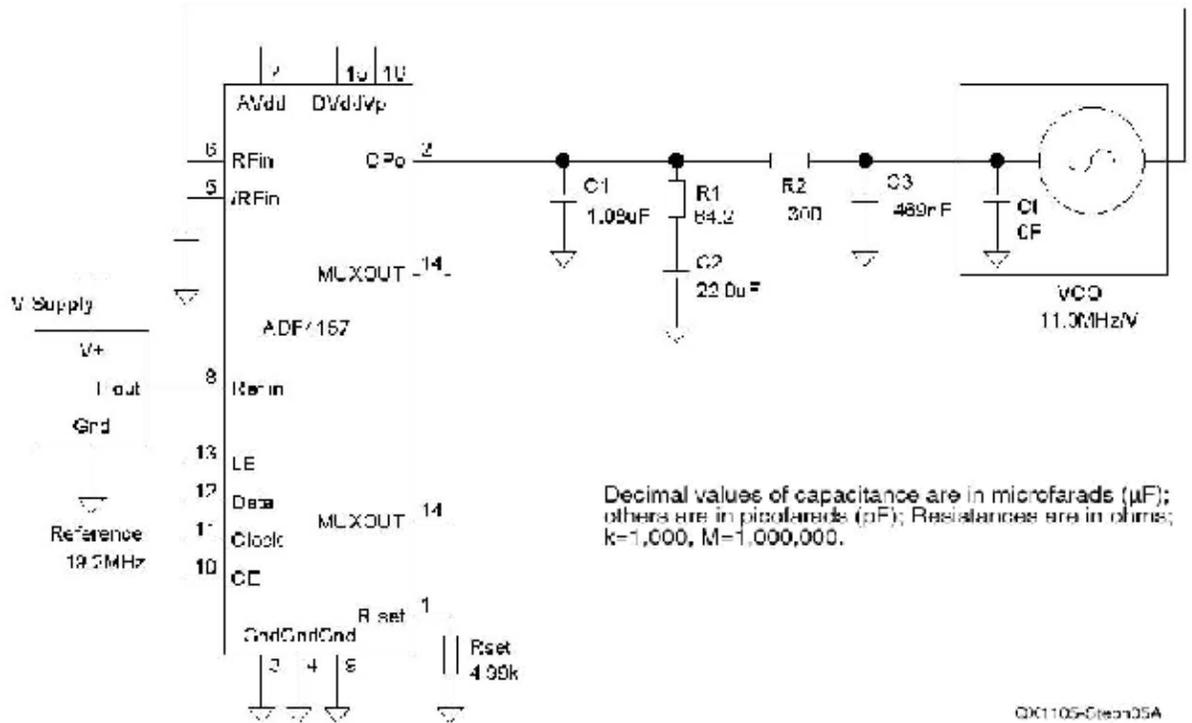
Figure 5 — Part A is the loop filter for a 0 to 5 V VCO and Part B is the loop filter circuit for a 0 to 15 V VCO.

to 6000 MHz frequency range are available from multiple manufacturers in the standard ½ × ½ inch package shown in Figure 2. They are powered by a 5 to 12 V dc supply, and the maximum tuning voltage ranges between 5 and 28 V. They are available on-line from distributors such as Mouser Electronics or directly from manufacturers such as Mini-Circuits.

The phase noise is adequate for narrowband modes, such as CW and SSB, through the 9 cm band, as shown in Figure 3. Phase noise is much lower than integrated VCO/PLLs. An ADF4350 operating at 412 MHz has 10 dB more noise at a 10 kHz offset and 20 dB more noise at a 100 kHz offset. At 3300 MHz, noise is 20 and 28 dB higher at 10 and 100 kHz offsets than the ROS-3200C-719 VCO module

PLL

The circuit board supports the Analog Devices ADF4100 series of PLLs in 16-pin TSSOP packages. These include integer-N and fractional-N devices, with upper frequency limits of 200 to 6000 MHz. The PLLs are controlled via a serial peripheral interface (SPI) port. This uses 3 pins (clock, data and latch enable) to serially transfer 32 bit data words into the PLL registers. The reference input connects to a frequency standard and the RF input to the VCO. Most have a 3 to 5 V charge pump output to control the VCO, although one integer-N model supports a 15 V charge pump at reduced phase detector frequencies.

The ADF4157 is the most interesting device for Amateur Radio use. It divides the VCO frequency by a 12 bit integer and 25 bit fraction, providing a tuning resolution of less than 0.06 Hz with a 2 MHz phase detector. This precision is obtained by rapidly changing the N counter between multiple values surrounding the desired fractional division ratio. The third-order interpolator, shown in Figure 4, is a sigma-delta modulator that shapes the resulting noise so that almost all is above the cutoff frequency of the loop filter, and does not appear on the VCO output.

Analog Devices provides free software (SimPLL) that may be used to design the frequency synthesizer. Given VCO parameters and desired loop bandwidth, it selects component values (Figure 5) and plots the expected phase noise and lock time (Figure 6). Only passive components are needed to control a VCO with 0 to 5 V tuning. For higher tuning voltages, an op amp integrator may be added to the circuit board to increase the maximum to as much as 30 V.

MCU

An Atmel ATtiny2313 microcontroller, shown in Figure 7, was chosen, as it is the least expensive part with a universal asynchronous receiver-transmitter (UART). This series also has 32 registers, so high precision multiply and divide routines can easily be implemented to control the PLL.

Software loaded into the microcontroller's flash memory calculates the correct values to load into each of the PLL registers, shown in Figure 8. The EEPROM contains constants that vary with the VCO type and frequency, such as the reference divider value, prescaler modulus, reference frequency and charge pump current. Flash memory is loaded via the ISP connector on the circuit board. The EEPROM is programmed by commands received on the serial port and an LED provides confirmation of each command. The command set is shown in Figure 9.

The integer and fractional parts of N loaded into R0 and R1 are calculated as follows:

$$N = f \times R / S$$

The output frequency is f, R is the reference divider value and S is the TCXO or external reference frequency. Both f and S are 32 bit integers and R is an 8 bit integer. N consists of a 16 bit integer quotient and a 32 bit fractional quotient. In order to minimize any spurs, the code alters the reference divider by subtracting one when N is within 0.5% of integer boundaries. N is then recalculated for the new phase detector frequency. Source code is available for download from the ARRL *QEX* website.[3]

Schematic Diagrams

The schematic diagrams of the two boards are shown in Figures 10, 11 and 12. Figure 10 is the microcontroller unit (MCU). U14 is a low-noise voltage regulator that supplies the MCU, PLL, TCXO and RS-485 transceiver. All can operate at 3 V ±10% except the RS-485 transceiver, which requires 3.3 V ±10%. A 7.3728 MHz crystal sets the microcontroller clock frequency at four times the UART sampling rate for the half-duplex 115,200 Baud serial port. The optional transistor array (Q3) is provided to
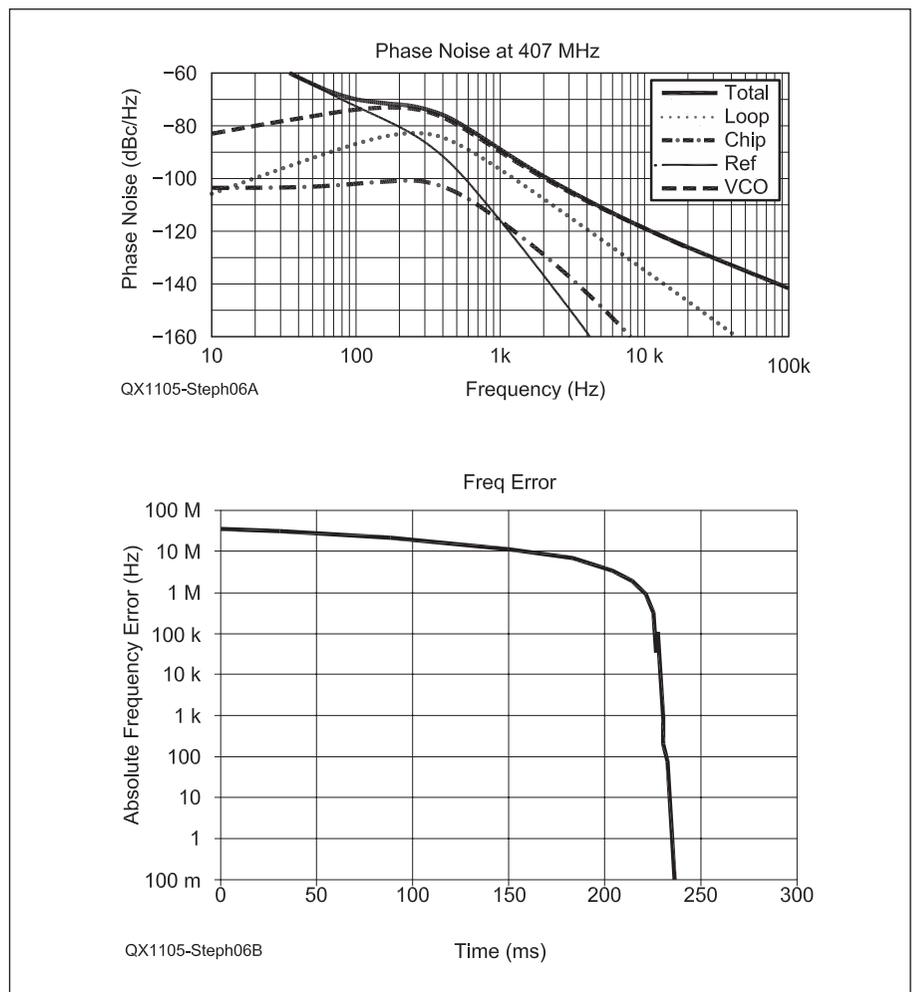
Figure 6 — Part A is a SimPLL frequency domain plot and Part B is a time domain plot.
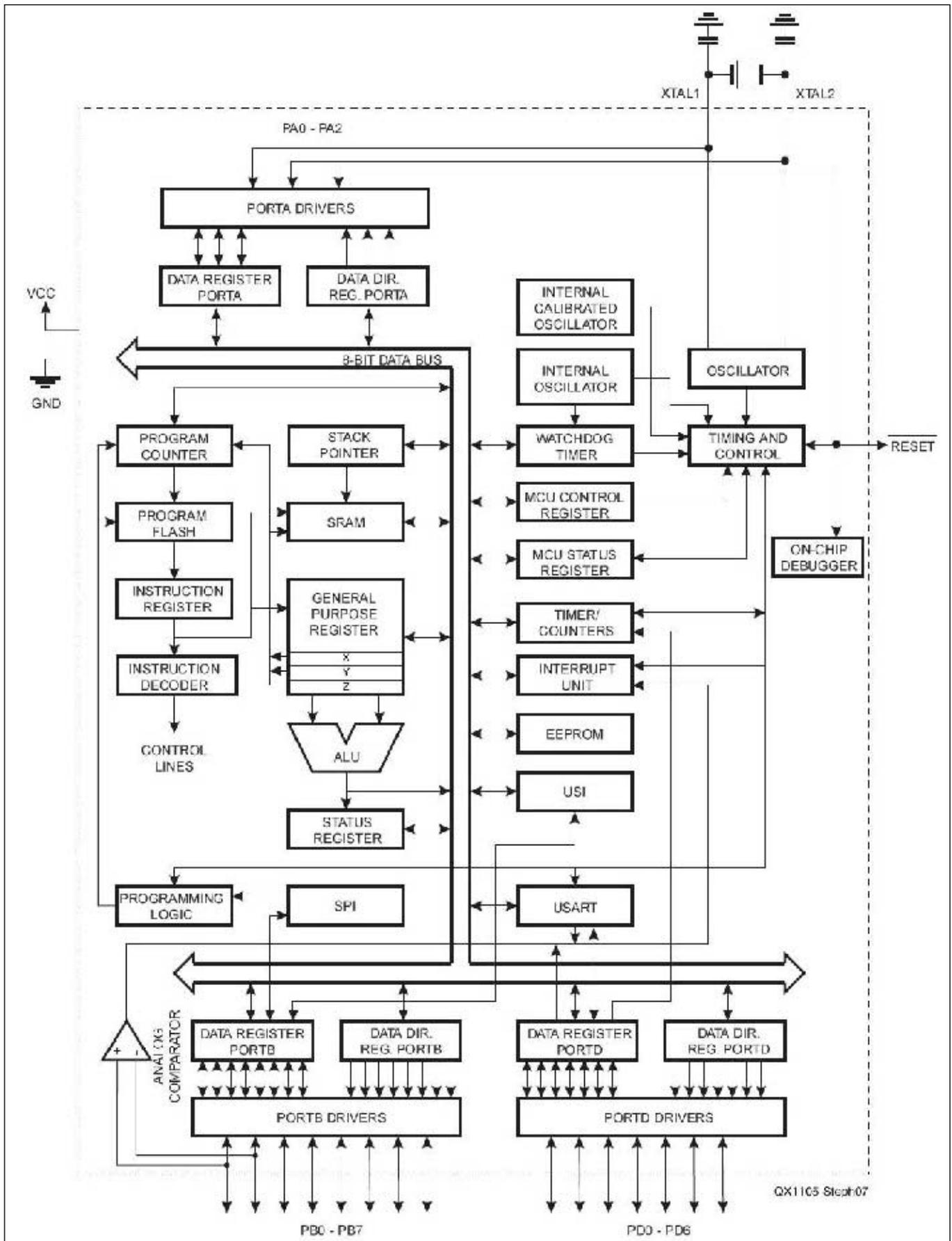
Figure 7 — This is the Atmet ATtiny2313 microcontroller block diagram

FRAC/INT REGISTER (R0)

QX1105-Steph08

RESERVED	MUXOUT CONTROL				12-BIT INTEGER VALUE (INT)												12-BIT MSB FRACTIONAL VALUE (FRAC)													CONTROL BITS		
DB31	DB30	DB29	DB28	DB27	DB26	DB25	DB24	DB23	DB22	DB21	DB20	DB19	DB18	DB17	DB16	DB15	DB14	DB13	DB12	DB11	DB10	DB9	DB8	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	M4	M3	M2	M1	N12	N11	N10	N9	N8	N7	N6	N5	N4	N3	N2	N1	F25	F24	F23	F22	F21	F20	F19	F18	F17	F16	F15	F14	C3(0)	C2(0)	C1(0)

LSB FRAC REGISTER (R1)

RESERVED				13-BIT LSB FRACTIONAL VALUE (FRAC) (DBB)													RESERVED												CONTROL BITS		
DB31	DB30	DB29	DB28	DB27	DB26	DB25	DB24	DB23	DB22	DB21	DB20	DB19	DB18	DB17	DB16	DB15	DB14	DB13	DB12	DB11	DB10	DB9	DB8	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	F13	F12	F11	F10	F9	F8	F7	F6	F5	F4	F3	F2	F1	0	0	0	0	0	0	0	0	0	0	0	0	C3(0)	C2(0)	C1(1)

R DIVIDER REGISTER (R2)

RESERVED	RESERVED	CSR EN	CURRENT SETTING (DBB)				RESERVED	PRESCALER	RDIV2 DBB	REFERENCE DOUBLER DBB	5-BIT R COUNTER (DBB)					RESERVED												CONTROL BITS		
DB31	DB30	DB29	DB28	DB27	DB26	DB25	DB24	DB23	DB22	DB21	DB20	DB19	DB18	DB17	DB16	DB15	DB14	DB13	DB12	DB11	DB10	DB9	DB8	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	C1	CPI4	CPI3	CPI2	CPI1	0	P1	U2	U1	R5	R4	R3	R2	R1	0	0	0	0	0	0	0	0	0	0	0	0	C3(0)	C2(1)	C1(0)

FUNCTION REGISTER (R3)

RESERVED																	SD RESET	RESERVED						LDP	PD POLARITY	PD	CP THREE-STATE	COUNTER RESET	CONTROL BITS		
DB31	DB30	DB29	DB28	DB27	DB26	DB25	DB24	DB23	DB22	DB21	DB20	DB19	DB18	DB17	DB16	DB15	DB14	DB13	DB12	DB11	DB10	DB9	DB8	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	U12	0	0	0	0	0	0	U11	U10	U9	U8	U7	C3(0)	C2(1)	C1(1)

TEST REGISTER (R4)

RESERVED							NEG BLEED CURRENT		RESERVED		CLK DIV MODE		12-BIT CLOCK DIVIDER VALUE												RESERVED				CONTROL BITS		
DB31	DB30	DB29	DB28	DB27	DB26	DB25	DB24	DB23	DB22	DB21	DB20	DB19	DB18	DB17	DB16	DB15	DB14	DB13	DB12	DB11	DB10	DB9	DB8	DB7	DB6	DB5	DB4	DB3	DB2	DB1	DB0
0	0	0	0	0	0	0	NB2	NB1	0	0	C2	C1	D12	D11	D10	D9	D8	D7	D6	D5	D4	D3	D2	D1	0	0	0	0	C3(1)	C2(0)	C1(0)

Figure 8 — Sigma-Delta PLL registers

Command	Syntax
Default frequency	D<0-3999999999>
Change frequency	F<0-3999999999>
Set charge pump current	C<0-15>
Invert charge pump output	I<0-1>
Prescaler configuration	P<0-7>
Set reference divider	R<0-31>
Frequency Standard	S<0-199999999>

Note: Commands terminated with carriage return or new line.

Figure 9 — Software control command set.
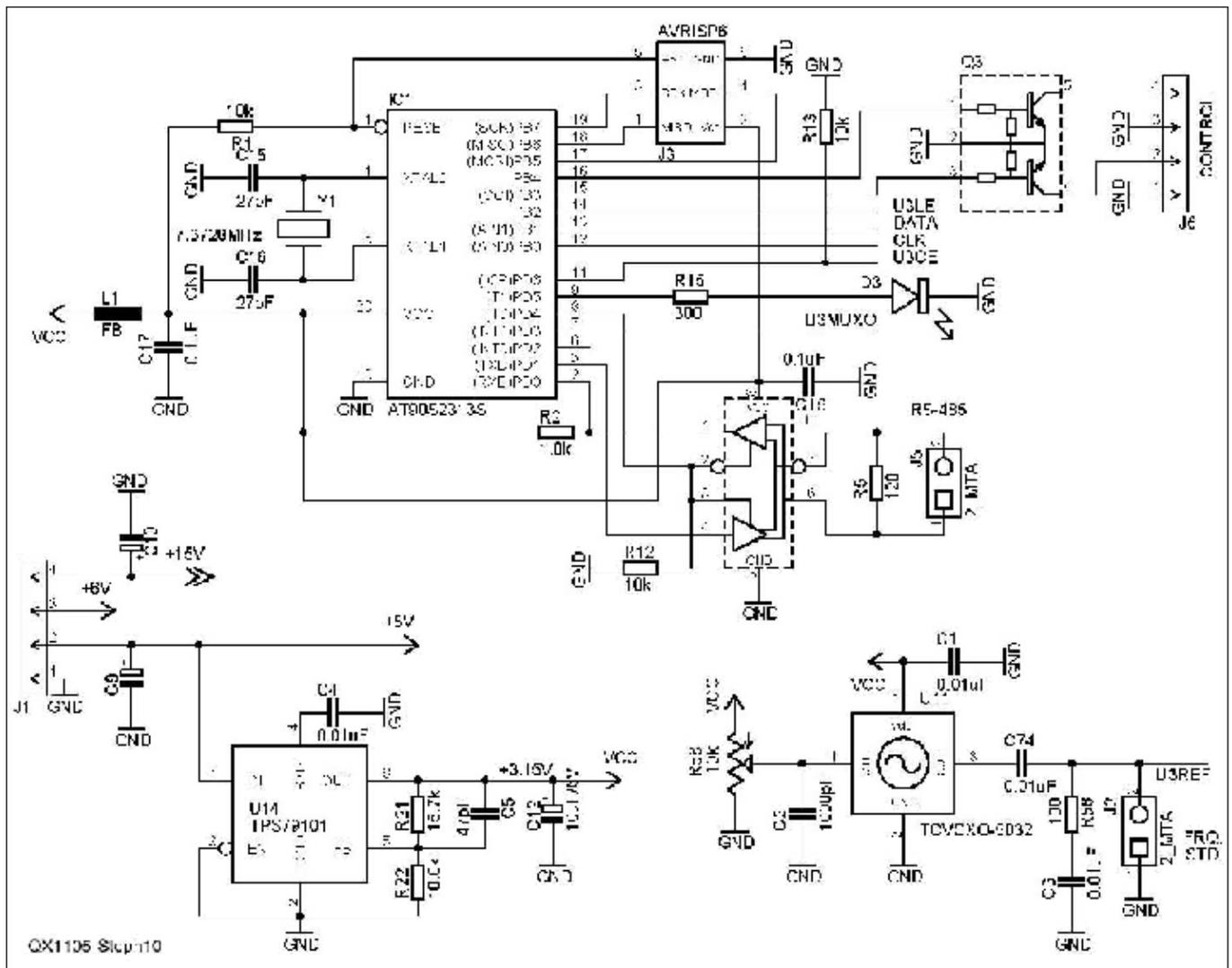
Figure 10 — This schematic diagram shows the microcontroller, the main voltage regulator and the TCXO.

support two high-voltage open-collector outputs. The TCXO (U11) is also optional and may be replaced by J2, C3 and R56 when using an external frequency reference. J3 is used to load firmware into the MCU.

Figures 11 and 12 show the RF portions of the low and high frequency circuit boards, respectively. The difference between the two is the inclusion of an LVDS buffer (U4) on the low frequency version. This squares the VCO output in order to meet the slew rate requirements of the PLL. Three MMBT5089 transistors (Q1, Q2 and Q4) are used as low-noise regulators for the op amp, VCO and PLL charge pump. The VCO regulator can be connected to the 8 or 15 V supply. The values of R4 and R49 may be altered to support VCOs requiring more than 5 V. The op amp and its regulator are installed if the tuning voltage is higher than 5 V. Component values for the loop filter (R6, R7, R10, R28,

R29, D9, C6, C14, C38 and C39) are selected to match the VCO used. R30, R31, R32 and R34 split the VCO output between the PLL input and the output buffer amplifier (U9) and may be altered to suit the output level of the VCO. The values shown provide 9.5 dB of attenuation at each output port. In order to support a wide range of MMIC amplifiers, the output amplifier may be connected to the 5 V supply via a jumper or the 8 V supply via R17.

UHF Synthesizer Performance

The outputs of two 390 to 425 MHz synthesizer circuit boards were mixed together as shown in Figure 13 in order to measure the phase noise. Since the output of the double balanced diode mixer is the sum of both synthesizers, the noise level is 3 dB higher than each individual synthesizer. The output was measured on a spectrum analyzer (Figure

14) using a 2 MHz offset and on a PC sound card with FFT software using a 20 kHz offset (Figure 15).

Phase noise levels below a 1 kHz offset are somewhat higher than predicted by SimPLL but less than that required for SSB and CW operation. This is probably due to the fact that SimPLL uses preliminary data for the ADF4157. Phase noise is approximately –68 dBc/Hz at a ±300 Hz offset, –81 dBc/Hz at a ±1 kHz offset and –106 dBc/Hz at a ±3 kHz offset. Phase noise is identical to a free-running VCO as the frequency offset increases.

VHF Synthesizer Performance

A VHF synthesizer was constructed for a 3 to 16 MHz HF transceiver. This uses a 75 to 150 MHz VCO as shown in Figure 16. The output is dived by 4 on the mixer circuit board. Note that an LVDS buffer is inserted
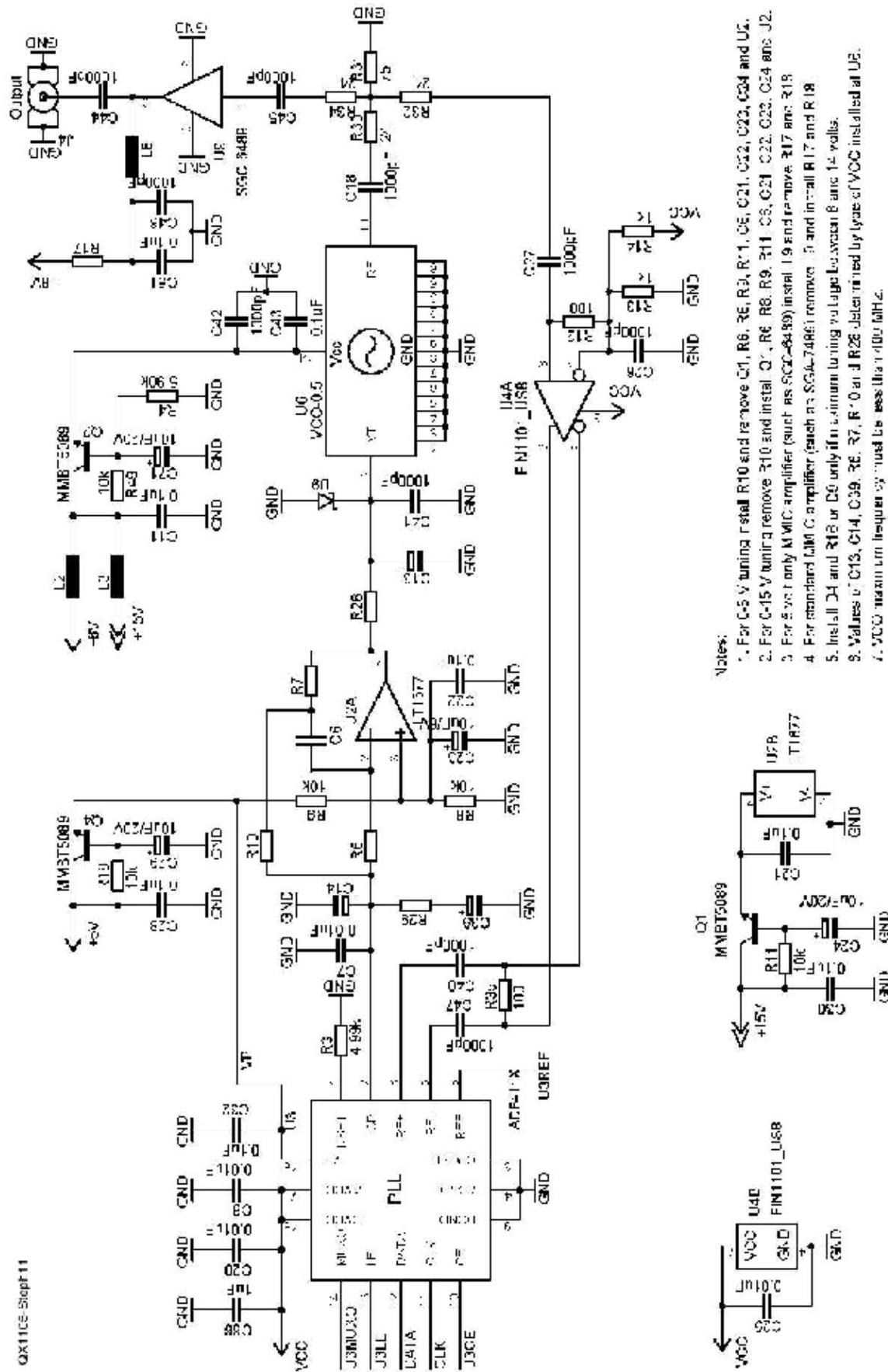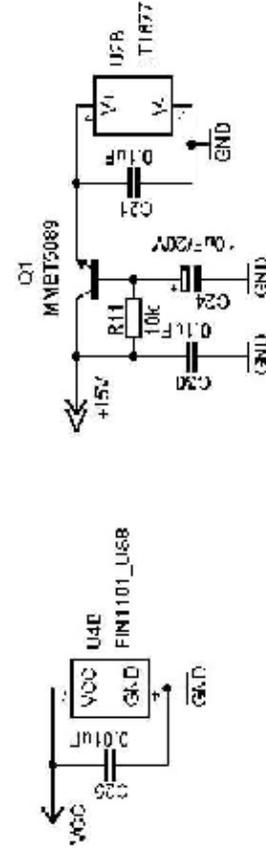
Figure 11 — Here is the schematic diagram of the VHF VFO.

