COMMUNICATIONS QUARTERLY

THE JOURNAL OF COMMUNICATIONS TECHNOLOGY

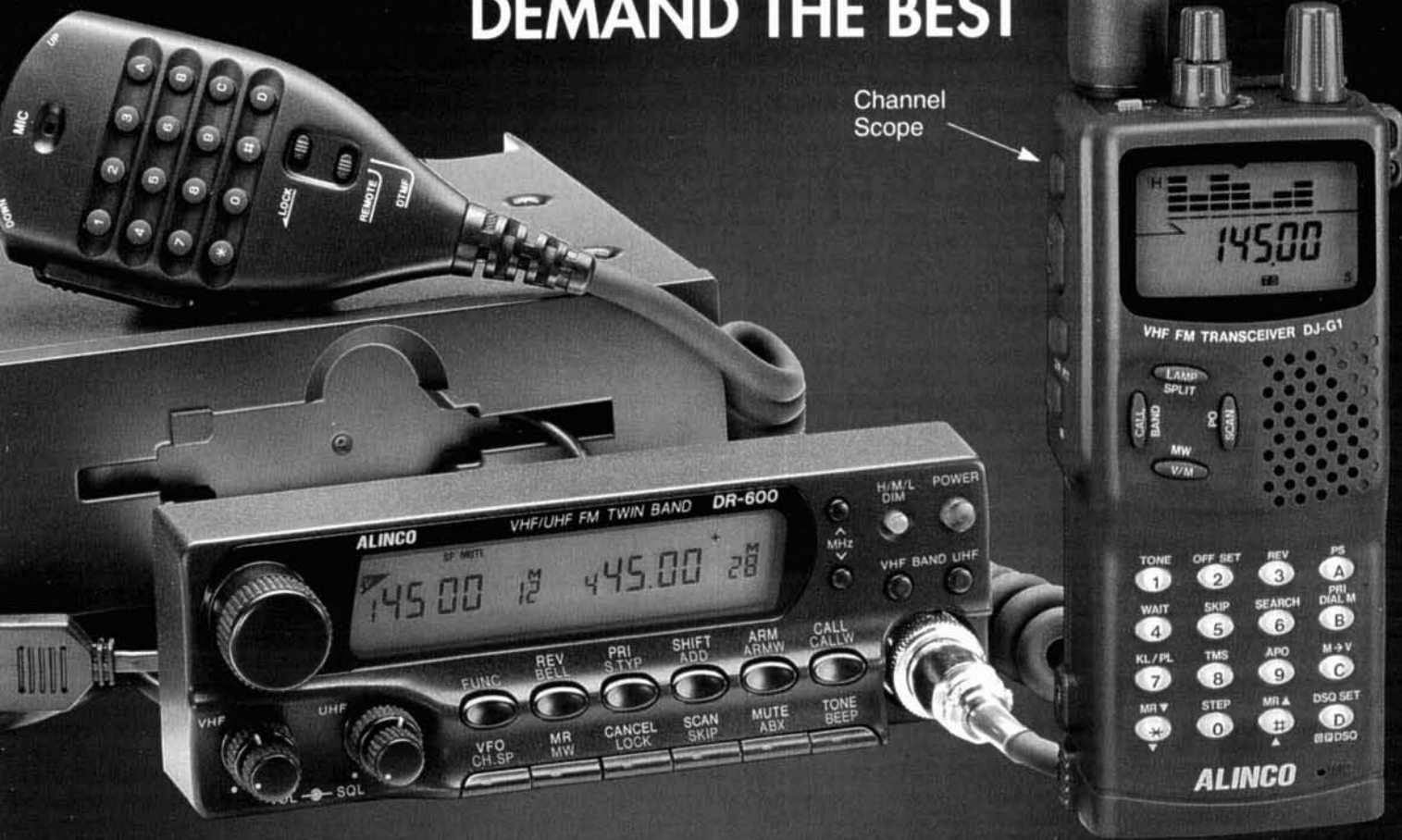Spring 1994

$9.95

COMMUNICATIONS QUARTERLY
THE JOURNAL OF COMMUNICATIONS TECHNOLOGY

CONTENTS

Fry, page 13

Littlefield, page 94

Cover Photo: In this issue, Bryan Bergeron, NU1N, explores Surface Acoustic Wave (SAW) technology. The cover illustrates facets of this subject. In the upper left, we see a SAW schematic, the grid represents the semiconductor wafer before etching, and the rectangular object in the foreground is a SAW filter/delay showing transducers and wave propagation. Photo by NU1N.

___EDITORIAL___

From The Future To The Past, And Back Again.

The operator looked out the window at the stark rose-colored horizon. He returned his attention to the radio in front of him and spoke into the microphone . . .

"KA1STC, KA1STC, KA1STC, this is K1ZJH calling from Mars Station Alpha. QSL?

On Earth, KA1STC crouched close to the speaker of her transceiver. She heard a loud crash of static, then out of the ether came the faint, but audible, call . . .

"KA1STC, KA1STC, KA1STC, K1ZJH."

She eagerly replied:

"K1ZJH, KA1STC, you are 5-by-2 on Earth."

After several minutes delay came the reply:

"5-by-2 Mars, Roger?"

KA1STC sent her final message into space, "Roger, 73," and awaited a final call from K1ZJH. "73," came the faint reply.

With that final 73, K1ZJH and KA1STC could lay claim to the first two-way QSO between Mars and Earth, and a special award— "The Planet Mars QSO Cup."

In fact, The Planet Mars QSO Cup is a real award that awaits the first two amateurs to complete a two-way amateur radio contact between Earth and the Red Planet. The Mars QSO Cup was the brainchild of two hams— Fred Johnson Elser, K3AA, and Stanley M. Mathes, K1CY—who created the award in 1928. The trophy is on display at The American Radio Relay League Headquarters in Newington, Connecticut, waiting to be claimed by the winners.

During the same decade that Elser and Mathes were designing their "heavenly" award, a group of prominent scientists were devoting their time to the design, construction, and operation of arc and spark transmitters. Many of you have probably heard of these types of transmitters; articles about them have appeared in the *AWA Journal* and *QST*. But perhaps you are not aware that, while arcs and sparks were developed within five years of each other and their fundamental circuits were almost identical, their operational characteristics were quite different. Bill Byron, W7DHD, has performed a detailed study of these transmitters, and presents his findings in this issue of *Communications Quarterly*. He begins his discussion by comparing arcs and sparks, and continues with an in-depth look at spark transmitters. A more thorough treatment of arc transmitters will appear in a later issue. To bring his work into the present, Bill uses a Fourier

Analysis program written in BASIC to calculate the frequency spectra of spark transmitters.

While the editorial staff of *Communications Quarterly* tends to concentrate on publishing information about current state-of-the-art technologies, we like to include historical material with present-day applications periodically. It is interesting to note that a lot of the technologies we think of as recent developments in the communications field, have roots reaching back to the earliest days of radio. A case in point: You may assume that wideband RF transformers were developed to meet the needs of broadband matching circuits for semiconductor-based circuits. Yet as early as 1924, companies such as Acme Aparatus Company were marketing wideband untuned transformers covering the 200- to 500-meter range. Intended for broadband interstage RF amplifier stages in reflex radio designs, today's broadband baluns and interstage transformers.

But let's return to Mars, and our intrepid radio operator. While it may seem far-fetched to even consider two-way communications with Mars, who's to say that it won't be a part of our future? After all, how many of us thought we'd be exchanging QSOs with space stations and shuttles orbiting the Earth? It's quite possible, that sometime in the future (granted it probably won't be in our lifetimes), men will visit Mars. Already, NASA and the European Space Agency (ESA) have plans to launch missions to the Red Planet at the end of this decade. These unmanned missions, the Mars Environmental Survey (MESUR) and MARSNET, would involve orbiting satellite relays gathering data from Mars-based rovers and fixed land stations for retransmission back to Earth. In their article, "HF Radio On Mars," Craig Fry, WL7C, and Robert Yowell, KC5BRG, discuss these missions, and put forth their own theory that HF radio would be a viable communications mode for future Mars explorers.

In all scientific endeavors, a knowledge of the past and the present intertwine allowing projections for future technologies. The dreams of the past are often the realities of the future. We must always bear in mind that what is past is prologue.

Terry Littlefield, KA1STC
Editor
Peter Bertini, K1ZJH
Senior Technical Editor

___TECH NOTES___

Of lawnmowers, sail boats, and radio equipment

We lead off this issue's Tech Notes with a project by Ron Mather, ZL2AXO: "A 230-Volt Generator from Scrap." Ron's generator, with its 230-volt, 50-cycle output lends a "down-under" air to this note. More importantly, the main components for this project are almost universally found taking up valuable space in a diminishing resource—our landfills. Our American readers will, of course, probably opt to use a junker 60-Hz, 115-volts AC induction motor for this project, along with suitable pulley ratios to match the optimum engine rpm to the induction motor's rated speed.

Peter Bertini, K1ZJH
Senior Technical Editor

A 230-Volt Generator from Scrap

Ron Mathers, ZL2AXO
Reprinted with permission from the September 1993 issue of Break-In.

A 230-volt, 50-Hz, 500-watt generator has been constructed primarily from discarded parts which would not be difficult to find for the keen homebrewer.

This generator is large enough to power relatively constant loads such as lighting, radio receivers, a TV set and, for instance, a 20-watt transceiver. However, it is not capable of supplying a 100-watt SSB/CW set, as the rapid fluctuations of load which the transceiver presents result in unacceptable swings in the generator terminal voltage. As well, the capacity is insufficient to run a refrigerator or an electric chain saw, though it will drive a 10-mm electric drill without problems.
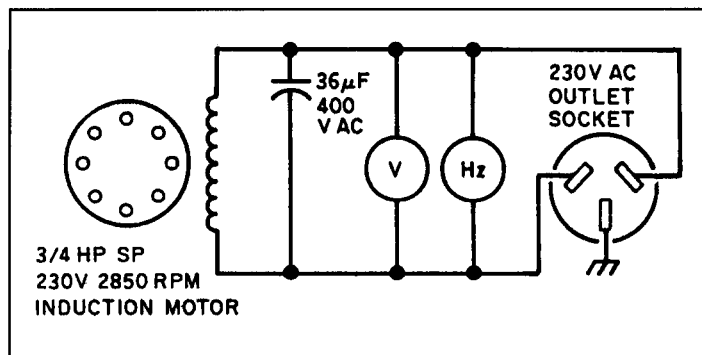
Figure 1. The induction generator.

A 3 HP lawnmower motor was brought out of retirement to play the part of the prime mover. This was coupled by a belt drive to a 3/4 HP single-phase induction motor used as an induction generator. The 3/4 HP motor was an old 2850 rpm pump motor salvaged from the local rubbish tip. Both bearings in the motor were renewed as it was found one tended to run warm. No modifications of any sort were done to the induction motor.

Capacitive excitation is used in the induction generator. Referring to **Figure 1** it can be seen that the connections are very simple indeed. It was found that a capacitance of 36 µF was a suitable size to give 230 volts, 50 Hz at no load. This was made up of three 12 µF, 400 volts AC condensers. The size of the capacitor required for capacitive excitation is approximately that which gives unity power factor when the motor is driven from the AC mains. Final adjustments to this figure can be made to give an output of 230 volts, 50 Hz at no load.

Changing the size of the capacitor changes the no-load frequency with the output voltage being kept at 230 by adjusting the throttle of the petrol motor.

Induction generators with capacitive excitation rely on the residual magnetism of the rotor to initiate the buildup of magnetic field in a somewhat similar way a DC shunt generator relies on the residual magnetism of its field poles. The magnetic field eventually rises to a value limited by the saturation curve of the machine. In earlier tests using a 1/4 HP induction motor as a generator it was found that if the generator was overloaded or short circuited, the residual magnetism was lost and, consequently, the generator was unable to restart unaided. The residual magnetism could be restored in one of three ways:

1. Connecting the generator to the AC mains and running it as a motor for a short period.
2. Discharging a charged condenser across the generator terminals while the machine is in operation.
3. Momentarily connecting a 6 or 12-volt battery across the generator terminals with the machine at rest.

The third method would probably be the most practical in most circumstances. As it turned out, the 3/4 HP motor gave no problems losing residual magnetism with short circuits or overloading. However, the possibility that a short circuit could result in the loss of residual magnetism must be kept in mind, and a proce-

dure for dealing with it available.

In operation, as load is applied to the generator, the terminal voltage needs to be kept constant by opening the throttle of the petrol motor. This increases the speed and results in an increase in the frequency. The frequency is proportional to the rotor speed minus the slip speed, which at full load is about 150 rpm. With a load of 500 watts, the frequency rises to 56 Hz. This rise in frequency is unlikely to be a problem in normal use apart from electric clocks running fast. To assist in the setting up of the generator, a frequency meter was built up using an old Jonan automobile tachometer driven by a 9-volt transformer.

The speed of the generator at no load is just over 3000 rpm. This needs to increase with load until at 500 watts output an estimated 3500 rpm is reached. On the petrol motor shaft, a pulley of 95 mm (3-3/4 inches) diameter drives a 1/2-inch vee belt to a 76 mm (3-1/2) diameter generator pulley.

The advantages of an induction generator include:

1. Low maintenance—no brushes, commutator, or slip rings to wear.
2. The output voltage is a sine wave.
3. The generator cannot be damaged by a short circuit.

Lights powered by the generator produce a noticeable flicker at a frequency equal to that of the power strokes of the petrol motor. It is caused by the low flywheel mass of the motor. However, it is not a serious problem and did not show up on the TV screen or affect the operation of other appliances.

BIBLIOGRAPHY
1. E.D. Basset and F.M. Potter, "Capacitive Excitation for Induction Generators," *Trans. AIEE*, May 1935, Vol. 4, page 540.

Moving on from lawnmowers to sail boats, Richard Cortis, VK2XRC, tracks down a mysterious electrical problem.

Beware of Dissimilar Metals

Richard Cortis, VK2XRC
Reprinted with permission from the October 1992 issue of Amateur Radio.

Apart from being an amateur radio operator, I occasionally attempt to pass myself off as a yachtsman. I am also the owner of a small sailing boat that's called "Mistress." Over the past year or so, my boat has been suffering from a lot of electrolysis.

Some electrolysis is always expected in boats, and zinc anodes are installed in various locations in an attempt to control the problem. However, the rate of loss of the sacrificial zinc anodes was particularly rapid on Mistress. Anodes can normally be expected to last in excess of 12 months, However, mine were lasting only about 6 months, requiring the installation of additional costly anodes and, even more importantly, the slipping of the boat, which is particularly costly. The problem was critical because the boat has an aluminum "sail drive" leg which protrudes through the bottom of the boat and looks somewhat like the bottom end of the leg of an outboard motor. This aluminum leg is particularly susceptible to corrosion when electrolysis is present. Accordingly, urgent action was required.

Possible Causes

All the usual things were investigated first. Stray currents from electric cables laid in the bilge were investigated and no likely cause was visible. A number of cables were replaced as a precaution. However, the electrolysis continued unabated. Further research was undertaken and there were suggestions that the copper based anti-fouling paint on the boat beneath the waterline may have had some effect. However,
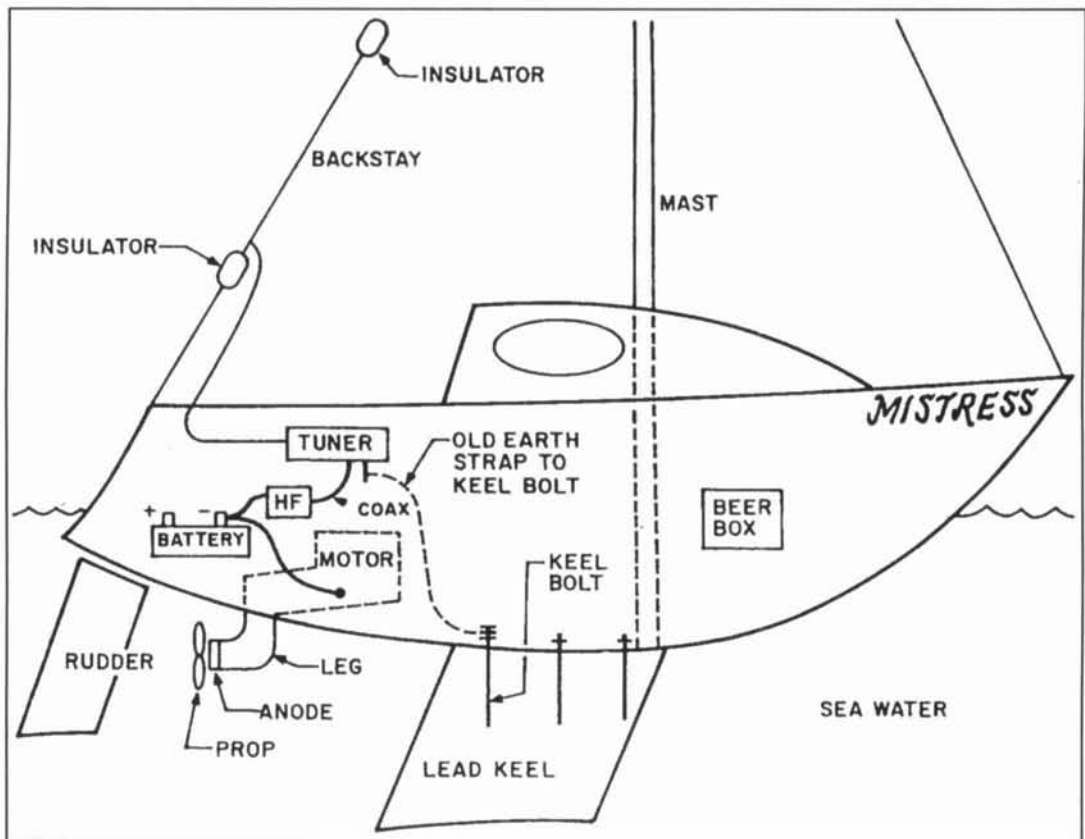
Figure 1. The radio installation setup aboard the Mistress.
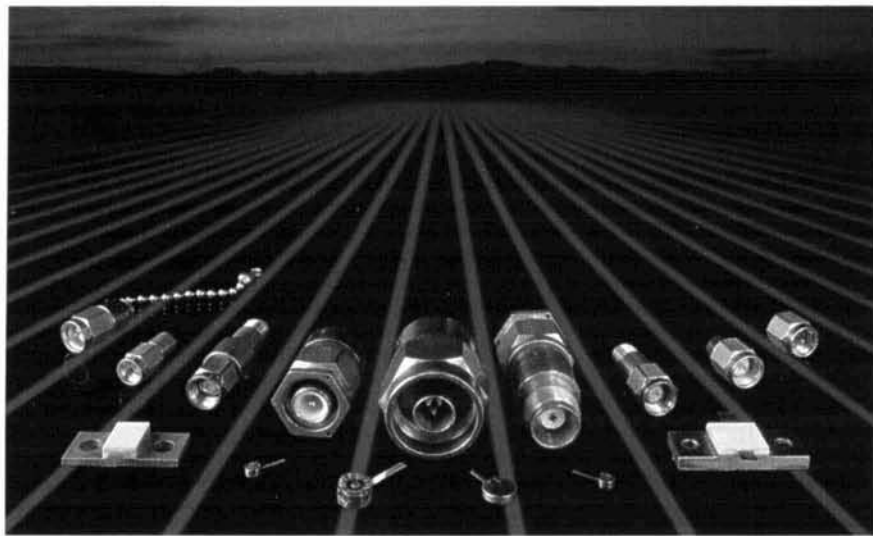
the area on the leg and on the bottom of the boat around the leg had been coated with a different anti-fouling, and the general opinion was that this was not affecting the electrolysis. I continued my search.

Recently I had the thought that the electrolysis problem first appeared at around the time the new HF marine radio was installed. **Figure 1** shows the radio installation on the Mistress. I commenced further investigation of the problem using my trusty salt-encrusted multimeter. The usual installation for an HF marine radio in a sailing yacht involves the linking of the ground connection from the tuner to a keel bolt using a heavy copper strap. This provides a substantial and effective connection to earth either by capacitance if the keel is encased in fiber glass, or by direct connection if not encased. During my investigation, the diver who cleans the marine growth from the bottom of the boat from time to time advised me that large areas of anti-fouling paint had become detached from the side of the lead keel. As this area had been carefully prepared and primed prior to the last anti-fouling, my suspicions were raised.

High Voltage Keel?

My investigation started at the accessible end of the earth strap between the tuner and the

keel. The earth strap was detached from the ground connection of the tuner and I used the multimeter to measure the voltage between the earth strap and the tuner. Initially, I set the multimeter to the lowest DC voltage setting, 0.25 volts, in order to get a meaningful reading. To my surprise, the needle on the meter swung across full scale. The next scale was selected and a reading in excess of 0.5 volts was indicated. I cannot give a more accurate reading than that as the multimeter is a cheap but robust item which lives on the boat and, because of salt spray and rough service, requires replacement quite regularly. Accordingly, I cannot afford a more sophisticated device. However, I find this meter provides a satisfactory level of service between replacements.

Having discovered this large potential difference, I then switched the meter from the voltage setting to the current setting, first selecting the 50-mA range. Again, the needle swung to full-scale deflection, and a higher range was selected. It appears that a DC current in the order of 250 mA was flowing between the ground connection of the tuner and the keel. This initiated further investigation and it was found that the negative connection on the battery was continuous through the HF radio and onto the outer braid of the antenna connection on the back panel. I also discovered the tuner was continuous from the outer braid connection of the coax from the HF through to the ground connection on the back of the tuner. The negative on the battery is connected by a heavy starter cable to the engine block, which is bolted to the gearbox, which is in turn bolted to the engine leg, which protrudes through the bottom of the boat into the sea water. I will not go into the details of the sealing of this engine leg through the bottom of the boat but, suffice to say it is durable and very flexible. The zinc anode is attached to the engine leg near the propeller.

Connections

So now we have it. The zinc anode is connected to the aluminum engine leg, which is connected to the gearbox, which is connected to the engine, which is connected to the negative of the battery, which is connected to the HF, which is connected to the tuner, which is connected to the earth strap, which is connected to the keel bolt, which is connected to the lead keel. And there we have it, a zinc anode at one end of a piece of wire connected to a lead cathode at the other end of the piece of wire. So why didn't I see it earlier? Well, thereby hangs a tale. The original HF radio installation in the boat had a different tuner which contained a capacitor in the ground connection. In the earlier radio, the negative from the battery was iso-

lated from the ground connection by this capacitor, which effectively blocked any DC current. So there was no electrolysis problem.

Now, having found the source of the problem, how do I remove the electrolysis? Happily, the solution was fairly simple and, perhaps, somewhat obvious. As the metal engine leg is of quite substantial dimensions beneath the water line, it was tried as a dyna-plate and proved to be generally satisfactory. No serious degradation in the level of the HF signal was detected. Accordingly, I made a more permanent connection between the ground connection of the HF tuner and the engine block. The connection between the tuner and the keel bolt was disconnected and removed.

Messages

There are a number of messages to amateurs and yachtsmen in this problem I have suffered. Firstly, one must never assume that items of equipment which look similar and provide the same function are actually constructed in a similar manner. The second and most important moral is that one must always avoid connections between dissimilar metals where any form

When we talked to technicians about servicing radios, they told us they were too busy to talk.

Now the new HP 8920A RF Communi-cations Test Set solves test problems in less time, so you have more.

- .Easy-to-use front panel

- Optional full-feature spectrum analyzer

- Rugged portable package

- Built-in computer with optional software for FM, cellular mobile and base station testing.

- Price $13,800

Keeping you ahead in communications.

For more information call: **1-800-344-3802** and ask for Charlie or Ken.

HEWLETT PACKARD

R-2600
"We take it to the mountain."

Craig D. "Ghee" Fry, *WL7C*
1308 Shallow Ford Road
Herndon, Virginia 22070
AND
Robert J. Yowell, *KC5BRG*
14614 Laurel Shadows Court
Houston, Texas 77062

HF RADIO ON MARS

*The solution to a difficult
communications problem?*

S ometime early in the next century, people will travel to Mars. They'll set up base camps and roam across the landscape to explore a new world. The success of their activities will depend upon how free they are to move about on the surface. Astronauts will need an adequate supply of air, water, and food to travel long distances, and they will need to communicate with their base stations.

Mission planners will have to tackle a difficult communications problem: how do you ensure the astronauts will be able to talk to each other over wide distances, say hundreds or thousands of miles? The typical UHF radio system used by NASA allows line-of-sight communication. Once beyond the horizon, the signal is lost. The visible horizon for a person standing on Earth is about 20 miles, but on
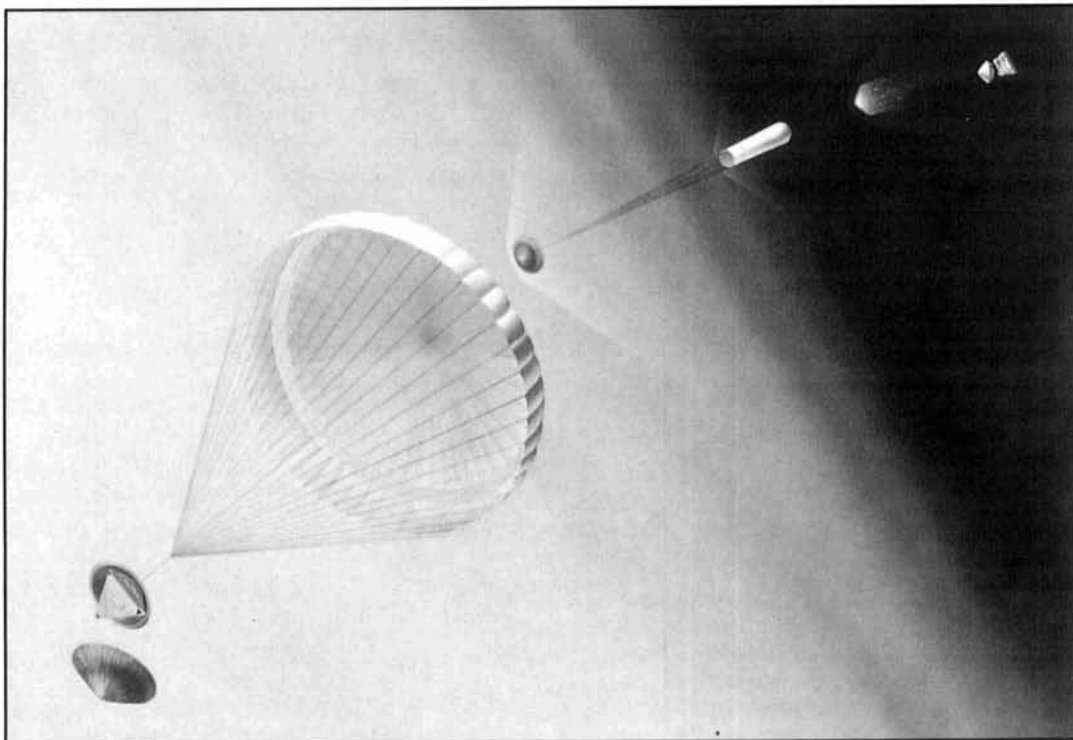
#	Site	Latitude	Longitude	Elevation (km)
1	Valles Marineris	−6.00	58.00	3.50
2	Valles Marineris	−5.00	54.00	3.50
3	Chryse Plan. VL-1	23.00	48.00	−2.20
4	Olympus Mons	13.00	130.00	2.00
5	Valles Marineris	−2.00	54.00	3.50
6	Hadriaca Patera	−32.00	268.00	2.00
7	NW Hellas	−40.00	310.00	−2.00
8	Argyre Planitia	−37.00	44.00	3.50
9	South Pole	−86.00	315.00	4.00
10	Sirenum Terra	−45.00	185.00	4.20
11	Northern Plains	60.00	50.00	−2.70
12	Gusev Crater	−15.00	185.00	0.50
13	Syrtis Major	5.00	295.00	3.50
14	North Arabia	38.00	309.00	2.20
15	North Pole	82.00	55.00	0.00
16	Chasma Borealis	−66.00	66.00	2.50

Table 1. Candidate MESUR observation sites. Latitudes are in degrees North, longitudes are in degrees East, and elevations are in kilometers.

