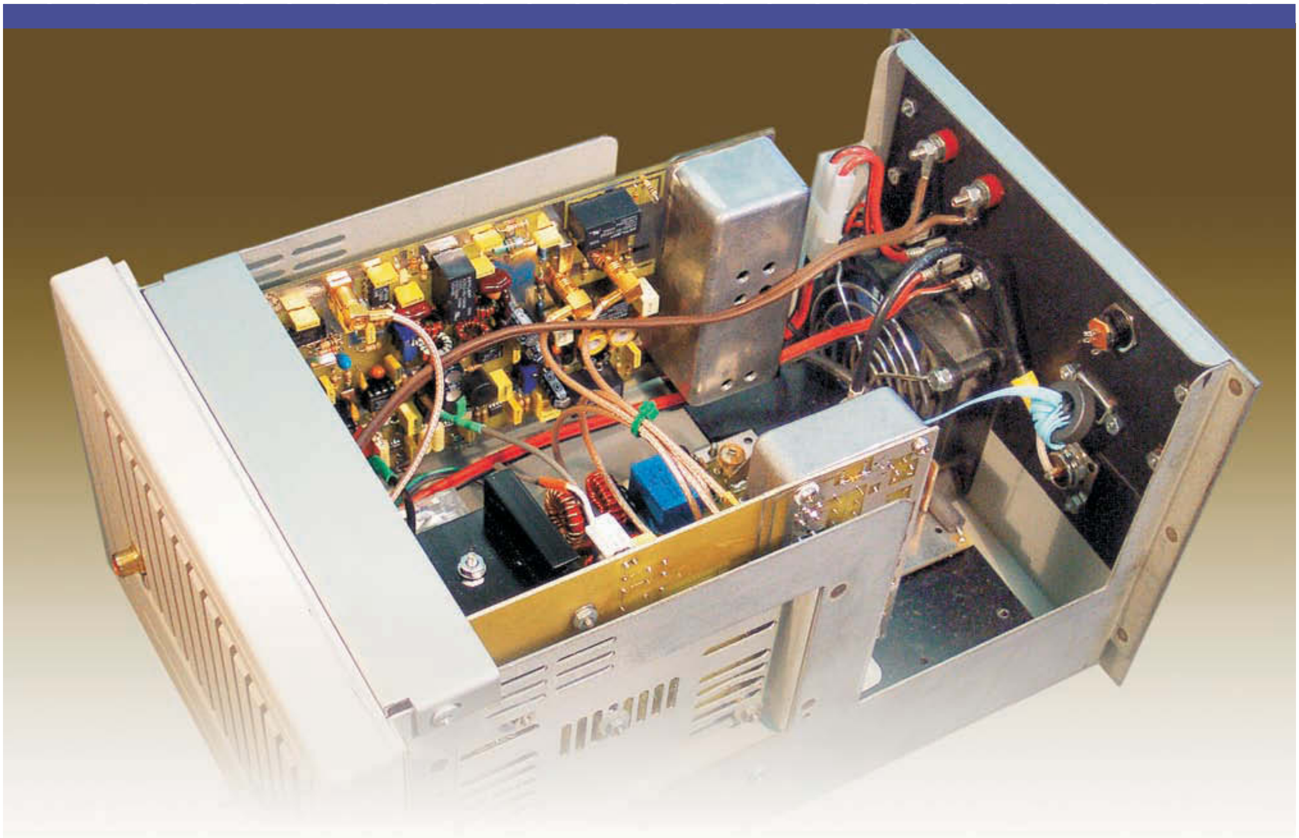
$5

QEX

INCLUDING: COMMUNICATIONS QUARTERLY

Forum for Communications Experimenters

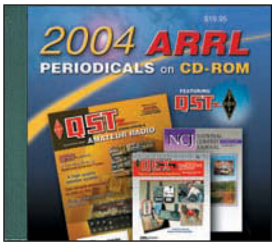
November/December 2005
Issue No. 233



KQ6F's Direct-Conversion, Phasing-Method SSB 100-W Transceiver for 40 and 75 Meters

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

CD-ROM Collections



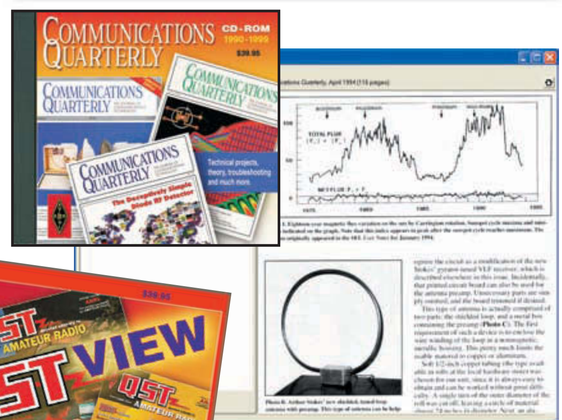
ARRL Periodicals on CD-ROM are fully-searchable collections of popular ARRL journals. Every word and photo published throughout the year is included!

SEARCH the full text of every article by entering titles, call signs, names—almost any word. SEE every word, photo (including color images), drawing and table in technical and general-interest features, columns and product reviews, plus all advertisements. PRINT what you see, or copy it into other applications.

System Requirements:
Microsoft *Windows*™. 1999, 2000, 2001, 2002, 2003, and 2004 editions support *Windows* and Macintosh systems, using the industry standard Adobe *Acrobat Reader*® (included).

Communications Quarterly CD-ROM

This CD-ROM collection covers volumes of *Communications Quarterly* published from 1990-1999. Gain access to advanced technical topics in articles which cover transmitter, receiver and transceiver projects, theory, antennas, troubleshooting and much more. High quality black-and-white page scans can be read on your computer screen or printed. Quickly **search** for articles by title and author, **select** specific year and issue, and **browse** individual articles and columns.
Requires Microsoft *Windows*™.



QST View

Every page of *QST*—all the ads, articles, columns and covers—has been scanned as a black-and-white image that can be viewed or printed.

BONUS! Latest, enhanced viewer—**AView**—included with any **QST View** purchase. **AView** is compatible with all sets of **QST View**.

ARRL CD-ROM Collections

ARRL Periodicals CD-ROM
(includes *QST*, *QEX*, *NCJ*)

#9396 Year 2004	$19.95
#9124 Year 2003	$19.95
#8802 Year 2002	$19.95
#8632 Year 2001	$19.95
#8209 Year 2000	$19.95
#7881 Year 1999	$19.95
#7377 Year 1998	$19.95
#6729 Year 1997	$19.95
#6109 Year 1996	$19.95
#5579 Year 1995	$19.95

QST View CD-ROM

#7008 Years 1915-29	$39.95
#6710 Years 1930-39	$39.95
#6648 Years 1940-49	$39.95
#6435 Years 1950-59	$39.95
#6443 Years 1960-64	$39.95
#6451 Years 1965-69	$39.95
#5781 Years 1970-74	$39.95
#5773 Years 1975-79	$39.95
#5765 Years 1980-84	$39.95
#5757 Years 1985-89	$39.95
#5749 Years 1990-94	$39.95
#8497 Years 1995-99	$39.95
#9418 Years 2000-2004	$39.95
#QSTV (all 13 sets). **Only $399** (Save $120)	

Communications Quarterly CD-ROM

#8780 (1990-1999)	$39.95

Ham Radio Magazine CD-ROM*

#8381 Years 1968-76	$59.95
#8403 Years 1977-83	$59.95
#8411 Years 1984-90	$59.95
#HRCD (all three sets)	$149.85

QEX Collection CD-ROM

#7660 (1981-1998)	$39.95

NCJ Collection CD-ROM

#7733 (1973-1998)	$39.95

www.arrl.org/shop
1-888-277-5289 (US)

*Ham Radio CD-ROM, © 2001, American Radio Relay League, Inc. Ham Radio Magazine © 1968-1990, CQ Communications, Inc.

Shipping & Handling charges apply: US orders add $6 for one CD, plus $2 for each additional CD ($12 max.). International orders add $5.00 to US rate ($17.00 max.). Or, contact ARRL to locate a dealer. Sales Tax is required for orders shipped to CA, CT, VA, and Canada.

INCLUDING: COMMUNICATIONS QUARTERLY

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

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

Issue No 232

Offices

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

Subscription rate for 6 issues:

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

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

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

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

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

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

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

About the Cover

Rod (KQ6F) Brink's 100 W phasing SSB transceiver begins on p 3.

Features

Columns

Nov/Dec 2005 *QEX* Advertising Index

Telephone: 860-594-0200

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

The purpose of *QEX* is to:
1) provide a medium for the exchange of ideas and information among Amateur Radio experimenters,
2) document advanced technical work in the Amateur Radio field, and
3) support efforts to advance the state of the Amateur Radio art.

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

Both theoretical and practical technical articles are welcomed. Manuscripts should be submitted on IBM or Mac format 3.5-inch diskette in word-processor format, if possible. We can redraw any figures as long as their content is clear. Photos should be glossy, color or black-and-white prints of at least the size they are to appear in *QEX*. Further information for authors can be found on the Web at **www.arrl.org/qex/** or by e-mail to **qex@arrl.org**.

Empirical Outlook

Emergency Preparedness and the New Federalism

In light of the ongoing Hurricane Katrina disaster, we offer a few words that we hope will provoke your thoughts and actions. This is not the "blame game" that so many reprobates have cited to deflect attention from their irresponsibilities. It's fact and opinion.

It's hard to imagine anyone who's not been touched in some way by the current devastation. Did we know it was coming? If so, what did we do about it in advance? The first answer is that we *did* know it was coming. The second answer is that we *didn't* do nearly enough in advance. Those answers should be self-evident.

Communications, command, control and intelligence (C^3I) are critical issues during any emergency. Radio amateurs are superb at providing communications when all else fails. All we have to do is arrive with the right equipment where communications are needed. We can benefit from improvements in our planning and execution, just as can those in command and control.

The ARRL *Amateur Radio Emergency Service Field Resources Manual* (*ARES FRM*) and *Public Service Communications Manual* (*PSCM*) are finely crafted documents but nowhere do they directly address the issue of transporting hams to disaster areas. Transportation is tacitly left to volunteers' personal capabilities or to other relief entities. The American Red Cross and The Salvation Army can get us where we need to be most of the time but we have to get ourselves to places like Montgomery, Alabama, first.

Our manuals refer to the *Federal Response Plan*, now the US Department of Homeland Security's *National Response Plan*, which was proudly announced in January 2005 by Tom Ridge, then Secretary of the department. The really odd thing is that this plan doesn't refer to Amateur Radio *anywhere* in its 426 pages! In fairness, we must state that ARRL is an affiliate of the Citizen Corps, which is mentioned. Citizen Corps works through a national network of state, local and tribal Citizen Corps Councils that bring together leaders from law enforcement, fire, emergency medical and other emergency management, volunteer organizations, local elected officials, the private sector and other community stakeholders. The League also maintains memoranda of understanding with many governmental and volunteer relief agencies. Those memoranda have helped us get our feet in the door, as it were; but they don't commit the Federal Emergency Management Agency (FEMA) or anyone else to provide us with transportation, food, shelter, flashlights and so forth.

What can be done? Before we get to that, check out a summary of the technical articles in this, our 233rd issue.

(Continued on p 62)

In This Issue

At last we have Rod Brink, KQ6F's direct-conversion phasing rig. 75- and 40-meter ragchewers take note! Rod documents his design of a good-sounding rig, backing up his claims with measurements. Gary Heckman, KC7FHP, takes us on a related journey through the theory behind quadrature phasing concepts and unwanted sideband reduction. Frank Carcia, WA1GFZ, presents a legal-limit power amplifier for those bands and for 160 meters. It's a solid-state unit that uses inexpensive switching FETs in class C at 50 V dc.

Mal Crawford, K1MC, describes his low-noise, differential-amplifier VFO. The oscillator operates in the 5.3-5.5 MHz range, which makes it a shoe-in for Heath and other popular rigs.

Sam Cowan, WØOAJ, has some ideas about reducing the bandwidth of analog phone emissions. It's a subject that's graced these pages before but Sam's take on it adds significant information. Yours truly has a slightly different way of reducing voice bandwidth that lends new meaning to the term, "simultaneous voice and data."

Contributing Editor L. B. Cebik, W4RNL, offers the second part of his look at antenna modeling software. Read about the similarities, differences, limitations and options of what's out there. □□

A Direct-Conversion Phasing-type SSB Rig

If carefully built and adjusted, a phasing-type SSB rig can sound better than a superhet—in both transmit and receive. Here's a simple transceiver for ragchewing on 75 and 40 meters.

By Rod Brink, KQ6F

In a recent ARRL publication,[1] Rick Campbell, KK7B, describes in detail his experience working with phasing-type transceiver circuits. I was intrigued by his work and surprised at its potential, thinking that phasing rigs were permanently "buried" decades ago by superhets with crystal filters. Campbell's work seemed to indicate that while a phasing transceiver might not have the same selectivity as a multi-conversion superhet with sharp filters, its minor shortcomings are offset by other advantages. These are: fewer adjustments, fewer spurs, and in many cases, better-sounding audio. This drew my interest, as I am an enhanced-SSB-audio fan. Inspired by his work, I set about to develop a simple rig for ragchewing on two of the low HF bands, 75 and 40 meters. These were chosen primarily because that's where I mostly hang out these days, since my favorite 17-meter band has faded nearly into oblivion with Cycle 23.

[1]Notes appear on page 15.

25950 Paseo de Los Robles
Corral de Tierra, CA 93908
kq6f@redshift.com

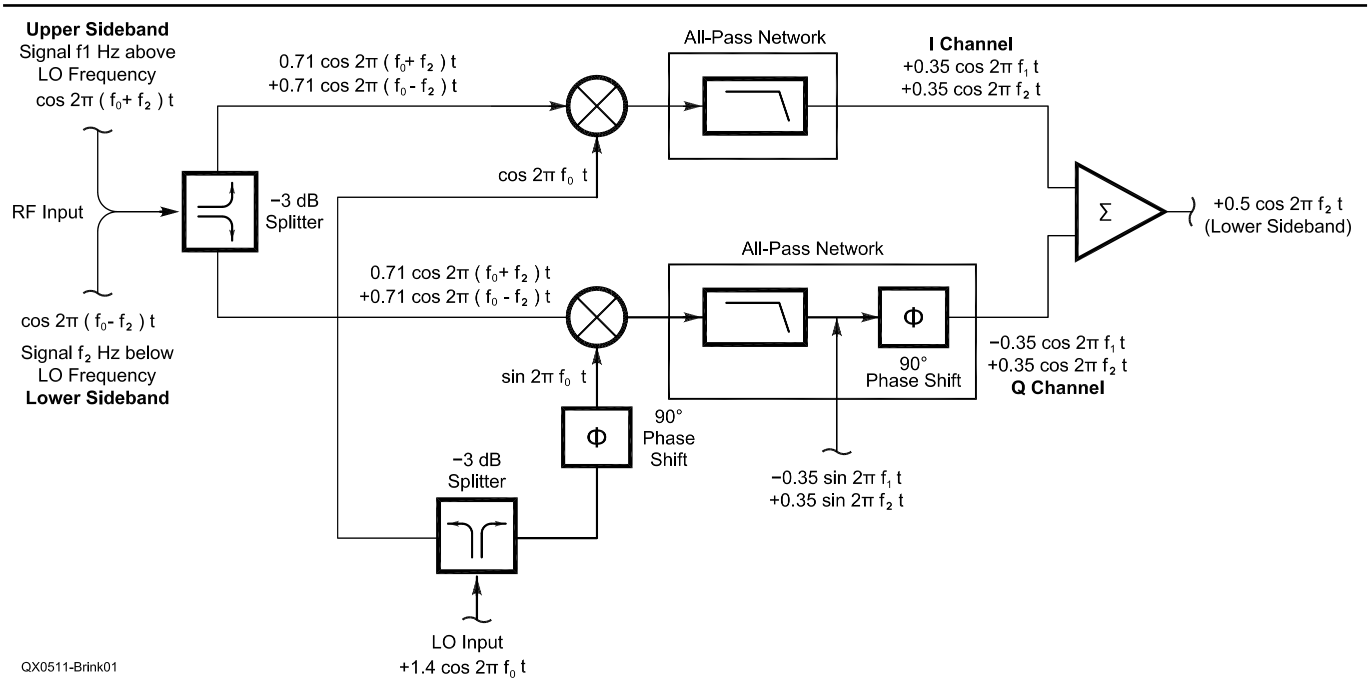


Fig 1—Block diagram of an Opposite-Sideband Reject Mixer (OSRM). Signals above the LO frequency cancel at the output, while signals below the LO frequency add to produce lower sideband. Mixer outputs to the low-pass filters may be reversed to produce upper sideband.

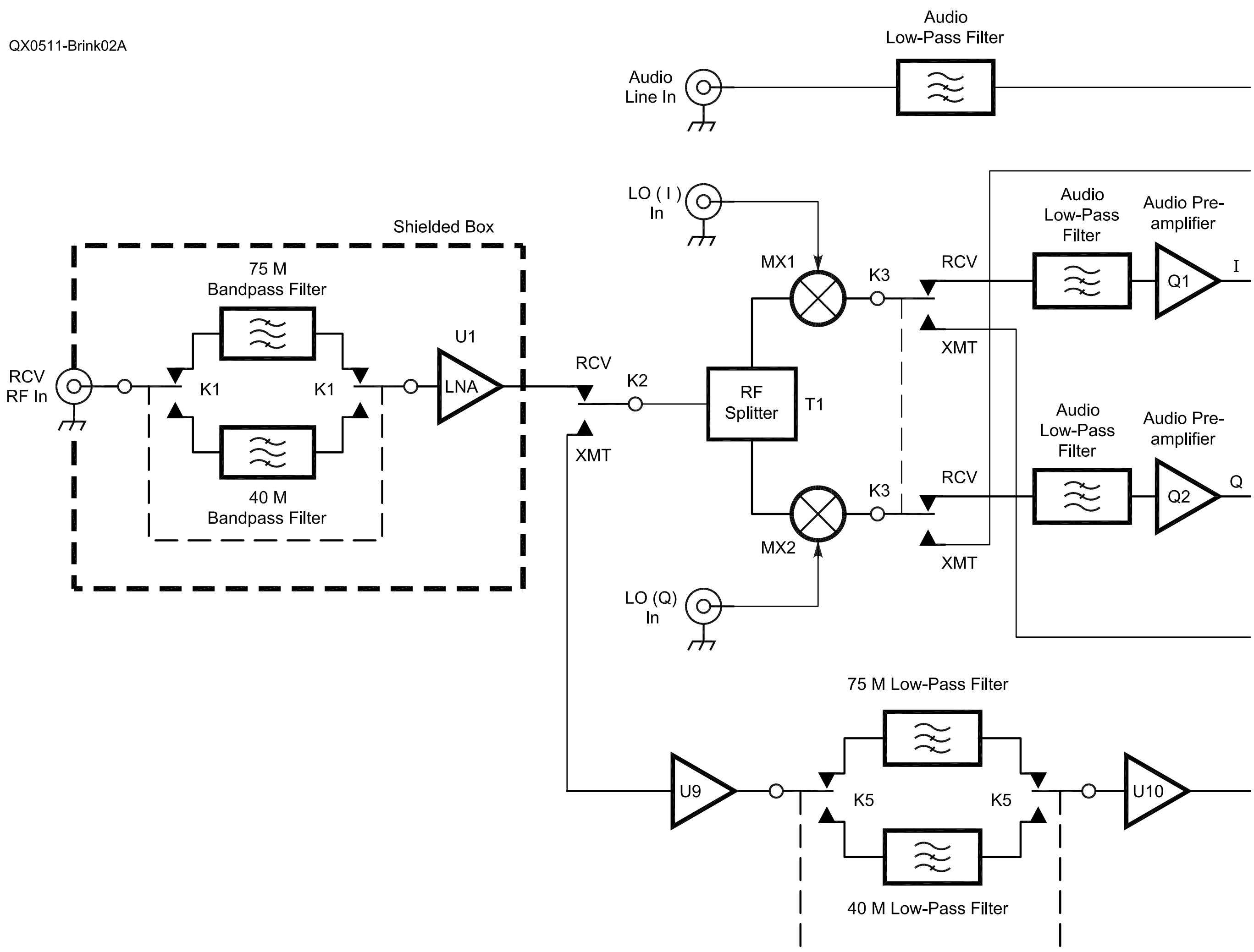


Fig 2—RF Board block diagram. Much of the circuitry is shared between transmit and receive.

Block Diagram—Receive Mode

The heart of the Direct-Conversion (DC) phasing-type receiver is the so-called Image-Reject-Mixer, or more accurately, Opposite-Sideband-Reject Mixer. In reality, it involves *two* mixers. It is shown as a block diagram in Fig 1, along with the mathematical relationships between the RF input and local oscillator. Signals summed at the output cancel for frequencies above the LO frequency and add for those below the LO to produce the lower sideband. Reversing connections at various places in the chain—including those at the mixer outputs—will produce the upper sideband.

Older hams will possibly remember with chagrin the phasing-type rigs of the '50s and '60s. Good opposite-sideband rejection requires amplitude balance to less than 0.1 dB and phase error less than 1° throughout the audio passband. Obtaining this kind of performance in a band-switched, vacuum-tube, paper-capacitor rig was nearly impossible. The rigs often sounded bad on the air. Many hams gave up trying to adjust their rigs and continued to transmit poor signals until SSB superhets with crystal filters came along.

The situation is different today. The required amplitude and phase tolerances are easy to obtain with modern components. In general, a properly designed phasing rig has fewer adjustments and fewer spurs than does a superhet with filters.

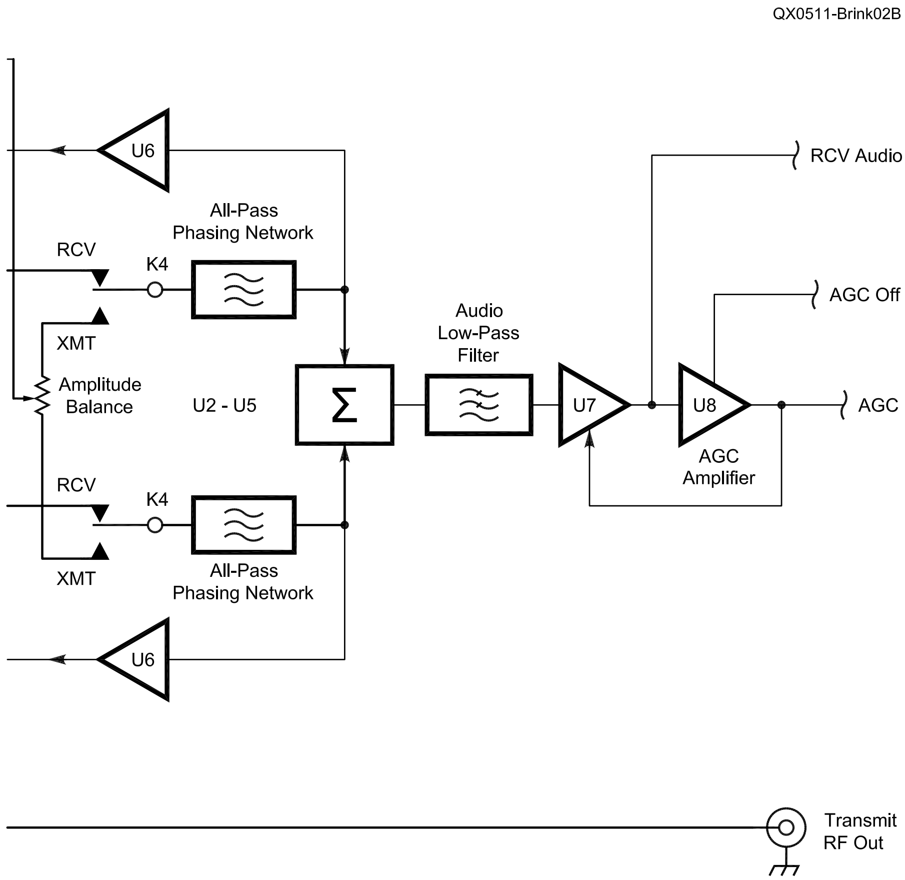
I built my transceiver into a surplus hard-disk enclosure. In order to get everything to fit, I subdivided things into four assemblies:

1. RF Board
2. Control Board
3. RF Low-Pass Filter Board
4. PA-Heat-Sink Assembly.

RF Board Block Diagram

The RF-Board block diagram is shown in Fig 2. With the unit in receive mode, the RF input passes through either the 75-meter or 40-meter preselection filters, is amplified by the LNA, then split by T1 and applied to the two mixers. Direct-conversion is employed, so the LO frequency is the same as the incoming RF. The mixers therefore act as block down-converters, heterodyning the incoming RF down to audio.

Subsequent amplification, processing, and bandwidth limiting are all done by audio circuits. This has both advantages and disadvantages.



Generally, the circuitry is easier to build and get working, with less expensive test equipment than that for RF circuits. On the other hand, audio circuits with very high gain must be designed carefully to avoid microphonics and stray pickup from nearby noise sources (especially raster sweep yokes in computer monitors—more on that later).

The mixer outputs pass through audio low-pass filters (important for reasons given later), are amplified 40 dB by audio preamplifiers and pass into all-pass phasing networks. We designate these two mixer outputs I and Q, for *In-phase* and *Quadrature-phase,* respectively. This is so because the LO signals to the mixer inputs are in quadrature, that is, phased 90° apart. The I and Q signals are identical except for the 90° phase difference. An *additional* 90° phase shift between the two occurs as the signals move through the all-pass networks. When the two outputs are summed, signals for one sideband are in phase with each other and those for the opposite sideband are 180° out of phase and cancel. The degree to which they cancel is a function both of their relative amplitudes and phase relationships. For *ideal* opposite sideband cancellation, the two LO signals would be *exactly* 90° apart, the relative phase shifts of the audio signals through the all-pass networks would likewise be *exactly* 90°, and amplitude variations would be *zero*. In practice, of course, these ideals are never achieved, only sought after.

The all-pass networks are so named because they pass all audio frequencies. Their ability to introduce 90° relative phase shifts is, however, limited to a specific range, which, in this case, is about 300-3500 Hz. Therefore, optimum opposite sideband suppression occurs only within this range.

To reduce off-channel interference, the summed output is passed through another audio low-pass filter and then amplified further by U7. AGC is applied at this point. I made provisions for defeating the AGC but always keep it active as a matter of personal preference.

Transmit Circuit Path

Much of the circuitry is shared between receive and Transmit modes. When the relays are set to transmit, the incoming transmit audio goes through a low-pass filter to accommodate the all-pass network limitations, then switched to the IF mixer ports, where it is multiplied with the I and Q local oscillator signals. The signals at the mixer RF ports, which were *inputs* in receive, are now *outputs* in transmit and pass into the RF splitter, which now functions as a power combiner. The RF transmit signal is amplified by U9, passes through either a 75 or 40 meter low-pass filter, gets further amplified by U10 and passes finally to the PA driver and PA for power amplification to +50 dBm (100 W).

Control-Board Block Diagram

The Control-board block diagram is shown in Fig 3. At the controller heart is the ubiquitous Stamp II module. This is a very handy and easy-to-use hybrid part that includes a microcontroller, 2 kbyte EEPROM for program storage, clock, power-supply regulator and serial interface, all in a 24-pin DIP package. Its internal firmware includes a *BASIC*-style interpreter that makes writing and debugging programs a snap. Its manufacturer has free downloadable software for program development on a PC. The instruction set includes some powerful commands that work flawlessly. In short, it's a dandy part.

The AD9835 from Analog Devices is another nice part. This is a complete DDS chip that runs from a 50-MHz clock and generates synthesized sine waves with 32-bit resolution.

There are spurs in the output that need to be cleaned up. Other people have used PLLs for this purpose, but I found that five-element low-pass filters did an adequate job and are, of course, much simpler. The filtered LO signals drive two quadrature transformers T2 and T3, which produce the required 90° phase shift for the I and Q channels. The transformers and their associated capacitors are frequency sensitive, so separate transformers are required for the two bands.

The following parameters are measured by an ADC chip and displayed on the computer monitor:

1. AGC voltage from the receiver is converted by software to S units.

2. Heat-sink temperature, which turned out to be unnecessary. I used an oversize heat sink for the PA and its temperature rarely gets over 98°F, even during windbag QSOs. There is a fan in the enclosure, but I disconnected it.

3. Forward and reflected power: these signals are generated by a directional coupler in series with the antenna lead on the low-pass filter board. They are run through peak-detection circuits before measurement.

The Control Board has another audio low-pass filter that receives the Receiver-Board output. This filter both improves receiver selectivity and reduces noise generated in preceding high-gain audio stages. Audio volume is controlled by a digital pot. The audio is then fed to an amplifier module that can drive 5 W into an 8-Ω speaker.

Detailed Circuit Description—RF Board

The receiver front-end schematic is shown in Fig 4. The RF input is

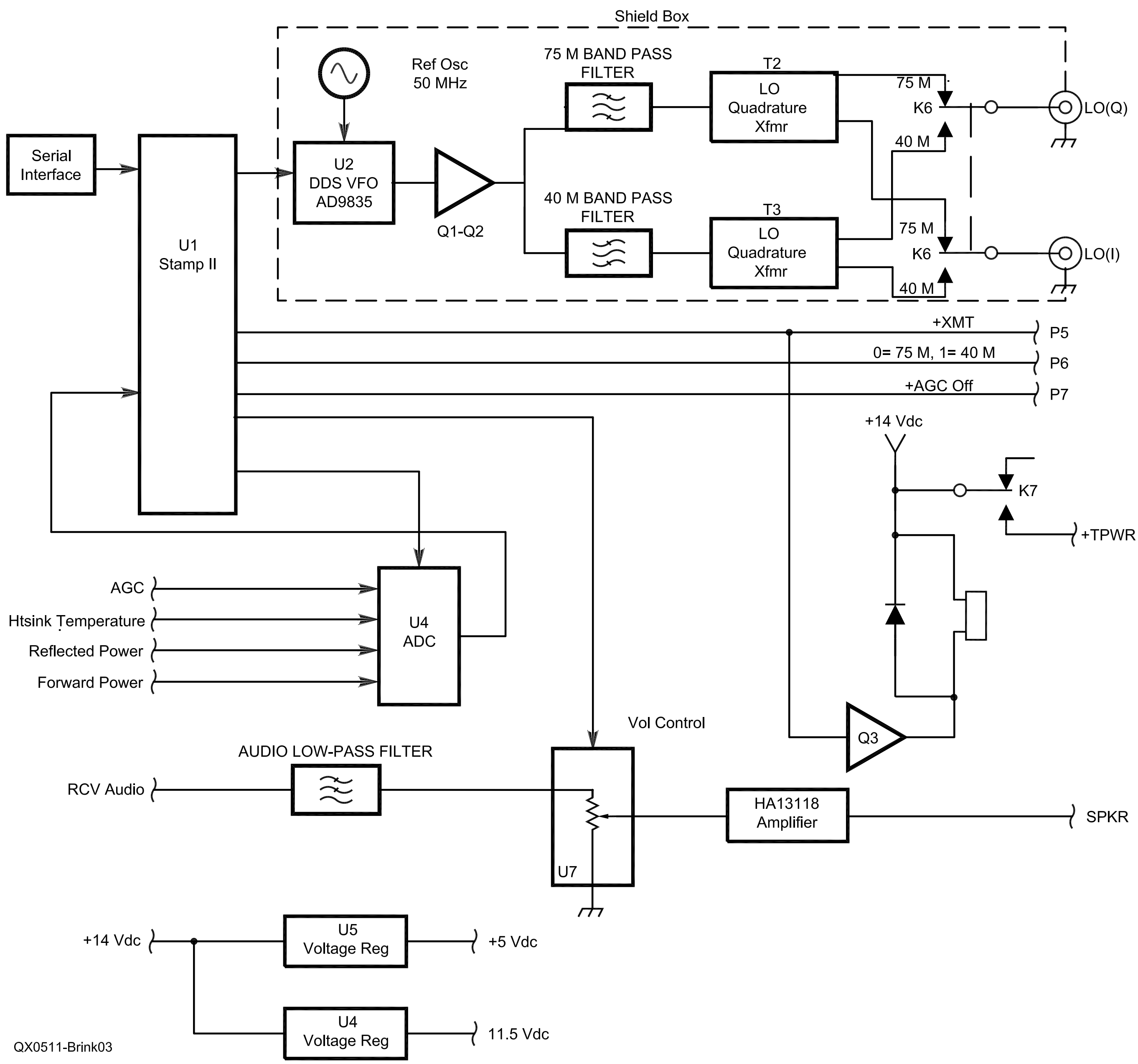


Fig 3—Control board block diagram. The local-oscillator components are located on the control board, away from the receiver input and contained within a shielded box to eliminate undesirable mixer unbalance (see text).

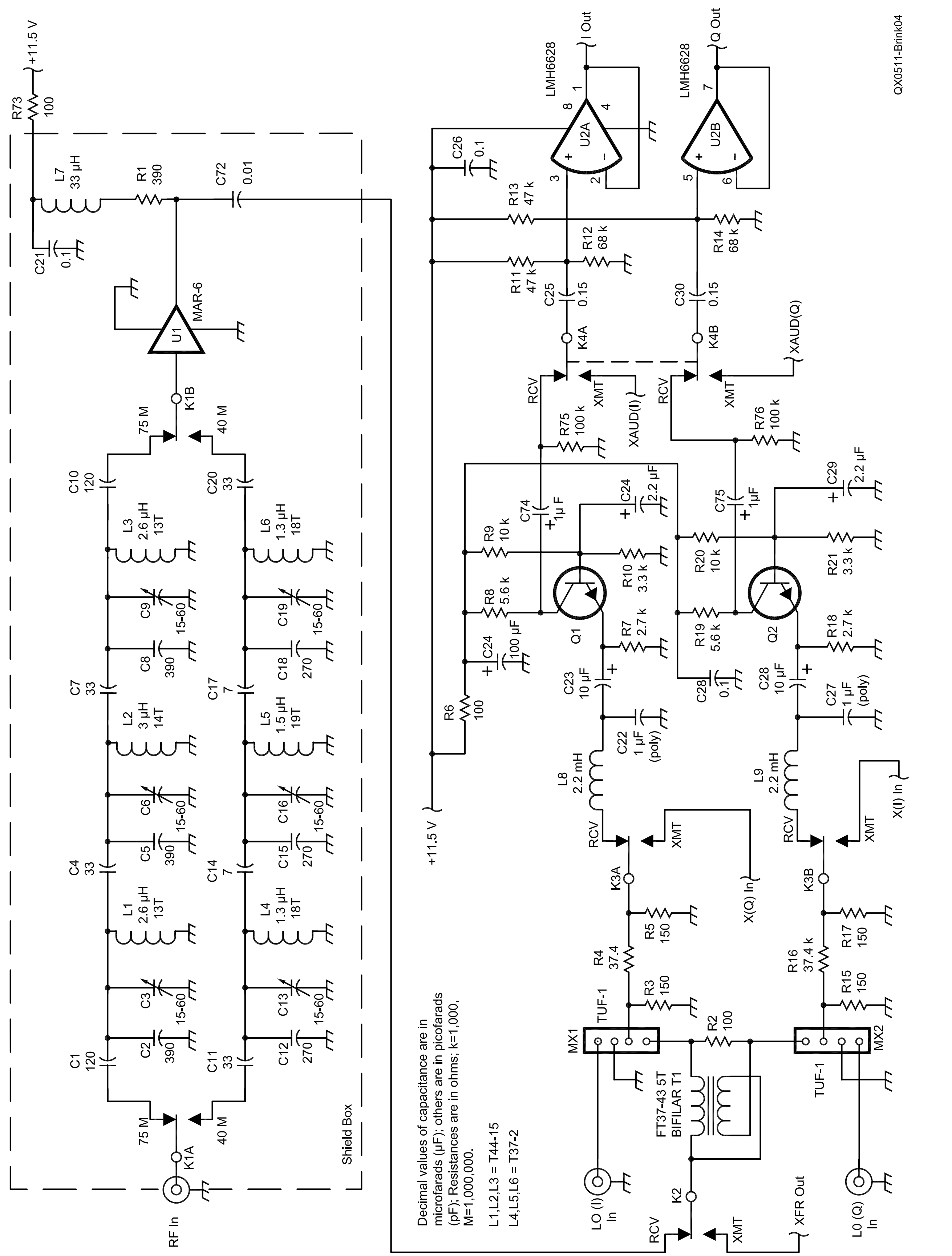


Fig 4—Receiver front-end schematic. The preselection filters and RF preamplifier are contained within a shield box, which was later found to be unnecessary.

Okay, but how does it Perform?

I approached this project with a casual attitude. I wanted to explore direct-conversion and learn firsthand about phasing-method SSB generation. I didn't want to build the most sensitive receiver in the world, nor was I interested in achieving sky-high third-order intercepts. Using high-level mixers didn't particularly appeal to me, at least not on this project. I just wanted to build something that satisfied my curiosities, with reasonably good performance, and above all, *sounds good.*

I live in California, where we hear loud signals on 75 meters and 40 meters, but at least they're not *bombarding* us from all directions like they would if I lived in, say, Kansas, or worse, in Europe. So some of the receiver parameters that others strive hard to perfect are of no particular interest to me, other than academically.

That said, I did run a few measurements: MDS = –125 dBm. I didn't bother measuring third-order intercept; the important thing is that I've never experienced any front-end overload or *any* other artifacts from extremely strong signals.

I also measured and plotted the audio responses of both the receiver and transmitter. The receiver response was measured by first setting a signal generator and the receiver to the same frequency, then stepping the signal generator down into the receiver lower-sideband region while measuring the sinusoidal audio-output amplitude. The transmitter response was measured by connecting an audio generator to the line input and measuring the RF output amplitude as the audio frequency was varied. Results are shown in Figs A and B.

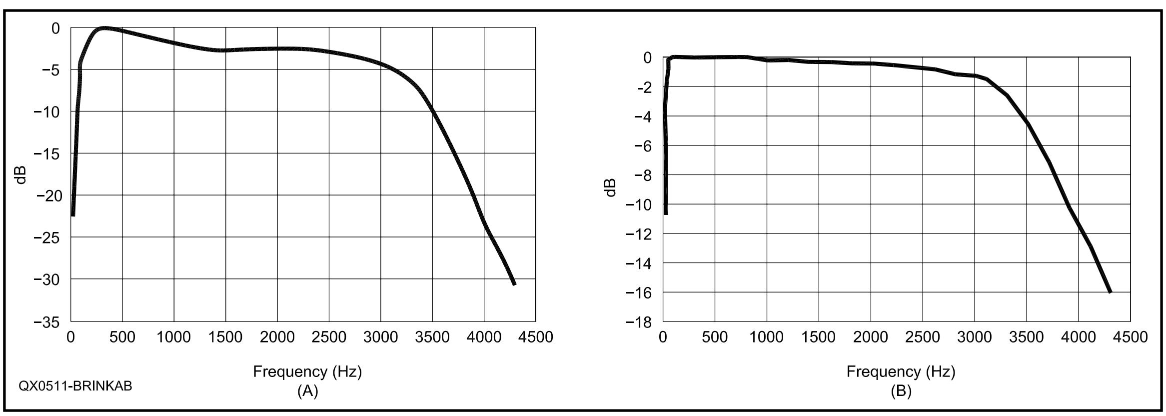
The transmitter has a little better response, because there is only one audio low-pass filter in its path, whereas the receiver has four, including the two at the mixer outputs. Consequently, there are more opportunities in the receiver for amplitude variations. Nevertheless, the receiver shows a –3-dB bandwidth of 100-2600 Hz, with upper-end audio extending to 3300 Hz, down –6 dB at that point.