Figure 12 — This schematic diagram shows the UHF VFO.
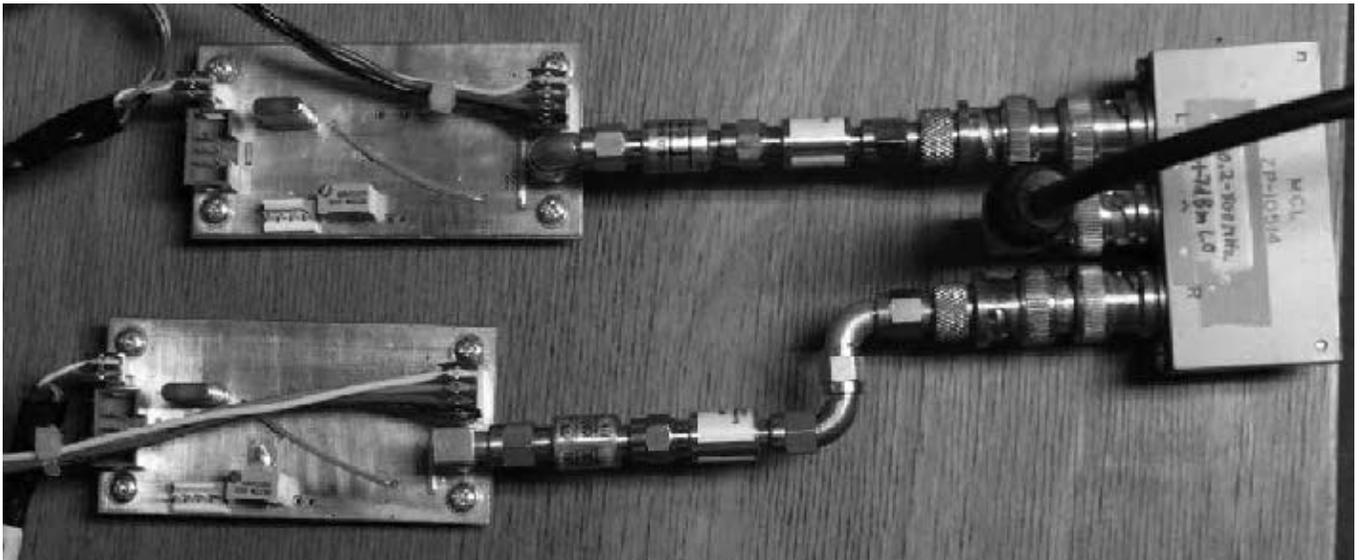
Figure 13 — This is a photo of the phase noise test setup. Outputs from two 390 to 425 MHz synthesizers were sent to a double balanced diode mixer, and the output was measured using a spectrum analyzer.
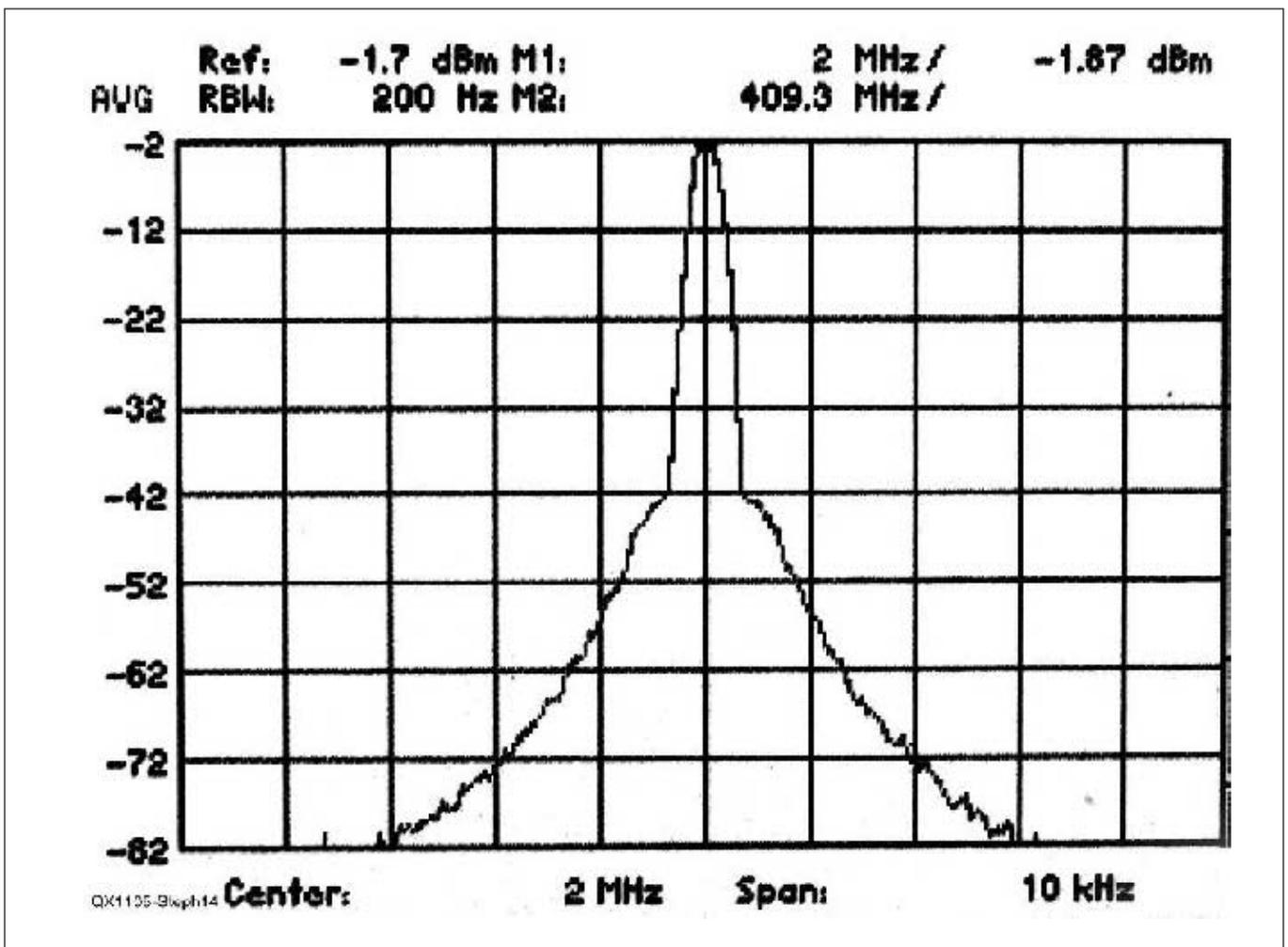
Figure 14 — This graph shows the spectrum analyzer phase noise measurement for the two 390 to 425 MHz synthesizers.
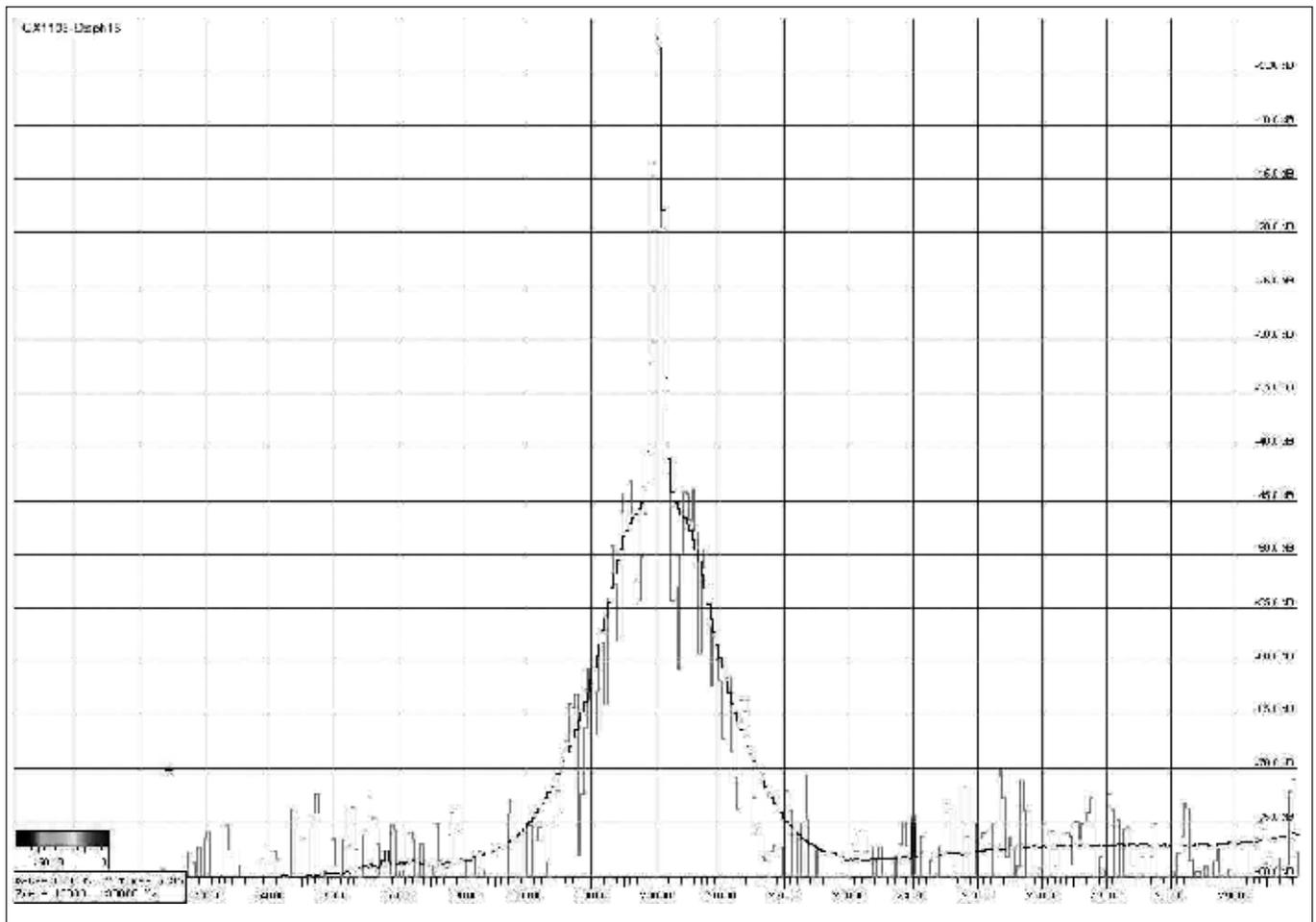
Figure 15 — A computer soundcard and DSP software were used to obtain this FFT display of the phase noise output from two 390 to 425 MHz synthesizers. The noise bandwidth was 60 Hz.

between the VCO output and PLL RF input to increase the slew rate of the signal. A low-noise rail-to-rail op amp is also added to increase the VCO tuning voltage to the range of 0 to 15 V. The large orange object is a 2.2 μF capacitor used to integrate the phase detector output.

The AF output of the receiver was measured with a 14.318 MHz low-noise crystal oscillator providing an RF input at the maximum signal level, and the output is shown in Figure 17. The light gray trace is the instantaneous FFT and the bold trace is the long-term average. Phase noise is approximately –72 dBc/Hz at a ±50 Hz offset. Note that the audio path is 16 bits with the AGC set point at –12 dBFS so the dynamic range of this measurement is only 84 dB.

John Stephensen, KD6OZH, became interested in radio at age 11, when his grandfather bought him a crystal radio kit. During the 1960s, he built several HF receivers using vacuum tubes and other parts procured from discarded black-and-white TV sets. After attending the University of California, he and two friends founded PolyMorphic Systems — a supplier of personal computer kits, and later manufactured computers — in 1975. In 1985, he was cofounder of Retix, a networking hardware supplier. John earned his Amateur Radio license in 1991, and has been active on bands from 7 MHz to 24 GHz, with interests including HF and microwave DXing and contesting. He has also been active on packet, satellites and on the HF bands using several digital modes. John has always designed and built his own Amateur Radio gear, some of which has been described in QEX.

Notes

[1]John B. Stephensen, KD6OZH, "The ATR-2000: A Homemade, High Performance HF Transceiver," Mar/Apr 2000 *QEX*, pp 3-15 and May/Jun 2000 *QEX*, pp 39-51.

[2]Oleg. Skydan, UR3IQO, "An All Digital Fractional-N Synthesizer", Nov/Dec 2003 *QEX*, pp 25-33.

[3]The source code file for this project is available for download from the ARRL *QEX* files website: **www.arrl.org/qexfiles**. Look for the file **5x11_Stephensen.zip**.

Figure 16 (right)— Top and bottom views of the 75 to 150 MHz synthesizer circuit board.
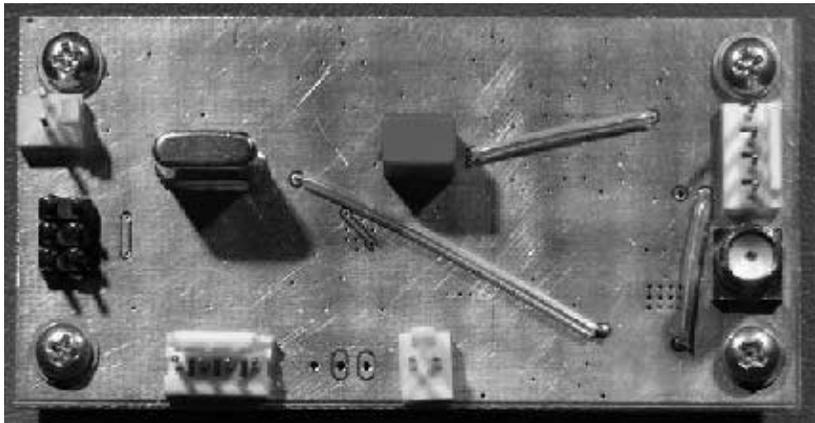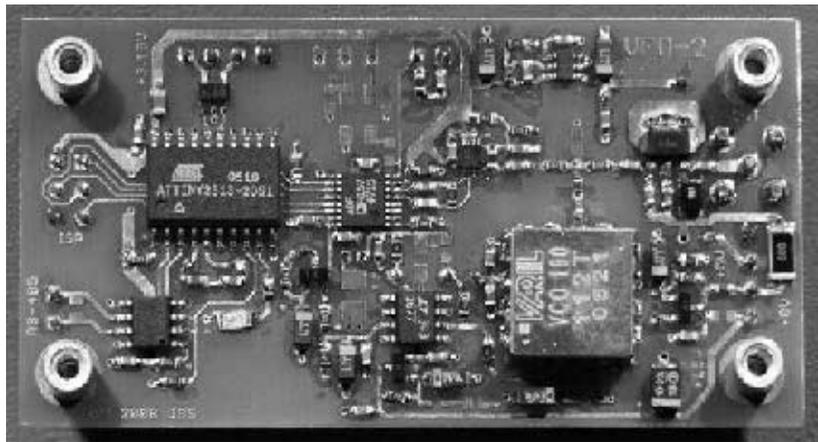
Figure 17 (below) — The HF receiver audio output was measured using a computer sound card and DSP software to plot this FFT graph, with 1.5 Hz noise bandwidth. The light gray trace represents the instantaneous FFT, while the bold trace represents the long-term average calculation.

Tom McDermott, N5EG

3950 Southview Ter, Medford, OR 97504; n5eg@tapr.org
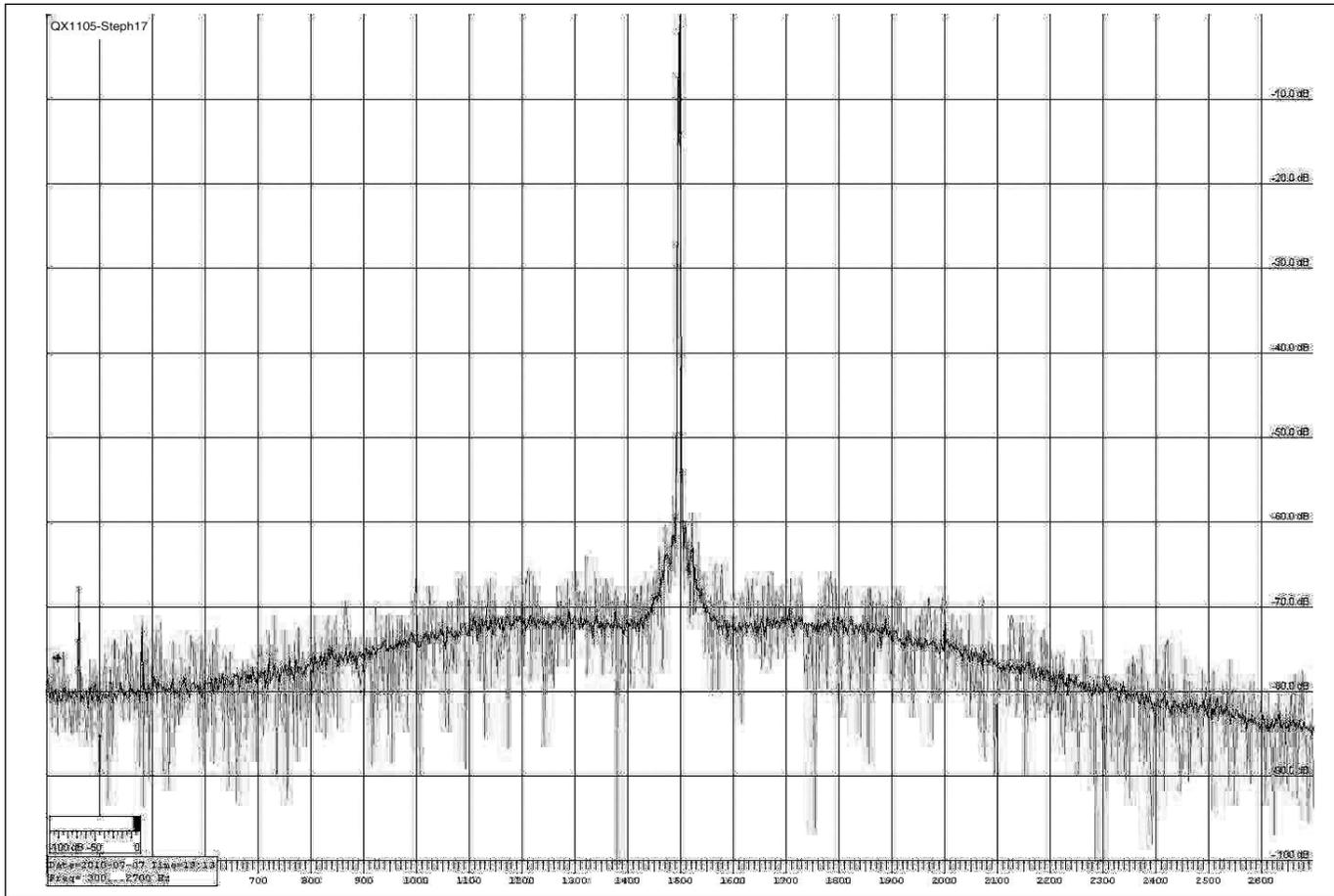
A Flexible 2-Port Network Calculator Tool

Computer program analyzes complex circuits using lab measurements of component networks.

Circuit simulation (such as SPICE) is used extensively in the design and optimization of analog circuits. For the large majority of circuit analysis problems it's an effective technique. However there are a few common RF analysis problems that are not easily addressed with circuit simulation. Another general technique of analysis has been developed based on 2-port networks. A 2-port is a network with two ports – input and output, each of which is usually ground-referenced. Commonly the 2-port will be drawn with two input connections (port + ground) and 2 output connections (port + ground). These 2-ports are commonly used for microwave circuit analysis, but the actual techniques are easily applied as well to HF and VHF/UHF RF networks commonly used by amateurs.

A modern Vector Network analyzer (VNA) can measure networks and extract the measured results represented as a 2-port network. Most commonly these are the S-parameters of a network, but it is also possible to extract the network transfer function as Z-parameters, Y-parameters and others. The VNA can measure the network response over a range of frequencies, and the response of the network is usually a set of the 4 parameters at each discrete measurement frequency. It is sometimes difficult to replace the measured response of the network with a simple circuit model that matches the response well over the measured frequency range. Instead the table of measured parameter values is many times the only practical description for the network.

Circuit simulators such as SPICE are difficult to use when we have empirical measured data tables that are frequency-dependent; in that case performing 2-port network computations may be easier and more indicative of the actual network values.

We sometimes know of stray circuit elements that are unavoidably present between our test equipment and the networks we are trying to measure. In this case it may be possible to subtract out the influence of the corrupting stray circuits mathematically, and this is especially so if the stray can be simply described (that is, it's not too complicated, and it does not bridge the network under test). While it is impossible to construct a physical network 2-port with negative component values, mathematically it's easy and actually fairly intuitive. Of course any such fictional compensation networks need to resemble as closely as possible the actual stray values in the circuit (except the sign), otherwise they will incorrectly compensate for the stray. If not, the compensation might be correct at one frequency but wrong at other frequencies. Interestingly, we can also use a model of a transmission line with negative length to provide compensation of both cable loss and phase delay.

On occasion we are able the measure the total response of a network, but we only know the values for some of the 2-ports that make up the network. It is possible to derive the response of the unknown 2-port block (or of the collection of unknown blocks assuming they are contiguous) by subtracting the known blocks from the overall measurement to leave just the unknown block.

These types of circuit analysis are difficult to do in circuit simulators such as SPICE, which may not converge well, may not handle negative component values well (or at all), are awkward to use with tabular parameter data, or don't compute the response of an unknown element easily.

Finally, we may want to specify a set of boundary conditions on 2-port networks (for example, the ratio and phase of drive currents to Z-parameter network). While this can be done in SPICE using a number of current sources, it is still difficult to closely model these conditions when the network itself is frequency-dependent (thus affecting the network port drive impedances) in a way that is modeled by simple circuit elements. In this case, 2-port analysis provides us some quite useful results. This case is particularly important when trying to model phased antenna arrays where there is a considerable amount of mutual impedance between the elements, and that mutual impedance changes with frequency.

Thus 2-port analysis can be a useful tool in a toolkit along with circuit simulation and other computation tools (such as a spreadsheet). This article will focus on 2-port modeling, and a flexible software tool written in Microsoft C# to make it fast, easy, and intuitive to do; then illustrate with three examples.

Definition of common 2-port Networks

Two-port networks have been described and used for many years. Some references that are or may still be available are Pozar[1], Matthaei, Young, Jones[2], and Gehrke[3]. The most commonly used 2-port networks are the S, Z, and Y networks, and the ABCD matrix representation. Each 2-port has a different set of port termination conditions used for defining and measuring each parameter. Each type can be mathematically converted to any of the other forms, but

[1]Notes appear on page 25.

the ABCD matrix turns out to be very easy to use computationally. In the included C# tool, each kind of 2-port network *Tile* (an object that the software uses to represent a single 2-port network) can have its ABCD matrix extracted. That extracted ABCD matrix is what is used for all calculations.

A 2-port impedance (Z) matrix relates the voltages measured at the ports of a 2-port network in response to the currents injected at the ports. Mathematically, we can describe the resultant voltages at ports 1 and 2 to the currents injected at ports 1 and 2 as:

$$V_1 = Z_{11}I_1 + Z_{12}I_2$$
$$V_2 = Z_{21}I_1 + Z_{22}I_2$$

Or, using matrix notation:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} * \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad \text{or} \quad [V] = [Z] * [I]$$

We measure each term in the Z matrix by first injecting a current into port 1 and open-circuiting port 2 (zero current), then measuring the voltage at port 1 (V1/I1 = Z11) and port 2 (V2/I1 = Z12). Then we reverse the network, injecting current at port 2 and open-circuiting port 1, and measure the voltage at port 1 (V1/I2 = Z21) and at port 2 (V2/I2 = Z22). With phased antenna arrays we many times are concerned with the element currents so an impedance matrix describing the antenna array is commonly encountered and measured using this method, or other methods. Similarly, the admittance matrix relates the resultant currents at each port to the voltages injected at the other ports:

$$[I] = [Y] * [V]$$

S-parameters are defined when all ports are terminated in a matched impedance, for example 50 Ω. The S-parameters relate the resultant voltage reflected (V–) out a port compared to the voltage injected (V+) into the various ports.

$$\begin{bmatrix} V_{1-} \\ V_{2-} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} * \begin{bmatrix} V_{1+} \\ V_{2+} \end{bmatrix}$$

ABCD Matrix description

An ABCD matrix describes the input voltage and current in terms of the output voltage and current of a 2-port network. This is backwards of the way that we normally think about circuit analysis — where we start at a generator and work our way towards the load. In analyzing ABCD matrices, we start at the load, and work our way towards the generator, asking essentially the question: "what voltage and current need to be supplied by the generator at the input to a 2-port network in order to cause a known voltage and current at the output of the 2-port network?" The terminology ABCD matrix is derived from the arbitrarily-chosen names of the 4 coefficients in the matrix. The two equations that describe the input node (V1, I1) in terms of the output node (V2, I2) are:

$$V_1 = A * V_2 + B * I_2$$
$$I_1 = C * V_2 + D * I_2$$

Where all the variables are complex numbers. It's convenient to represent this in matrix form as:

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} * \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

In the software tool the {V, I} pair representing a node is called a *NodeSet*, and the matrix holding the parameters of a two-port is a *ParameterSet*. The ParameterSet contains the 4 complex numbers of the ABCD representation of a Tile, at one single specific frequency. Then NodeSet(in) = ParameterSet * NodeSet(out). By using a matrix and vector to hold the values, and by defining a multiply operator, this reduces the description of the computation in the source code to a single multiply operation – just like the expression above because the software has defined matrix and complex operations for the *ParameterSet* and *Complex* object types.

An example derivation of a simple component network shows how to describe a 2-port via an ABCD matrix. Assume that the network is a simple series impedance between port 1 and port 2, with no impedance to ground. The voltages and currents are shown in Figure 1. The definition of direction of the currents shown in this figure allows direct chaining of the computations. The output current must be the same as the input current, and the input voltage must be equal to the output voltage plus the I*Z drop across the series impedance.

By inspection, the term 'C' must be zero and 'D' must be one so that the input and output currents are the same. The term 'A' must be one and the term 'B' must be equal to the series impedance Z. This yields the input node pair values as:

$$V_1 = 1 * V_2 + Z * I_2 \qquad\qquad V_1 = V_2 + Z * I_2$$
$$I_1 = 0 * V_2 + 1 * I_2 \quad \text{which simplifies to} \quad I_1 = I_2$$

So the ABCD matrix for a series impedance is:

$$\begin{bmatrix} 1 & Z \\ 0 & 1 \end{bmatrix}$$

Z can be composed of a combination of series or parallel components (resistors, inductors, and capacitors), or it could be the impedance presented by a transmission line. Thus Z generally is frequency-dependent, and we need to run the network chaining computation at each frequency of interest.

For a shunt network, the voltage on the input and output terminals are the same, but the input current is the sum of the shunt current and the current out of the network. Letting Z be the shunt impedance, then Y is the shunt admittance, equal to 1/Z. This yields the input node pair values as:

$$V_1 = 1 * V_2 + 0 * I_2 \qquad\qquad V_1 = V_2$$
$$I_1 = (1 / Z) * V_2 + 1 * I_2 \quad \text{which simplifies to} \quad I_1 = Y * V_2 + I_2$$

So the ABCD matrix for a shunt impedance is:

$$\begin{bmatrix} 1 & 0 \\ Y & 1 \end{bmatrix}$$

Using an ABCD Matrix

Essentially, the ABCD matrix tells us that if we know the output *NodeState* (the voltage and current at the output of the tile), we can derive what the voltage and current at the input of the tile would have to be in order for that known tile output voltage and current to be what they are. By computing and storing the node state at each node, we can quickly select the display of impedance, return loss, or other node parameters at any node, and similarly we can compute the voltage or power transfer between any pair of nodes by walking the computation from the output of the last tile to the input of the first tile.

The various types of 2-port networks can be converted back and forth between forms. From a computational perspective, it's easy to use the ABCD format for chaining the calculations across multiple 2-port devices. While it is possible to chain other format calculations (such as using the chaining rules for S-parameters) the computation is straightforward when using an ABCD matrix.

When using ABCD matrices, the total network response is simply

the cascade of the various 2-port networks, achieved by progressive matrix multiplication of all the 2-port networks starting at the last 2-port. For the C# tool described in this article, the approach chosen was to save the computations for all of the nodes in the network. The { voltage, current } pair are computed at each node and saved, and the input nodestate of the previous computation becomes the output nodestate of the next. This is accomplished by matrix multiplication of the network ABCD matrix times the output node voltage, current, resulting in the input node voltage, current, and then iterating to the front of the network (saving all the node pairs along the way).

The software defines an object known as a *ParameterSet* which contains the frequency of the dataset, and the 4 complex parameters of a tile at that one frequency. The ABCD parameters are extracted from the *Tile* depending on what kind of 2-port network the set is describing. The tile can optionally contain a collection (in a list) of *ParameterSets*, one per frequency, or it can be described at just one frequency if the tile is easily described analytically and the frequency-dependent behavior can be easily computed (such as a resistor, capacitor, inductor, or combinations of them). Additionally a *NodeState* object contains the complex voltage and currents at a node.

The calculation computes a list of NodeStates, one for each node in the network (a network has one more node than there are tiles). The ParameterSet multiplication operator is defined so that a 2x2 matrix times a 2x1 vector (ABCD matrix times NodeState) or 2x2 matrix times a 2x2 matrix (ABCD matrix times ABCD matrix) multiplication is performed depending on the types of the objects being multiplied. This makes the ABCD chaining calculations look simple in the source code while in reality a significant quantity of matrix multiplication and complex number computations are happening as a result of that single multiply instruction.