Mars it's about half that far. Radio repeater stations could be set up on the surface to extend the communication range, or orbiting relay satellites might receive the weak signals from the surface and retransmit them over the horizon. A third method of communicating would be to use Earth stations to relay signals between Martian explorers. This idea may not be too practical unless the exploring astronauts have lots of transmitting power and tremendous patience. Depending on how close the two planets are in their orbits, the travel time for a radio signal from Mars to Earth and back ranges from 9 to 40 minutes!

We'd like to suggest that Mars explorers may be able to communicate with each other by bouncing their signals off the Martian ionosphere, a natural "radio mirror." But before people travel to Mars, robots will first make the trip.

Background

Near the end of this decade, NASA and the European Space Agency (ESA) plan to send unmanned spacecraft to Mars. These spacecraft will carry landers, to be deployed as observing stations on the surface. Remote vehicles, or "rovers," and balloons may also be used to study a larger portion of the Martian surface and atmosphere.

NASA's plans call for placing a chain of small stations on Mars, called the Mars Environmental Survey, or MESUR (pronounced "measure"). This program will collect meteorological, seismological, and other surface and atmospheric data.[1] Table 1 shows the planned locations of the 16 MESUR landers. These sites were selected to provide extended coverage of the Martian globe, and at the same time, allow the study of intriguing geological sites. A precursor mission called MESUR Pathfinder could head for Mars as early as 1996. Photo A shows an artist's sketch of the MESUR Pathfinder lander descending to the surface of Mars. Several of the MESUR landers may carry "microrovers," such as Rocky IV shown in Photo B. Rocky weighs in at about
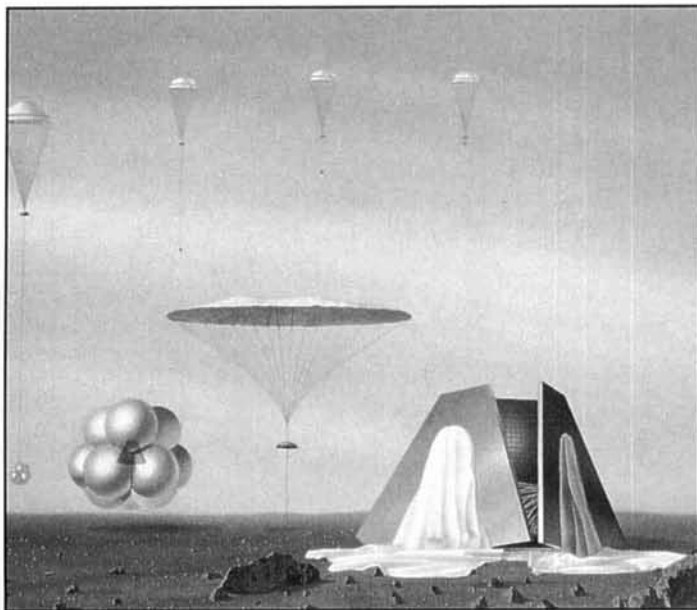
Photo A. NASA plans to send a series of small spacecraft to Mars at the end of this decade. This project, called the Mars Environmental Survey, or MESUR, will place 16 landers and several microrovers on the planet's surface. The details of how MESUR will be linked together to form a network are still being worked out. This artist's conception shows the prototype MESUR Pathfinder descending to the Martian surface.

16 pounds (11 kilograms), and will be smart enough to maneuver on his own. He will carry a visible-light spectrometer and color camera, and tools to chip rocks and scoop soil samples. Rocky will be linked to his lander by low power UHF radio.

The ESA is developing a mission called MARSNET to put three surface stations on Mars at about this same time.[2] The MARSNET sites form a triangle about 3000 kilometers on a side and will complement the MESUR network. If these missions fly as planned, they'll provide a network to give us a global picture of the Martian environment for the first time.

Both MESUR and MARSNET missions will likely rely upon orbiting relay satellites to send the data collected at the multiple sites back to Earth. It's possible the data may be transmitted directly from each site to Earth instead. But both of these relay concepts present problems. Relay satellites are expensive, require a lot of fuel to send to Mars, and make up a sizable portion of the overall cost of the mission. Direct transmission of data to Earth requires NASA to aim the high-gain antennas of its Deep Space Network (DSN) toward Mars to pull in the weak signals. Considering the large number of sites in the MESUR mission, the requirement to collect data from each site may place heavy scheduling demands upon the DSN.

On the other hand, if the data and telemetry from missions such as MESUR could be forwarded to two or three central sites on the Martian surface, for subsequent retransmission to Earth, a much smaller number of Earth stations would be needed, and would reduce the overall cost of the communication portion of the mission budget. This is where the Martian ionosphere comes in. At the Case for Mars conference in Boulder, Colorado this in June 1993, we proposed using, ionospheric refraction for over-the-horizon communication on Mars.[3] Here, we look at the feasibility of using shortwave (or HF—high frequency) radio to connect a chain of remote sites on Mars for the relay of collected data. We use typical observed characteristics of the Martian ionosphere and the MESUR site locations for illustration purposes. We'll keep in mind that current HF data communications technology won't support the total data rates required by the MESUR mission—10 MBit per day from 8 sites, and 1 MBit per day from the remaining sites.

Over-the-horizon propagation

First, consider radio propagation here on Earth. People have been using short-wave radio to communicate for over 90 years. An HF radio wave transmitted upward from the surface is refracted back toward the Earth as it travels

Photo B. The MESUR Pathfinder lander and its microrover, Rocky IV, on the Martian surface. Room for an HF mobile antenna?

through the ionosphere. The radio wave may eventually travel downward to a distant point on the ground, several hundreds or thousands of kilometers from the transmitter. This mechanism provides a fairly reliable means of communicating. The radio wave may also be reflected off the ground at the distant point and back into the ionosphere to be refracted again. In this way, around-the-world communication occurs via "multi-hop" propagation.

Even with the advent of satellites, shortwave radio continues to play a major role in long distance communication on Earth. This is due to short-wave radio's low equipment costs and fairly consistent performance, at least at low bandwidths (several (kHz). Even with low power transmitters and simple antennas, communication usually occurs during part of each day, with the right choice of frequency. Often times, the power needed for good communication ranges down to the 1-watt level.[4] Also, new technologies are pushing usable bandwidths higher and higher.

It seems reasonable to suspect that the same HF radio propagation might occur in any planetary ionosphere, as long as the electron density is sufficient to refract the radio wave back to the surface, and attenuation due to absorption isn't too great. Early space probes discovered Mars has an ionosphere, with electron densities
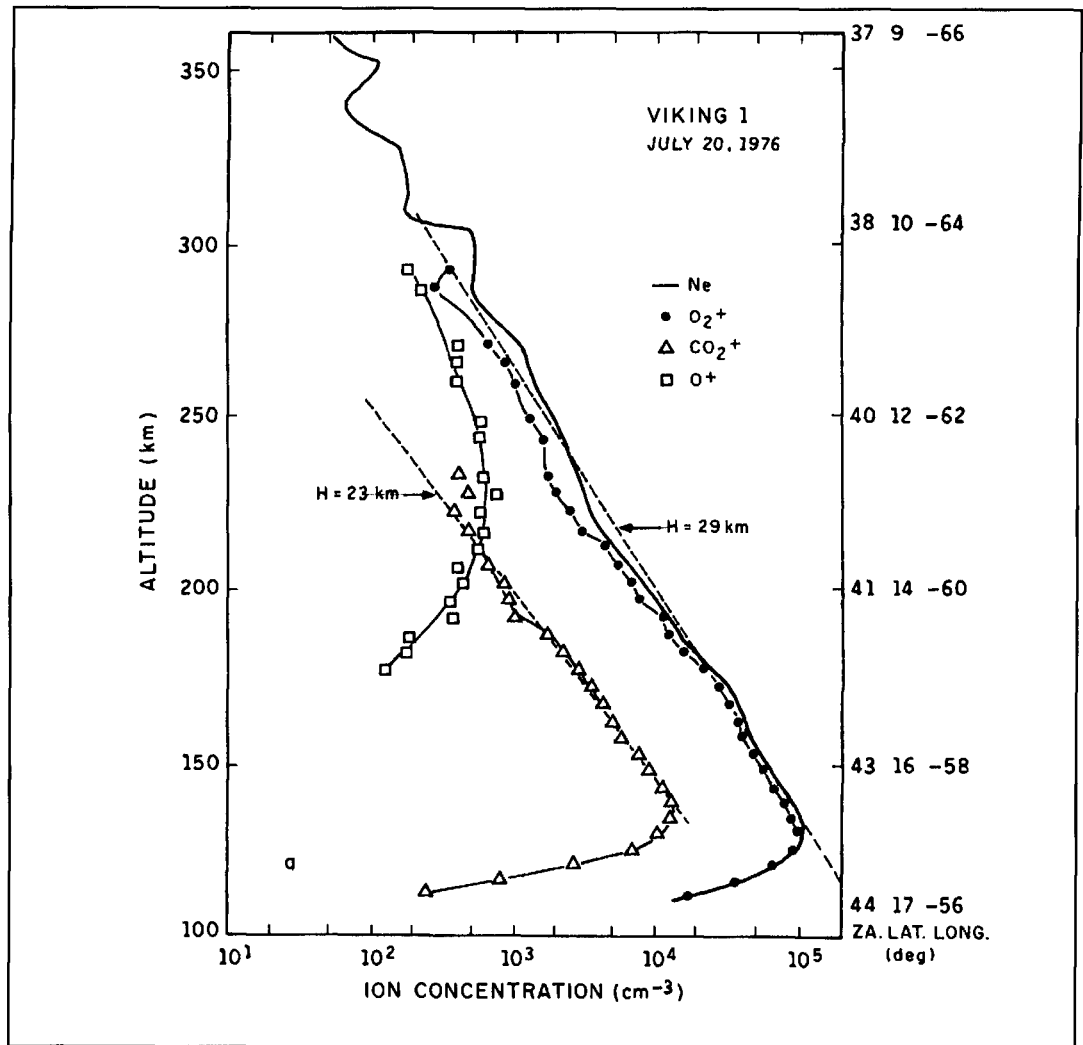
Figure 1. Observations of ion concentrations in the Martian ionosphere during the Viking I lander descent. The solar zenith angle, labeled ZA, ranges from 37 degrees to 44 degrees in this data set. This figure is from Hanson et al.[5]

similar to Earth's ionosphere. This leads us to suspect over-the-horizon communication should be possible on Mars.

The Martian ionosphere

Let's look at the properties of the Martian ionosphere. Peak electron densities are about 1 to 2×10^5 cm^{-3} during the Martian daytime, and fall to about 10^3 cm^{-3} at night. **Figure 1**, from Hanson et al., shows the data obtained from the Viking 1 lander as it descended to the surface in 1976, during solar minimum.[5] The neutral Martian atmosphere is mostly CO_2, and the chief ionized component is O_2^+, which peaks in density at about 135 kilometers. The electron density equals the sum of the densities of all of the ion species. **Figure 2**, from Zhang et al.,[6] shows a series of electron density profiles obtained from radio occultation measurements, in this case from the Mariner 9 orbiter, during

the 1971 in the declining phase of the solar activity cycle. The bottomside ionosphere, that portion below the ionized density peak, and of interest to HF communicators, is fairly consistent during the daytime. For an in-depth discussion of our current understanding of the Martian ionosphere, see the review by Kliore[7] and the papers by Zhang et al.[6,8]

The Martian ionosphere has an electron density-versus-height profile similar to the so-called Chapman profile, in that it increases steeply in density with an increase in altitude up to the peak density, and falls off more slowly (roughly exponentially) above the peak. Observations from a number of spacecraft have shown the bottomside ionosphere is well established and fairly consistent from observation to observation. Whether the main layer is more like the Earth's E, F1, or F2 layers is still under debate. The Martian ionosphere probably has a D layer, and a thin E layer has been observed, but the details are still undetermined because of

present measurement difficulties. Martian day-time ionospheric densities are comparable to the summer hemisphere conditions at mid-latitudes on earth. The Martian ionosphere doesn't totally disappear at night. Peak electron densities often remain above 10^3 cm^{-3}, but the values are much more variable than during the daytime.

In radio propagation jargon, the critical frequency is the highest frequency of a vertically traveling radio wave that will be reflected back to the surface.[9] The frequency, f_c, in Hz, is related to the electron density, n_e, in cm^{-3} by the formula:

$$f_c = 9000 \sqrt{n_e} \qquad (1)$$

From all this we conclude that the critical frequency of the Martian ionosphere is about 3 to 4 MHz during the daytime and drops to a few hundred kHz at night. The highest frequencies that will be returned back to the surface are somewhat higher than the critical frequency for non-vertical propagation, and are highest for low radiation angles. For a good discussion of terrestrial HF propagation, refer to any recent edition of *The ARRL Handbook for the Radio Amateur*.[10]

HF network study

The locations of the 16 MESUR sites listed in **Table 1** were used to calculate the great circle distances between sites, as shown in **Table 2**. Read **Table 2** as you would a mileage chart on an interstate map. For example, the distance between sites 1 and 2 in Valles Marineris near the Martian equator is 243 km, and between sites 2 and 5 is 178 km.

Computerized radio propagation studies usually involve modeling (obtaining a good numerical representation of) the ionospheric

Figure 2. Martian ionospheric electron density profiles obtained from Mariner 9 radio occultation observations. This figure is from Zhang et al.[6]

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Valles Marineris	0															
2	Valles Marineris	243	0														
3	Chryse Plan. VL-1	1813	1694	0													
4	Olympus Mons	4382	4594	4603	0												
5	Valles Marineris	335	178	1521	4551	0											
6	Hadriaca Patera	7845	7750	8506	8124	7887	0										
7	NW Hellas	5904	5768	6541	9062	5885	2048	0									
8	Argyre Planitia	1986	1971	3561	5606	2144	5904	4138	0								
9	South Pole	5029	5072	6702	6336	5249	3277	2725	3141	0							
10	Sirenum Terra	6541	6729	8220	4522	6873	3755	4841	5376	2822	0						
11	Northern Plains	3927	3854	2193	4370	3677	8438	7614	5753	8890	8860	0					
12	Gusev Crater	7311	7551	8216	3625	7618	4519	6216	6891	4596	1777	7371	0				
13	Syrtis Major	7301	7063	6451	9277	7052	2668	2788	6405	5404	6374	5790	6550	0			
14	North Arabia	6434	6212	4898	7640	6102	4723	4620	6821	7345	8193	3716	7451	2095	0		
15	North Pole	5213	5153	3498	4445	4975	7616	7795	7056	10170	8278	1305	6518	5273	3233	0	
16	Chasma Borealis	3569	3647	5333	5438	3824	4770	3744	1877	1522	3528	7496	5174	6527	8000	8772	0

Table 2. Great circle distances between each of the 16 candidate sites listed in *Table 1*, in kilometers. Read this table as you would a highway mileage chart. The intersection of a row and column gives the distance between the corresponding two locations.

features you want to study, and ray tracing (determining the effects of the ionosphere on the radio wave and its path of travel through this ionosphere). Usual parameters of interest are the ground range, and group and phase delay of the radio wave. In order to study Martian HF propagation, we represented the Martian ionosphere using a "quasi parabolic" ionosphere model as described by Croft and Hoogasian.[11] **Figure 3** is from their paper, but modified for the Martian ionosphere. In this model, the electron density profile is defined by the equation:

$$N_e = N_m \left[1 - \left(\frac{r - r_m}{y_m} \right)^2 \left(\frac{r_b}{r} \right)^2 \right] ; \; r_b < r < r_m \left(\frac{r_b}{r_b - y_m} \right)$$

(2)

$$N_e = 0 \qquad\qquad ; \text{ elsewhere}$$

where:

N_m	= peak electron density
r	= distance from the center of Mars
r_o	= radius of Mars
r_m	= height of peak ($= h_m + r_o$)
y_m	= semi-thickness (one half the layer thickness)
r_b	= $r_m - y_m$

Ray tracing is accomplished by integrating the refractive index (related to local plasma density) along the radio ray path, taking the bending due to refraction, and other factors, into account. **Figure 3** shows the ray path geometry used in the calculations. The quasi-parabolic equation has the advantage of yielding an exact solution when integrated. This means that if the quasi-parabolic model is a good representation of the bottomside Martian ionosphere, the ray tracing should be representative of expected behavior on Mars.

We hand-fit quasi-parabolic curves to observed Martian electron density profiles, and calculated single-hop propagation distances, or ground ranges ($= 2\theta_t$), along the Martian surface. **Figure 4** shows the daytime and night-time electron density profiles for the quasi-parabolic model ionosphere. For the daytime conditions, we used a peak electron density of 1.7 x 10^5 cm^{-3}, a peak height of 135 km, and a semi-thickness parameter of 35 km. For the Martian night, we used a peak density of 1 x 10^3 cm^{-3}, a height of 150 km, and semi-thickness of 50 km. **Figure 5** shows the ground range of a refracted ratio wave for various take-off angles (β_o in **Figure 3**), and includes curves for several frequencies. Note that for a given frequency, two take-off anglescan give the same ground range, and a transmitted signal can arrive at a receiver from two different

angles. This means that multipathing can degrade signal quality, just as on Earth. **Figure 6** shows the results of the nighttime calculations. The daytime and nighttime figures are similar, except that the frequencies that produce the same ground ranges for a given take-off angle are lower at night by an order of magnitude.

Our results indicate that with the right choice of transmitting frequency, HF radio can achieve 2000 km ground ranges during both Martian day and night. Referring to **Table 2**, a number of MESUR sites are within the one-hop 2000 km range. For example, data could be relayed from site 9 at the south pole to site 3 at 28 degrees north via sites 16, 8 and either site 1 or 2 in Valles Marineris. In fact, data could be relayed between sites 1, 2, 3, 5, 8, 9 and 16.

Spacecraft observations indicate the height of peak density rises to 150 km in the daytime near the day/night terminator. This implies that the one-hop ground range might be as much as 2200 km when the sun is low on the horizon. Under these conditions both the north and south pole MESUR sites could communicate with the Valles Marineris sites, and in that way, with each other. The network would look as follows:

$$15 \Leftrightarrow 11 \Leftrightarrow 3 \Leftrightarrow (1,2,5) \Leftrightarrow 8 \Leftrightarrow 16 \Leftrightarrow 9$$
$$6 \Leftrightarrow 7$$
$$10 \Leftrightarrow 12$$
$$13 \Leftrightarrow 14$$

Only site 4 at Olympus Mons would be out of range of all other stations.

If two-hop propagation were possible, the ground range would increase to 4000 km. This would allow all sites to be connected by HF radio and communications would be global.

In this study we ignored small scale irregularities in the Martian ionosphere, and the existence of any D or E layers. We assumed the ionosphere was spherically symmetric and didn't have any horizontal gradients. We also didn't include a Martian magnetic field, and didn't consider the effect of absorption on the radio signal. Signal attenuation should be included in future studies. Although it isn't intuitively obvious, signal attenuation has no effect on ground range, only on the signal strength at the receiver end of the path. In other words, the signal will travel wherever the ionosphere lets it, but it may not be heard at the receiver because it is attenuated below the noise floor of the receiver.

HF communications requirements

For HF communication to be practical on Mars, a number of conditions have to be met. We've listed the important ones here.

1. The peak Martian ionospheric electron

densities have to be sufficient to refract the transmitted radio waves back to the surface at the receiver end of the link.

2. The ionosphere must not cause excessive absorption of the radio signal. On Earth, absorption can cause a 0 to 20 dB attenuation of the signal per hop.

3. The one-hop or multi-hop coverage zones have to allow mutual access between relay sites. That is, every site must be within at least one or two hops of another site to allow relay of data.

4. The upper and lower usable frequencies must be far enough apart in the radio spectrum to allow the entire bandwidth of the modulated signal to propagate.

5. Ionospheric irregularities and traveling disturbances must not cause excessive scintillation, dispersion, or multipathing to such an extent that unacceptable distortion or loss of signal occurs.

6. The ionosphere has to be stable enough to support propagation at a particular frequency for the period of time necessary to allow the transfer of data and telemetry at the rates needed.

7. Radio noise generated by equipment or natural sources must be acceptably low at the receiver sites. That is, signal-to-noise ratios at the receiver sites have to be sufficiently large for reliable communications.

8. Ground reflection losses must not be too great if we're relying on multi-hop propagation to bridge the link. On Earth, the losses from ground reflections can range from 0.5 to 3 dB per hop.

9. The benefits of using HF radio have to outweigh the shortfalls. The required hardware and software complexity, mass, and power consumption must be traded off against other communications options.

Actually, in some ways HF propagation might be better on Mars than on Earth. Most of the observations of the Martian ionosphere were obtained during the declining phase of the solar activity cycle, or at solar minimum. During solar maximum, when MESUR flies, the ionosphere should be more dense than the measured values, and the critical frequencies will be higher. This means radio propagation should be even better than at solar minimum. A short-wave DXer's rule of thumb says that the higher the critical frequency, the higher the allowable transmitter frequency and the better the signal quality will be.

The interaction of the solar wind with the Earth's ionosphere deposits large amounts of energy in the ionosphere, and geomagnetic activity often causes large variations in ionospheric conditions. These "ionospheric storms" can degrade HF communications. The Earth has a strong magnetic field that causes birefrin-
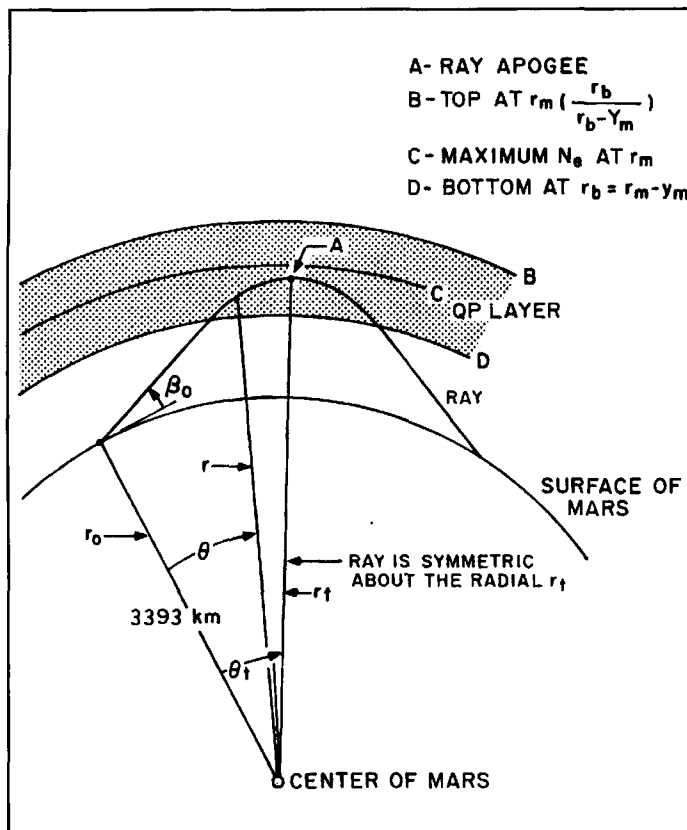
Figure 3. Ray path geometry for the quasi-parabolic Martian ionosphere model, adapted from Croft and Hoogasian's paper.[11]

gent refraction. This splits radio waves into two components, the ordinary and extraordinary waves, and also complicates terrestrial propagation. The Earth can be a noisy radio environment. Auroral, atmospheric, and equipment-generated sources contribute to the noise. The Earth's HF spectrum is also crowded with many users, and interference can be quite a problem. Ask anyone who's worked a pile-up on 20 meters.

On the other hand, Mars doesn't have a strong magnetic field, and in spite of the topside ionosphere being exposed directly to the solar wind, the bottomside is fairly stable. The Martian magnetic field is either very weak or nonexistent, so birefringent refraction will be much smaller than on Earth. There are some potential radio noise sources on Mars: extraneous emissions from electronic equipment onboard the landers or rovers, and electrostatic discharge due to wind-driven dust or ice. Noise sources external to the planet, such as solar and galactic radio emissions shouldn't be a problem. However, it's possible that galactic radio noise might pass through the nighttime ionosphere at a low angle and pass over the horizon and be refracted at the bottom of the dayside ionosphere. This could contribute to the received noise at a sunlit site.
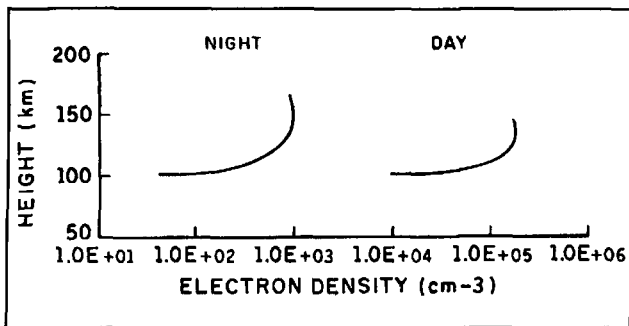
Figure 4. Electron density profile for the quasi-parabolic iono-sphere model, with daytime parameters of $N_m = 1.7 \times 10^5$ cm^{-3}, $y_m = 35$ km, and $h_m = 135$ km, and nighttime parameters of $N_m = 1 \times 10^3$ cm^{-3}, $y_m = 50$ km, and $h_m = 150$ km.