It's interesting that since the mixer output frequencies extend literally to dc, receiver response *could* be extended lower by simply increasing the size of the various series coupling capacitors in the audio chain. *Why do this?* For one reason, there are practically no stations transmitting audio that low. For another, extending the audio bandwidth lower would just increase noise (60 and 120 Hz, in particular).

Making the Transmitter Sound Good

On the other hand, a little different attitude may be taken with respect to the transmitter. For openers, the audio response can be extended in both directions, provided it's not extended *too far* on the upper end. My transmitter response extends all the way down to 25 Hz at –3 dB. This may seem a bit extreme, given that nobody's voice (certainly not mine) extends that low. Nevertheless, the effect is to provide plenty of "headroom" on the bottom, so that external EQ (equalization) can be "pumped up" there if desired to enhance and "sweeten" the transmit audio.

In practice, I take this idea even further. I use condenser-type studio quality microphones and run them through commercial audio processing equipment to add EQ, slight compression, and downward expansion (set to act as a "soft" noise gate to kill background room noise). Besides "souping up" the low end with the EQ, I similarly increase the high end, as much as 12 dB. This overcomes the slight sag in the transmitter response and makes for crisper, more intelligible audio.



switched to either the 75 or 40 meter preselection filter. Each is a stagger-tuned, three-resonator band-pass filter. Fixed-value toroidal inductors were chosen here, for their higher Q compared to slug-tuned inductors. Trimmer capacitors are adjusted to obtain the desired band-pass response. The response curve for the 75 m filter (shown in Fig 5) is typical. The filters each have about 6 dB of insertion loss and the RF preamplifier has +18 dB gain, so the net gain is +12 dB. As a precaution, I placed the front end within a shielded box to isolate it from radiated LO energy that could unbal-

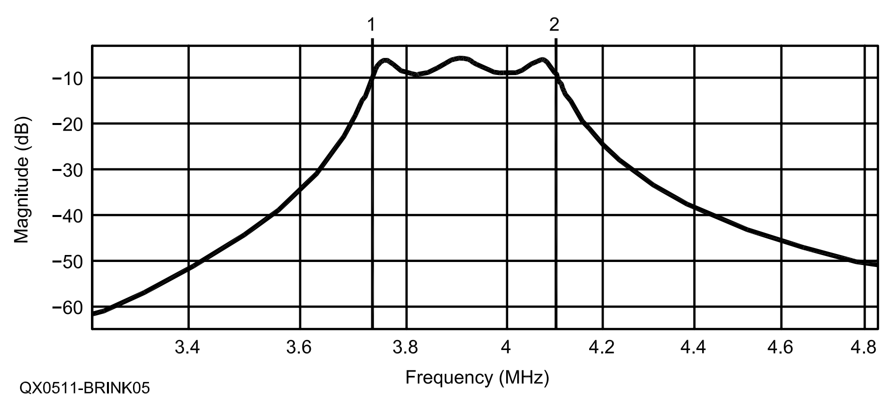


Fig 5—75 meter preselection filter response curve.

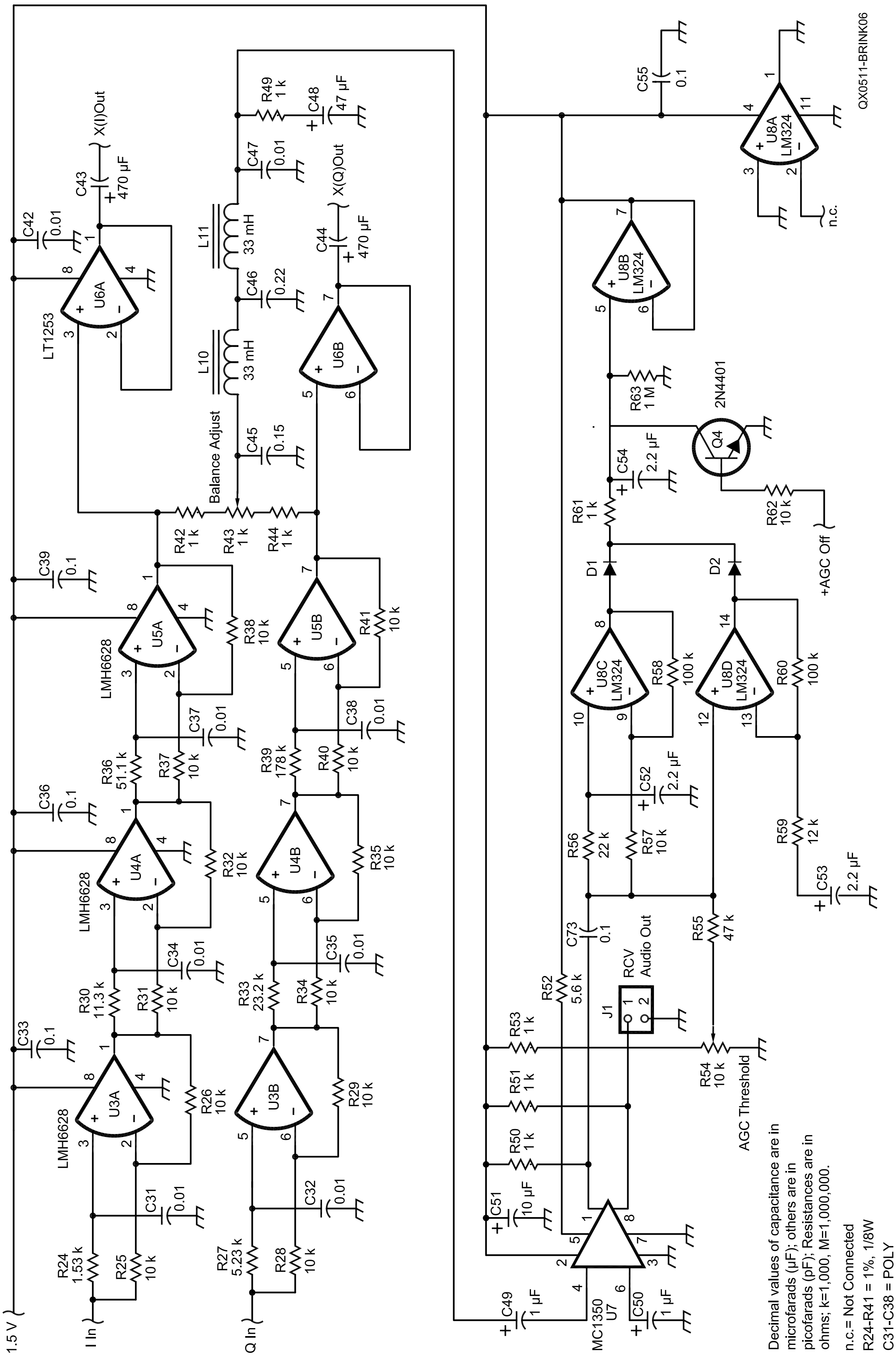


Fig 6—All-pass filter networks and AGC schematic. If ±1% capacitors are used in the all-pass networks, nearly 60 dB of opposite-sideband rejection can be achieved.

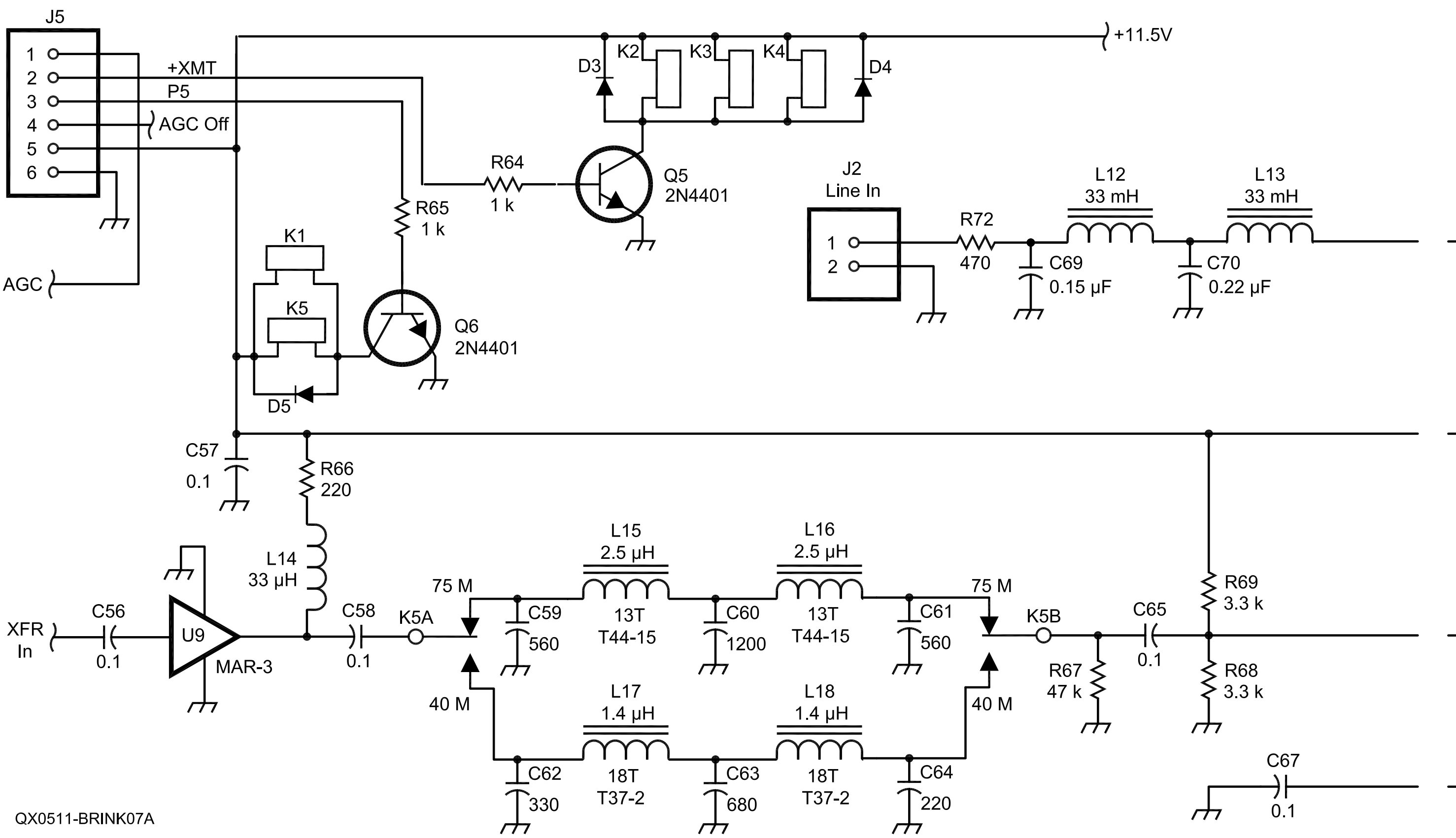


Fig 7—Relays, RF pre-driver, Transmit audio-filter schematic. The RF low-pass filters between U9 and U10 could be omitted, but they keep harmonics out of the PA.

ance the mixers and produce dc offsets at their IF output ports. (For this case, RF input frequency = LO frequency. Difference = 0 Hz = dc.)

The RF preamplifier output feeds T1, a T-splitter implemented with a 1:1 transformer. Mixers MX1 and MX2 are TUF-1 mixers, chosen over the more common SBL-1 because of better LO-to-RF port isolation. The SBL-1 performance would have been adequate, however. Good LO-to-RF port isolation is important for the same reason that stray LO pickup should be avoided: Both serve to unbalance the mixers and produce dc offsets.

The termination networks on the mixer IF-port outputs are very important. First, they must present a 50-Ω termination impedance (or close to it) to prevent reflections that would degrade opposite-sideband suppression. Second, they must function as audio low-pass filters to prevent strong off-channel signals from overloading the sensitive first-stage audio amplifiers and producing AM demodulation.[2] Others have used complicated diplexers to provide this needed termination impedance, but I found that a –6-dB pad followed by a simple LC

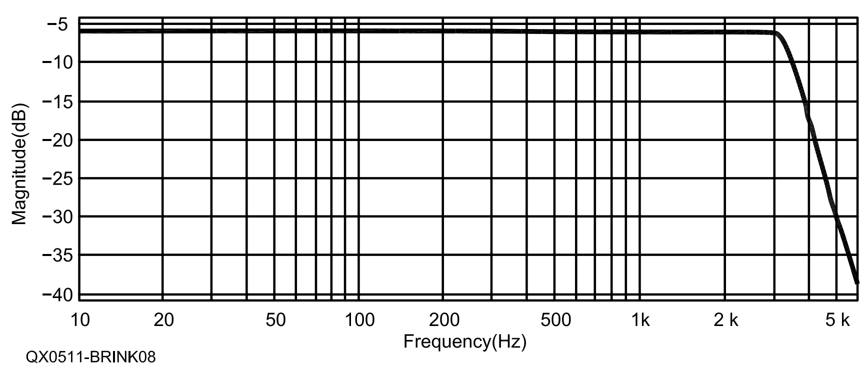
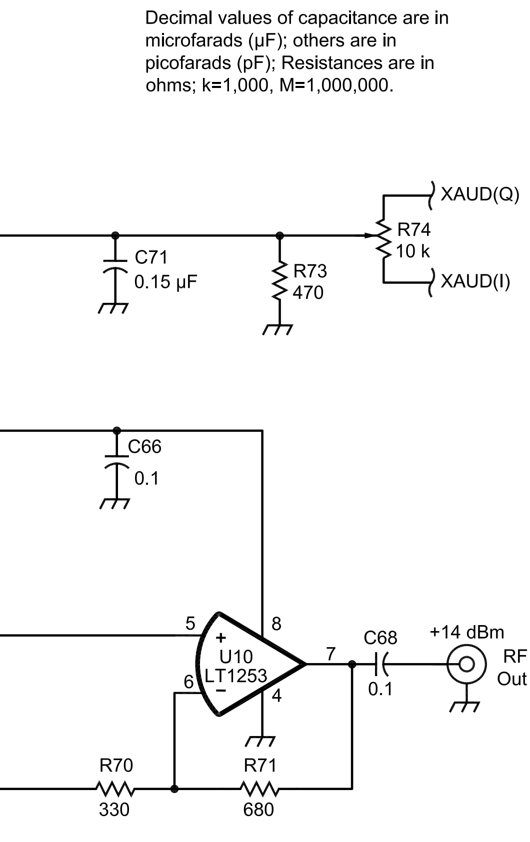


Fig 8—Audio low-pass filter response.

low-pass filter seemed to work well.[3]

From here, the signals feed common-base amplifiers with gains of about 40 dB. They are then switched through relay K4 to unity-gain buffers, U2. The buffer outputs go to the input of the all-pass networks consisting of U3, U4, U5 and associated components, shown in Fig 6. These networks pass all audio frequencies and shift the phase 90° relative to each other over a range of about 270-3600 Hz. Notice the use of 1% resistors. If the capacitors were within 1% tolerance (they aren't), the phase errors over this range would be small enough to achieve 58 dB of opposite-sideband rejection, provided everything else in the receiver was perfect. (Of course, it isn't, either.) After adjusting amplitude imbalance and tweaking the 90° LO phasing (more on that later), I was able to achieve about 45 dB of sideband suppression.



This has proven good enough.

The I and Q signals at the outputs of the all-pass networks are summed into a balance pot, where slight amplitude unbalances occurring upstream are nulled out. At this point, the signal is lower-sideband audio. It passes through a low-pass filter (L10 and L11 and associated capacitors), which improves selectivity and reduces noise before U7 provides up to 50 dB additional gain as controlled by the AGC.

AGC

There is no IF in a direct-conversion receiver, so AGC must be derived from audio. In general, audio-derived AGC is more difficult to get working correctly than IF-derived AGC, because there are fewer sample points in a given period. This is especially true with low-frequency audio. Consequently poor attack-time characteristics can cause problems with the leading edges of loud low-frequency

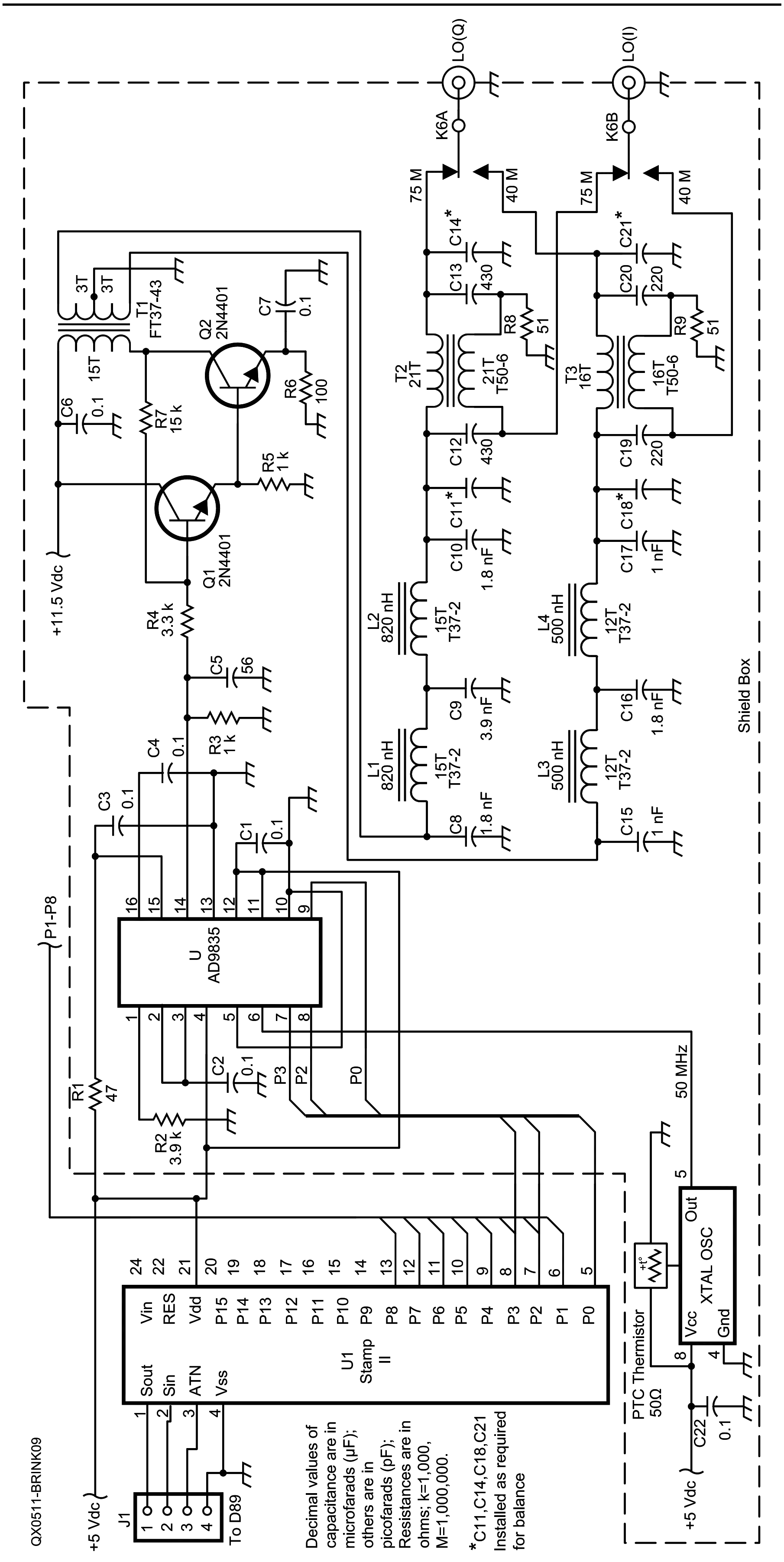


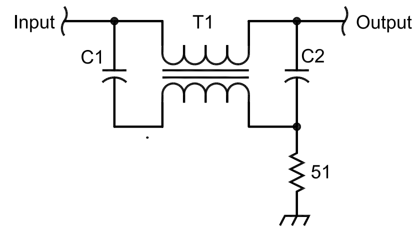
Fig 9—Controller, DDS, LO filters, phasing transformers. All LO components are contained in a shielded box to minimize LO radiation.

audio bursts after periods of silence. To some extent these difficulties can be ameliorated by employing full-wave rectification of the audio signal instead of the more familiar half-wave technique. In addition, we must exercise care in setting up the correct gain and time constants in the AGC loop. I spent an inordinate amount of time getting these values right. I included the ability to shut off AGC by switching U8B pin5 to ground with Q4. Of course, while running in this mode, I must reduce the volume to prevent eardrum blow-out from very strong signals.

The output from U7 is amplified 20 dB and inverted by U8C, while U8D provides equal amplification *without* inversion. The two outputs are summed through diodes that positively charge C54, with the attack time set by R61 and decay determined by R63.

Transmit Circuit Path

Because the all-pass filters are comprised of several components, I decided to time-share them between receive and transmit modes. The mixers are relatively expensive and they are shared, too. Sharing is accomplished by small PC-mount relays K2, K3 and K4. It takes about 20 milliseconds to switch the relays and allow for them



T1 WINDING INDUCTANCE = $\frac{50}{2\pi f_0}$, where f_0 = design center frequency

C1, C2 REACTANCE = $\frac{1}{100(2\pi f_0)}$

QX0511-BRINK10

Fig 10—Design formulas for quadrature hybrids.

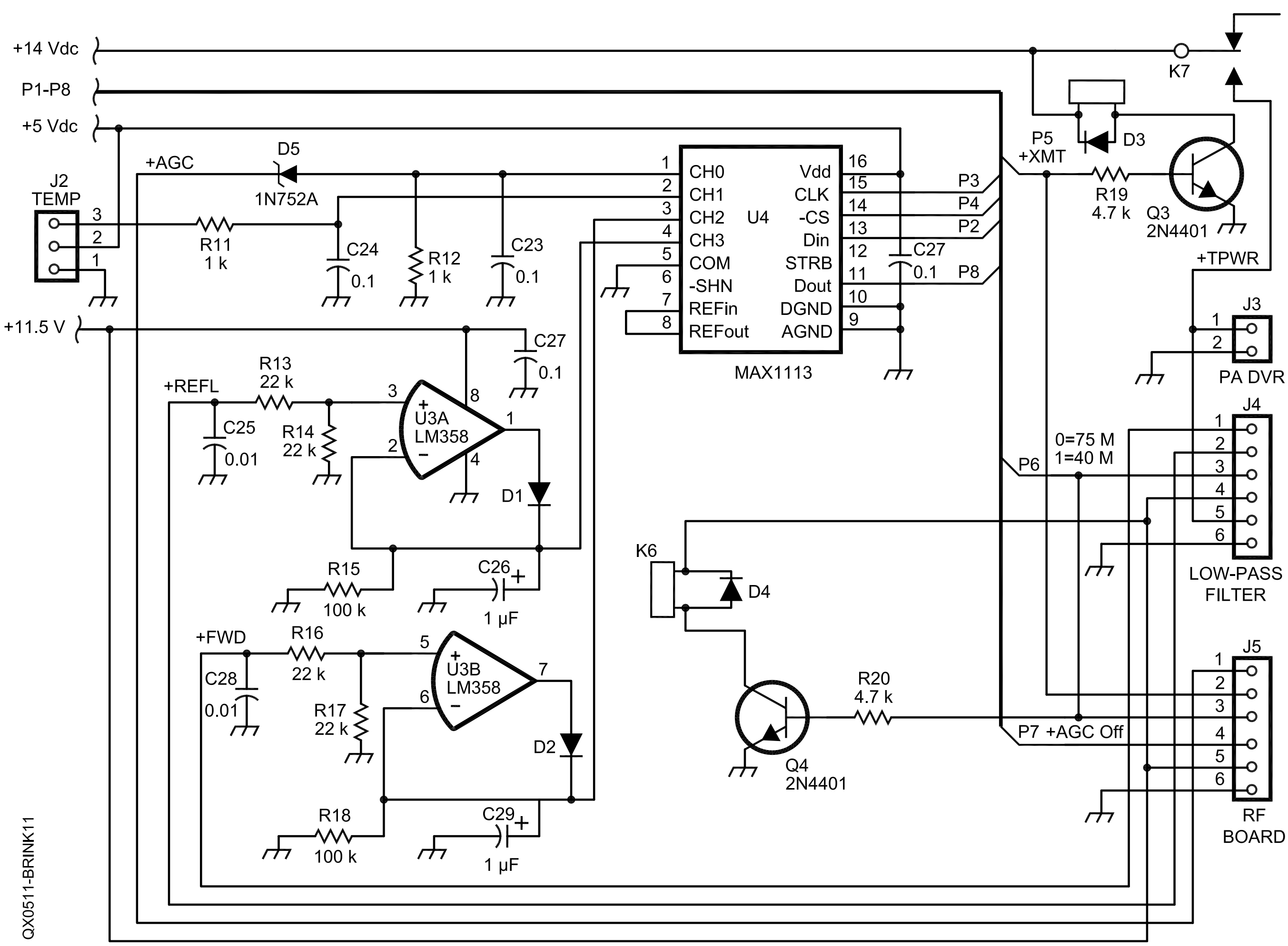


Fig 11—Analog measuring circuits. Reflected- and forward-power voltages are peak detected before measurement.

to settle, but this is completely unnoticeable when operating SSB. To prevent "pops" in the speaker, the receiver audio is muted during this interval by "turning down" the digital volume pot (U7 in Fig 3) to zero.

These relays plus K1 and K5, which provide band switching, are shown in Fig 7.

Line-level audio enters the transceiver and passes through a low-pass audio filter that rolls off the upper frequencies just above 3 kHz (see Fig 8). The audio is then split by pot R74 and routed through relay K4 to the all-pass networks. R74 is adjusted to compensate for amplitude imbalances that occur downstream and to maximize opposite-sideband rejection in the transmit signal. (I performed this step with a spectrum analyzer, but it could also be done just as effectively by modulating the transmitter with 1 kHz audio, tuning a nearby SSB receiver to the opposite sideband and adjusting for minimum tone.) The I and Q outputs from the all-pass network are amplified 6 dB by U6 (Fig 6) and pass back to the mixers through relay K3. The mixers up-convert them to RF and T1 combines their powers.

There are a few harmonics in the RF signal that appears at the output of U9, and I chose to suppress them with low-pass filters that are selected by relay K5 before further amplification by U10. Without these filters, the harmonics would have been present (and amplified) by all following stages. They probably would be suppressed to within legal levels by the output low-pass filters that follow the PA, but I didn't like having them "rattling around" in the PA driver and PA stages.

Detailed Circuit Description—Control Board

The Stamp II module shown in Fig 9 is the heart of the Control board. It connects to the PC via a 9600-baud serial port. Its 16 I/O pins can be programmed as either inputs or outputs. Pins P0, P2 and P3 connect to the DDS chip, U2. This chip contains various registers, including four eight-bit holding registers that accept the 32-bit frequency word. These four registers must be reloaded each time the carrier frequency is changed. The sequence is as follows:

1. Operator changes the frequency at the PC.
2. PC software issues a frequency change command to the Stamp II.
3. Stamp II lowers P0 to signal the DDS that a new frequency word is to follow.
4. Stamp II generates an 8-bit control byte followed by the first 8 data bits of the 32-bit word and shifts these 16 bits out on P2 along with 16 clocks on P3. The control byte tells the DDS which of the four holding registers is to be loaded.
5. The Stamp II repeats step 4 until all 32 frequency bits are loaded.

Thus, a total of 64 bits is required to change the DDS frequency.

The DDS is clocked from a 50-MHz crystal oscillator module. The module case is kept warm to about 95°F with a positive-coefficient (PTC) thermistor to minimize frequency drift.

With 32-bits resolution and running at 50 MHz, the DDS can be

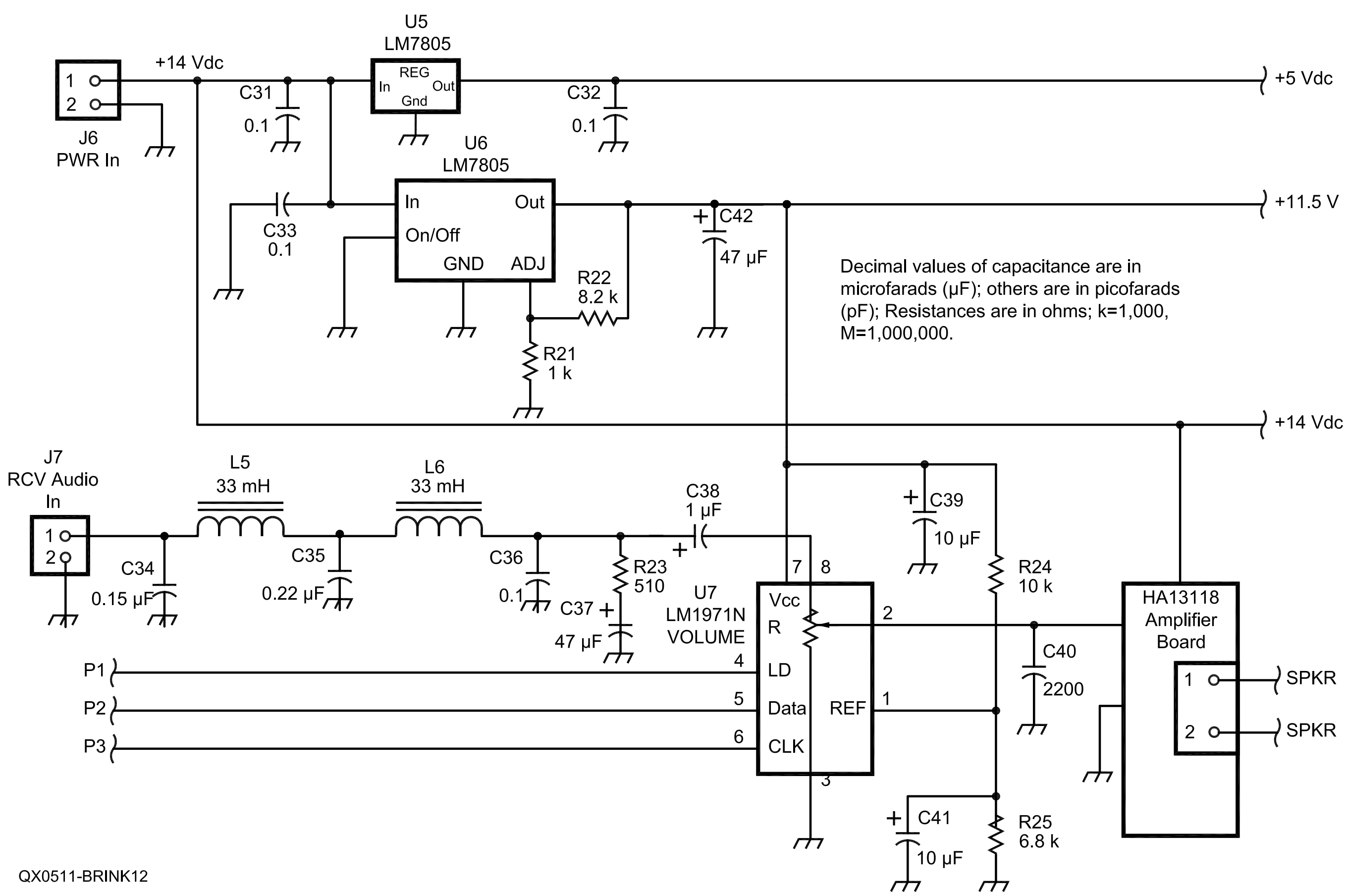


Fig 12—Power regulators, low-pass audio filter, digital potentiometer, speaker amplifier. The low-pass filter reduces noise and improves on-channel selectivity.

stepped in 0.01164 Hz increments. This is much greater resolution than is required. In my software, the minimum step value is 10 Hz. Load resistor R3 sets the level to 1.5 $V_{p\text{-}p}$ at the input to buffer amplifier Q1, Q2. The buffer output is filtered with selectable low-pass filters to remove aforementioned spurs developed in the DDS, and presented to two quadrature hybrids, one for each band. The hybrids generate the two I and Q local-oscillator signals that are phased 90° apart. The 90° phase shift is not perfect owing to differences in component values, so I made provisions in the PC board layout for additional "tweak" capacitors at both inputs and outputs of each hybrid. In practice, I found that just two capacitors were required—C14 and C21—to maximize opposite-sideband rejection. Design formulas for the hybrids appear in Fig 10.

Fig 11 shows the analog measuring circuits. The AGC voltage, which varies with receive signal strength, is dropped down by Zener D5 to fit within the 5-V measurement range of ADC chip U4. Temperature of the PA heat sink is measured by an LM34 sensor glued onto the heat sink with epoxy. Its output passes to the ADC through noise filter R11 and C24. Voltages proportional to reflected and forward power generated on the PA low-pass filter board are divided with resistive dividers and peak detected by U3. The +14 Vdc power input is switched on through relay K7 in transmit mode to energize the bias circuits for the PA driver and key the TR relay on the low-pass board. Relay K6 is energized in 40 meter mode to select the 40 meter preselection filter.

The AGC-controlled amplifier U7 on the RF board provides up to 50 dB gain in the audio path, but it also introduces some noise (not a lot, but some). To reduce the effect and to improve on-channel selectivity, another audio low-pass filter is placed after it. Fig 12 shows it connected ahead of digital volume control U7. This chip varies the volume from 0 to –62 dB in 1-dB steps. It has an audio-taper characteristic and is controlled by the Stamp II. It feeds a speaker amplifier that I bought as a small kit. The amplifier is capable of up to 5 W into an 8-Ω load. C40 suppresses audio clicks from U7 when the volume is being changed. Also shown in Fig 12 are two regulators for developing +5 Vdc and +11.5 Vdc.

RF Driver and PA Stages

The RF driver is shown in Fig 13. It consists of two cascaded push-pull stages and provides about 25 dB of gain. Emitter feedback in both stages helps stabilize dc biasing and broaden the frequency response. R4 and L1 help flatten the response.

Bias compensation is provided by diodes D1 and D2, temperature coupled to the Q4 and Q5 heat sinks. Quiescent bias levels are determined by potentiometer R1 and are set to 35 mA for Q4, Q5.

The power amplifier, shown in

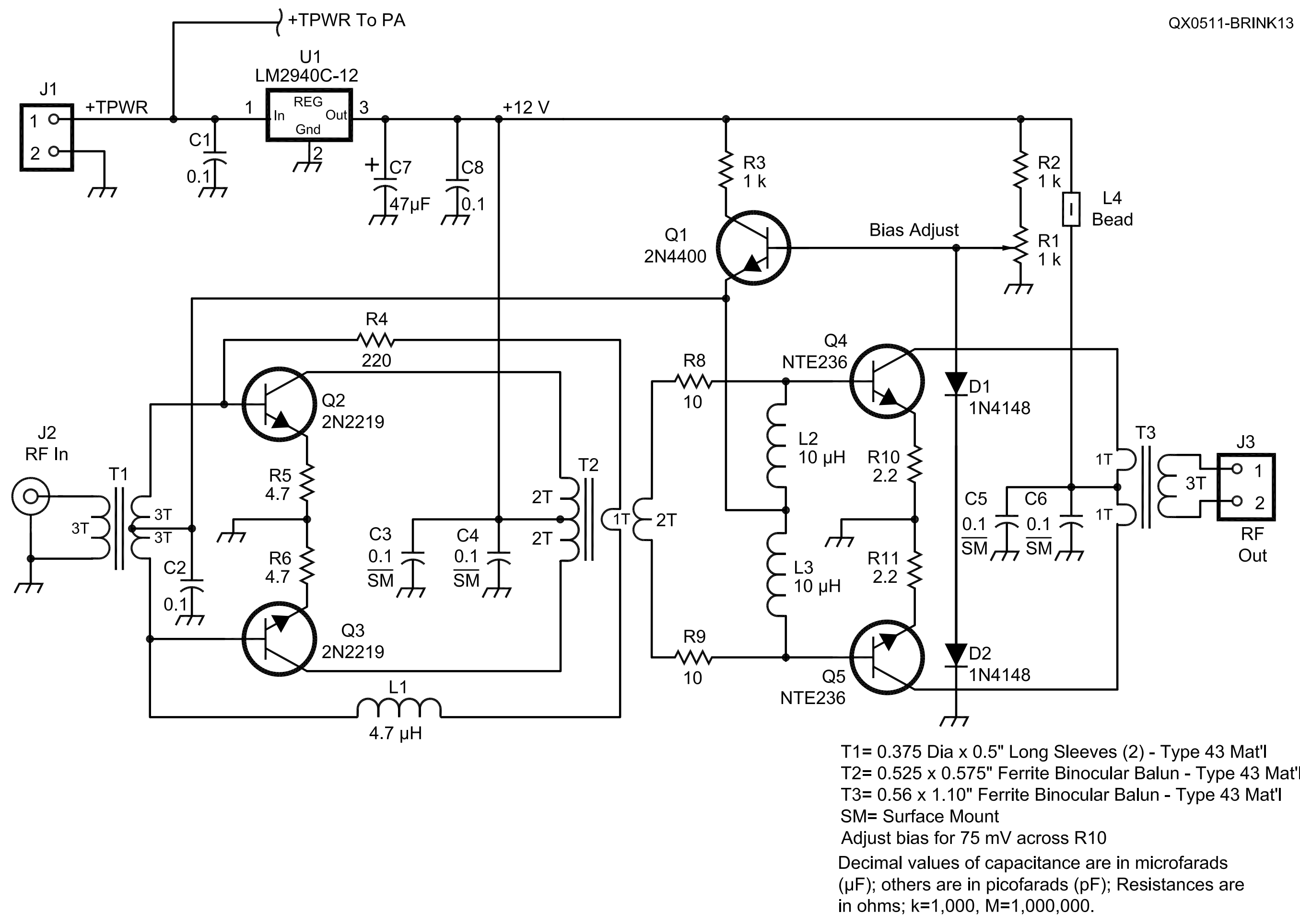


Fig 13—PA driver schematic. R4 and L1 were selected to optimize response flatness. Diodes D1 and D2 are placed in thermal contact with the Q4, Q5 heat sinks.