Computing Network Response

To compute the network response (a cascade of tiles) we start at the last tile in the network, and set the output voltage and current (the nodestate after the last tile) to {V=1+j0, I=0+j0}. This means that the output voltage is 1 volt at a phase angle of zero degrees, and the output current is zero — essentially we terminate the last tile into an open circuit. Then we work our way back towards the source (at node zero) through all the tiles computing each node voltage and current pair as we go, and saving them as a list of NodeStates. At any point, the voltage transfer function is simply the last node voltage (1+j0) divided by the node voltage at the present point. We can compute the impedance at any node as V node / I node.

The power transfer can be different than the voltage transfer, because the impedance at each point in the network can be different. When dealing with complex voltages and currents, we need to take into account the phase angle of each. The power delivered into a load is:

$$P = \frac{1}{2} Re\left(V * I^*\right)$$

The real part of the product of the voltage V and the complex conjugate of the current I is the power into the load. The factor of one-half arises because V and I represent the peak voltages of a sinusoidal wave since they are complex vectors. We can compare the power at any point in the network to any other point — except the last node. The last node (which is after the last tile) is an open circuit, and no power can be delivered to an open circuit (because the current is zero). The C# tool will signal an error if we try to compute the power transfer from some node to the last node, since it is not possible to deliver any power into that open circuit.

Terminating Tile

In fact, the last tile is somewhat special in the network analysis. Many times we will make an S-parameter measurement of a complicated load (such as an antenna). The resulting S-parameter matrix including the antenna does not have a defined set of output terminals (it's really a one-port measurement), and thus no port-to-port transfer function exists. S12, S21, and S22 are many times just set to zero in the resulting measured S-parameter file by the VNA when a 1-port measurement is made. We cannot derive an input voltage and current that can cause the output voltage and current of 1+j0, 0 because there's nothing actually connecting port 2 to anything (we would attempt to compute infinite voltage times zero transfer).

The software tool makes a special case for the last tile – the input term S11 (or Z11 or Y11) is derived based on the output terminal pair not being used, and uses these modified values in the extracted ABCD parameters set rather than the normal 2-port extraction. This turns the last tile into a simple one-port termination of the network that is being analyzed, which is actually what we likely measured with our test equipment anyway.

The pseudo-code for calculating the entire network response is then:

```
Foreach (frequency)
{
    // open circuit after last tile
    Set NodeState(out) = {V=1+j0,I=0+j0}
    Foreach (tile, starting at the last tile
            and working forward)
    {
        //Calculate {Vin,Iin} from {Vout,Iout}
        // using the ABCD matrix
        //  of the tile at the current
        // frequency
        NodeState(in) = ParameterSet(in ABCD form) *
                NodeState(out)
        Append NodeState(in) to the list
                of node values for that frequency
        Update the value of NodeState(out) =
                NodeState(in) and continue
    }
}
```

Deriving the Input Impedance

An important class of problems to analyze deals with determining the input impedance of a 2-port where there is mutual impedance between the input and the output, and some driving condition between the input and output is known. This occurs for example in the case of a vertical antenna array where multiple elements are driven with various known currents but at (possibly) different phase angles. The relationships between the antenna can be described as a Z-parameter matrix (likely frequency-dependent) where the self-impedance of the elements (Z11, Z22, etc.) are known, and the mutual impedance of the elements to one another (Z12, Z21, etc.) have been previously measured. These mutual impedances change the driving point impedance of each antenna depending on the excitation magnitude and phase at the other antennas and thus affect the performance of the array, it cannot be ignored.

To determine the input impedance of the 2-port in this case requires setting a forcing condition on the Z-parameter network during analysis. Normally we compute [V1,I1] based on the [V2,I2], but in this case we may know both I2 and I1 magnitude and phase (or their complex ratio), while not knowing V2 and V1. We must specify I1 in relation to I2. Since we do not know I2 when we are describing

the network, we specify I1/I2, as a ratio. This ratio must include the amplitude of the currents driving the two antennas (the two ports) as well as the phase relationship of the currents to one another (are we driving the antennas in-phase, out-of-phase, phase-quadrature, or something else).

Given the ratio, we have two unknowns and two equations, and can thus compute the node voltages, and from that the network input and output impedances. If we set I2 equal to 1+j0, then specifying a condition that the ratio I1/I2 = –1 would tell us that the two elements are driven in-phase but with the same current magnitude (recall the negative sign in defining the current convention). Setting I1/I2 = 0–j1 for example would set the driving current magnitudes equal but 90 degrees out of phase. The C# program specifies the ratio in dB, and phase angle in degrees to make the ratio entry more intuitive. The program extracts at each frequency the ABCD matrix from the Z-parameter matrix (or Y-parameter or S-parameter matrix).

First we derive V2 knowing both I1 and I2:

$$I_1 = CV_2 + DI_2 \quad thus \quad V_2 = \frac{I_1 - DI_2}{C}$$

Then we can compute V1 knowing both I2 and V2 using the ABCD formula we already know:

$$V_1 = AV_2 + BI_2$$

Now the analysis of the 2-port network can proceed normally since we have computed the node conditions at node 1 and it can cascade forward in the standard manner. In order to force some current at I2, there must be a shunt network terminating the output of the tile. It really does not matter too much what that shunt network is, because the input current is defined as a ratio of the output current, and V2 will be computed as necessary to force the value of shunt current, but it's convenient to use a shunt resistor of large value. Additionally, appending a shunt network after the Z-parameter network means that it is not the last tile, and so the full set of 4 Z-parameters (or Y, or S) will be used (which is what we want).

Transmission line

The response of a transmission line is described by several fundamental parameters: loss, velocity factor, and characteristic impedance. The latter two properties can be used in an equation to describe the input impedance of a transmission line that is terminated in an arbitrary load impedance. Zo is the characteristic impedance of the line, Zl is the load impedance, and beta is the phase delay characteristic of the line per unit length. A simple cable approximation is:

$$Z_{in} = Z_o \frac{Z_L + jZ_o \, tan(\beta l)}{Z_o + jZ_L \, tan(\beta l)}$$

However, this does not account for the loss of the transmission line, just its phase. A more general formula that includes line loss can be used if we are able to utilize complex arguments to hyperbolic trig functions. A side benefit is that the resulting equation appears simpler than the one above. In most of the classic math textbooks, the complex variable z (not impedance) is defined as x + j y. Then the trigonometric identities (in terms of x, y, and z) are:

$$sinh(z) = sinh(x) * cos(y) + j \, cosh(x) * sin(y)$$

$$cosh(z) = cosh(x) * cos(y) + j \, sinh(x) * sin(y)$$

$$tanh(z) = \frac{sinh(z)}{cosh(z)}$$

If we define a new constant gamma, $\gamma = \alpha + j\,\beta$, which includes both the loss (alpha) and phase delay (beta) properties per unit length, then $\gamma * l = \alpha\,l + j\,\beta\,l$ represents the loss and phase delay for the entire length of the cable. $\alpha*l$ is the loss of the line section, in Nepers; $\beta * l$ is the phase delay of the line section in radians. The general transmission line input impedance equation including loss then becomes:

$$Z_{in} = Z_o \frac{Z_L + Z_o \, tanh(\gamma l)}{Z_o + Z_L \, tanh(\gamma l)}$$

Where Zin, Zl, and Zo are complex impedances, using the previous identities to compute the hyperbolic tangent functions. The above formulation is usually called the low-loss cable approximation.

The COMPLEX number library included in the C# program provides regular and hyperbolic trig definitions for complex arguments, which makes the source code much more readable. *Visual Studio* C# 2008 does not include complex numbers, so a COMPLEX module was written to perform the common operations.

The Zin of the transmission line computed above is inserted into the ABCD matrix. For a series transmission line, the resultant ABCD matrix is:

$$\begin{bmatrix} 1 & Zin \\ 0 & 1 \end{bmatrix}$$

Where Zin is the result of the previously derived load to the transmission line Zl. Series and shunt stubs are modeled similarly to the series transmission line, except that the termination impedance of the stubs must be known as a circuit value (rather than being computed as in the case of the load terminating a series transmission line). The ABCD matrix for a series stub is the same as above, while the ABCD matrix of the shunt stub is:

$$\begin{bmatrix} 1 & 0 \\ 1/Zin & 1 \end{bmatrix} \quad \text{also can be shown as} \quad \begin{bmatrix} 1 & 0 \\ Yin & 1 \end{bmatrix}$$

The program uses a simplified stub termination impedance consisting of a resistor in series with an inductor and the series combination of those two in parallel with a capacitor. This matches pretty well to the cases where a low value of termination resistance has some series lead inductance, or when an open circuited line has some fringing capacitance.

As with other RLC loads, setting C=0 means 'no capacitance'. The program traps out cases where setting a reactance to zero would cause a divide-by-zero error, and instead removes that device from the model.

The loss of a transmission line can be approximated above a few hundred kilohertz from an equation that fits two coefficients to a function of frequency, one directly proportional to frequency, and the other proportional to the square-root of frequency. The program provides a drop-down list of a few commonly used coaxial and twin-conductor types of cables and pre-populates the characteristic impedance (both real and imaginary parts), the loss coefficients, and the velocity factor, you must supply the length of the cable. The values used in this program come from VK1OD's on-line transmission line loss calculator[4]. The pre-populated values can be over-ridden with other values of the parameters for a particular case, if known. The characteristic impedance of many coaxial cables starts to change a lot below one megahertz becoming increasingly more reactive, so the above equations need to be used with appropriate caution.

Series and Parallel RLC Networks

Series and shunt impedances are pretty straightforward to model as 2-port networks. A series impedance could be composed of a series connection of RLC parts, or perhaps as a parallel connection of RLC parts. At each frequency we compute the complex impedance of the collection of parts, and substitute that complex impedance number into the series ABCD matrix. Similarly a shunt impedance could be either series RLC, or parallel RLC. Again the complex impedance at each frequency is computed and substituted into the shunt ABCD matrix. These are the same matrix format as shown above.

Many times we will only have one or two of the RLC elements, and a common short-hand notation is to use a component value of 0 to signify that no component is present. Thus the calculator tool needs to handle a few special cases of zero value and instead remove that device from the circuit. For example, a value of 0 µH in a parallel circuit would short out the circuit, when what we really mean is that there is no inductor present (thus, the inductor has an infinite parallel impedance rather than a zero parallel impedance).

Transformers

A perfect transformer is easily modeled just from the turns ratio, but is not very useful since they don't exist and the assumption of perfect coupling is usually poor. Good transformer models can be quite complex yet still have significant error compared to the actual RF devices. A simple compromise model was selected that requires describing the primary winding inductance, the turns ratio, and the mutual coupling between the primary and secondary, but ignores winding resistance, stray capacitance and other effects. The symbols are defined as:

L1 = primary inductance
L2 = the secondary inductance
N = the turns ratio = Secondary turns / Primary turns
k = coupling coefficient

Traditionally it ranges from zero to plus one, but in this model we allow it to range from –1 to +1. A negative value indicates that the secondary winding is opposite in phase compared to the primary, thus a value of negative 1 indicates perfect coupling between windings but the transformer inverts the polarity of the output compared to the input. A value of +1 indicates the output winding has the same polarity as the input and is perfectly coupled. A value of zero indicates no coupling between the two windings at all. M is the mutual inductance computed from the above parameters. Since the model requires L1 to be defined, we compute the value of L2 (secondary inductance as):

$$L_2 = L_1 N^2$$

And the mutual inductance is computed as:

$$M = k\sqrt{L_1 L_2} = kNL_1$$

This shows that M has the same algebraic sign as k.

We then translate the transformer model into a T network with 3 inductors having inductance values as shown in Figure 3. Note that some of the inductors will have negative inductance values depending on the value of k. Then we compute the impedance of each element at each frequency of interest. Z1 = impedance of (L1–M), Z2 = impedance of (L2–M), and Z3 = the impedance of the mutual inductor (M). The ABCD parameters are derived from those impedances as:

$$ABCD = \begin{bmatrix} 1 + \dfrac{Z_1}{Z_3} & Z_1 + Z_2 + \dfrac{Z_1 Z_2}{Z_3} \\ \dfrac{1}{Z_3} & 1 + \dfrac{Z_2}{Z_3} \end{bmatrix}$$

Reading Data From a Measured Network

An important capability of the tool is to read in 2-port networks that have been measured by external test equipment. An industry-standard file format for 2-port networks is the Touchstone S2P format[5]. Most Vector Network analyzers (VNAs, and some other types of equipment) are capable of exporting an S2P formatted file. This tool is able to read an S2P file, decide what kind of 2-port network format it represents, parse the file, and store the 2-port parameters as a function of frequency in a data structure for that network tile. Each network block independently stores its own retrieved data, as the data for each network may have different formats, frequency ranges or frequency step sizes.

Parameter Frequency Interpolation

Since the 2-port data are captured by the test equipment at specific discrete frequencies, and the 2-port analysis may occur at frequencies that are not exactly aligned with the measured frequencies, some means to interpolate the retrieved data is needed. Since a 2-port parameter is in general a complex number, it is important to account for the fact that the phase could wrap almost 360 degrees between adjacent samples. The simplest way I have found to interpolate these samples is to convert the parameters to real + imaginary format, and then perform two one-dimensional linear interpolations in frequency (one for the real component, and one for the imaginary component of the parameter). This eliminates complications related to wrapping of coordinates.

It's possible to describe the output node in terms of the input node (backwards from what we have been doing so far) by using matrix inversion, although with the proviso that sometimes the ABCD matrix cannot be inverted (sort of like trying to divide by zero). We multiply both sides by the inverse matrix to work backwards.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} * \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} * \begin{bmatrix} A & B \\ C & D \end{bmatrix} * \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

thus

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$$

since a matrix times its inverse is the unity matrix (similar to multiplying by one).

The C# program defines an inversion operator for the ParameterSet matrix because it turns out to be useful when trying to remove a known ABCD matrix's impact when we have an unknown matrix. Additionally, an "IsSingular" test is implemented for the matrix inversion. A matrix which is singular cannot be inverted, this test allows us to abort a computation before error.

This has been left in as a hook if a capability to derive unknown 2-port network properties is added to the program in the future.

C# Program Implementation

The 2-port tool was written in Microsoft *Visual Studio* 2008 C#, which supports the NET 3.5 runtime (which is needed by the chart control). It has also been tested with *Visual Studio* 2010.

The Microsoft C# language starting with version 2.0 provides a "generic" capability, similar to the C++ "template." This allows the use of data structures, such as *List*, that can apply to a wide range of

data types. The List<Type T> data structure is very easy to use in C#. The tiles, parameter sets, node sets, etc. are each contained in lists, which do not have specific size constraints. There are ways to search through and re-order lists, and to identify an element in a list by index (position in the list) and to retrieve an index or a copy of the list element itself.

The design element types are marked as "*Serializable*" in the software. This allows a very simple implementation for streaming (saving) the data out to disk and retrieving it back. However each and every type that ultimately is sent to disk (even when embedded within another type) must be serializable. For example the *Complex* number type must be serializable in order to make a *ParameterSet* (which contains 4 complex numbers) serializable.

Additionally, each data structure has "Properties" associated with most of the internal object variables. By explicitly marking *Property* meta-data in the source code, the structures are "browsable" by an object browser control which accesses various Properties meta-data. The object browser is a tool that can be added to a windows form to permit run-time inspection and change of object variables. It's a little clunky to need to implement a property for each variable, but the ease of testing and debugging with the object browser is usually a good trade-off. In earlier editions of *Visual Studio*, the object browser tool needs to be loaded into the toolkit, as it is not installed by default. The toolkit is used only when compiling programs. When running the executable part of the program the browser tool is embedded within the EXE.

Required Runtime Packages

To run the software, the .NET 3.5 SP1 framework needs to be downloaded and installed. The NET 3.5 SP1 framework is a free download.[6]

Microsoft released a free very flexible chart control for .Net applications [MS Chart][7]. The chart control also needs to be downloaded from the Microsoft Web site and includes examples, help, etc. The control comes in two pieces: the runtime package, and the *Visual Studio* designer integration package. Only the runtime portion is needed to run the software.

Required Compiler Packages

A free version of C# 2008 (Express Edition) is available from Microsoft. There is a newer version available, C# 2010 Express, and the package has been compiled and tested under both versions[8]. The code in this tool was developed using the 2008 professional version. The free version does not support building an MSI installer package, but this 2-port tool does not interact with device drivers or the underlying Windows API's and does not have any specific installation requirements – thus an installer does not need to be built. The free Express Edition tool is all that's needed to change, compile, and experiment with the source code. One limitation of *Visual Studio* 2008 compared to predecessor versions is that it no longer produces EXE files that work with Windows 98, ME, NT, or 2000. The compiled code only works with Windows XP, Vista, and Windows 7.

C# 2008 Express does not support the integration of the MSChart control within the *Visual Studio* designer (thus allowing placing the control on a form, changing the size, etc.) but it will compile the code provided with this article since the control is already embedded in the designer resource. Thus you can change the code, but cannot interact with the control very much via the designer. If you have *Visual Studio* 2008 (not the free 2008 C# Express Edition), then you can also download the MSChart designer integration package which allows you to make all design-time changes on the chart control.

The express edition does not install the *Visual Basic* Power Packs 2.0 (*Visual Studio* 2005 version) by default. You will need to install version 2.0 of the control[9] so that the Microsoft.VisualBasic. PowerPacks.VS namespace reference can be found. Different versions of C# may or may not link properly to this reference. If it cannot resolve the reference (compile errors, and an error exclamation mark in the references folder of the solution explorer), first delete the reference, then add it back in again which will prompt you to provide a link to the proper location on disk.

Using the Program

The Design tab of the program is where designs are entered. Initially the program brings up a grid of 24 slots. The slots are in series, from upper left to lower right, across each row left-to-right then down to the next row. The last (right-most) tile on the first row is just before the first (left-most) tile on the second row. Right-clicking the mouse allows inserting or deleting tiles. The first tile can be inserted using either Insert Right or Insert Left, and it will be inserted between nodes 0 and 1. The insert operation brings up a menu that allows selection of the type of tile (2-port) that is to be placed. Tiles that have been placed on the design tab can be selected by clicking the mouse over the desired tile, and that tile will be highlighted. The selected tile can be deleted by right-clicking the mouse. The selected tile can be browsed in the tile browser tab, and its parameter type, and component values can be quickly changed. Once a tile is selected, a new tile can be inserted to the left of or to the right of the currently selected tile. If a new tile is inserted when no tile is selected, it will be inserted at the first possible position (between nodes 0 and 1) and all the rest of the tiles will get bumped to the right. The node numbers are shown on the background of the design tile to make it easier to identify and remember node numbers when specifying them on the analysis tab.

The tile selection menu allows direct input of the parameters, selection of a forcing current ratio (for Z-parameter antenna matrices usually), and whether a 4-parameter set for a (S, Y, Z, or ABCD) tile is fixed in frequency, or is frequency-dependent. If it is frequency dependent, a menu pops up allowing the selection of an S2P file to load it from. A design can be saved to a file, the design cleared, and a design retrieved from a file. When saving a design, the entire state of the design tiles are saved, including all the frequency-dependent parameters (they lose their association to the original S2P file but all the data are saved). A List data structure holds the frequency-dependent list of parameters values, so the program does not set arbitrary limits on the number of frequency points that any tile can hold until the program runs out of memory or disk space. Designs with over a thousand frequency points have been tested (and it results in a several hundred KB file).

One limitation is that the frequency points must be in the original S2P file in increasing frequency order, but I've not yet found any S2P files that violate this ordering. The actual frequency of each analysis point does not need to line up with the frequencies read in from the S2P file, nor do different tiles need to align in frequency with any other tiles. The software searches each tile and finds the bracketing set of frequency parameters, then does a linear-in-frequency interpolation between them to the actual analysis frequency. This works well except in the case where the parameter values have a large deviation from a linear approximation between the two frequency points. In that case, the S2P file that the parameters were read in from needs higher measurement resolution and more points or a narrower frequency sweep to minimize the difference between adjacent measured samples.

The analysis parameters page allows editing the analysis parameters (start and stop frequency, selected displays, number of analysis points, etc.) in an object-browseable format, it duplicates much of the functionality of the controls on the chart tab.

The MSChart control is used in the C# program to display the

results of the analysis over frequency on the analysis tab. The chart does not have a polar-mode that can work as a Smith chart, so it's only used to display the parameters in rectangular format. The zoom, pan, and other modes of the chart are active so you can drag the mouse to zoom into the chart, or push the zoom-reset button on each axis to go back to the full chart. It's faster to do than to explain.

The Analysis Tab of the program brings up the chart control and a few buttons and check boxes to enable displaying the impedance (in several formats) at a selected node, and the voltage and power transfer functions between a pair of nodes. The analysis button will turn pink when the analysis results in some type of error (such as trying to measure the power into the last node after the last tile). The values are clipped to –308 dB when a zero-value is converted to dB to avoid a run-time exception with the chart. If any values display as –308 dB, then it probably represents a value of minus infinity. There's a control or two to set auto scale or manual scale and max/min values for the Y-axis (except phase, which has a fixed axis of –180 to +180 degrees), and to select linear or logarithmic frequency axis.

The Tile Browser allows changing the values of the selected tile (the one with a black box around it). This is very convenient for changing tile values, types, or drive conditions quickly.

Source Code

The complete C# source code tree with all designer resources, the compiled executable file, and the examples have been placed on the ARRL QEX Web site[10]. If you just want to run the program, simply download the ABCDmatrix EXE file, unzip it, and execute it. You should install the CHM (help) file into the same folder as the executable program.

The source can be edited, compiled, and built in the *Visual C# Express* 2008 integrated design environment (IDE). Make sure to check that all references are present, and resolve any that are missing (such as the visual Basic Powerpack, the MSChart control, etc.).

Figure 1 – Series Impedance 2-port.

Figure 2 – Shunt Impedance 2-port.

Figure 3 – 2-port model of a transformer with mutual inductance.
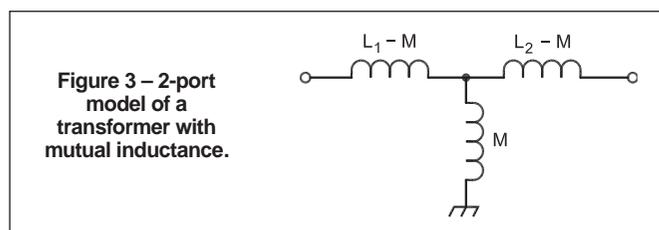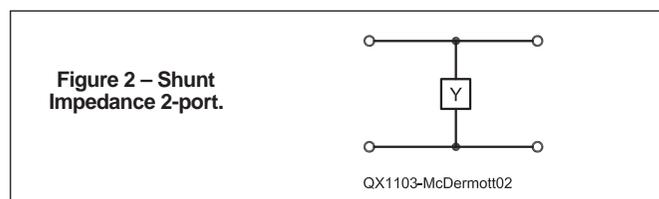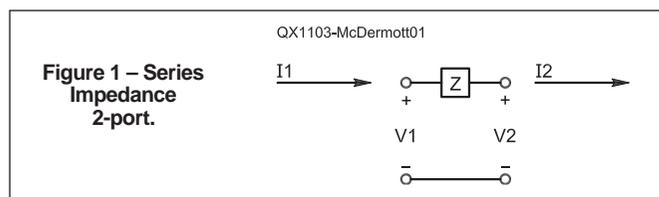
Menus & Operation

The program has a pretty simple interface. Select the Network Design Tab, and Right-click to bring up a context menu. When there are no tiles on the design surface, or no tile has been selected, either InsertRight or InsertLeft will insert a new tile at the first location (between nodes 0 and 1) and bring up a menu to specify it. The menu allows fixed values or frequency-variable values read from a file (for S-, Z, and Y- parameters). Left-click a tile to select it. The selected tile can be edited on the Tile Browser tab. The selected tile can be deleted, copied to the clipboard, or cut (and copied to the clipboard). InsertRight and PasteRight will insert a tile after the selected tile. InsertLeft and PasteLeft will insert a tile before the selected tile. Delete will remove a tile without copying it to the clipboard, cut does copy it to the clipboard.

The Analysis Parameters tab allows changing the analysis parameters, but largely duplicates the checkboxes and numeric windows on the analysis tab. The Analysis tab displays the results of an analysis. The 'from' box selects the node where impedance (or S11) is displayed, and is the starting point for a Power transfer or Voltage transfer analysis. The 'to' box is the terminating node for voltage or power transfer. Note the power transfer can not be computed at the output of the last tile because no power can be delivered to an open circuit. The tab will catch and correct errors in the values of the 'to' and 'from' tab (except for power analysis) and will try to select the whole design when blank or out-of-range values are selected. The Analyze button will update the display. It will turn pink if a gross analysis error is present (for example trying to display the power into an open circuit).

A design can be saved to and retrieved from disk, it uses the suffix 'til' (for 'tile'). 'New' erases all the tiles on-screen and the clipboard.

Note that you can use the mouse to drag/select an area of the chart and the view will zoom into that rectangular area. Slider controls allow scrolling horizontally or vertically, and buttons allow reverting back to un-zoomed axis.

Examples

Here are three examples and solutions worked out using the tool. The examples have been chosen to illustrate some common functions of the software tool, especially where one or two of the uses might otherwise be a bit confusing.

Example 1 – shunt and series coaxial stubs.

The first example is a 50 Ω series source feeding a shunt stub, then a series stub, then a 50 Ω shunt load resistor. Figure 4 is a screen
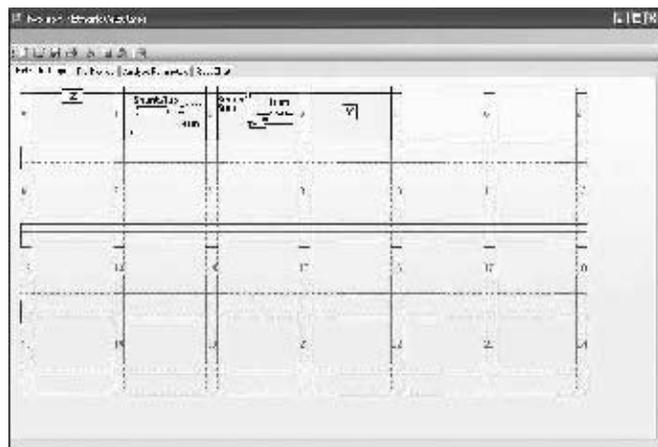
Figure 4 – Network Design tab shown the circuit of example 1, with the shunt stub selected.

shot of the Network Design Tab. As each tile is inserted, a pop-up menu allows selecting the type of tile, and the parameters of that tile. Later on, an existing tile on the Network Design tab can be copied, which copies all its parameters, or an existing tile can be selected and then edited on the tile browser tab. The first element has a series RLC of 50 Ω + 0 μH + 0 pF. The 0 pF means "no capacitor", so the first series element is 50+j0. The shunt stub has a termination R=0.2 Ω, L= 0.02 μH, and C = 1 pF. It's made from RG-58, and has a length of 10 meters. The velocity factor, characteristic impedance and loss parameters are filled in automatically when selecting the RG-58 cable type. The series stub is also 10 meters of RG-58, but has R=1 MΩ, L=0 μH (is open-circuited instead of shorted), and has 1 pF of capacitance. The shunt load is series RLC of 50+j0. The shunt stub is selected.

Figure 5 shows the Tile browser tab. Since the shunt-stub is selected, it's what is being browsed by the tile browser. All of the tile parameters can be changed using the tile browser. The Cable length field is highlighted, and the stub is 10 meters long. Figure 6 shows the Rect Chart tab — the analysis results in rectangular chart format. The voltage transfer from node 0 to node 4 is shown, magnitude in dB, and the phase in degrees. We can see that the series and shunt stubs, although both of exactly the same length, aren't exactly on the same frequency due to slight differences in their stray termination impedances. Also, the insertion loss increases slightly with fre-

quency (more readily seen using the power transfer from node 0 to node 3 (not 4 as the transfer to node 3 is the power transfer into the 50 ohm shunt load).

Figure 7 shows the input impedance at node 0, in real ohms, imaginary ohms format. If you change the series source resistor from 50 ohms to 0 ohms, a much different transfer function results. Y-auto
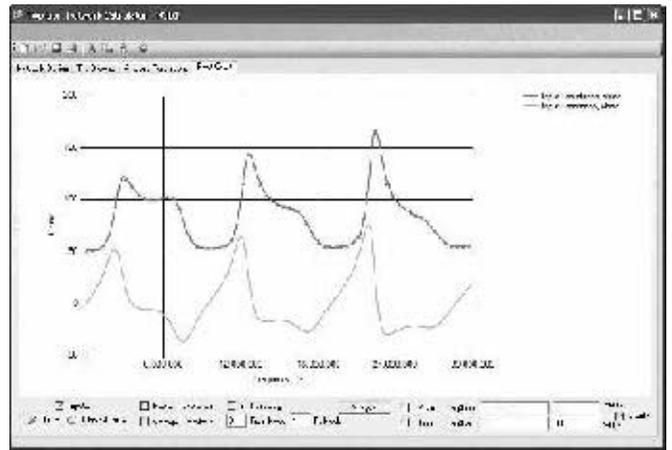
Figure 7 – Example of data input and output on the Rect Chart Tab showing real and imaginary impedance at Node 0.