Future studies

What kinds of studies would be useful to determine the feasibility of using HF radio as a communications mode on Mars? There are a number of radio experiments that could be included on Mars missions such as MESUR. One extremely low-cost piece of equipment is the relative ionospheric opacity meter, or riometer. This instrument is a self-calibrating HF receiver used to measure absorption by monitoring galactic radio noise that passes through the ionosphere at frequencies above the critical frequency. Absorption is determined from the difference in the received signal and the signal that would have been received if there were no absorption. Active monitoring of the bottomside ionosphere can be accomplished using sounders. A vertical incident sounder, or ionosonde, transmits radio waves over a broad frequency range straight up toward the ionosphere and receives the reflected radio waves.
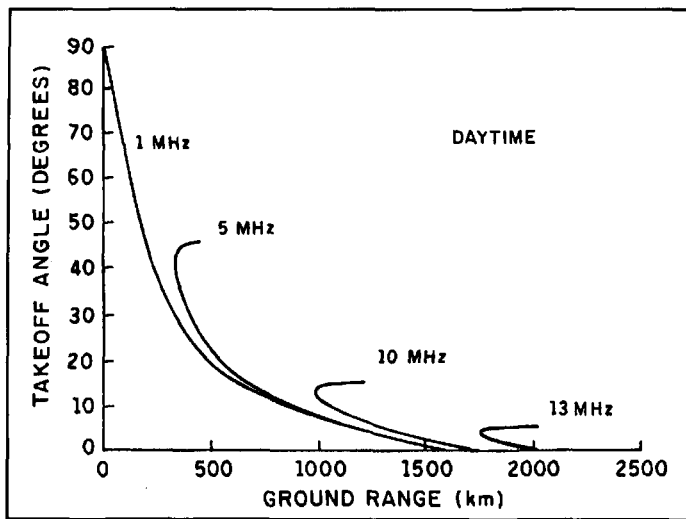
Figure 5. Ground ranges as a function of take off angle (β_0 in Figure 3) for 1, 5, 10, and 13-MHz radio waves, using the daytime electron density profile shown in Figure 4. Peak electron density is 1.7×10^5 at 135 km.

Ionospheric structure can be determined from the returned signal. Oblique incidence sounders such as the chirp sounder are similar to the ionosonde, but use the geometry shown in Figure 3 to probe the ionosphere. Chirp sounders that operate in the 6-watt transmitter power range are available commercially.[12]

Several Earth-based investigations can be conducted before we actually send HF radio experiments to Mars. These include a) HF radio propagation computer modeling and network studies, and b) equivalent terrestrial radio propagation experiments under conditions similar to those expected on Mars. This would give us insight into whether the Martian ionosphere can support reliable long-distance communication.

a. HF propagation simulations. Recent research in the space physics community concerning Mars has focused on several areas: the interaction of the Martian ionosphere with the impinging solar wind; whether Mars has an intrinsic magnetic field or one induced by solar wind interactions; and the structure of the top-side ionosphere. However, the bottomside ionosphere is important for HF radio propagation because radio waves are refracted back to the surface only at altitudes below the height of the electron density peak.

An ionospheric model that characterizes the spacecraft measurements and current theoretical understanding of the Martian ionosphere, should be coupled with an accurate radio ray tracing algorithm. Several researchers have developed Martian ionosphere models.[13,14,15] Realistic studies should use three-dimensional (3-D) ionospheric models, which not only include vertical variations of plasma density, but horizontal changes as well. Zhang et al.,[6] in their paper on the daytime Martian ionosphere, presented a simple model based upon a modified Chapman layer. This model fit their reanalysis of the Mariner and Viking radio occultation data fairly well. A thorough propagation study should also include a full 3-D numerical or piece-wise analytical ray tracing algorithm. Candidate ray tracing algorithms are described in Jones and Stephenson[16] and Reilly and Strobel.[17]

A realistic simulation must account for signal losses due to absorption. Unfortunately, theoretical absorption models based upon collision processes aren't known for accurately predicting observed absorption conditions in the terrestrial ionosphere.[18] Agarwal[19] looked at absorption effects on radio propagation in the Martian ionosphere. He assumed the dominant ion species was O^+, and looked at radio propagation in the frequency range below 100 kHz or so. We now know that O_2^+ is the main ion species at peak density heights, so his analysis should be reaccomplished, looking at effects of

The Planet Mars QSO Cup

By the end of the 1920s most amateur records had been set, and after the initial smashing of waves upon waves of records, there seemed to be few worlds left to conquer. Technology had evolved in a few short years to the point where arc and spark transmitters were obsolete, and superhet receivers replaced the finicky regenerative in many amateur shacks.

Enter Fred Johnson Elser, W6FB/W7OX and Stanley M. Mathes, ex-7OE and then K1CY, who conceived of the idea of a trophy that would outdistance any short-term advancements in radio communications—the Planet Mars QSO Cup.

In 1928, Elser had just returned home to Manilla in the Philippines from a pilgrimage to ARRL headquarters in Newington, Connecticut. While there, he met Percy Maxim, W1AW, ARRL founder and president. During time spent with Maxim, he learned of W1AW's fascination with the planet Mars.

At the same time, Mathes (Lieutenant Commander USN) was stationed in the Philippines, which were then under United States jurisdiction, and also part of the ARRL field organization. After a visit Mathes paid to Elser's shack, the two avid radio hams came up with a unique trophy for the first amateur two-way communication with Maxim's favorite planet, Mars.

After settling on the rules for their award, Elser and Mathes set out to find a suitable trophy. In a store in Baguio, where many examples of Igorot native woodcarving were on display, they found their prize. The trophy is a large curved bowl, standing about 12 inches high. Seated figures on the trophy base, which stands for the planet Earth, signify the people of the Earth. The standing figures holding the bowl, which represents the planet Mars, are amateurs who will bridge the gap between the two planets. The Mars Cup is presently on display at the ARRL headquarters in Newington, Connecticut, where it has resided since 1929.

Two plates are affixed to the base of the cup. One bears the legend "First Amateur Radio Two-Way Communication Earth and Mars,"

The Elser-Mathes cup. This trophy for the first two-way amateur communications between Earth and Mars conceived by Col. Fred Elser, K3AA, and Lt. Comdr. Stanley Mathes, K1CY, while in the Philippine Islands in 1928. Photo courtesy of the Amateur Radio Relay League and *QST* magazine, Newington Connecticut.

with as yet unfilled spaces for the names, calls, and dates for those who would claim the trophy. The other plate simply bears the words [the] "Elser-Mathes Cup." While it was the award creators' fondest wish to see the cup claimed during their lifetime, it seems that was not to be. However, with amateurs routinely active from space, and the surface of the moon long broached, we can only hope that we will all soon see the day when amateurs lay claim to this unique award, and Messrs. Elser and Mathes dream is at last realized.

Peter X, K1ZJH
Senior Technical Editor

Information on the Planet Mars QSO Cup was taken from "That Planet Mars QSO Cup," by Fred Johnson Elser, W6FB/W7OX. The article appeared in QST magazine.

---

absorption on the HF part of the radio spectrum. Much is yet to be learned about the behavior of the Martian ionosphere, and the roles played by solar wind interactions, ionospheric convection, neutral winds, atmospheric gravity waves, and airborne dust.

b. Equivalent terrestrial HF experiments. We can run HF radio experiments on Earth to predict the behavior of HF radio propagation and communication on Mars. These experiments should be conducted at sites and under environmental conditions that match those on Mars as closely as possible. Candidate sites include regions where ground conductivity and dielectric constants are similar to those expected on Mars, and where ionospheric conditions are similar to those observed by previous Mars missions. Antarctica is probably the best match
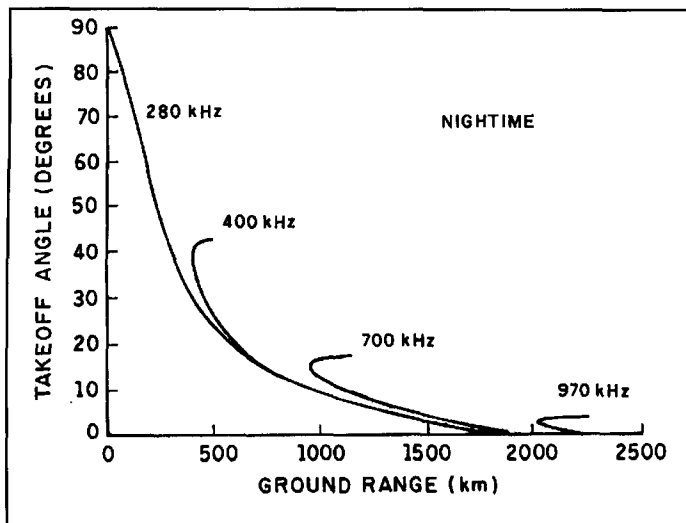
Figure 6. Ground ranges as a function of take-off angle for 280, 400, 700, and 970-kHz radio waves, using the nighttime electron density profile shown in *Figure 4.* Peak electron density is 1 x 10³ at 150 km.

radio for over-the-horizon propagation on Mars. The existence of an ionosphere on Mars implies that HF propagation should be possible. A tentative study using a quasi-parabolic approximation with exact ray tracing shows that one-hop propagation out to 2000 to 2200 km along the Martian surface is possible. We used the MESUR landing site scenario to look at the possibility of conducting radio science experiments on Mars and the possibility of establishing an HF network on the Martian surface. Approximately one-half of the MESUR landers could be connected via HF radio during at least part of the day in the one-hop case. All the landers could communicate if two-hop propagation occurred. Data could then be relayed from remote sites to two or three Earth-link stations. This would reduce or eliminate the need for orbiting relay satellites or the direct transmission of data back to Earth from each site.

The tentative results presented in this paper, using the MESUR scenario of 16 landers, are encouraging. For HF propagation to be practical for network communication on the Martian surface, it must be competitive with other options, such as the option to use relay satellites, or to beam the data directly from each site to Earth. Further investigations and link margin analyses require 3-D raytracing through realistic ionospheric models. Equivalent terrestrial propagation experiments would be beneficial. The results should be analyzed in terms of the implications for Martian radio science experiments, and in terms of HF radio as a viable communications mode on Mars. These studies will help determine the expected ionospheric circuit characteristics for likely sites.

Although we're proposing a low-tech approach to planetary radio experiments and communications, the idea has never been tested on another planet. NASA certainly wouldn't rely on HF radio for their MESUR surface network backbone, but preliminary studies do suggest that HF radio experiments would be useful additions to the manifest on future planetary missions. These types of experiments can provide valuable information about the Martian ionosphere and about signal propagation characteristics. Researchers have been using HF receivers and low power transmitters to study the Earth's ionosphere for over 75 years. It's time we move this proven technology off planet, and support NASA's new direction to do things cheaper, faster, and better. ■

from this perspective. Sites should be selected that support single-hop propagation with geometry similar to those anticipated on Mars. Again, since Mars has half the diameter of the Earth, geometry with 1000 to 4000 km propagation paths on Earth correspond to equivalent Martian separations of 500 to 2000 km.

Recent advances in digital communication capabilities have improved the data rates and performance of HF radio. Adaptive digital modulation techniques have come on the scene, and error-free data transfer rates of 750 Bps are achievable, using modulation techniques such as CLOVER.[20] The picture continues to improve. Recently, the Advanced Research Projects Agency issued a Small Business Innovation Research Program Solicitation for the development of an HF network system capable of 64 kBps.[21] If this research pans out, the new technology would make HF radio a stronger contender for network communications on Mars.

In planetary missions, the average transmitted power must be kept low because electrical power is a limiting factor. This impacts communications because lower power equates to decreased signal quality. However, there is a popular misconception that HF radio requires high power. The low power chirp sounder discussed above supports automatic link establishment and data transfer. Keep in mind that higher power communications systems are still a possibility, as long as the transmission duty cycles can keep power consumption low enough. This should also be investigated further.

Summary and conclusions

This article discussed the concept of using HF

REFERENCES
1. G.S. Hubbard, "Mars Environmental Survey Concept," *Mars: Past, Present and Future*, edited by E.B. Pritchard, American Institute of Aeronautics and Astronautics, Washington, D.C., 1992, pages 141–152.
2. G.E.N. Scoon, "MARSNET—A Mars Network of Surface Stations," *Mars: Past, Present and Future*, edited by E.B. Pritchard, American Institute of Aeronautics and Astronautics, Washington, D.C., 1992, pages 153–168.
3. C.D. Fry and R. Yowell, "Over-the-Horizon Communications on Mars Via

How Are We Alike?

Just how similar is Mars to Earth? In spite of being half the size and 1.5 times farther from the Sun, in many ways Mars is remarkably similar to Earth, as the table below shows. The length of a Martian day and the tilt of Mars' rotational axis are nearly the same as Earth's. This means that Mars has seasons as on Earth, but the Martian seasons last twice as long as ours. The Martian air, mostly composed of CO_2, is too thin to breath but dense enough to allow steep temperature changes to kick up strong winds. These winds create global dust storms which can blot out our view of its surface for months on end.

Mars-Earth Comparison

	Earth	Mars
Average distance from Sun (millions of km)	149.6	227.9
Radius at equator (km)	6378	3393
Surface gravity relative to Earth	1.0	0.4
Average density (gm/cm^3)	5.5	3.9
Number of days in a year	365.26	686.98
Length of day	23H56M	24H37M
Tilt of rotational axis (degrees)	23.4	25.2
Atmospheric composition	N_2, O_2	CO_2
Atmospheric pressure at surface (millibars)	~1000	~6–8
Number of moons	1	2
Temperatures at the surface (°C)	±40	+20/–100
Ionosphere:		
Peak electron densities (day) (cm^{-3})	$\sim10^5-10^6$	$\sim10^5$
Peak electron densities (night) (cm^{-3})	$\sim10^5$	$\sim10^3$
Height of peak densities (day) (km)	~200–300	~130–150
Height of peak densities (night) (km)	~300–400	~150

HF Radio Propagation," *Case for Mars V*, Boulder, Colorado, June 1993, in press.

4. D. Martin and D. Newkirk, "Transmitter Power: What It Is, What It Does and How to Use It," *QST*, March 1993, pages 28–31.

5. W.B. Hanson, S. Sanatani, and D.R. Zuccaro, "The Martian Ionosphere as Observed by the Viking Retarding Potential Analyzers," *Journal of Geophysical Research*, Vol. 82, 1977, pages 4351–4363.

6. M.H.G. Zhang, J.G. Luhmann, A.J. Kliore, and J. Kim, "A Post-Pioneer Reassessment of the Martian Dayside Ionosphere as Observed by Radio Occultation Methods," *Journal of Geophysical Research*, Vol. 95, 1990, pages 14,829–14,839.

7. A.J. Kliore, "Radio Occultation Observations of the Ionospheres of Mars and Venus," *Venus and Mars: Atmospheres, Ionospheres, and Solar Wind Interactions*. Geophysical Monograph 66, American Geophysical Union, Washington, D.C., 1992, pages 265–276.

8. M.H.G. Zhang, J.G. Luhmann, and A.J. Kliore, "An Observational Study of the Nightside Ionospheres of Mars and Venus with Radio Occultation Methods, *Journal of Geophysical Research*, Vol. 95, 1990, pages 17,095–17,102.

9. K. Davies, *Ionospheric Radio*, Peter Peregrinus Ltd., London, 1990, Chapter 6.

10. *The ARRL Handbook*, 66th Edition, American Radio Relay League, Newington, Connecticut, 1989, Chapter 22.

11. T.A. Croft and H. Hoogasian, "Exact Ray Calculations in a Quasi-Parabolic Ionosphere with no Magnetic Field," *Radio Science*, Vol. 3, 1968, pages 69–84.

12. BR Communications, *Tactical Chirpcomm (R) Model 4280 Transceiver Product Brochure*, 1249 Innsbruck Drive, Sunnyvale, California 94089, 1991.

13. A. Tan and S.T. Wu, "Post-Viking Martian Ionospheric Model," *Acta Astronomica*, Vol. 28, 1978, pages 187–194.

14. H. Shinagawa and T.E. Cravens, "A One-Dimensional Multispecies Magnetohydrodynamic Model of the Dayside Ionosphere of Mars," *Journal of Geophysical Research*, Vol. 94, 1989, pages 6506–6516.

15. T.E. Cravens, "Ionospheric Models for Venus and Mars," *Venus and Mars: Atmospheres, Ionospheres, and Solar Wind Interactions*, Geophysical Monograph 66, American Geophysical Union, Washington, D.C., 1992, pages 277–288.

16. R.M. Jones and J.J. Stephenson, *A Versatile Three-Dimensional Ray Tracing Computer Program for Radio Waves in the Ionosphere*, Office of Telecommunications Report 75–76, Department of Commerce, Washington, D.C., October 1975.

17. M.H. Reilly and E.L. Strobel, "Efficient Ray Tracing Through a Realistic Ionosphere," *Radio Science*, Vol. 23, 1988, pages 247–256.

18. L.F. McNamara, *The Ionosphere: Communications, Surveillance, and Direction Finding*, Krieger Publishing Company, Malabar, Florida, 1991.

19. D.C. Agarwal, "Electromagnetic Wave Propagation in the Martian Ionosphere, *Institute of Telecommunications Journal*, Vol. 18, New Delhi, India, January 1972, pages 26–28.

20. G.W. Henry and R.C. Petit, "HF Radio Data Communications: CW to CLOVER," *Communications Quarterly*, Spring 1992, pages 11–24.

21. The Department of Defense, *FY 1994 Small Business Innovation Research (SBIR) Program Solicitation 94.1*, Washington, D.C., September 1993, pages ARPA 14.

PRODUCT INFORMATION

Jan Crystals Announces Publication Of Design Guide And Catalog

A new design guide and catalog of frequency control crystals is now available from Jan Crystals, a manufacturer of custom designed quartz crystal components.

The twelve page publication is oriented to electronics designers, engineers, and manufacturers and includes charts, graphs, formulas, and technical information.

The publication is available, by contacting the Design Engineering Department of Jan Crystals, at 813-936-2397.

Arthur J. Stokes, Sr., *N8BN*
Box 398
Hudson, Ohio 44236

A GYRATOR TUNED VLF RECEIVER

This VLF monitoring receiver uses a gyrator circuit to eliminate LC-type tuning

Previously published articles[1,2] have described two different types of VLF receivers used to monitor very low frequency stations for sudden ionospheric disturbances. These effects have been discussed in **References 1** and **2**. In the spring 1993 issue of *Communications Quarterly*, Peter Taylor and I described a VLF receiver with three separately tuned states using 88-mH toroid coils and air variable capacitors. This arrangement worked well but was somewhat cumbersome.

After publication of the spring 1993 article, I received a communication from Phillip Eide, KF6ZZ, suggesting a gyrator circuit for tuning that eliminated the LC type of tuning used in the previous receivers. I immediately started to experiment with the gyrator approach. This article is the result of my work with a gyrator tuned VLF receiver.

Note that the gyrator directly replaces the cumbersome toroids and slug-tuned coils used in earlier receiver circuits. The Q of a gyrator greatly exceeds that of the iron/copper induc-

tances. Furthermore, the "L" value is easily changed by a resistance. With this type of circuit, it's easy to construct a complete SID receiver around a single 89 cent IC package. Another advantage of this new design is that there is only one stage of LC tuned circuit as opposed to several in previous designs.

Gyrator circuits

I'll make no attempt here to get into gyrator theory. It's sufficient to say that a gyrator is a simple circuit topology that superbly simulates an inductor. The theory is well covered by Moschytz.[3] A basic gyrator is illustrated in **Figure 1**.

The inductance is calculated using the following formula:

$$L = \frac{R1 \; R3 \; R5}{R2 \; 1/C4} \text{ for } R2 = R3 \quad L = R1 \; R5 \; C4$$

In this circuit, I have chosen R1, R2, and R3 as 3.3 k. R5 is a series combination of 3.3 k fixed and 10 k variable. With C4 as 0.001 µF, the inductance range of the gyrator is from 43.9 mH to 10.9 mH depending on the setting of the 10 k pot. This, with a parallel capacitance of 0.002 µF, provides a tuning range of 17 to 34 kHz, a useful range for our purpose. These values will vary somewhat due to the tolerance of the components.

Following the tuned input circuit is a variable gain stage with a gain up to 40 dB feeding into a precision rectifier with another gain of 40 dB. The final capacitor, C6, provides a long time constant to smooth the trace of the *Rustrak*
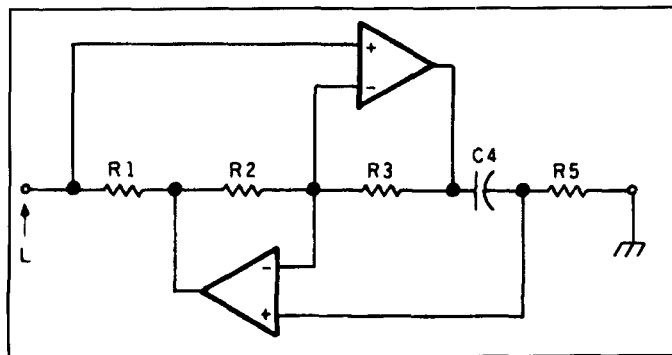

Figure 1. A basic gyrator circuit.
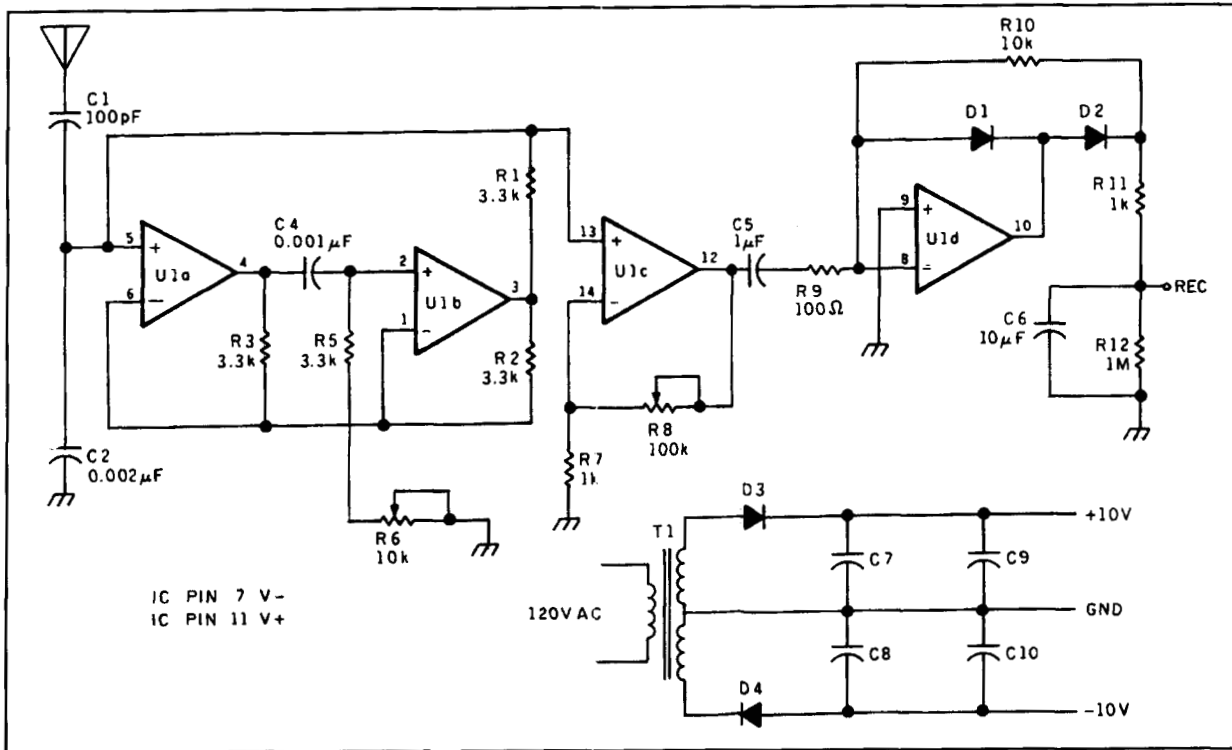
Figure 2. Receiver schematic.

recorder. The schematic is shown in **Figure 2**.

Potentiometer R6 controls the tuning, and overall gain is controlled by R8. A center-tapped 12-volt transformer is used for the bipolar power supply.

Construction

After some experiments with wire-wrapped circuits and a few printed circuit board layouts,

I settled on the arrangement shown in **Figure 3**. There was some tendency to oscillate in a few of the tests, but I eliminated this problem with bypass capacitors on the power supply leads and shielded leads to the pots. Most of the components are Radio Shack parts. A list is provided with **Figure 4**. I recommend the 4136 op amp as best for this purpose. It's well suited for the gyrator and also the final amplifiers.

The circuit board is mounted on standoffs in

Figure 3. Circuit board layout.

Parts List: VLF Receiver

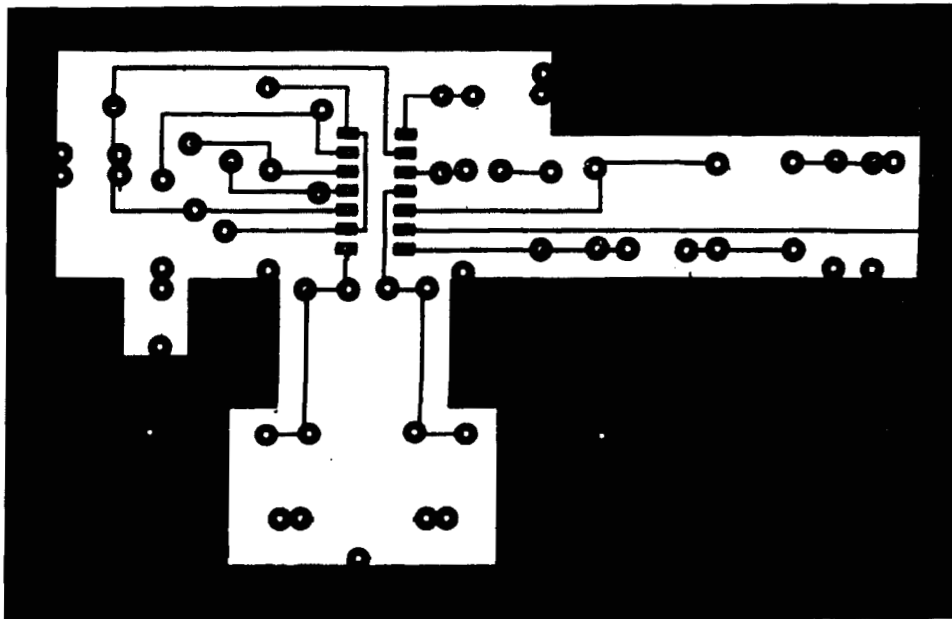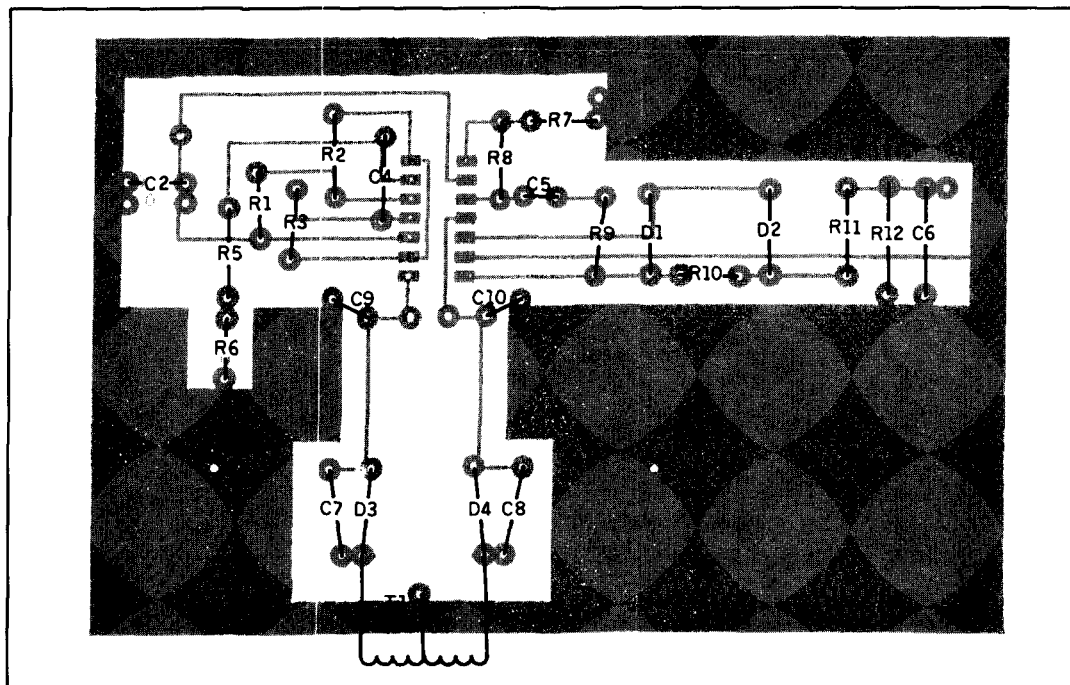|			
R1,R2,R3,R5	3.3-k resistors, 1/4 watt	C4	0.001-µF ceramic capacitor
R7,R11	1-k resistor, 1/4 watt	C5	1-µF electrolytic capacitor
R9	100-ohm resistor, 1/4 watt	C6	10-µF electrolytic capacitor
R10	10-k resistor, 1/4 watt	C7,C8	470-µF electrolytic capacitors
R12	1-meg resistor, 1/4 watt	C9,C10	0.1-µF ceramic capacitors
R6	10-k linear pot	D1,D2	1N34 diodes
R8	100-k linear pot	D3,D4	1N4001 diodes
C1	100 pF ceramic capacitor	U1	4136 integrated circuit chip
C2	Two parallel 0.001 µF ceramic capacitors	T1	12-volt CT transformer
		Metal cabinet and miscellaneous hardware.	