Fig 14, develops 100 W into a 50 Ω load. It's a straightforward push-pull design similar to one from the late Helge Granberg, K7ES, of Motorola. It uses two Toshiba 2SC2879s, available from various sources.[4] I mounted the transistors to an oversize heat sink to eliminate the need for a cooling fan. Transformer-coupled feedback is employed to achieve a broadband response. Transistor Q4 is mounted on the heat sink to provide bias compensation. Regulator U1 makes the bias very stable. R9 is adjusted for about 300 mA total collector current, which is required for good IMD performance.

To key the PA, voltage +TPWR is switched on at the Control Board and supplied to U1 through R11.

Output Filter, Directional Coupler, TR relay

Refer to Fig 15. These circuits are located on a separate PC board. The output filters are quasi-elliptical designs. Capacitors C2 and C9 steepen the roll-off characteristic, not tuned to second or third harmonics as in conventional elliptical designs.

The directional coupler is straightforward. C13 is adjusted for a null when operating into a 50-Ω resistive load (zero reflected power).

Firmware/Software

The firmware that runs in the Stamp II employs a *BASIC*-style instruction set and was developed using the PC-based editor supplied by Parallax, the Stamp II manufacturer. There are 115 lines of code, including the variable-statement field.

The PC software was developed using Microsoft *Visual Basic*. The Graphical User Interface (GUI) is a virtual transceiver control panel and is shown in Fig 16.

Interconnection Diagram

All connections to the transceiver are on the rear panel. Interconnections inside the enclosure are shown in Fig 17.

Packaging the Transceiver

This is where I really "lucked out." In a Bay-area surplus warehouse aptly named "Weird Stuff," I found a nice little chassis that once held a couple of external hard drives. There was a small switching power supply in the bottom and a fan on the rear panel. I saw that my PA/heat sink assembly would fit nicely in place of the power supply and in front of the fan (which I later found unnecessary, so it remains disconnected). A built-in shelf would later hold the PA low-pass filter board, and there is plenty of space along the sides for mounting the RF and Control boards. See Figs 18 and 19. What a perfect fit! The cost for the chassis was $10.

This was an enjoyable project. The results are described in the sidebar "Okay, but how does it Perform?"

Notes

[1]R. Campbell, KK7B, et al, *Experimental Methods in RF Design* (Newington: ARRL, 2003), Ch 9.

[2]I learned firsthand about AM demodulation by failing to insert the low-pass filters on my first prototype. What resulted was low-level sounds of strong off-channel signals that

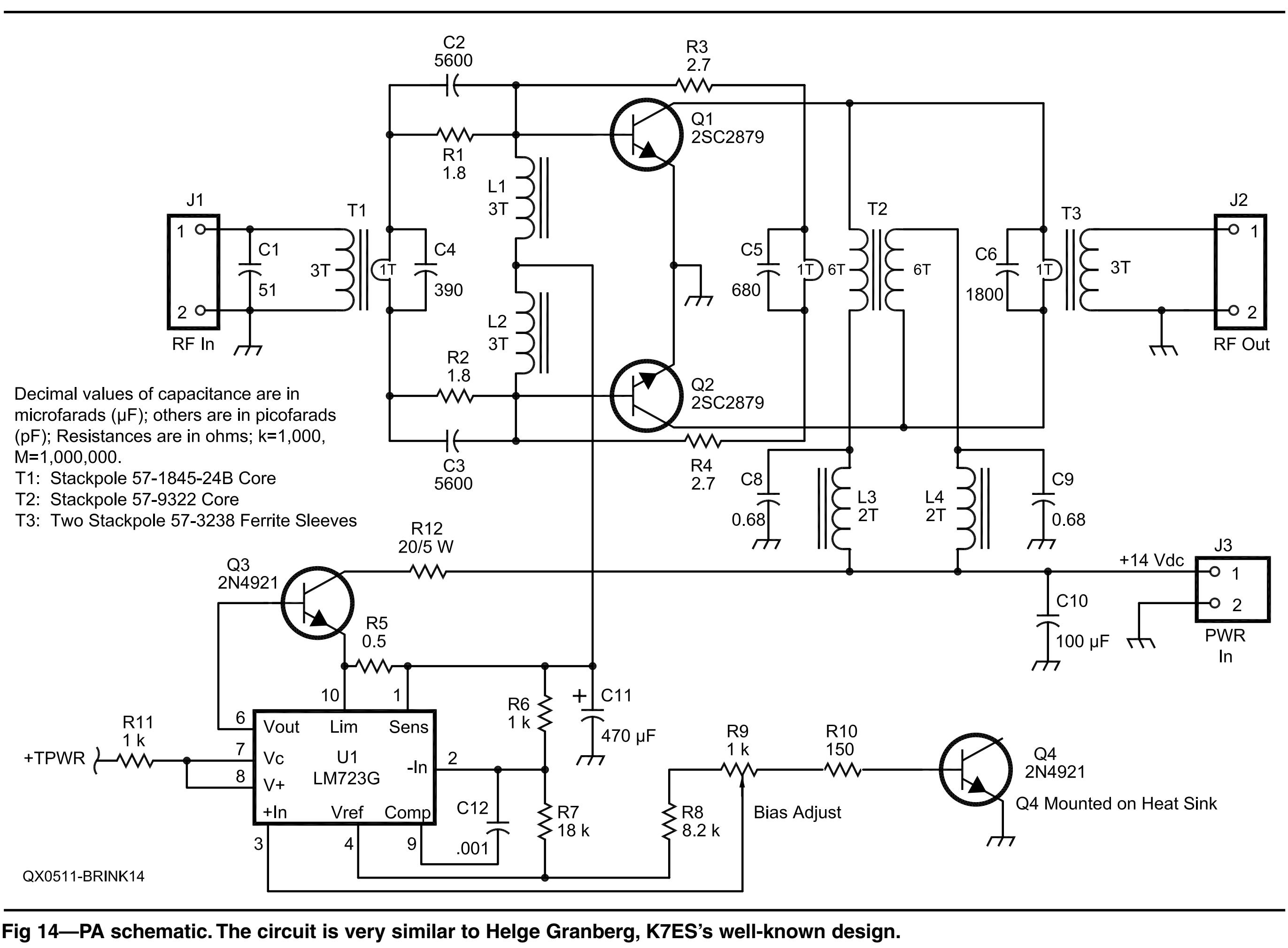


Fig 14—PA schematic. The circuit is very similar to Helge Granberg, K7ES's well-known design.

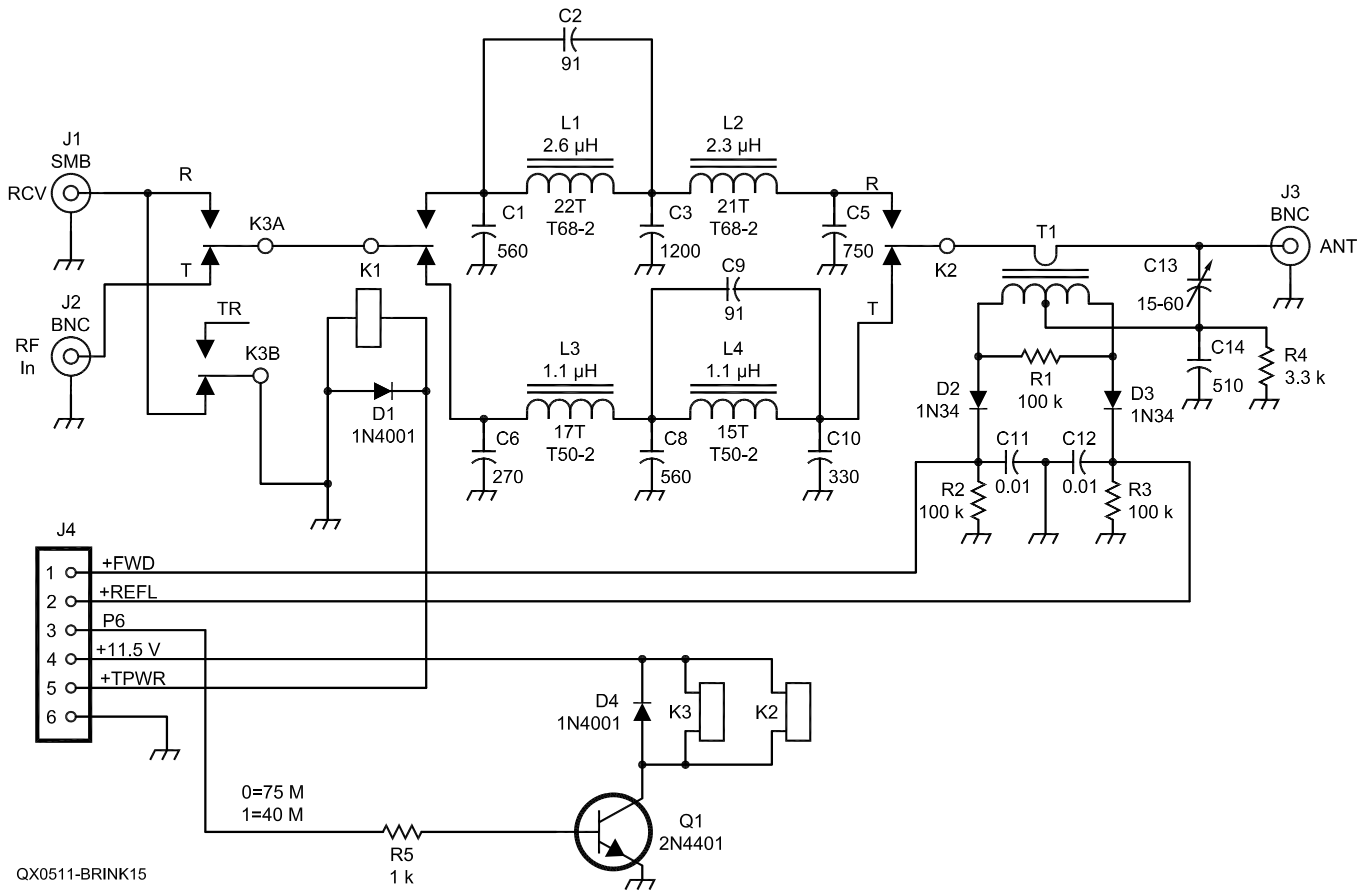


Fig 15—Output filters, directional-coupler schematic. With a 50-Ω load, C13 is adjusted for minimum +REFL voltage.

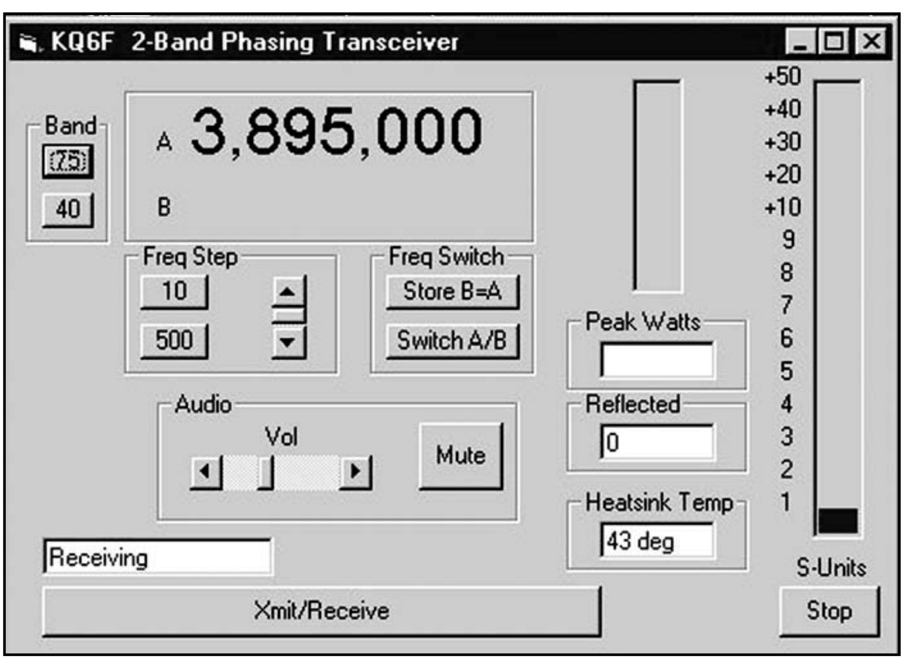


Fig 16—Graphical User Interface (GUI). PC software was written in *Visual Basic* and runs under *Windows*.

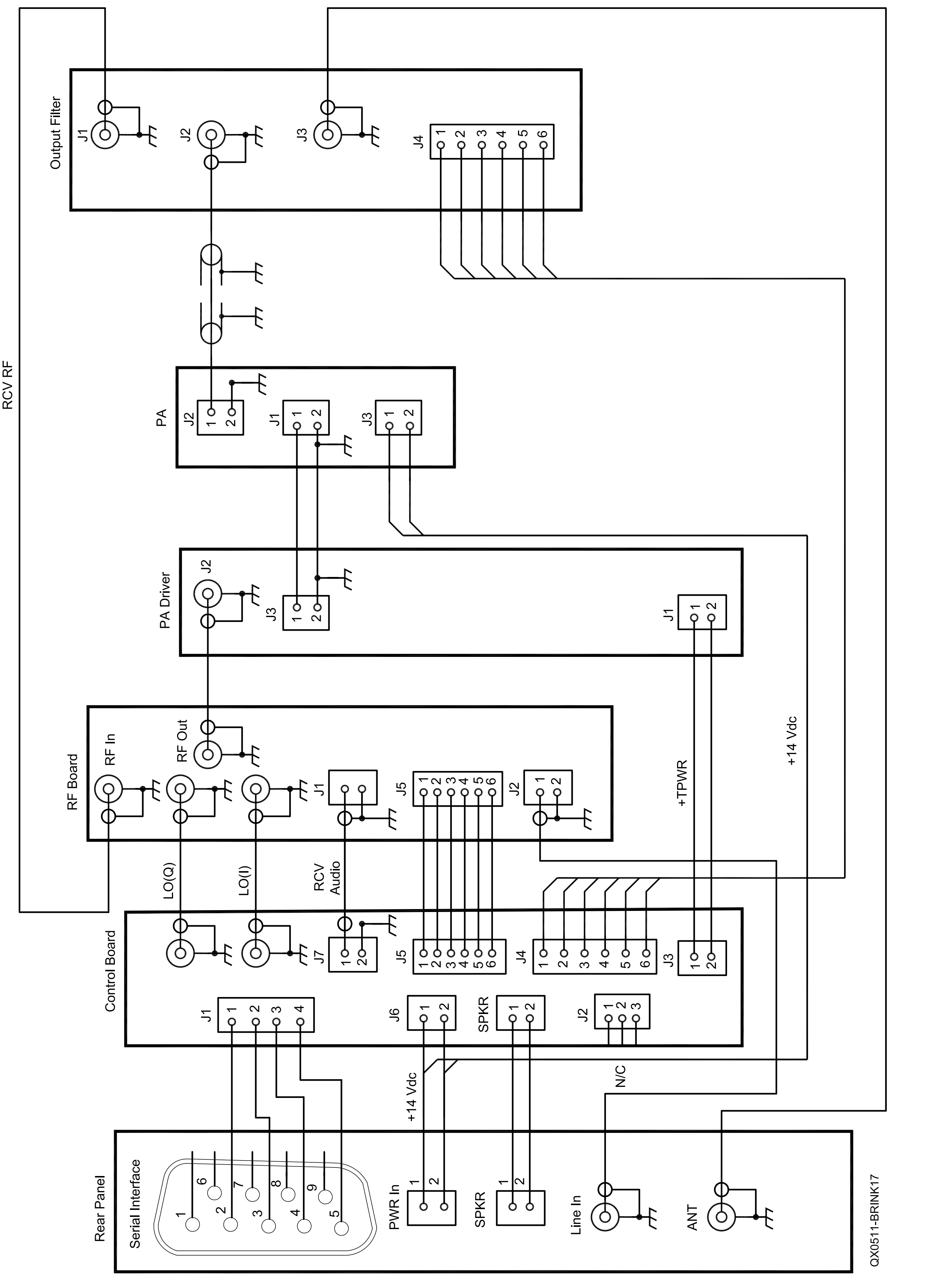


Fig 17—Interconnection diagram.

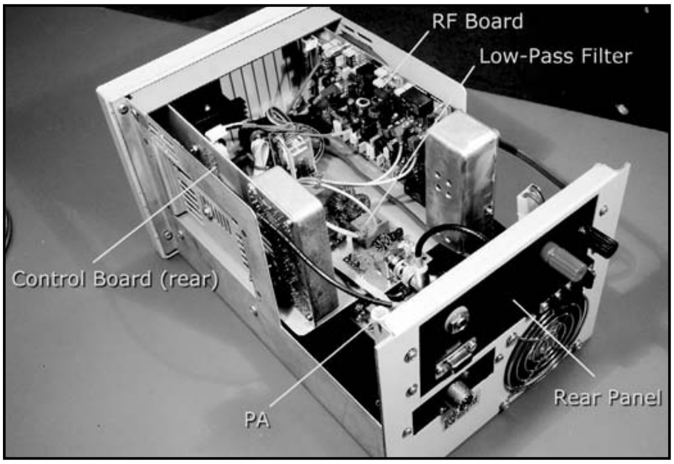


Fig 18—Interior view with covers removed.

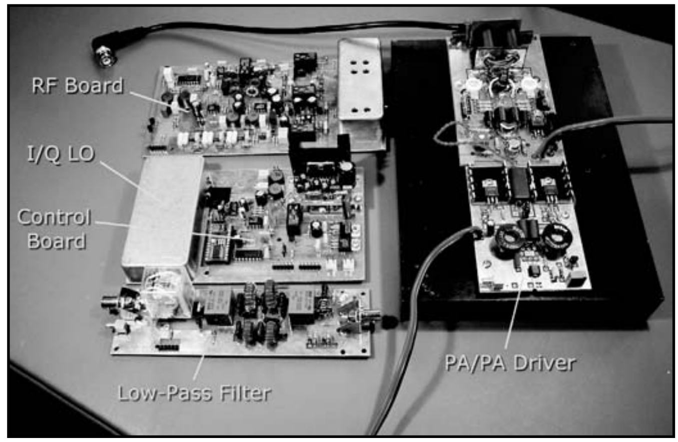


Fig 19—View of individual subassemblies removed from chassis.

sounded like what you get when you tune an SSB signal with an AM receiver. I struggled with this intractable problem for weeks before KK7B showed me the light in an exchange of e-mails. The problem occurs because the mixers are down-converting the whole range of signals within the selected band. Signals out to a few kilohertz are present at the mixer IF ports. The audio low-pass allows only the on-channel signals to reach the amplifier first stage.

[3]The inductors in this filter, given they are comprised of many turns and are at the input of a very high-gain audio amplifier, are extremely sensitive to stray magnetic fields. The raster sweep yoke in my computer monitor creates a buzz in the receive audio if placed too close to the rig.

[4]2SC2879 transistors are available from RF Parts and others. The original design called for the Motorola MRF454s, and they are still available, although no longer manufactured by Motorola. The entire PA can be purchased as a kit (AN762) from Communication-Concepts Inc. The kit includes the MRF454s. The 2SC2879s give equivalent performance and are cheaper.

Rod Brink was first licensed in the early 1950s, but got busy with school and forgot all about Amateur Radio. After retiring in 1995, he re-licensed as KE6RSF, then AD6TV and now KQ6F. He holds BS(EE) and MS(EE) degrees and worked 37 years for various electronic companies in Silicon Valley, serving as Vice-President Engineering for the last 20 years of his work life. He now spends much of his leisure time designing and building computer-controlled radios. □□

Quadrature Phase Concepts

A mathematical analysis of phasing techniques to reduce a particular sideband.

By Gary Heckman, KC7FHP

Why Math?

This article will discuss the mathematical analysis behind the phasing, multiplexing, and summation techniques used to reduce a particular sideband. This article will also discuss some techniques to eliminate part of a received signal, and why they don't work. While math is not everyone's delight, in some instances it can show how and why things work in ways that circuit explanation can not.

This article will be of particular interest to hams who build or want to build their own receivers. The problem is that once you build a receiver, you listen to it, and it sounds fine; you don't really know if the principles behind the receiver's design are performing or not. When you hear other received stations, you assume all is okay. This article encourages the builder to beware.

If the math shows that a receiver's design will work, chances are that you can make it work. If the math shows that a receiver's design will not work, there is very little chance that you can get it to work the way you expect, even if you do hear the talk of other received stations.

302 E 4350 S
Washington Terrace, UT 84405-5647
kc7fhp@arrl.net

The Basics

There are four basic equations that we will use in this analysis. These equations can be found in college trigonometry books. I went to the library and looked in the reference section to find a likely book that had the equations.[1] These equations are the starting place for this type of analysis. See Table 1.

Notice how in Equation A the Sine of the sum of two quantities results in the sum of Sine and Cosine products while in Equation C the Cosine of the sum of the same two quantities results in the difference of Cosine products and Sine products. That is the way the relationships are, and it is not a mistake. You can prove the accuracy of the relationship identities with a calculator. Plug in some angle for "C" and another angle for "A" and you will find that they are accurate.

Eliminating a sideband

Fig 1 shows how to generate a particular Single Sideband (SSB) signal in a transmitter. The audio signal and the carrier (oscillator) signal are shown being phased +45 degrees and –45 degrees because it is generally easier to physically maintain the levels and perform a small phase

[1]Notes appear on page 23.

difference on each signal than it would be to phase one signal by 90 degrees. The result in either case, however, would be the same.

The audio signal to the upper multiplexer (shown as an "X" with a circle around it) will be a Cosine-wave signal at some point in time. The audio signal to the bottom multiplexer will be a Sine-wave signal at that same point in time due to the phasing. A Cosine waveform and a Sine waveform are each 90 degrees or one quadrant different in phase. Please note: It is *not* easy to phase an audio signal the same amount over a band of frequencies 300 Hz to 3000 Hz, unless digital principles are used.

The carrier signal is likewise phased. It is easy to phase a constant frequency like a carrier. Have the carrier signal to the upper multiplexer be a Cosine-wave signal and have the carrier signal to the lower multiplexer be a Sine-wave signal. The result of the multiplexer action is to multiply the two signals applied to it. This means we have Cos(C)*Cos(A) as the output of the top multiplexer in Fig 1. Please be mindful of the simplified notation: "C" represents 2π times time "T" times the carrier frequency "c." "A" represents 2π*T times the audio frequency "a." The Cosine of a specific frequency will yield a specific number. We want the continually varying waveform, typically with hills and valleys. To get that kind of continuously varying waveform we need time in the function. The 2π value is one radian cycle. The use of capital letters to substitute for something else simplifies the math. I will show the math in general terms in this article using letters but the examples in the Figs will show specific frequencies. Look at the equations in Table 1 to see what we have to do to get a Cos(C)*Cos(A) result. We have to add Equation C and Equation D in Table 1 and we get

$$2 * \text{Cos(C)} * \text{Cos(A)} = \text{Cos (C + A)} + \text{Cos (C – A)} \text{ or,}$$

$$\text{Cos(C)} * \text{Cos(A)} = \frac{\text{Cos(C + A)} + \text{Cos(C – A)}}{2} \qquad \text{[Eq 1]}$$

We expect this. This is the familiar carrier plus and minus the audio signal. It is the upper sideband added to the lower sideband. In a typical multiplexer, the carrier will sneak through and the result will be an amplitude modulated signal with carrier. AM is a typical result of multiplexing an audio signal with a carrier.

Now look at the lower multiplexer. What do we have to do with the basic equations to get a Sin(C)*Sin(A) result? We have to subtract Equation C from Equation D in Table 1 and we get

$$2 * \text{Sin(C)} * \text{Sin(A)} = \text{Cos(C – A)} - \text{Cos(C + A)} \text{ or,}$$

$$\text{Sin(C)} * \text{Sin(A)} = \frac{\text{Cos(C – A)} - \text{Cos(C + A)}}{2} \qquad \text{[Eq 2]}$$

We have a lower sideband okay, but we have an upper sideband that is inverted or 180 degrees out of phase. In practice, tuning off frequency would be required to detect this strange AM signal.

At the point in Fig 1 where the results of the upper multiplexer are added to the results of the lower multiplexer the result is the same as adding Equation 1 and Equation 2. The right side of the sum of Equations 1 and 2 is Cos(C)*Cos(A) + Sin(C)*Sin(A). The left side of the sum of Equations 1 and 2 is Cos (C – A). The result Cos(C–A) is only the lower sideband and the upper sideband is missing.

This is an acceptable way to generate a single sideband signal. It does not appear to be as popular as using crystal filtering to generate SSB,[2] but it is mentioned in literature.[3]

Likewise, instead of adding the results of the two signals from the multiplexers, we subtract them, we get Cos(C+A). This is only the upper sideband and the lower sideband is missing. So by adding or subtracting the results of the two multiplexers we get the lower or upper sideband, exclusively.

This is how and why the configuration of Fig 1 works.

Table 1—Four Trigonometric Identities

Equation A	Sin(C+A)	=	Sin(C)*Cos(A)+Cos(C)*Sin(A)
Equation B	Sin(C–A)	=	Sin(C)*Cos(A)-Cos(C)*Sin(A)
Equation C	Cos(C+A)	=	Cos(C)*Cos(A)-Sin(C)*Sin(A)
Equation D	Cos(C–A)	=	Cos(C)*Cos(A)+Sin(C)*Sin(A)

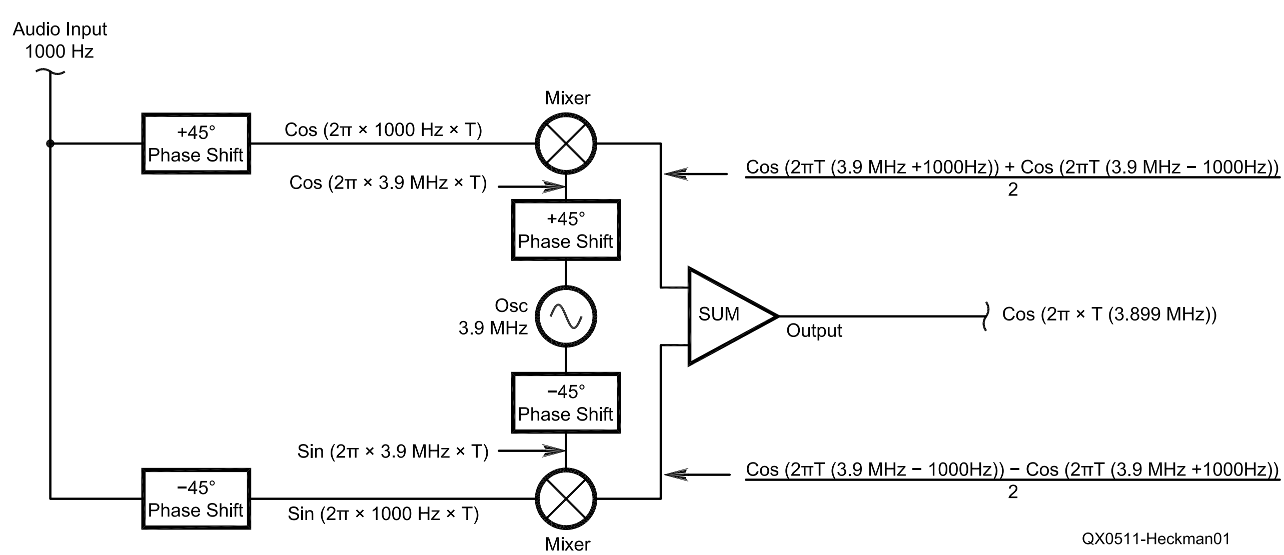


Fig 1—SSB quadrature phasing that provides either USB or LSB output

Trying to Eliminate Part of the Incoming RF Spectrum

Ready for a little more complexity? Consider Fig 2. Assume that the input RF signal is a broad band of frequencies and we are interested in a particular frequency represented by the letter "f." Again F=2πfT. Further assume that the RF frequencies are Cosine signals. We can make this assumption because at some time each will be a Cosine signal. The RF signal can then be represented by Cos(F). The upper phase shifted oscillator signal can be represented by Cos(O) at some time. At that same time, the bottom phase shifted oscillator signal can be represented by Sin(O).

The output of the upper multiplexer in Fig 2 is then Cos(F)*Cos(O). Looking at the trigonometric identities in Table 1 we see that the upper multiplexer output can be represented by

$$(\cos(F+O) / 2) + (\cos(F - O) / 2).$$

To do this, I am expecting you to mentally exchange "F" for "C" and "O" for "A" in the Table 1 equations. The high frequency result of F+O can be filtered out, leaving

$$\cos(F - O) / 2 \qquad \text{[Eq 3]}$$

as the result of the upper multiplexer.

The output of the lower multiplexer is the product of the two inputs or Cos(F)*Sin(O). Looking at the trigonometric identities in Table 1 we see that the lower multiplexer output can be represented by

$$\sin(F+O) / 2 - \sin(F - O) / 2.$$

The high frequency result of F+O can again be filtered out, leaving

$$-\sin(F - O) / 2 \qquad \text{[Eq 4]}$$

as the result of the lower multiplexer.

The top signal, Equation 3, is phase shifted by + 45 degrees, so the result is Cos(F–O+45) / 2. Using Equation C of the trigonometric identities, exchanging (F–O) for "C" and 45 degrees for "A," this signal is the same as

$$\cos(F - O)*\cos(45) / 2 - \sin(F - O)*\sin(45) / 2 \qquad \text{[Eq 5]}$$

result of the upper phase shifted multiplexer. Remember to replace "A" by 45 degrees. Both Sin(45) and Cos(45) are $1/\sqrt{2}$.

The bottom signal Equation 4 is phase shifted by –45 degrees, so the result is –Sin(F–O–45) / 2. Using Equation #B of the trigonometric identities, this signal is the same as

$$-\sin(F - O)*\cos(45) / 2 + \cos(F - O)*\sin(45) / 2 \qquad \text{[Eq 6]}$$

The result of lower phase shifted multiplexer.

Both Sin(45) and Cos(45) are $1/\sqrt{2}$. Did you catch that the negative second term of Equation B is multiplied by the negative sign of the Sine function making it a positive? We started out with –Sin (F – O – 45) / 2, right?

When we add the results of the upper and lower phase shifted multiplexer signals, add Equations 5 and 6, we get

$$-\sin(F-O)/\sqrt{2} + \cos(F-O)/\sqrt{2}$$

Assuming F>O so F–O is positive, or assuming F<O so F–O is negative, it does not matter. No part of the incoming frequency spectrum "f" is eliminated.

What we Wish we Had

Using direct conversion, f=o. What we want is something like

$$\sin(F - O) + |(\sin(F - O)|$$

A sideband of "f" that is higher in frequency than "f" ends up positive when "o" is subtracted from "f" and comes through. A sideband of "f" that is lower in frequency than

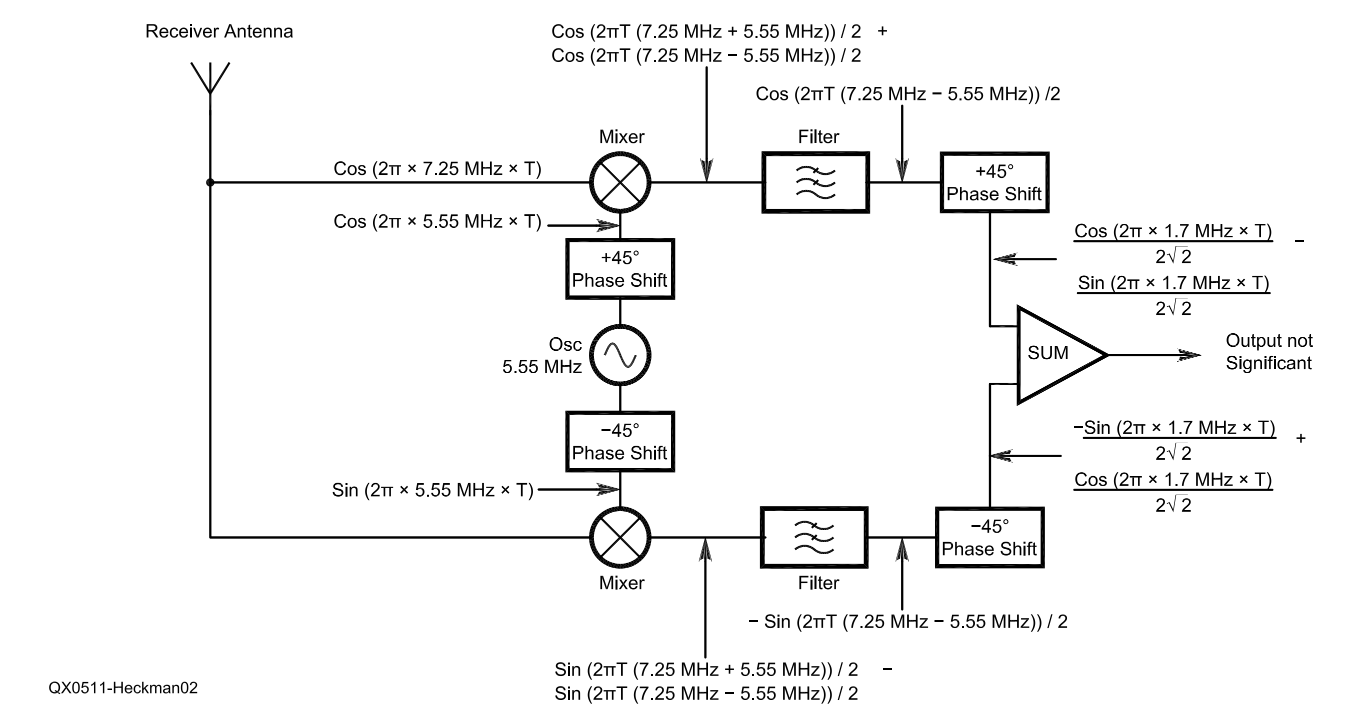


Fig 2—An attempt at reducing a portion of the incoming RF signal

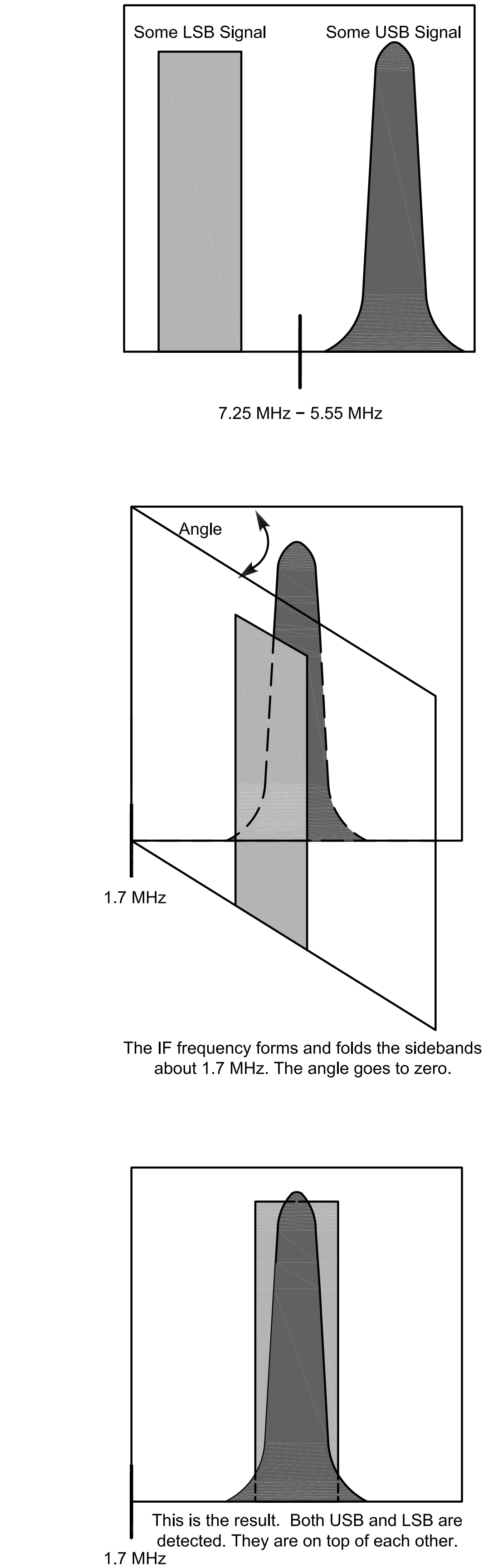


Fig 3—A visual explanation of baseband folding.

QX0511-Heckman03

"f" ends up negative when "o" is subtracted from "f." The Sin of a negative number is negative, but the absolute value is still positive and the two Sine terms cancel. This means that part of the incoming RF spectrum is eliminated. That is what we wish we had, but it is not there.

What we end up with is the baseband folded about f–o. See Fig 3. A humped signal above "f" in frequency or a square wave signal below "f" in frequency is folded about f–o when we multiplex by "o." We get both signals coming through the IF passband.

For example: Suppose the received signal is 7.25 MHz + a hump signal higher in frequency than 7.25 MHz. Further suppose the oscillator is at 5.55 MHz. The IF frequency is 1.7 MHz + the hump signal. Filtering eliminates the upper sideband of 7.25MHz + 5.55 MHz = 12.8 MHz. The IF passband will also accept a RF signal at 7.25 MHz—a square signal lower in frequency than 7.25 MHz. The result is an IF passband at 1.7 MHz plus a hump signal and a square signal simultaneously as seen on a spectrum analyzer. Unless additional filtering is employed, a signal at 3.85 MHz will also be detected (5.55 MHz – 3.85 MHz = 1.7 MHz). The multiplexer process will also produce 9.4 MHz (5.55 MHz + 3.85 MHz = 9.4 MHz). The important point is that no part of either the hump signal or the square-wave signal is eliminated.

Direct Conversion

Suppose that the passband about the incoming frequency "f" has a USB added to it and a LSB subtracted from it. The result from the upper multiplexer is:

$$\text{Cos(F+USB – O)*Cos(45) / 2 – Sin(F+USB–O)*Sin(45) / 2} \qquad \text{[Eq 7]}$$

See Equation 5.

$$\text{+ Cos(F – LSB – O)*Cos(45) / 2 – Sin(F – LSB – O)* Sin(45) / 2}$$

This is then added to the result of the lower multiplexer:

$$\text{Sin(F+USB – O)*Cos(45) / 2+Cos(F+USB – O)*Sin(45) / 2} \qquad \text{[Eq 8]}$$

See Equation 6.

$$\text{Plus –Sin(F – LSB–O)*Cos(45) / 2+Cos(F – LSB – O)* Sin (45) / 2}$$

Because O=F, the equations simplify to:

$$\mathrm{Cos(USB)} / (2\sqrt{2}) - \mathrm{Sin(USB)} / (2\sqrt{2}) \text{ plus}$$

$$\mathrm{Cos(-LSB)} / (2\sqrt{2}) - \mathrm{Sin(-LSB)} / (2\sqrt{2}) \qquad \text{[Eq 9]}$$

added to

$$-\mathrm{Sin(USB)} / (2\sqrt{2}) + \mathrm{Cos(USB)} / (2\sqrt{2}) \text{ plus}$$

$$-\mathrm{Sin(-LSB)} / (2\sqrt{2}) + \mathrm{Cos(-LSB)} / (2\sqrt{2}) \qquad \text{[Eq 10]}$$

The result is:

$$\frac{\mathrm{Cos(USB)} - \mathrm{Sin(USB)} \text{ plus } \mathrm{Cos(-LSB)} - \mathrm{Sin(-LSB)}}{\sqrt{2}}$$

Again there is no cancellation of either the imbedded USB signal or the imbedded LSB signal.