Figure 5 – Example of data input on the Tile Browser Tab for the selected shunt stub of Figure 4.
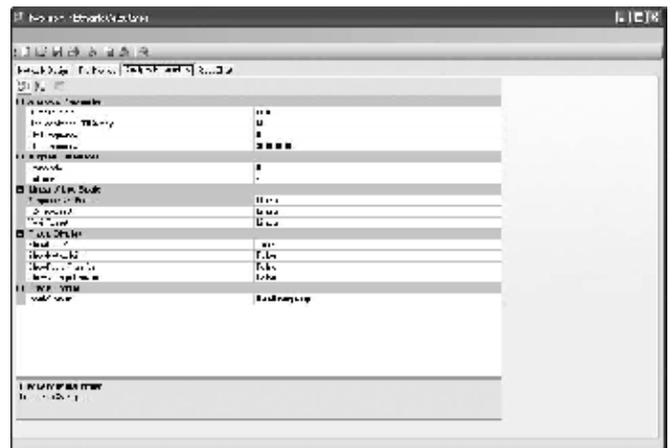
Figure 8 – Example of data input on the Analysis Parameters Tab.

Figure 6 – Example of data input and output on the Rect Chart Tab showing magnitude and phase.
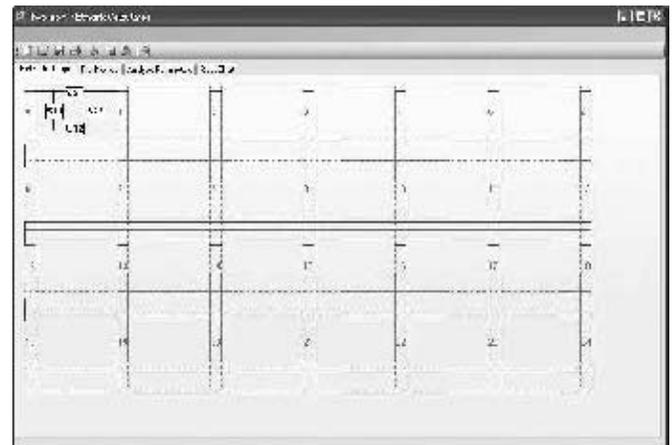
Figure 9 – The Network Analysis Tab showing S parameter input for a Tile.

scale is selected, analysis range is zero to 30 MHz.

Figure 8 shows the Analysis Parameters tab. Normally most analysis parameters can be selected from the 'Rect Chart' tab. Sometimes however the defaults may not be what are desired. This tab allows changing for example the Reference resistance for S11 analysis (the default is 50 Ω), or the number of points in the analysis (this example uses 400 points), as well as other parameters.

Example 2 - Subtracting networks

This next example shows how to mathematically remove known text fixture strays from a measured network response. The solution is hypothetical since the compensation consists of networks that in fact cannot be realized. However, its purpose is to let us discover the underlying network properties.

Figure 9 shows an S-parameter tile, which is loaded from data measured by a VNA of a termination load through some RG-58 coaxial cable. Figure 10 shows the S11 analysis at node 0, which is exactly the same as what's displayed on the network analyzer (as it should be). The number of points and the start/stop analysis range have been set to the same as the VNA measurement. Figure 11 shows the addition of a coaxial line segment in front of the measured result. The length of the coaxial cable is set to negative 1.04 meters, and the characteristic impedance is set to 50.75 – j0.4 Ω. This approximately subtracts out the effect of the cable feeding the termination load (but ignores some other error sources), letting us see that the load imped-

ance is roughly 50.5 Ω (with some uncompensated error).

Figure 12 shows the added cable compensation parameters. This allows experimenting with a wide range of parameters values for the selected tile's values. Figure 13 shows the compensated S11 measurement. Figure 14 shows a close-up of the resultant real part of the load impedance.

Figure 12 – The compensating transmission Line data shown in the Tile Browser.
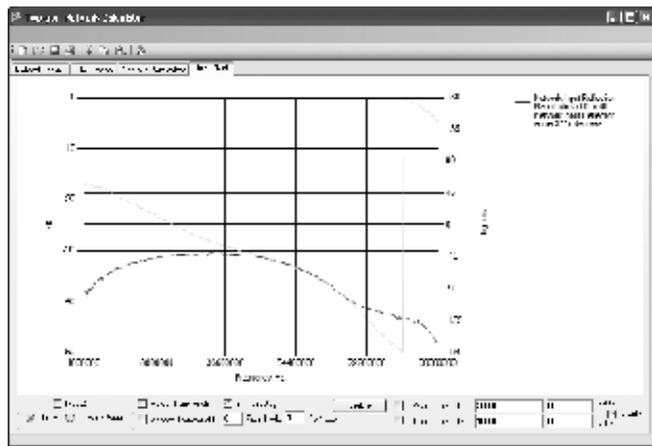
Figure 10 – Display of S11 for Tile in Figure 9.
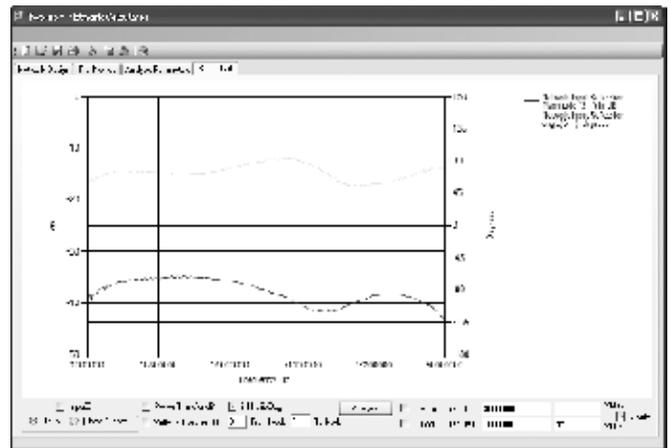
Figure 13 – The resulting S11 data after the compensating transmission line is added.
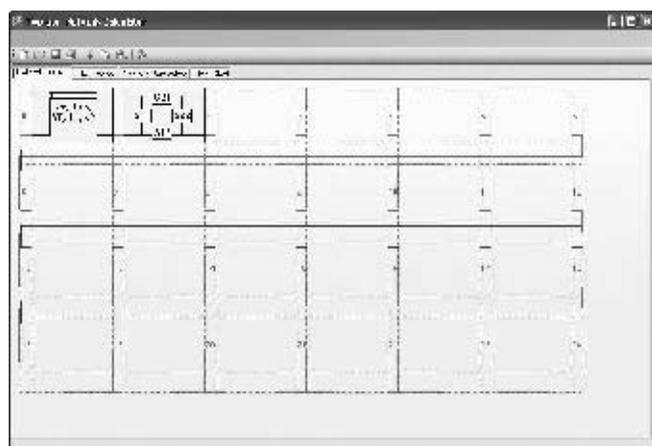
Figure 11 – Display of Network Design with a compensating transmission line added to the S parameter Tile.
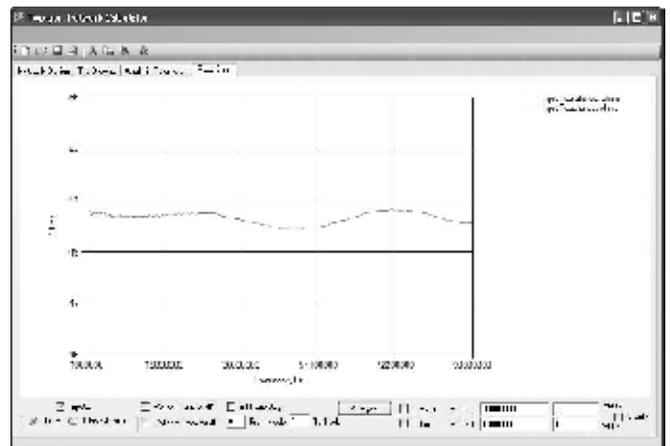
Figure 14 – A zoomed version showing the real part of the compensated S11 value.

You can experiment with the parameter values on the 'Tile browser' tab with the cable compensation tile selected, and look almost instantly at the impact to S11, Zin, etc. on the Rect chart tab. Of course just a negative-length cable can't compensate for some of the other measurement errors.

Example 3 – Phased vertical antenna

This example demonstrates the changing input impedance when trying to feed a two-element phased vertical array under varying conditions. It illustrates the impact of mutual coupling on the feed network design for the verticals. Due to the lack of available measured data, a 2-element quarter-wave vertical antenna array (each with 16 radials) over average ground was modeled in NEC2. The self (Z11) and mutual (Z12) impedances over a range of frequencies were derived, and an S2P then file created containing Z11, Z12, Z21, and Z22 for the two element array at each frequency. Because this is a theoretical model, symmetry was assumed, thus Z12=Z21 and Z11=Z22 which would not happen in a real network due to slight differences between the two elements.

We can then compute the input impedance of one of the antenna elements for a variety of feed conditions at the other element, and graph the various results. We need to place a dummy terminating resistor at the output of the array, typically a shunt element of large resistance when using a current-forcing condition for the output of the Z-parameter network — otherwise we would be attempting to force current into an open circuit (and thus creating infinite voltage) and the analysis indicates an error due to an attempted divide by zero; attempting to force zero current into an infinite impedance still results in the same error. It's easiest to just place a high-value shunt load at the output of the Z-parameter network as all practical drive conditions can be modeled without needing to change the topology.

The file "Monopole_z.s2p" contains the Z-parameter data for the modeled two-element array from 3.5 to 4.0 MHz in 50 kHz steps. We load it into the tools by selecting the "INSERT" Block Type Z, Frequency Table for Matrix Elements selections from the tile input form. Then we tell the tool which s2p file to read, and all the data points from that Z-parameter network are inserted.

Figure 15 shows the network connectivity for Example 3. In order to measure the input impedance of one antenna element with essentially zero coupling to the second element, the current forcing of the Z-parameter network is set to FALSE. Then the high value of the shunt network assures that virtually no current flows out of the network. This means that we will measure just Z11 of the network. Figure 16 shows the resultant self impedance of one of the antenna elements.

The resistance varies from just over 30 Ω (at 3.5 MHz) to just under 50 Ω (at 4.0 MHz) and the reactance varies from about –25 Ω to +15 ohms, with resonance near 3.8 MHz. If we now set the current forcing function to TRUE, and set IMAG current ratio = 0 and the REAL forcing current ratio to -1 we are feeding the two
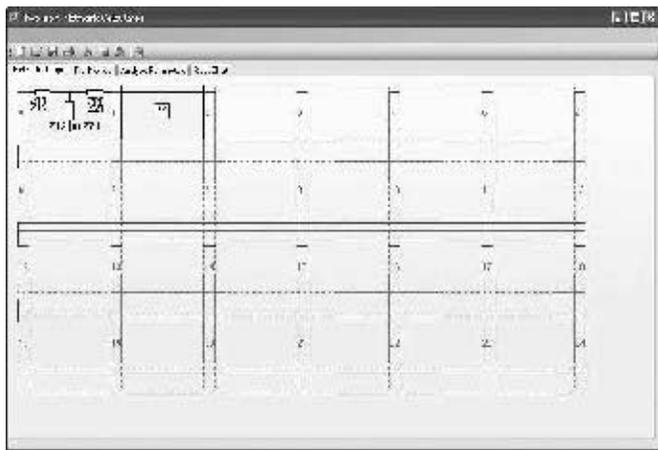
Figure 15 – The Network Design for Example 3.
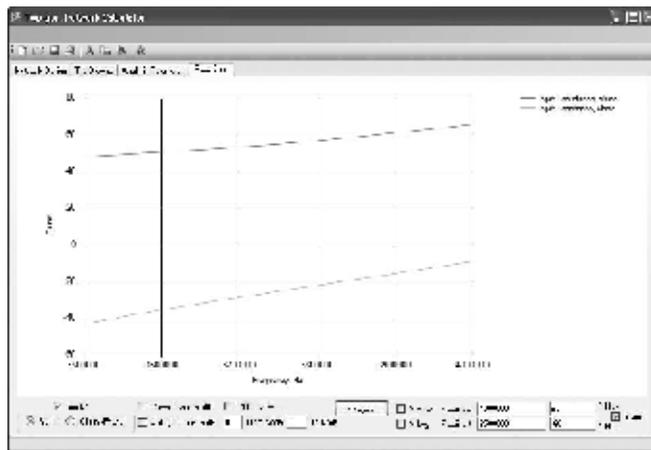
Figure 17 – The resultant input impedance of one antenna element when the two elements are fed in-phase with equal current.
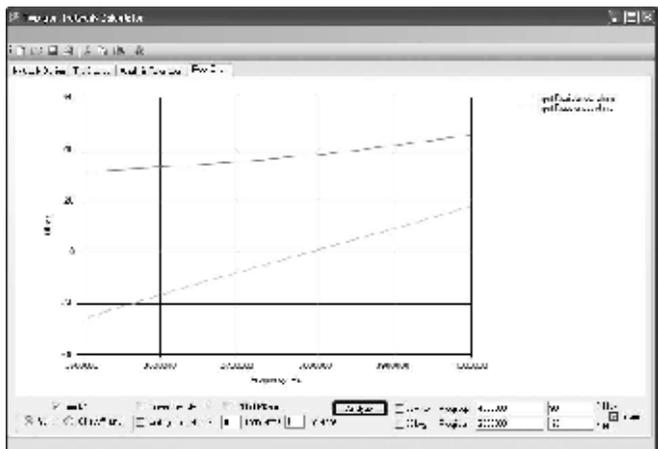
Figure 16 – The resultant self impedance of one of the antenna elements in Example 3.
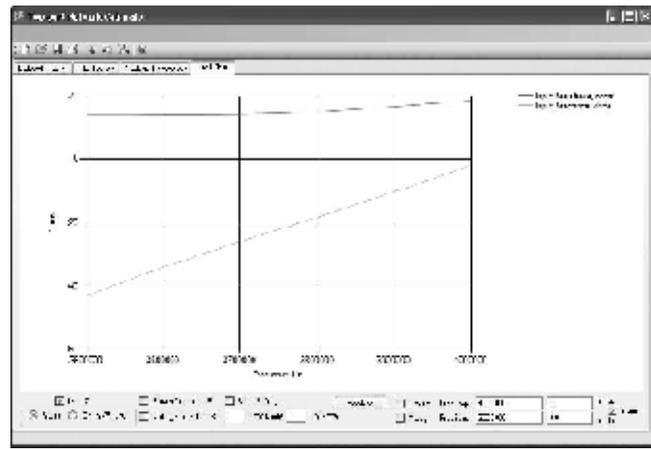
Figure 18 – The resultant input impedance when the two elements are fed 90-degrees to each other.

elements IN-PHASE (the negative sign is because the two port currents are defined with opposite signs to allow chaining the calculations, recall Figure 1 above). The input impedance to the first element changes dramatically as a result.

Figure 17 shows the resultant input impedance of one antenna element when the two elements are fed in-phase with equal current. The input resistance now varies from 50 to 70 Ω, and the reactance from –45 to –10 Ω, and the element does not have a resonance between 3.5 and 4.0 MHz. If we feed the two elements with a 90-degree phase (forcing IMAG = –1, REAL = 0) the input impedance drops a lot. Changing the sign of Forcing_IMAG alters which element is fed in leading phase to the other, and the impedance similarly changes a lot.

Figure 18 shows the resultant input impedance when the two elements are fed 90-degrees to each other. In this case the input resistance of the fed element is around 15 Ω across the entire frequency range. If we change which element is leading in phase compared to the other, the impedances again change.

Figure 19 shows the resultant input impedance when the two elements are fed with the opposite –90-degrees phase relationship. Here the apparent resonant frequency of the fed element appears to be about 3.6 MHz, and the resistance varies from about 50 to 70 Ω.

The point of this exercise is to show that the feed networks for a phased vertical array must contend with quite different feed-point impedances as the phasing conditions of the array are altered. With these conditions now quantifiable, we can insert our desired phasing and impedance matching networks in front (to the left) of the Z-parameter network of example 3 and model whether it accomplishes delivering the desired current into element one of the array in the proper phase, and presents a proper load impedance to the transmitter, all over the desired frequency range.

For Further Work

This tool has already proven to provide extremely quick feedback for passive circuit and network ideas, sometimes much faster than trying to set up a model in the normal circuit simulator tools and providing the appropriate control statements. Rapid change of 2-port values via the tile browser allows quickly iterating through a number of design alternatives.

Further work could be done to extend this tool. Some possibilities are:

1. Provide a Smith Chart view of the input return loss, S11. While it is straight-forward to convert S11 to a reflection vector X and jY components, the Microsoft chart control does not readily allow the addition of the corresponding appropriate background axis for the polar chart.
2. Provide the ability to extend the Z-parameter matrix from 2-ports to a larger number of ports (such as 4-ports) that would allow the entry of mutual impedance and self impedance data for antenna arrays with more than 2 elements.
3. If the response of an entire network has been measured, but one (and only one) of the elements of the cascade of 2-port networks is unknown, it is possible to compute what the response of that unknown network must be.

Tom has been a licensed amateur for 40 years, and is a life member of the ARRL. He enjoys HF, SSB and RTTY operation, various technical topics in amateur radio mainly involving computers. He has a Bachelor's degree in Electrical Engineering from the University of California, Berkeley. His professional background is in high speed fiber optic transmission and switching equipment. Tom holds 12 patents and is a member of the IEEE and Internet2.
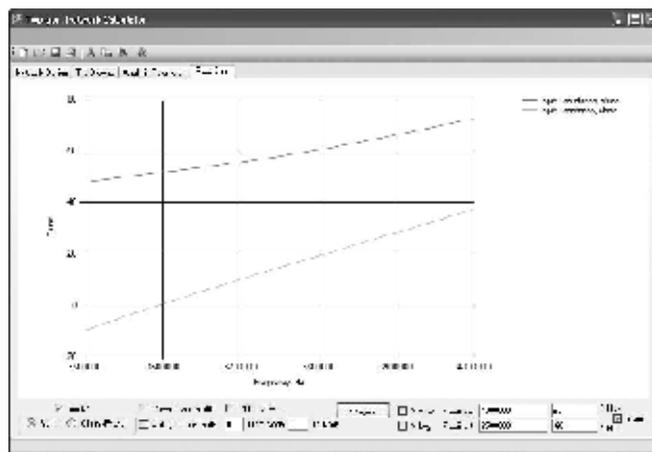
Figure 19 – The resultant input impedance when the two elements are fed with the opposite -90-degrees phase relationship.

Notes

[1] *Microwave Engineering*, Third Edition. David M. Pozar, 2005 John Wiley & Sons, ISBN 978-0-471-44878-5.

[2] *Microwave Filters, Impedance-matching Networks, and Coupling Structures*, Matthaei, Young, and Jones, 1980, Artech House Books, ISBN 0-89006-099-1, Sections 2.06 through 2.13.

[3] Vertical Phased Arrays: Part 5, Forrest Gehrke, K2BT, *Ham Radio Magazine*, December 1983, pp 59-64.

[4] VK1OD transmission line loss calculator is located at: **vk1od.net/calc/tl/tllc.php** The k1 and k2 (frequency-dependent loss parameters) for a few cable types were extracted from this tool.

[5] Touchstone® File Format Specification Rev 1.1, Copyright © 2002 by the EIA/IBIS Open Forum. Available from **www.vhdl.org/pub/ibis/connector/touchstone_spec11.pdf**

[6] The Microsoft NET 3.5 SP1 Framework can be downloaded (as of September 2010) from: **www.microsoft.com/downloads/details.aspx?FamilyID=AB99342F-5D1A-413D-8319-81DA479AB0D7&displaylang=en**

[7] The Microsoft Chart Controls for .NET Framework 3.5 can be downloaded (as of September 2010) from: **www.microsoft.com/downloads/details.aspx?FamilyId=130F7986-BF49-4FE5-9CA8-910AE6EA442C&displaylang=en**

[8] Microsoft C# 2008 Express Edition can be downloaded (as of September 2010) from: **www.microsoft.com/express/downloads/**

[9] The Microsoft Visual Basic 2005 Power Packs 2.0 can be downloaded (as of September 2010) from:

www.microsoft.com/downloads/details.aspx?FamilyID=92faa81e-e9c1-432c-8c29-813493a04ecd&displaylang=en

[10] All files have been placed on the ARRL Web site at: **/www.arrl.org/qexfiles.** The file set contains the tool executable (EXE file), a help file, the tool source code (released under the GPL license), the three example design files, and the s2p files for the examples. It does not contain the Net frameworks or chart tools from Microsoft.

QEX⌐

Al Christman, K3LC

Grove City College, 100 Campus Dr, Grove City, PA 16127-2104

A Study of Tall Verticals

The author conducted a study of quarter wavelength vertical antennas used on bands higher in frequency than that for which they were designed. Here are the results of his analysis.

Some time ago my friend Joe Johnson, K3RR, asked me how his quarter wavelength 160 meter vertical would perform on 80 meters. We both knew that on 80 meters the antenna height would be about 0.5 λ, so we expected the gain to be higher, with the peak of the main radiation lobe occurring at a lower take-off angle. Neither of us had a good idea of what the exact numerical results might be, however.

With that in mind, this article was written in order to answer several questions: First, "How well does a vertical antenna that was designed for one band perform when it is operated on a higher frequency band?" The second (related) question: "Is it worth the effort to put up a ⅝ λ radiator?"

Computer Simulations

Full-size 0.25 λ radiators on 160 meters are well over 100 feet tall, so it was decided to utilize a triangular tower as the basis for the top-band antenna model. A tower whose face width is 12 inches was employed, although the bottom end was "tapered" from three legs down to one single tubular conductor, to simplify the construction of the feed point (see Figure 1). The operating frequency was set to 1830 kHz, and a ground-screen consisting of sixty λ/4 no. 14 AWG radials (length = 134.4 feet) was installed. These radials were buried to a depth of 3 inches in "average" soil having a conductivity of 0.005 siemens per meter and a dielectric constant of 13.

I simulated all of the antennas described in this article using the *EZNEC* software package, which is available from Roy Lewallen, W7EL.[1] A real antenna system would probably be made from zinc-coated steel tower sections and copper radials. *EZNEC* allows the use of only a single type of conductor, so I selected aluminum since its conductivity is better than that of zinc, but inferior to copper.

Results for a Quarter Wavelength 160 Meter Vertical

Table 1 shows what happens when the full-size λ/4 top-band antenna is used on 160 meters and also on the 80 meter band. Initially, the length of the tower was

[1]*EZNEC* antenna-simulation software is available from Roy Lewallen, W7EL, PO Box 6658, Beaverton OR 97007.
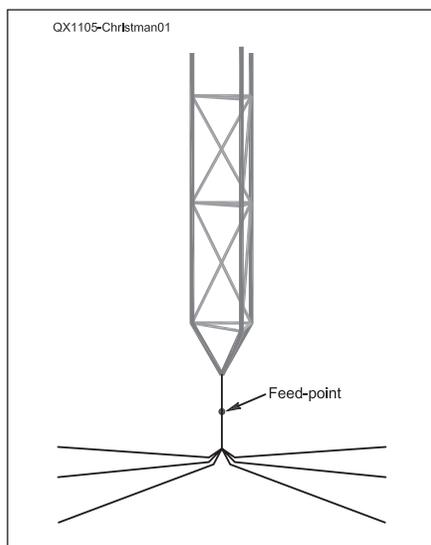
Figure 1 — Close-up of the base region of the full-size 160 meter vertical antenna system, showing how the three tower legs are tapered down to a single conductor at the lower end of the monopole. For clarity, only six of the 60 buried radials in the ground screen are shown.

Table 1

Performance of a full-size quarter-wavelength vertical monopole antenna designed for 160 meters, when used on that band and also on 80 meters. The monopole is built from triangular tower sections (12 inch face width), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 134.4 feet (0.25 λ at 1830 kHz). The vertical element is tuned to resonance at 1830 kHz, with a resulting antenna height of 125.418 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND	
	160 Meters	*80 Meters*
Operating Frequency (kHz)	1830	3650
Input Impedance (Ω)	36.0	706 – *j* 310
SWR (50 Ω ref.)	1.39	16.9
Peak Gain (dBi) and Take-off Angle (°)	1.20 at 22.9	1.14 at 18.7
Gain (dBi) at 5° Take-off Angle	–3.42	–3.45
Gain (dBi) at 10° Take-off Angle	–0.36	–0.13
Gain (dBi) at 15° Take-off Angle	0.75	0.95
Gain (dBi) at 20° Take-off Angle	1.15	1.12
Half Power Beamwidth (°)	43.7	31.6
Efficiency (%)	41.2	33.0

adjusted in order to resonate the entire system at a frequency of 1830 kHz. The input reactance fell to (approximately) zero when the overall tower height was "pruned" to just 125.418 feet. Key performance parameters — such as gain, take-off angle, input impedance, SWR, and efficiency — were then recorded from the *EZNEC* output.

Next, the operating frequency was changed to 3650 kHz, in the 80 meter band, and the computer analysis was repeated. An examination of Table 1 reveals that the peak gain of the tall monopole is similar on both bands, although the elevation angle is about four degrees higher on "Top Band." The SWR is quite high on 80 meters, so an impedance-matching network would be needed there. Figure 2 displays the accompanying elevation-plane radiation patterns. At take-off angles of 20° or less, the gain is very similar on both bands, but the signal strength at higher elevation angles is reduced on

80 meters, due to the greater electrical height of the tower on this band.

Notice that the system efficiency is significantly lower on 80 meters, as compared to 160 (33% versus 41.2%). This is probably because the region on the tower where maximum current occurs is more than 60 feet up in the air, and the corresponding displacement current strikes the earth on a portion of the ground screen where the radials are spread relatively far apart. To confirm this, the tower was shortened to a height of only 61.3 feet, which allowed the antenna system to resonate at 3650 kHz (with the ground-screen left unchanged). When using the λ/4 tower in combination with a λ/2 ground-screen, the efficiency on 80 meters rose to 40.7%

Results for Two Different Quarter Wavelength 80 Meter Verticals

The same kind of triangular tower sections (12 inch face width) were employed to construct an *EZNEC* model of a λ/4 monopole for the 80 meter band. The ground-screen was composed of sixty no. 14 AWG radials, each with a length of 67.368 feet, which is equal to λ/4 at a frequency of 3650 kHz. The vertical element was then adjusted to resonate the antenna system at this same frequency, which required an overall height of 61.163 feet. Recall from the previous discussion that the resonant height for a similar tower with a larger ground screen (134.4 foot radials) was 61.3 feet.

The outcome from the *EZNEC* simulation is posted in Table 2, not only for usage
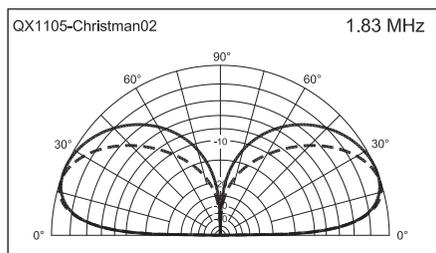
Figure 2 — Elevation-plane radiation patterns for a resonant ¼ λ 160 meter vertical antenna (made from tower sections) when used on both 160 and 80 meters.
Solid trace = operating on 160 meters (1830 kHz)
Peak gain = 1.20 dBi at 22.9° take-off angle
Dashed trace = Operating on 80 meters (3650 kHz)
Peak gain = 1.14 dBi at 18.7° take-off angle

Table 2
Performance of a full-size quarter-wavelength vertical monopole antenna designed for 80 meters, when used on that band and also on 60 and 40 meters. The monopole is built from triangular tower sections (12 inch face width), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 67.368 feet (0.25 λ at 3650 kHz). The vertical element is tuned to resonance at 3650 kHz, with a resulting antenna height of 61.163 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND		
	80 Meters	*60 Meters*	*40 Meters*
Operating Frequency (kHz)	3650	5367	7150
Input Impedance (Ω)	33.7	155 + j197	602 – j71.1
SWR (50 Ω ref.)	1.48	8.31	12.2
Peak Gain (dBi) and Take-off Angle (°)	0.53 at 25.1	0.75 at 23.7	1.05 at 20.9
Gain (dBi) at 5° Take-off Angle	–5.13	–4.98	–4.34
Gain (dBi) at 10° Take-off Angle	–1.61	–1.34	–0.70
Gain (dBi) at 15° Take-off Angle	–0.23	0.09	0.64
Gain (dBi) at 20° Take-off Angle	0.36	0.65	1.04
Half Power Beamwidth (°)	44.2	40.7	34.9
Efficiency (%)	35.1	35.3	34.5

Figure 3 — Elevation-plane radiation patterns for a resonant ¼ λ 80 meter vertical antenna (made from tower sections) when used on 80, 60, and 40 meters.
Solid trace = operating on 80 meters (3650 kHz)
Peak gain = 0.53 dBi at 25.1° take-off angle
Dashed trace = operating on 60 meters (5367 kHz)
Peak gain = 0.75 dBi at 23.7° take-off angle
Dotted trace = operating on 40 meters (7150 kHz)
Peak gain = 1.05 dBi at 20.9° take-off angle

Table 3
Performance of a full-size quarter-wavelength vertical monopole antenna designed for 80 meters, when used on that band and also on both 60 and 40 meters. The monopole is built from lengths of tapered metal tubing (diameter varies from 3 inches to 0.75 inch), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 67.368 feet (0.25 λ at 3650 kHz). The vertical element is tuned to resonance at 3650 kHz, with a resulting antenna height of 67.482 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND		
	80 Meters	*60 Meters*	*40 Meters*
Operating Frequency (kHz)	3650	5367	7150
Input Impedance (Ω)	36.0	207 + j311	828 – j505
SWR (50 Ω ref.)	1.39	13.7	22.7
Peak Gain (dBi) and Take-off Angle (°)	0.77 at 24.7	0.79 at 22.9	1.07 at 19.6
Gain (dBi) at 5° Take-off Angle	–4.83	–4.83	–4.05
Gain (dBi) at 10° Take-off Angle	–1.33	–1.20	–0.46
Gain (dBi) at15° Take-off Angle	0.05	0.19	0.79
Gain (dBi) at 20° Take-off Angle	0.63	0.72	1.07
Half Power Beamwidth (°)	43.7	39.3	31.9
Efficiency (%)	37.0	34.9	32.7

on 80 meters, but also for both 60 and 40 meters. Notice that the peak gain and the input resistance rise continually as the operating frequency is increased, while the take-off angle falls by several degrees. The radiation-pattern plots (which are similar to one another) are given in Figure 3.