Figure 4. Parts placement diagram.

a metal box with the potentiometers on the front panel. A BNC connector for the antenna and a 1/8-inch jack for the recorder are mounted on the back panel.

Inasmuch as the receiver was intended to be used for SID monitoring on a continual basis, the output is designed to operate a 100 µA Rustrak Recorder. The tuning potentiometer is calibrated on the front panel by feeding a signal from a signal generator through a 100-k resistor to the antenna connector.

Summary

I've described my experiments with a somewhat novel method of tuning a VLF receiver. It has performed very well at my location in northern Ohio on 24 kHz and 21.4 kHz signals from NAA and NSS in Annapolis, Maryland and Cutler, Maine, respectively. Various antennas have been described in other articles. I generally use an insulated 24-foot vertical aluminum tube as an antenna. This is connected to the receiver through the center conductor of a coax line. The shield is grounded to an earth ground. A small diamond shaped loop antenna 14 inches on a side with 100 turns of no. 26 enamelled copper wire has also proved very effective on these stations. The loop was tuned with the proper capacitance in parallel to resonate at each of these frequencies. No ground was needed with the loop.

Acknowledgment

My sincere thanks to Phillip Eide, KF6ZZ, for his basic suggestion of the gyrator and the precision rectifier. Without his encouragement and assistance this project would not have been possible.

Those who wish , may purchase the circuit board for this project from Far Circuits, 18N640 Field Court, Dundee, IL 60118, for $7 plus $1 shipping and handling. ■

REFERENCES
1. Peter Taylor and Arthur Stokes, "Recording Solar Flares Indirectly," Communications Quarterly, Summer 1991, page 29.
2. Peter Taylor, "The Solar Spectrum," Communications Quarterly, Spring 1993, page 51.
3. G.S. Moschytz, Linear Integrated Networks, Van Nostrand Reinhold Company, New York, New York, 1975.

W. J. Byron, *W7DHD*
P.O. Box 2789
Sedona, Arizona 86339

ARCS AND SPARKS: PART 1

How Did They Work?

A rc and spark transmitters were exactly contemporary. The techniques for each developed simultaneously, although the origins of spark technology preceded those of the arc by about five years. Even though the arcs got a rather late start, they later surpassed the sparks both in effectiveness and power output. Arcs eventually operated at a power input of over one million watts; two such transmitters (two duplicate converters in one station) were finished and were operated in France, starting in 1920.[1] They ran up until the mid-thirties.

There were always many more spark transmitters operational than arcs, for three main reasons. First, spark transmitters were easier to construct and to operate. Second, sparks could be operated up into the MF range, whereas arcs could not be operated much above 200 kHz, and that was only possible with a special design. The arcs were limited to VLF generally, particularly the high-powered versions—and that precluded their use by hams. The third, and perhaps the most effective reason, was that commercial arc stations in this country were almost entirely manufactured by the Federal Telegraph Company, which possessed the U.S. rights to the patented Poulsen system. Federal made most of the large arc stations in the Western Hemisphere, and also manufactured the most effective (and the most powerful) transmitters in the world. Federal also was "late in the game." It was perhaps as late as 1910 before the company made much of an impact. A decade later, Federal was out of the arc business; ousted, as all others were, by the imminent appearance of high-vacuum thermionic high-power transmitting tubes.

That the fundamental circuits of arc and spark transmitters are almost exactly alike, belies the fact that their operational characteristics were nearly *diametrically opposite* of each
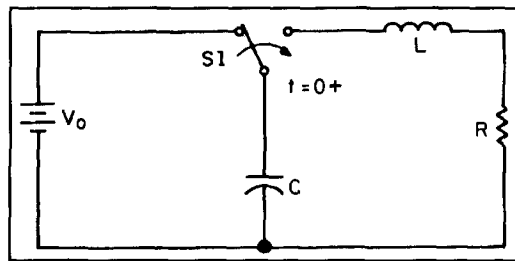

Figure 1. The basic circuit.

other. They didn't look alike, either. Spark transmitters operated from high-voltage, low current supplies, usually AC, and arcs *required* low-voltage high-current DC. The gaps of very large spark transmitters (100+ kW) were always exposed to the atmosphere, and thus were visible and audible; *very* audible. However, the quenched-gap spark transmitters were relatively quiet. They were enclosed for a reason other than for noise-reduction. Arc gaps were always shielded—this was necessary for controlling the atmosphere around the arc, which was a hydrogen-rich "fluid" of one kind or another— and they were very quiet. Arcs always required a magnetic field; sparks did not. Sparks could be keyed by interrupting the primary power source; arcs required an *un*interrupted power source. Spark discharges continued to conduct for many RF cycles; arc gaps were extinguished for a portion of *every* RF cycle. As a matter of fact, there was nothing alike in their operation or appearance except that they both required current to jump across an ionized gap. The following paragraphs describe the main attributes, advantages and disadvantages of each.

Some spark folklore

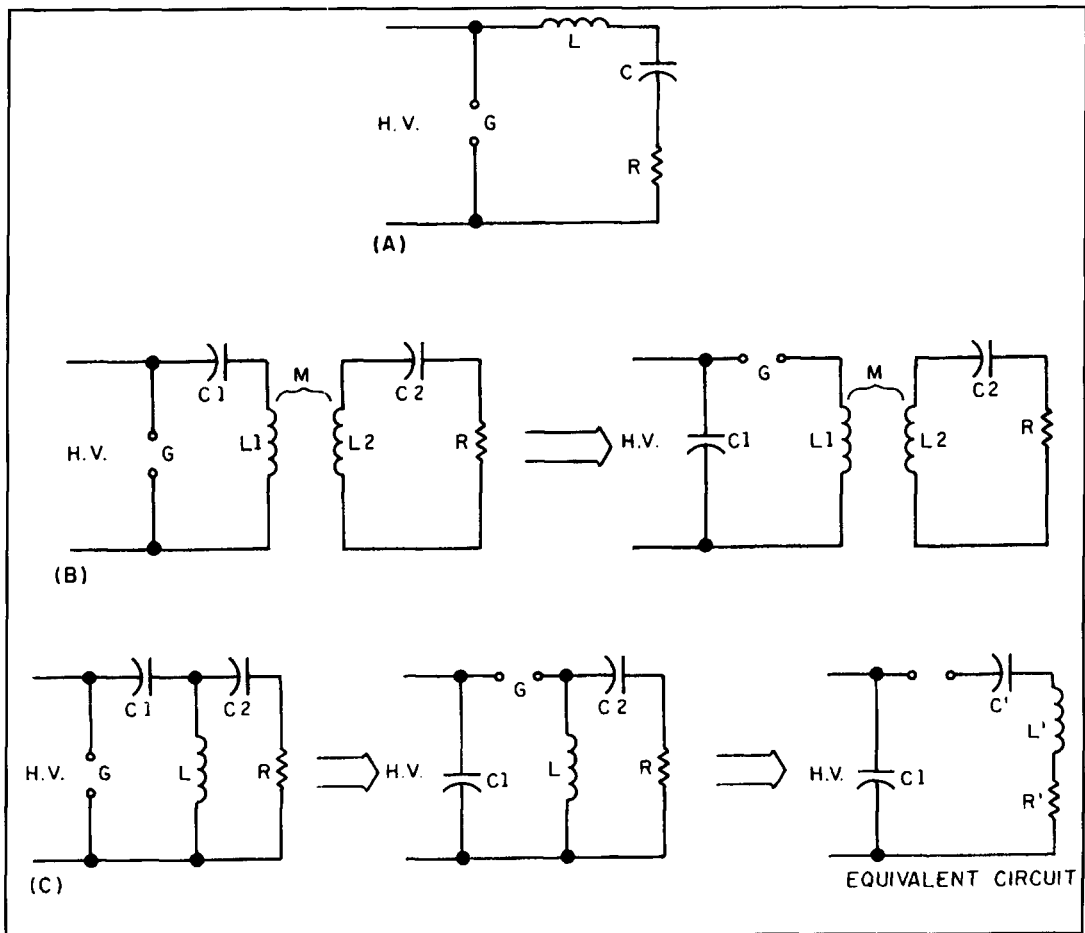Heinrich Hertz started it all. The very first

Figure 2. (A) General circuitry for the demonstration of the basic technique. (B) The Braun transmitter. (C) The Wien transmitter.

transmission of electromagnetic waves in what we call the communications band occurred in the latter part of the 19th century. He used the spark, and it produced waves in the HF/VHF part of the spectrum. Hertz was a first-class physicist. His proof of the existence of electromagnetic waves gave confirmation to James Clerk Maxwell's theory, which had been criticized and rejected by his peers for a decade.

There were several experimenters who continued with the art; Popov and Marconi were among them. Gradually, practical techniques were developed. All of them employed the only means of transmission known at the time: *spark generators*. The most effective impetus, of course, was provided by Marconi and his staff in the first trans-Atlantic transmission. The most successful application of spark methods required operation in the LF/VLF bands. The reasons will become obvious later.

The history itself, when not being presented by professional historians, has been massaged over and over by many of us who are really too young to remember; we just pass on the "old wives' tales." Most of the surviving operators

who ran spark transmitters seriously are now easily in their 80s or 90s. Then there are the slightly younger ones (like me) who at one time or another have tried duplicating spark transmitter operation with the most common device available: the Ford Spark Coil. We were aided and abetted by publications (in the '30s, of course), which still carried residual articles and circuits. This was done even though spark transmitters had been unlawful (for hams) since 1926. It didn't bother us, though! We went on about our business of destroying the family receiver's broadcast reception, by connecting a wire "antenna" to one side of the gap, and a ground rod to the other. None of us had any idea what frequency we were on. We didn't even know the theory of operation. We only knew that there was a circuit drawn in a popular publication of some kind, (possibly in one of Hugo Gernsback's magazines) and we duplicated it—"sort of." Looking back on some of those articles, it's apparent that even the *authors* sometimes didn't really understand, either! Fortunately for us, we were fighting a losing battle. Our signals were so wide that the

energy per kHz was minuscule—particularly because the spark coil itself was incapable of more than a few watts.

Could the arrangement described above have been successful? Yes, it could. The hook-up was a simpler version of the Marconi method used for quite a few years. All the absolutely essential parts were there: the spark gap, an antenna, and a ground connection. We just didn't know enough about the procedure to carry it out properly. Had the antenna been properly loaded with an inductance to bring the total resonance into the broadcast band, had the "ground" been composed of a rectangular spider-web-like counterpoise, had the Ford spark-coil been capable of more power, and had the gap been composed of properly adjusted spherical balls, we could have been in a "heap o' trouble." It would have been heard all over town, and over most of the broadcast band.

The professional equipment was quite sophisticated for the time. The mechanical equipment was very complex even by today's standards. It also was effective, and was the *only* practical means of wireless communication for about a decade. Spark transmitters were used on shipboard, in land stations, and even by the military on both sides of the lines in World War I.[2] Woodrow Wilson's "Fourteen Points"—the terms of German surrender—were transmitted from the United States to Germany (via NSS to POZ, near Berlin) by means of an arc transmitter. POZ, the German receiving station, used the spark. The following paragraphs will explain their operation in considerable detail.

Spark technology

The method of producing an oscillatory wave from a static direct-current source is familiar to all who have studied alternating current circuit analysis.[3,4] The basic circuit appears in **Figure 1**. With switch S_1 in the left-hand position, the capacitor will charge to the potential, V_0. If the switch is thrown to the right at t=0, it connects the charged capacitor to the series combination of R and L.

If R is too large, the resulting current in the circuit is non-oscillatory, and represents special cases—that of "critically damped" to "over-damped." Below a certain value of R, however, the current wave-shape will follow that of an exponentially decaying sine wave of the form:

$$i(t) = \frac{V_0}{\omega_n L \sqrt{1 - \zeta^2}} e^{-\zeta \omega_n t} \sin(\omega_n \sqrt{1 - \zeta^2}\, t) \tag{1}$$

where:

ζ = the damping factor (< 1 here) and,

ω_n = the natural resonant frequency of LCR,

which is an *under-damped* case, and it reduces to a general equation:

$$i(t) = a_0 e^{-a_1 t} \sin(a_2 t) \tag{2}$$

That form will appear later in this section. It is the form of one train in a series of such pulses that were transmitted for years by spark transmitters of all makes. Some later modifications of the shape will also be covered; they represent the ultimate development of spark circuitry.

The production of spark transmitter signals followed the principles described above. The main differences between the circuit in **Figure 1** and those of practical transmitters lie in the methods of charging the capacitor and in the magnitudes of the voltages and currents.

The general circuitry for demonstration of the basic technique appears as **Figure 2A**. We start with the power input designated as "H.V." For this explanation, we'll consider it to be direct current. It doesn't make any difference whether it is AC or DC. Both were used. However, the more successful supplies were always AC.

The spark-gap is designated "G." It is "open" (no spark) at the beginning of each pulse. At the time of voltage application, the capacitor "C" is charging, following the voltage of the source. When the charging potential reaches the breakdown voltage of the gap, a spark is produced. Now the gap is conductive with a low intrinsic resistance, and the rest of the circuit is isolated from the source. The circuit GLCR oscillates as in **Figure 1**, decreasing exponentially to a potential too low to sustain the spark in the gap. The energy in the system is dissipated in R.

This circuit was used for several years. The inductance L (**Figure 2A**) included both the inductance of the antenna and whatever loading inductance was necessary for the wavelength in use. C was the capacitance of the antenna itself, and R represented all the losses in the system—including that of the gap, wiring, ground resistance, and the all-important one, *radiation resistance*. The very early Marconi transmitters were of this type.

Power inputs to these early transmitters were limited. The capacitances of even large antennas of the era were usually on the order of 0.01 μF (later antennas were much larger). The maximum charge was thus limited. The actual power input to the system can be computed by:

$$\text{Power Input} = \delta \frac{1}{2} CV^2 \tag{3}$$

where δ is the charging rate (per second), C is

CONTINUOUS WAVE

SPECTRUM, CW, 33.3 kHz, PULSED ONCE PER msec

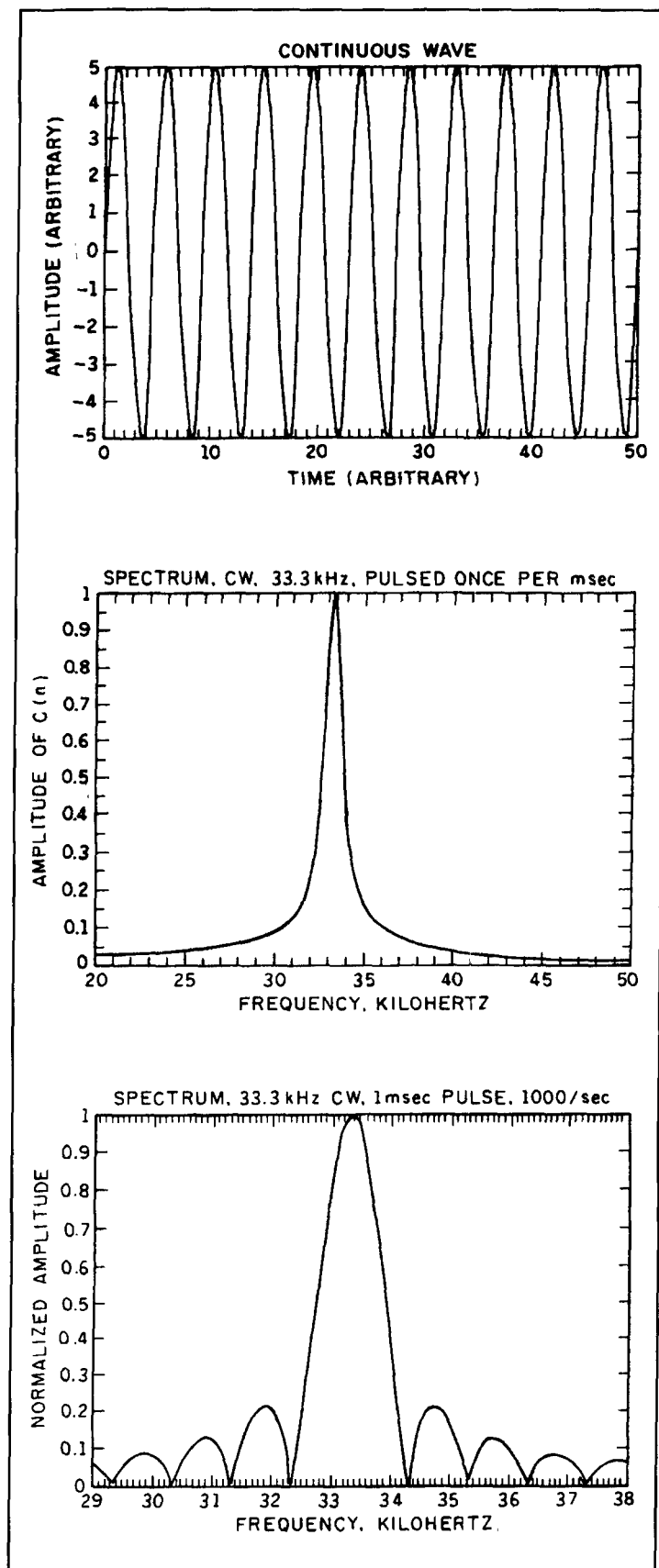**SPECTRUM, 33.3 kHz CW, 1msec PULSE, 1000/sec**

Figure 3. (A) Continuous wave log decrement. (B) CW spectrum, 33.3 kHz, pulsed once per msec. (C) Spectrum, 33.3 kHz 1 msec pulse, 1000/sec.

the capacitance of the antenna in Farads, V is the peak voltage on the capacitor at the instant of gap breakdown, and L and R are as previously defined. It is the small capacitance that limited the power of the early transmitters, even with voltages of 10,000 or more—many, many times more.

The currents in all the circuits of **Figures 1** and **2** decay exponentially, and that created practical problems. If the current doesn't decay, the wave shape is that of a continuous wave, and occupies only one frequency in the spectrum. The shorter the duration, i.e, the higher the damping, the broader the signal. It was recognized early, and the U.S. Department of Commerce (then regulating transmitters) restricted U.S.-licensed transmitters in 1919 to a maximum log decrement of 0.2.

The log decrement (which isrelated to the damping factor in **Equation 1**) is defined as the natural logarithm of the ratio of any cycle amplitude divided by the amplitude of the next occurring cycle:

Log Decrement = ln (A2/A1)   (4)

Figures 3, 4, 5, and **6** are waves of several different log-decrements.

Figures 2B and **2C** are the basic circuits of F. Braun (who also invented the Braun Tube, the precursor of the oscilloscope), and Max Wien. The Braun circuit was the more successful; it was patented in 1898.[4]

All the transmitters of the form of **Figure 2A** suffered power limitation because of the relatively low fixed capacitance of the antenna. Voltages as high as 80,000 were applied to those systems; that voltage appeared on all exposed antenna components. It also challenged the insulators. High-voltage RF insulator research and development was necessary because of this problem—which was common to both arcs and sparks.

Braun solved the power limitation problem very neatly. His basic circuit (**Figure 2B**) is the one that nearly all of the subsequent transmitters used—from the lowest-powered ham transmitters to the very large military, governmental, and commercial transmitters. The latter peaked at approximately 350 kW.

There were other advantages in this system; it reduced some harmonic radiation (when properly adjusted; otherwise it could broaden the signal significantly) and further enabled large energy storage in capacitor C_1. It was critical in adjustment, requiring variable coupling between L_1 and L_2 (a form of what was known as "loose coupling"). Various forms of loose coupling appeared in both transmitters and receivers. The Braun circuit was double-tuned, and could actually produce *two* resonances

simultaneously if mistuned—almost always by overcoupling; hence, various forms of "loose coupling" were developed. Not only did the "double-frequency" reduce the quality of the signal, it was unlawful per U.S. regulations.

M. Wien sought to eliminate the characteristic tuning difficulties in the Braun system by combining the two inductances, L_1 and L_2, and their mutual inductance into one, while retaining large energy storage. The circuit is shown in **Figure 2C**.

Efficiencies of these transmitters usually were between 60 and 85 percent. Of the power ultimately reaching the antenna circuit, less than a tenth actually was radiated; antenna efficiencies were notoriously poor. They still are. Who has room to lay down more than one hundred radials, each one many kilometers in length? Adequate ground systems were, and are, the most important and limiting factors in the performance of Marconi verticals; particularly VLF Marconis. In general the Wien system was the more efficient, but it suffered a disadvantage because it couldn't divorce the log-decrement of the antenna from that of the storage circuit. Although POZ (a *very* famous German station) used the Wien system for a while, it reverted to the Braun system until it was shut down after World War I. One of its successors was at Eilwese, which used a Goldschmidt alternator. The Goldschmidt was of the same general power level as the Alexandersons, and was developed to circumvent the Alexanderson-Fessenden patents.

By the time hams were using the Braun circuit, they were under the "200 meters and down" (1,500 kHz and *up*) dictum. Although the storage capacitance (C_1) enabled much more energy storage than the capacitance of the antenna alone, the power input still followed **Equation 3**. It was difficult, if not impossible, to achieve the one kilowatt maximum allowable input if ordinary 60-Hz line power was used in combination with a single ball-gap or even with the rotary gap. The charging rate δ, had to remain at twice the line frequency (120 Hz). The capacitor could be charged only at that rate.

The circuit was known as the "oscillation circuit", and the coil in amateur transmitters usually was constructed of quarter-inch copper tubing, and consisted of four or five turns about six or eight inches in diameter. It was a "high-C" combination. The larger was "C," the smaller "L" had to be—and that meant a correspondingly more difficult energy transfer to the antenna circuit.

The two graphs, **Figures 7** and **8**, were calculated on the bases of the largest practical storage-capacitor (somewhere around 2800 pF for a four-turn coil) and a frequency of 1,500
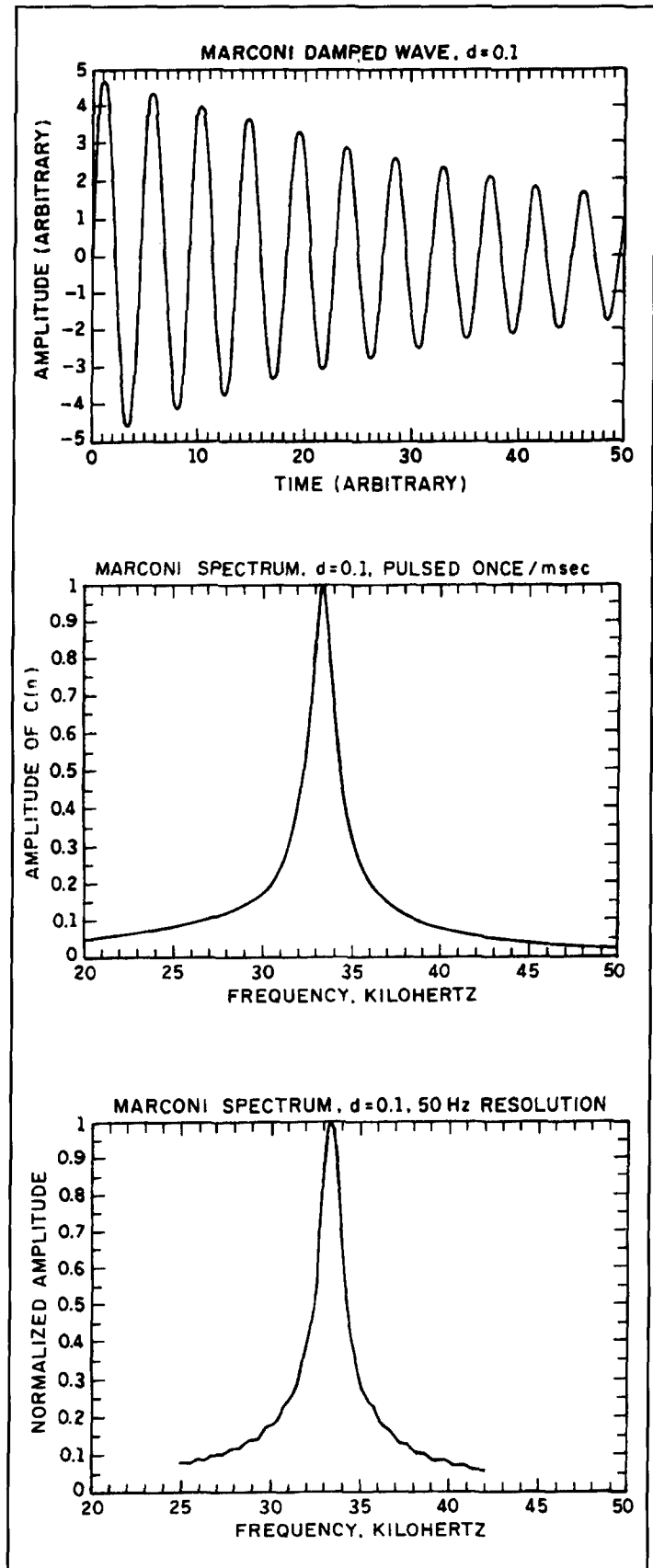

Figure 4. (A) Marconi damped wave, d = 0.1. (B) Marconi spectrum, d = 0.1, pulsed once/sec. (C) Marconi spectrum, d = 0.1, 50-Hz resolution.
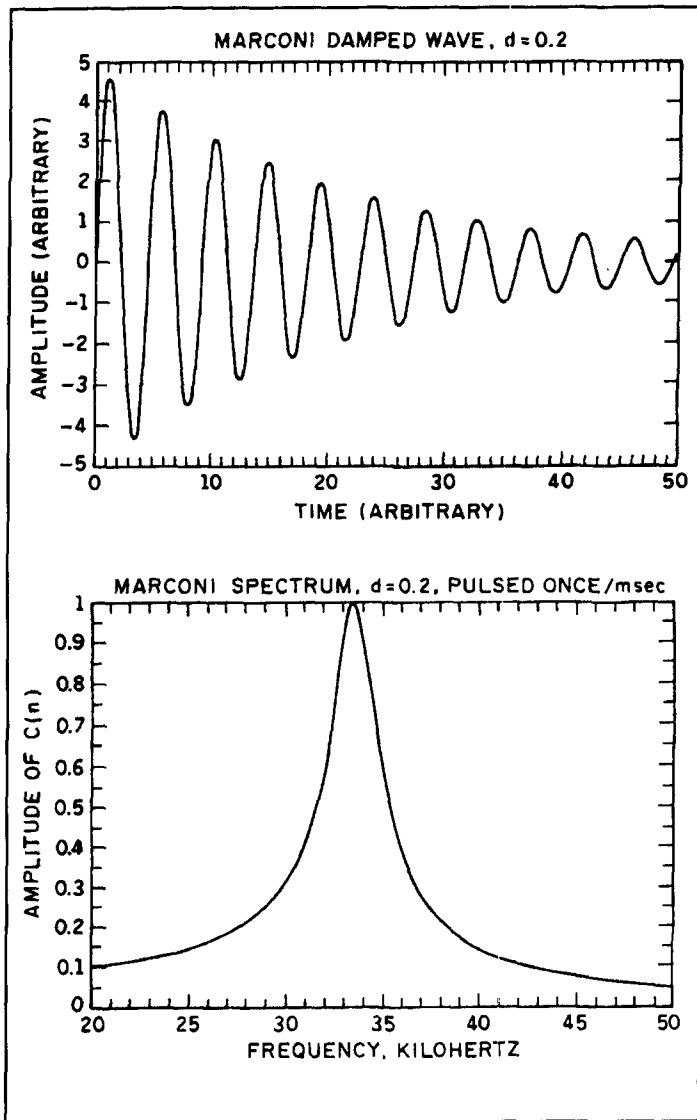
MARCONI DAMPED WAVE, d = 0.2

MARCONI SPECTRUM, d = 0.2, PULSED ONCE/msec

Figure 5. (A) Marconi damped wave, d = 0.2. **(B)** Marconi spectrum, d = 0.2, pulsed once/msec.

kHz (200 meters). These were computed for 2,814 pF. The graphs show very well just how difficult it was to achieve the maximum input for amateur transmitters following the forced band-change. It would require approximately 77,000 volts (peak) for a transmitter using 60-Hz line current, and it would require a peak voltage of 27,000 using a synchronous pulse-frequency of 1000/second. Either would be hard to manage in the typical ham's basement or garage. They are hard to manage safely *anywhere*. It would be an unusual ham who could build circuits—and provide insulation—for 27,000-volt systems, and even more unusual for a 77,000-volt transmitter.