Last Try to Eliminate Part of the Incoming RF Spectrum

Will more quadrature phasing help? Suppose we modify Fig 2 and make Fig 4. I am going to speed up the math so that I don't annoy the reader to death.

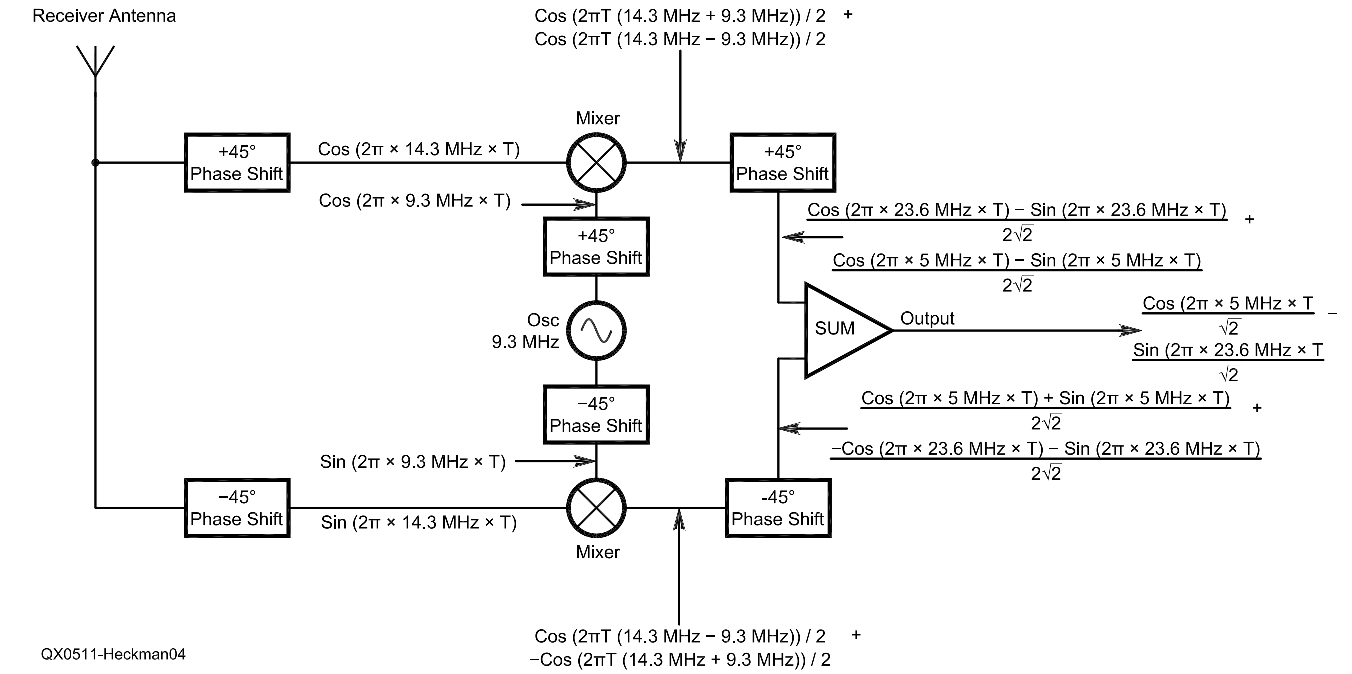


Fig 4—A last attempt at reducing a portion of the incoming RF signal.

Cos(F)*Cos(O) = Cos(F + O) / 2+ Cos(F – O) / 2
Sin(F)*Sin(O) = Cos(F – O) / 2–Cos(F + O) / 2

Notice here that if these two equations are added, the upper sideband resulting from the multiplexing process is eliminated. However, no part of the incoming RF band "f" is eliminated. In other words, we can add and eliminate (F + O) but f is still there in (F – O). But we are not ready to add them just yet; we have yet to phase shift the signals from each multiplexer.

Cos(F + O + 45) / 2 = Cos(F + O)*Cos(45) / 2 –Sin(F + O) *Sin(45) / 2 [Eq 11]

Cos(F – O + 45) / 2 = Cos(F – O)*Cos(45) / 2 – Sin(F – O) *Sin(45) / 2 [Eq 12]

Cos(F – O – 45) / 2 = Cos(F – O)*Cos(45) / 2 + Sin(F – O) *Sin(45) / 2 [Eq 13]

and

–Cos(F + O – 45) / 2 = –Cos(F + O)*Cos(45) / 2 – Sin(F + O)*Sin(45) / 2 [Eq 14]

Adding the resulting equations we get

Cos(F – O)*Cos(45) – Sin(F + O)*Sin(45) or
$\mathrm{Cos}(F - O) / \sqrt{2} - \mathrm{Sin}(F + O) / \sqrt{2}$

What we expected to see is a cancellation of terms when F – O became negative. It is not happening. The imbedded USB and LSB signals about "f" are not canceled. It regrettably will not work.

Conclusion

We can see why an audio signal properly phased and multiplexed with a carrier cancels a sideband. It is due to the change of sign of a term in the trigonometric identities in Table 1. But if we phase and multiplex an incoming RF signal with a carrier or BFO and even phase the signals again we can not eliminate the imbedded USB or LSB signal. All we accomplish is folding the RF signal about the resulting IF signal at f – o. It appears that the only way to eliminate an imbedded LSB signal is by narrow IF passband filtering. As a fellow ham, I would suggest that you not waste your time with these configurations. Yet, I am not going to be so arrogant as to say it is impossible to cut off part of the incoming RF signal mathematically; but it does not appear that math will support eliminating part of an incoming RF spectrum using quadrature phasing techniques. If there is a reader out there who knows how to do it, please let me know. I would love to be able to eliminate part of the incoming RF spectrum without a bunch of filtering.

Acknowledgment

I am grateful to Bill Whitworth for his friendship and insight, which was instrumental in developing this paper. Over the years we have had many great conversations and his suggestions are highly regarded.

Notes

[1]Ovid W. Eshbach & Mott Souders, *Handbook of Engineering Fundamentals*, John Wiley & Sons, 1 Wiley Dr, Somerset, NJ 08875-1272, p 274.

[2]Doug DeMaw, W1FB, "The Principles and Building of SSB Gear," *QST*, Sep 1985, p 18.

[3]John P. Froehlich, *Information Transmittal And Communicating Systems*, Published by Holt, Rinehart and Winston, Inc, 383 Madison Ave, New York, NY 10017, p 139.

Gary Heckman, KC7FHP, became interested in radio when his brother got a crystal radio kit for Christmas in the '50s. In 1963 Gary took a National Radio Institute course and paid for it by painting three of his father's buildings. Gary went on to obtain a degree in Electrical Engineering from the University of Nebraska. During a stint in the Army, Gary repaired radios, direction finding equipment, and even a digital clock. Gary then worked Civil Service for the Air Force for 31 years developing procurement documents and Statements of Work for installation of tens of millions of dollars' worth of microwave radios, fiber optic data transport systems, and radar equipment for a Bombing and Gunnery Range. In spite of life's pressures, Gary obtained his ham radio license in 1994.

A Single-Stage 1500 W Amplifier for 160 / 80 / 40 Meters

Do you want lots o' watts per buck? Here's your project!

By Frank A. Carcia, WA1GFZ

This article demonstrates the ease of designing and building a solid-state, high-efficiency power amplifier using inexpensive switching-power-supply FETs. (Use static-protective methods while working with FETs.—*Ed*) Although there is a wide range of possible operating voltages the user may select, the final goal here is to generate 1500 W output from about 50 V dc input. The gates are presently operated at zero bias, but there is provision for a positive bias input to operate the amplifier in class AB modes. This may be tried later, but the present application is for AM and CW modes only. Fifty to 60 W of drive power is required on 80 meters.

The input and output transformers are designed to cover 160 through 40 meters and handle at least 120 V peak input voltage during modulation peaks. The amplifier has been tested with three different output-transformer turns ratios:

- A 1:2 ratio produces about 500 W of output power,
- A 1:3 ratio produced about 800 W,
- A 1:4 ratio produces about 1200 W.

All of these tests were performed in the class E mode. Drain voltage was limited to 40 V during the testing. The shunt-loading control in the tank circuit allows a wide range of output power for each output transformer ratio.

The circuit layout has evolved over 15 years to minimize parasitic reactance. This is an important consideration in low-impedance RF circuitry. This layout has been scaled up from a 160-meter transmitter that I built in 1996 using 14 IRF840 FETs in push-pull parallel running 1500 W peak power on AM. The efficiency of that final is 90%. It has been in constant use since 1997, using PWM amplitude modulation with a two-phase current-mode feedback converter.

The amplifier is built on a 10×11-inch heat sink 3 inches tall with a total surface area of about 1200 square inches to minimize temperature rise. A slot machined in the center allows the output transformer secondary winding to pass between the two halves of the output transformer primary. This heat sink was sized to handle the extra power dissipation if the FETs are to be positive biased to operate in the linear-amplifier mode. A quick analysis of

181 Columbia Rd
Enfield, CT 06082
francis.carcia@hs.utc.com

temperature rise yields "130°C / W / square inch." This means there would be about 22°C of temperature rise if the final were dissipating 200 W. Because of the high efficiency of this amplifier, this heat sink has a considerable safety factor even at 1500 W output. Fig 1 shows the finished heat sink with slot and tapped holes for mounting the PC board and components.

Two 3×4×1/8-inch-thick copper heat spreaders connect the FET drains (via the case of each FET) to the 1/2-inch copper tubing of the output-transformer primary. The heat spreaders are silver-soldered to the transformer primary tubes. This arrangement greatly reduces shunt inductance, thereby increasing efficiency.

Class E operation requires a shunt capacitor on each FET drain to store and release energy over the cycle. The heat spreaders form one plate of these capacitors and the heat sink is the other. A sheet of 0.003-inch Kapton serves as the dielectric. The FET mounting screws also attach the spreaders to the heat sink and Nylon inserts are used at each FET to electrically insulate the mounting screw from the heat spreader. The heat spreaders also serve as low impedance connections to the drains and output-transformer primaries and help keep the temperature of each FET constant, so they share the load properly. Fig 2 shows the primary of the output transformer where it connects to the spreaders.

The output capacitance of the FETs plus the distributed capacitance of the heat spreaders provides enough shunt capacitance to operate class E on 80 meters. Although this amplifier could be used on 160 meters, additional external shunt capacitance would be necessary to limit the peak voltage on the FET drains. Further, the Kapton insulator would need to be thicker if this final was tuned to 40 meters.

The FETs used for this project are Fairchild FQA11N90 rated for 900 V_{ds} (drain-source voltage) and cost about $3 each in small quantities. Their power dissipation is 300 W at 25 °C, and their current rating is 11.4 A. Operating class E on 80 meters, the peak drain voltage during the "off" part of the operating cycle is about 175 V at 40 V dc on the final. Twenty-two of these devices are used in the amplifier, which provides a considerable safety factor, and the total cost of these parts was under $100. The cases of these devices are barely warm after a couple minutes of constant carrier output at 1000 W without a fan on the heat sink. Although I have not tried this amplifier on 40 meters, other amateurs have constructed 40-meter amplifiers using these FETs.

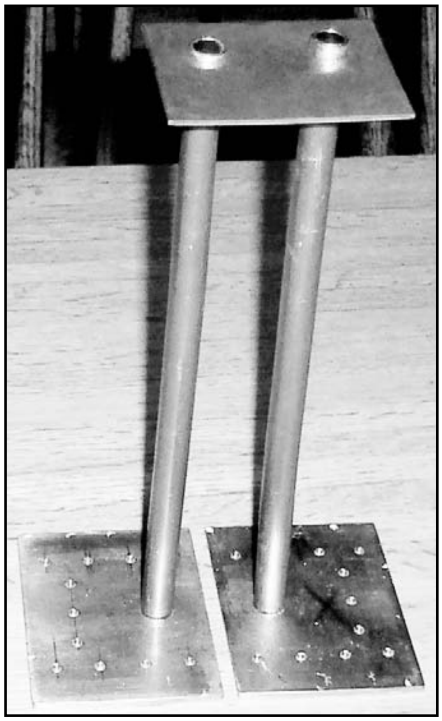


Fig 2—Primary of the output transformer where it connects to the heat spreaders.

The input capacitance of the FET gates is about 3000 pF each (resulting in an extremely low-impedance circuit), it is impractical to connect the gate drive with conventional hook-up wire. A PC board was built with two transmission lines on one side (one for each phase) and a ground plane on the other. A small area was left for the driver-transformer secondary center tap. This area is connected to ground but could be used as a point where positive bias is injected. Fig 3 shows the PC board.

The broadband input transformer has a 4:1 turns ratio to match the 50 Ω drive to the amplifier input. The end of the transmission line opposite the drive transformer is loaded with two 50 Ω 10 W non-inductive resistors connected in parallel. This reduces the reactance of the input circuit and presents a more constant load to the driving source. The input SWR is about 1.5:1 across the 80-meter band. Additionally, a small two-turn inductor tunes the gates to peak the drive signal. Although this inductor can be eliminated, the load resistance would need to be lower to match the drive to 50 Ω, with a corresponding increase in drive power requirements. The match would have been better in this configu-

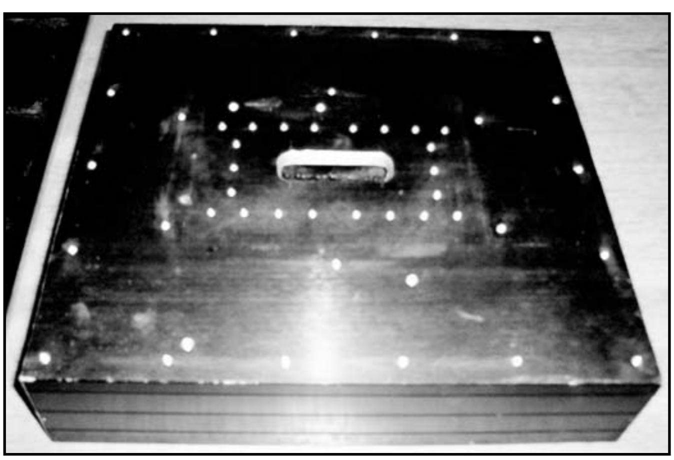


Fig 1—Heat sink with slot and tapped holes for mounting the PC board and components.

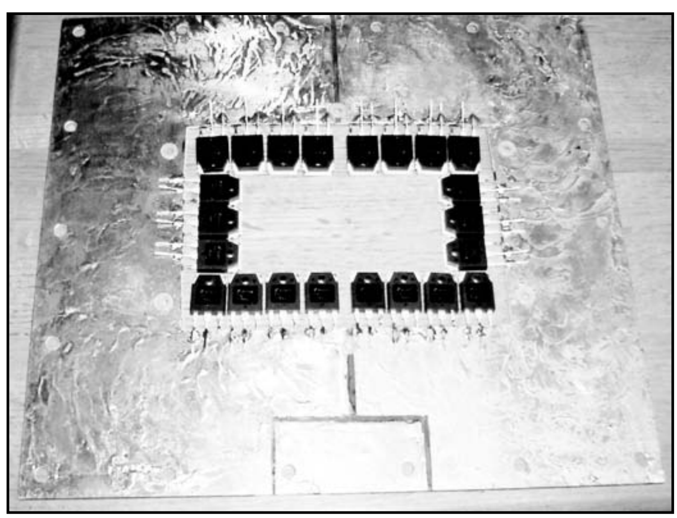


Fig 3—The PC board.

ration if there were a little more inductance between the drive transformer secondary and the PC board.

The gate drive needs to arrive at each FET simultaneously for the devices to share power properly and, furthermore, should have a relatively quick rise time to minimize the time the FET operates in its linear region. The delay between the first and last FET was measured at about 2 ns, which appears to be acceptable. The equal temperature rises of the FETs is an indication that they are evenly sharing the load.

The FET leads are soldered to the PC board with the gate leads on the top layer (transmission line) and the sources on the bottom layer (ground). The drain leads are not used; as mentioned earlier, the FET drain is connected to the heat spreader through its body. Both sides are tinned with 3% silver solder to minimize the resistance. The PC board is attached to the heat sink with #6-32 cap screws and nylon step washers (on the top) to insulate the gate traces. The board is spaced off the heat sink with 24 3/16-inch spacers connecting the source plane to the heat sink. The FETs are attached with 22 #4-40 cap screws. The assembly is aligned by installing four #4-40 threaded rods in the corner holes of the heat sink, so the Kapton is not damaged during assembly. My first insulator failed after a FET mounting screw dug a hole in the insulator while I was trying to find the hole in the heat sink. The threaded rods are replaced with cap screws after the assembly is aligned and the other cap screws are installed. A small amount of thermal grease was used on the heat sink, heat spreaders and FETs to fill any voids. Fig 4 shows the assembled amplifier.

Great care is required to protect the Kapton dielectric of the shunt capacitor during assembly. It is also very important to remove all burrs and dirt to prevent damaging the Kapton as the FETs are tightened down. Kapton has a 7,000 V/mil voltage rating, and the insulator in the 160-meter rig has never failed running at about 80 V peak dc input. The peak positive voltage has been measured at about 400 V during testing with the amplifier mistuned. High voltage is also present on the secondaries of the step-up output transformers. The output transformer's secondary lead is #10 (AWG) Teflon insulated wire, which provides sufficient, if not optimum, insulation. The final configuration will have a sleeve of Kapton on the inside of the transformer primary tubes to provide better voltage insulation.

A quick test was performed to determine if it was worth the time to package the amplifier into a chassis. The test verified that the design would work well and was worth the investment of material and time to complete the project. Fig 5 shows the test set up. Gratifyingly, the amplifier worked very well under these sub-optimum conditions. Several CW contacts and good reports provided the motivation to continue the project. No test is complete without a failure of some sort, and true to tradition, the series tuning capacitor failed during these tests, so a vacuum-variable cap was installed to complete the tests. The circulating current is quite high in the tank circuit, so the final design uses vacuum caps for both the tuning and the loading capacitors.

The next logical test step was to try a larger dc power source, because my test supply was severely overloaded. This amplifier was tested in the presence of several hams after the 2003 Marlboro, Massachusetts hamfest. Al, K1JCL, contributed a large dummy load and RF power meter and Tom, K1JJ, brought his peak-reading Bird 43 meter. Steve, WA1QIX, provided the test shack and modulator. We modified Steve's modulator by disabling his overload interlocks so we could see what this amplifier would dish out. My goal was to see how much power could be extracted from the amplifier with 50 V on the drains and the peak power under modulation.

We found the amplifier would make 1000 W at 50 V but required additional shunt caps for 80-meter operation. The class E drain waveform was cleaned up when four 1000 pF doorknob capacitors were added between the drain busses and ground, so as to increase the shunt capacitance of each phase by 2000 pF. This increased shunt-capacitor value allowed lower-impedance operation.

The result of this test proved that the amplifier would easily produce 1000 W of carrier. This was the threshold of stability for the tank component values because of the output-transformer ratio. The heat sink temperature was warm but not hot during these tests. We had no way of accurately measuring the dc input current, the heat sink didn't get warm enough to worry about.

This test indicated that a bit of work on the output network was necessary. The 1:4 turns ratio of the output transformer introduced too much leakage inductance and was later reduced to 1:2 which, in turn, requires a change in the output-tank component values. In this case, the inductor value decreases while the capacitor values must increase to maintain the proper class E wave-

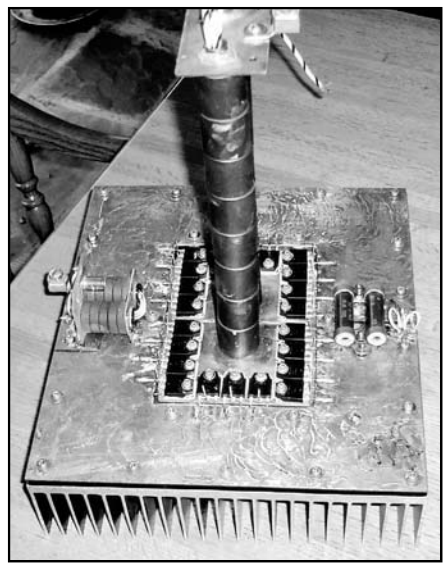


Fig 4—The assembled amplifier.

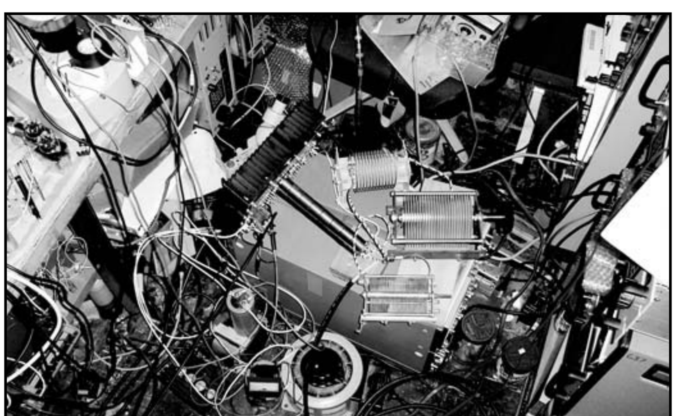


Fig 5—The amplifier test setup.

form. The new total shunt-capacitor value is about 8000 pF per phase, to allow a push-pull dc operating impedance under 2 Ω. The $R_{DS\,ON}$ of each bank of FETs is about 0.09 Ω; efficiency will decrease if the operating impedance is decreased further.

This amplifier was scaled to use larger output transformer cores than I had in my junk box. The next larger size of commercially-available core (1.5-inch OD, shorter, and having a lot more core area) would be a much better design choice. Using these larger cores, the height of the output transformer could be reduced by 50% with a corresponding reduction in leakage inductance with a better aspect ratio. However, the tall stack of cores never changed case temperature during testing.

An applications engineer suggested A material should be operated under 150 gauss at 4 MHz to keep losses low, which seems like a very conservative number. Type-43 material has similar performance with a little higher permeability. The A637 cores used were flea market loot. They are older switching-power-supply cores that work well up to 80 meters. Type-43 comes in the same size: 1.125×1.125 inches.

Fig 6 shows the schematic of the final amplifier. The parts count is quite low but the layout is very critical to insure power sharing of the FETs. Figs 7 and 8 show the amplifier module mounted into a chassis for additional testing.

The amplifier will produce a lot more power at higher voltages, but the goal was to operate at 50 V on the drains to provide a comfortable safety margin on the voltage rating of the FETs, shunt capacitor insulation and output transformer insulation. This amplifier can be scaled to operate on 40 meters without additional shunt capacitors, but half as many cores should be used in the output transformer. The amplifier can also be modified to operate on 160 meters with additional shunt capacitors and/or a larger heat spreader to increase the distributed capacitance value. This amplifier would also make an efficient CW transmitter final if a slope-controlled series switch were to be included in the dc power source. It cannot be used as a CW linear amplifier because key clicks will be generated as the driver turns on and off. AM is also possible with this amplifier using a series modulator.

The adjustment of a Class E amplifier is somewhat different than for other modes, and efficient operation depends largely on obtaining a clean waveform. An oscilloscope should be used to monitor the drain waveform and the gate drive until the user is comfortable operating a class E final. The gate drive should be at least 24 $V_{P\text{-}P}$, which corresponds to 50 to 60 W of RF input power on 80 meters.

This amplifier was operated without a load by mistake without any protection or failure. There is at least a 2:1 voltage safety factor in the drain waveform while running 1200 W output at 50 V input. After the initial tank capacitor failure, no other components have failed after extensive testing over a one-year period.

Conclusion

This amplifier will produce 1500 W output at 50 V on the drains, but the efficiency will be lower than expected. Sixty or 70 V would be a better choice. Steve, WA1QIX, set a rule of thumb at 50 W carrier per device as a good choice for high efficiency. My testing agrees with this choice. A few more FETs would allow 1500 W at 50 V. My next design will use a larger heat spreader to eliminate the doorknob caps added to this configuration.

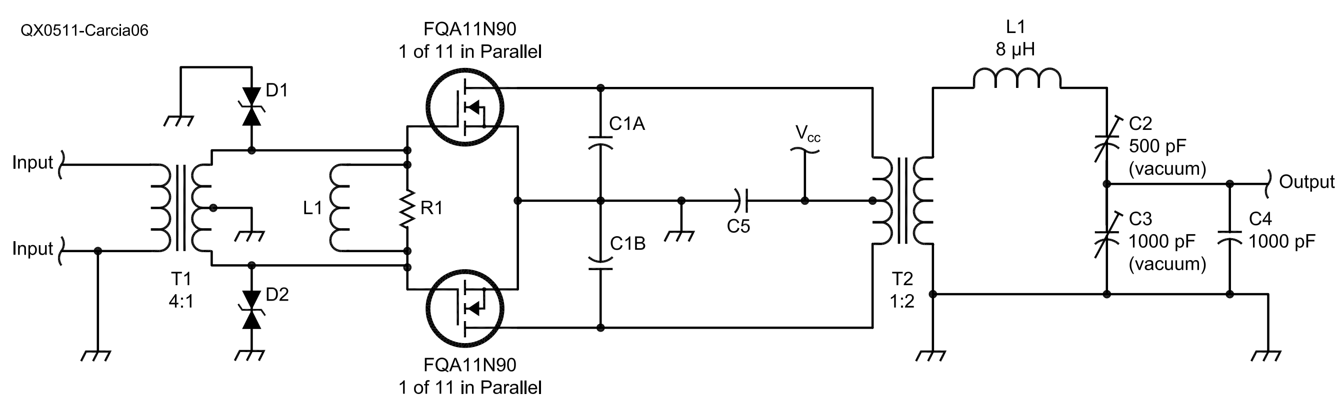


Fig 6—A partial schematic of 80-meter class-E amplifier (FQA11N90). One of 11 parallel connected circuits shown. Unless otherwise specified, use 1/4 W, 5%-tolerance carbon composition or film resistors.

C1A—3×4-inch copper plate 1/8-inch thick. FET drains and primary of output transformer common point. This plate is isolated from the heat sink with a 0.003-inch thick Kapton sheet.

C1B—Same as C1A. A capacitor is formed by the two heat-spreader plates and the heat sink. The value of each capacitor is about 3000 pF. This is in parallel with about 3000 pF of output capacitance from the 11 parallel connected FETs. This is the class-E shunt capacitance. Later 4000 pF was added to the shunt value by connecting two 1000 pF doorknob capacitors across C1A and C1B.

CR1, CR2—15-V, 1500 W Transorbs for FET gate-transient protection.

C4—Parallel combination of film and disc capacitors providing a low-impedance dc bypass.

L1—2 turns #14 AWG bare copper wire 1/2-inch ID, tuned to peak the drive signal on 80 meters.

L2, C2 and C3 are typical values for 80 meters. Double them for 160 meters and halve them for 40 meters. (C2 and C3 are vacuum-variable capacitors.)

R1—(2) 50 Ω, 10 W non-inductive resistors connected in parallel.

T1—Broad bandwidth input transformer with 4:1 impedance, with center tapped secondary. (Source: Communication Concepts, Inc; www.communication-concepts.com or an old solid-state amplifier transformer.)

T2—16 Chrometrics A637 cores (type-43 is a good substitute). Turns ratio is 1:2, 1:3 or 1:4. 1 turn 1/2-inch OD copper tubing primary and #10 AWG Teflon insulated wire secondary. This is sized to work on 160 meters with 120 V dc (peak) input voltage.

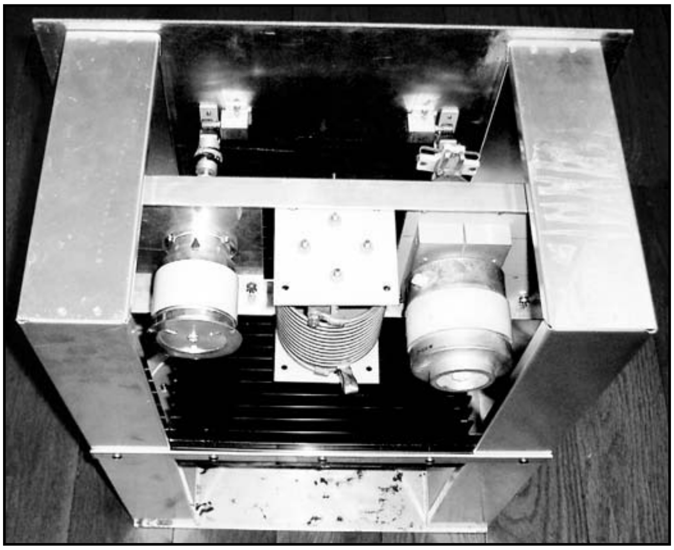
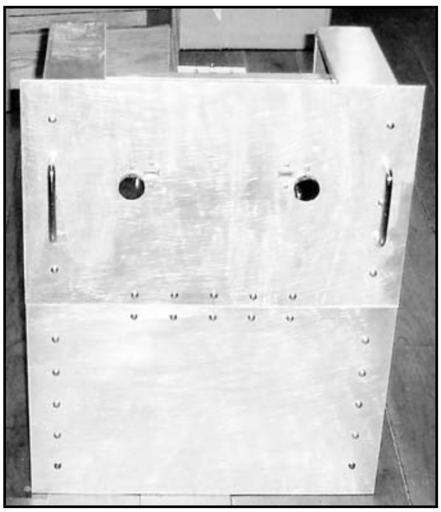


Fig 7—A view of the amplifier module mounted into a chassis for additional testing.



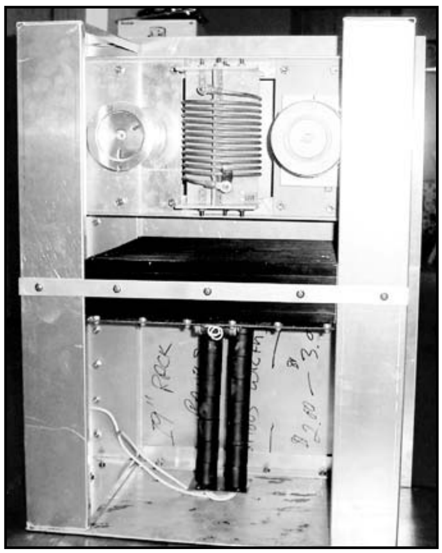


Fig 8—Views of the amplifier module mounted into a chassis for additional testing.

This amplifier is a work in progress, and the present configuration provides 1000 W of CW carrier at around 46 V, with an efficiency of 88%. The output inductor runs hot. An inductor of lower value wound with copper tubing would improve efficiency. An AM modulator providing 100 V peak voltage would easily produce 1500 W peak power with plenty of safety factor. The output power was pulsed to 5 kW peak output using a perfect 50 Ω load with 130 V on the drains as a final test. There was no sign of problems, but the output inductor was too small for this power level.

I recently found a six-meter linear using power FETs that used TO220 FETs. There was about a dozen configured in push-pull parallel with positive bias on the gates. This proves linear amplification is possible. A bank of RF FETs with over 5000 W of dissipation would not be possible for $100 in parts.

Nathan Sokal, WA1HQC, has provided several *QEX* articles covering the design of class-E power amplifiers (*QEX* Jan/Feb 2001, p 9; Mar/Apr, p 60; May/Jun, p 60).

Steve, WA1QIX, has a Web site covering several Class E projects others have built and he helps to simplify Nathan's design choices. This is a good reference for determining final tank component values. There is also good information on tuning class E amplifiers (**www.classeradio.com**). Steve has links to Nathan's information and design software.

The Class E FORUM (**http://classe.monkeypuppet.com)** is another good Web site for finding information and contacting amateurs building solid-state transmitters. It is normal to measure close to 90 % efficiency on a class E amplifier. This is very nice during the summer when the heat from a tube amplifier wants to chase you out of the shack after a short time.

Thanks to the Following

Edited by John, W3JN

Tests: Steve, WA1QIX; Al, K1JCL; Tom, K1JJ

Frank Carcia, WA1GFZ, has been a licensed amateur since 1966. He has been an avid homebrewer for 40 years building receivers, transmitters, linear amplifiers and beam antennas. He is mainly interested in high-performance receivers.

Frank worked in communications and control electronics for 30 years doing design and testing. □□

Low-Noise-Sideband Differential Amplifier VFO

This high-performance low-noise-sideband VFO operates in the widely used 5.5 MHz to 5.3 MHz frequency range, with a tuned circuit that can be easily modified for operation at other frequencies.

By Mal Crawford, K1MC

In the late 1970s W7ZOI and W1FB[1] described a high-performance low-noise-sideband variable frequency oscillator (VFO) designed by K7HFD that used a differential amplifier in the oscillator circuit. The description of the oscillator has also appeared in recent editions of the Handbook.[2] I've recently completed the design and construction of a heterodyne exciter to replace a classic vacuum-tube Heath HG-10B VFO that has been used with my homebuilt CW transmitter for over two decades. The excellent noise sideband performance of the differential amplifier VFO design made it a logical selection for the tuned oscillator of the new Heterodyne Exciter. The tuning range of my Differential VFO was selected to be 5.5 MHz to 5.3 MHz to allow transceive operation with the 5.5 MHz to 5.0 MHz tuned linear master oscillators (LMO) in my Heath SB-301 and SB-303 receivers. The 200 kHz tuning range of the VFO is wide enough to cover the CW sections of the HF bands. The tuned circuit component values can be easily changed for operating at other frequencies around 5 MHz, such as in the 60 m amateur allocation, or to reduce or expand the tuning range.

The Differential VFO is built on two different circuit cards to electrically and thermally isolate the sensitive VFO circuitry from the voltage-regulator circuitry. The VFO circuit card is mounted in a die-cast aluminum box that provides electrical shielding, mechanical rigidity, and thermal stability. The second circuit card contains the output low-pass filter, voltage regulator, and bias adjust circuitry.

[1]Notes appear on page 37.

19 Ellison Rd
Lexington, MA 02421
Malcolm_Crawford@Raytheon.com

Differential Amplifier Operation

The small-signal gain of a differential amplifier is determined by the collector resistance and the combined emitter current of the transistor pair. A true differential amplifier uses a current source in the emitter circuit to set the current in the transistor pair. The Differential VFO uses resistors in the emitter circuits of the transistors to approximate the characteristics of a current source. The emitter resistor approach simplifies the circuitry and allows operation with a single supply voltage. The emitter currents in the transistors are determined by the transistor base bias voltage and the emitter resistance. The small-signal gain of the differential amplifier is slightly over 14, with a combined emitter current of 6.2 mA and a collector load resistance of 230 Ω that results from the tank circuit coil to link turns ratio and a tuned circuit inductor Q of 200.

The first transistor in the differential pair is used as the oscillator with the second transistor acting as a limiting amplifier stage that produces a current-limited square wave signal at the collector. When the oscillator transistor Q1 is ON it draws all the emitter current and forces the limiting amplifier transistor Q2 into the OFF state by reverse biasing the emitter-base junction. On the other half of the waveform cycle Q1 is biased OFF by the transformed tuned circuit signal, which allows Q2 to turn ON and conduct all the emitter current. Low noise sideband operation performance is obtained by having the maximum Q1 ON collector current set by the sum of the base-bias voltage and the tapped tank circuit signal voltage and the emitter resistance. Controlling the maximum current prevents the transistor from going into saturation and lowering the load resistance seen by the tank circuit. The smaller load resistance associated with transistor saturation conditions lowers the loaded Q of the tank circuit and widens the noise sideband pedestal. The Q2 collector waveform is a sharp square wave whose amplitude is set by the current swing and the collector resistance value. The limiting action in Q2 provides another benefit of reducing the amplitude modulation (AM) noise sidebands well below the phase modulation (PM) noise sidebands. The Q2 collector resistor value was selected to allow a bifilar transformer to convert the output impedance to the 50 Ω value needed to terminate the input of the low-pass filter.

Oscillator Operation

The emitter follower configuration of the differential amplifier results in a high base input resistance. The high base resistance is then multiplied by the turns ratio of the tuned circuit inductor tap to minimize the resistive loading on the tuned circuit to prevent lowering the tuned circuit Q. The tuned circuit equivalent parallel resis-

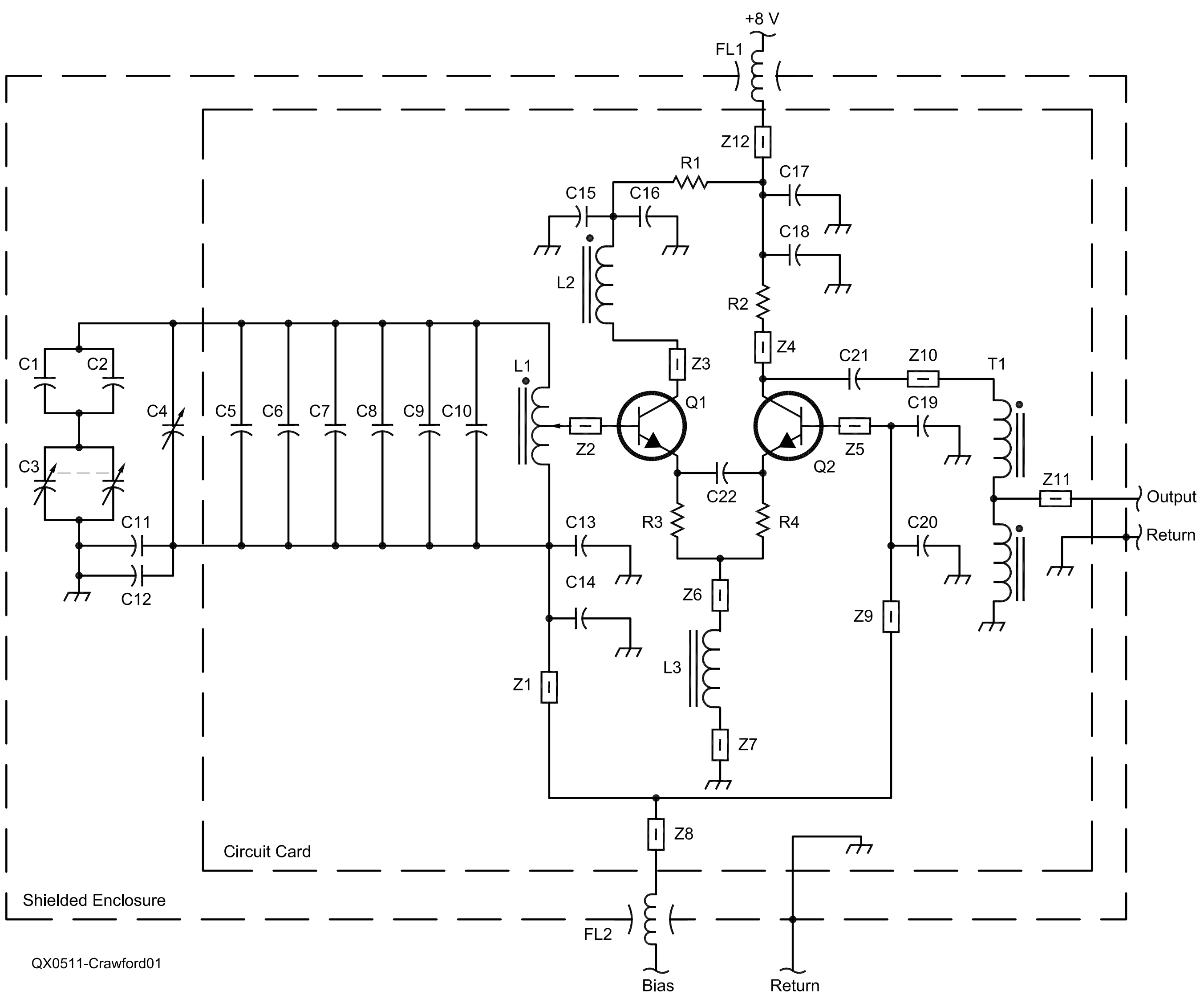


Fig 1—VFO schematic.

tance is then stepped down by the inductor to link winding turns ratio into collector circuit of Q1. The stepped-down tuned circuit parallel resistance becomes the collector load for the input stage of the differential amplifier. The gain around the oscillator loop from base to collector to tank circuit to base must be greater than unity for oscillation to start. The portion of the oscillator loop gain through the tuned circuit is set at one half by the ratio between the number of turns in the inductor tap (1 turn) and the link winding (2 turns).