A second λ/4 80 meter vertical element was also designed, but this time the monopole was constructed from lengths of aluminum tubing, with the diameters of the various sections tapering down from 3 inches to a minimum of 0.75 inch. This vertical element was placed over the same ground screen as before (sixty no. 14 AWG aluminum radials, each 67.368 feet long), and immersed in the same type of "average" soil. Table 3 shows the key performance parameters calculated by *EZNEC*, on the three bands of interest (80, 60, and 40 meters). It can be seen that the values for gain and elevation angle are similar to those obtained earlier when the monopole was built from tower sections.

Results for a Quarter Wavelength 60 Meter Vertical

A full-size λ/4 monopole was then designed for operation on the 60 meter band, using lengths of tapered aluminum tubing. The diameter of the largest section is 3 inches, with subsequent portions gradually reduced in size, reaching 1.5 inches at the tip. The length of the 60 buried no. 14 AWG aluminum radials is 45.816 feet, or λ/4 at a frequency of 5367 kHz. The overall height of the vertical element was trimmed to 44.884 feet, which provided resonance at the center of the band (5367 kHz). This antenna can be utilized effectively on three bands (60, 40, and 30 meters), and the outcome from the computer modeling is summarized in Table 4. As before, we find that raising the frequency yields an increase in peak gain and feed-point resistance, along with a reduction in the take-off angle of the main radiation lobe.

Results for a Quarter Wavelength 40 Meter Vertical

Table 5 illustrates what happens when a λ/4 monopole constructed from aluminum tubing is employed on the 40 meter band. Here the diameter of the tapered sections of tubing ranges from 2 inches down to 0.75 inch, and the ground screen consists of sixty no. 14 AWG aluminum radials whose length is 34.39 feet, which is λ/4 at 7150 kHz. The vertical element was adjusted to resonate the antenna system at this same frequency, leading to a final height of 33.736 feet.

The outcome from the computer analysis is displayed (Table 5) for applications on both 30 and 20 meters, along with the
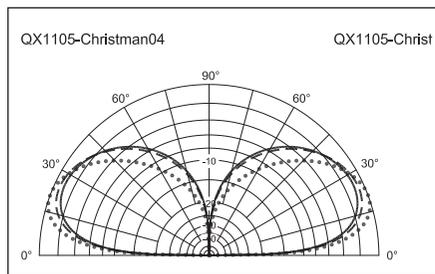
Figure 4 — Elevation-plane radiation patterns for a resonant ¼ λ 40 meter vertical antenna (made from tapered sections of tubing) when used on 40, 30, and 20 meters.
Solid trace = operating on 40 meters (7.15 MHz)
Peak gain = 0.29 dBi at 26.4° take-off angle
Dashed trace = operating on 30 meters (10.125 MHz)
Peak gain = 0.69 dBi at 24.7° take-off angle
Dotted trace = operating on 20 meters (14.175 MHz)
Peak gain = 1.28 dBi at 20.5° take-off angle

40 meter band. As was seen previously, the peak gain and the input resistance rise continually as the operating frequency is increased, while the elevation angle falls by several degrees. Figure 4 gives us the principal radiation patterns. The plots for 40 and 30 meters resemble each other, while that for 20 meters is clearly more compressed, peaking at a lower take-off angle.

Results for a Quarter Wavelength 30 Meter Vertical

A computer model of a full-size λ/4 monopole was then created for operation on the 30 meter band, again using lengths of tapered aluminum tubing. The diameter of the largest section is 1.5 inches, with subsequent pieces gradually reduced in size to a final value of 1 inch at the tip. The length of the 60 buried no. 14 AWG aluminum radials is 24.286 feet, or λ/4 at 10.125 MHz.

Table 4

Performance of a full-size quarter-wavelength vertical monopole antenna designed for 60 meters, when used on that band and also on both 40 and 30 meters. The monopole is built from lengths of tapered metal tubing (diameter varies from 3 inches to 1.5 inches), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 45.816 feet (0.25 λ at 5367 kHz). The vertical element is tuned to resonance at 5367 kHz, with a resulting antenna height of 44.884 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND		
	60 Meters	*40 Meters*	*30 Meters*
Operating Frequency (MHz)	5.367	7.15	10.125
Input Impedance (Ω)	37.2	112 + *j*182	898 − *j*68.1
SWR (50 Ω ref.)	1.34	8.48	18.1
Peak Gain (dBi) and Take-off Angle (°)	0.27 at 25.8	0.61 at 24.6	1.11 at 21.1
Gain (dBi) at 5° Take-off Angle	−5.80	−5.42	−4.43
Gain (dBi) at 10° Take-off Angle	−2.11	−1.68	−0.72
Gain (dBi) at 15° Take-off Angle	−0.61	−0.17	0.66
Gain (dBi) at 20° Take-off Angle	0.05	0.46	1.09
Half Power Beamwidth (°)	44.1	41.5	34.7
Efficiency (%)	32.9	34.5	34.8

Table 5

Performance of a full-size quarter-wavelength vertical monopole antenna designed for 40 meters, when used on that band and also on both 30 and 20 meters. The monopole is built from lengths of tapered metal tubing (diameter varies from 2 inches to 0.75 inch), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 34.39 feet (0.25 λ at 7150 kHz). The vertical element is tuned to resonance at 7150 kHz, with a resulting antenna height of 33.736 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND		
	40 Meters	*30 Meters*	*20 Meters*
Operating Frequency (MHz)	7.15	10.125	14.175
Input Impedance (Ω)	35.9	152 + *j*235	866 − *j*318
SWR (50 Ω ref.)	1.39	10.6	19.7
Peak Gain (dBi) and Take-off Angle (°)	0.29 at 26.4	0.69 at 24.7	1.28 at 20.5
Gain (dBi) at 5° Take-off Angle	−6.03	−5.49	−4.19
Gain (dBi) at 10° Take-off Angle	−2.25	−1.68	−0.48
Gain (dBi) at 15° Take-off Angle	−0.68	−0.13	0.89
Gain (dBi) at 20° Take-off Angle	0.03	0.53	1.28
Half Power Beamwidth (°)	44.2	41.3	33.0
Efficiency (%)	33.0	34.9	35.0

The over-all height of the vertical element was pruned to just 23.491 feet, providing resonance at mid-band. This antenna can be utilized effectively on five bands (30, 20, 17, 15, and 12 meters) and the outcome from the *EZNEC* simulations is posted in Table 6. We find (as expected) that raising the frequency yields a boost in the peak gain, accompanied by a lowering of the elevation angle of the main radiation lobe

Results for a Quarter Wavelength 20 Meter Vertical

Table 7 reveals what happens when a λ/4 monopole built from aluminum tubing is employed on the 20 meter band. Now the diameter of the tapered sections of aluminum tubing ranges from 1.0 inch to 0.75 inch, and the ground screen consists of sixty no. 14 AWG aluminum radials whose length is 17.347 feet (λ/4 at 14.175 MHz). The vertical element was adjusted to resonate the entire antenna system at this same frequency, leading to an overall height of 16.535 feet.
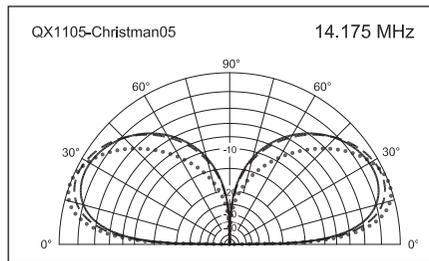
Figure 5 — Elevation-plane radiation patterns for a resonant ¼ λ 20 meter vertical antenna (made from tapered sections of tubing) when used on 20, 15, and 10 meters. Solid trace = operating on 20 meters (14.175 MHz) Peak gain = 0.56 dBi at 27.3° take-off angle Dashed trace = operating on 15 meters (21.225 MHz) Peak gain = 1.13 dBi at 25.2° take-off angle Dotted trace = operating on 10 meters (28.7 MHz) Peak gain = 1.48 dBi at 20.8° take-off angle

In Table 7, the outcome of the computer analysis is depicted for utilization on five amateur bands, from 20 through 10 meters. As was seen previously, the peak gain and the input resistance rise continually as the operating frequency is increased, while the take-off angle decreases by several degrees. Figure 5 summarizes the radiation-pattern plots for operation of the antenna on the 20, 15 and 10 meter bands. Notice that the lobe shapes are similar on 20 and 15 meters, but that for 10 meters is lower and more compressed.

Is a 5/8 Wavelength Antenna Really Superior?

In this part of the article, we will examine the electrical performance of vertical antennas as their height is varied from ¼ to ⅜ to ½ to ⅝ λ. Each antenna model is constructed from aluminum, using a single no. 10 AWG conductor for the monopole, along with 60 buried radials made from no. 14 AWG wire.

Table 6

Performance of a full-size quarter-wavelength vertical monopole antenna designed for 30 meters, when used on that band and also on 20, 17, 15 and 12 meters. The monopole is built from lengths of tapered metal tubing (diameter varies from 1.5 inches to 1.0 inch), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 24.286 feet (0.25 λ at 10.125 MHz). The vertical element is tuned to resonance at 10.125 MHz, with a resulting antenna height of 23.491 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND				
	30 Meters	*20 Meters*	*17 Meters*	*15 Meters*	*12 Meters*
Operating Frequency (MHz)	10.125	14.175	18.118	21.225	24.940
Input Impedance (Ω)	33.0	135 + *j*208	658 + *j*299	571 −*j*468	93.9 −*j*282
SWR (50 Ω ref.)	1.52	9.40	15.9	19.1	19.3
Peak Gain (dBi) and Take-Off Angle (°)	0.52 at 26.9	0.95 at 25.3	1.40 at 22.8	1.65 at 19.7	2.02 at 16.2
Gain(dBi) at 5° Take-off Angle	−6.02	−5.40	−4.54	−3.69	−2.42
Gain (dBi) at 10° Take-off Angle	−2.16	−1.54	−0.73	0.01	1.06
Gain (dBi) at 15° Take-off Angle	−0.54	0.05	0.77	1.34	2.00
Gain (dBi)at 20° Take-off Angle	0.22	0.75	1.33	1.64	1.75
Half Power Beamwidth (°)	44.5	42.0	37.6	30.9	24.3
Efficiency (%)	34.9	37.4	39.1	36.2	37.0

Table 7

Performance of a full-size quarter-wavelength vertical monopole antenna designed for 20 meters, when used on that band and also on 17, 15, 12 and 10 meters. The monopole is built from lengths of tapered metal tubing (diameter varies from 1 inch to 0.75 inch), while the ground system is composed of 60 no. 14 AWG wire radials, whose length is 17.347 feet (0.25 λ at 14.175 MHz). The vertical element is tuned to resonance at 14.175 MHz, with a resulting antenna height of 16.535 feet. All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	BAND				
	20 Meters	*17 Meters*	*15 Meters*	*12 Meters*	*10 Meters*
Operating Frequency (MHz)	14.175	18.118	21.225	24.940	28.7
Input Impedance (Ω)	32.3	79.5 + *j*147	174 + *j*278	479 + *j*403	953 − *j*66
SWR (50 Ω ref.)	1.55	7.53	12.6	16.4	19.2
Peak Gain (dBi) and Take-Off Angle (°)	0.56 at 27.3	0.90 at 26.3	1.13 at 25.2	1.33 at 23.5	1.48 at 20.8
Gain (dBi)at 5° Take-off Angle	−6.12	−5.64	−5.25	−4.76	−4.12
Gain (dBi) at 10° Take-off Angle	−2.22	−1.75	−1.37	−0.92	−0.36
Gain (dBi) at 15° Take-off Angle	−0.57	−0.12	0.23	0.62	1.05
Gain (dBi) at 20° Take-off Angle	0.22	0.63	0.93	1.22	1.47
Half Power Beamwidth (°)	44.7	43.2	41.8	38.8	33.3
Efficiency (%)	35.2	37.4	38.9	39.1	36.7

Table 8

Performance comparison between vertical antenna systems of varying height, when operating on 80 meters at a frequency of 3650 kHz. The monopoles are made from no. 10 AWG wire, with a ground screen composed of 60 buried no. 14 AWG radials (radial length = monopole height). All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	¼ λ System	⅜ λ System	½ λ System	⅝ λ System
Monopole Height and Radial Length (ft)	67.368	101.05	134.74	168.42
Input Impedance (Ω)	$41.4 + j\,24.4$	$229 + j\,605$	$2324 - j\,1425$	$86.1 - j\,479$
SWR (50 Ω ref.)	1.75	36.8	64.0	55.5
Peak Gain (dBi) and Take-off Angle (°)	0.39 at 24.7	0.79 at 21.7	0.96 at 17.6	0.42 at 13.3
Gain (dBi) at 5° Take-off Angle	−5.21	−4.34	−3.42	−2.81
Gain (dBi) at 10° Take-off Angle	−1.70	−0.91	−0.14	0.06
Gain (dBi) at 15° Take-off Angle	−0.32	0.35	0.85	0.34
Gain (dBi) at 20o Take-off Angle	0.25	0.76	0.89	−0.63
Half Power Beamwidth (°)	43.7	38.0	29.0	20.3
Efficiency (%)	33.8	34.3	29.6	29.8

Table 9

Performance comparison between vertical antenna systems of varying height, when operating on 40 meters at a frequency of 7150 kHz. The monopoles are made from no. 10 AWG wire, with a ground screen composed of 60 buried no. 14 AWG radials (radial length = monopole height). All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	¼ λ System	⅜ λ System	½ λ System	⅝ λ System
Monopole Height and Radial Length (ft)	34.391	51.586	68.781	85.976
Input Impedance (Ω)	$39.9 + j\,25.0$	$235 + j\,570$	$1937 - j\,1247$	$81.9 - j\,436$
SWR (50 Ω ref.)	1.81	32.5	54.8	48.7
Peak Gain (dBi) and Take-off Angle (°)	0.15 at 26.2	0.68 at 23.3	0.89 at 19.1	0.68 at 14.5
Gain (dBi) at 5° Take-off Angle	−6.15	−5.15	−4.13	−3.12
Gain (dBi) at 10° Take-off Angle	−2.38	−1.44	−0.56	0.08
Gain (dBi) at 15° Take-off Angle	−0.82	0.02	0.66	0.67
Gain (dBi) at 20° Take-off Angle	−0.11	0.59	0.88	0.04
Half Power Beamwidth (°)	44.1	39.3	30.7	22.3
Efficiency (%)	31.9	34.0	30.4	31.7

The length of these radials will always be adjusted so they are equal to the height of the vertical element.

80 Meters

The findings for 80 meter operation are revealed in Table 8. Peak gain climbs smoothly as the height of the monopole increases from ¼ to ⅜ to ½ λ, but actually diminishes somewhat for the ⅝ λ antenna. The elevation angle where maximum gain occurs, however, continually falls as the system is made taller. Figure 6 shows the principal radiation-pattern plots. Notice that the ⅝ λ element generates *slightly* more gain than the λ/2 vertical at the very lowest take-off angles. For general DX applications on this band, it seems that a monopole height in the vicinity of λ/2 may be the best choice, because of the large amount of high-angle radiation generated by the ⅝ λ element.

40 Meters

Table 9 lists the results for operation on the 40 meter band. As was true on 80, here the maximum-gain value rises continually as the height of the antenna increases from ¼ to ⅜ to ½ λ but once again we see a small reduction for the ⅝ λ system. Nevertheless, the elevation angle where maximum gain occurs constantly decreases as the element is made

taller. Plots of the principal-plane radiation patterns are posted in Figure 7. Here, we can see that the ⅝ λ vertical is superior to the λ/2 monopole at *any* take-off angle up to 15°. According to *EZNEC*, at an elevation angle of 5° the ⅝ λ antenna performs better than

Figure 6 — Elevation-plane radiation patterns for vertical antennas of various heights, when operating on 80 meters (3.65 MHz). The monopoles are made from no. 10 AWG wire and the ground screens are composed of 60 buried no. 14 AWG radials (radial length = monopole height).
Solid trace = ¼ λ system
Peak gain = 0.39 dBi at 24.7° take-off angle
Dashed trace = ⅜ λ system
Peak gain = 0.79 dBi at 21.7° take-off angle
Dotted trace = ½ λ system
Peak gain = 0.96 dBi at 17.6° take-off angle
Dash-dotted trace = ⅝ λ system
Peak gain = 0.42 dBi at 13.3° take-off angle

the λ/2 version by a full decibel, so the taller system appears to be the best selection for DX work on this band.

20 Meters

The outcome for operation on 20 meters
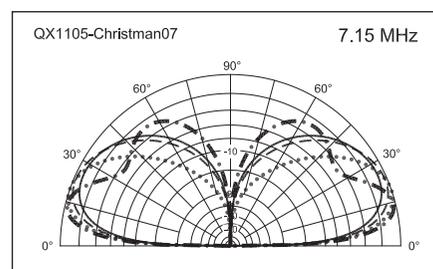
Figure 7 — Elevation-plane radiation patterns for vertical antennas of various heights, when operating on 40 meters (7.15 MHz). The monopoles are made from no. 10 AWG wire and the ground screens are composed of 60 buried no. 14 AWG radials (radial length = monopole height).
Solid trace = ¼ λ system
Peak gain = 0.15 dBi at 26.2° take-off angle
Dashed trace = ⅜ λ system
Peak gain = 0.68 dBi at 23.3° take-off angle
Dotted trace = ½ λ system
Peak gain = 0.89 dBi at 19.1° take-off angle
Dash-dotted trace = ⅝ λ system
Peak gain = 0.68 dBi at 14.5° take-off angle

Table 10

Performance comparison between vertical antenna systems of varying height, when operating on 20 meters at a frequency of 14.175 MHz. The monopoles are made from no. 10 AWG wire, with a ground screen composed of 60 buried no. 14 AWG radials (radial length = monopole height). All conductors are aluminum, and the soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	¼ λ System	⅜ λ System	½ λ System	⅝ λ System
Monopole Height and Radial Length (ft)	17.347	26.020	34.694	43.367
Input Impedance (Ω)	39.0 + j28.4	247 + j536	1595 – j1070	77.4 – j392
SWR (50 Ω ref.)	1.97	28.3	46.3	41.8
Peak Gain (dBi) and Take-off Angle (°)	0.29 at 27.1	0.91 at 24.3	1.16 at 19.9	1.21 at 15.0
Gain (dBi) at 5° Take-off Angle	–6.35	–5.28	–4.18	–2.86
Gain (dBi) at 10° Take-off Angle	–2.46	–1.45	–0.49	0.48
Gain (dBi) at 15° Take-off Angle	–0.81	0.11	0.84	1.21
Gain (dBi) at 20° Take-off Angle	–0.04	0.76	1.16	0.70
Half Power Beamwidth (°)	44.4	40.4	31.5	22.8
Efficiency (%)	32.9	36.3	32.9	34.7

is given in Table 10. This time, the peak gain *always* rises whenever the system is made taller, accompanied by a simultaneous decrease in the take-off angle where maximum gain occurs. Figure 8 depicts the four elevation-plane radiation patterns, and an examination of these plots indicates that a ⅝ λ element yields the most desirable performance for 20 meter DX applications.

Losses

Notice that the input resistance for all of the resonant λ/4 elements is generally in the low-to-mid 30 Ω range, providing SWR values on the order of 1.5:1 in most cases. The SWR is much higher when using a taller monopole, however, so an impedance-matching network of some kind must be included as part of the antenna system. Such networks will dissipate a certain amount of power, mainly in the inductor(s), but this factor has been omitted from the present review. When this power loss is properly taken into account, the actual improvement in performance achieved by using a taller radiator will be less than what is shown in the tables.

I mentioned at the beginning of this article that all of the models were constructed using aluminum conductors. Table 11 illustrates what happens when either copper or zinc is substituted. *EZNEC* analysis indicates that copper works slightly better, while zinc is a bit worse. Variations in peak antenna gain amount to no more than 0.06 dB, and the efficiency changes by just 0.6% at most. The greatest impact takes place on 80 meters.

Conclusions

This article has discussed the use of λ/4 vertical antennas on bands that are higher in frequency than the one for which they were primarily designed. Computer analysis

Table 11

Antenna performance as a function of the type of metal employed. In each case, the monopole height and radial length are fixed at 0.25 λ. The monopole is constructed of no. 10 AWG wire and the 60 buried radials are made from no. 14 AWG wire. The soil is "average" (conductivity = 0.005 siemens/meter and dielectric constant = 13).

	Aluminum	Copper	Zinc
80 Meters (3.65 MHz)			
Peak Gain (dBi) and Take-off Angle (°)	0.39 at 24.7	0.43 at 24.7	0.37 at 24.7
Efficiency (%)	33.8	34.2	33.6
40 Meters (7.15 MHz)			
Peak Gain (dBi) and Take-off Angle (°)	0.15 at 26.2	0.18 at 26.3	0.13 at 26.3
Efficiency (%)	31.9	32.1	31.8
20 Meters (14.175 MHz)			
Peak Gain (dBi) and Take-off Angle (°)	0.29 at 27.1	0.31 at 27.0	0.27 at 27.1
Efficiency (%)	32.9	33.1	32.8
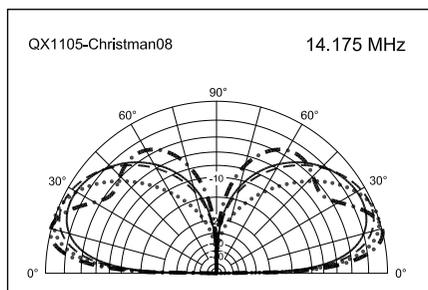
**Figure 8 — Elevation-plane radiation patterns for vertical antennas of various heights, when operating on 20 meters (14 175 MHz). The monopoles are made from no. 10 AWG wire and the ground screens are composed of 60 buried no. 14 AWG radials (radial length = monopole height).
Solid trace = ¼ λ system
Peak gain = 0.29 dBi at 27.1° take-off angle
Dashed trace = ⅜ λ system
Peak gain = 0.91 dBi at 24.3° take-off angle
Dotted trace = ½ λ system
Peak gain = 1.16 dBi at 19.9° take-off angle
Dash-dotted trace = ⅝ λ system
Peak gain = 1.21 dBi at 15.0° take-off angle**

reveals that an extended-length monopole can provide additional gain at a lower elevation angle, but the resulting input impedance is often far-removed from 50 Ω, so an impedance-matching network will be needed in order to present a low SWR to the transmitter. The losses that are generally present in such networks may partially negate the increase in gain generated by the taller monopole. Being able to use the same antenna structure on multiple bands, however, yields definite savings in terms of cost and real estate.

The ⅝ λ vertical has a reputation for generating a lot of extra gain at a much-lower take-off angle, when compared to a conventional λ/4 monopole. *EZNEC* studies suggest that a ⅝ λ antenna may be worthwhile on 40 and 20 meters (provided that ⅝ λ radials are also utilized in the ground screen), but the optimum monopole height on 80 meters may be closer to λ/2.

Kazimierz "Kai" Siwiak, KE4PT

10988 NW 14th St, Coral Springs, FL 33071; ke4pt@amsat.org

An Optimum Height for an Elevated HF Antenna

What is the best height for your antenna?
The author considers factors that can help you decide.

There are two ways to think about antenna and propagation problems in linear media: in transmit mode and in receive mode. By the reciprocity theorem both methods will predict the same performance. We will view the problem of finding an optimum height for HF antennas in receive mode rather than in transmit mode, because this reveals very interesting insights. For example, the field-strength at the receiving location is the result of an interference pattern between waves that arrive by a direct path added to the wave reflected from the earth's surface. The addition of these two waves results in a standing wave versus height for the field strength at the receiving location. Because this vertical standing wave has peaks and can have deep nulls, there is an optimum placement for an antenna. In the equivalent transmit mode point of view, far-field transmit patterns are calculated as an interference pattern between the direct wave and a ground reflected wave, but as *The ARRL Antenna Book* explains, that point of view obscures the physical meaning of "take-off" angle, so we can't directly appreciate what happens when an antenna is elevated.[1] By viewing the problem in receive mode, however, we see, among other things, that waves arriving from the lowest arrival angle do not always result in the best link margin to a DX station. We can also see that low antennas can work surprisingly well for DX, and that the best height for vertically polarized antennas is not the same as for horizontally polarized antennas.

With this analysis it is easy show that the optimum antenna height depends on frequency, polarization, properties of the earth at the reflection point, and on the arrival angle from the wave source in the ionosphere. While surface roughness is considered, there is also a terrain dependence, which for simplicity will not be considered here; see Dean Straw's terrain analysis program HFTA in the 21st edition of *The ARRL Antenna Book*. Furthermore, since the apparent wave earth reflection point is usually distant from the antenna, it is not important what the earth looks like directly under an elevated antenna. What is important is the earth's properties at the reflection point — typically hundreds to thousands of meters distant from the tower. This is an idealized problem where we allow for surface roughness, but we assume an earth that is smooth enough so that we can apply spherical earth geometry.

We begin by laying a foundation based on a spherical earth geometry for the propagation of waves to the receiving location. The reflection properties of ground and sea water are shown to affect how the

reflected wave combines in constructive and destructive interference with the direct wave. Optimum heights are found for desired ranges of arrival angles and for multiple bands. Finally, path link margins are estimated for multi-hop propagation. We discover that a range of "take-off" angles must be accommodated for optimum performance.

Spherical Earth Geometry

Because we are dealing with distances that approach the earth's horizon, we calculate the direct and earth-reflected paths using spherical-earth reflection geometry. The solution to the spherical earth geometry given in Chapter 2 of M. I. Skolnik's *Radar Handbook* involves a cubic equation to find the arc distance G_b to the reflection point.[2]

$$2G_b^3 - 3GG_b^2 + \left[G^2 - 2a_e \left(h_{ant} + h_i \right) \right] G_b + 2a_e h_{ant} G = 0$$

[Eq 1]

where:
h_{ant} is the height at the receiving antenna,
a_e is the earth's radius,
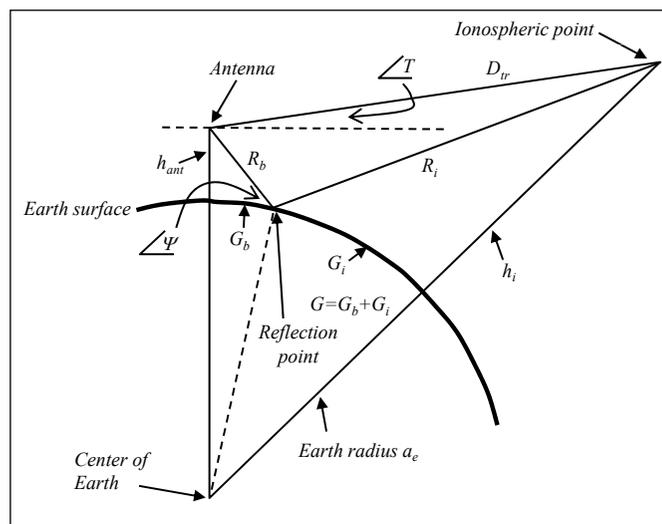
Figure 1 — Spherical earth geometry, shown with an exaggerated height dimension. Source: based on information from *Radar Handbook* (see Note 2).

and the distances G and G_b are functions of the angle T between the local horizon and the direction to the wave source point at height h_i in the ionosphere. Figure 1 shows the spherical earth reflection geometry and identifies all of the parameters.