That wasn't the end of the effort. There were some hams who reduced the primary inductance to a minimum—and correspondingly increased the capacitance to a maximum. The

practical culmination of this was the construction of *one-turn* primary inductances, and capacitances of about 10,000 pF. Correspondingly, somewhat lower voltages were required, and some hams were operating this way.

Capacitors in all of the spark transmitters were required to withstand at least 15,000 volts in smaller transmitters, and 80 to 100,000 volts in the largest, while simultaneously containing from 0.1 μF to as much as 1.5 μF capacitance. Remember that the wavelengths usually were *very* long; 12 kilometers was a common wavelength for the very large transmitters, corresponding to a frequency of 25 kHz. They also had to be of high quality, because they were handling large circulating currents.

It would be difficult (at least expensive) even now to produce a single capacitor of 1 μF size at 100,000 volts that would handle several thousand RF amps. On some very large spark transmitters, the initial discharge (first cycle) current was nearly 10,000 amps.

The most common small-transmitter capacitor dielectric was glass. Commercial plate-glass was much used by hams, though other dielectrics were tried—mica being one. Glass was used by many of the commercial-sized units, too. Here, however, the design ranged from actual latter-day Leyden-jar configuration with glass dielectric, to pressurized CO_2 dielectrics in steel pressure-vessels. Both these were used, always in parallel combinations. Oil-immersed glass plate dielectric capacitors were used in latter-day large Marconi transmitters employing the Braun circuit.[5]

Pressurized capacitors can be identified by their flanges and bolts. Internal construction of some of the latter resembled fully-meshed modern-day Jennings vacuum variables. Others were often stacked interleaved parallel plates, insulated from each other and pressurized as were the cylindrical types. High-vacuum techniques were not yet developed to the point necessary for that technology, and techniques that were essential to the soon-to-come high-vacuum thermionic tubes were yet to appear.

It was understood that if the energy per pulse was multiplied by the pulse-rate, higher power would result. It was known also that the ear, as well as the earphones of the day, were most sensitive to frequencies around one kilohertz. Except for an occasional special circumstance (read about station MUU, later) the usual alternator supply frequency was established at 500 Hz, giving the 1-kHz pulse rate. The rotary gap was used in most of these systems, and it was used in *all* of the stations of 100-kW power levels and above.

In those cases where the gap speed was synchronized with the line frequency, a "synchro-

nous gap" was produced. There were hundreds of synchronous ship- and land-based transmitters in use. Most were manufactured by Marconi's companies. Fessenden (and his company NESCO) had several in service, too. The gap rotor was on the generator shaft, assuring synchronism. The note of the signal (actually a sort of "interrupted CW") was therefore one kilohertz, as was δ.

The electric utility system frequency in the United States largely was fixed at 60 Hz. If an amateur wished to use a rotary gap, he was forced to accept either a synchronous pulse-rate of 120 per second (if he were wealthy enough to obtain a synchronous motor) or to use an ordinary motor of relatively high speed. This type of motor invariably produced a "warbling" tone because it would be passing in and out of synchronism continuously. It was a non-synchronous combination. Most hams operated in just this mode.

They also tried to take advantage of increased pulse-rate within the 60-Hz supply restraint. The rationale can be explained partially by **Figure 9**. One cycle of a 60-Hz sinusoid is the solid curve. The hypothetical rotor has ten points, giving a "tooth passage" every 36 degrees in a synchronous sense. They are first labeled A through K, unprimed. K actually represents the "A" of the next cycle. By adding 18 degrees to each point, we have the primed points, A' through J', in which a K' (not shown) would be the A' of the next cycle. The dotted zone is the area in which no spark is possible. The gap is too large, or the voltage too low; either is correct.

No spark is possible at points A, F, and K. However, it is possible at B, C, D, and E, and on the other half-cycle, at G, H, I, and J. The power input for the first case would be four times the energy at the voltage of B, say, plus four times that of point C—all points multiplied by the appropriate δ, of course.

If the phase is shifted just 18 degrees, sparks are possible at all primed points—including two (C' and H') at the highest voltage of the sine. For any other phase-angle, the resultant power input will drift between these values. Either case would theoretically increase the power input over a synchronous pulsing at points C' and H'. And it *did*. Hams were able to achieve the maximum 1-kW input with 60-Hz power. The problem was that the sparks did not occur exactly when they were supposed to, and the voltages did not follow the nice, clean sine wave. In fact, they occurred at considerably lower values than are indicated in **Figure 9**. The increase in power input and output was evident, but the effectiveness of the signal wasn't. It actually decreased.
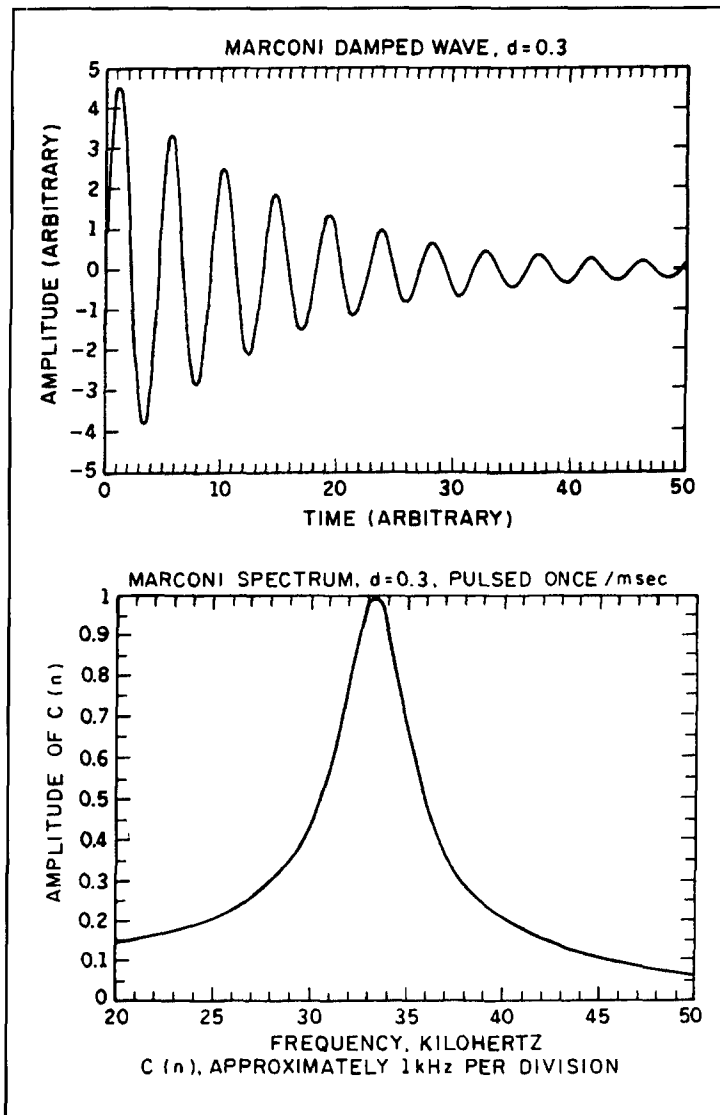
Hams of the day had a habit of turning the

Figure 6. (A) Marconi damped wave, d = 0.3. (B) Marconi spectrum, d = 0.3, pulsed once/msec.

gap motor off during the "dah-dit-dah" on sign-over, allowing the spark-rate to drift down and through true synchronism at 120 Hz. The signal readability actually increased even though the measured power (real, by the way) was decreasing. What was happening? This question was the subject of at least one *QST* article. The authors' conclusions were essentially correct, but for the wrong reasons.

The transformers in use then would produce many thousands of volts; but a transformer of say, 25,000 volts rms, would deliver only 40 mA at 1 kW! They also possessed many turns of fine wire (necessary for the voltage increase), and much higher insulation requirements. Comparatively large spaces brought about by the insulation requirement also meant decreased flux-linkage. Those transformers suffered from both high ohmic wire resistance and leakage reactance.
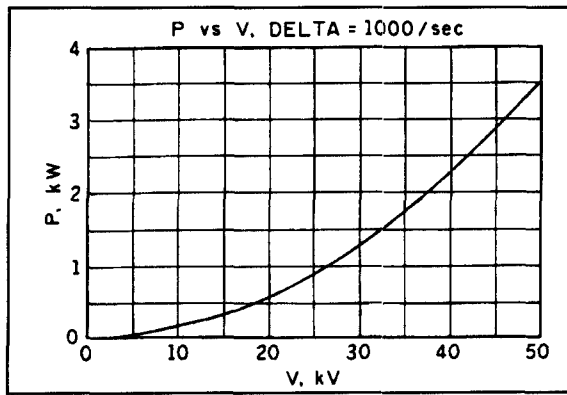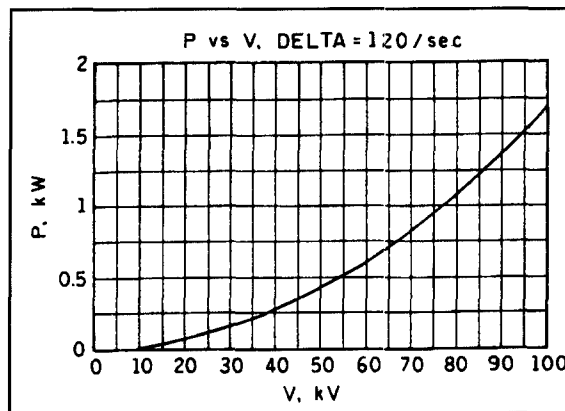
Figure 7. P versus V, delta = 1000/sec.

Figure 8. P versus V, delta = 120/sec.

It is not unreasonable to assume that a 10 percent voltage drop occurs in such a transformer when fully loaded by a resistance. One can estimate the combined effects of leakage reactance and resistance (they will appear as a combined resistance) by dividing the voltage change by the current change; 2500 V/0.04 A. The resultant resistance, 62,500 ohms, is now the "source resistance" of the transformer. Every pulse in spark transmitters is preceded by a charging cycle. The capacitor must be charged from the source. The voltage on the capacitor follows the usual inverse exponential, $V_c = V_0 (1 - e^{-t/\tau})$, where τ is the time-constant, RC, R is the internal dynamic resistance of the transformer, and C is that of the RF primary circuit, in Farads. It is 0.625 millisecond for C = 10,000 pF and R = 62,500 ohms. Referring to **Figure 9**, it is seen that there are only 1.67 ms between pulses. It is also true that three "τs" (1.875 ms) are necessary for a capacitor to assume 90 percent of its full charge. In short, the capacitor *never* gets charged. There are enough pulses, though, to achieve the kilowatt average input, but the peak power never reaches the equivalent power of fully-charged capacitors. It is exacerbated by the sinusoidal

shape of the source voltage—or at least it *tries* to be sinusoidal. The only loads those transformers saw were charging surges. They were grossly nonlinear loads. It was difficult duty.

Detectors used in the reception of spark signals were *envelope* detectors. Even regenerative detectors on the verge of oscillation (the most sensitive state) were envelope detectors. They all were sensitive to p.e.p.—*peak envelope power*. Sound familiar? As the spark rate decreased at the signover, more time was available for the capacitor to charge, and its p.e.p. increased. There is a modern lesson here, by the way, for SSB operators infatuated with excessive use of processors. Even product detectors are envelope detectors in the general sense.

Large commercial installations were plagued by the same problems. There were methods of resonating the source (at 1 kHz) simultaneously with the charging voltages. It is an interesting concept, but is not appropriate for systems with 60-Hz supplies. A good description of the principle can be found in **Reference 5**.

The "quenched gap" was devised to prevent "arcing" at the gap. The power sources for transmitters ranged from fractions of kilowatts to hundreds of kilowatts. Once a ball-gap became ionized, it had a tendency to continue conducting long after the oscillating current in the circuit had ceased. Sometimes it did not quench at all. The quenched gap—one of the solutions—was a static device. It was composed of a number of parallel circular plates, each insulated from the next by an annular ring. Eight or ten such plates were common. The overall gap therefore consisted of several shorter ones. One could select just how many stages he desired by tapping onto any particular stage. It had the property of discouraging arcing, but it was limited in its heat-handling capability, so it was used in transmitters of relatively low power. It was quiet in operation, and where there was a choice (switch-selected) mostly in small 2-kW transmitters, like the American Marconi P-4, it was the preferred gap. Quenched gaps were used in some larger transmitters (mostly in Germany) but they were limited in heat-dissipating capability, even with forced-air cooling.

All of the large commercial transmitters used the rotary gap. When the gap is required to pass thousands of amperes at tens of thousands of volts, heat becomes a severe problem. So does noise and (probably) the production of ozone. Rotary gap discharges in the very large commercial transmitters could be heard for miles, and the radio rooms of ship-board transmitters were frequently sound-proofed, particularly on passenger liners.[6]

The motion of the rotor not only quenched the gap, it cooled the rotor points. The stator
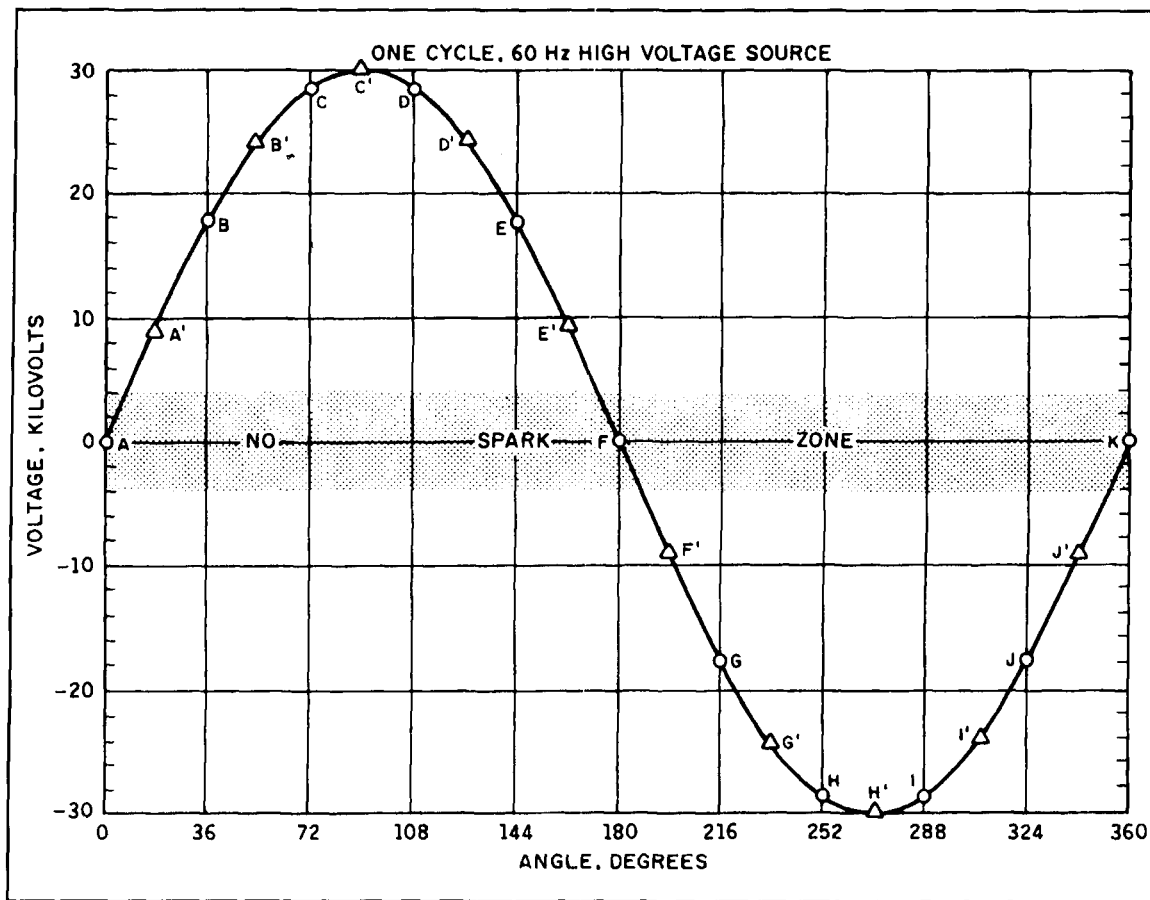
Figure 9. Partial explanation of rationale behind taking advantage of increased pulse-rate within the 60-Hz supply restraint.

points on very large rotaries often were water-cooled; Fessenden used self-contained (and insulated) thermal siphons with small tanks.

Adjustment of the Braun circuit was somewhat tedious. There was a transfer of the energy from the oscillation circuit via the "oscillation transformer"—actually L_1 and L_2 in **Figure 2B**, to the secondary circuit, and the two currents are presented in **Figures 10** and **11**. Note that the antenna current still decays exponentially with the log-decrement of the particular antenna and ground system (beyond t = 20) in the figure. The abrupt rise in antenna current (at t = 0+) produced by the circuit in **Figure 2A** is missing, and its absence contributes to a narrower signal width.

It was easy to over-couple the Braun "oscillation transformer," and serious degradation of the signal always resulted. It is very easy to see how this happened. All of the total energy per pulse (at t = 0) is contained in the primary circuit CGL. That energy is transferred to the secondary system, starting with the first swing of the primary current. If the coupling is too close, the field established by the current in the secondary system is reflected back into the prima-

ry system with enough energy to reignite the spark gap, and the process continues back-and-forth until the remaining energy is too low to reignite the gap.

A simulation of the effect appears as **Figures 12** and **13**. The trick was to adjust the coupling "just so"—far enough apart so that there would not be enough energy fed back to restart the spark, yet close enough to maintain good efficiency. Observation of the currents in real time could occur only with the aid of laboratory instruments; in practice, the operators used what were for the time conventional instruments. "Hot-wire" ammeters, d'Arsonval (dynamic) wattmeters, wave-meters, and regular voltmeters sufficed.

Notice that the currents appear to be "modulated." Indeed, that is just what is happening. **Figure 14** shows the currents in **Figures 11** and **12** recalculated to reveal the decaying sinusoidal (and cosinusoidal) envelopes. Note that when one is at a maximum, the other is at a minimum—right down the line. If the actual signal frequency were in the audio range, it would appear to have "beat-notes"—a not-so-subtle indication of a "second frequency" that
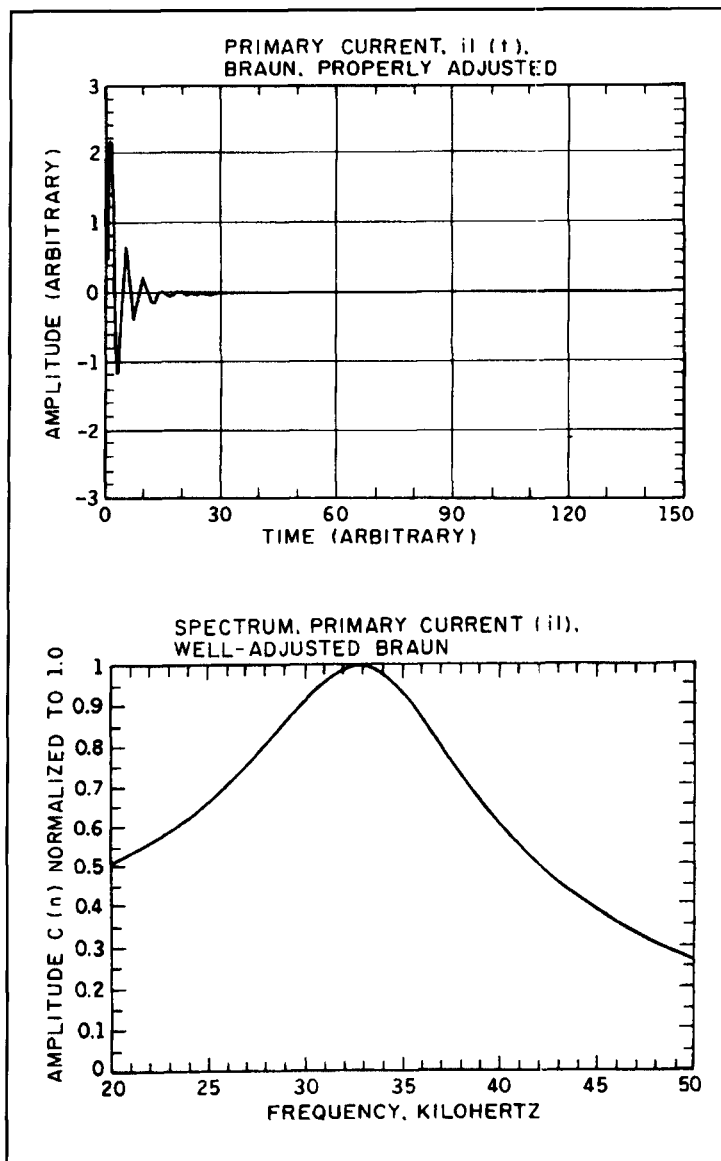
Figure 10. (A) Primary current, i1(t), Braun, properly adjusted. (B)
Spectrum of the primary current (i1), well adjusted.

was mentioned earlier. The envelopes reflect the combined log-decrement.

The figures show both primary and secondary currents to have a common log-decrement. Zenneck goes into great detail about the differences of decrement between coupled circuits both with and without a spark-gap.[7] If the primary gap quenches by giving up all of its energy to the secondary, and if the coupling is loose enough so there is no reignition of the spark, the secondary—which is the antenna circuit in this discussion—behaves as if there were no primary circuit at all, and it decays by its own unique decrement. Coupled decrements will not be addressed.

Marconi designed and operated very advanced spark transmitters. There were others,

too, but his are perhaps the most thoroughly documented. Some very exotic electrical designs and equipment were required to overcome certain problems. One of his transmitters at Carnarvon, Wales (callsign MUU), was a 300-kW spark transmitter. Very thorough discussion of the design of this transmitter can be found in the AWA Review.[5] A few characteristics follow.

None of the damped wave-shapes presented later on show a log-decrement of less than 0.1. MUU and several others, however, emitted a damped wave of exceedingly small decrement—0.029 being rather representative of those systems. Such decrements were completely out of the reach of hams. They occurred in large, expensive, professional systems with very good ground setups, and they were for waves of approximately 10,000 meters. Both were beyond the pocketbooks, the knowledge, and later, the law, for hams ("200 meters and down"). The Carnarvon antenna (for 11,000 meters) was a trapezoidal flat-top, varying from 450 to 600 feet wide by 3100 feet long. It was supported at a height of 400 feet by ten masts. Its feedpoint resistance was 1.4 ohms, and the capacitance of the antenna was 0.039 μF.[5] Its spectrum would no doubt show a lot of the same "fine structure" present in the CW spectrum (Figure 3).

An alternator supplied 150-Hz power to the transmitter. At one spark at each of the positive and negative peaks, the "tone" (spark-rate or δ) was 300 Hz. It wasn't an arbitrary selection. With a decrement of about 0.03, over seventy oscillations in the antenna current were required to reduce it to 10 percent of its initial value. Thus, there was always a remaining oscillating current in the antenna circuit—and it was present at the next primary pulse. Even though the gap was synchronized at the alternator frequency, there was still "jitter" inherent in all those systems, so synchronizing the primary initial current rise to the desired 90 degree lag with respect to the antenna current was extremely difficult. It was impossible to guarantee. Rather than abuse transformer, gap, and other components, a synchronous pulse rate was selected to assure that the remaining oscillating current in the antenna was reduced to about 10 percent. In this way, the initial pulse phase could vary from +90 degrees to –90 degrees without damaging components. Remember that the initial current in the primary transformer was 9,000 amperes.

The maximum capacitor voltage was 35,000. Primary circuit capacitance was 1.5 μF. Even at 11,000 meters (27.3 kHz) the total primary inductance could not exceed 22.7 *microhenries*. That meant that "interconnecting" conductors could not add much inductance. They were of

copper, quarter-inch thick, 24 inches wide, and several feet long. The primary of the oscillation transformer was three turns, with an inner diameter of about 4 feet. The "turns" were composed of a sheathed 12-inch diameter cable of Litz-wire. The size of components—both electrical and physical—by this time had limited further power escalation of spark transmitters.