The net oscillator loop gain of 7 is determined by multiplying the differential amplifier gain of 14 by the 0.5 tap-to-link turns ratio of the tank coil. The output amplitude of the oscillator keeps increasing once dc power is applied until it reaches a steady state level. At this point the amount of energy fed back into the tuned circuit only needs to be large enough to make up for the losses in the tuned circuit.

Differential Oscillator Circuit Details

The main changes in the Differential VFO from the original K7HFD design are in the tuned circuit and the dc biasing. A schematic of the Differential VFO circuit is shown in Fig 1. The original 10 MHz operating frequency of the K7HFD design was cut in half by doubling the tuned circuit inductance and capacitance. NP0 (C0G) zero temperature coefficient ceramic capacitors are used in the tuned circuit to reduce frequency variation due to ambient temperature changes. Multiple capacitors in parallel divide the tuned circuit circulating currents, reduce heating effects, and average out the tuned circuit capacitance value tolerances and temperature coefficients.

A toroidal coil was selected over an air-core solenoid inductor because of its self-shielding properties and smaller size. Two cores were stacked to reduce the core magnetic flux density level to less than a third of the 25 gauss 100% permeability reference level. The large core volume also eliminates any heating effects from the 3 mW of power dissipated in the coil. The tuned circuit toroidal inductor uses -7 (white) or -6 (yellow) iron powder core material because these two mixes have the lowest temperature coefficients. I used one -7 core and one -6 core in my circuit because I had them on hand. Two -7 cores or two -6 cores can be used for the tuned circuit inductor by changing the value of one of the capacitors in the tuned circuit to compensate for the slightly different A_L values of the iron powder cores. Change the value of C5 from 30 pF to 42 pF if two -6 cores are used and change the value of C5 from 30 pF to 18 pF if two -7 cores are used. Recentering the fixed capacitor value allows the full range of the trimmer capacitor to compensate for capacitor and inductor value tolerances.

The values of the fixed capacitors in the tuned circuit of the Differential VFO can be changed to allow the use of available tuning capacitors. The total tuned circuit capacitance must increase from 349 pF at 5.5 MHz to 375 pF at 5.3 MHz with the 2.4 µH inductor. To use another tuning capacitor, compute the values of C1 and C2 needed for the composite capacitance of C1, C2, and C3 to change by 26 pF when going from minimum tuning capacitance to maximum tuning capacitance. The values of the fixed tuned circuit capacitors C5 through C10 are then changed to set the total tuned circuit capacitance at 349 pF with C3 at its minimum capacitance. A smaller change in the composite capacitance of C1, C2, and C3 will give you a smaller tuning frequency range if desired. A capacitance increase of only 20 pF is needed for a smaller 150 kHz tuning range. The Differential VFO design can be used at other frequencies by first scaling the tuned circuit inductance and total capacitance from the 5.5 MHz values. Use the scaled inductance and capacitance values to compute the values of C1 and C2 for the desired tuning range as described above. Then compute the values of the fixed tuned-circuit capacitors for the desired operating frequency.

The trimmer capacitor can be eliminated to reduce the size of the oscillator by adding a 30 pF NP0 capacitor to the tuned circuit to maintain the operating frequency range. The operating frequency range can be set to within 12 kHz of the desired frequency by changing the value of one or more of the fixed value NP0 capacitors in the tuned circuit. The nominal frequency versus capacitance scale factor for the tank circuit is –8 kHz per picofarad for an L1 inductance of 2.4 µH. The 2% tolerance NP0 capacitor value step sizes are approximately 3 pF for values below 50 pF, which will produce steps of 24 kHz in the operating frequency.

A base bias voltage +1.4 V dc for the transistors in the circuit is used to set the quiescent collector currents at 3.1 mA each. The transistor collector currents were reduced in the Differential VFO design compared to the 15 mA per transistor in the K7HFD design to lower the output amplitude and limit power dissipation heating effects in the oscillator circuit components. The output amplitude of the Differential VFO is –6 dBm compared to the +17 dBm in the original K7HFD design. The output power level of the Differential VFO can be increased by changing the base bias voltage supplied to the transistors. Increasing the base bias voltage from +1.4 V dc to +2.1 V dc will increase the collector current of each transistor to 5.8 mA and quadruple the output power to 0 dBm. Higher output power levels can also be obtained by lowering the emitter resistor values to increase the collector current while keeping the base bias voltage at +1.4 V dc. The supply voltage should be increased to +10 V dc at the high output power levels used in the K7HFD design to avoid transistor collector saturation conditions with the larger signal voltage swings. Output power can also be increased by another 3 dB by eliminating the two way power divider (U2) at the output of the low-pass filter.

Separate emitter resistors were used for each transistor to balance the quiescent collector currents and eliminate the need for matched transistor pairs. Having Q1 in the ON state is necessary at power turn-on so that it can provide the small signal linear gain needed for the oscillator to start. The measured emitter voltages of the two unmatched transistors were within a millivolt when the tuned circuit was shorted to ground to prevent the circuit from oscillating. A single emitter resistor can be used as is done in the K7HFD design if a balanced dc state for test and troubleshooting purposes isn't desired and the transistors are selected so that the oscillator starts reliably. The collector circuit of Q1 contains a small value resistor to improve the decoupling of the +8 V dc regulated power source. The collector circuit of Q2 contains a 200 Ω resistive load that is transformed by T1 to 50 Ω to provide the correct source impedance for the low-pass filter.

Two 2N2222A metal-case TO-72 transistors were used in the Differential VFO. Similar plastic-case NPN transistor types such as the PN2222 and the 2N3904 used in the original K7HFD design are also suitable for operating frequencies up through 10 MHz. For high power operation the larger TO-5 case 2N2219A is recommended. Transistors such as the 2N2369A with an f_T greater than 500 MHz should be used at higher operating frequencies up to approximately 20 MHz.

Parallel ceramic capacitors with values two decades apart were used for bypassing. The two capacitor approach provides better bypassing over a wider range of frequencies than is possible with a single capacitor because of component self resonances. Ferrite beads were used liberally to attenuate VHF parasitics resulting from lead inductance and stray capacitance resonances in the circuitry. Without the beads the very sharp edges of the square wave in the output half of the differential amplifier produced a small VHF signal burst on each cycle, due to parasitic resonances.

Low-Pass Filter Circuit Details

A five-element 0.1-dB-ripple Chebyshev low-pass filter with a cut off frequency of 5.5 MHz is used on the output of the oscillator to reduce the harmonics of the sharp square waveform at the collector of Q2. A schematic of the low-pass filter is shown in Fig 2. The low-pass filter attenuates the square wave harmonics that extend out through the VHF range to produce a sinusoidal output signal. The low-pass filter is recommended even when the Differential VFO is used to drive the local oscillator port of a balanced mixer. The filtering insures that the mixer LO drive will have a 50% duty cycle to maintain the balance of the mixer switching.

The low-pass filter component values were determined by scaling the values for the 15.087% reflection coefficient filter in the *ARRL Handbook* Chebyshev Filter Design Normalized Tables. In the 2005 edition of the *Handbook*, this is Table 12.2 found on page 12.11. Two standard-value mica or NP0 capacitors are connected in parallel to match the input and output capacitor design values of 664 pF and the middle capacitor design value of 1144 pF. The two high-Q inductors in the low-pass filter use the same T50-6 iron powder toroidal cores as the tuned circuit inductor. T50-2 (red) iron powder cores can be used for L4 and L5 by reducing the number of turns from 22 to 20 to account for the higher A_L. The low-pass filter component values can be scaled if the Differential VFO operating frequency is changed from 5.5 MHz.

Voltage Regulator Circuitry

An adjustable LM317 linear voltage regulator is used to provide a steady voltage to the Differential VFO and filter out any noise coming in on the supply line. A schematic of the voltage regulator circuitry is shown in Fig 2. A nominal operating voltage of +8.0 V dc is used to keep oscillator heating to a minimum and allow the regulator to be used with input voltages between +11 V dc and +35 V dc. Diodes are included in the voltage regulator circuit for input short circuit protection and input reverse volt-

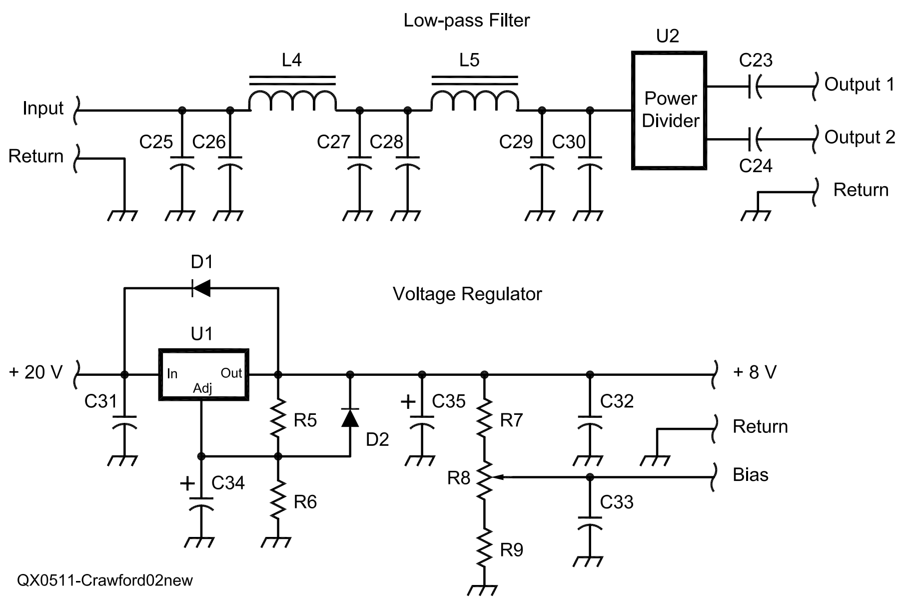


Fig 2—Low-pass filter and voltage regulator schematic.

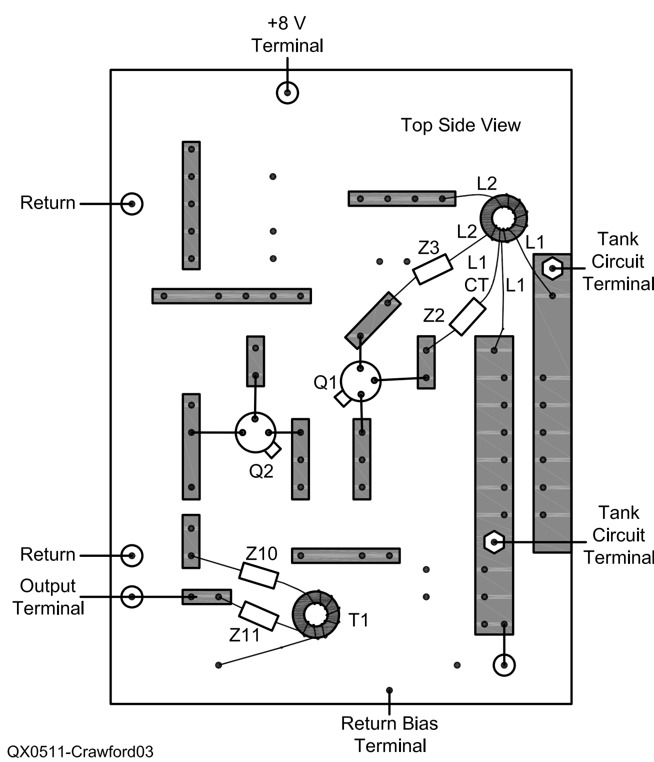


Fig 3—VFO circuit card layout, as seen from component side.

age protection. Large-value polarized capacitors are used on the adjustment and output terminals of the voltage regulator to reduce output noise levels and improve transient response. Smaller-value ceramic capacitors on the input and output of the regulator improve stability in case the power supply is not located close to the regulator. The circuit card copper cladding provides adequate heat sinking for the TO-220 case of the voltage regulator IC with the low oscillator and bias circuit load currents.

A resistor divider network with a multiple-turn potentiometer is used to set the oscillator base bias voltage at around +1.4 V dc. A multiturn potentiometer is used to provide a smooth and stable adjustment for the base bias voltage. The base bias voltage can be set with a voltmeter or adjusted to set the output signal amplitude at the desired level. Low-tolerance metal-film resistors should be used in the regulator voltage adjustment divider and the base bias voltage divider networks to provide stable voltages to the oscillator over the ambient operating temperature range and eliminate the need for an adjustment to set the regulator output voltage. The measured regulator output voltage was +8.05 V dc with the fixed resistor values shown in the schematic. To increase the Differential VFO output from –6 dBm to 0 dBm, change R9 from 412 Ω to 1.00 kΩ in order to recenter the adjustment range around +2.1 V dc.

Building the Differential VFO

VFO Circuit

The hardest VFO component to find is the main tuning capacitor. The one I selected for C3 is a Millen double-bearing type used by Heath in their HG-10 VFO. Hamfest flea markets are a good source for high-quality air-variable tuning capacitors. Two commercial sources of tuning capacitors are Ocean State Electronics and Surplus Sales of Nebraska. The value of the tuning capacitor is not critical since the values of C1, C2, and the tuned-circuit capacitors can be changed to provide the desired tuning frequency range. Capacitors C11 and C12 can be replaced by a direct connection to the tank circuit if both the rotor and stator of the tuning capacitor are not grounded. A gear reduction drive should be used with the tuning capacitor to obtain a comfortable and smooth tuning rate. The Heterodyne Exciter uses an Eddystone 898 slow motion dial with a 100:1 reduction ratio for tuning the Differential VFO.

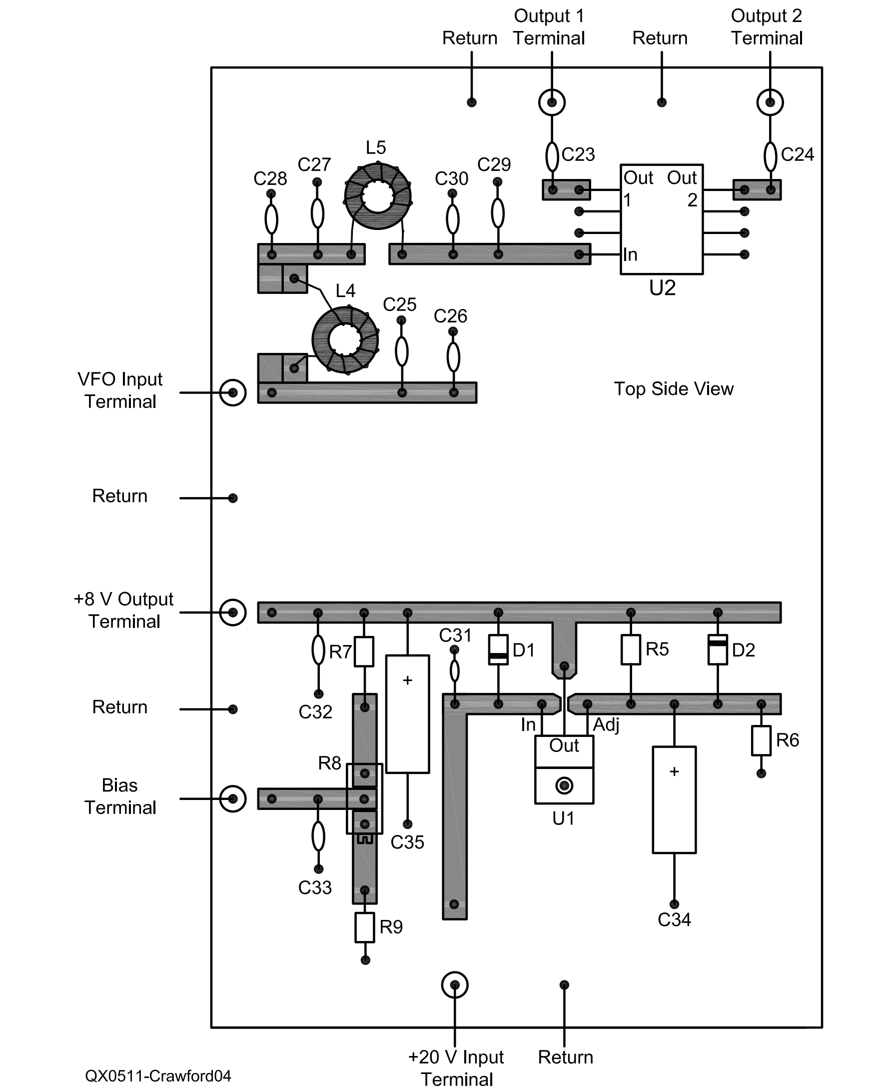


Fig 4—Low-pass filter and voltage regulator circuit card layout, as seen from component side.

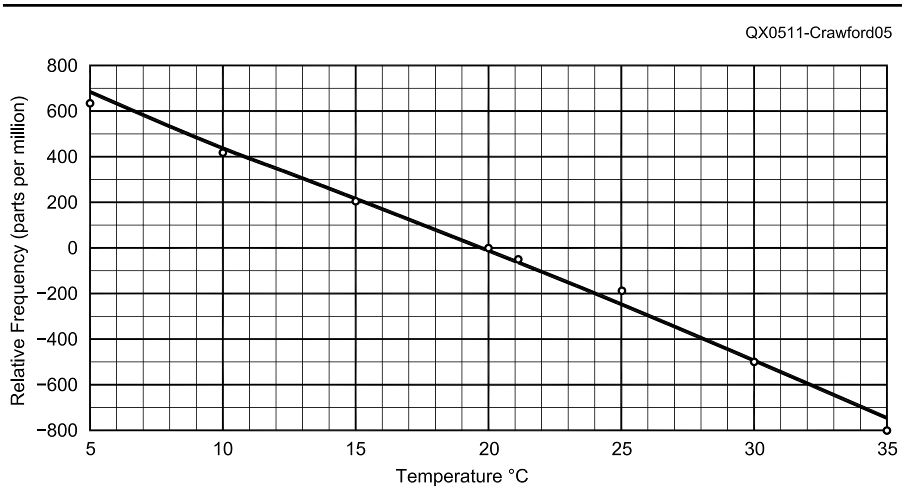


Fig 5—VFO temperature stability.

The trimmer capacitor used for C4 was a Hammarlund APC-51 air variable. A high-quality glass or ceramic piston trimmer capacitor mounted on the circuit board would be a suitable substitute for fine tuning the oscillator frequency. Low-cost ceramic trimmer capacitors should not be used in the tuned circuit because of their very high temperature coefficients. A mid-range value of 30 pF for the APC-51 trimmer capacitor was used in computing the value of the fixed capacitors in the tuned circuit. The values of the fixed capacitors in the tuned circuit should be adjusted if a substitute trimmer capacitor with a different mid-range capacitance is used.

The oscillator circuit is built on a 2.4×3.3 inch unclad glass epoxy board. A layout of the oscillator circuit-card components and the copper-tape wiring runs is shown in Fig 3. The layout uses the standard printed wiring board drawing convention of showing the reverse side traces projected through to the top side. There is one jumper wire to carry the bias voltage from the circuit-card terminal trace to the Q2 base circuitry. Unclad board is used to eliminate parasitic capacitors formed between circuit traces on one side of the board and ground plane on the other side. The oscillator circuit board and the two air-variable capacitors are mounted in a Hammond 1590C (3.70×4.70×2.06 inch) die-cast aluminum box that provides good shielding, mechanical rigidity, and thermal transient isolation. Alternate parts that are slightly different in size are the Bud AN-1304 (3.54×4.53×2.17 inch) and LMB Heeger (3.69×4.69×2.06 inch) die-cast aluminum boxes. The two iron-powder toroidal cores used in the tuned circuit inductor were bonded together using Duco® Cement. The coil is supported above the circuit card by its five leads. The oscillator and base bias voltages are brought into the enclosure through filter feedthroughs and connected to the circuit card by twisted wire pairs. Insulated feedthrough terminals can be used for the oscillator and base bias-voltage inputs if filter feedthroughs are not available. The decoupling provided by the ferrite beads and capacitors on the oscillator circuit card should be adequate to keep any external signals out of the oscillator. A Teflon® insulated feedthrough terminal is used for the 5.5 MHz output signal of the oscillator. Ground studs and ground lugs are used for the oscillator voltage, base bias voltage, and output signal return connections. The VFO picture in Fig 8 was taken with the enclosure cover off to show the circuit card construction details and the two variable capacitors.

Low-Pass Filter and Voltage Regulator Circuit

These two circuits were built on a 4×5 inch double-sided, copper-clad glass epoxy board that is mounted next to the oscillator assembly. A layout of the low-pass filter and voltage regulator circuit components and wiring runs is shown in Fig 4. A 5 inch length of 3/4 inch aluminum angle stock is used as a mounting bracket for the board. A short piece of miniature 50 Ω coaxial cable is used to connect the 5.5 MHz signal output of the VFO to the input of the low-pass filter. The mounting tab on the LM317 adjustable voltage regulator is connected to the output of the regulator, and must be electrically insulated from ground. A round mica washer with thermal grease was used between the regulator tab and the circuit card to provide the electrical isolation. A nylon screw and a nylon nut were used to fasten the tab and mica washer snugly to the circuit card to aid heat conduction. The oscillator and base bias voltages are carried to the VFO by twisted pairs of #24 AWG wire. The component side of the circuit card is shown in Fig 8.

A mechanical approach was used to fabricate the low-pass filter/voltage regulator circuit card because of the small number of traces that they each contained. Photocopies of the layouts were made and used as guides for

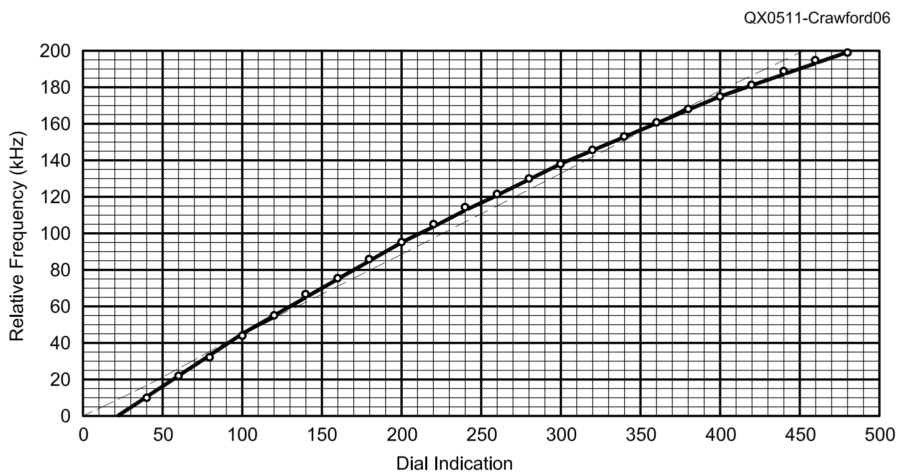


Fig 6—VFO tuning linearity.

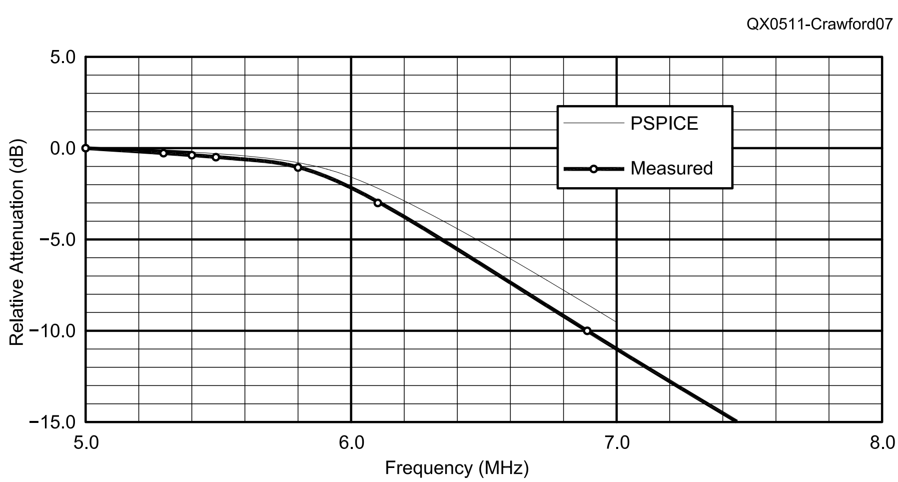


Fig 7—Low-pass filter frequency response.

routing out the unwanted copper by gluing them to the boards with rubber cement. A small hobby drill with an engraving cutter bit was used to cut away the copper on the boards. Copper tape was put over the edges of the circuit card to connect the ground planes on the two sides. The copper tape was first burnished with a wooden dowel and then soldered to the ground planes on both sides. A 3/8 inch diameter drill was used to clear the ground plane for component leads. The wiring traces on the circuit card and the copper tape are shown in Fig 9. Be sure to wear eye protection when routing the copper cladding and drilling the component lead holes if you use this method to fabricate the circuit boards.

Setup

Setting up the Differential VFO only takes two steps. The first is to set the tuning range of the oscillator with the trimmer capacitor using a frequency counter or general coverage receiver with an accurate display. Adjust trimmer capacitor C4 until the VFO output frequency is 5.500 MHz with the plates of C3 barely meshed. The second step is to set the base bias voltage at +1.4 V dc by adjusting R8.

My Oscillator Only Amplifies

There are three likely causes if the Differential VFO fails to oscillate. The first is that the sense of the feedback coil (L2) winding is reversed. The feedback coil is wound in the same direction as main coil (L1) with the starting end connected to the collector of Q1 and the finish end connected to the 36 Ω current limiting resistor. The sense of the two coil windings is indicated by the dots on the schematic drawing. The second cause of non-oscillation is low tuned-circuit Q. The operation of the oscillator depends on an inductor Q of approximately 200 to set the oscillator small-signal loop gain at 7. The equivalent parallel resistance of the tank inductor is coupled into the collector circuit by the turns ratio of L1 and L2. If the collector resistance is too small, the gain of the differential amplifier may be lowered to the point where the loop gain drops below unity and oscillation can't be sustained. Low resonator Q can result from shorts between turns on the tuned circuit inductor. The grounding of the tuning and trimmer capacitor frames can inadvertently short the tuned circuit and prevent oscillation. Both sides of trimmer capacitor C4 must float to avoid shorting the base bias voltage. Check the wiring of C3 to make sure that the C1-C2 side isn't grounded by the frame and mounting hardware.

Circuit Test Results

Several sets of measurements were made on the completed Differential VFO once the tuning frequency range was set. The first test was a power on frequency change measurement using a laboratory grade counter to determine the effects of self heating as the Differential VFO warmed up at a room ambient temperature. The relative frequency change was only –10 Hz after a 30 minute warm-up with a peak frequency error of –125 Hz as shown in the following table.

Time	*Relative Frequency*
Power on	0 Hz
1 minute	–31 Hz
5 minutes	–125 Hz
10 minutes	–109 Hz
30 minutes	–10 Hz

The frequency counter was also used to determine tuning linearity by measuring the output frequency as a function of dial rotation. The plotted results in Fig 6 show the slight bowing expected with a conventional tuning capacitor having a capacitance directly proportional to shaft rotation. The plot also shows the best straight-line fit of 0.44 kHz/dial division to the tuning characteristic to show the de-

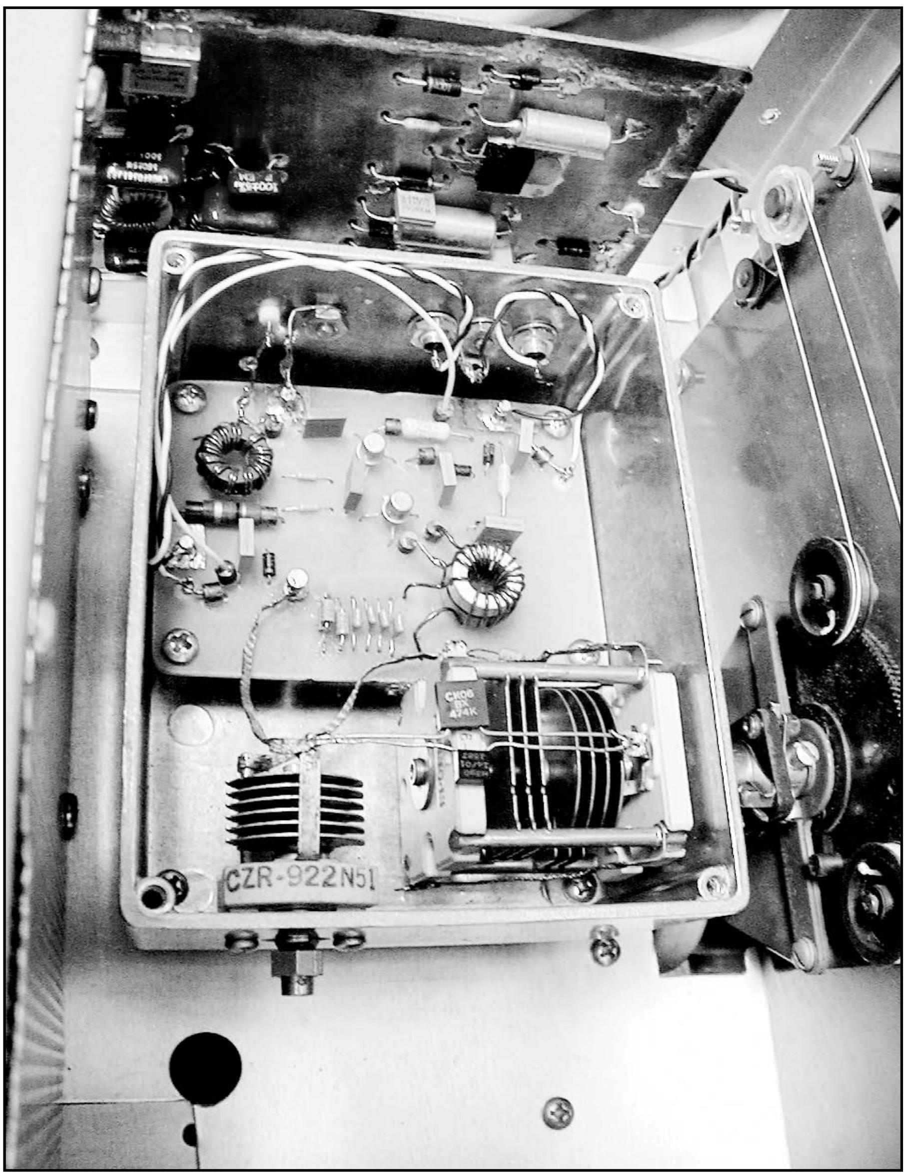


Fig 8—VFO and component side of LPF-VR circuit card.

viation from straight line tuning.

Another set of test measurements used a spectrum analyzer to determine the harmonic levels at the output of the five-section low-pass filter. All harmonics were 60 dB or more below the 5.5 MHz fundamental output signal as shown in the following table.

Harmonic	*Relative Amplitude*
Second	–61 dBc
Third	–65 dBc
Fourth	Below analyzer noise floor
Fifth	–83 dBc
Sixth	Below analyzer noise floor
Seventh	–85 dBc

The frequency response of the low-pass filter was also measured and is plotted in Fig 7 along with the frequency response predicted by a PSPICE circuit simulation. The measured frequency response shows the slightly lower corner frequency resulting from parasitic capacitances on the circuit board and component lead inductance. The measured frequency response closely matched the predicted attenuation out to 21 MHz where leakage paths around the filter start limiting the filter attenuation. The leakage paths resulted in a measured attenuation of –65 dB at 28 MHz compared to –73 dB predicted by the simulation.

Frequency stability over temperature was the last set of tests run on the Differential VFO. The oscillator, voltage regulator, low-pass filter, and the separate Heterodyne Exciter ac power supply were placed in a temperature chamber and tested in 5°C (9°F) steps from +5°C (+41.2°F) to +35°C (+95.2°F). The frequency temperature coefficient over this range was determined by plotting the measured output frequency versus temperature data and fitting a straight line. The slope of the straight-line fit to the data was –46.7 ppm/°C over this 30°C temperature range, or –252 Hz/°C. The –46.7 ppm/°C temperature coefficient of the oscillator frequency is slightly larger than the +32 ppm/°C temperature coefficient of the iron-powder toroidal core and includes the contributions of the air-variable trimmer and the 0 ± 30 ppm/°C temperature coefficients of the NP0 capacitors. (The signs of the frequency and inductor temperature coefficients are reversed because the resonant frequency of the tank circuit goes down as the inductance and capacitance go up.) The plot in Fig 5 shows the frequency change with respect to the frequency at 20°C at the seven measurement-temperature data points and the best straight-line fit of the data points.

The VFO tank circuit voltage couldn't be measured accurately with an oscilloscope because it loaded the high Q tank even when a high impedance probe was used. The signal voltage measured at the base of Q1 was 1.2 V_{P-P}, which would scale up by the turns ratio of the tank circuit inductor to 20 V_{P-P} across the tank circuit. This is lower than the 50 V_{P-P} resonant circuit voltage in the original K7HFD design, but is large enough to provide low FM noise-sideband levels. The large voltage swing of the tank circuit

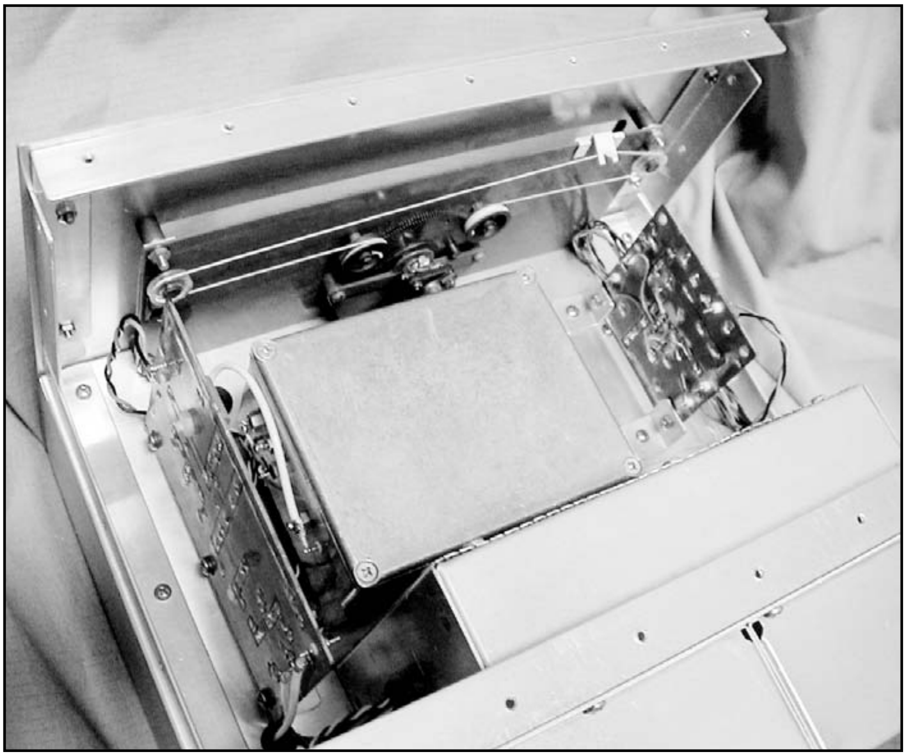


Fig 9—VFO with cover on and wiring side of LPF-VR circuit card.