The angle T is also called the "take off angle" and the "local elevation angle." See the ARRL website files update to *The ARRL Antenna Book*.[3] The direct wave arrives along path D_{ir}, and the reflected path includes distance R_i from the ionosphere to the earth reflection point and R_b from the reflection point to the receiving location. The reflection occurs at the arc distance G_b from the base of the antenna tower, and as the direct wave arrival angle T deceases, then the arc distance to the reflection point increases. Our chief concern is with the difference in the path lengths,

$$\Delta R = (R_b + R_i - D_{ir}) \qquad \text{[Eq 2]}$$

and with the surface reflection coefficient at the reflection point because these determine the nature of the field variation versus height, h_{ant}.

Reflection Coefficients and Combined Waves

The plane wave reflection coefficients Γ_H for horizontal and Γ_V for vertical polarization are used to find the reflection from land or sea on a spherical earth. (See Chapter 6 of *Radiowave Propagation and Antennas for Personal Communications*.[4]) The reflection coefficient is modified by the divergence factor D and surface roughness S_r factor. The wave divergence factor is:

$$D = \left[1 + \frac{2G_b G_i}{a_e G \sin \psi} \right]^{-1/2} \qquad \text{[Eq 3]}$$

where ψ is the angle of incidence on the earth's surface. The surface roughness factor is:

$$S_r = \exp(-r) I_0(r); \quad r = 2\left(k h_{sd} \sin(\psi)\right)^2 \qquad \text{[Eq 4]}$$

where:
I_0 is the modified Bessel function
$k = 2\pi f / c$ is the wave number
f is the signal frequency in Hz
c is the speed of light in m/s.

The roughness factor for the reflected wave is based on a roughness factor originally derived for a ratio of rough-sea to smooth-sea reflection, and is applied here generally to an earth reflection. The surface roughness parameter h_{sd} is the standard deviation of the surface height distribution in the reflection region. The complete reflection coefficients are thus $\Gamma_H S_r D$ and $\Gamma_V S_r D$ for a rough spherical earth. The reflected term fields are also multiplied by $d = D_{tr} / (R_b + R_i)$ to account for the difference in free space loss due to the differential distance between the direct and reflected waves.

For this study we will assume that horizontally polarized power is added to vertically polarized power in a ratio, P_{HV}. For substantially horizontally polarized waves, P_{HV} is chosen here to be between 10 and 20, and for substantially vertically polarized waves, P_{HV} is between 0.005 and 0.01. The polarization impurity primarily results in a slight reduction of the depths of nulls in the vertical standing wave patterns. The two polarization components are added as power because the polarization is decomposed by the ionosphere into elliptical polarization, (see *Ionospheric Radio Propagation*[5]) and reflections from a rough surface are generally random and time-variable. The expression for the signal power, P normalized to the free space value, of the combined waves at the receiving height, h_{ant} is:

$$P = \frac{P_{HV}\left[1 + \exp(-jk\Delta R)\Gamma_H S_r D d \right]^2 + \left[1 + \exp(-jk\Delta R)\Gamma_V S_r D d \right]^2}{1 + P_{HV}}$$

$$\text{[Eq 5]}$$

The unity terms in each of the brackets represent the direct wave amplitude, and the remaining terms are the reflected wave, each in ratio to the free space value. The phase difference, $k\Delta R$, along with the phase of the reflection coefficients conspire to produce the vertical standing wave pattern of the field strength at the receiving location. *This is before any antenna is placed at the receiving location*. Since the earth's radius is large compared with the height of the ionosphere, angles T and ψ are nearly the same value, despite the exaggerated view in Figure 1. Since antenna free space elevation patterns for a level antenna are essentially symmetrical in elevation about the local horizontal plane, the direct wave entering the antenna from angle T above the horizontal plane is weighted by the same antenna pattern gain value as the reflected wave entering the antenna from angle ψ below the horizontal plane. Note also that the earth's horizon is *slightly below* the elevated antenna horizontal plane.

Expected Angles of Arrival

We will be optimizing our solution over a desired range of arrival angles. Expected arrival angles T for waves from the ionosphere for HF Propagation are available in *The ARRL Antenna Book* product notes files on the ARRL website for HF (see Note 3). For example, the combined 80 m to 10 m arrival angle statistics between Florida (FL) or Massachusetts (MA) and all regions of the World are shown in Figure 2.

Those statistics show that half the arrival angles are less than 6°, and that 90% of the arrival angles are smaller than 16°. So for HF cases, we will confine our interest to arrival angles between 2 to 16°. Viewed in transmit mode, this is the *range* of "take-off" angles that must be accommodated. Similar curves may be derived for 6 m band sporadic-*E* propagation. Notably, in the July and August 2009 "World Above 50 MHz" *QST* column, Gene Zimmerman, W3ZZ, comments on the work of Joe Kraft, CT1HZE, suggesting that arrival angle probabilities for 6 m band sporadic-*E* are bimodal, with one peak at ~5° and another at ~10° with very little below 3° or 4° or above ~13° or 14°.[6, 7] Thus, arrival angles of 3° to 14° emerge as a range of interest for 6 m sporadic-*E* operations. Also see my article, "Optimum Height for an Elevated Communications Antenna," in *DUBUS* magazine.[8] While different from HF in the specifics, the angle ranges of interest are similar, and justify the range between 2° and 16°.

Location of the Reflection Point

The distance G_b to the reflection point on the earth's surface is solved by Equation 1 as a function of receiving point height. There is only a very
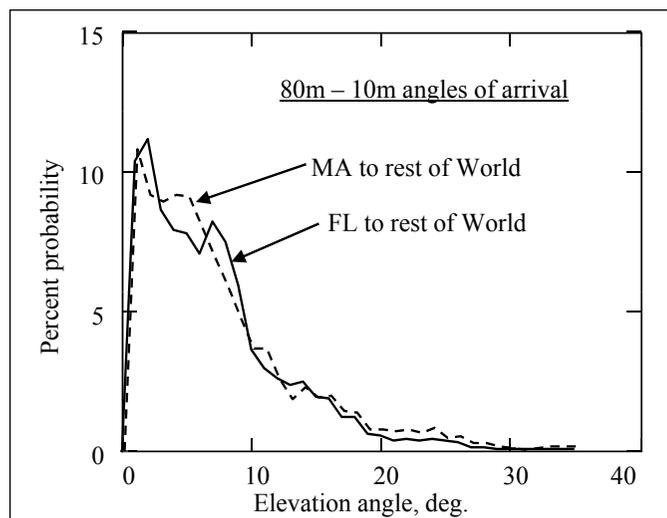
Figure 2 — Composite probability of arrival angles.
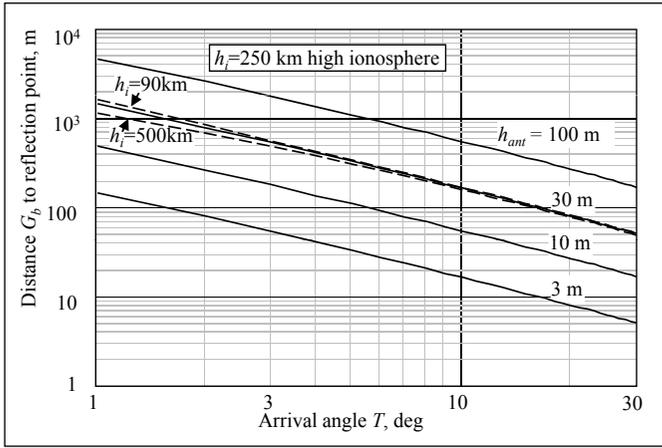
Figure 3 — Distance to the reflection point is tens to thousands of meters.
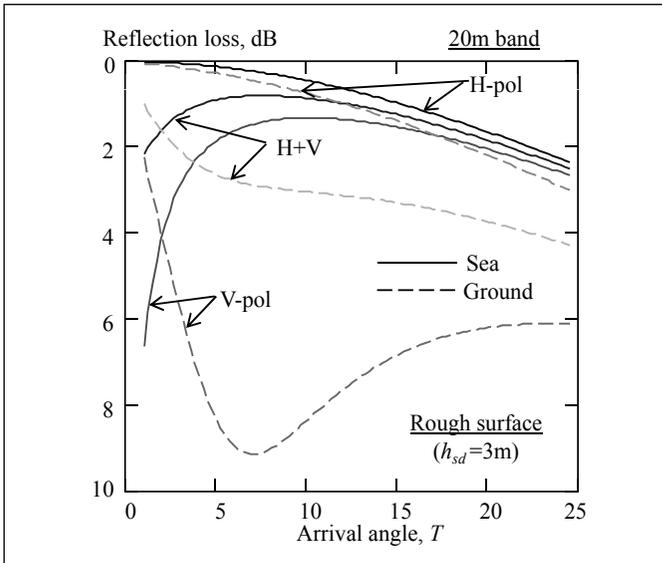
Figure 4 — Reflection coefficient with surface roughness, 20 m band.

weak dependency on the height of the ionosphere; heights from 90 km to as much as 500 km, the range of heights for the E, E_S, and F layers of the ionosphere, give very nearly the same geometrical result. There is, however, a strong dependency on the receiving height location. Figure 3 shows the distance to the reflection point versus the arrival angle for several receiving heights between 3 and 100 m with a 250 km high ionosphere. The 30 m high antenna distances are also shown (dashed lines) for 90 km and 500 km high ionosphere. *Since the reflection point is typically from a few kilometers to tens of meters away the ground immediately below the antenna does not affect elevated antenna performance.* A very good approximation to the reflection point distance is:

[Eq 6]

$$G_b = \frac{55 h_{ant}}{T}$$

where:

h_{ant} is the antenna height in meters
T is the arrival angle in degrees.

The reflection point given by Equation 6 is the same as for the transmit case; please see "The Effect of Ground in the Far Field" in Chapter 3 of *The ARRL Antenna Book* (see Note 1). It should be noted that transmit patterns computed in the presence of the ground often quoting a "take off angle," *implicitly assume that, the ground is flat to beyond the distance given by Equation 6*. Here, in contrast, recall that we have allowed for a ground roughness factor.

Earth Reflection Loss

The ground or sea reflection loss, L_{earth} in dB for *multiple hop paths* can be found by setting the direct wave "1" terms to zero in Equation 5 and expressing the result in decibels. Figure 4 shows the loss in the 20 m band for horizontal, vertical and a 50% mix of the polarization, for reflection from the sea and from a medium earth ($\varepsilon = 12$) versus the angle T. The reflection includes a surface roughness factor of 3 m. For $2 \leq T \leq 16°$ this reflection loss can amount to more than 1 dB for horizontal polarization, but as much as 9 dB for vertical polarization reflected from earth ground.

Optimum Antenna Height

We can now solve Equation 5 at various frequencies, polarizations, ground constants and as a function of the height of an antenna. The specific antenna pattern — that is, the free space pattern — is not important as long as the elevation plane beamwidth is sufficient
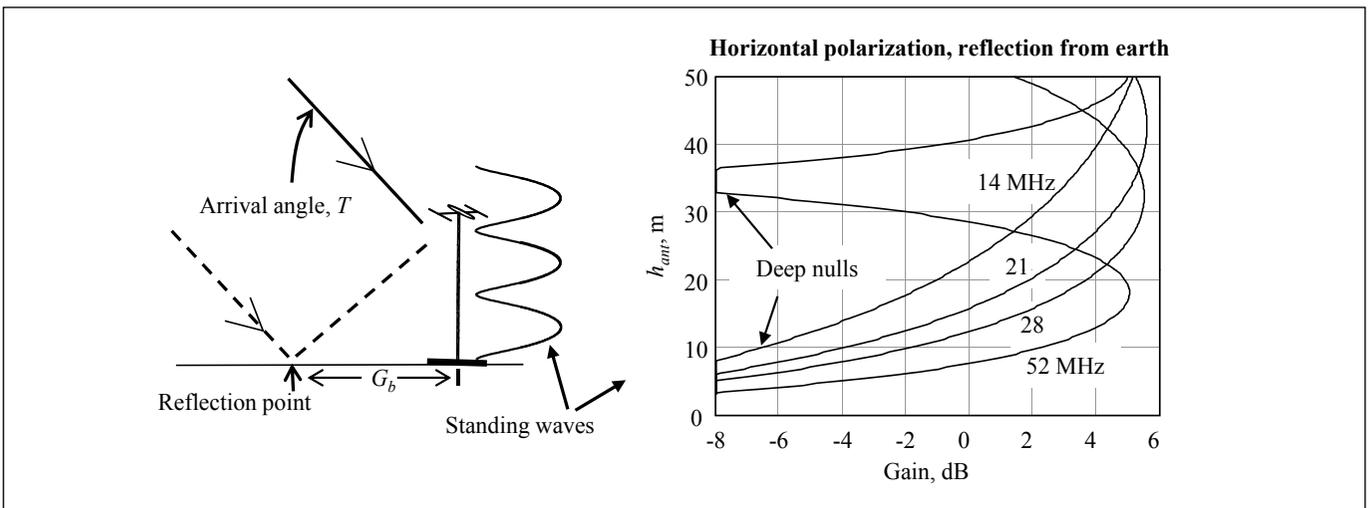
Figure 5 — Horizontal polarization (P_{HV} = 20), earth ground, T = 5°, roughness is 3 m.
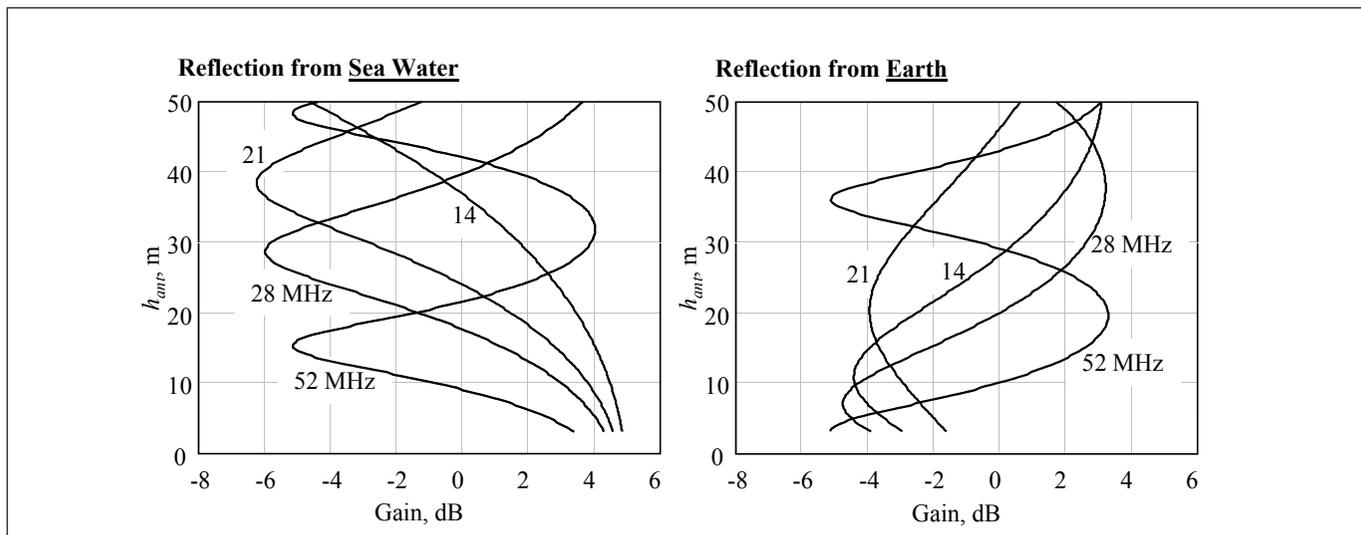
Figure 6 — Vertical polarization ($P_{HV} = 0.05$), $T = 5°$, roughness is 3 m, reflections from (left) sea water and from (right) earth ground.

to include the important angles of arrival, both above and below the local horizontal plane. We do note, however, that as the angle T increases, the waves arrive in pairs above and below the main beam peak, so that the full antenna gain for directive antennas cannot be always be realized — especially for very high gain (narrow elevation plane beamwidth) antennas.

Figure 5 shows the geometry and the calculated vertical standing wave patterns produced by the interaction of the direct and earth reflected waves for earth ground parameters $\varepsilon = 5$ and $\sigma = 0.005$ S/m. The standing wave peaks and nulls depend on frequency and on arrival angle, here 5°. This suggests placing the antenna at the signal peak, which is one definition of the optimum antenna height.

Results for horizontally polarized waves reflected from the sea differ primarily in the depth of nulls compared with earth ground reflected results of Figure 5. There are transmitter mode equivalents to the receive mode standing wave patterns shown in Figure 5. The transmit mode patterns are computed in the presence of a ground, and usually a peak "take-off angle" is identified; see for example Figure 3 in the companion article in the June 2011 issue of *QST*.[9] Clearly the transmit mode patterns do not make it easy to identify the best height for the antenna.

Figure 6 shows the vertical polarization performance for reflection from sea water $\varepsilon = 70.6$ and $\sigma = 4.54$ S/m, on the left and from ground with $\varepsilon = 5$ and $\sigma = 0.005$ S/m on the right. The saline water model is from *Radiowave Propagation and Antennas for Personal Communications* (see Note 4). The sea-reflected, vertically polarized case has an optimum at sea level. This is why vertically polarized antennas on the beach are so effective on some DXpeditions such as during the VP6DX operation. Note that the optimum heights per frequency for vertically polarized antennas with the reflection from earth ground are not the same as for horizontal polarization. Ground mounted vertical antennas with a reflection from earth ground will have negative height gains of –1 to –5 dB. The gains shown in Figures 5, 6 and 7 are in addition to any free space directive gain provided by the antenna system. Results in Figures 5 and 6 are exactly analogous to the results that have been predicted and measured to within a decibel at open air test sites in the 30 to 932 MHz range. See Section 6.3 in *Radiowave Propagation and Antennas for Personal Communications* (see Note 4).

Concentrating now on the 20 m band, Figure 7 shows field-strength signal levels relative to the free space value for reflections from the ground. *These are not antenna patterns but rather signal*
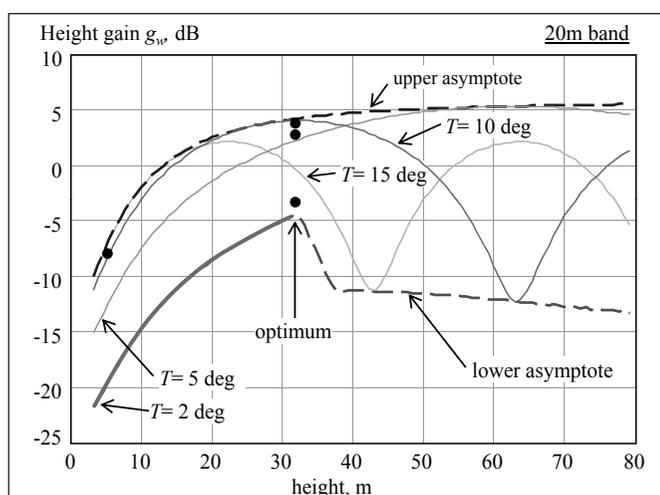
Figure 7 — Height gain for horizontal polarization in the 20 m band.

field strength levels that are then sampled by an antenna. The axes have been flipped compared with the previous figures. The upper dashed asymptote is the *maximum constructive interference* for the continuum of all arrival angles between 2 and 16°. Specific results for 2°, 5°, 10° and 15° are shown by the embedded curves. The lower dashed asymptote is defined by the *destructive interference* for the continuum of arrival angles. The lower asymptote intersects the 2° arrival angle curve at a cusp, which defines an optimum antenna height for that frequency. At that elevation, the height gain, g_w has the smallest variation versus the range of arrival angles, and its minimum gain value is the highest. When an antenna is placed there, the actual free space antenna gain, at the pattern elevation angle, T, adds to this field strength height gain. Antennas that are higher than the optimum height will encounter degraded performance at the higher angles of arrival because the nulls defining the lower asymptote to the right of the cusp are likely to be a factor. This is why in some cases a lower antenna can significantly outperform a higher antenna. If we had chosen a higher minimum required arrival angle, the optimum height would decrease. Similar curves can be drawn for other HF bands or combinations of bands, and optimum heights can be found.

Multiband Considerations

Since the geometry of the reflection point, including divergence and surface roughness, are fixed in physical dimensions, the vertical interference patterns don't quite scale with wavelength. Thus, the optimum height does not scale exactly with frequency. Some multiband Yagi beams can cover the 40 m to 6 m bands in a single structure. Raising and lowering such an antenna is not usually desirable, so knowing an overall optimum height could be very useful. A family of curves like the 20 m band curves in Figure 7 can be calculated for any frequency band or any combination of frequency bands.

One effective strategy for finding an overall optimum over multiple bands is to choose the best height for the highest frequency band of interest. That somewhat sacrifices the performance for the lowest arrival angles at the lower frequency bands, but more gently than the destructive interference loss of height gain for higher arrival angles if a higher antenna were chosen.

The optimum heights for various frequency bands between 7 and 54 MHz are shown in Figure 8. The three curves are for three different minimum angles, the upper curve shows optima for a 1° to 16° arrival angle range, the middle curve for 2° to 16°, and the lower curve for 3° to 16°. The middle curve slopes from about 1.5 to 1.6 wavelengths between 7 and 29 MHz.

If operation anywhere in the 10 m to 40 m bands is of equal interest, the "best" height works out to about 19.9 m. That height is suitable for arrival angles as low as 1° in the 10 m band, and is also suitable for angles above about 4° in the 20 m band. In the 40 m and 30 m bands the results are "best effort," but as will be shown in the next section, paths at higher arrival angles may exist, but with an increased number of earth-ionosphere hops. If the 20 m band is to be optimized, then the best height is about 32 m. If 6 m band operation is important then the optimum height is about 15.3 m. The heights

between about 15 m and 32 m (50 to 105 ft) emerge as a good range of compromise choices for multiband HF and 6 m band operations.

This analysis also provides some insight into the physical basis for the operation of phased Yagi antennas mounted at different heights on a tower. By combining the signals from the two or more Yagis using phase shifters, it is possible to enhance gain in the direct-wave path while minimizing the destructive interference from the earth reflection. Possibly significant performance improvement might be realized.

Path Link Considerations

Many details are important in calculating a path link at HF, but for illustration here we examine a simplified path where both ends of the link are located on relatively flat (but not smooth) terrain, and the ionosphere and earth are suitable for the needed reflections along the path. Path link margin depends on the height of the ionosphere, h_i as well as on the arrival angle, T. Figure 9 shows the hop distances for several ionospheric heights as a function of the arrival angle over a spherical earth. For our example we will assume that the ionospheric refraction and reflection occurs at an effective height of 250 km. So a 10,000 km path might be traversed with 3, 4 or 5 hops, each 3,333 km or 2,500 km or 2,000 km respectively. Other paths are possible as well, as Davies described in *Ionospheric Radio Propagation* (see Note 5). The three different hops are marked by the shaded circle in Figure 9, with corresponding marks in Figure 7. Different hop distances mean different arrival angles, which affects the total path loss.

The wave interference gain, or height gain, g_w in dB shown in Figure 7 applies to each end of the link. Ionospheric reflection/refraction loss is L_{ion} in dB and can be as little as 2 to 5 dB.[10] In this simplified example, we will use 3 dB to account for polarization decomposition, as described by Davies (see Note 5). The free space loss is $27.6 + 20 \log(2 D_{tr} \times f)$ dB for one hop, where the frequency,
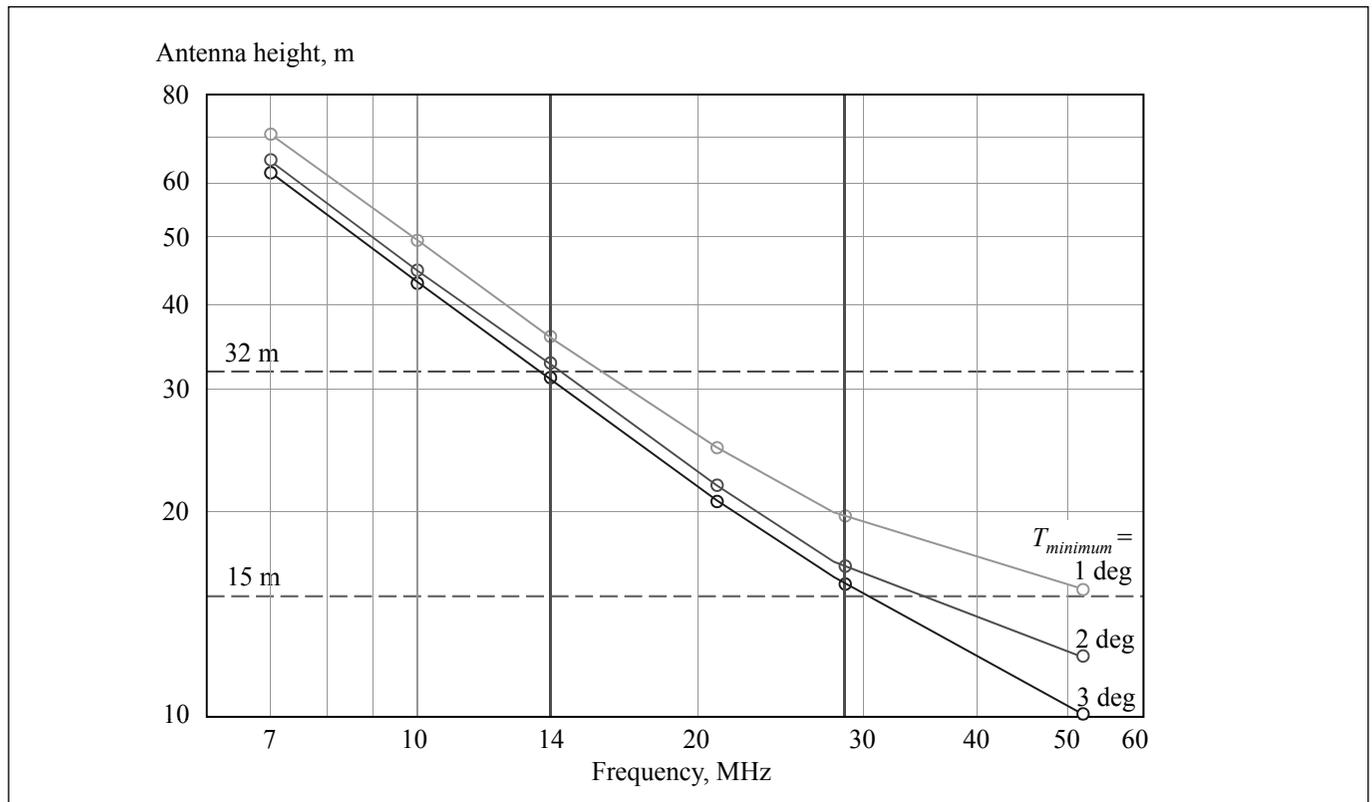
Figure 8 — Optimum antenna heights over even terrain for various frequencies.

Table 1

Path Losses in a 10,000 km Path for Different Numbers of Hops.