There were about 450 spark transmitters of various sizes in operation in the U.S. at the height of their era. This compared with perhaps a dozen large arcs, most of which were in the hands of both the U.S. Navy and Federal's commercial enterprises. The spectrum was crowded. Operation with the minimum log-decrement was sought by everyone, and there were several different practical methods employed by the operators for determining it, which, as stated above, was restricted to 0.2 after 1919. They faced even *another* restriction, also in the Department of Commerce regulations—affecting both arcs and sparks. If the transmitted signal contained more than one frequency (measurable by a sensitive wave-meter), the second (unwanted) signal could not exceed twenty percent of the main signal. An exception was in the transmission of emergency and emergency-related signals. Operation with less-than-lawful signals, by the way, is still permitted under those circumstances.

The problem of crowding was compounded by the fact that the common detectors were envelope detectors, still using tuned circuits of relatively low selectivity. Crystals (Galena, Silicon Carbide, and some others—with "cat's whiskers") as well as electrolytic and thermionic diodes (Fleming Valves) were used most often. Heterodyning detectors were of little use in receiving spark signals, and diode-detectors were equally useless in most cases in the detection of pure CW (then usually arc) signals. All of these detectors were passive. They had no signal amplification. The *transmitters* had to supply the power necessary to actuate operators' earphones—operators who were sometimes 3,000 miles or more away from the transmitter. Improvements brought about by the development of both "gassy" (i.e., De Forest) and high-vacuum triodes (hence regenerative detectors, too) did enable signal enhancement, but much later. They appeared in the latter part of the spark transmitter era—principally in its last decade. They were better suited to the reception of CW signals.

Operation of large spark transmitters in the near vicinity of receiving posts using crystal detectors was an incompatible combination. During the several years of development—principally to higher powers—occasional "welding" of cat's-whiskers to the crystals occurred, even with the T/R switch in the "receive" position.
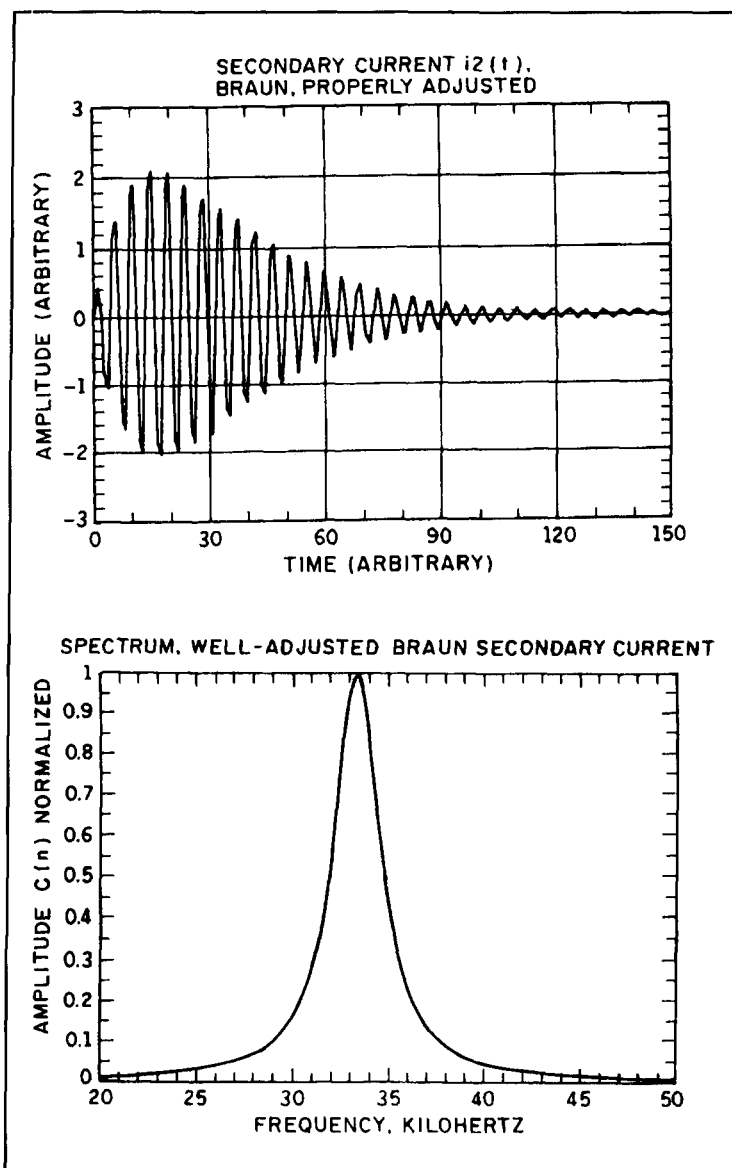
Figure 11. (A) Secondary current I2(t), Braun, properly adjusted. (B) Spectrum of the well-adjusted Braun secondary current.

That is just *one* of the reasons for high-powered stations placing their receiving locations many miles away from the transmitters. Keying was accomplished by telephone lines—"hard-wired." That was true of all large stations, no matter what type of equipment was used. It was a more efficient use of personnel and facilities, such as telephone and telegraph, to have traffic nets serviced from one location. Often those operators were controlling several transmitters simultaneously on different frequencies.

Some analyses

Fortunately, spark-generated wave-shapes are easy to reconstruct. The correctly-coupled and over-coupled Braun transmitter cases faithfully

OVERCOUPLED BRAUN PRIMARY CURRENT (i1)

SPECTRUM, OVERCOUPLED BRAUN PRIMARY CURRENT
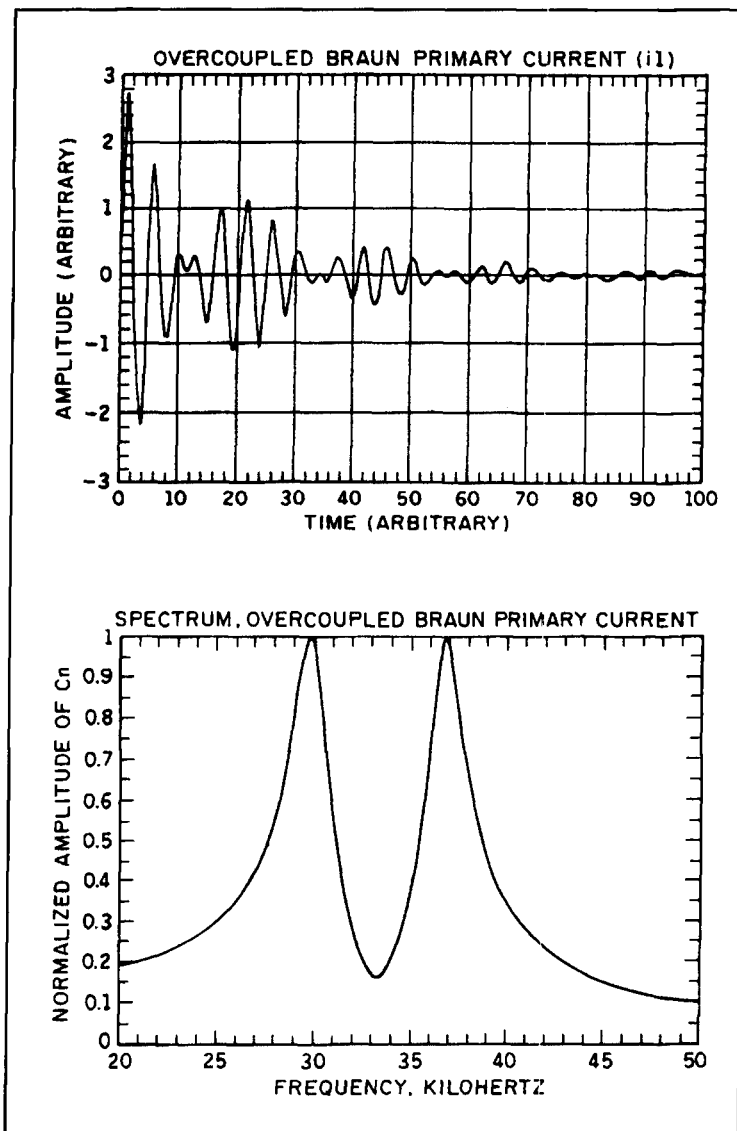
Figure 12. (A) Overcoupled Braun primary current (i1). (B) Spectrum of the overcoupled Braun primary current.

reproduce the shapes of published wave-shapes from Zenneck, dated 1915. The waves simulated in this paper are adequate for the purpose.

The preceding paragraphs have presented the shapes of virtually all the transmitted waves and some that are not, like the Braun primary currents. They are paired with their frequency spectra, which in turn are related to Federal regulation of the log-decrement of spark signals.

What were these signals like? Since there are few people alive today who actually have operated spark transmitters, we must use other means to evaluate them. One good way is to calculate their frequency spectra. It also removes subjective bias from the evaluations.

We are fortunate to have many tools for bandwidth investigation. The most valuable analytical tool for use in this case is Fourier

Analysis, because we have clear analytical expressions with which to work.

They are:

$$i(t) = a_0 \, (\sin(a_1 t)) \, e^{-a2\,t}, \tag{5}$$

$$i(t) = a_0 \, e^{-a1\,t} \, (\sin(a_2 t)) \, (1 - e^{-a8\,t}), \tag{6}$$

$$i(t) = a_0 \, e^{-a1t} \, (\sin(a_2 t)) \, (\sin(a_3 t)) \, (1 - e^{-a4\,t}), \tag{7}$$

$$i(t) = a_0 \, e^{-a1\,t} \, (\sin(a_2 t)) \, (\cos(a_3 t)) \, (1 - e^{-a4\,t}), \tag{8}$$

and,

$$i(t) = a_0 \, (\sin(a_1 t)) \tag{9}$$

Equation 5 is identical to **Equation 2**, and is repeated for convenience. It describes both the Marconi wave-shape and the Braun primary current when properly coupled. **Equation 6** is of the form of the Braun Secondary (antenna) Current, properly coupled. **Equations 7** and **8** describe the overcoupled Braun secondary and primary currents, respectively. **Equation 9** is that for a steady-state carrier. One spectrum for a pulsed carrier is included. The programs all yield line-amplitudes (C_n), but the spectra are easier to judge if the envelopes are shown— rather than the individual lines—so most of them show the envelopes.

They all lend themselves well to numerical integration, and that is how the BASIC programs—generated for this paper—did it. The advent of fast personal computers, particularly those with math coprocessors operating on compiled programs, makes it relatively easy. All of the spectra were calculated with a minimum of 1,500 products per point, and most of them encompassed 50+ points (frequencies in this case). For those interested, uncompiled BASIC operation on the (CW) integral required 14 seconds per point; when compiled and with the coprocessor invoked, it probably would require less than a second. All these calculations were performed (uncompiled) by an 80386 microprocessor operating at 25 MHz.

Some standardization was required to put all these spectra on the same footing. I chose a spark operating frequency of 33.3 kHz, for one of the bases. All of the curves reflect that frequency. There's nothing special about 33.3 kHz. It was in the right range, and was selected before this paper was begun. One time unit in all the figures represents 6.6667 microseconds. The repetition rate (δ) was assumed to be 1000/second. That repetition rate is employed in *all* the figures, including the one-ms pulsed CW case. When the integration was first performed on a CW carrier pulsed every millisecond, it did *not* reveal a single thin line; it had

finite width. This solution resulted from my taking 1-kHz "cuts" at the functions; it did not have the required resolution. The two spectra depicting pulsed carriers (**Figures 3** and **15**) were produced by specifying 0.05 kHz steps. Notice the narrowing of the signal width as the pulse length is increased. The spark-generated signals of d = 0.2 or more did not show any "fine structure"—they were just plain *broad*!

It is important to remember that the programs automatically start the function from zero at t=0+, and at the end of the pulse, if it has not damped out, is similarly dropped again to zero—something that the operating transmitter gap cannot do. The spectra therefore must contain the frequencies necessary to define both the leading- and trailing-edges of the pulses. They have been normalized to 1.0 maximum amplitude for easy comparison. Identification and further descriptions are found within the figure captions.

The preceding paragraphs have been dedicated to the apparatus and techniques of 70 to 90 years ago. Are there any similar signals transmitted today? Certainly! A recent issue of *Communications Quarterly* contains a figure showing the shape of a modern LORAN-C pulse.[8] It is reproduced here as **Figure 16A**. Notice that it strongly resembles the shape of a "well-adjusted" Braun antenna current. The pulse simulation (**Figure 16B**) is not exact, but is so similar that the analysis probably can be trusted. The equation for the simulated pulse is:

$$i(t) = a_0 (e^{-a_1 t}) (\sin(a_2 t)) (1 - e^{-a_2 t_3})\qquad(10)$$

This equation may look like one of the others, but there is a t^2 in the exponent of the last term. None of the others have it; it was necessary in order to simulate the LORAN pulse.

These are different times, and the LORAN-C system exists for an entirely different purpose. The pulse has been subjected to the same exercises as have all the others in this article—for curiosity's sake. It would have been "just inside" what the law would have allowed; its log-decrement is about 0.15, and its spectrum is broader than that of the well-adjusted Braun antenna current, which for this paper (and usually in practice, too) had a smaller decrement.

The LORAN pulse carrier frequency has been translated to our 33.3 kHz "standard," and is pulsed at the 1 kHz rate. Its spectrum can be laid right over that of a Braun secondary current pulse of log-decrement 0.15. There is a small difference, however. The leading edge rise of the LORAN pulse is slightly gentler than that of the Braun. It is reflected in the spectrum, which shows lower-amplitude frequency components away from the carrier. Otherwise they are practically identical.
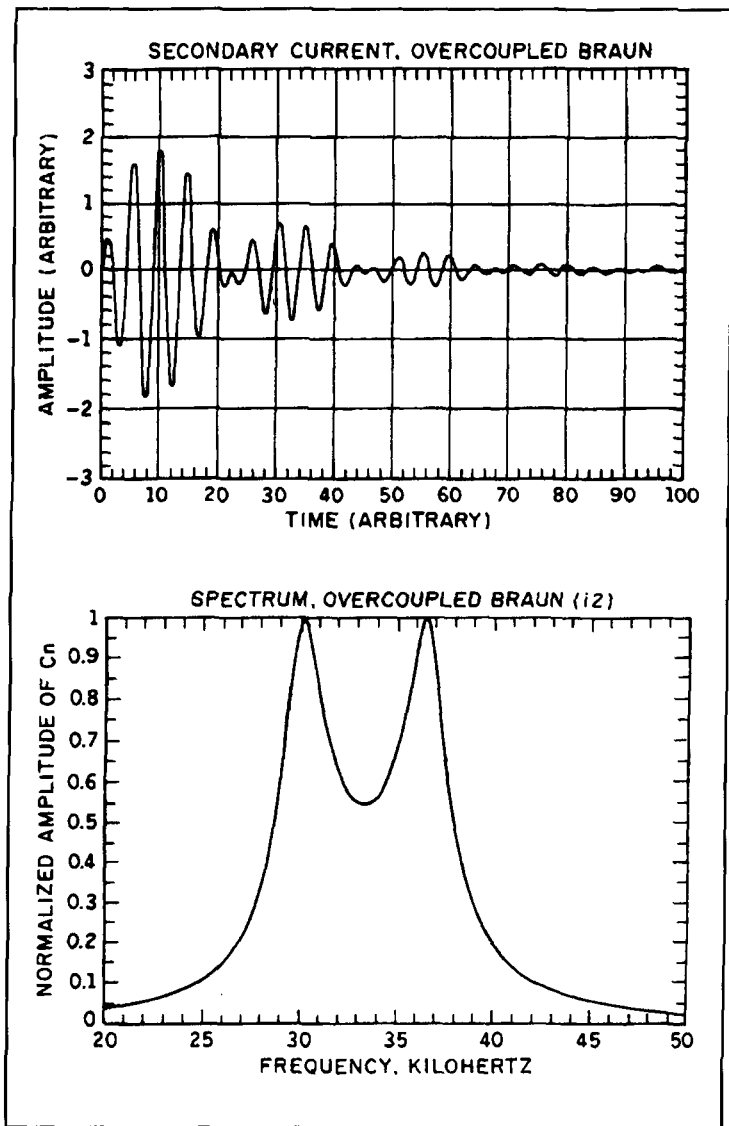
Figure 13. (A) Secondary current, overcoupled Braun. (B) Spectrum, overcoupled Braun (i2).

One last pulse-spectrum was calculated. To demonstrate just how much of the spectral width results from the steepness of both the leading and trailing edges of the pulse, a waveshape that has gentle slopes on both ends was synthesized. Its defining equation is:

$$i(t) = 5(\sin(1.3963\ t))(\sin(2.068E\text{-}2\ t))^2\qquad(11)$$

Figures 17A and **17B** are the wave and its spectrum. It is the narrowest spectrum reproduced in this article—except for the two "CW" spectra. There are very low-level peaks on either side of the carrier, however, and they occur at multiples of the "modulation" frequency (the simulation required that the carrier be modulated sinusoidally, plus have a repetition rate of the same base frequency—1 kHz). They
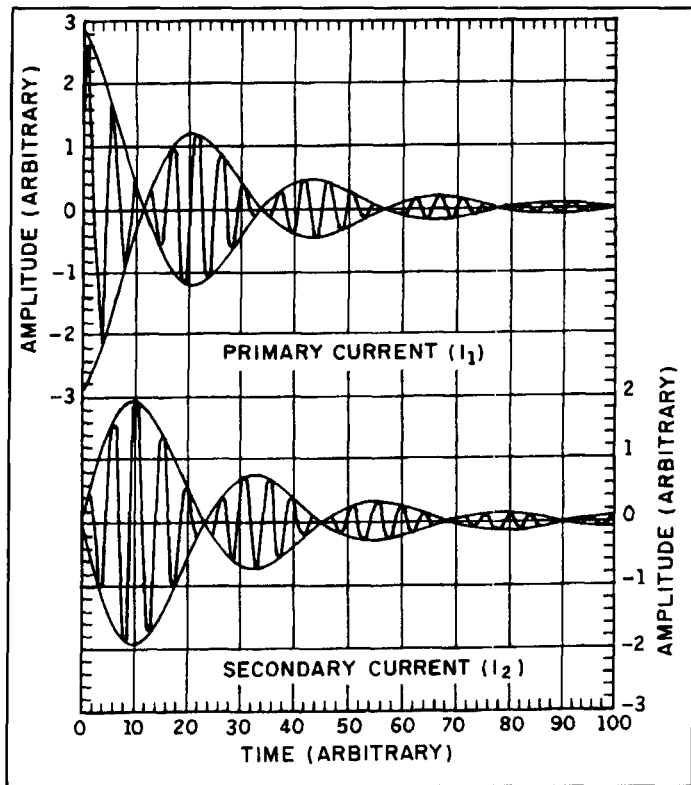
Figure 14. Currents in Figures 12 and 13 recalculated to reveal the decaying sinusoidal envelopes.
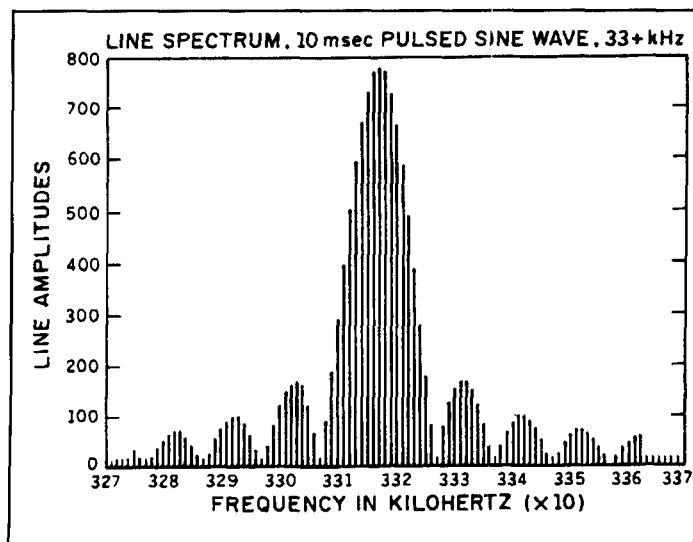
Figure 15. Line spectrum, 10 msec pulsed sine wave, 33+ kHz.

would be much more prominent if the pulse were to last for several cycles of the modulation envelope. The pulse was engineered to use only one full cycle of the envelope, and many cycles of the 33.3 kHz "carrier," which is therefore much more prominent in the spectrum. Otherwise there would be *three* prominent peaks: one for the carrier, and two for the mod-

ulation—one appearing on each side of the carrier peak. We all have seen that type of spectrum elsewhere.

The BASIC program

A technique often used for determining the Fourier coefficients of unknown functions (such as one that may appear on an oscilloscope face) was adapted for these waves. Where one normally must measure the ordinates of figures recovered from an oscilloscope, we merely calculate the ordinates; we already have the analytical expressions. It is a simple program, containing merely 23 lines, of which only 15 are really necessary to determine the coefficients. It appears as **Figure 18**.

Absolute values of the coefficients are deemed unimportant for this study. All spectra have been normalized to 1.0 maximum, and therefore the original values of the coefficients disappear in the normalization. The magnitudes of A_n and B_n (and thus C_n) are first computed for their peak values (a few points—here, around 33.3 kHz), and then the normalizing factor—unique for each spectrum—is entered as a divisor on Line 171. It is the "/10.01" in **Figure 18**. It is valid only for that spectrum.

The coefficients A_n and B_n in the program descended from the equations contained in any one of several *Circuit Analysis* textbooks.[3,9] Their defining expressions are:

$$A_n = \frac{1}{\pi} \sum_0^{2\pi} y \sin nx\, \Delta x, \qquad (12)$$

$$B_n = \frac{1}{\pi} \sum_0^{2\pi} y \cos nx\, \Delta x, \text{ and} \qquad (13)$$

$$|C_n| = \sqrt{A_n^2 + B_n^2}, \text{ since} \qquad (14)$$

$$C_n = A_n + jB_n \qquad (15)$$

A_0, the DC offset present in an asymmetrical wave-form, is missing in these calculations because of inherent symmetry. C_n and n are the numbers tabulated in the program output. They are labeled "CN" and "N."

The BASIC program is not universal. Each spectrum requires its describing function be entered manually into the program, along with certain program modifications. As it appears in **Figure 18**, it is set up to calculate the coefficients for a CW wave. It is programmed to evaluate those coefficients *every 50 Hz* (see Line 172: N = N +.05). Most of the spectra

were calculated with N set equal to 1.0 or 0.5. It's your choice.

The spectrum for the 1-ms pulsed CW (see **Figure 3B**) reveals much "fine structure." The spectrum of the wave with d=0.1 was generataed using 50-Hz resolution to see whether there would be similar "fine structure." **Figure 4B** shows that it does, indeed, though it is of limited significance in the overall appearance. This version, by the way, calculated 340 points—at 14 seconds per point. It was done to generate confidence in the spectra calculated for all the other waves in this paper. There is very little significant difference in the spectrum as calculated with 50-Hz resolution compared with that using only 1-kHz resolution—but there *is* some fine detail.

All the wave-shapes were computer-generated, and no doubt are "cleaner" than those generated by an operating transmitter. We can stipulate that they all start precisely at zero when t = 0, and end when we want them to. Practical transmitters *all* suffered "jitter;" an uncertainty in the timing of the start of a spark. Arc transmitters exhibited the same phenomenon (for a different reason) that will be revealed when arc technology is discussed in Part 2.

Marconi's CW spark transmitter

The heading seems contradictory; the bulk of this article has covered most of the methods used in spark technology, with emphasis on the damped-wave character of spark transmitters. There was an attempt, partly successful, to produce a transmitter using spark technology, for continuous waves.

Marconi and his very successful companies (several existed, including the American Marconi Company) were all based on the rotary spark, transmitting at very high power levels. It was obvious to all in the late 'teens and early 'twenties, that spark technology no longer stood at the leading edge of radio development. Nevertheless, Marconi's entire career had been in the development of just that technique. He had developed it (nearly) to its practical limit.

The wave of the future was to be in continuous waves, and everybody knew it. Marconi was not idle, however; he had been developing a method of using sparks to generate nearly pure CW, by employing an ingenious idea—the "timed spark."

The transmitter consisted of several disks (rotary gaps) on one shaft, which were arranged to produce a spark discharge in a rigidly timed sequence. The idea was to synchronize the shaft and the gaps so each primary impulse was followed by the next, adding to the coupled currents in sequence, thus producing continuous waves. **Figures 19** and **20** show the method
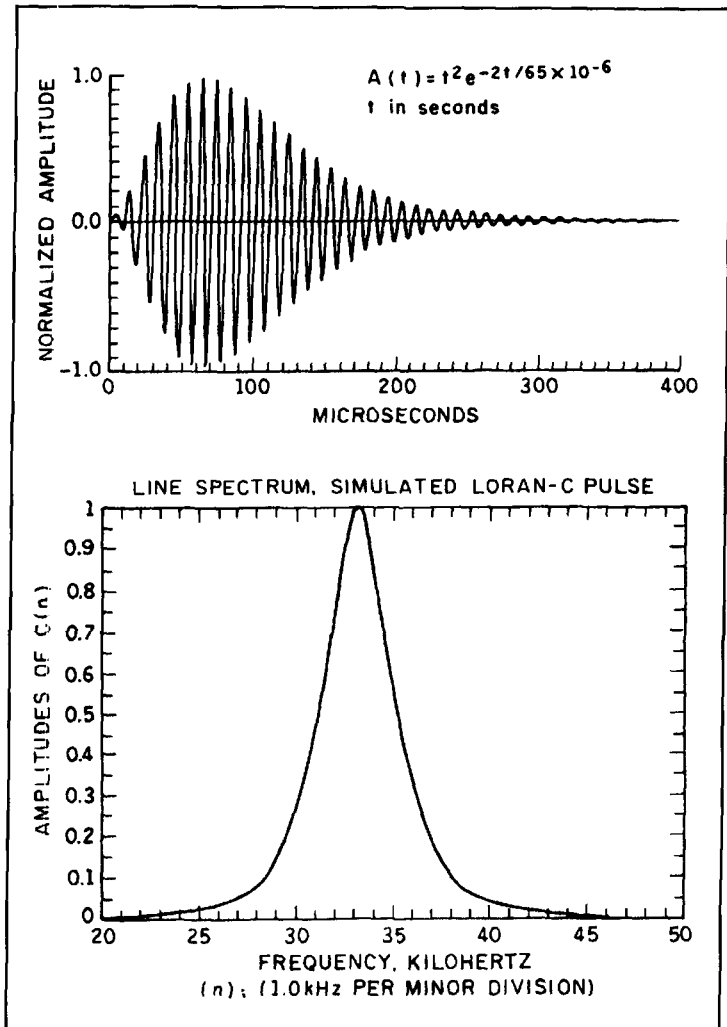
Figure 16. (A) Shape of a modern LORAN-C pulse reproduced from Evans and Evans, *Communications Quarterly*, Fall 1991. (B) LORAN-C pulse simulation.

and its resulting waveforms. Elmer Bucher said that the exact circuits shown in his book and reproduced here were not the final configuration.[6] At the time of its publication (1921) Marconi was reluctant to divulge design details.