Parts List

C1, C6, C7, C8, C9, C10—56 pF NP0 ceramic, 2%, 50 V.
C2—62 pF NP0 ceramic, 2%, 50 V.
C3—air variable tuning capacitor 10 pF - 50 pF.
C4—Air variable trimmer capacitor 5 pF - 50 pF.
C5—30 pF NP0 ceramic, 2%, 50 V.
C11—0.47 µF ceramic, 50 V.
C12—0.047 µF ceramic, 50 V.
C13, C15, C17, C19, C21, C22, C23, C24—0.1 µF ceramic, 20 V.
C14, C16, C18, C20—1000 pF ceramic, 20 V.
C25, C29—560 pF mica or NP0, 5%, 20 V.
C26, C30—100 pFmica or NP0, 5%, 20 V.
C27—1000 pF mica or NP0, 5%, 20 V.
C28—150 pF mica or NP0, 5%, 20 V.
C31, C32, C33—1 µF ceramic 50 V.
C34—10 µF electrolytic, 20 V.
C35—47 µF electrolytic, 20 V.
D1, D2—50 V, 1 A diode (1N4001).
FL1, FL2—Filter feedthrough.
L1—2.4 µH 17 turns #24 enameled wire on T50-7 and T50-6 stacked iron-powder toroid cores, tap one turn up from the ground end.
L2—2 turns #24 AWG wound over L1.
L3—47 µH RF choke (J.W. Miller 9250-473).
L4, L5—1.98 µH, 22 turns #24 enameled on T50-6 iron-powder toroid core.
Q1, Q2—Silicon NPN transistor fT > 200 MHz (2N2222A).
R1—36 Ω, 5%, 0.1 W.
R2—200 Ω, 5%, 0.1 W.
R3, R4, R5—249 Ω metal film, 1%, 0.1 W.
R6—1.27 kΩ metal film, 1%, 0.1 W.
R7—3.74 kΩ metal film, 1%, 0.1 W.
R8—1.0 kΩ, 24 turns, 3/8 inch leaded trimmer potentiometer (Bourns 3296W-1-102, Spectrol 64W, Murata 3102W).
R9—412 Ω metal film, 1%, 0.1 W.
T1—12 turns #24 AWG enameled wire, bifilar wound on FT-5061 ferrite toroid core.
U1—1 A positive adjustable voltage regulator (LM317) TO-220 case.
U2—Two-way power divider (Merrimac PDF-2A-100).
Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8, Z9, Z10, Z11—FB-1 mix 43 ferrite bead. (0.14 in OD, 0.05 in ID, 0.12 in long).
Iron powder toroidal cores, ferrite toroidal cores, and ferrite beads were obtained from Palomar Engineers.
Other components are available from Future/Active, Digi-Key and Mouser.
Tuning and trimmer capacitors are available from Ocean State Electronics and Surplus Sales of Nebraska.

demonstrates why shielding and extensive bypassing and decoupling were used in the Differential VFO. The waveform at the emitters of the two transistors was a partial sinusoid indicating Q1 was operating in Class C. The output waveform at the collector of Q2 was slightly trapezoidal rather than square because of the filtering action of the low-pass filter input capacitors (C25 and C26).

Measurements of the FM noise-sidebands of the Differential VFO have not been made to date. A version of the K7HFD design was designed, built, and tested by former co-worker Mark Lanciault.[3] Mark's version of the oscillator used 2N5109 transistors, operated near 30 MHz, and was tuned by varactor diodes rather than an air-variable capacitor. PM noise-sidebands at frequency offsets greater than 200 kHz from the carrier were measured at –156 dBc/Hz using two identical oscillators and a Hewlett-Packard 3048A phase-noise measurement system. This matches the –155 dBc/Hz PM noise sidebands measured at 10 kHz from the carrier for the K7HFD design. The noise pedestal in Mark's design was wider due to the lower Q of the tuned circuit caused by the higher operating frequency and the varactor tuning diodes.

The PM noise sideband levels for the Differential VFO are expected to be –145 dBc/Hz beyond 15 kHz from the carrier based on the two sets of oscillator measurements. The measured PM noise-sideband levels of Mark's design were scaled up by +10 dB to account for the lower collector current and resulting smaller output signal amplitude in the Differential VFO design. Mark's oscillator was biased for 10 mA in each transistor compared to the 3.1 mA in the Differential VFO. The width of the noise sideband pedestal of the Differential VFO is based on an estimated tuned-circuit inductor Q of 200. Running the Differential VFO with higher current levels wasn't considered because it would not result in any improvement in the PM noise-sidebands of the Heterodyne Exciter that uses the Differential VFO. The Differential VFO output is mixed with a direct digital synthesizer (DDS) signal with spurious signal levels between –80 dBc and –90 dBc in the region 100 kHz on either side of the carrier. The estimated PM noise-sidebands of the Differential VFO would be nearly 30 dB lower in a 500 Hz bandwidth than the spurious outputs of the DDS.

Notes

[1]W. Hayward, W7ZOI, and D. DeMaw, W1FB (SK), *Solid State Design for the Radio Amateur*, p. 126, Figure 26, K7HFD low noise oscillator.

[2]*The ARRL Handbook for Radio Amateurs*, 2006, p. 10.18, The K7HFD Low-Noise Oscillator.

[3]M. Lanciault, *Doppler VCO Status Report for Low Frequency Section of the On Board Reference IDP*, April 1 1991 (Unpublished internal company technical report).

Mal Crawford, K1MC, was first licensed as a Novice in 1959 with the call sign WV2IPC. Since completing his active military service in the U.S. Army Signal Corps in 1972, he has lived in New England and worked in the field of missile and radar electronics. Mal enjoys hiking and volunteer trail maintenance and construction activities when not designing, building and operating homebuilt equipment. His recently completed HF transmitter encompasses a half century of electronic technology, ranging from a neutralized class C vacuum-tube amplifier to an integrated-circuit direct digital synthesizer. □□

United States Postal Service

Statement of Ownership, Management, and Circulation

1. Publication Title	2. Publication Number	3. Filing Date
QEX	0 8 8 6 - 8 0 9 3	September 27, 2005
4. Issue Frequency	5. Number of Issues Published Annually	6. Annual Subscription Price
Bimonthly in: Jan, Mar, May, Jul, Sep, Nov	6	$24.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*
225 Main Street, Newington, Hartford County, CT 06111-1494
Contact Person: Katherine Capodicasa
Telephone: (860) 594-0257

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*
225 Main Street, Newington, CT 06111-1494

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*
Publisher *(Name and complete mailing address)*
Harold Kramer, 225 Main Street, Newington, CT 06111-1494
Editor *(Name and complete mailing address)*
Doug Smith, 225 Main Street, Newington, CT 06111-1494
Managing Editor *(Name and complete mailing address)*
Robert Schetgen, 225 Main Street, Newington, CT 06111-1494

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
American Radio Relay League, Inc.	225 Main Street, Newington, CT 06111-1494

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → XX None

Full Name	Complete Mailing Address

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:
XX Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

PS Form **3526**, October 1999 *(See Instructions on Reverse)*

13. Publication Title: QEX		14. Issue Date for Circulation Data Below	
		Sep/Oct 04 - Jul/Aug 05	Sep/Oct 05
15. Extent and Nature of Circulation		Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*		6,983	6,900
b. Paid and/or Requested Circulation	(1) Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	5,511	5,348
	(2) Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)*	0	0
	(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	257	259
	(4) Other Classes Mailed Through the USPS	721	524
c. Total Paid and/or Requested Circulation *[Sum of 15b. (1), (2),(3),and (4)]*		6,489	6,131
d. Free Distribution by Mail *(Samples, compliment ary, and other free)*	(1) Outside-County as Stated on Form 3541	122	122
	(2) In-County as Stated on Form 3541	0	0
	(3) Other Classes Mailed Through the USPS	36	9
e. Free Distribution Outside the Mail *(Carriers or other means)*		79	58
f. Total Free Distribution *(Sum of 15d. and 15e.)*		237	189
g. Total Distribution *(Sum of 15c. and 15f)*		6,726	6,320
h. Copies not Distributed		257	580
i. Total *(Sum of 15g. and h.)*		6,983	6,900
j. Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*		96.48%	97.01%

16. Publication of Statement of Ownership
XX Publication required. Will be printed in the Nov/Dec 05 issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner
[signature] Publisher
Date: September 27, 2005

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.

PS Form **3526**, October 1999 *(Reverse)*

Reduced Bandwidth Voice Transmission

A modest proposal

By Sam Cowan, WØOAJ

Hams who received their first licenses in the late '50s were eyewitnesses to the struggle to implement single sideband for use on the ham bands. The "sidewinders" were relegated to the few kilocycles at the bottom and top ends of the voice sub-bands, and there were some spectrum-usage wars.

Eventually, SSB won because of the technical advantages of the mode. Chief among these is reduced bandwidth. SSB uses approximately one half as much spectrum as an amplitude-modulated transmission. This allows more stations to use the same allocated spectrum. If proper filtering is used, it allows for weaker signals to be copied through the noise and interference.

For voice to be intelligible, audio frequencies from 300 Hz to 3000 Hz should be transmitted. The bandwidth required for an SSB signal is then 2.7 kHz. If the bandwidth is reduced to less than 2.7 kHz, the audio begins to become unintelligible. This article

Box 266
Stromsburg, NE 68666-4420
sc92228@alltel.net

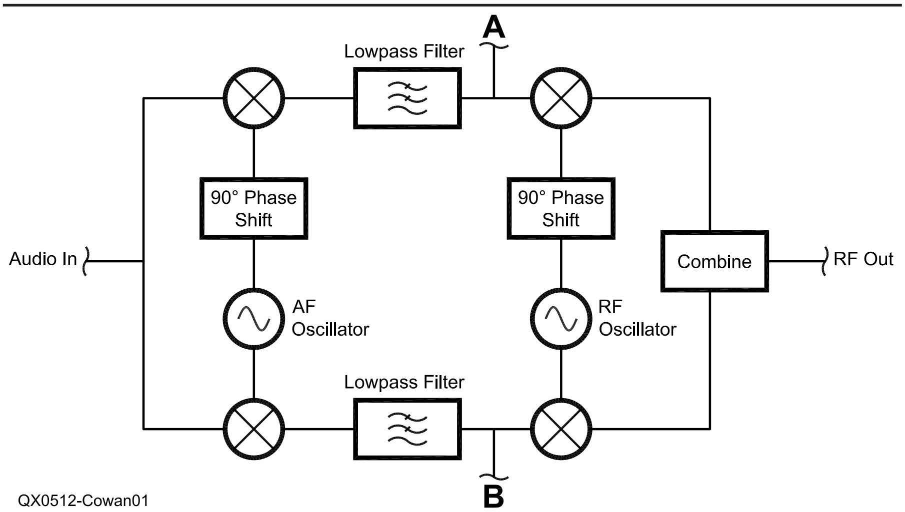


Fig 1—The "Third" or Weaver method of generating SSB signals. The low pass filters cut off at 1350 Hz, such that the "higher" mixing frequencies are eliminated.

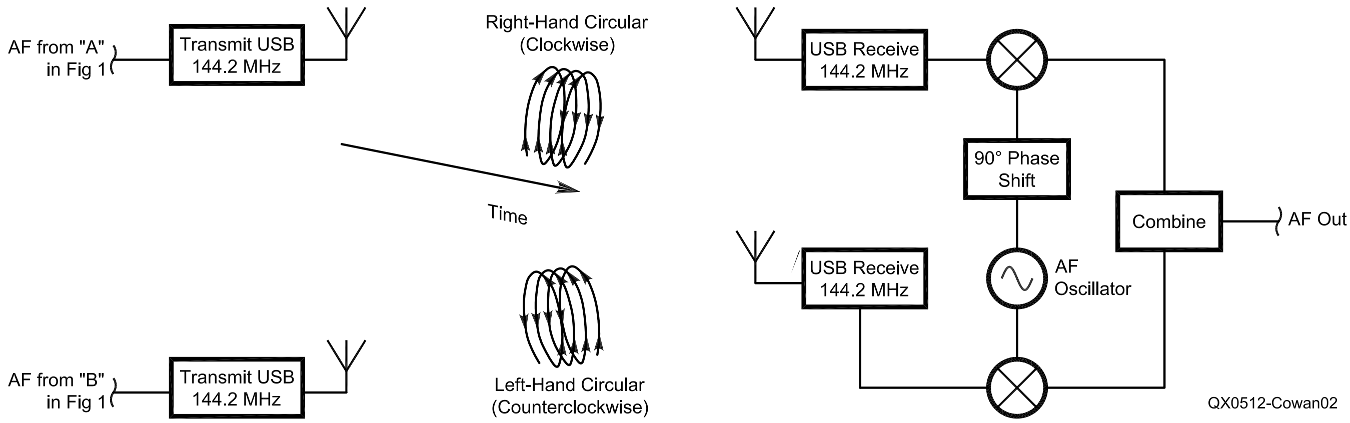


Fig 2—The proposed method uses two transmitters and two receivers, all set to the same frequency.

proposes a method whereby audio (or any other information) can be transmitted using only one-half the spectrum normally required.

The Proposal

Most hams know that there are two different methods used to generate SSB signals (the filter method and the phasing method). The first one used on the ham bands was the phasing method (based on an article in *G. E. Ham News),* but when Collins Radio Company produced a very sharp mechanical filter, the filter method (using mechanical or crystal-lattice-filters) soon dominated.

Many hams don't know that a *third method* of SSB generation was developed, called the "Weaver method" or just "the Third method." This phasing method does away with the need for a wide-bandwidth audio phase-shift network.

Fig 1 shows a block diagram of this method. The audio at points A and B in this circuit contains frequencies from 0 Hz to 1350 Hz. The higher audio frequencies have been folded back on the lower frequencies and phased such that the original audio can be reproduced.

The Basic Scheme

To accomplish my proposed feat, audio from the Third method (points A and B in Fig 1) is used. As shown in Fig 2, two different SSB transmitters (on the same frequency) are used to transmit the two separate audio signals. Since the transmitters are on the same frequency, we need a way to isolate them. This is

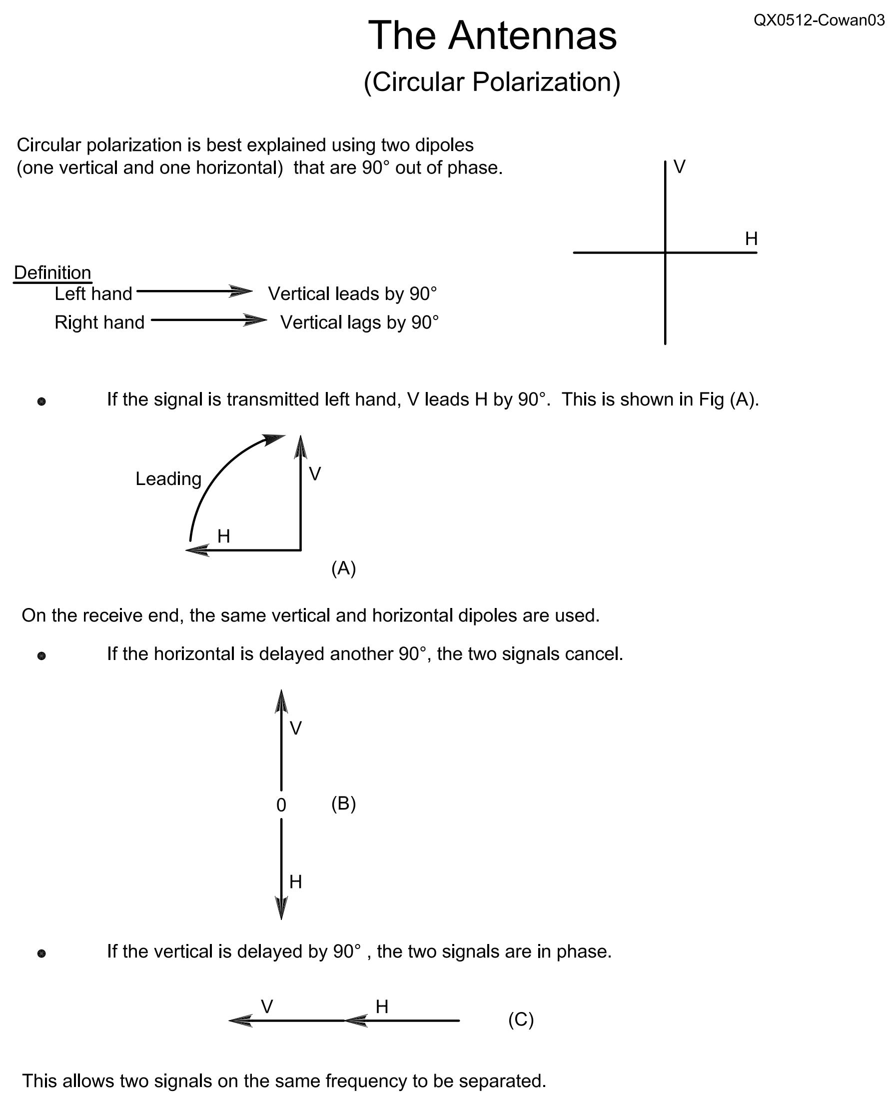


Fig 3—Circular polarization allows two signals on the same frequency to be separated by 20 dB.

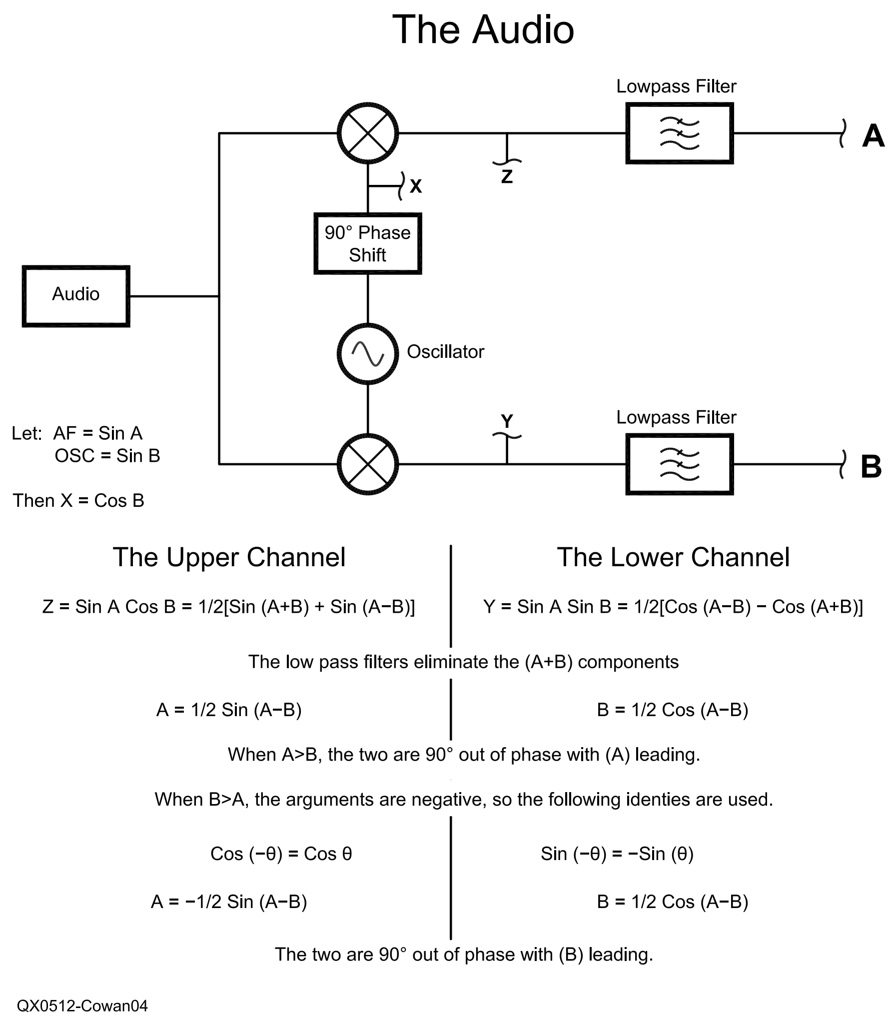


Fig 4—The audio components of the Third method consist of an upper and lower channel 90° out of phase.

accomplished by using right-hand and left-hand circular polarization at the antennas. The response to right-hand circularly polarized signals is reduced (diminished 20 dB—*Ed.*) in a left-hand circularly polarized system. Similarly, the response to left-hand circularly polarized signals is greatly reduced in a right-hand circularly polarized system.

Fig 2 shows that two receivers are also used. The audio output of these receivers will be the same as the audio at points A and B of Fig 1. These outputs can be recombined as shown to produce the original audio. Depending on the phasing, the audio could be either normal or frequency inverted. If inverted, simply switch which balanced modulator gets the 1350 Hz, that is 90° phase shifted.

This scheme should transmit full audio using only 1.35 kHz of spectrum, one-half of that normally required. Although the example uses audio, the same basic scheme could be used to reduce the bandwidth for digital data, analog video and so on.

Sam Cowan, WØOAJ, was first licensed as KNØORA in 1958. He is a retired engineer and electronics instructor, he is the author of one previous QEX *article, "Error Detection and Correction Codes" (Nov 1986) and two books for Prentice-Hall. He earned a BS in Physics and Math from Nebraska Wesleyan University, and attended the Rochester Institute of Technology.*

Source Coding and Digital Voice for PSK31

Source coding more than doubles the throughput of PSK31 and provides an avenue to digital voice at extremely low bit rates.

By Doug Smith, KF6DX

My friend Joe Taylor, K1JT, provided excellent insight into source coding and its benefits in his article in our last issue.[1] It occurred to me that source coding is equally applicable to other existing RTTY formats and perhaps to digital voice, as well. Let me begin with a general definition of source coding.

Joe wrote: "Shorthand radio messages have been widely employed since the days of spark and land-line telegraphy; the familiar Q-signals are another universally understood type. They are simple forms of what in communication theory is called the 'source encoding' of messages." The goal is to minimize the information transmitted while maximizing implied meaning. One obvious example is *QRZ?*, meaning "Who is calling me?" It is clearly more efficient to send 4 characters than 18 in any format. In ASCII code, for example, that constitutes a compression ratio of 18/4=4.5. The compression ratio may be more or less in other formats. It may be found by computing the number of elements or bits used in each message, or by computing the time taken to transmit each.

The chief drawback of source coding is that to maintain good compression ratios, we must limit our lexicon—the number of discrete source codes available. A secondary drawback may arise when considering how to encode many different languages, rather than just one. I will show how that secondary drawback can be ameliorated by pre-coding and post-decoding text processing.

[1]Notes appear on page 49.

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

Source Coding for PSK31

Varicode

I chose PSK31 as my virtual test bed because it employs a clever scheme that its progenitor, Peter Martinez, G3PLX, calls Varicode.[2] The encoding of ASCII characters is shown in Table 1. Since the code "00" is used as the letter separator, no character's code contains that sequence. No more than 12 "1"s in a row are allowed to occur in any transmission. That allows the PSK receiving system to remain synchronized with the transmitter by providing a phase transition at least that often, since a "0" signifies a phase transition and a “1” signifies no phase transition.

Another useful feature of Varicode is that the bit lengths of the encoded characters are roughly proportional to the inverse of their frequency of use in English words. The word space (ASCII character 32) is the most common character and uses only one bit—or three if you include the two-bit letter space that accompanies all characters. Letter “e” is the next most frequently used, followed by “t” and “o,” and so forth. (Actually, “a” is the third most frequently used letter in most English writing. Other languages are different.) It might be argued that in Amateur Radio, the Arabic numerals “0” through “9” are very often used, but Peter elected to code them at lengths of 8 and 9 bits.

Word-Based Source Coding

Character frequency has been well exploited in Varicode but now I am going to ignore it in favor of coding entire words instead. Why? Well, it

Table 1—The Varicode alphabet.

The codes are transmitted left bit first, with “0” representing a phase reversal on BPSK and “1” representing a steady carrier. A minimum of two zeros is inserted between characters. Some implementations may not handle all the codes below 32.

*ASCII**	*Varicode*	*ASCII**	*Varicode*	*ASCII**	*Varicode*
0 (NUL)	1010101011	+	111011111	V	110110101
1 (SOH)	1011011011	,	1110101		
2 (STX)	1011101101	-	110101	W	101011101
3 (ETX)	1101110111	.	1010111	X	101110101
4 (EOT)	1011101011	/	110101111	Y	101111011
5 (ENQ)	1101011111	0	10110111	Z	1010101101
6 (ACK)	1011101111	1	10111101	[	111110111
7 (BEL)	1011111101	2	11101101	\	111101111
8 (BS)	1011111111	3	11111111	]	111111011
9 (HT)	11101111	4	101110111	^	1010111111
10 (LF)	11101	5	101011011	_	101101101
11 (VT)	1101101111	6	101101011	‘	1011011111
12 (FF)	1011011101	7	110101101	a	1011
13 (CR)	11111	8	110101011	b	1011111
14 (SO)	1101110101	9	110110111	c	101111
15 (SI)	1110101011	:	11110101	d	101101
16 (DLE)	1011110111	;	110111101	e	11
17 (DC1)	1011110101	<	111101101	f	111101
18 (DC2)	1110101101	=	1010101	g	1011011
19 (DC3)	1110101111	>	111010111	h	101011
20 (DC4)	1101011011	?	1010101111	i	1101
21 (NAK)	1101101011	@	1010111101	j	111101011
22 (SYN)	1101101101	A	1111101	k	10111111
23 (ETB)	1101010111	B	11101011	l	11011
24 (CAN)	1101111011	C	10101101	m	111011
25 (EM)	1101111101	D	10110101	n	1111
26 (SUB)	1110110111	E	1110111	o	111
27 (ESC)	1101010101	F	11011011	p	111111
28 (FS)	1101011101	G	11111101	q	110111111
29 (GS)	1110111011	H	101010101	r	10101
30 (RS)	1011111011	I	1111111	s	10111
31 (US)	1101111111	J	111111101	t	101
32 (SP)	1	K	101111101	u	110111
!	111111111	L	11010111	v	1111011
“	101011111	M	10111011	w	1101011
#	111110101	N	11011101	x	11011111
$	111011011	O	10101011	y	1011101
%	1011010101	P	11010101	z	111010101
&	1010111011	Q	111011101	{	1010110111
‘	101111111	R	10101111	|	110111011
(	11111011	S	1101111	}	1010110101
)	11110111	T	1101101	~	1011010111
*	101101111	U	101010111	127	1110110101

*ASCII characters 0 through 31 are control codes. Their abbreviations are shown here in parentheses. For the meanings of the abbreviations, refer to any recent *ARRL Handbook*.

turns out a significant compression ratio can be obtained that way. Let me show you how.

English has roughly 1 million words. Obviously I can spell any of them with Varicode but words like *floccinaucinihilipilification* (the act of estimating as worthless!) are, in fact, worthless. Studies have revealed that the typical 3-year-old has a vocabulary of around 500 words; he learns about 10 words a day until graduating from high school with about 60,000 between his ears.[3] That seems a good lexicon size for which to shoot during source-coding development. It would be nice to have a few thousand left over for specialized radio terms, abbreviations and so forth.

A person with a 64,000 word vocabulary typing at 75 words per minute (wpm) is generating about 20 bits/s maximum, since each word can be respresented by 16 bits and:

$$bit\ rate = \frac{(75\ wpm)(16\ bits/w)}{60s/m} = 20\ bits/s \quad \text{(Eq 1)}$$

Seventy-five wpm is very fast for a typist but it is not exactly rapid-fire speech. It is about a quarter as fast as the fastest speakers can understandibly go. A bit rate of 32 bits/s would support 120 wpm but none of that accounts for the need for letter and word spaces. PSK31 supports about 50 wpm as it is. We can do much better with word-based source coding using Varicode-coded characters.

My premise is to use two- and three-character combinations of Varicode codes to represent words. Single Varicode codes would still be used to represent individual letters, numerals, punctuation and symbols. Computing the number of available two- and three-code permutations seems like a grungy task in statistics. Armed with some rather simple relationships, though, I programmed my desktop to discover them, along with some related performance figures.

Refer to Table 2. It shows the number of unique two- and three-code permutations, *including the "00" letter spaces but not the word space*, sorted by bit length. The word-space code is reserved for its originally intended use and is not included in the permutations. The three-code permutations add only the code corresponding to "e" ("1100" including the letter space sequence). The "e" code combines with all the two-code permutations; it may come either first, second or third in the three-code combinations. Hence, we get exactly three times as many of those three-code permutations as two-code permutations, but each is four bits longer than its corresponding two-code partner. The total of 71,156 permutations seems to cover the selected lexicon size nicely.

To minimize transmitted data rate, I assign the shortest code combinations to the most frequently used words in written English. The most common word is "the" and it takes the sole length-8 combination. The longest combinations are 28 bits and those would take about a full second to transmit. That would be a good excuse to assign those codes to words that take a second to pronounce, regardless of their usage frequency. Table 3 shows a few samples from three decades of word frequency rankings,[4] along with bit lengths from plain Varicode and source-encoded Varicode. Note that those 32 words comprise over 25% of written English (!) and only 10 of them are drawn from the top 100, only 21 from the top 1000. The bit-length averages at the bottom of the table are of only marginal interest but they do hint that a compression ratio of two is probable.

To find the actual compression ratio using typical written English, we must take both bit length and word frequency into account. Since the most frequently used words have the shortest codes, we might suspect that the actual compression ratio would be greater than two. I have not compiled a list of 71,156 words yet but I do have a list of the top 2000 words ranked according to decreasing frequency of use (see Ref 4). Those 2000 words constitute 94.24% of written English and so form a reasonably good basis for performance estimates.

To do the computation, I normalized each word's frequency (which is based on a million-word lexicon) to 94.24%. Then I multiplied the normalized frequency by the source-coded bit

Table 2

Showing permutations of two- and three-letter Varicode symbols by bit length, including the "00" letter separators.

Length (bits, w/"00" letter separator)	*# of unique Varicode symbols of this length*	*# of unique 2-letter combos of this length*	*# of unique 3-letter combos of this length**	*Totals*
3	1	0	0	1
4	1	0	0	1
5	2	0	0	2
6	3	0	0	3
7	5	0	0	5
8	8	1	0	9
9	13	4	0	17
10	21	10	0	31
11	34	38	0	72
12	40	65	3	108
13	0	130	12	142
14	0	330	30	360
15	0	390	114	504
16	0	592	195	787
17	0	782	390	1172
18	0	1241	990	2231
19	0	1490	1170	2660
20	0	1955	1776	3731
21	0	2468	2346	4814
22	0	3981	3723	7704
23	0	2720	4470	7190
24	0	1560	5865	7425
25	0	0	7404	7404
26	0	0	11,943	11,943
27	0	0	8160	8160
28	0	0	4680	4680
Totals	*128*	*17,757*	*53,271*	*71,156*

*Three-letter permutations add only Varicode symbol "e" (length 4) to the two-letter combinations listed, in one of three positions: first, second or third.

length, calling that the weighted bit length. I did the same for plain Varicode for each word.

I simply added all the weighted bit lengths to get an average bit length for both Varicode and source-coded Varicode. I added the frequencies as a check. The result was a 2.64 ±0.2 compression ratio. The uncertainy is relatively high on that number because the plain Varicode average bit length was taken directly from Peter's article and only two significant figures were given for the average weighted bit length of Varicode codes.

Details of Possible Implementation

As in plain Varicode, words would be separated by a word space. Single Varicode symbols would be used to transmit numerals, punctuation not inside a word, symbols and so forth. They would have also to be set off by a word space or by multiple letter spaces so that software could recognize them for what they are. A word space would signal the decoder at the receiving end that a source-coded symbol had been received.

Look-up tables would be used at both transmitter and receiver for coding and decoding, respectively. Words not found in the look-up table at the transmitter would have to be encoded as strings of single Varicode symbols, with the appropriate spaces so as not to be confused with source-coded combinations.

Languages other than English could be supported with different look-up tables. Note that letters with diacritics, such as ü, å and é do not form roadblocks to the system since they are parts of the look-up tables and not of the coded symbols. Communicators would simply have to agree to use the same language tables. Perhaps some protocol at the beginning of each transmission could be established to indicate which language was coming and call signs "in the clear" but that might not be necessary.

It would be nice if each word and symbol combination in the English coding scheme translated directly to other languages; that is, if the code for English "and" was the same as for German "und," Spanish "y," and so forth. That way, even if the transmitting station were working in Spanish, you would receive him in English. That is impossible directly, though. Unlike English, many languages use noun, pronoun and article declensions that change with word gender and from singular to plural. German and Spanish are once again handy examples. In addition, many languages have verb conjugations that are gender-, number- or object-dependent. English has relatively few of those but a few persist (I go, you go, he goes). I can think of two solutions to the problem; one seems simple at first and the other, complex.

The simple solution would be to park those unique non-English words in a reserved area of the coding map. If source-coded Varicode were to be optimized for English, then those non-English codes would reside in long bit-length locations that would compromise non-English performance significantly. Having to accommodate two or more verb forms for every tense, as are found in many languages, would take up a tremendous number of codes. English seldom encounters that situation. For instance, the past tense of "run" is "ran" whether I ran, he ran, you ran, we ran or they ran. In many languages, the form of the verb depends on the object for active tenses and on the subject for passive tenses.

The complex solution involves processing of received text prior to presentation to the user. Software would be designed to parse sentences for their objects (or subjects) and the associated verb. It would have to correct the verb form for a given tense based on that parsing before display-

Table 3

A sampling of words from three decades of frequency ranking, comparing their lengths in plain Varicode and in source-coded Varicode. For the 2000 most commonly used English words, the plain Varicode average length is actually 39.38 bits (unweighted).

Frequency rank	*Word*	*Frequency (%)*	*Varicode bit length*	*Source coded bit length*
1	the	7.308	17	8
2	of	3.803	13	9
3	and	3.014	20	9
4	to	2.732	10	9
5	a	2.440	6	6*
6	in	2.101	12	9
7	that	1.043	24	10
8	is	0.9943	13	10
9	was	0.9661	22	10
10	he	0.9392	12	10
100	down	0.0881	28	13
101	should	0.0874	41	13
102	because	0.0869	46	13
103	each	0.0863	26	13
104	just	0.0858	31	13
105	those	0.0837	29	13
106	people	0.0834	36	13
107	Mr.	0.0826	26	13
108	how	0.0823	22	13
109	too	0.0820	15	13
110	little	0.0818	34	13
1000	current	0.0102	45	16
1001	spent	0.0102	30	16
1002	ate	0.0102	15	16
1003	covered	0.0102	46	16
1004	role	0.0102	23	16
1005	played	0.0102	42	16
1006	I'd	0.0102	28	16
1007	date	0.0102	23	16
1008	council	0.0101	48	16
1009	race	0.0101	25	16
1010	Charles	0.0101	49	16
Totals		*26.383*	*26.0 avg.*	*13.1 avg.*

*"a" is always transmitted as a single Varicode symbol even when used as an indefinite article in writing. The same is true of "I."

ing a decoded message to the user. Such software must obviously be tailored to the grammatical dictates of each individual language and it is a tricky business, to say the least. Nonetheless, we do have grammar-checking software out there these days along with spell-checkers. It is not entirely unreasonable to think that passing raw received text through such a filter could be done. In that case, it ought to be possible to match English and foreign words exactly by their source codes. In addition, we do see language translation software these days that does well. We shall have another look at text preprocessing later as part of the following discussion of source-coding for extremely low bit rate digital voice.

Source Coding and Digital Voice

As we have a system for transmitting a little over 120 wpm over PSK31, the method presents itself as an extremely low-bit-rate digital voice system. Speech recognition and synthesis software is required. At two words per second, the rate is just high enough to be tolerable. You will not call any auctions or horse races at that rate but you will be able to transmit voice at signal-to-noise ratios (SNRs) well under 0 dB—with proper filtering and signal processing. The quest for that level of performance is my main motivation and is the goal of others with whom I have consulted. Existing digital voice systems reproduce natural-sounding speech but they occupy 2.7 kHz or more. My experience has been that they start degrading below 20 dB SNR and become virtually unusable at around 7 dB SNR.

For this scheme, I evidently have to abandon my notion that to be acceptable, a digital voice system must sound natural and must carry the nuances of human voices that allow you to tell who is speaking, whether they have a cold, and so on. In this system, only the word information will be transmitted and none of the pitch, formant and other important traits of individual voices.

Will it sound like a robot? Yes, to some extent; but there are certain things that can be done at the receiving end to enhance naturalness that I will try to explain. Many of you have heard the NOAA weather radio broadcasts that use text-to-speech software to generate a human-sounding voice. That is an efficient way for them to go but I wonder if it put someone out of a job. In any case, the burden of generating all the voice inflection, word emphasis (or de-emphasis) and other nuances is placed on the transmitter in the NOAA system. The receiver can be expected to do none of it. In my source-coded digital voice system, it is the opposite: The burden is placed entirely on the receiver.

Block Diagram

Fig 1 is a block diagram of the system. At the transmitter resides a computer with a sound card and microphone. The user speaks and speech-recognition software running on the computer matches recognized words to source codes for transmission. If the user speaks faster than 150 wpm then a buffer holds coded data for transmission. Note that with appropriate delays in place, the user would have the opportunity to correct speech-recognition errors and to delete “ums” and “ers” in his speech before transmitting the data. VOX operation would be possible in the traditional way.

At the receiving end of the link, speech-synthesis software generates speech from received text. The

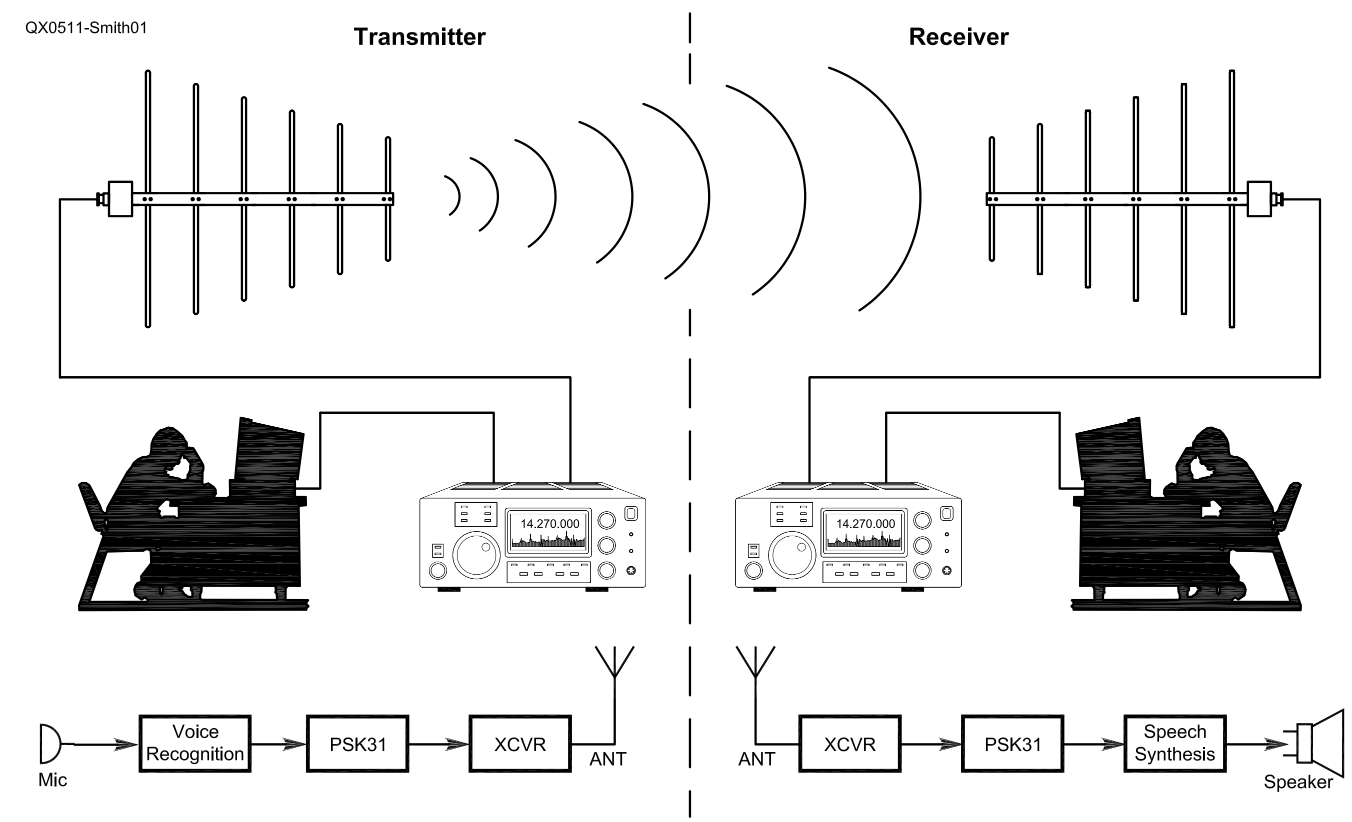


Fig 1—Block diagram of a PSK31-based digital voice system.

software incorporates text preprocessing to ascertain sentence grammar and other linguistic traits so it can impart inflection and other nuances as appropriate. For languages other than English, it also performs grammar checking to correct article and verb tense forms according to the number or gender of object or subject.