Hops	T (deg)	First hop loss (dB)	Height gain (dB)	Rest of hops loss (dB)	PL (dB)	S-units
3	2.8	[126.1 + 3]	–[–4 – 4]	{ 9.6 + 7.1}	153.8	3.6
4	6.9	[123.7 + 3]	–[+3 + 3]	{10.6 + 8.1 + 7.1}	146.4	4.8
5	10.4	[121.8 + 3]	–[+4 + 4]	{11.7 + 9.2 + 8.2 + 7.6}	157.8	2.9
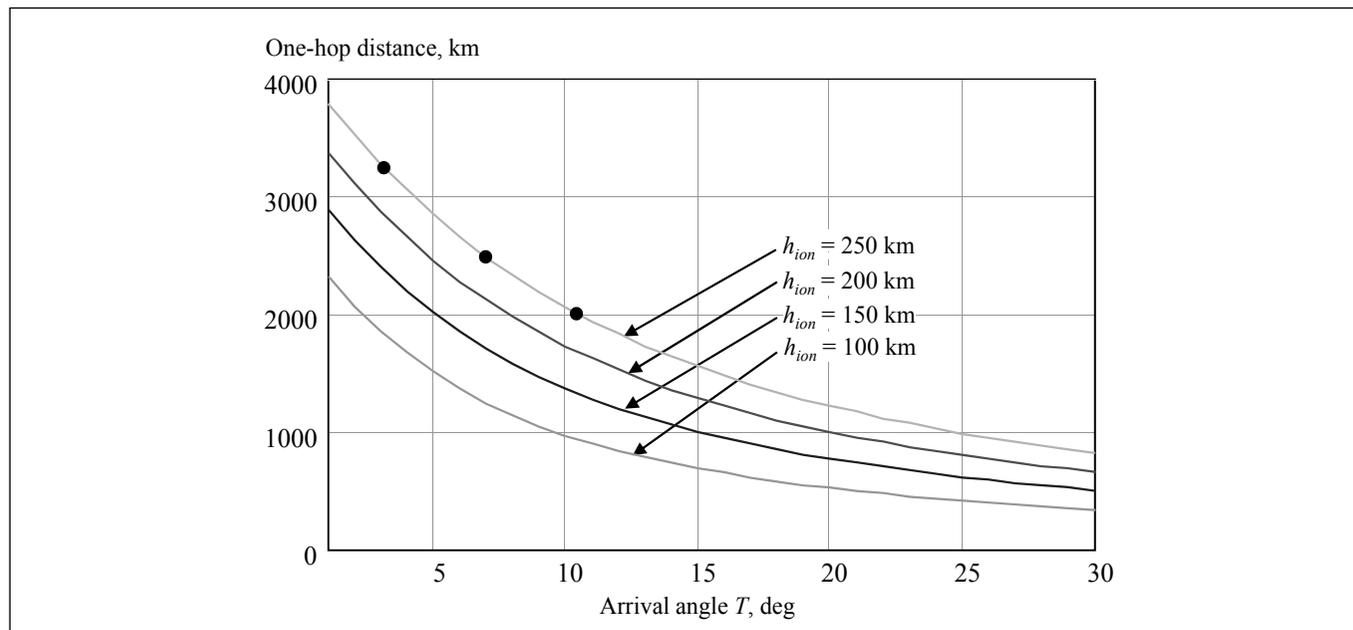
Figure 9 — Hop distances, with the 3, 4, and 5 hop points marked for a 10,000 km path.

f, is in MHz and the distance, D_{tr} is in meters. Each additional j^{th} hop adds an incremental free space loss, an earth reflection loss, $L_{earth,j}$ (from Figure 4), and another ionospheric reflection loss, $L_{ion,j}$. The path loss for *n* hops is written in Equation 7 so that the bracketed terms are for a single hop or first hop, including wave interference at the link ends *A* and *B*. The braces contain additional losses for hops 2 through *n* if present.

$$L_{path} = \left[27.6 + 20\log(2D_{tr}f_{MHz}) + L_{ion} - g_{w,A} - g_{w,B} \right] + \cdots$$

$$+ \left\{ \sum_{j=2}^{n} \left(20\log\left(\frac{j}{j-1}\right) + L_{ion,\, j} + L_{earth,\, j} \right) \right\}$$

[Eq 7]

Our example path in the 20 m band with a 250 km effective ionospheric height might require 3 to 5 or more hops to traverse a 10,000 km path. The various gains and losses for this idealized example are listed in Table 1. In general, several of these as well as other possible paths will exist, causing fading and signal variations as the ionosphere changes. Table 1 shows the path losses and estimated received S-units for 50 W transmitted power (approximately 100 W PEP for CW or processed SSB) and with 32 m high dipoles at each end. Gain antennas will improve signals in proportion to the antenna gains. The bracketed and braced terms in Table 1 correspond to the same terms in Equation 7.

Notice that the four-hop path has a stronger signal by over an

S-unit more than the example three-hop path because the increased height gains g_w of a combined 8 dB at the higher arrival angle (the difference between the top and bottom solid circles at the optimum height in Figure 7) at both ends of the link more than compensate for the additional reflection losses of an additional hop. The height gain is the intersection of the arrival angle, *T*, with the antenna height in Figure 7. The four-hop 6.9° arrival angle results in less destructive interference by 7 dB at each end of the link than the three-hop 2.8° arrival angle. *The lowest arrival angle path is not always the best!* Agonizing over a lower "take-off angle" is futile. This effect justifies a compromise lower limit for the angle of arrival at lower frequencies when choosing a compromise height for a multiband antenna. The five-hop path suffers additional earth and ionospheric reflection losses, but still results in a respectable S = 2.9 signal.

Suppose that the antenna at one end of the link is lowered to 5m. The height gain, g_w becomes –17 dB for the 2.8° three-hop path, so that path is not viable. The gain is –8 dB for the four-hop path, however, which is 12 dB lower than at the optimum height, resulting in an S = 1 reading. That is still a –115 dBm signal, which is suitable for CW as well as SSB. This result helps to explain the occasional spectacular DX results possible from low and indoor attic antennas.[11] If the arrival angle is, say >5°, the low antenna captures signals that are not dramatically worse than from a high antenna. Indeed, KE4PT has earned WAS-TPA and DXCC, now with 200 confirmed entities as well as a 6 meter VUCC from southern Florida, using just an indoor antenna.

Uncertainties in the ionospheric reflection/refraction loss

increase as the number of hops increases, and Equation 7 represents a best case value. Link reliability can be estimated by attaching variances to the several propagation loss components and by using the method of Hagn described in Section 8.4 of *Radiowave Propagation and Antennas for Personal Communications* (see Note 4).

Summary and Conclusions

Constructive and destructive wave interference from a direct path and an earth reflected path causes a vertical standing wave at the antenna location. The standing wave pattern details depend on the wave angle of arrival, polarization, on whether the reflection point was ground or sea water,

and on the terrain profile (not considered here). Optimum antenna heights are largely governed by the lowest arrival angle deemed important at the highest desired frequency. Antennas that are placed too high can suffer from significant wave destructive interference at desired higher arrival angles. The earth reflection point is typically several kilometers away for low arrival angles, but can be tens of meters for very high arrival angles, so the condition of the ground immediately below an elevated antenna is of little importance. Because height gain can be significantly greater for higher arrival angles, the lowest arrival angle path (fewest hops) does not always result in the best link margin for paths that can be closed with different

numbers of earth-ionosphere hops. Optimum height is 1.5 to 1.6 wavelengths for any one band, or a compromise height can be found for a multiband antenna operating over several bands by using the optimum for the highest frequency. Keeping in mind that this analysis was limited to rough, but not locally mountainous earth, nor a dense urban region, antenna heights in the range of 15 m to 32 m (50 to 105 ft) are found to be reasonable compromise choices for multiband antennas operating from a fixed height.

Kazimierz (Kai) Siwiak, KE4PT, earned a PhD from Florida Atlantic University, Boca Raton, FL, and his BSEE and MSEE from the Polytechnic Institute of Brooklyn, Brooklyn, NY, specializing in antennas and propagation. He founded TimeDerivative Inc., a wireless technology consultancy in 2003. He is a registered Professional Engineer and Senior Member of IEEE. Dr. Siwiak holds 38 US patents, has authored many peer-reviewed papers, four textbooks, and has contributed chapters to other books. His work appears in ARRL publications and in QST. He holds an Extra Class Amateur Radio operator license and is a life member of AMSAT and a member of ARRL, where he serves on the RF Safety Committee. He is also an ARRL Technical Advisor. He is an avid DXer, and was involved with SAREX (Space Amateur Radio Experiment) as a team member, including many SAREX operations and school contacts. His interests include flying (instrument and multiengine commercial pilot), hiking and camping.

Notes

[1] R. Dean Straw, N6BV, Ed., *The ARRL Antenna Book*, 21st Edition, 2007.

[2] M. I. Skolnik, *Radar Handbook*, Second Edition, McGraw-Hill Professional, 1990.

[3] See the *Antenna Book* tab at **www.arrl.org/product-notes**, updated statistical elevation-angle files, accessed 29 December 2010.

[4] K. Siwiak, Y. Bahreini, *Radiowave Propagation and Antennas for Personal Communications*, Third Edition, Artech House, Norwood MA, 2007.

[5] K. Davies, *Ionospheric Radio Propagation*, National Bureau of Standards Monograph 80, Washington DC, 1965.

[6] Eugene Zimmerman, W3ZZ, "Long Distance E$_s$ Propagation on 50 MHz at Solar Cycle Minimum — Part 1," The World Above 50 MHz, *QST*, July 2009, p 88.

[7] Eugene Zimmerman, W3ZZ, "Long Distance E$_s$ Propagation on 50 MHz at Solar Cycle Minimum — Part 2," The World Above 50 MHz, *QST*, August 2009, p 86.

[8] K. Siwiak, "Optimum Height for an Elevated Communications Antenna," *DUBUS Magazine*, Vol. 39, 3rd Quarter 2010, pp 86-99.

[9] K. Siwiak, "What's the Optimum Height for an HF Antenna?," *QST* June 2011.

[10] Recommendation ITU-R P.534-4.

[11] K. Siwiak, KE4PT, "You Can Enjoy DXing with a Modest Station," *CQ Amateur Radio*,

QEX

Alan Payne, G3RBJ

Laurel Bank, Sand Road, Wedmore BS28 4BZ, Great Britain; paynealpayne@aol.com

A New Theory For the Self Resonance, Inductance and Loss of Single Layer Coils

A new theory is presented that explains the self resonance in coils but without reference to "self capacitance." This new theory treats the coil as a transmission line, and is able to explain the changes that occur in both the inductance and loss with increasing frequency.

Introduction

All coils show a self resonant frequency (SRF) rather like a parallel tuned circuit. Knowledge of the SRF is important because as this frequency is approached, the parameters of the coil change, including its inductance, resistance and Q.

The conventional theory to explain these effects assumes that the coil has self capacitance, which along with the coil inductance produces this resonance, but nowhere in the coil can this capacitance be measured or deduced. Because of this it has been suggested over the years that the resonance could be due to a transmission-line mode on the wire, and this paper produces such a theory. This proposes that a coil has two modes, the classical "coil" mode which dominates at low frequencies and a transmission-line mode which dominates at high frequencies.

A significant result of the new theory is that the inductance and resistance really do increase with frequency, rather than being an apparent increase as predicted by the self-capacitance theory.

Measurements

The theoretical analysis in this article has been supported by measurements on the following coil. It had a length (l_c) of 21 mm, and was wound with 17 turns of enameled copper wire of diameter 0.5 mm, onto a tubular plastic form 0.5 mm thick and of outside diameter 11.2 mm. The ratio of the wire diameter to pitch, d_w / p, was therefore 0.4 and this was chosen because it is close

to the optimum value for minimum ac resistance. In other words, a larger diameter wire could actually increase the ac resistance. See "The Proximity Effect in Systems of Parallel Conductors and Electrically Small Loop Antennas" by G. Smith, Figure 1-13.[1] The measured inductance at 3.4 MHz was 1.52 µH.

For the measurements of inductance and SRF, the coil rested on a Perspex (clear acrylic plastic) shelf raised 15 cm away from the bench and any metalwork. It was connected to two metal terminals that protruded through the Perspex shelf, and these terminals were connected to an Array Solutions AIM 4170 analyzer via two 1.6 mm diameter copper wires, each about 16 cm long. The impedance of these wires was calibrated out at the connection terminals. Despite the fact that the "cold" end of the coil was not actually connected to ground or any metalwork, the λ/4 mode was still excited (see later).

The measurement of series resistance is more difficult, because it has to be measured in the presence of the very large series inductive reactance, and this leads to significant errors. This reactance was tuned out with high quality air variable capacitors, having silver plated vanes and contacts, and ceramic insulation. This loss was probably low but it is not possible to verify this, and so the resistance measurements include any series resistance caused by these capacitors. A more

significant effect is that a capacitor increases the self capacitance *just by virtue of its size*, so it should be physically as small as possible. Nevertheless, it is difficult to make the effective capacitance less than about 1 pF, and this, along with the additional path length of the connecting leads and the capacitor, reduces the SRF of the example coil from 163 MHz down to 88 MHz.

Current Theory

Self Capacitance and SRF

The current theory assumes that the SRF of a single layer coil is due to its self capacitance and that this can be approximated by a capacitor across its ends. The most extensive work on the capacitive approach is in a paper by R. G. Medhurst in the Feb and March 1947 issues of *Wireless Engineer*.[2] Medhurst *measured* a number of coils and derived an empirical equation for their parallel self capacitance, which matched his measurements (dimensions are in cm):

$$C_0 = 0.1126 \, l + 0.08 \, D + 0.27(D^3 / l)^{0.5} \text{ pF}$$

[Eq 1]

Attempts have been made to produce a *theoretical* basis for this capacitance, the most notable being by A. J. Polermo in the *Proceedings of the IRE*, in 1934.[3] He produced an equation based on the concept of inter-turn capacitance, which was supported by his own experiments. His theory was demolished by Medhurst, who came very close to accusing Polermo of adjusting his

[1]Notes appear on page 44.

results to fit his theory.

For the coil given above, Medhurst's equation gives $C_0 = 0.55$ pF. Notice how small this capacitance is, and so a very small stray capacitance due to the measuring jig can have a significant effect on the measured SRF, as can any leads connecting the coil to the jig. This is addressed later.

Taking the usual equation for the resonant frequency of a lossless parallel resonant circuit:

$$f_r = 1 / [2\pi (LC_0)^{0.5}] \qquad \text{[Eq 2]}$$

then 0.55 pF resonates with the inductance of 1.52 µH at $f_r = 174.6$ MHz. For comparison the SRF was *measured* as 163 MHz, and this was done by grounding one end via a short lead, leaving the other open and coupling into the coil with a two turn loop. So, correlation with experiment was reasonable given the problems of accurate measurement at these high frequencies, and the susceptibility of the measurement to the small stray capacitance of the lead. There is also some uncertainty in the value of D to be used in Equation 1, since the precise diameter of the *current* path is not known.

Effect of Self Capacitance on Inductance and Resistance

If a coil with a parallel capacitance is used in a parallel tuned circuit, the effect of this capacitance (whether self capacitance or added capacitance) is merely to reduce the capacitance necessary to resonate at any particular frequency. If the coil is used in a series tuned circuit, the effect of the parallel capacitance is to reduce the Q, increase the *apparent* inductance and to increase the *apparent* series resistance. In *The Theory and Design of Inductance Coils*, V. G. Welsby gives Equations 3, 4 and 5, below.[4] In these equations Q, L and R are the low frequency values for the inductor and Q_c, L_c and R_c are the values that would be measured at a frequency f, when there is a self resonant frequency of f_r:

$$Q_c = Q [1 - (f / f_r)^2] \qquad \text{[Eq 3]}$$

$$L_c = L / [1 - (f / f_r)^2] \qquad \text{[Eq 4]}$$

$$R_c = R / [1 - (f / f_r)^2]^2 \qquad \text{[Eq 5]}$$

These equations are widely accepted but it is important to understand what they mean. They are not designed to describe the parameters of a coil at any frequency, as we can see by examining Equation 5. It seems to say that the basic resistance R does not change with frequency, and that the *measured* value R_c only changes because of the effect of the self capacitance (as Welsby says "the presence of the capacitor modifies the current into the winding"). Clearly R *does* change with frequency due to skin effect, so what does Equation 5 mean? Well, Welsby needed an equivalent circuit of a coil that he could

use in his network analysis, and the accepted view at the time was that coils *did* have self capacitance so his equivalent circuit included this. A measurement of coil resistance at say, frequency f, however, gave the resistance R_c, whereas for his equivalent circuit he needed to know R. So Equation 5 gave him R, and the equation is meant to be a correction factor on a coil *at the measurement frequency only for a coil that has already been measured*. It is not a prediction of what R or R_c will be at any particular frequency.

If we want a prediction of the resistance at any frequency we can modify Equation 5 with an equation for the change of R with frequency due to skin effect and proximity:

$$R = R_{dc} \varphi H \qquad \text{[Eq 6]}$$

where:

R_{dc} is the dc resistance of the wire, $\rho l / A$

$H = r^2 / (2r\delta - \delta^2)$ (for $d / \delta > 2$)

$\delta = 66.6 / f^{0.5}$ mm for copper

r is the radius of the wire

Inserting this into Equation 6 we get:

$$R_c = [R_{dc} \varphi H] / [1 - (f / f_r)^2]^2 \qquad \text{[Eq 7]}$$

H is a multiplier due to the skin effect (see the link to the file **Zint.pdf** at **www.G3YNH.info/zdocs/comps/part_1**.**html** for an excellent explanation)[5] and φ is a multiplier that accounts for the increase in resistance caused by the magnetic field from adjacent turns — the proximity effect. There is much dispute about the validity of the theoretical analysis of proximity effect (such as Butterworth), and this is beyond the scope of this article. There is general agreement, however, that Medhurst's *measured* results provide good accuracy. For the coil parameters used here, Medhurst gives $\varphi = 1.5$

So, for the coil described previously,

Equation 7 gives the solid curve in the graph of Figure 1. The dashed curve shows the measured values. In this case, the series tuning capacitor reduced the SRF (see the Measurements section of this article), but it is difficult to calculate by how much, and so it was measured and found to be 88 MHz. This is the value of f_r used for the plot of Equation 7 in Figure 1. The correlation is seen to be very good, given the measurement problems.

Now we turn to Equation 4 for inductance. This is also designed as a correction for the *measured* inductance, but interestingly it also provides a good prediction of the inductance at any frequency. So if we take the coil described earlier, we find that Equation 4 gives very good agreement with measurements, as shown in Figure 2.

In Figure 2, the measured SRF reduced to 138 MHz because of the additional capacitance of the connecting leads (see later), and it is this frequency that was used as f_r in Equation 4, rather than Medhurst's SRF.

λ/4 and λ/2 Modes

The Medhurst/Welsby equations can give a good prediction of the inductance and with suitable modifications (Equation 7), also the resistance with frequency. As we have seen, however, they are based on a fictional self-capacitance of the coil.

The alternative explanation is that there is no self capacitance, and resonance occurs because the wire forming the coil acts as a transmission-line, which then resonates when the wire length corresponds approximately to $n\lambda/4$, where n is an integer 1, 2, 3 and so on.

Normally, one end of the coil is grounded and experiments show that this normally leads to the λ/4 mode, with the grounded terminal providing an effective transmission-line short circuit. This is also the mode when
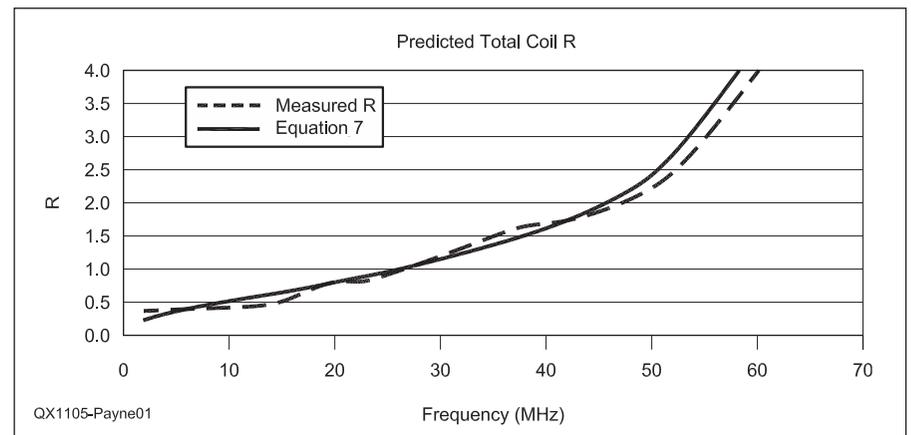
Figure 1 — This graph compares the predicted total coil resistance as calculated by Equation 7 (solid line) with the author's measured values (dashed line) for the sample inductor.

the coil is inserted into a test jig for the measurement of inductance, resistance and SRF.

The λ/2 mode can also be excited. For this, the coil is suspended away from other objects and ground, and energy coupled into it via a coupling loop of a few turns, located at the center of the coil. The whole arrangement is now balanced against ground. Resonance in the coil can be determined by measuring the input impedance of the loop, and this will peak at the SRF.

End Effect And Phase Velocity

Before addressing the transmission line theory of inductance it is necessary to look at two key parameters of transmission lines: end effect and phase velocity.

When a two wire transmission line, short circuited at the far end, has a length approximately equal to $\lambda_g/4$ it will resonate like a parallel LC circuit (λ_g is the wavelength of wave propagation on this line). Indeed at VHF and UHF, λ/4 transmission lines are often used in the place of conventional LC circuits.

I have been careful to stress that the line length is only *approximately* $\lambda_g/4$ for resonance and this is because of the "end effect," whereby the transmission line appears to be longer than its physical length, by up to 15% or even more.

Of course the coil is not a two wire transmission line since it has only one conductor, but standing waves can also appear on single wires, and the quarter wavelength monopole antenna and the half wavelength dipole are good examples. It has been found that the end effect appears here also, and can be calculated from the capacitance of the end disc formed from the conductor cross section, along with a capacitance caused by charge accumulation at the end of the line. For conventional wire antennas, where the conductor diameter is small, this capacitance is very small, and so the end effect is around 2%. If the antenna is short and fat, however, the end effect can be very large. For example, in *Antennas, Theory and Practice*, Figure 11-2, S. A. Schelkunoff shows a curve where the end effect is 15% when the antenna diameter is equal to λ/50.[6] You might think that if the end of a fat antenna was hollowed out to form a tube, the end capacitance would reduce noticeably, and with it the end effect. Early workers in this area, however, were surprised to find "not the slightest difference…" See "Experimentally Determined Impedance Characteristics of Cylindrical Antennas" by Brown and Woodward in the April 1943 issue of the *Proceedings of the IRE*.[7] This brings us to the helix, which of course is also hollow.

Measurements by the author indicate that when the helix dimensions are much less
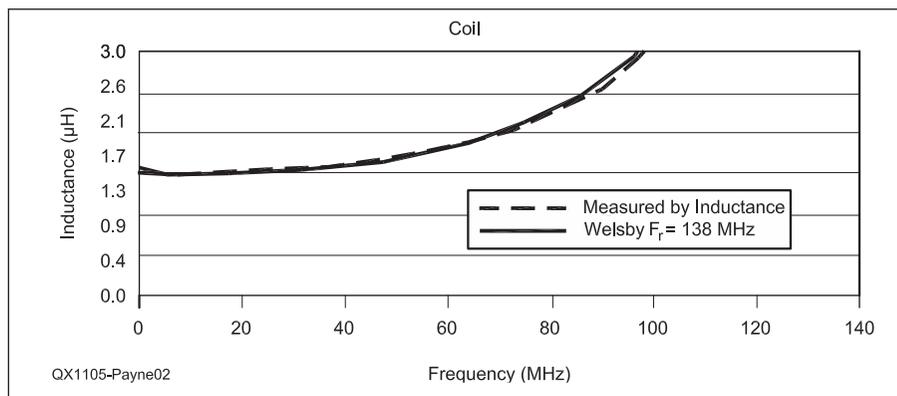
Figure 2 — This graph compares the inductance of the coil predicted by Equation 4 (solid line) with the author's measured values (dashed line).
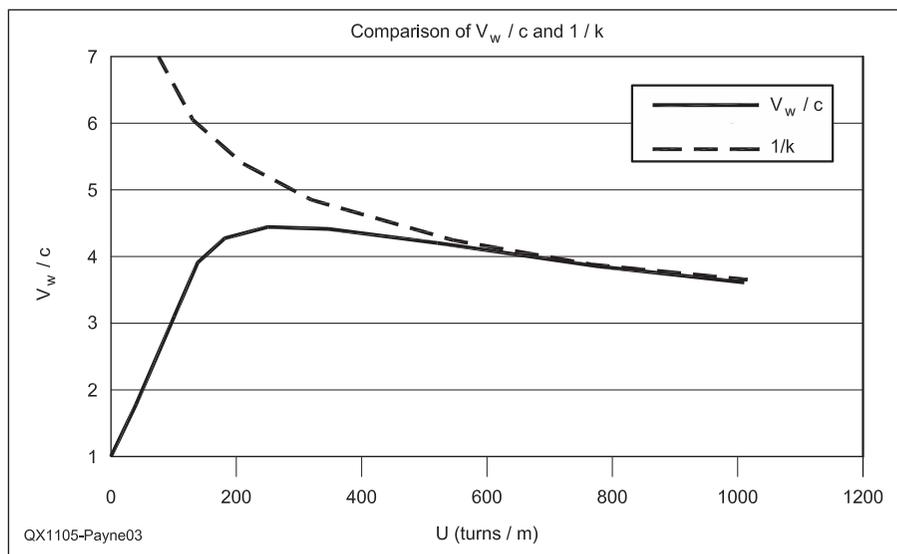
Figure 3 — The solid line shows the relative wave velocity along the coil as the coil pitch changes. The dashed line represents the reciprocal of *k*, as defined with Equation 12.

than a wavelength, the end effect is on the order of the radius of the helix. That is, the helix seems to be extended at each end by a distance of approximately half its radius. Interestingly this result has been known for some time, in a different guise, and is known as Nagaoka's constant. He gave the following equation for the low frequency inductance of a solenoid:

$$L = K_n \mu_0 N^2 A / l \qquad \text{[Eq 8]}$$

where:

K_n is Nagaoka's constant and is given by Welsby (see page 42 of Note 4) as approximately:

$$K_n \approx 1 / [1 + 0.45 D / l – 0.005 (D / l)^2] \qquad \text{[Eq 9]}$$

If D / l is less than unity this can be approximated as $K_n \approx 1 + 0.45 D / l$.

So, we can rewrite Equation 8 as:

$$L \approx [\mu_0 N^2 A] / [l (1 + 0.45 D / l)] \qquad \text{[Eq 10]}$$

Notice that in effect, Nagaoka is saying that the apparent length of a coil is increased by 0.45 D, or 0.225 D at each end. To emphasis this point, the end effect is proportional to the *coil* diameter, not to the *wire* diameter as in an antenna. In the subsequent calculations it is Nagaoka's constant that will be used to define the end effect.

To apply this to the transmission-line mode we note that with the λ/2 mode, the end effect is applicable at both ends because it is double ended and balanced. In the λ/4 mode, however, one end is effectively shorted and so has no end effect, and the other has 0.225 D.

It is convenient to express the end effect

as a fraction of the total length, so that the apparent length of the coil l_c' is given by $l_c' = l_c (1 + \delta)$, where for the $\lambda/2$ mode $\delta = [0.45 D / l_c]$ and for the $\lambda/4$ mode $\delta = [0.225 D / l_c]$. So taking the coil being used as an example for this article, where $D = 11.7$ mm and $l_c = 21$ mm, the end effect is 0.25 (25%) and 0.125 (12.5%) respectively. Notice how large these are. Although the end effect is calculated from the coil diameter and length, the apparent wire length is also increased by the *same proportion*, so we have:

$l_w' = l_w (1 + \delta).$

Turning now to phase velocity, if we take a two-wire transmission line (with air between the lines) or a straight single wire such as an antenna, EM waves travel down the lines with a velocity very close to that of the speed of light, c. If the straight wire is made into a helix the wave follows the helical path along the wire at the speed of light if the diameter of the helix is large, comparable to a wavelength. When the diameter is small, the wave again follows a helical path but this time with a pitch somewhat larger than that of the wire — up to four times larger or more. One way to look at this is that the phase velocity *down the wire* exceeds that of the velocity of light, c. A. G. Kandoian and W. Sichak have studied this theoretically and shown that the velocity down the wire, relative to the speed of light V_w / c, depends upon the pitch of the winding, its diameter and the frequency of operation.[8] Their equation can be written as:

$V_w / c = V_w' = [(1+ x^2) / (1 + (k\,x)^2)]^{0.5}$
[Eq 11]

where:
$x = \pi D / p$
$k = \sqrt{20}/\pi [D^2 f / (300 p)]^{0.25}$
D is the coil diameter in meters
p is the winding pitch in meters
f is in MHz.

Notice that the numerator gives the path length of the wire (per meter of coil) and the denominator that of the wave.

If $(kx)^2 \gg 1$ then this simplifies to $V_w' \approx 1 / k$, so:

$V_w' \approx 1 / k = [D^2 f / (73 p)]^{-0.25}$ [Eq 12]

Generally for normal coils in the HF range this approximation applies when $D / p > 4.5$, to give accuracy generally better than 10%.

As an example, for $D = 11.7$ mm and $f = 3.4$ MHz, Figure 3 shows the relative velocity plotted against U (turns per meter, equal to $1/p$). Also shown is the approximation $V_w' \approx 1 / k$.

Combining all the above equations, including the approximation for V_w', we get the following expression for *half wave* resonance:

$f_r^{1.25} \approx [300 \times 0.5 / (l + \delta)] / [D^2 / (73 p)]^{0.25}$ MHz
[Eq 13]

where:
δ is the fractional end effect.
For *quarter wave* resonance:

$f_r^{1.25} \approx [300 \times 0.25 / (l + \delta)] / [D^2 / (73 p)]^{0.25}$ MHz
[Eq 14]

One important result from these equations is that the quarter wave resonance is not at half the frequency of the half wave resonance, because at this lower frequency the velocity will have changed. For instance, for the coil considered here the half wave resonance is calculated to be at 265 MHz, and the quarter wave resonance at 153 MHz, not including any leads.