There were apparently two generators. One was operated successfully from station MUU with the excellent antenna system there. The other station was to be at Marion, Massachusetts, Marconi's terminus in the U.S. MUU was using the timed-spark transmitter in daily operation. Bucher wrote, "With a power of less than 100 kW, the author has copied signals from the Carnarvon station in New York City during the daylight hours with an aerial 100 feet in length. The signals were considerably stronger than those obtained from other foreign stations using radio-frequency alternators of greater power. The wave length employed during those tests was close to 10,000 meters."[6]

5*(sin(1.3963*x))*(sin(2.06E-2*x))^2

TIME; (150 UNITS = 0.001 SECOND)

SPECTRUM, f(x) = 5*(sin (1.3963*x))* (sin (2.06E-2*x))^2

FREQUENCY, KILOHERTZ
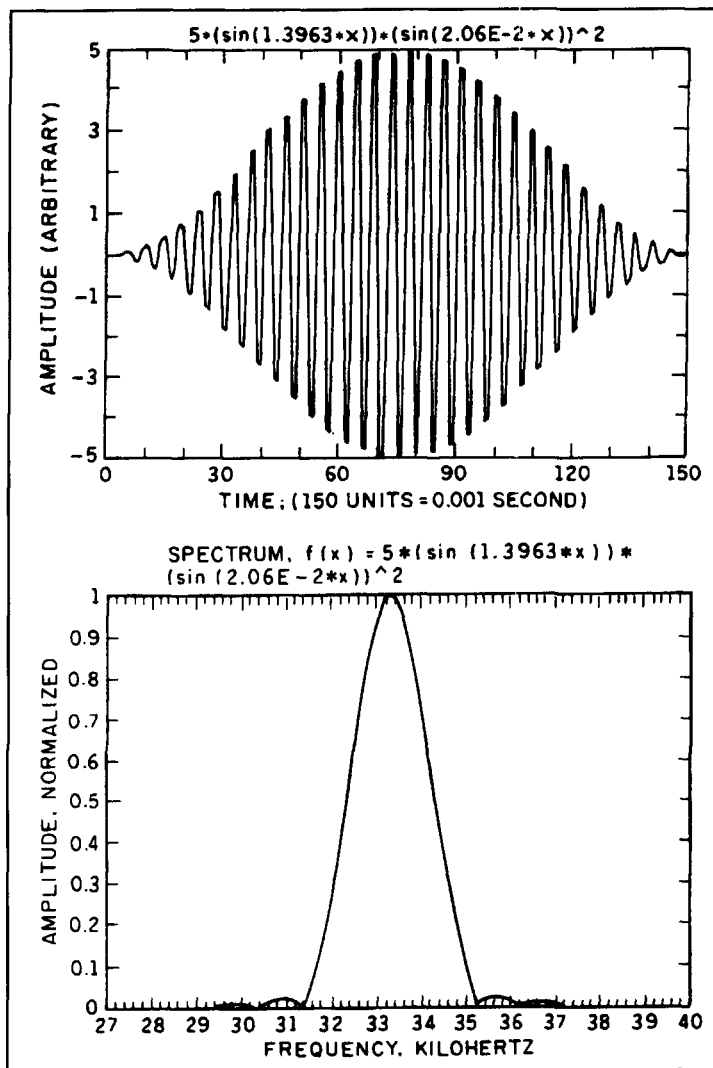
Figure 17. (A) Synthesized wave-shape with gentle slopes on both ends. (B) Spectrum of the synthesized wave.

The other generator was actually installed at Marion, Massachusetts. It never was successful.[10] The antenna at Marion was an inverted-L one mile in length. It also used the Alexanderson multiple-tuned feed system. It was that antenna system, when the "star" type grounding system was used earlier, that produced brush fires. The soil underneath the antenna was very sandy, and obviously had a low conductivity, despite the fact that it was within sight of the ocean. Many transmitting sites, by the way, were built on shorelines, or actually on the beaches themselves.[11] The antenna system could very well be the reason for failure of the spark CW transmitter; that would seem to be the only difference between the two sites. MUU has a system that operated with a log-decrement of 0.029. By examining **Figure 19** again, it seems obvious that if the oscillation transformers (P_1, C_1, S_1, etc.) had a large combined log-decrement, the currents could not combine properly to produce a good continuous wave. High ground resistance has the *strongest* influence on the log-decrement of VLF antennas.

The CW generator was the last major development using the spark. It was developed as a possible answer to the other competition— several alternator types, and of course, the very high-power arc transmitters, and the thermionic vacuum-type, which was destined to obsolete them *all*. It was the spark's last gasp.

REFERENCES
1. Captain L.S. Howeth, USN. *History of Communications-Electronics in the United States Navy*, Washington, U.S. Government Printing Office, 1963.
2. David Kahn, *The Codebreakers- The Story of Secret Writing*, Macmillan Publishing Company, Inc., New York. (Tells of intercepts of field transmitters on both sides in World War I and their decryption.)

```
20      A1 = 0:B1 = 0
30      CTR = 0
40      DEFDBL A, B, Y, W
50      INPUT "ENTER STARTING N" ; N
51      LPRINT"Y = 2*(SIN(1.3963*X))
60      PRINT "WAIT - PROGRAM RUNNING." :PRINT:
61      PRINT TIME$
70              FOR X = 0 TO 150 STEP .1 ' (Asks for 1,500 computations)
80                  Y = 2* (SIN(1.3963*X))
90                      W = (6.2832*X)/150
100                     Y1 = Y*SIN(N*W)/150
110                     A1 = A1 + Y1
120                     YY = Y*COS(N*W)/150
130                     B1 = B1 + YY
140                     CTR = CTR+1 '(This is to keep track of 1,500 integration steps)
150             NEXT X
160     LPRINT USING"##.##" ; N
161     PRINT USING "##.##" ; N
170     LPRINT "CN = " ; SQR (A1^2+B1^2)/10.01:LPRINT
171     PRINT "CN =" ; SQR (A1^2+B1^2)/10.01
172     A1 = 0: B1 = 0: N = N+.05
181     PRINT CTR-1 ' (Erases the last incrementing in line 140)
190     CTR = 0: IF N <= 38 THEN GOTO 60 ELSE END
```

Figure 18. Program listing.
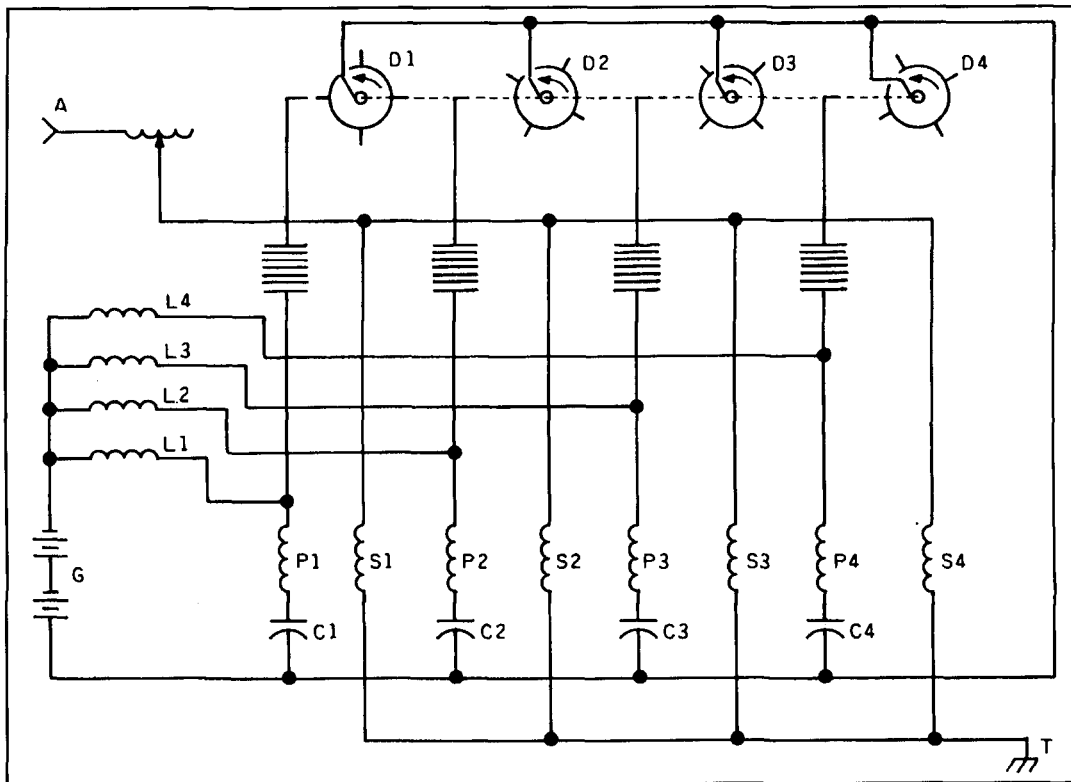
Figure 19. Marconi's method of generating continuous oscillation. (From Bucher).
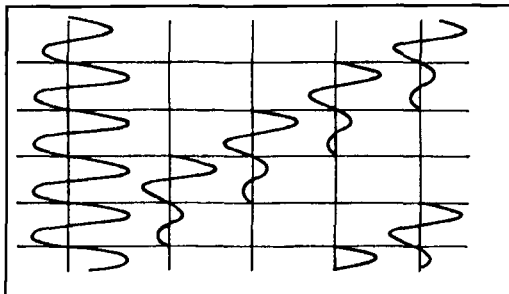
Figure 20. How groups of damped oscillators can be made to overlap one another. (From Bucher).

3. M.E. Van Valkenburg, *Network Analysis*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1955.

4. J. Zenneck, *Wireless Telegraphy*, McGraw-Hill Book Co. Inc., New York, 1915.

5. A.C. Goodnow, "Across the Gap; An Appraisal of Spark Radiotelegraph Engineering," *The AWA Review*, Vol. 2, 1987.

6. Elmer E. Bucher, "Practical Wireless Telegraphy," *Wireless Press*, New York, 1921.

7. As [4], Article 76 et seq.

8. Randy Evans, KJ6PO, and David Evans, N6UEZ, "Loran-C Frequency Calibrator," *Communications Quarterly*, Fall 1991, page 20.

9. R.M. Kerschner and G.F. Corcoran, *Alternating Current Circuits*, John Wiley and Sons, Inc., New York, 1951.

10. Bruce Kelley, W2ICE, personal communication. Bruce witnessed the final days and the dismantling of this and other stations. He also is Curator of the AWA Museum.

11. Irving Vermilya, W1ZE, "Wireless Cape Cod (WCC)", *QST*, February 1942.

Peter O. Taylor
AAVSO Solar Division
4523 Thurston Lane, #5
Madison, Wisconsin 53711-4738

THE SOLAR SPECTRUM
Another index of solar activity

In a recent issue of the Space Environment Laboratory *User Notes,*[1] Frank Recely of SEL and the National Solar Observatory (NSO) described the solar synoptic program which utilizes the vacuum telescope on Kitt Peak, Arizona. Virtually all active solar phenomena, sunspots, flares, coronal holes, etc., are related to magnetic influences on the Sun (**Image A**), so this work is of interest to a broad range of researchers and scientists. According to Mr. Recely:

"The vacuum telescope, one of twenty-two on the mountain, is a coelostat-fed 76-centimeter aperture reflector with a 36-meter focal length. It was built in 1972, with partial funding from NASA to support the *Skylab* mission of the early 1970s, and has been collecting data ever since. The telescope was designed by NSO scientist emeritus, William Livingston.

"The current program consists of collecting a daily full-disk, 1 arc-second per pixel solar magnetic map at 868.8 nanometers (nm), and a full-disk helium 1083.0 nm spectroheliogram. (The helium image is a good proxy for an extreme ultraviolet image of the Sun, and allows us to detect coronal holes and mass ejections.) These data are processed daily in real or near-real time and distributed to the solar community.

"Space Environment Services Center (SESC) receives these data in the form of daily 512 x

Image A. Learmonth Solar Observatory digital map of a portion of the solar magnetic field. Many, if not all of the Sun's active phenomena are related to magnetic influences. Eric A. Kihn of the National Geophysical Data Center supplied this image.
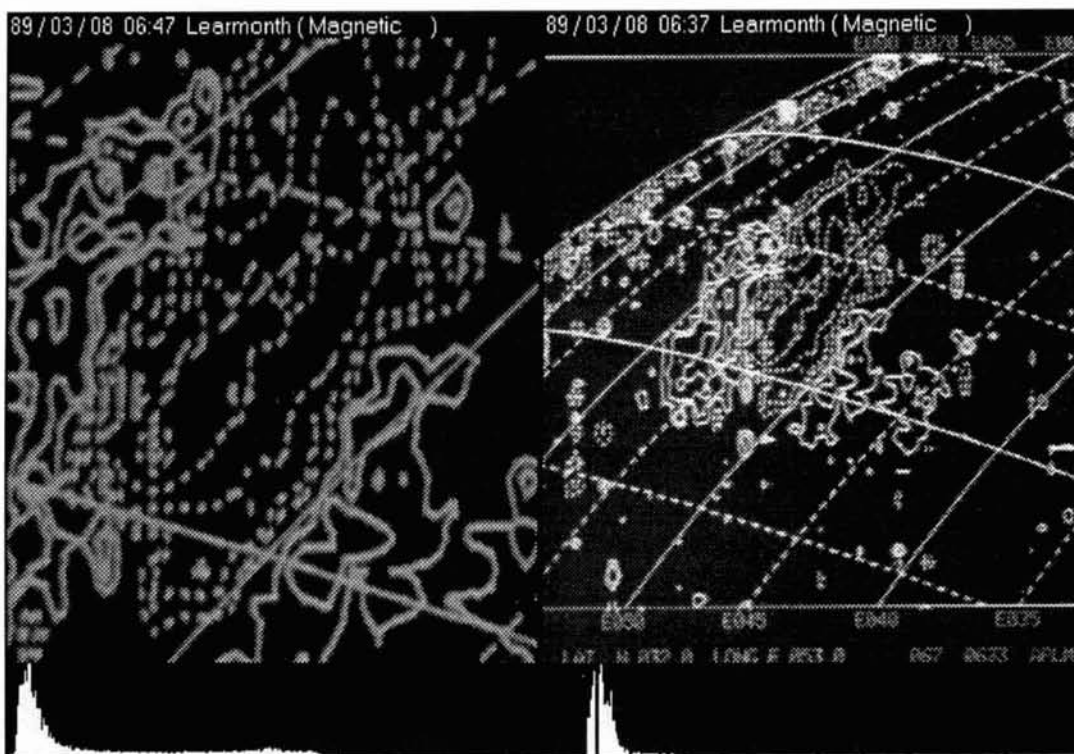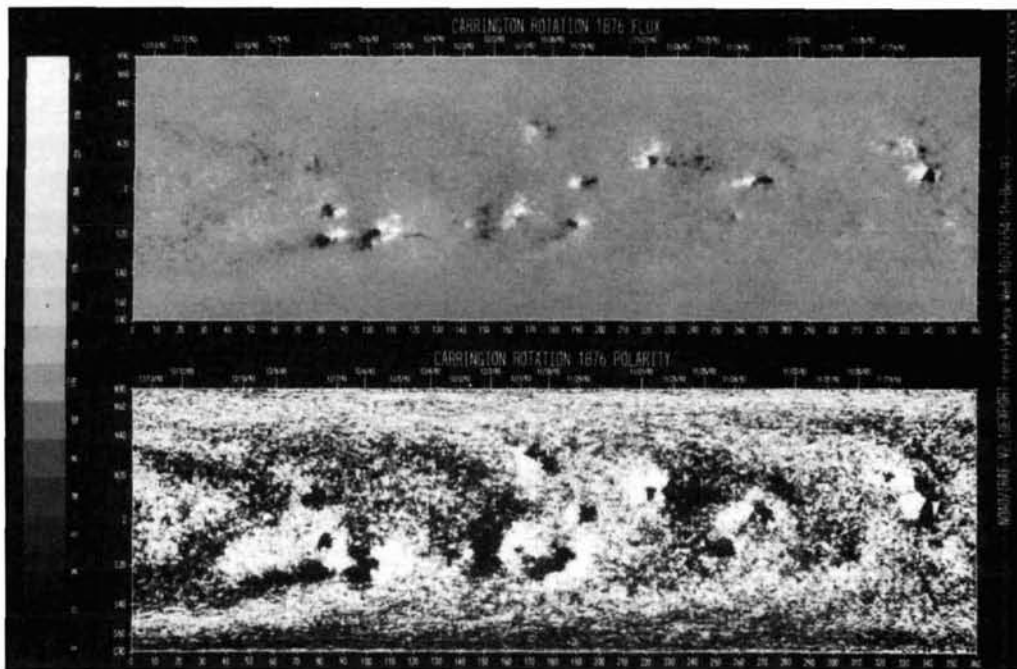
Photo A. Kitt Peak Solar Magnetic Field Synoptic Map. Daily full-disk solar magnetograms can be computer processed into Carrington Rotation maps. Maps such as the one shown here correspond to a single mean rotation of the sunspot zone, according to the method described by Richard Carrington in 1853. In the photo, bright areas represent positive polarities, while dark zones are negative. Note the opposite "leading" (more westerly) polarity for sunspot groups in the Northern and Southern Hemispheres. Northern leaders normally possess positive polarity during odd-numbered cycles, and southern leaders are negative; a feature that reverses each sunspot cycle. (Photo courtesy of Frank Recely, National Solar Observatory—Kitt Peak.)

512 pixel (4 arc-second resolution) digital images, which are sent automatically to Boulder through both Internet and SELSIS along with a line drawing of coronal holes inferred from 1083.0 nm images. These data—from November 1992 to present—also reside on an anonymous **ftp** disk in Tucson and are available to researchers.

"In addition, we collect a daily full-disk Calcium II K (393.4 nm) image of the Sun from which a *calcium brightness function* is measured as a parameter of solar activity. These data are not available in real time but are furnished at a later date. In the near future, we plan to provide additional images to the Internet community. Our immediate goal is to include the magnetic and 1083.0 nm maps at 1000 x 1000 pixels (2 arc-second accuracy) and to add an 868.8 nm intensity image (a possible first step to deriving sunspot area as a percent of total disk).

"Each month, the daily data are processed into Carrington rotation maps such as that shown in **Photo A**, and distributed to the World Data Center and SESC. These maps are also available on NSO's anonymous disk for users. Dr. Karen Harvey of NSO computes an average total magnetic flux value from the Carrington map. A plot of the past 18 years appears in

Figure 1, with sunspot maxima and minima indicated."

A few inferences from **Figure 1**:

• "The magnetic flux increases fivefold from minimum to maximum. This is a far larger increase than in the more common solar activity parameter, 10-centimeter radio flux.

• "The magnetic flux is closer in pattern to the 10-cm flux than to the sunspot number, as both reach a maximum after the sunspot maximum.

• "Cycle 22 is rapidly declining. When it drops below 30×10^{14} Webers, there is a three-plus year period of low solar activity."

We are indebted to Frank Recely for his explanation of this productive program, and for his description of the variation in magnetic activity that accompanies each solar cycle.

A shielded, tuned-loop antenna with preamplifier

A few contributors to our program of recording ionospheric anomalies caused by solar flares[2] (myself included) have called attention to problems with interference on the signal to the VLF receiver. Some of this results from
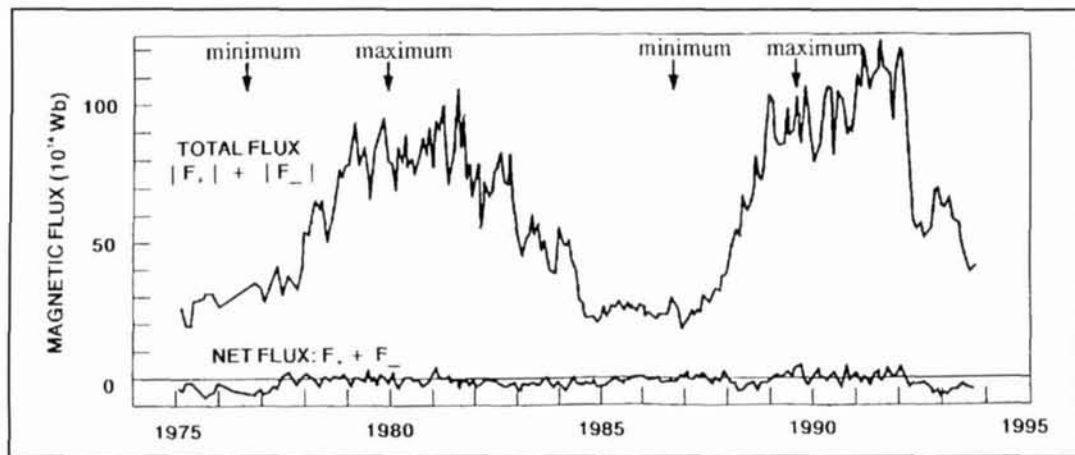
Figure 1. Eighteen-year magnetic flux variation on the sun by Carrington rotation. Sunspot cycle maxima and minima are indicated on the graph. Note that this index appears to peak after the sunspot cycle reaches maximum. The diagram originally appeared in the SEL *User Notes* for January 1994.
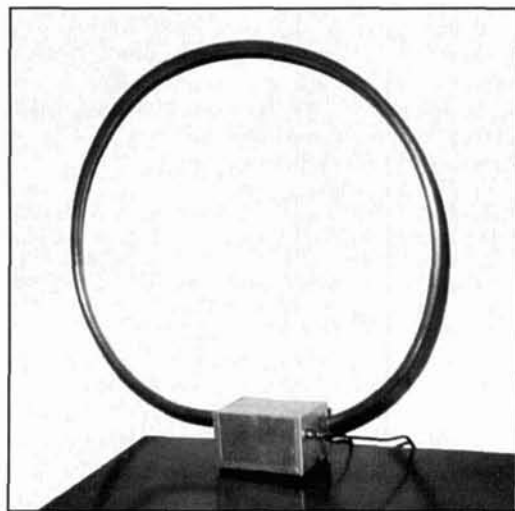
Photo B. Arthur Stokes' new shielded, tuned-loop antenna with preamp. This type of antenna can be helpful in locations where the VLF signal is weak.

exterior electrical noise, but other common sources are television sets and computers—especially older-model computer monitors. This situation can be particularly aggravating if one uses a small loop antenna inside the house. In fact, the noise emanating from some monitors can substantially affect TV reception. When I encountered similar problems, I did as I often do when an electronic difficulty rears its ugly head: I turned to Arthur Stokes for a solution.

Arthur suggested that we try to overcome my dilemma by exchanging my usual loop antenna for a shielded, tuned-loop with preamplifier (**Photo B**). These units generally incorporate similar design principles, and a number have been described in previous articles. The schematic for our design in shown in **Figure 2**. A parts list is included. The alert reader will rec-

ognize the circuit as a modification of the new Stokes' gyrator-tuned VLF receiver, which is described elsewhere in this issue. Incidentally, that printed circuit board can also be used for the antenna preamp. Unnecessary parts are simply omitted, and the board trimmed if desired.

This type of antenna is actually comprised of two parts: the shielded loop, and a metal box containing the preamp (**Photo C**). The first requirement of such a device is to enclose the wire winding of the loop in a nonmagnetic, metallic housing. This pretty much limits the usable material to copper or aluminum.

Soft 1/2-inch copper tubing (the type available in rolls at the local hardware store) was chosen for our unit, since it is always easy to obtain and can be worked without great difficulty. A single turn of the outer diameter of the roll was cut off, leaving a circle of material almost 24 inches in diameter. Next, an aluminum box measuring 3 x 4 x 5 inches was obtained. A box of this size could contain all the wire connections and preamp, and form a base for the shield. These boxes are usually supplied in two sections that fit neatly together. The cover required no modifications, so this part was set aside.

Two 3/4-inch holes were cut into opposite ends of the remaining section, and a 1/2-inch electrical conduit adapter was inserted through one opening. At this point, the loop of copper tubing should be adjusted by carefully shaping the material so that one end fits into the conduit fitting while the other is placed into the opposite opening, leaving about 1/2-inch of tubing extending into the box. The screw on the conduit fitting is then tightened so it clamps the copper tubing securely in place.

A short section of plastic tube about 1-1/2 inches long should be slipped over the free end of the copper tubing in order to insulate it from
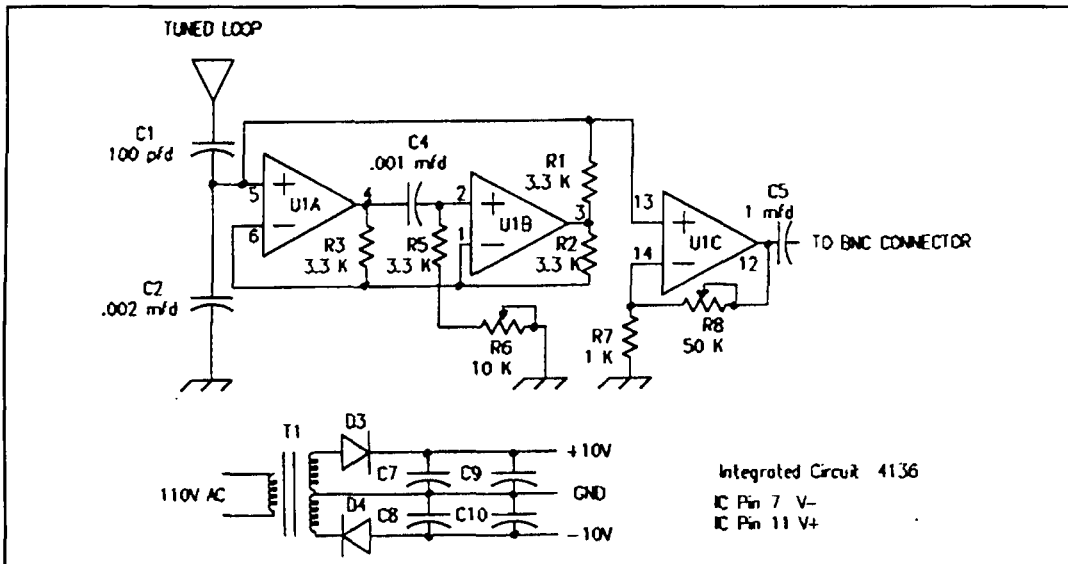
Figure 2. Schematic drawing of the antenna preamplifier. This unit utilizes the same printed circuit as the new Stokes' gyrator-tuned VLF receiver, which is described elsewhere in this issue.

the metal box, since the copper shield must not form a shorted loop! To attach the coil to the box even more firmly, a bit of fiber glass resin can be applied around the seams where the loop joins the box, and allowed to cure.

The next step is to prepare the wire loop and insert it into the copper shield. The material that forms our loop was salvaged from a 50-pair telephone cable. Such cables are comprised of 24 gauge insulated copper wires that are conveniently color coded. Incidentally, a bundle of 50 wires is about all that can easily be pulled through a copper tubing shield of this diameter.

The packet of wire—measuring about 7 feet in length—should be trimmed so that only about 4 inches protrudes from each end of the shield. The wire ends are stripped of insulation and soldered together to form a single 50-turn length, and short pieces of electrical tape are wrapped around the soldered joints. Care must be taken to identify each separate wire during this process, ensuring that a single continuous coil results.

The two ends of the wire loop are connected to a terminal strip, with a fixed capacitor shunted across the coil. The value of this component should be accurately adjusted by using a signal generator and oscilloscope to set the resonant frequency of the loop to that of the desired station. In our case, a 0.016 µF capacitor resonated at exactly 24.0 kHz—the frequency of VLF station NAA located in Cutler, Maine. One end of the loop is grounded to the shield and aluminum box and the other is connected to a 100-pF capacitor input to the preamp.