Speech Recognition: Speech to Text

I do not propose to write speech-recognition software, since several very good programs are already out there.[5] Still, it might be instructive to explore some of the highlights of how it is currently done.

In many speech recognition applications, a machine need only understand single words, short phrases or simply part of what is spoken to carry out its mission. That often occurs in control systems where the vocabulary is severely restricted and can be referred to as *isolated word* or *connected word* mode. However, dictation systems have existed for at least 10 years that do a creditable job of converting speech to text. There, a large lexicon must be supported along with a control vocabulary that allows the user to correct errors without actually touching the machine. During dictation, speech to be converted to text can be called *continuous speech* mode and the control functions use one of the other two modes. The same is true for a digital voice application.

A second characteristic of speech recognition systems has to do with knowledge of a speaker's voice patterns. Systems may be *speaker-dependent*, *speaker-independent* or *speaker-adaptive*. Speaker-dependent systems are customized to one or more individual speakers whose speech patterns are stored in machine memory for later reference. Those include systems that must be trained to recognize words before putting the system to use. Speaker-independent systems are tailored to work with a broad variety of speakers, about which the systems have no previously stored knowledge. Adaptive systems alter stored speech pattern information using feedback based on errors they make while the system is in use.

Even with a single speaker, variations in pronunciation and the use of words and phrases make speech recognition inherently difficult. Variability across many speakers is certainly even greater and includes regional accents, different voice pitches and different speaking speeds. Another deleterious factor may be the speaking environment. Background noises, echoes and other stray stimuli may easily destroy one's chances for success.

A third issue relates to how much linguistic knowledge is instilled in the coding algorithm. A practical dictation system would have to know the difference between homophones like "red" and "read", "write" and "right" and so forth to put the correct word in context. The system might have to engage in a dialogue with the user to clarify ambiguities.

Basic Algorithm Classes

Let us look at three basic algorithm classes for speech recognition: artificial intelligence algorithms, phonetic algorithms and pattern-matching algorithms. Each of those has been tried with varying degrees of success.

Artificial intelligence algorithms combine acoustic features of speech with linguistic knowledge and other knowledge of human speech patterns. So-called expert systems (neural networks) are generally employed that attempt to mimic what we know about how speech recognition is accomplished by human beings. In other words, expert systems take everything the game gives them and try to emulate the human ear-brain combination. This approach is very computationally intensive and has not found its way into practical commercial systems as yet.

In a phonetic algorithm, various discrete speech sounds are sought in the input and labels are assigned to them. The idea is that we can define a finite set of phonetic units or phonemes and pick them out of continuous speech one by one. Then the phoneme labels are combined to find a match with words from a vocabulary stored in machine memory. As example, the word "ditch" contains three phonemes: initial consonant "d," short vowel "i" and final diphthong "tch." Only one match for that labeled phoneme combination would be found in the stored lexicon. Ambiguity could conceivably still arise, though, because some phonemes are very close to one another in sound, or are identical. The "tch" sounds identical to "ch" in almost every case I can think of. Letter "j" sounds identical to soft "g" and "s" is identical to soft "c." Fortunately, English has avoided virtually all that ambiguity. Dich (ditch) is not a word, nor is gelly (jelly). You would not find a match for joos (juice) or fâs (face). You would, however, still have trouble distinguishing homophones without using some linguistic knowledge or user feedback to process the text before transmission.

Pattern-matching algorithms are found in almost all modern speech-recognition systems. They constitute a superset of phonetic algorithms because they may employ not only phoneme matching but matching of entire words or phrases. They are explicitly associated with speaker-dependent algorithm training, which may take place before the system is put into use, or during use adaptively.

Pattern Training and Matching Details

Fig 2 is a block diagram of a non-adaptively trained pattern-matching system. With the switch in the training position, speech is analyzed and significant data about its features are stored either as templates, statistical models, or both, in machine memory. With the switch in the recognition position, speech is compared with the stored templates or models to provide data to a decision-making engine. The stored comparison data are referred to as templates if they are non-parametric representations of the sound to be matched to a spoken pattern. For example, they might be time-domain waveforms recorded during training. Stored comparison data are referred to as statistical models if they contain information about parameters of speech that are part of a speech

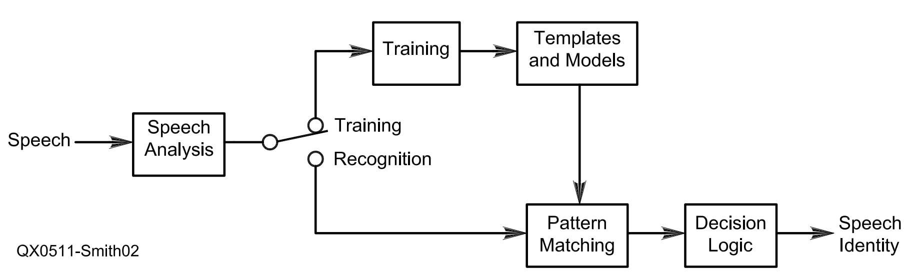


Fig 2—A non-adaptively trained speech recognition system.

modeling system. For example, they might contain information about spectral content, voice tract length and pitch, or parameters for a linear predictive coding (LPC) algorithm. For either templates or statistical models, the time dimension is crucial. How a sample of speech changes over time is all we have to go on.

As I have pointed out on these pages before, speech does not change much over time frames of about 10-30 ms. So the speech analysis block of Fig 2 must provide samples only every 20 ms or so when statistical models are being used. Templates may require a higher sampling rate and therefore may be less memory-efficient; although zero-counting algorithms, which have been used with some success, are very memory-efficient. Studies over the years have shown that the human hearing system response is logarithmic with sound intensity. They have also shown that it responds in proportion to the spectral power distribution rather than amplitude distribution. Further studies reveal that a time-spectral representation (how the spectrum changes over time) contains all information necessary to fully characterize speech. Thus we set aside purely time-domain templates in favor of a time-frequency approach using logarithmic power spectra that change with time.

The Fourier transform of a logarithmic power spectrum is called the cepstrum. It got that funny name in a paper with a very funny name.[6] The cepstrum provides a very efficient way of characterizing short segments of speech for phonetic identification and matching. An individual phoneme may be recognized in the speech analysis engine as a contiguous set of cepstral samples that do not change much over several successive 20 ms periods. I refer to such a contiguous set as a stable sample block. Words and phrases can be characterized as a series of connected phonemes, but some words and phrases do not fit that model as well as others. For example, the vowel sound in "mine" is one in which the shape of the mouth and position of the tongue change continuously—along with the spectral content. The specrum is never really stable but you can almost reproduce it with the concatenation of two phonemes: "ah" and "ee": "ah-ee." Placement of a few intermediate phonemes would make that crude example into a better model. On the other hand, from many speakers in states of the southeastern U.S., "mine" comes out as m-ah-n, which is simpler by at least one step.

Now enter a statistical model called the hidden Markov model (HMM).[7] The HMM is used in conjunction with phoneme templates during training to create a sequence of values and probabilities for word or phrase characterization. Refer to Fig 3. Each stage in the sequence corresponds to one phoneme or stable sample block. Since none of the stages is associated with absolute certainty of phoneme recognition (that's the hidden part), values and uncertainties are assigned for each stage that correspond to the cepstral energy distribution, to the stage duration and a few other things not shown in the figure. During training, those values and coefficients derived from them are stored, or combined with previously stored data in the case of multiple training opportunities. With a speaker-dependent system, the combining may take the form of simple averaging; a speaker-independent or speaker-adaptive system may have to use more sophisticated techniques. In any case, a variety of algorithms is available for assigning the stored values. They range from simple to esoteric. I will not go into them here.

Pattern matching consists of finding the stored statistical data (the reference obtained from the HMM during training) that most closely matches those of the spoken word or phrase (the unknown obtained from the HMM during live speech). A big hurdle in the way of finding a match occurs when part or all of the speech to be recognized is not spoken at the same rate as it was during training. A procedure called dynamic time warping[8] may be employed to stretch or shrink either the reference or the unknown until a best fit is found. Then there is the matter of absolute time alignment between reference and unknown. Alignment may be optimized by shifting either reference or unknown backward or forward in time.

One shudders to think that the foregoing matching procedure would have to be performed between the unknown word or phrase and every reference word or phrase, but that is not neces-

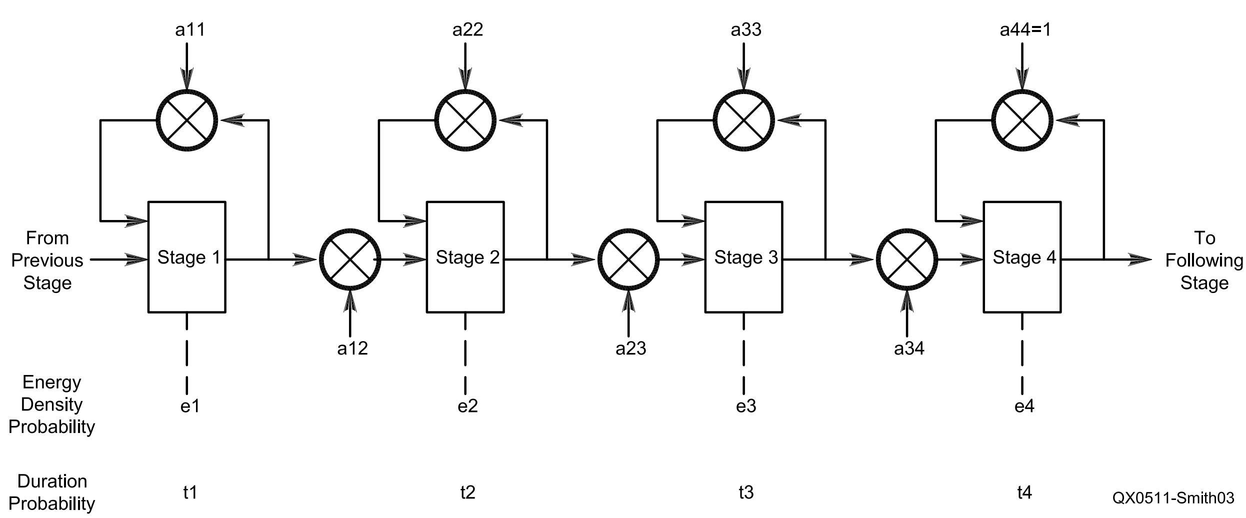


Fig 3—Simplified block diagram of a four-stage hidden Markov model. A detailed block diagram is difficult to draw since it needs to look three-dimensional where the coefficients are determined.

sarily true. You can initially try to match the first phoneme of the unknown with only those reference data that correspond to words or phrases starting with that phoneme. Then you try to match the second, third and so on until a best fit is obtained. In fact, the search order may be optimized in other ways. Yes, the processing power required is significant; but if it were easy, everybody would be doing it!

Upon achieving a satisfactory match, a textual representation of the word or phrase may be taken from a look-up table and placed in a buffer. Linguistic preprocessing is usually necessary to avoid the wrong homophone for those words that have homophones. You might have to wait for another few words to come along before your preprocessor can decide.

Performance

You might likewise have to wait for a few of the previous and following words to go by before your pattern-matching machine could do its job, too, because in fluently spoken language, patterns of connected words or syllables must be recognized, not just the individual words themselves. Considerable linguistic knowledge must often be applied to determine the correct words. As recently as 10 years ago, error rates for isolated word mode sometimes ran as high as 4%.[9] Connected-word and continuous speech modes typically saw error rates as high as 10%. Significant improvement has been reported since then but to get to 0.1% territory, only an adaptive system or a system whose entire vocabulary is speaker-dependent will do. An adaptive system learns its mistakes through user feedback. It would drive you nuts at first but in the long run, the frequency of errors would diminish toward a lower limit.

Speech Synthesis: Text to Speech

Unlike speech recognition algorithms, speech synthesis has a long history. A speech synthesizer was demonstrated at the 1939 World's Fair, for instance. I can only imagine what it sounded like, but I bet it was far from natural.

If you were to build the speech recognition engine described above and train it, then you would automatically have all the statistical acoustic information to replay words as speech. You know all the phonemes and their durations. It would be a relatively simple matter to translate written words to audio and just concatenate them. But you would certainly get the sound of Robby the robot; native speakers do not sound like that. They learn to speak fluently, to place emphasis and de-emphasis appropriately within sentences, to pause occasionally between phrases and sentences, and to use inflection (change of intonation). Even ably punctuated writing cannot convey all the information necessary to impart those characteristics to a spoken version. Human beings rely on their knowledge of grammar and meaning to do the job. It follows that we have to put some of that knowledge into speech synthesis software to approach natural-sounding speech.

A machine can mimic a knowledge of grammar to a considerable degree but it cannot be said to understand the meaning of a sentence, especially within any large context. Speech synthesis algorithms thus must use chiefly grammatical cues and pauses, where possible, to determine what timing, emphasis and inflection to use. Some additional cues outside of grammar and punctuation are possible in a very limited number of instances.

Preprocessing of text to analyze grammar, source timing and any other cues is clearly the first step in speech synthesis. That naturally causes a delay in the speech output because software must parse words and phrases both before and after the word currently being considered. A delay might not be all bad, though, because it could allow sentences to be synthesized and played at faster than 120 wpm, aiding naturalness, with the corresponding longer pauses between phrases or sentences.

The first step in text preprocessing would be to find certain delimiters within the text, such as pauses or full stops that would be signified by two or more contiguous word spaces in the source-coded scheme. Those silent periods are good boundaries between phrases or sentences for block preprocessing. A second step might be to analyze syntax to categorize words by their parts of speech: nouns (object, subject, neither), verbs, adjectives, adverbs and so on. That information can be quite useful in determining

Table 4

Showing weighted bit lengths for plain Varicode and source-coded Varicode, to compute average bit lengths for the 2000 most commonly used English words. Those 2000 words constitute 94.24% of written English. The average length for plain Varicode words was taken from Peter Martinez' PSK31 article and was not computed by me.

Frequency rank	*Word*	*Relative frequency (%)*	*Weighted Varicode length (bits)*	*Weighted source coded length (bits)*	*Compression ratio*
1	the	7.308	1.2424	0.5846	2.125
2	of	3.803	0.4944	0.3423	1.444
3	and	3.014	0.6028	0.2713	2.222
4	to	2.732	0.2732	0.2459	1.111
5	a	2.440	0.1464	0.1464	1.000
•	•	•	•	•	•
•	•	•	•	•	•
•	•	•	•	•	•
1996	games	0.00573	0.0019	0.0010	1.9
1997	cultural	0.00573	0.0032	0.0010	3.2
1998	plenty	0.00573	0.0022	0.0010	2.2
1999	mile	0.00573	0.0014	0.0010	1.4
2000	components	0.00573	0.0036	0.0010	3.6
Totals		*100.0 ±0.1*	*32.5 avg ±2.0*	*12.311 avg ±0.01*	*2.64 avg ±0.2*

accentuation, phrasing and intonation. In the sentence, “If I were you, I’d put in for a vacation,” emphasis might be placed on “you” or “I” and on the second syllable of “vacation.” Emphasis in this instance means a slight loudness increase and an increase in the fundamental voice frequency or pitch. Words or syllables emphasized in that way are said to be accented. Nouns, verbs, adjectives and sometimes adverbs tend to be accented. Other words or syllables may need to be de-accented. In general, function words, including auxiliary verbs and prepositions, are de-accented. The third type of word or syllable would be unaccented or cliticized. Cliticized words lose all their stress and are generally monosyllabic.

Really, there is a lot more that goes into a good accenting algorithm. Take for example the sentence, “I’m not a water skier, I’m a snow skier.” The half-baked accenting rules above might place accents on “skier” in two places rather than on “water” and “snow” where they belong. The two instances of “skier” should actually be de-accented. Algorithms have been developed that force de-accenting of words that are repeated within a phrase or sentence and accenting of the words that precede them but some seemingly intractable situations still exist. Try the sentence, “I’d like to help my sister, not hurt her.” Most of today’s accenting algorithms are not good enough to accent “help” and “hurt” because words are not repeated anywhere. The only clue a machine might have is that those words are antonyms and the second word is preceded by “not.” Work is ongoing regarding such things but just as soon as you think you have a new rule, someone comes with a sentence which fails to satisfy, such as this: “This is not a court of justice, it is a court of law.” Admittedly, that sentence could stand without the accents but you get the idea.

All dictionaries show information about pronunciation and accenting of syllables in polysyllabic words. What they do not show is such information about open compound words. That information could be stored separately but then we would be getting away from general rules into specific cases.

Some syntactic parsers use certain key words to break sentences into logical phrases for the purpose of accenting and intonation. That approach has met with the same kind of limited success as the others. It seems there are always exceptions to the rule. Pauses or full stops in speech are currently the only reliable indicators of phrase boundaries.

In most speech-synthesis systems, an intonation curve is computed for each phrase identified by the syntactic parser. The curve depends on the cadence of the phrase and the placement and number of accented, de-accented and unaccented words. Computing it is a complex task that is beyond my ken but Bell Labs and others evidently have done considerable work on the problem.

Conclusion

I have shown that source-coding of standard Varicode can boost PSK31 throughput to over 125 wpm on average, which is enough for digital voice at a casual speaking rate. My digital voice system would use currently available technology. Note that it would provide both written and acoustic output at the receiving end without modification of the protocol. The possibility exists of universal language compatibility. Have I built and tried it? No, but I have tried speech recognition and synthesis systems, including every time I dial information or my bank’s automated teller on my telephone.

My digital voice system should be able to set some voice contact records on every band because of its noise performance advantage. It might be a bit impractical (or not) but it was interesting to me to be able to show that it could be done. Comments are welcome!

Notes

[1] J. Taylor, K1JT, “The JT65 Communications Protocol,” *QEX*, Sep/Oct 2005.

[2] P. Martinez, G3PLX, “PSK31: A New Radio-Teletype Mode,” *QEX*, Jul/Aug 1999.

[3] J. Kaminski, J. Call and J. Fischer, “Word Learning in a Domestic Dog: Evidence for ‘Fast Mapping’,” *Science*, June 11, 2004.

[4] List from the Virtual Language Centre, Hong Kong, **http://vlc.polyu.edu.hk**.

[5] For examples, try **www.voicerecognition.com**.

[6] B. Bogert, M. Healy and J. Tukey, “The Quefrency Alanysis of Time-Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum and Saphe Cracking,” *Proc. of the Symposium on Time Series*, John Wiley & Sons, New York, 1963.

[7] L. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc IEEE*, Feb 1989.

[8] H. Sakoe and S. Chiba, “Dynamic programming optimization for spoken word recognition,” *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Jun 1981.

[9] L. Rabiner and B. Juang, “Speech Recognition by Machine,” *The Digital Signal Processing Handbook*, V. Madisetti and D. Williams, eds. CRC Press, Boca Raton, FL, 1998.

Antenna Options

By L. B. Cebik, W4RNL

Modeling Software, Part 2

In Part 1 (Sep/Oct, p 54) of this parade through the options facing the new modeler or the modeler wishing to upgrade his or her capabilities, we examined a number of fundamental choices among the available calculating cores for round-wire modeling activities. Besides noting a few (but certainly not all) of the differences between *NEC* and *MININEC* programs, we also looked at some differences between *NEC-2* and *NEC-4*. As we came ever closer to sorting among the available commercial implementations of each core, we found variations in the model file formats used among programs. Finally, we explored the need for a modeler to consider both short- and long-range modeling activities when deciding on the segmentation limitations that attach to some programs.

As if these were not enough options for one to consider when investing in modeling software, we shall explore in this episode a number of other options and what they may (or may not) mean to the modeler. Throughout, we are bypassing explanations of how modeling programs do what they do, since that information appears in many other articles and books. As well, we are not recommending any one or more individual program, but trying to set forth the considerations that each program purchaser should think about when exploring the available programs at their Web sites. Last time, we presented a table of programs and the URLs to use when exploring the possibilities.

The Availability of Commands

MININEC appeared with an abbreviated set of commands relative to *NEC*-2. The available implementations of both public domain *MININEC* and the proprietary version known as *Expert MININEC* have added to the original command list. In general, however, the number of added commands is small.

NEC is another matter. New modelers often select entry-level software that internally restricts the number of commands relative to the full list accepted by the core. The general purpose in the restriction is to provide the modeler with a user-friendly interface for setting up a model and for examining the output. Among the most notable examples of this practice are standard *EZNEC* and *NEC-Win Plus*, although each program uses a unique interface.

To get a handhold on the range of commands that are usable by the *NEC* cores, examine Table 1. It presents a complete list of commands for *NEC-2* and *NEC-4*. Note that some commands are specific to each core and that some commands change formats when moving from one core to the other. The *NEC* manuals for both versions tend to divide the commands into those that specify the geometry structure of the antenna and those that control the model afterwards, either by introducing modifications—such as loads or transmission lines—or by specifying the desired output data.

Entry-level programs tend to restrict the number of geometry commands to just three: GW, GS, and GE.

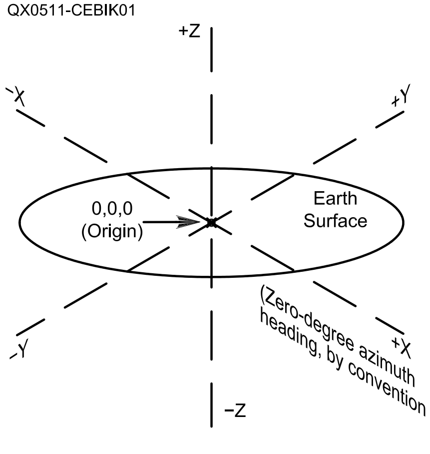
1434 High Mesa Dr
Knoxville, TN 37938-4443
cebik@cebik.com

GW specifies a wire's coordinates, number of segments, and radius (although a program may list the entry as the wire diameter). In most cases, the actual user entry is invisible with respect to the command name. GS is a necessary entry that many entry-level programs automatically insert to convert the user's unit of measure into meters, the unit required by the core for calculation. GE simply marks the end of the geometry portion of the model.

Full *NEC* programs that permit ASCII entry of commands allow the use of all of the commands applicable to the core in use. Such programs are *4NEC2*, *NEC-Win Pro*, and *GNEC*. Some of these programs permit multiple modes of entry. For example, the listed Nittany Scientific programs have an assist screen for each new command so that the user does not need to worry about the line format. In addition, they have an insert—essentially, the *NEC-Win Plus* program—to allow wire, source and load creation in a manner identical to the entry-level program. Finally, the programs have a wire-assist function that aids in transferring model geometries created on spreadsheets into the *NEC*-model format.

The command set is useful in creating complex geometric structures in a compact form. Some programs—*EZNEC* especially—have wire-formation functions that replicate many of the commands in the geometry list. For example, one can form length-tapered elements (GC), helices (GH), along with wire-grids and radials systems (GM). The results are a list of individual wires, each equivalent to a GW entry. Consider a set of monopoles spaced $^1/_2$-λ apart above a wire-grid simulation of a ground-plane surface. Depending on the number of monopoles and the outer dimensions of the rectangular wire grid, the *EZNEC* construct might have from hundreds to thousands of individual wire entries. In contrast, using the GW entry for the first wire in each group and the GM command to replicate it the required number of times, we might set up any monopole array and its ground surface in no more than six entries. Mastering the full command set for either *NEC-2* or *NEC-4* requires a far longer learning curve than becoming able to produce useful results from the core with a more restricted command set.

The control commands are equally restricted in entry-level programs. Among the structure modifying commands, the EX or source command used in entry-level programs is least versatile. The user may specify a voltage source placed on a segment of the geometry. These programs do create an indirect current source, if the user wishes, a function not directly available in range of selections available under the EX label. Among the excitation variations unavailable to the entry-level program user are those involving plane wave excitation —both linear and elliptical. These



The Cartesian Coordinate System With an "Earth Surface" Shown

Fig 1—The Cartesian coordinate input system.

Table 1—The *NEC* Commands

Command	*Command Name*	*Notes*
Structure, Geometry	***Input Cards***	
CM, CE	Comment Cards	
CW	Catenary Wire	*NEC-4* only
GA	Wire Arc Specification	
GC	Continuation Data for Tapered Wires	
GE	End Geometry Input	
GF	Read Num. Greens Function	
GH	Helix-Spiral Specification	Form differs between *NEC-2/NEC-4*
GM	Coordinate Transformation	
GR	Generate Cylindrical Structure	
GS	Scale Structure Dimensions	
GW	Wire Specification	
GX	Reflection in Coordinate Planes	
SP	Surface Patch	
Program Control Cards		
CP	Maximum Coupling Calculation	
EK	Extended Thin-Wire Kernel	*NEC-2* only
EN	End of Run	
EX	Excitation	
FR	Frequency	
GN, GD	Ground Parameters	
IS	Insulated Wires	*NEC-4* only
JN	Junction Charge Distribution	*NEC-4* only
KH	Interaction Approximation Range	*NEC-2* only
LD	Loading	
LE, LH	Near Fields along a Line	*NEC-4* only
NE, NH	Near Fields	Form differs between *NEC-2/NEC-4*
NT	Networks	
NX	Next Structure	
PL	Data Storage for Plotting	*NEC-4* only
PQ	Print Control for Charge on Wires	
PT	Print Control for Current on Wires	
RP	Radiation Pattern	
TL	Transmission Lines	
UM	Upper Medium Parameters	*NEC-4* only
VC	Include End Caps	*NEC-4* only
WG	Write Num. Greens Function File	
XQ	Execute	

commands are useful in modeling activities that analyze the receiving and scattering properties of antennas.

At the lowest entry level, the user has access only to far-field data in both tabular and graphical forms. The graphics are not a function of the core, but are fairly standard functions that programs provide to ease user understanding of the data. Some programs provide tabular near-field data, and others provide—at a more advanced level—surface-wave data. Missing from the list are other options attached to the RP command, and totally missing are the receiving (PT) and mutual coupling (CP) options for output requests. These options are only available in full *NEC-2* or *NEC-4* implementations.

The GF and WG commands associated with reading and writing numerical Green's files are especially useful to those who must repetitively use portions of large model files. Consider a series of wire-grid reflector structures having various dimensions. Then consider having to test each reflector with a variety of driver structures, each of which may require changes of position relative to the reflectors. The modeler can initially create a series of reflectors and create (WG) Green's files for each one. Then, by recalling the appropriate file (GF), the modeler can add the driver and run the entire model in a fraction of the time needed to re-run the entire reflector portion from scratch.

Whether your modeling—both now and in the future—requires the additional commands available in the full *NEC* programs is more than an idle consideration. You may well not need to make a very long-term decision at the beginning. Software makers do offer discounts for upgrading, but only when the move is from one program to another in a single product line.

Input/Output Style and Presentation

The input and output systems of coordinates often take some users by surprise. The surprise factor is often a function of the system with which the user is most familiar: the compass rose versus the Cartesian coordinate system. Fig 1 shows the Cartesian coordinate system as used by both *NEC* and *MININEC*. In all cases, the +X-axis corresponds to the 0° heading used in the output reports for radiation patterns. +Y corresponds to 90°. Of course, –Z values are valid only for free space models in all of the cores and only in *NEC-4* for models using a ground.

Note that the system effectively counts counterclockwise. For most antenna structures, the direction of counting presents no problem. However, when constructing arrays of multiple antennas, such as an AM broadcast tower set, not only are directions from one tower to the next important, but the model may be used employing a compass rose or clockwise orientation for the field geometry. So far as I know, only *Expert MININEC* offers an option for input values using true compass or azimuth headings.

Fig 2 shows the circles for the output conventions. Inherently, *NEC* and *MININEC* use the phi or counterclockwise system for headings in the X-Y plane, along with a Z-axis theta system that counts from the overhead or zenith angle downward. Some software systems offer only this option for the presentation of radiation patterns, but other implementations also offer (and refer to) azimuth and elevation headings. In most cases, conversion of theta angles to elevation angles (angles above the horizon) is simple. However, the phi-to-azimuth (clockwise) conversion is more complex. Hence, some software simply calls the phi patterns azimuth patterns. Other software, like the sample standard 2-D azimuth pattern in Fig 3 from *NEC-Win Plus*, modifies only the outer ring values without flipping the pattern itself. For symmetrical

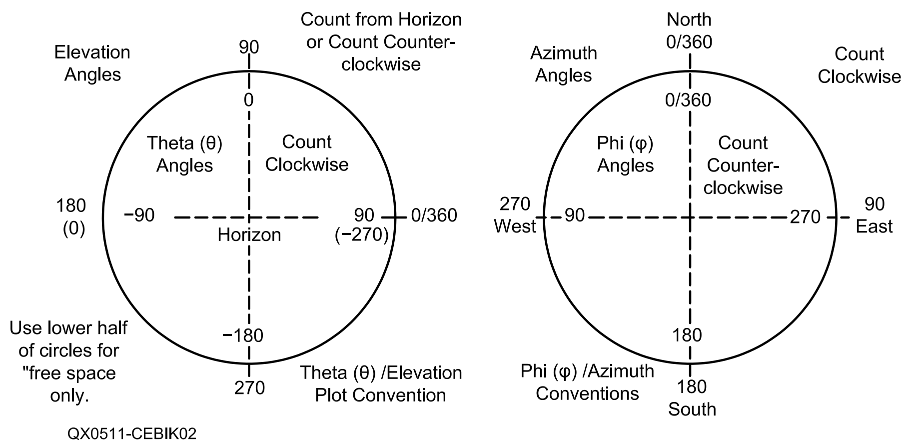


Fig 2—The phi/theta vs the azimuth/elevation radiation pattern systems.

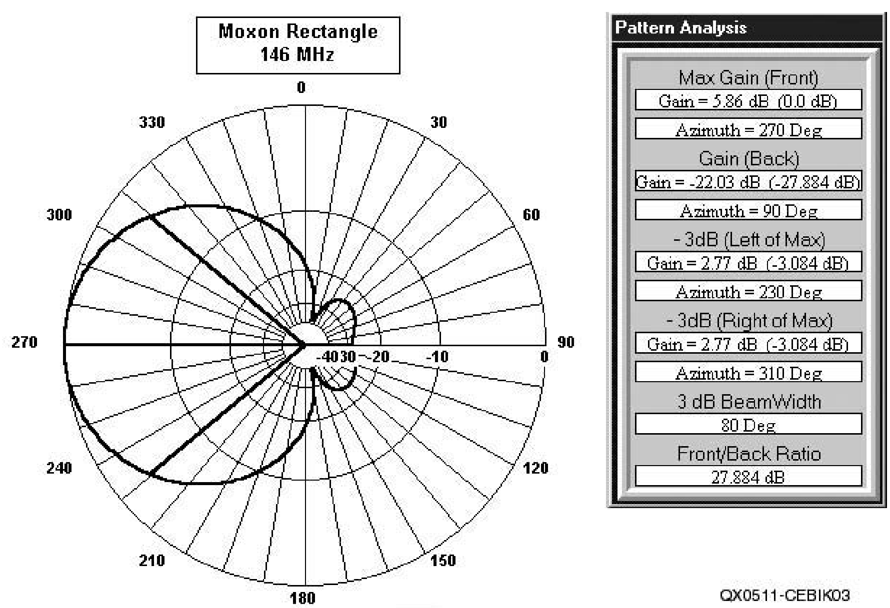


Fig 3—A standard type of 2-D output "azimuth" pattern (from Nittany Scientific software).

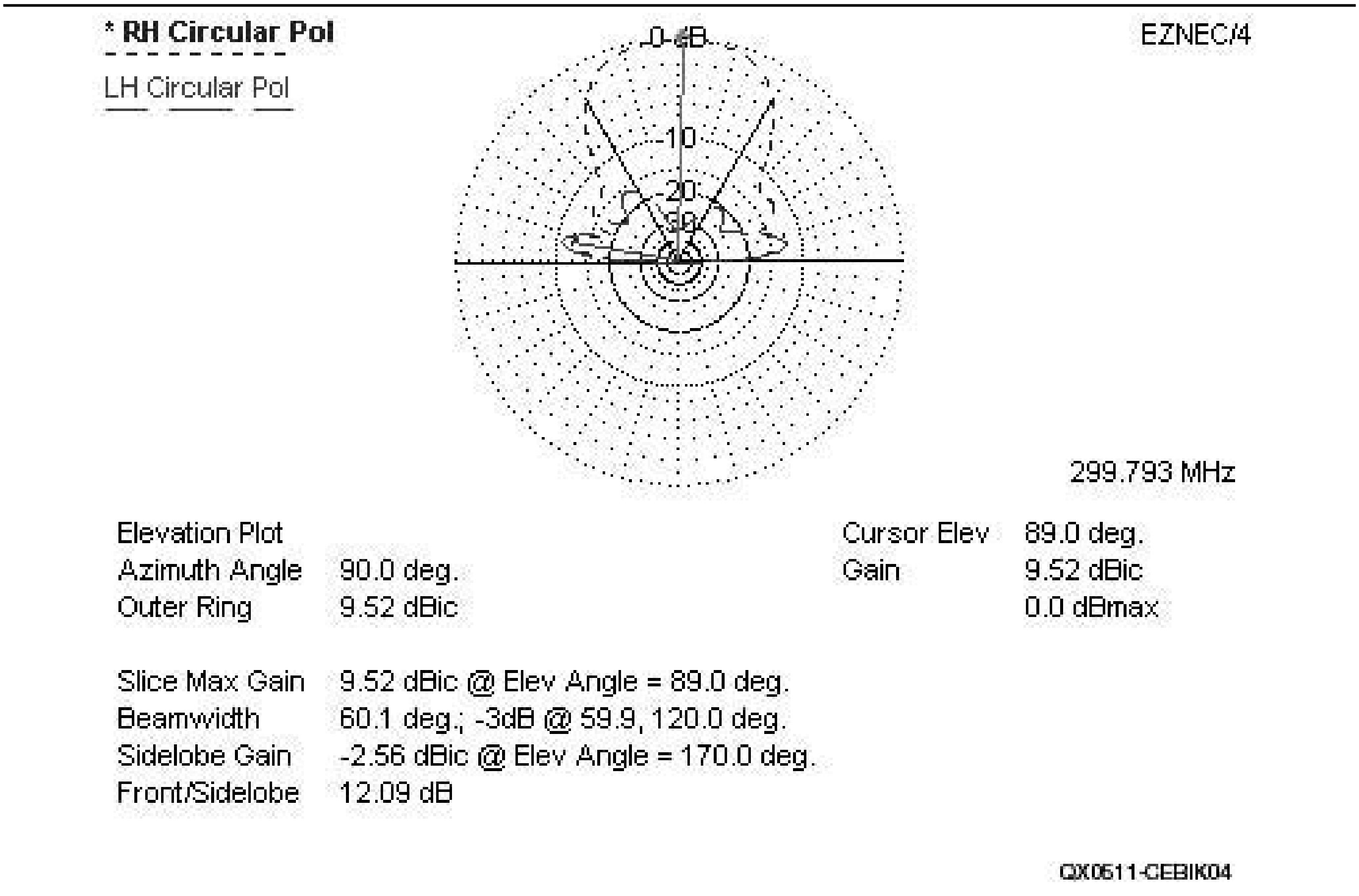


Fig 4—An *EZNEC* pattern showing circular polarization components.

patterns, this practice makes no difference, but may create confusion for non-symmetrical patterns if the user forgets what conversion process is at work.

In the realm of 2-D patterns, there are many more options for presentation than simply the total pattern. Post run calculations can sort out the left-hand and right-hand circular components of a pattern, as in the *EZNEC* sample shown in Fig 4 and also available in the Nittany Scientific software *Multi-Plot* facility. Besides showing a pattern, the software may also make supplementary data available in various forms, as shown in the side-box in Fig 3. Of special interest is the *NEC4WIN* presentation in Fig 5; it adds a supplemental calculation of the headings for each lobe in the pattern.

Most modeling software provides a 3-D pattern for the user with a presentation that will vary slightly with each software package. Fig 6 shows the Nittany Scientific version, which allows a conventionalized but not scaled representation of the antenna in the typical line pattern. The antenna view and pattern change together as one rotates the image. Some

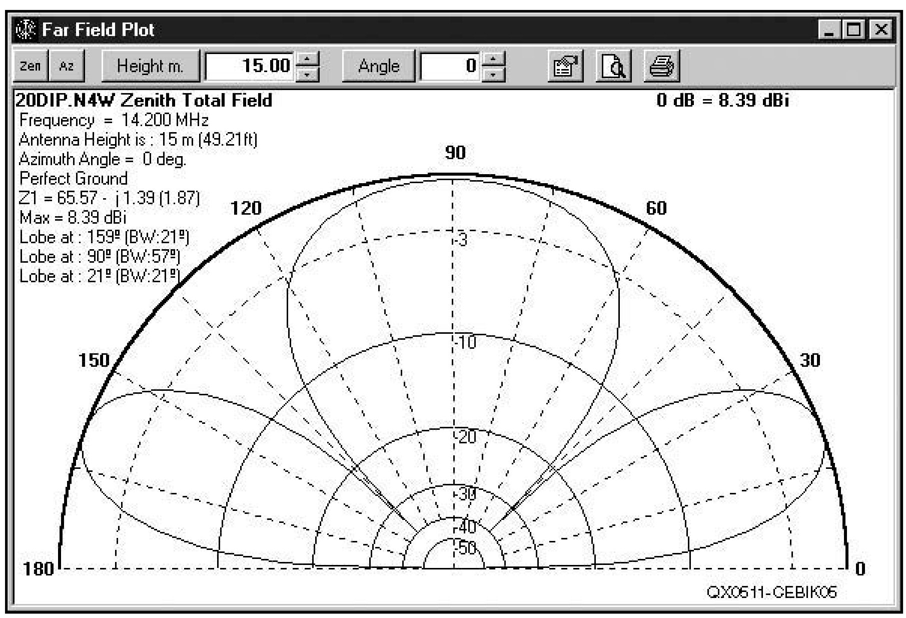


Fig 5—A *NEC4WIN* elevation pattern with lobe data.

packages use a standard separation of lines, representing increments (in degrees) between points in the far-field sample. Other programs allow user control of the increments, although one must always strike a balance between wide line spacing and little pattern definition on the one hand, and small increments for detail that may create an unreadable dark

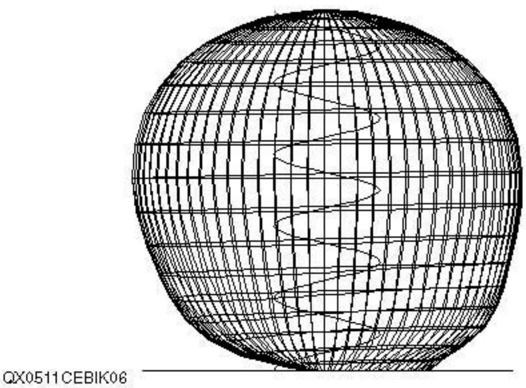


Fig 6—A typical 3-D far-field pattern (from Nittany Scientific), with a conventionalized (not to scale) representation of the antenna geometry.