So in summary, for a given wire length, the end-effect reduces the self resonant frequency, whereas the increased velocity increases it. For some dimensions of coil and winding pitch these two effects can be equal and therefore cancel, so that the length of the wire will be exactly equal to $\lambda/4$ or $\lambda/2$.

New Theory

The new theory is based upon the following physical model: it is assumed that the inductance of the solenoid is produced by the *sum* of two inductances. The first is the normal low frequency inductance, described by familiar equations such as Equation 8 and is the result of the flux that flows down the core of the solenoid, and out of one end and into the other end. This we will call the "coil inductance." The second is a result of the flux around the *wire*, and this flux follows the path of the wire (approximately) and causes a transmission-line mode. This we will call the "transmission line inductance." This inductance is normally negligibly small, but the model assumes that it increases greatly at frequencies above about $f_r / 2$, and at the same time the "coil" inductance reduces because the flux turn to turn is no longer in phase. At frequencies above about $f_r / 2$ the transmission-line flux dominates. So as the frequency increases we go from a conventional "coil" mode being dominant to a transmission-line mode being dominant.

Reducing Coil Inductance

First we will look at the decreasing "coil inductance." At low frequencies the flux from each turn of the coil is in phase with all the others, and adds to give the overall flux. At higher frequencies there is a phase shift turn to turn, and the resultant flux (and thus the inductance) will decrease, and this decrease is given by the normalized sum of n vectors (where n is the number of turns) each

with a phase shift of ψ per turn (see *Antennas* by J. D. Kraus, page 78)[9]:

$E = 1 / n \times [\mathrm{Sin}\; n\psi / 2] / [\;\mathrm{Sin}\; \psi / 2]$
[Eq 15]

Now $\psi = 2 \pi (l_w' / n) / \lambda_g$, where (l_w' / n) is the apparent wire length per turn, and λ_g is the wavelength on the wire. The factor $\psi / 2$ is always small, so $\mathrm{Sin}\; \psi / 2 \approx \psi / 2$, and we have:

$E = [\mathrm{Sin}\; (\pi l_w' / \lambda_g)] / [\pi l_w' / \lambda_g]$ [Eq 16]

Now $\lambda_g = v_g / f = V_w c / f = V_w' \; 300 / f$, where f is in MHz. So:

$E = [\mathrm{Sin}\; \{(\pi l_w' f) / (300 V_w')\}] / [(\pi l_w' f)/ (300 V_w')]$
[Eq 17]

The *flux* φ in the coil will therefore decrease according to Equation 17. The *inductance* is given by $L = N\, d\varphi/di$, so this will also reduce as φ for a given number of turns N. But the turns are now not in phase and so the effective N also reduces as Equation 17. Overall then we can expect the inductance to reduce according to the square of Equation 17 :

$L_{coil} = L_0 [(\mathrm{Sin}\; \theta) / \theta]^2$
[Eq 18]

where:
L_0 is the low frequency inductance given by Equation 8
$\theta = \pi f l_w' / (300 V_w')$
$l_w' = l_w (1 + \delta)$
f is in MHz
V_w' is given by Equation 12.

Inductance of Transmission Line

Now we will turn to the transmission-line inductance. Normally, one end of a coil is grounded, and so we can consider its wire as a transmission line with one end shorted. Resonance will occur when the line, including its end effect, is a quarter wavelength long. Assuming a lossless line, the input reactance at the open end, at frequency f will be:

$X_{line} = Z_0\; \mathrm{Tan}\; [2\pi f l_w' / V_w]$ [Eq 19]

Below resonance the line will have an inductive reactance, with an inductance given by: $L_{line} = X_{line} / (2\pi f)$. If the frequency is in MHz, and $V_w / c = V_w'$, then:

$L_{line} = Z_0 / (2\pi f) \times \mathrm{Tan}\; [\; 2\pi f l_w' / (300 V_w')]$
[Eq 20]

where:
l_w' is the apparent length of the line including the end effect $= l_w (1+ \delta)$
f is in MHz and inductance is in microhenrys.

This would be the inductance if the wire were straight. When it is coiled into a solenoid, however, there is another effect, and to explain this we must look at the flux around the wire, in particular the flux between adjacent turns. In this area the flux from one wire is cancelled by that from its neighbor, at least

at low frequencies and so the transmission-line mode is suppressed. As the frequency is increased there is a phase shift between the turns, and the cancellation is then not complete, and this allows the transmission-line mode to propagate. So the effect is the opposite to that of the coil inductance (Equation 8), and the line inductance in microhenrys becomes:

$$L_{line} = [Z_0 / (2\pi f)] [1 - \{(\text{Sin } \theta) / \theta\}^2 \text{ Tan } 2\theta] \quad \text{[Eq 21]}$$

where:
$l_w' = l_w (1+ \delta)$
$\theta = \pi f l_w' / (300 \, V_w')$
f is in MHz.

For Equation 21 we need the characteristic impedance of the wire, Z_0, and this is somewhat problematic. A good start is Kandoian and Sichak (see Note 8) who give the characteristic impedance of a helical monopole antenna (a quarter wavelength above ground) of length h and diameter D as:

$$Z_0 = 60 [\ln (4 \, h / D) -1] (c / v_a) \quad \text{[Eq 22]}$$

They say "The first factor is the familiar expression, which predicts the reactance of ordinary antennas. The (c / v_a) factor takes account of the fact that the axial velocity of the helical antenna is less than the ordinary antenna"). Now the axial velocity, in terms of the wire velocity V_w / c is:

$$(v_a / c) = V_w / c \times l_c / l_w \quad \text{[Eq 23]}$$

where:
$l_c' = l_c (1 + \delta)$
l_w is the wire length
l_c is the coil length.

Equation 22 is related to the input impedance of a monopole antenna, that is, the impedance between the bottom end of an antenna and a large conducting ground plane. Here we are measuring at the top of the coil and so we can expect Z_0 to be larger, and good agreement with experiment (on the one coil) is given by :

$$Z_0 = 180 [\ln (4h / D) - 1] (c / v_a) \quad \text{[Eq 24]}$$

For the coil we have h = $l_c' = l_c (1 + \delta)$ and D is equal to the winding diameter at the center of the wire. So :

$$Z_0 = 180 [\ln (4 \, l_c' / D) - 1] (c / v_a) \quad \text{[Eq 25]}$$

where:
$(v_a / c) = V_w / c \times l_c / l_w$
l_w is the wire length
l_c is the coil length.

Connection Leads

The self resonant frequency is very sensitive to small stray capacitance, and because of this Medhurst carried out experiments to measure the capacitance of his leads. In principle we could also use the capacitance of the leads to calculate the foreshortening effect on

the transmission line, but there is an easier method. Instead of the leads adding capacitance, they must be considered as adding length to the coil wire. The phase velocity on the leads will be close to c, however, whereas the coil wire has velocity V_w'. So the length to be added is:

Apparent length of Connection leads = $V_w' (l_1 + l_2)$ $\quad$ [Eq 26]

Now if the leads are short compared with the wire in the coil, we can make a gross assumption that $V_w' \approx 2$ and will not lose too much accuracy as a result. So we can say:

Apparent length of Connection leads $\approx 2 (l_1 + l_2)$ $\quad$ [Eq 27]

where:
l_1 and l_2 are the lengths of the two connecting leads.

Experimental Confirmation of Inductance

Using Equations 18, 21 and 25 (in other

words, ignoring the leads for the moment) we get the graph shown in Figure 4.

The dotted curve is the rising inductance of the transmission line, and the dashed curve the falling inductance of the coil. Added together they give the dot-dash curve, which should be compared with the solid line measured curve. We see that the predicted curve falls below the measurements, but we have not included the length of the connection leads. Adding these in according to Equation 27 gives the graph of Figure 5.

So the new theory seems to give a good prediction of the inductance, and therefore some confirmation of the underlying modes. It has only been checked against the one coil, however, so further experimentation and research is needed.

Transmission-line Loss Resistance

We now need to know the resistive loss when the coil is in the transmission-line mode. This has been analyzed for a half wave dipole antenna by Schelkunoff and Friis (see
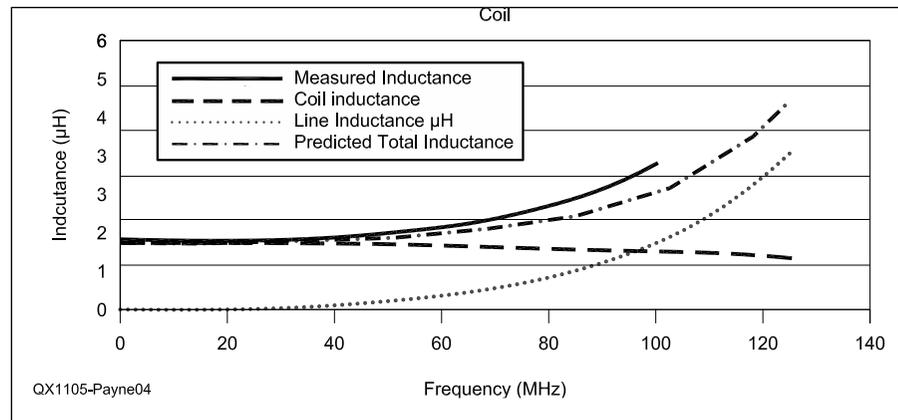
Figure 4 — The solid line on this graph represents the measured inductance as it varies with frequency. The dotted line represents the increasing inductance value using the transmission-line mode while the dashed line shows the decreasing inductance value of the coil with increasing frequency. The dot-dashed line represents the combination of transmission-line mode and coil inductance values.
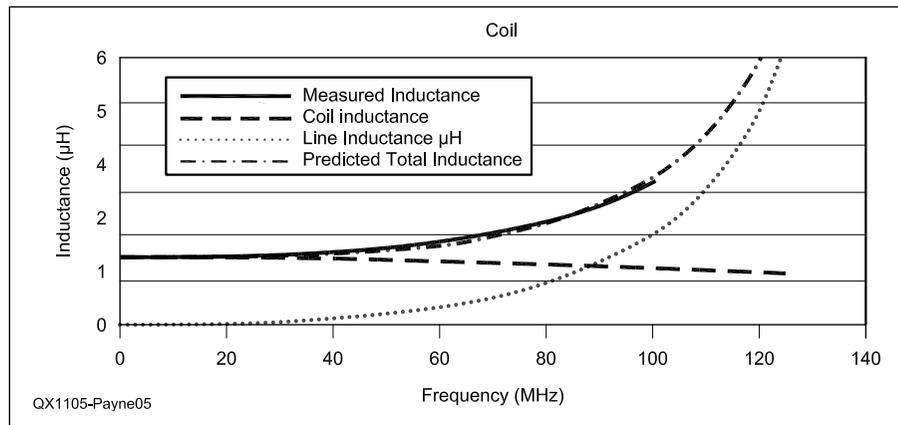
Figure 5 — This graph compares the measured inductance values, shown with a solid line, with the values of transmission-line mode inductance (dotted line), coil inductance (dashed line) and the combined values of the predicted total inductance (dot-dashed line).

Note 6, page 338), assuming that the current is sinusoidal. Following their method, the input resistance of a quarter-wavelength monopole having a wire resistance of R_0 Ω/m and length l_w will be:

$R_{in} = 0.5\ R_0\ l_w\ [1 + \{\mathrm{Sin}\ 2\ \beta\ l_w\ (1 + \delta) - \mathrm{Sin}\ 2\beta\delta\ l_w\} / \{2\ \beta\ l_w\}\ ] / \mathrm{Cos}^2\ \beta\ l_w\ (1+\delta)$ [Eq 28]

where:
$\beta = 2\pi / \lambda_g$
$\lambda_g = 300\ V_w' / f$
f in MHz.

[Note: There are two ways to define the length extension δ. First, as here, we can define it as a fraction of the length l_w, so $l_w' = l_w\ (1 + \delta)$, (so δ is, say 0.1 or 10%). Second, we can define it as a an actual length, in meters, so $l_w' = l_w + \delta$. Schelkunoff defines δ as a *length*, so his terms in δ reflect this.]

For $R_0\ l_w$ in the above we can use Equation 6 ($R_0\ l_w = R_{dc}\ \varphi\ H$). Then for the coil described earlier the resistance versus frequency, as given by Equation 28, is shown in Figure 6.

Also plotted in Figure 6 is Equation 7. Notice that this over estimates the theoretical loss by a progressively larger factor as the SRF is approached. Both curves are for an SRF of 153 MHz, calculated from Equation 14 for δ = 12.5%.

Equation 28 is rather complex, and an empirical equation that gives very similar results, at least up to $f_r / 2$ is:

$R_c = [R_{dc}\ \varphi\ H] / [1 - (f/f_r)^{2.9}]^2$ [Eq 29]

where
H is the skin effect = $r^2 / (2r\delta - \delta^2)$ (for $d/\delta > 2$)
δ = 66.6 / $f^{0.5}$ mm for copper
φ is the proximity factor given by Medhurst
f_r is given by Equation 13.

This simple equation agrees with Equation 28 to within ±5% for δ from 0 to 0.67 for all frequencies up to $f_r / 2$ and for all values of phase velocity V_w' (and thus f_r).

Equations 28 and 29 give the resistive loss in the transmission-line mode only, so in principle we should add in the loss in the coil mode, factoring each according to their relative values. This has not been done because it seems to add a layer of complexity that is not warranted, given the fact that the above equations seem to give a good estimate of the low frequency loss, as well as the high frequency loss.

So we now have two equations for the loss, Equation 7 given earlier and Equation 29 above. Which is correct? Equation 7 has the advantage that it agrees well with the measurements but these included the (unknown) losses from both the tuning capacitor and the form on which the coil was wound, so it seems to be fortuitous that it gave the right answer. A more extensive series of experiments could resolve the matter.
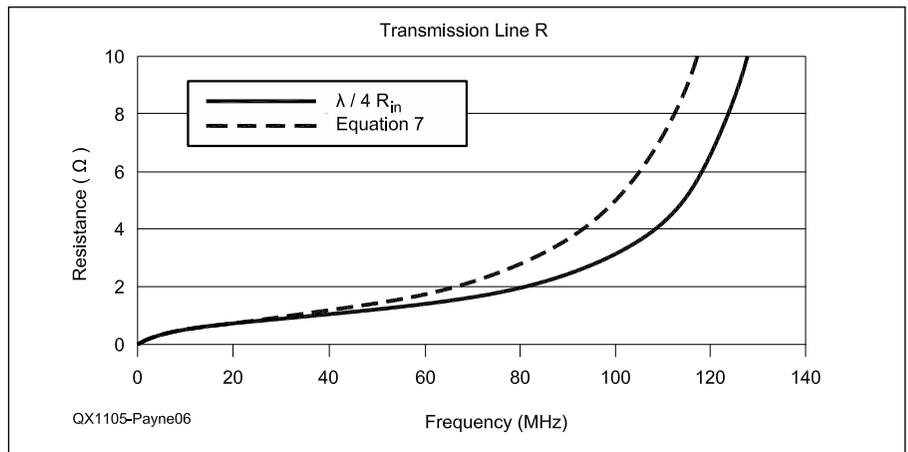
Figure 6 — The solid line shows the λ/4 transmission-line mode resistance as the frequency increases. The dashed line represents the values of coil resistance calculated by Equation 7.

SUMMARY

A new theory for the operation of coils has been given, which proposes two modes, the classical "coil" mode that dominates at low frequencies and a transmission-line mode that dominates at high frequencies.

It has been shown that the inductance and resistance increase with frequency, and that these are real changes and not *apparent* changes as predicted by the current self capacitance theory.

The inductance can be well approximated by the following empirical equation, and this also agrees well with experiment:

$L_c = L_0 / [1 - (f/f_r)^2]$ [Eq 30]

For the loss resistance a theoretical equation is given, but a good approximation is given by the following empirical equation:

$R_c \approx [R_{dc}\ \varphi\ H] / [1 - (f/f_r)^\alpha]^2$ [Eq 31]

where:
a value of α equal to 2.9 gives a value of resistance close to the theoretical transmission-line model, whereas α equal to 2 gives values closer to that measured, which includes the unknown losses in the tuning capacitor and in the winding form.

For the above equations the resonant frequency f_r is given by:

$f_r^{1.25} \approx [300 \times 0.25 / (l + \delta)] / [D^2 / (73\ p)]^{0.25}$
MHz [Eq 32]

The theoretical work has been supported by measurements but only on one coil, and more measurements would be needed to confirm the theoretical model put forward.

Alan Payne, G3RBJ is an Electrical Engineer with an MIEE degree. For most of his working life he was in Radar and EW design, working in both industry and for the United Kingdom Ministry of Defence. He is now retired. He has been a licensed Amateur Radio operator since 1961, and has had the same call sign for all 50 years. Alan's primary Amateur Radio interest has always been antennas, especially for the lower frequencies, where size is a serious constraint.

Notes

[1] G. Smith, "The Proximity Effect in Systems of Parallel Conductors and Electrically Small Loop Antennas," Tech. Rep. 624, Div of Engineering and Applied Physics, Harvard University, Cambridge, Mass, Dec 1971.

[2] R. G. Medhurst, "HF Resistance and Self Capacitance of Single-Layer Solenoids," *Wireless Engineer*, Feb 1947, pp 35-43 and March 1947, pp 80-92.

[3] A. J. Polermo, "Distributed Capacity of Single layer Coils," *Proceedings of the IRE*, 1934, Vol 22, p 897.

[4] V. G. Welsby, *The Theory and Design of Inductance Coils*, Macdonald & Co 1960.

[5] David Knight; Find the link to the document **Zint.pdf** at **www.G3YNH.info/zdocs/comps/Part_1.html**.

[6] S. A. Schelkunoff and T. H. Friis, *Antennas, Theory and Practice*, John Wiley and Sons, Inc, New York, 1950.

[7] Brown and Woodward "Experimentally Determined Impedance Characteristics of Cylindrical Antennas," *Proceedings of the IRE*, April 1943, pp 257-262.

[8] A. G. Kandoian and W. Sichak, "Wide-Frequency-Range Tuned Antennas and Circuits," *IRE Convention Record*, 1953 Pt 2 pp 42-47.

[9] J. D. KRAUS, *Antennas*, McGraw-Hill Book Company, Inc, 1950.

QEX

Letters to the Editor

Design, Construction and Evaluation of the Eight Circle Vertical Array for Low Band Receiving (Mar/Apr 2010)

The Lost 6 dB

Concerning the article by Joel Harrison, W5ZN and Bob McGwier, N4HY, the gain numbers for the eight circle vertical array are 6 dB lower than reported in the article.

This is due to the combining method used, which does not provide isolation between the antenna elements. Signals arriving on any one antenna element are dissipated not only in the 75 Ω build-out resistor of the element in question, but also in the 75 Ω build-out resistors of the other three antennas. This 6 dB loss can be avoided by using 0 degree hybrid combiners (sometimes referred to as "magic tees"). Use of three hybrid combiners to combine the four antenna elements, such as the MiniCircuits PSC 2-1-75 combiners, in this application provides approximately 25 dB of isolation between the antennas and also provides a 75 Ω termination to both ends of the 56° delay line (when used with the MiniCircuits PSC 2-1-75 only one 56° line is required since the impedance is 75 Ω at all ports). This improves the broadband performance of the system by providing good return loss over a much wider bandwidth for the delay line.

As the authors state, they believe the receive system gain (with the 6 dB loss) was adequate in their application. With that in mind, if the 0 degree hybrid combiners are used, the additional 6 dB of gain could be used up by building out the impedance of each antenna to 300 Ω. This approach lowers the Q by a factor of four and further improves the bandwidth by the same factor. The receive system gain is reduced by 6 dB, but the hybrid combiners improve the gain by 6 dB, so the net system gain change is 0 dB. This does require the addition of a 4:1 transformer at each antenna element to match the impedance to the 75 Ω transmission lines. Not only does this improve the bandwidth, it reduces the possible degradation of the antenna elements not tracking each other with time and environmental changes

— 73, Robye Lahlum, W1MK, 45 Brookview Rd, Boxford, MA, 01921; **w1mk@arrl.net**

Hi Robye,

Thank you for your informative feedback on our article and we concur with your conclusions. While gain is not the primary objective in this array design, individual element and array stability is very important as detailed in the article. Our approach was to provide a very simple feed system with good stability and excellent performance and your feedback takes that approach to the next level. We greatly appreciate you sharing it!

— Joel Harrison, W5ZN, 528 Miller Rd, Judsonia, AR 72081; **w5zn@w5zn.org**

— Bob McGwier, N4HY, 64 Brooktree Rd, East Windsor, NJ 08520; **rwmcgwier@ gmail.com**

A Simple Path to Complex Impedance (Sep/Oct 2010)

Dear Larry,

A few errors crept into my article, and here are the corrections:

On page 7, in Table 1A, Program Line # D084, the INSTR should read "x" not "X."

On page 13, in Table 3B, the Average Error (%) for 10.0 SWR should read "3.40" not "5.05."

On page 13, in Table 3C, Max Error (%) should read "Max Error (Deg)", and Average Error (%) should read "Average Error (Deg)," and Standard Deviation (%) should read "Standard Deviation (Deg)."

On page 15, in Table 4C, Max Error (%) should read "Max Error (Deg)", and Average Error (%) should read "Average Error (Deg)," and Standard Deviation (%) should read "Standard Deviation (Deg)."

Throughout the text I refer to the coupler that I used as "bidirectional." This is the correct name of the type of coupler that I built and used. The bidirectional circuit in Figure 5 on page 5, is what I built, and it is correctly named in that figure's caption. In several other figures and photos, however, the coupler is incorrectly referred to as, or is incorrectly labeled as, a "dual directional coupler," or a "DDC." Technically this is an error because a dual directional coupler is a different circuit from the bidirectional coupler circuit that I had built, although either type of coupler works well in this application.

— 73, Michael Bowman, KG2MG, 5077 Old Bald Hill Rd, N. Hemlock, NY 14466; **kg2mg@ arrl.net**

Hi Michael,

Thank you for sending those corrections. Our apologies to you and our readers for not catching and correcting those errors in the article.

— 73, Larry Wolfgang, WR1B, QEX Editor; **lwolfgang@arrl.org**

Upcoming Conferences

The 2011 IEEE Microwave Theory & Techniques Society (MTT-S) International Microwave Symposium (IMS2011)

June 5-10, 2011
Baltimore, MD Convention Center

IMS2011 will be held in Baltimore, Maryland, as the centerpiece of *Microwave Week 2011*, from Sunday, June 5 through Friday, June 10, 2011.

IMS2011 will cover developments in microwave technology from nano devices to system applications with a ***Microwaves for the World*** theme. Technical paper sessions, interactive forums, plenary and panel sessions, workshops, short courses, industrial exhibits, application seminars, historical exhibits, and a wide array of other technical activities will be offered. The Awards Banquet and Crab Feast are but two of the highlights of the many social activities planned. Not only are we arranging a great Guest Program, but there are many activities and attractions within Baltimore and the surrounding area that are sure to entice you to extend your visit.

In addition to IMS2011, the technical program of *Microwave Week* includes the Radio-Frequency Integrated Circuit Symposium (**www.rfic2011.org**) and the Automatic Radio-Frequency Techniques Group Conference (**www.arftg.org**).

For details about submitting papers, student design competitions, area hotels or any other information about the symposium, see the IMS2011 website at **ims2011.mtt.org**.

Society of Amateur Radio Astronomers 2011 Annual Conference

June 25-28, 2011
National Radio Astronomy Observatory
Green Bank, West Virginia

Note that this is a date change from what was reported in this column in the Jan/Feb 2011 issue of *QEX*.

Deadlines for submitting manuscripts for the proceedings have passed, but you can still attend the Conference. Presentations will be on radio astronomy hardware, software, education, research strategies and philosophy. For additional information about the Conference, and to register, visit the Society of Amateur Radio Astronomers website at **www.radio-astronomy.org**.

45th Annual Central States VHF Society Conference

July 29-30, 2011
The Westin, Dallas-Fort Worth Airport
4545 W John Carpenter Freeway
Irving, TX 75063

The Central States VHF Society, Inc is soliciting papers, presentations and poster displays for the 45th Annual CSVHFS Conference on 29-30 July, 2011. Papers, presentations, and posters on all aspects of weak-signal VHF and above Amateur Radio are requested. You do not need to attend the conference, nor present your paper, to have it published in the Proceedings. Posters will be displayed during the two days of the Conference.

Presentations at Central States aren't necessarily technical — they cover the breadth of the VHF/UHF ham radio hobby. Highlights in past years have been demonstrations of Software Defined Radio and LASER Communication beyond line-of-sight. Presentations generally vary from 15 to 45 minutes and step you through the highlights of the topic at hand, with complete texts published as articles in the *Proceedings*.

Topics of interest include:

- Antennas including Modeling/Design, Arrays, and Control
- Test Equipment including Homebrew, Using, and making measurements
- Construction of equipment, such as Transmitters, Receivers, and Transverters
- Operating including Contesting, Roving, and DXpeditions
- RF power amps including Single and Multi-band Vacuum Tube and solid-state
- Propagation including Ducting, Sporadic E, Tropospheric and Meteor Scatter, and so on
- Pre-amplifiers (low noise)
- Digital Modes, such as WSJT, JT65, and others
- Regulatory topics
- EME
- Software-defined Radio (SDR)
- Digital Signal Processing (DSP)

Non-weak signal topics, such as FM, Repeaters, packet radio, and similar topics are generally not considered acceptable. There are always exceptions, however. Please contact the folks below if you have any questions about the suitability of a topic.

Strong editorial preference will be given to those papers that are written and formatted specifically for publication, rather than as visual presentation aids.

Deadline for submissions:

- For the Proceedings: Monday, May 1st, 2011
- For Presentations delivered at the conference: Monday, 28 June 2011
- For notifying the Conference organizers that you will have a Poster to be displayed at the conference: Monday, 27 June 2011. Bring your poster with you on July 29.

Electronic formats (preferred)

Via e-mail

Upload to a web site for subsequent downloading

On media (CD/DVD, USB stick/thumb drive)

Contact: Kent Britain, WA5VJB; **wa5vjb@flash.net**, 1626 Vineyard, Grand Prairie, TX 75052.

There is more information about this year's conference available on the Central States VHF Society website at **www.csvhfs.org/2011conference/index.html**.

The 30th Annual ARRL and TAPR Digital Communications Conference

September 16-18, 2011
Four Points by Sheraton BWI Airport
7032 Elm Road
Baltimore, Maryland 21240

Mark your calendar now and start making plans to attend the premier technical conference of the year, the 30th Annual ARRL and TAPR Digital Communications Conference to be held September 16-18, 2011, in Baltimore, MD. The conference location is the Four Points by Sheraton BWI Airport, Baltimore, MD.

We recommended that you book your room prior to arriving. TAPR has reserved a block of rooms at the special DCC room rate of $90.00 single/double. This special rate is good until August 15, 2011. After that you will pay the regular room rate. To book your room, use the link on the TAPR website under Conferences (**tapr.org/dcc.htm**) or call the hotel directly (Reservations: 1-800-368-7764 Phone: 410-859-3300) and mention the group code **DCC** (Digital Communications Conference) when making reservations. *Be sure to book your rooms early!*

The ARRL and TAPR Digital Communications Conference is an interna-

tional forum for radio amateurs to meet, publish their work, and present new ideas and techniques. Presenters and attendees will have the opportunity to exchange ideas and learn about recent hardware and software advances, theories, experimental results, and practical applications.

Topics include, but are not limited to: Software defined radio (SDR), digital voice (D-Star, P25, WinDRM, FDMDV, G4GUO), digital satellite communications, Global Position System (GPS), precision timing, Automatic Packet Reporting System® (APRS), short messaging (a mode of APRS), Digital Signal Processing (DSP), HF digital modes, Internet interoperability with Amateur Radio networks, spread spectrum, IEEE 802.11 and other Part 15 license-exempt systems adaptable for Amateur Radio, using TCP/IP networking over Amateur Radio, mesh and peer to peer wireless networking, emergency and Homeland Defense backup digital communications, using *Linux* in Amateur Radio, updates on AX.25 and other wireless networking protocols and any topics that advance the Amateur Radio art.

This is a three-Day Conference (Friday, Saturday, Sunday). Technical and introductory sessions will be presented all day Friday and Saturday.

Join others at the conference for a Friday evening social get together. A Saturday evening banquet features an invited speaker and concludes with award presentations and prize drawings.

The ever-popular Sunday Seminar focuses on a topic and provides an in-depth four-hour presentation by an expert in the field. Check the TAPR website for more information: tapr.org.

Call for Papers

Technical papers are solicited for presentation and publication in the *Digital Communications Conference Proceedings*. Annual conference proceedings are pub-lished by the ARRL. Presentation at the conference is not required for publication. Submission of papers are due by 31 July 2011 and should be submitted to: Maty Weinberg, ARRL, 225 Main Street, Newington, CT 06111, or via the Internet to **maty@arrl.org**.

QEX-