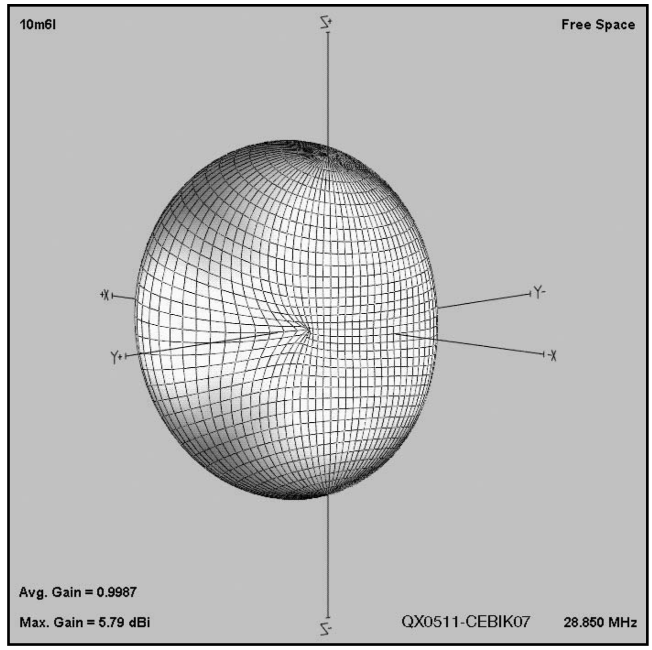


Fig 7—The colorized 3-D far-field pattern used in *Antenna Model*.

graphic. Antenna Model adds color to the pattern for easier reading.

Although virtually all antenna modeling software packages offer graphical far-field representations that show the same data, some presentations may be better suited to specific applications than others. Hence, the prospective user should consider the options in the appearance of pattern graphics as well as the number of different pattern types that may be available. Far-field patterns have major and minor axes, and the RPO command allows the user to specify these data in place of the more conventional vertical and horizontal components of the total field. We have also noted that one can calculate the circular components of a total field. Some programs, such as *EZNEC Pro*, may also allow the presentation of surface-wave (RP1) fields, while others may plot near-field (NE or NH) data.

In addition to polar plots, some software implementations provide an array of rectangular plots. At the low end of the scale, *EZNEC* provides an SWR plot across a specified frequency range, while *NEC-Win Plus* plots both SWR and the impedance components. *NEC2GO* provides a graph with multiple data lines for frequency sweeps. *Antenna Model* supplies separate gain, 180° front-to-back ratio, worst-case front-to-back ratio, resistance, reactance, and SWR curves for its frequency sweeps. At the upper end of the range of available rectangular

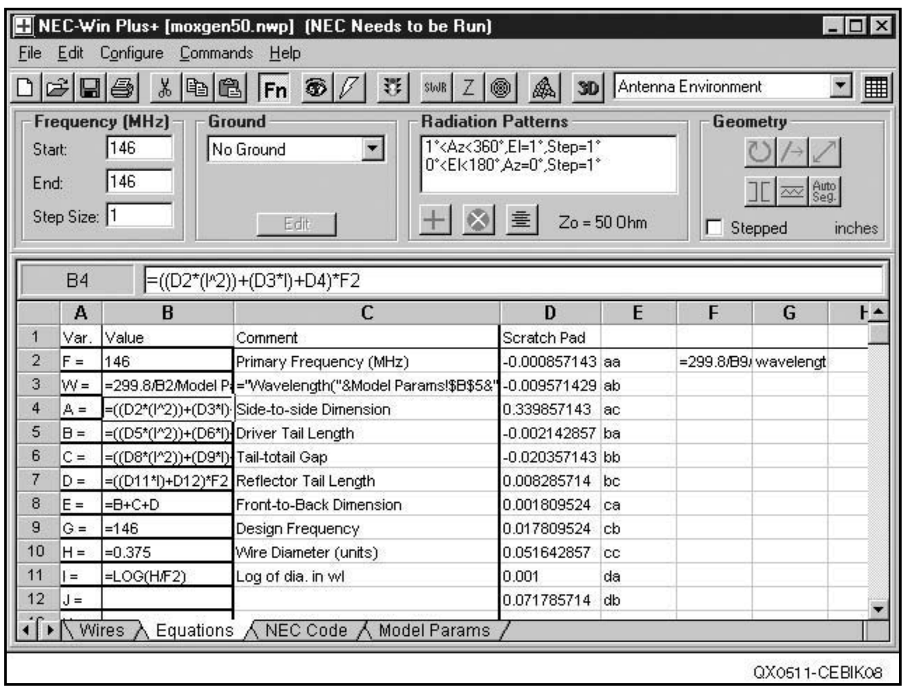


Fig 8—The "Equations" page of a *NEC-Win Plus* model of a Moxon rectangle.

plots are *NEC-Win Pro* and *GNEC*, which provide not only graphs for frequency sweeps, but also current and receiving data plots.

To increase the available graphical representations of data from entry-level programs, there are supplemental programs. For example, *EZPlots* from AC6LA provides extensive graphing of *EZNEC* frequency sweep data, which emerges from the core program in tabular form only. More flexible is AC6LA's *MultiNEC*, which works with a considerable number of *NEC* and *MININEC* programs to provide versatility on both the input and outside sides of the core. It allows batch runs and the use of variables on the input side, and produces a number of polar and rectangular plots for the output data. Both of these programs are *Excel* applications rather than standalone programs.

Auxiliary Functions

Any calculation made by an antenna-modeling program outside of the core is an auxiliary function. Since the SWR corresponding to a reported source impedance requires a user specification of a reference impedance, it is technically an auxiliary calculation. Likewise, calculation of the circular components of a total far field fits the same category. However, these calculations are so intimately connected to the core calculations that we tend to think of them as inherent to the modeling process.

A function that is clearly auxiliary is the "model-by-equation" facility or the use of variables, represented by the *NEC-Win Plus* "equations" page shown in Fig 8. (Part 1 showed a sample—Fig 3—of the *NEC-Win Plus*"wires" page revealing the use of variables rather than numbers.) In *NEC-Win Plus*, the equations are preserved only in the .NWP format file, but saving the model in .NEC format saves only the resulting numbers.

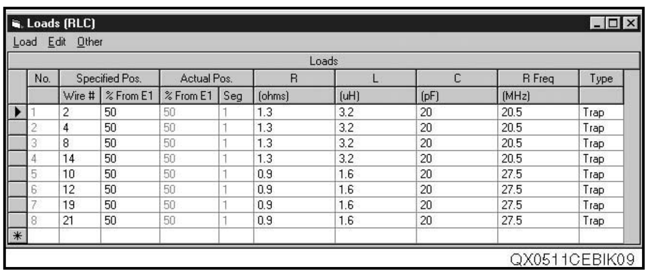
Some auxiliary functions are extensions of existing commands that the core already provides. Fig 9 shows the screen from *EZNEC* in which the user would specify loads using R-X, or series or parallel R-L-C values. Placing a trap in an antenna element normally requires external calculation of a parallel L and C value, including the inductor series resistance (calculated from the known or estimated coil Q). Hence a series-to-parallel conversion is in order. The trap is also designed to be resonant at a certain frequency. Hence, the net resistance and reactance at other frequencies varies in accord with how far off resonance the trap may be. Rather than calling for an external recalculation of the load assembly for each new frequency, *EZNEC* recalculates the load value for that frequency based on the input data provided by the user.

NEC2GO provides a different type of integrated auxiliary calculation. Fig 10 shows the matching network selection screen. The user selects the desired line impedance and the network type—only viable networks are active relative to the source impedance derived from an initial core run. There are other matching network programs available, but the unique part of the *NEC2GO* system is that it converts the selected network and values into an NT command and adds it to the model, moving the source to a new wire created by the auxiliary function.

The facility of some programs to create various wire structures within the program has been noted in passing. Radial systems (whether used for a ground plane or as a "top hat") and wire-grid rectangles are the most common, although in *EZNEC*, one can also create helices and circles (or, more correctly, straight-wire approximations of circles). For more complex structures, Nittany Scientific offers a package called *NEC-Win Synth* to synthesize the geometry section of models. The output can be saved in both .NWP and in .NEC formats for use in a wide variety of *NEC* software. The package includes over 50 preset shapes ranging from simple geometric shapes, like an open or closed cylinder, to generic vehicle shapes, such as a sedan, van, or pickup truck.

True auxiliary functions are integrated into the software package that implements a given calculation core. Numerous programs have independent modules to perform various calculations, but often, they are conveniences more than integrated parts of the modeling process. For most of the independent calculation modules, one can find external programs to perform the same calculations. Hence, these functions may have lesser status than integrated auxiliary functions when evaluating one's options in the selection of antenna-modeling software.

Our samples of auxiliary functions, like every other aspect of this survey of options, incompletely represent what is available in the commercial implementations of *NEC* and *MININEC*. The goal is to make you aware of the range of enhancements that may be available to the antenna modeler. In deciding what software is best for both immediate and future needs, you will have to decide which features of which soft-



Loads (RLC)

Load Edit Other

Loads

No.	Specified Pos. Wire #	Specified Pos. % From E1	Actual Pos. % From E1	Actual Pos. Seg	R (ohms)	L (uH)	C (pF)	R Freq (MHz)	Type
1	2	50	50	1	1.3	3.2	20	20.5	Trap
2	4	50	50	1	1.3	3.2	20	20.5	Trap
3	8	50	50	1	1.3	3.2	20	20.5	Trap
4	14	50	50	1	1.3	3.2	20	20.5	Trap
5	10	50	50	1	0.9	1.6	20	27.5	Trap
6	12	50	50	1	0.9	1.6	20	27.5	Trap
7	19	50	50	1	0.9	1.6	20	27.5	Trap
8	21	50	50	1	0.9	1.6	20	27.5	Trap

QX0511CEBIK09

Fig 9—The load entry screen in *EZNEC*, showing an example of traps.

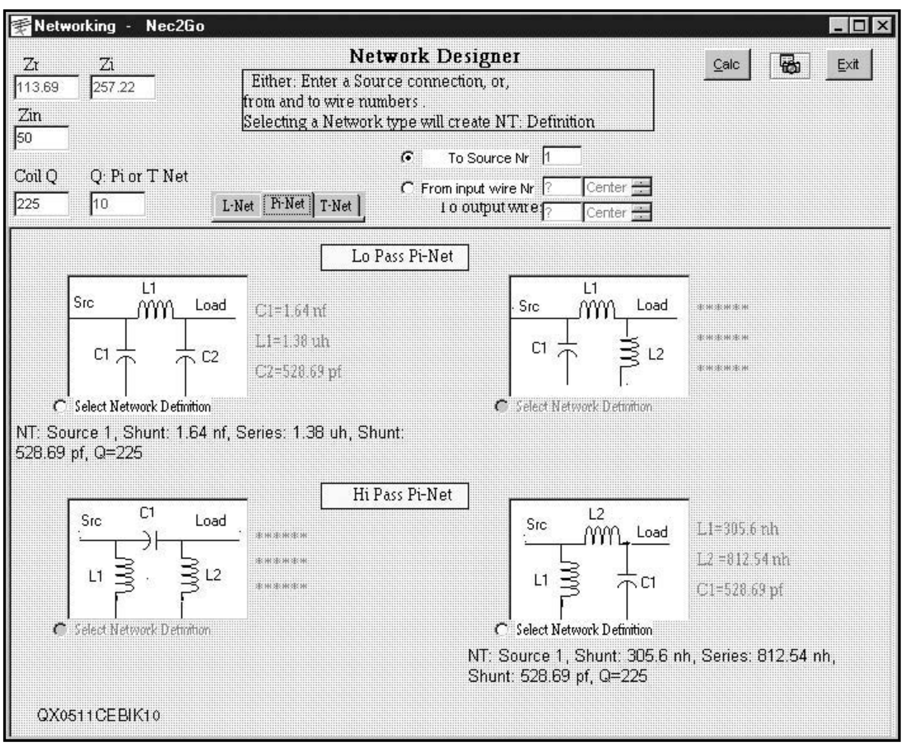


Fig 10—The matching network page from *NEC2GO*.

ware best mesh with those needs.

Cost

I would be remiss if I did not take cost into account as one of the option sets facing the antenna modeler. Software is an inherently unstable commodity, with both the available features and the cost changing over time. As operating systems evolve, changes in software are inevitable. The PC software that we have sampled is generally compatible with operating systems up through *Windows XP*. What the near future may bring, such as 64-bit CPUs, etc, and the effects that future developments may have on antenna modeling software appear in a crystal ball that I do not have. Therefore, these notes are necessarily dated and subject to change.

In thinking about cost, the prospective user must consider the program capabilities that come with the price and the support available from the program developer. Typically, but not always, free programs come without any commitment to support. Support has at least two dimensions: bug fixes and tutorial material. All of the developers noted in this exploration of options want to know about bugs and are interested in user suggestions. However, suggestions for change and enhancement are always subject to programming feasibility, and a given suggestion might not show up in a software package until many versions down the line—if at all.

Table 2—Approximate Software Costs for the Programs Mentioned

All values rounded, and costs are in US dollars. Costs are for single copies and do not include any applicable shipping and/or handling charges. In some cases, there may be multiple-copy discounts and site licensing. The list includes only programs available for under $1000.

MININEC	
MMANA	free
NEC4WIN95	$ 60
NEC4WIN95VM	$ 90
Antenna Model	$ 85
Expert MININEC Professional	$390
Expert MININEC Broadcast Professional	$790
NEC-2	
4NEC2	free
NEC2GO	$ 40
EZNEC (standard)	$ 90
EZNEC Plus	$150
EZNEC Pro/M	$450
NEC-Win Plus	$150
NEC-Win Pro	$425
NEC-4 (plus cost of license)	
EZNEC Pro/4	$600
GNEC	$795
Adjunct Programs	
EZPlots	$ 20
Multi-NEC	$ 40
NEC-Win Synth	$100

Tutorial materials take many forms. Many packages have attached tutorials to guide the user through at least the initial stages of modeling with the software. Manuals come in two forms: via the "help" facility and/or in a printed manual. In some cases, the program developer may provide or update a manual with documents that require user printing. Articles appear from time to time in various journals, such as *QST*. My own series on "Antenna Modeling" appears monthly in *antenneX* and at my Web site, **www.cebik.com/amod/modeling.html**, and covers topics relevant to the use of both *NEC* (*-2* and *-4*) and *MININEC*. Of all the cores, *NEC-2* enjoys the most tutorial support. The ARRL *Antenna Modeling* correspondence course is a 31-lesson introduction to antenna modeling with exercise models provided. Nittany Scientific offers both basic and intermediate modeling tutorials, including exercise models.

Nevertheless, the bottom line is dollars for digits. Table 2 provides the software prices as derived from on-line sources at the time of writing. Remember that *NEC-4* requires a prior license from LLNL as well as the implementing software, so include all costs when counting your resources. The numbers are for general guidance and are rounded. Be certain to get exact prices, including shipping and handling, from vendors before making your investment in antenna modeling software.

Whether you are considering an entry-level program or an advanced version of the latest software, these notes are aimed to expand the range of factors that you should consider as you approach a decision. Not all of the available options would fit the space available for these notes, so expect to do much more detailed analyses in preparing to write a check. □□

2005 Index

Issues 228 through 233
(All dates are 2005 unless otherwise indicated.)

Out of the Box (Mack)

RF (Lau)

Sidebars

Tech Notes

Technical Articles

Upcoming Conferences

Letters to the Editor

Measuring Cable Loss (May/Jun 2005)

Frank,

I have read your cable-loss article with interest. I note that your preferred measurement parameter using the MFJ-259B is reflection coefficient, and that you make the point that the two-digit display limits the accuracy of your method of cable-loss measurement.

When you made those measurements with the 259B, you would also have had results for resistance and reactance. If you were to put those values into the remote-impedance spreadsheet that I described in the Jul/Aug 2004 issue of *QEX*, it would give you a value for loss as well as Z_o and effective electrical length. I would be interested to know how the loss results would compare with those derived from the reflection-coefficient reading.

73, *Ron Barker, G4JNH;* **g4jnh@onetel.com**

Hi Ron,

Thanks for your letter. The MFJ-259B instrument has a resistance bridge in it with a reference resistance of 50 + *j*0 ohms. Hence, I believe it directly measures reflection coefficient magnitude by computing the ratio of some voltages. The evidence I have for this conclusion is that I reverse-engineered the MFJ-259, which preceded the MFJ-259B. The SWR and return loss are calculated from the reflection coefficient results, but the algorithm used by MFJ leaves a lot to be desired. I don't know how MFJ obtains *R* and *X* values, but I would like to know.

If the accuracy of *R* and *X* measurements is comparable to that of reflection coefficient magnitude, the resolution might be better if *R* and *X* are used to compute reflection coefficient magnitude (with a spreadsheet). I don't know, however.

It's good to see that others such as yourself are interested in milking as much valid data out of instruments that are available to us.

73, *Frank Witt, AI1H;* **ai1h@aol.com**

Frank,

Thanks for your reply. I suspect that all of the antenna analyzers use a resistance bridge of the type shown in the 17th edition of the *ARRL Antenna Book* on p 27-5 (Fig 5).

Resistance and reactance (but not the sign of the reactance) can be derived from the magnitude of reflection coefficient and the magnitude of impedance as AC6LA showed in the *ARRL Antenna Compendium, Volume 7* (p 44, Eqs 10 and 11). I note that you have an article in that volume, so I guess that you must have a copy. Dan didn't show the derivation of Eq 10; it's an interesting bit of algebra that I enjoyed working out. I could send it to you if you so wish. I think that Dan's review of antenna analyzers in that article was one of the best articles that I've seen in a long time.

Going back to the resistance bridge as shown in the Fig 5 that I referred to above, the magnitude of the reflection coefficient is merely V_{ref}/V_{fwd}. The magnitude of impedance is the ratio of the voltage at J2 to the voltage across R_s times R_s, so only one further voltage measurement is required: that across J2. The voltage across R_s is then $2V_{fwd}$ –V at J2. It is my guess that the 259B measures these voltages and has an algorithm to compute the results that it displays.

73, *Ron, G4JNH*

Receiver Measurements: How to Evaluate Receivers (Jul/Aug 2005)

Doug,

Refer to p 9, in the left-hand column, and the calculation of IIP3 (the third-order input intercept point) and OIP3 (the third-order output intercept point). I was taught that the difference between them is the gain of the device (loss being negative gain); that is, OIP3 = IIP3 + G.

In the first example, that of Fig 15, OIP3 is calculated correctly at 25 dBm. Thus, IIP3 would be 25 dBm + 8 dB = 33 dBm. If the –74 dBm output spurious frequency level is transferred to the input by +8 dB, just as are the main signals, this equals –66 dBm. Using this value in the equation yields 33 dBm.

The second example is the same way. The OIP3 is calculated correctly at 42 dBm. The IIP3 would be 42 dBm –8 dB = 34 dBm. Transferring the output spurious frequency signal level of –60 dBm to the input gives –68 dBm. Using this value in the equation for IIP3 gives 34 dBm.

Regards, *Larry Joy, WN8P;* **ljoy@kantronics.com**

Hi Larry,

You're right. Ulrich sent us the following corrections to the equations you cited:

"From Fig 15 we can see that the input intercept point ($IP_{3\,(IN)}$) can be calculated as follows:

$$IP_{3(IN)} = \frac{3\times 0-(-74+8)}{2} = 33\ \text{dBm}$$

Similarly, the output intercept point is calculated to be:

$$IP_{3(Out)} = \frac{3\times(-8)-(-74)}{2} = 25\ \text{dBm}$$

If the assumed mixer now is active instead of the passive mixer (+8 dB loss) and has 8 dB gain, and the IMD products are at –60 dBm, we now calculate:

$$IP_{3(IN)} = \frac{3\times 0-(-60-8)}{2} = 34\ \text{dBm}$$

and

$$IP_{3(Out)} = \frac{3\times 8-(-60)}{2} = \frac{24+60}{2} = 42\ \text{dBm}$$

73, *Doug Smith, KF6DX,* QEX *Editor;* **kf6dx@arrl.org**

Hi Doug,

N1UL's article is a little out of date in its references to CEPT and FTZ standards. In Europe, national standards were progressively withdrawn after the formation of ETSI, the European Telecommunications Standards Institute (**www.etsi.org**) in 1987. There are now quite a number of standards and methods of measurement applicable to radios, including those for commercially available Amateur Radio equipment, available from ETSI. In general, there are various groupings: Land Mobile Radio, which include standards for analogue and digital, both with external and integral antennas; Short-Range Devices; Fixed Links; Maritime (HF and VHF); Fixed Wireless Access (FWA); various WLAN standards; and so forth. Most of those have measurement methods applicable to today's needs. All the old CEPT standards had gone by some 15 years ago, and even the FTZ no longer exists as such. It had become the RegTP until about June this year, but has metamorphosed yet again!

It should be noted that in European testing, the high-level blocking requirements for receivers can quite easily be subsumed by the EMC stress levels, although the applicability is dependent upon which EMC immunity classification is used.

One important factor that N1UL did not touch upon is that of measurement uncertainty, which for some measurements, especially intermodulation, is potentially rather large. The ETSI standard for a certified lab is the measurement of the applied power to ±0.75 dB for a given measurement confidence level (95%), and that uncertainty is one upon which all the major measurement institutions and manufacturers agreed. Once we do an IMD test where we have two sources of uncertainty of 0.75 dB, the result is, of course, the root of the sum of the squares, which is about ±1.0 dB. Throw in the uncertainty in the coupler loss, uncertainty in its mismatch loss, the losses and mismatch losses of any attenuator and the uncertainty in the mismatch loss of the device under test, and you can see that under "Type Approval" lab conditions, the measurement uncertainty in a 3rd-order IMD measurement is unlikely to be less than 3 dB, and 6 dB is much more likely. The real test of this is to take an equipment round to a number of laboratories—it was from such a series of measurements that ETSI have a measurement uncertainty on radiated power measurement of ±6 dB! This business of uncertainties in IMD measurements was, of course, discussed in *QEX* a year or two ago.

A very useful document on the subject is ETR028, from ETSI; but it's not light bedtime reading! Downloads from ETSI are free, although you have to register. Unfortunately, few of the test levels used in ETSI standards are very applicable to the highest-performance amateur receivers today. A further complication is that the EU has decided in their wisdom that to stimulate (!) competition, only essential receiver parameters will be included in future standards. Essential parameters are those such as receiver radiation or, where a malfunctioning receiver would cause problems to a network, those parameters required for the network to continue to function correctly. This makes spectrum engineering very difficult. So ETSI standards now have a Part 1 and a Part 2. Part 2 includes those parameters to which conformance leads to a presupposition that the equipment meets the essential parameters of the Radio & Telecommunications Terminals Equipment Directive—a somewhat oversimplified explanation, but now you can understand why I'm looking forward to retiring in 21 months at age 60!

73, *Peter Chadwick, G3RZP;* **g3rzp@g3rzp.wanadoo.co.uk**

Hi Doug,

In my paper, I stressed a variety of important measurements, regardless of their vintage and regardless of whether the institutions which proposed this do still exist. It would have been inappropriate to associate these measurements with more modern standards. As to the measurement uncertainties, one certainly has to use calibrated signal generators, hybrid couplers and others. The real life is that we compare intercept points of 10, 20, 30, or even 40 dBm and a few dB make no practical difference. It is more important that all measurements on different devices under test are done with the identical test set up and there may be an absolute error, but the relative error is small. This allows comparison.

I think Peter Chadwick is trying to be more catholic than the pope. The EMC stress level is unrelated to laboratory measurements and has no business here. There are many more things one can measure, but the ones I have outlined have the top priority.

Another reader, from Italy, commented on the input and output intercept point of mixers. He did not realize that the examples taken, on purpose, were not the same. Having the same test condition, the difference between input and output IP3 is the loss or gain of the stage.

Best regards, *Ulrich L. Rohde, N1UL;* **ulr@synergymwave.com**

Hi Doug,

I don't understand, "It would have been inappropriate to associate these measurements with more modern standards." Why? And, of course, the measurement techniques described are wholly inappropriate for use with systems using digital modulation. But it does show that there is more way than one of skinning a cat!

As far as EMC tests are concerned, I'm afraid that commercially available equipment in Europe has certain requirements; and since those were based on IEC determinations of average RF levels, it would seem that applying them has some advantages. Of course, the old-fashioned radio with its tuned preselector wins out here.

73, *Peter, G3RZP*

Gentlemen,

Time to drag out that old saw: The nice thing about standards is that there are so many from which to choose. We'll have much more debate on this in future issues.*—Doug, KF6DX*

***Octave* for Signal Analysis (Jul/Aug 2005)**

Hi Doug,

Following are responses to the thoughtful comments of Ray Mack, WD5IFS, on my article.

The function abs() in *Octave* differs from the function abs() in C, but is essentially identical in mathematical functionality to abs() in C++ and FORTRAN. The function fft() in *Octave* returns magnitude and phase information for the coefficients of a Fourier series and not power density spectral values. *Octave* is compatible with other implementations of fft() in that regard. A more detailed discussion of each of these issues follows.

When one moves from one programming language to another, it is important to pay attention to the differences between the syntaxes of the two languages. Variations such as sqr() versus sqrt() and the difference in row and column order between FORTRAN and C matrices have plagued folks for years and are always worthy of a note of caution.

The function abs() in *Octave* is certainly different from abs() in the C language. In versions of C without the complex data type, abs() accepts only integer arguments and the function fabs() operates on floats and doubles. The function abs() in FORTRAN and C++, however, behaves very much as does abs() in *Octave*. In FORTRAN 77 and FORTRAN 90, abs() accepts REAL, DOUBLE, INTEGER, and COMPLEX arguments. In C++, abs() is overloaded as double abs(const complex) to accept complex arguments.

In both FORTRAN and C++, abs() returns the magnitude of the polar representation of a complex argument and in each case, it does it by computing the square root of the sum of the squares of the real and imaginary components just as does *Octave*. FORTRAN and C++ are thus closer to *Octave* in this regard than they are to C.

Since *Octave* code can look much like C code, it is tempting to compare or contrast features of *Octave* with comparable features of C. I think, though, that it might be more useful to compare the features of any function in a mathematical language like *Octave* with the same functionality provided by other math utilities like *Mathcad*, *Mathematica*, *Maple* and *Matlab*.

Note that *Octave* is a language oriented toward vector and matrix calculations. abs() thus operates on all the elements of a vector if presented with a vector argument, something that would require a for() loop in many languages. When presented a single scalar argument, integer, real, or complex, *Octave*'s abs() behaves exactly as does abs() in FORTRAN or C++.

The function fft() does produce a vector of complex numbers which represent the coefficients of a Fourier series for the time domain sequence that is the argument of fft(). The coefficients of the Fourier transform of a periodic sequence of real numbers are always represented by a sequence of complex numbers and the Fast Fourier Transform is no exception. At each frequency, the complex value encodes both amplitude and phase information. You can see how that works by noting, as did Ray, that using the cos() function in the article, the amplitude is represented predominantly by the imaginary component of the vector element. If you change the time domain function to sin() in the *Octave* code, the magnitudes of the real and imaginary components of the corresponding fft() output will be interchanged.

Note that the values output by fft() are complex amplitude values containing both magnitude and phase information and are not power density values. You can confirm that by doubling the amplitude of the time domain function and watching the amplitudes of the frequency domain representation double, rather than increasing by a factor of four as would be the case for samples of the signal's power density spectrum.

Other implementations of fft() and ifft() are compatible with *Octave* but, being a little cautious, I did some testing of *Octave* against a couple of other programs when I first obtained *Octave*. As is the case with text and reference books, various implementations differ as to where the constant coefficients that are used as multipliers of fft() and ifft() are placed. Each implementation can be made equivalent to another by simply adjusting the constants. I chose constants that would transform amplitudes (voltage or current) in the time domain into peak amplitudes of spectral components, the scheme usually used in Fourier transforms for communications work. With proper adjustments of the constants, *Octave* yielded the same results that I got from a C++ fft() implementation and from *Mathcad* 2001.

Phase information is often useful, but I omitted it in my plots since it doesn't matter for the purposes of the article. I plotted the magnitude of each spectral component, which can be obtained by taking the square root of the sum of the squares of the real and imaginary components, using abs().

Although both components are squared in the process, taking the square root of their sum prevents them from representing a function of higher-order than amplitude.

If I need to obtain the phase angle of a spectral component, I use arg() or angle() to calculate it. Together, abs() and arg() (or angle()) provide a conversion to polar coordinates of the component represented in rectangular coordinates by the output of fft(). A family of such functions, often real(), imag(), abs(), and arg(), are generally provided by math languages or tools for converting values from one system of coordinates to the other.

Where, in the article, I invited folks to experiment with transforming s_cxr back to the time domain using ifft(), I believe that I should have emphasized that it won't work to use ifft() on trunq, the vector that is truncated for plotting. That vector is missing the entire negative half of the spectrum along with much of the positive half when it matters, and will not yield the original time domain signal if it is passed through ifft().

73, *Maynard Wright, W6PAP;* **w6pap@arrl.net**

Dear Doug,

Wow! A very nicely written explanation of the subtleties of this aspect of *Octave*. It is some very useful additional information. The only change is that cos(x) will give real values in the FFT and sin(x) gives the imaginary ones.

I fell into the trap Maynard mentioned because I am neither a C++ or FORTRAN programmer. It is good to have the perspective of where *Octave* has its roots.

73, *Ray Mack, WD5IFS,* QEX *Contributing Editor;* **wd5ifs@arrl.org**

A Better Antenna-Tuner Balun (Sep/Oct 2005)

Hi Andrew,

Congratulations on your fine article. Your analytical approach to this problem and the results are impressive. Certainly, many will take advantage of your contribution.

I propose that you carry your experimental evaluation one step further by measuring the loss and imbalance using the techniques I described in the *QEX* article you referenced in yours, and in "Baluns in the Real (and Complex) World," *ARRL Antenna Compendium*, *Volume 5*, (pp 171-181). This technique employs a geometric resistance box, a network analyzer like the MFJ-259B and a low-loss antenna tuner. Your results would be a good subject for another *QEX* article.

You could use the technique to compare the voltage balun and current balun you describe in your article with the hybrid balun made up of those two components. If I can be of any assistance or if you have any questions, please let me know.

73, *Frank Witt, AI1H;* **ai1H@aol.com**

Hi Frank,

Many thanks for your e-mails and comments. I feel truly honored to be complimented by someone of your stature. I think your suggestion of applying your measurement techniques to measure the loss and imbalance of the voltage and current of hybrid baluns is excellent and I will get working on it without delay. In fact, I've already "pre-spent" my author's remuneration for the article on a Palstar ZM-30, which should be ideal for this. With your permission, I would like to send you a draft of the article for review before submitting it to *QEX*.

73, *Andrew Roos, ZS1AN;* **andrew@exinet.co.za**

Help for Oscillator Failure in the HP-8640B (Sep/Oct 2005)

Hi Doug,

I'm a real fan of the HP-8640B, so the article on cavity oscillator repair was good data. I have a couple working units and two hanger queens for parts. They are actually quite easy to repair. The transition to hanger queen happens when the plastic gear train breaks down.

Enclosed is a list of HP part number conversions to industry-standard (JEDEC) parts. Feel free to share it with the masses.

Frank Carcia, WA1GFZ; **francis.carcia@hs.utc.com**

Hi Frank,

And thanks for that cross-referenced list. We've posted it on our Web site at **www.arrl.org/qexfiles/300-hpxref.pdf**.*—Doug, KF6DX*

Al Williams, VE6AXW, SK

Author Allen Ross "Al" Williams, VE6AXW, of Edmonton, Alberta, Canada passed away in the early morning hours of July 8, 2005, from complications of cancer. He is sorely missed by his wife and children, and by us.

Al wrote the innovative article, "A Regulated 2400-V Power Supply," *QEX*, Jul/Aug 1999. He also wrote the book, *Bush and Arctic Pilot*, Hancock House Publishing, 1998, ISBN: 0-88839-433-0. Al was an accomplished pilot and technologist. □□

Empirical Outlook

(continued from page 2)

The published PSCM was last updated in 1996; the ARES FRM, in 1995. We think it's time for revisions. Neither manual mentions the Salvation Army Team Emergency Radio Network (SATERN), for example. ARES is supposed to issue credentials honored by local authorities so we can fulfill our primary fundamental purpose as a voluntary, licensed communications service. However, we're far from getting the recognition and respect that the Red Cross and Salvation Army do.

We should leave the command and control of localized disaster responses to the locals as much as possible. Put them in charge and give them what they need; otherwise, get out of the way. If we cannot depend on our federal government to provide funding, personnel and materiel promptly as needs arise, then let's either get under the hood and fix the problem or simply cut the bureaucrats out of the picture. It's easier for us to establish and maintain relationships with our own community's officials than with those at the national level. Let the mayors coordinate with the governors and the governors, with the president.

As to preparedness from an engineering perspective, Congress and the president clearly need to get their acts together. 2004 was a bad hurricane season but 2005 flood-control funding for New Orleans, LA was cut by the steepest level in history. (Source: Congressional Budget Office.) Diversion of funds to Iraq was cited as a reason. Our guess is that in future, you'll hear less of the phrase "saving lives here at home" regarding overseas actions.

Since New Orleans is in a "bowl" below sea level, water must constantly be pumped out of the city, even with the levee system. A lot of city land consists of wetlands that were drained and filled with sand long ago. That's one reason that the city—and its levees—are slowly sinking.

The city's pumping system has the capacity to remove about 1200 m^3/s of water. The city's land area is about 470 km^2. (Source: US Geological Survey.) The other half of the area within the city limits is normally underwater. Peak precipitation events are predicted every 10 years or so that are expected to dump about 20 cm of rain on the city in a 12-hour span. (Source: NOAA.) That is 2200 m^3/s of water or almost double the pumping capacity. A few inches of standing water wouldn't destroy the city but a category-4 or -5 hurricane produces worse effects: a surge from the Gulf, elevated levels in the Mississippi River and so forth. It's been well-known that any significant breach or overflow of the levee system would knock out electricity and submerge the pumps, rendering them useless. Engineers knew what to do to prevent that and how to do it but they didn't have the money to proceed.

UC Berkeley Professor of Civil Engineering Robert Bea has studied the situation for well over 40 years. His own home was destroyed under four meters of water during Hurricane Betsy in 1965. The same levee system breached by Katrina failed then, too. We wonder how anyone could declare the system sufficient for a 200-year span, as Lt Gen Carl Strock, Commander of the US Army Corps of Engineers, did. (Source: Reuters.) Today's challenges far outweigh anything from 1965.

Bea began his career with the Corps of Engineers in 1954. His conclusions remain unchanged since that era: The levees are insufficient as designed to protect the city and the current system is untenable. (Source: UC Berkeley Graduate School of Journalism.) We can hardly disagree with what has been—and still is—so obvious.

The new federalism seems to have bitten off more than it can chew. We believe a serious and thorough recasting of emergency-preparedness roles is in order. Rebuild the hierarchy from the bottom up and let the makers of such outstanding organizations as the Internal Revenue Service, the US Postal Service, FEMA and so forth do what they do best: shovel money. But let's get them to shovel it in the right directions.—*73, Doug Smith, KF6DX,* **kf6dx@arrl.org**. □□

Upcoming Conferences

20th Annual Southwest Ohio Digital Symposium, 14 January 2006

The 20th Annual Southwest Ohio Digital Symposium will be held Saturday 14 January 2006, Registration 0800, technical sessions 0900-1600 EST.

Location is the Middletown Campus, Miami University, 4200 N University Blvd, Middletown, Ohio. This is the second longest continuously running digital symposium in the United States, surpassed only by the ARRL/TAPR Digital Communications Conference.

We will discuss all types of digital modes, including 802.11b 2.4 GHz communications, use of digital modes in recent Hurricane relief efforts, experimentation with APRS, PSK-31, MFSK, etc.

We typically accept papers until a week or so before the event, so if you have a pet mode or new experimental results of an existing mode, please contact Hank Greeb, **n8xx@arrl.org**. See **www. swohdigi.org/** for more details.

This symposium is free—no admission charge. Optional catered lunch at cost.

Talk-In Freq: 146.61, 224.96 and 444.825 MHz with standard offsets. □□

In the next issue of QEX/Communications Quarterly

In our next issue, Randy Evans, KJ6PO, presents "*PSpice* for the Masses." It's a thorough look at what is perhaps the most popular circuit design and analysis program, in the form of *LTSpice* from Linear Technologies.

Check this out because the price is right: It's available for download at no cost. □□

QEX Subscription Order Card

ARRL
225 Main Street
Newington, CT 06111-1494 USA

QEX, the Forum for Communications Experimenters is available at the rates shown at left. Maximum term is 6 issues, and because of the uncertainty of postal rates, prices are subject to change without notice.

For one year (6 bi-monthly issues) of **QEX**:

In the US
- ☐ ARRL Member $24.00
- ☐ Non-Member $36.00

In the US by First Class mail
- ☐ ARRL Member $37.00
- ☐ Non-Member $49.00

Elsewhere by Surface Mail
(4-8 week delivery)
- ☐ ARRL Member $31.00
- ☐ Non-Member $43.00

Canada by Airmail
- ☐ ARRL Member $40.00
- ☐ Non-Member $52.00

Elsewhere by Airmail
- ☐ ARRL Member $59.00
- ☐ Non-Member $71.00

Subscribe toll-free with your credit card **1-888-277-5289**

☐ Renewal ☐ New Subscription

Name________________________ Call ____________

Address________________________

City________________ State or Province ________ Postal Code ________

☐ Payment Enclosed to ARRL

Charge:

☐ MasterCard ☐ VISA ☐ American Express ☐ Discover

Account #________________________ Good thru________

Signature________________________ Date________

Remittance must be in US funds and checks must be drawn on a bank in the US.
Prices subject to change without notice.

06/01