The pc board was mounted on 1/2-inch standoffs in the bottom of the box. A 12-volt,

Parts List: Preamp/Loop Antenna

R1,R2,R3,R5	3.3-k resistors, 1/4 watt
R6	10-k linear potentiometer or trimpot
R7	1-k resistor, 1/4 watt
R8	50-k linear potentiometer or trimpot
C1	100-pF ceramic capacitor
C2(2)	Parallel 0.001-µF ceramic capacitors
C4	0.001-µF ceramic capacitor
C5	1-µF electrolytic capacitor
C7,C8	470-µF electrolytic capacitors
C9,C10	0.1-µF ceramic capacitors
D3,D4	1N4001 diodes
U1	4136 integrated circuit chip
T1	12-volt CT transformer

Metal cabinet and miscellaneous materials (see text).

center-tapped transformer supplies power to the circuit. Of course the preamp can also be constructed on a small section of perfboard, with a similar placement of components. If that method is chosen, all wiring should be point-to-point with short pieces of fine wire used throughout.

Since only an initial tuning and gain adjustment is necessary, multi-turn trimpots were selected for R6 and R8. A fixed, 10-k resistor can be used for R8 to give a gain of 10, which seems to be a convenient value for my location in Southern Wisconsin.

The only adjustments necessary are to tune R6 of the gyrator for maximum signal strength, and R8 to the desired gain. When the recording system is set up, the loop should be placed a few feet away from the receiver in order to minimize any possible interaction. The two units can be connected with a 4 to 5 foot length
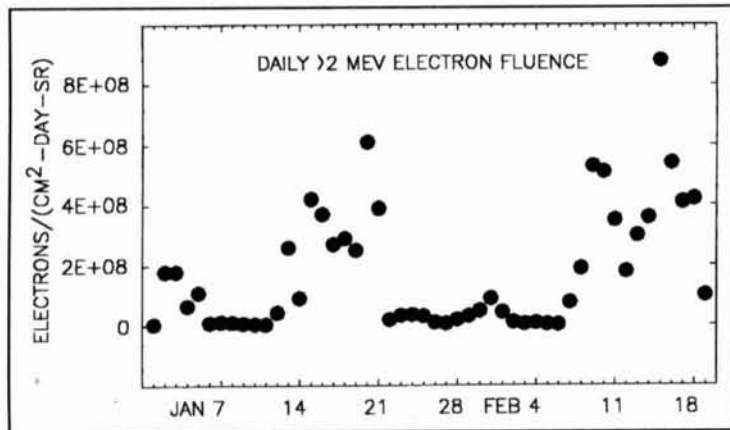
Figure 3. Enhanced electron levels from 1 January through 19 February 1994, as measured by instruments aboard the *GOES* spacecraft. Each of the plotted values has been integrated over the UT day.
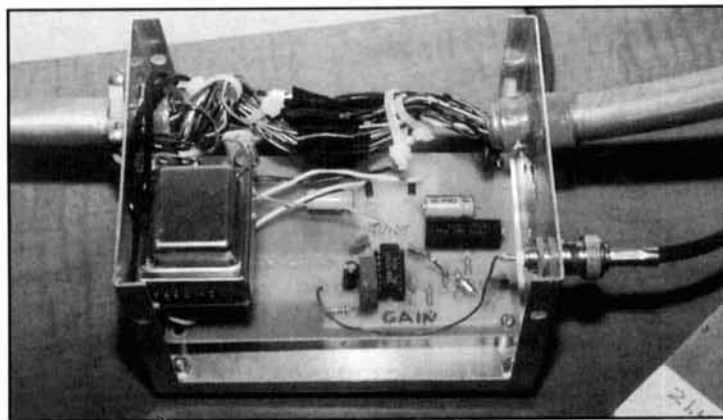
Photo C. The arrangement of copper tubing shield and wire loop in conjunction with the preamp's electronic components.

of coax. If a signal generator is not available, a little patience will be required to tune both the preamplifier and receiver to the exact same frequency. The preamp has a maximum gain of 50; but it is suggested that this be held to perhaps 10 or 20.

The Stokes' loop antenna works beautifully in my location. When I moved to Wisconsin last summer, the VLF signals that I had monitored in the past—NSS in the Washington, D.C. area and NAA—were weak and plagued with local interference. Since implementation of the new antenna systems, both produce excellent recordings allowing me to effectively monitor the ionospheric effects of many solar flares as the eruptions occur.

The January spacecraft anomalies

During mid-January the Canadian communications satellites, Anik E1 and E2, experienced failures and other spacecraft anomalies were reported at geosynchronous altitude. At the time, a number of media and other quasi-offi-

cial reports attributed these problems to a "massive solar flare" or even a "tear in the Sun's surface"! Such statements were erroneous to say the least. No eruptive events (flares, disappearing filaments, etc.) occurred that could have accounted for these difficulties.

In the Summer 1992 issue of *Communications Quarterly*, we discussed some basic features of the solar wind and the low density areas of the Sun's atmosphere known as "coronal holes." To our knowledge the precise mechanism is not fully understood, but after certain of these solar particle streams "connect" with the Earth's magnetosphere, electrons can be accelerated to moderate and high energies. Electrons with energies greater than 2 million electronvolts (MeV) were, in fact, detected by the *GOES* spacecraft during mid-January (**Figure 3**). Moreover, in this case relatively high energy levels were maintained over an unusually long interval.

Although not certain at this juncture, one explanation for the failure of Anik and the other problems in the near-Earth environment suggests an association with a phenomenon called "deep-dielectric charging," wherein energy penetrates down into the body of a spacecraft. Satellites can slowly discharge such currents, but damage to delicate electronics can occur if energy is deposited at a rate faster than it escapes.

Recurrent coronal holes are not necessarily geoeffective during each appearance, but they do have longer lifetimes and greater stability during the declining phase of a sunspot cycle. The coronal hole(s) that are suspected of involvement in the January incidents, rotated back into a position to affect the Earth's environment at the end of the first week of February. As can be seen from **Figure 3**, another long period of enhanced electron flux similar to January's high-energy flow was recorded.

A final point: The severe geomagnetic disturbance which began on February 20th-21st and resulted in a complete radio blackout over HF transpolar paths, was not related to the situation described above. Rather, these conditions—including a number of satellite magnetopause crossings, a Forbush decrease, and a sharp surge in reports of aurorae—were caused by the shock wave and plasma cloud from a solar flare and mass ejection that occurred early on the 20th. Protons reached a peak (>10 MeV) of near 10,000 particle flux units shortly after 9:00 UT on the 21st as the fast-moving outflow passed the Earth. This was the first proton event to be recorded since March 1993. ■

REFERENCES

1. SEL *User Notes*, Number 4, January 1994.
2. "Recording Solar Flares Indirectly," *Communications Quarterly*, Summer 1991.

Marv Gonsior, *W6FR*
418 El Adobe Place
Fullerton, California 92635

TX HIGH-PASS FILTER APPLICATION

Improve low-pass filters performance by combining them with absorptive high-pass filters

When operating into reactive loads, low-pass filters (LPF) may exhibit degraded characteristics compared to those obtained when working into a pure resistive load. One reason for this is that interaction may take place at harmonic frequencies where the reflection to the input impedance of the LPF causes its performance to deteriorate significantly. This phenomenon can be alleviated, if not cured, with an absorptive high-pass filter (HPF) incorporated in shunt and resistively terminated, so the LPF sees a resistive load in its stopbandconsistent with its design requirement. Although much has been published regarding the design and application of LPFs, I found very little information about transmitter HPFs used in this application. Two overseas companies are currently manufacturing what I will call a "compound filter" of this nature; one in Japan and the other in Germany. The Japanese filter goes one step further. It consists of a shunt HPF to an LPF to a shunt HPF. The HPFs are terminated in 50 ohms. The company is reputable and claims significant performance specifications. Their power rating is limited to one kW. Unfortunately, I have no data on the filter being marketed in Germany.

In the spring of 1992, I published[1] a survey of LPF performance. While preparing that article, I became aware of previous work on the use of a HPF in a shunt configuration with a LPF.[2] A brief write-up and schematic of the same application appears in the *The ARRL Handbook.*[3] I thought the idea of combining an LPF with an HPF was most intriguing, so it became a goal to learn about this concept and put it into practical application. My results are presented here.

When reviewing the earlier work, I noted that a rather simple four-element HPF was used in conjunction with a four-element LPF. Results were reasonable for 1968, and TV IFs weren't then in the 45-MHz region. Today, it's possible to obtain LPFs with good attenuation at 45 MHz and this is sometimes necessary. In 1968, a rather simple HPF could be used in conjunction with an LPF because its shape was much less demanding from 30 MHz to the first TV frequency of concern—56 MHz. To combine a high-pass and good-quality low-pass filter today, an exceedingly good shape is a necessity. An equal number of elements is required for the two filters to complement each other. It became readily apparent that a high-pass/low-pass combination would necessitate a
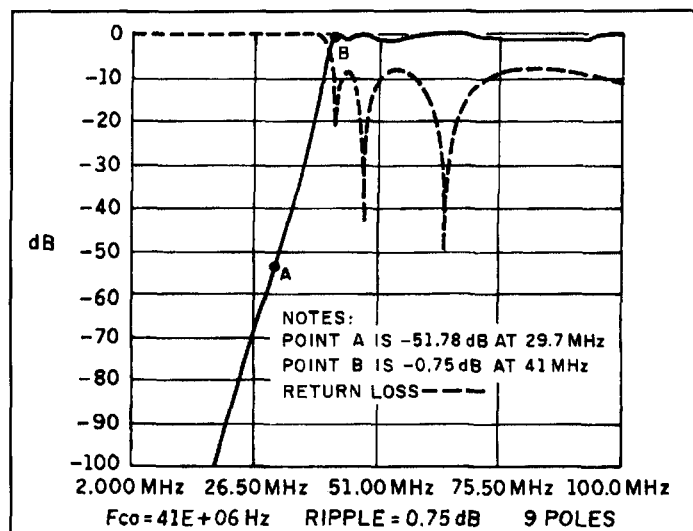
Figure 1. Theoretical HPF performance.
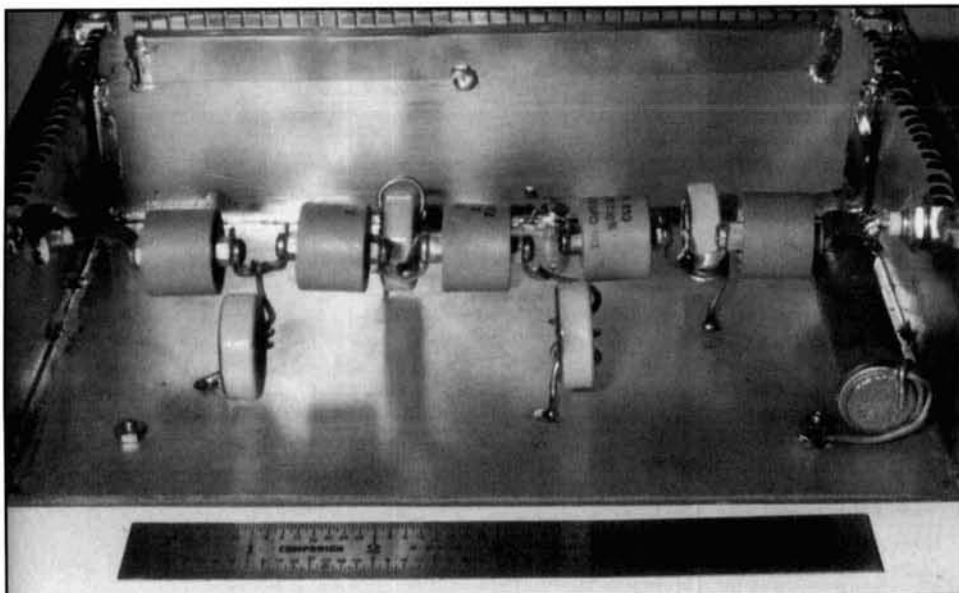
Photo A. High-pass filter configuration.

1:1 element relationship, causing me to quickly progress from a five to a nine-element Cheby-chev filter! However, my computer program,[4] into which I input some very demanding speci-fications, indicated that a twelve-element filter would be ideal. I compromised by increasing the ripple factor from 0.1 to an insignificant 0.75 dB in the passband, while lowering my sights to nine elements. This design was quite satisfactory. **Figure 1** is taken from the com-puter program print out of the theoretical result.

The design concept

I wanted to avoid the need to compensate the filter from cover-on to cover-off due to capaci-tive effects. I looked into the use of toroids to eliminate this problem, which is very dramatic in most LPFs I've measured. If the problem couldn't be resolved, I would need to be able to tune the filter from the outside, and this meant air variables, etc. Also, I wanted to use standard values of relatively high performance, non-inductive capacitors. First, I considered using etched Teflon™ printed circuit board, which would be ideal. Unfortunately, it didn't physi-cally lend itself in the HPF configuration, as it would in an LPF where one side of the capaci-tors is grounded. I decided to use Centralab type 850/S units because they essentially fulfill all the requirements—including very reason-able current and a large overkill on voltage. (Should you wish, you can substitute high volt-age ceramic NPO discs in parallel for the Centralabs.) The results were most gratifying. With the use of an oversize enclosure (more on that later) and the toroid inductors, I found no discernible response change when the filter was

covered, partially uncovered, or totally uncov-ered when on the breadboard.

The toroids are easy to wind and the entire assembly is very straightforward.

Some construction details

To minimize inductance between the capaci-tors, I used 1/4-inch wide copper strap. You could substitute 1/2-inch pieces of 6-32 thread-ed brass stock to interconnect the capacitors. This would make assembly very neat and more simple. Keep all leads short and solder them directly to the enclosure, which is constructed from 1/16-inch thick double-sided printed cir-cuit board whose dimensions are 8 x 4.75 x 2.5 inches. The two copper planes *must* be fully joined together using methods such as screws, connectors, and solder. My results indicate that the enclosure could probably be smaller, but I wasn't going to expend the effort to find out! I included a BNC connector at the terminated end of the filter for two reasons. The first was to allow for testing on the spectrum analyzer and tracking generator; the second was for pur-poses of adding a detector at that point for some performance evaluation. I leave the use of the connector at the input end up to you. **Photo A** was taken when a BNC was needed at both ends for testing. The 50-ohm terminating resis-tor wasn't connected in the circuit for spectrum analysis, as the output of the tracking generator to the filter input is 50 ohms. Therefore, the fil-ter is properly terminated in its characteristic impedance at both ends because the HP141T is also a 50-ohm system. Centralab type 850/S capacitors are available at flea markets and some surplus dealers at fairly reasonable prices.

The inductors are made from Micrometals T94-6 toroids (0.942-inch OD and a number 6 mix) that require only three turns of number 14 wire 3-3/4 inches long, including the leads. Micrometals' number 6 iron-powder mix is specified to be most effective in the 10 to 50 MHz range, and is color-coded yellow. The two end inductors are close wound and the intermediate ones are all wider spaced. **Figure 2** gives their dimensional details. I measured the coils on a Hewlett Packard model 4265B Universal Bridge, although I found that high precision inductances weren't mandatory. Consequently, if one wished to duplicate this filter, I feel that only very slight variations would occur with only minor effects on the performance. I tested shielding between filter elements and found no measurable change. This led me to conclude that the use of toroids was reasonable. It's certain that air-wound inductors would have higher Q, but such inductors would result in enclosure effects combined with the probable need for internal shields. As a precaution, I isolated the toroids from each other. This was solely for the purpose of eliminating capacitive coupling, because they are relatively free of inductive coupling due to their inherent self-shielding characteristic. In earlier breadboarding, I noted considerable performance degradation from apparent capacitive coupling. This was eliminated simply by physical isolation.

Test results

In all configurations, I used a computer program for modeling. The final HPF calculations and schematic obtained using the program are illustrated in **Figure 3**. My conventional spectrum analyzer and tracking generator test setup is shown in **Figure 4**.

Figure 5 shows the performance characteristics of the HPF. The attenuation at 29.7 MHz is -52 dB—essentially the same value provided by the computer program based upon an infinite Q for all elements. This translates to a power loss of only approximately 9 mW out of 1500 watts. The attenuation at 41 MHz is desirably minuscule, which is consistent with the computer calculation and design objective. Also, note that my capacitors are slightly greater than the theoretical ones due to the manufacturer's standard values. This means the filter is probably overdesigned in power capability unless one were to have some very serious harmonics being generated or coupled to the wrong antenna! **Figure 5** demonstrates the result of using a compound filter with a LPF, in this case a Drake TV-3300LP (regrettably no longer in production), that has good rejection at 41 MHz. Both scales on the spectrum analyzer are 10 MHz and 10 dB per division. In all cases of discrete harmonic mea-

Figure 2. Toroid inductors.

surements with the analyzer, I use a -54 dB T notch filter centered at 14,195 kHz. This gives me the ability to place a very large fundamental signal into the analyzer without any risk of overloading its mixer, letting me obtain good harmonic measurements.

I tested the high-pass filter for ten minutes at 1500 watts continuous at 29.7 MHz, without an LPF, in shunt into a 50-ohm Bird dummy load. The carrier was relatively rich in harmonics. It exhibited no temperature rise during this type of operation as measured by a Fluke 8020A meter in conjunction with a Fluke 80T-150

N	SERIES C	SHUNT L	HP FILTER
1	39.23E-12		Fco = 41E+06 Hz
2		163.43E-09	RIPPLE = 0.75 dB
3	26.72E-12		RG = 50 Ω
4		152.93E-09	RL = 50 Ω
5	26.23E-12		Fs = 30E+06 Hz
6		152.93E-09	As = 51.78 dB
7	26.72E-12		
8		163.43E-09	
9	39.23E-12		

Figure 3. High-pass filter schematic.
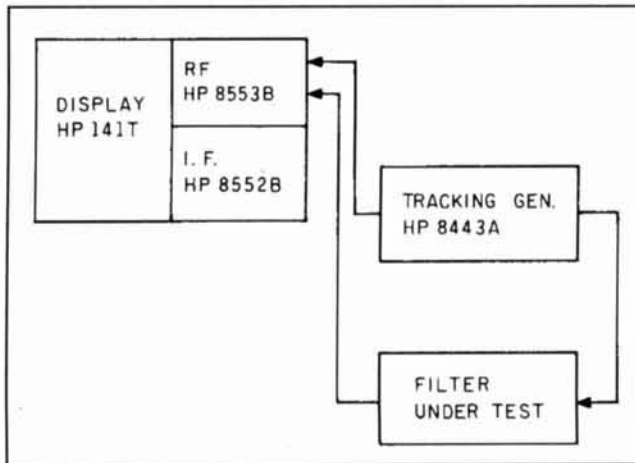
Figure 4. Measurement system. Note: The HP-141T/8443A is a 50-ohm system.
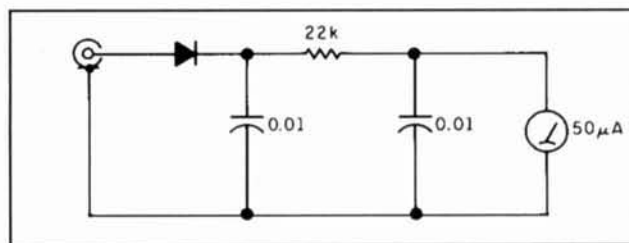
Figure 5. Detector circuit.

Photo B. Compound filter response. Note: Scales are 10 MHz and 10 dB/division.
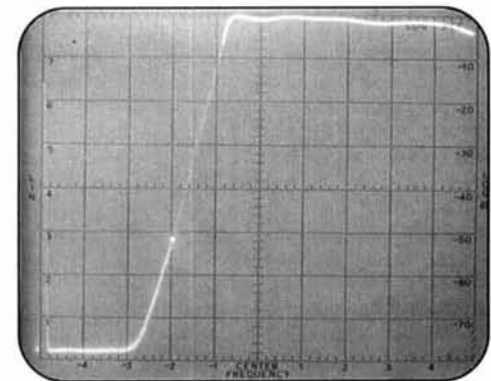
Photo C. High-pass filter response. Notes: Scales 10 MHz and 10 dB/division. 29.7 MHz is -52 dB. The tracking generator marker is at 29.7 MHz.

temperature probe. While I expected this theoretically from the measured response curve as shown in **Photo B**, it did provide proof that the toroids (I had no doubts about the capacitors!) had sufficient capability for the application. Based upon those results, it would appear that the Micrometals T80-6 (0.80 inch versus 0.96 inch) size would probably be adequate (and save a little money). Of course, this would mean reconfiguring the coils. You could try Micrometals' number 10 mix (black color code) for slightly better Q in the 50 to 100 MHz region. However, their number 6 mix appeared to be reasonable since, from my test results, a higher Q didn't appear necessary.

I then tested the filter under the same conditions, as noted above, but in conjunction with an LPF (a Drake TV-3300LP) into a 2:1 VSWR. I did this by paralleling two 50-ohm Bird dummy loads to simulate what might occur in some real-life situation. There was an increased relative output across the HPF internal dummy load in accordance with the detector schematic, as shown in **Photo C**. There was no measurable temperature rise in the HPF as a result of my tests, but my poor LPF became a bit warm at times!

Conclusions

A compound filter of this type can only yield

improved harmonic attenuation by virtue of its nature, because its response curve should provide a good 50-ohm termination to all frequencies above 41 MHz. This prevents significant amounts of that energy in that spectrum from being reflected back to the LPF, potentially degrading its performance. Those who operate in a fringe area, may find that the additional effectiveness from an LPF provided by this HPF, without any modification, will offer a distinct advantage. Relative voltage readings, taken by the detector at the output of the filter, confirm its ability to terminate HF harmonics into a 50-ohm load.

Acknowledgments

I wish to express my sincere appreciation to the Micrometals Corporation, W6NBI, N5AR, and K2SZE for their contributions, which made my work possible.

REFERENCES

1. Marv Gonsior, W6FR, "Low-pass Filter Performance," *Communications Quarterly*, Spring 1992, page 75.
2. Weinrich and Carroll, "Absorptive Filter for TV Harmonics," *QST*, November 1968, page 20.
3. "An Absorptive Filter," *The ARRL Handbook*, 1989 edition, page 40-9.
4. Filter program by Tim Murphy.

L.B. Cebik, *W4RNL*
1434 High Mesa Drive
Knoxville, Tennessee 37938-4443
BITNET: CEBIK@UTKV

COMPARING MININECS

A guide to choosing an antenna optimization program

Although the MININEC code (currently Version 3, revision 13) serves as the basis for available commercial and non-commercial antenna modeling implementations, the resultant programs for amateur use show some significant differences. Consequently, no one program is best for all possible applications. Notices and advertisements for various MININEC enhancements do not give many clues as to which program may be best for an individual ham, regardless of whether he or she is just exploring the possibilities of antenna modeling or is highly experienced with various forms of computer modeling.

Differences in programming and capabilities are both worth investigating. Ease of use and adequacy of results have no absolute criteria. Rather, each is application dependent. Therefore, in this user examination of three well-advertised versions of MININEC, I shall relate each evaluative remark to one or more applications. This procedure should preclude the

```
                        Free Space
                        14.175 MHz

                           Options

 Field            » Total H V RC LC MaxLin MinLin AxialRatio
 Segmentation Density              10 segments / halfwave
 Auto Segment-Length Tapering      Off On
 Frequency-Offset Correction       Off On
 Bent-Wire Correction              Off On
 X-Axis Pattern Symmetry           No Yes
 2-D Pattern Resolution            1° 2° 4°
 El Angle for Az Pattern            0.0°
 Az Angle for El Pattern            0.0°
 SWR Reference Impedance           50.0 Ω
 Save Options & Settings           No Yes


 Gain & F/B    View     Sources     2-D Plot    Antenna
 Far Field     Edit     Frequency   3-D Plot    Options
 Near Field    Quit     Optimizer   RUN File    Colors

                        Enter Esc
```

Figure 1. AO's main menu with the "Options" menu overlaid.

```
 File      Edit      View      Change      Solve for      Options
I____I
┌─────────────────┐
│ Load...    F3   │
├─────────────────┤
│ Exit            │
├─────────────────┤
│ Create... F2    │         A N T E N N A      M O D E L
├─────────────────┤
│ Rename...       │
├─────────────────┤
│ DOS Shell       │
└─────────────────┘

        ────────────────────────────────────────────────────────
        Copyright (C) 1992 Teri Software Co.    All Rights Reserved

                          Version 1.2

        L. B. Cebik, W4RNL                         Serial 1I000003
        ────────────────────────────────────────────────────────




 There is enough memory for approximately 252 pulses
```

Figure 2. AM's opening screen with the "File" menu displayed.

impression that I favor one program over another. At the same time, relating features to applications will make some sense of the inevitable rank ordering of the programs under comparative scrutiny.

I have looked extensively at the following programs: ELNEC 3.02 (W7EL), AO 5.0 (an especially embellished version of MN by K6STI), and Antenna Model or AM 1.02 (Teri Software).[1] Other versions of MININEC have been and are still available.[2] MN and ELNEC have been around for several years, while AM is quite recent. For the record, I have purchased each of these programs at full price and have no overt or covert obligations to their designers. I am solely responsible for the emphases and applications-related evaluations that follow.

In looking at MININEC programs, I assume the reader has a basic familiarity with both the history of the code, its relationship to NEC, and the general limitations of its use, where these are common to all three programs.[3] The limitations especially have received significant mention in amateur literature.[4] Only those factors relevant to differences among the programs will be mentioned in the program descriptions. I have not included any sort of critical examination of the basic MININEC code.

MININEC subjects all its implementations to numerous general limitations. For example, the matrices used in MININEC's basic calculations grow exponentially with the number of wire segments or pulses, thus setting limits on the size or geometry of the antenna one may model. Wire segments should be greater in length than 10^{-4} wavelength and shorter than 10^{-1} wavelength. Segment lengths should also be greater than the wire diameter, perhaps at least 1.25 times greater. Results for horizontal antennas lower than 0.2 wavelength above a perfect or real ground are generally unreliable, or more precisely, depart from the results given by NEC by an amount that makes calculations and expectations unreliable. Adjacent segments should have a length ratio of 2:1 or less. As the frequency is increased, antenna wires with a diameter that is a significant part of a wavelength begin to yield results that depart from empirical models. Antenna geometries calling for nonlinear junctions of wires (bent wires; that is, elements with significant angles) yield inaccurate results without special treatment. Closely spaced wires, as in parallel transmission lines or folded dipole elements, also demand special treatment and may surpass the ability of MININEC to model accurately. Source impedances and resultant SWR figures are calculated over perfect ground and may vary from real antenna to real antenna, depending upon the nature of the soil below it. Even

though MININEC permits the modeling of elaborate ground structures, these models find use in only some of MININEC's calculations. This list of limitations is far from exhaustive.

Implementations of MININEC attempt to correct, compensate for, or overcome only some of these limitations. Thus far in the history of the code, wire segment ratios, angular antenna geometries, close-wire geometries, and frequency-related errors have been most subject to modifications of MININEC or to special antenna description procedures. At present, the remaining limitations appear to have one of two statuses: either they are not correctable within the limits proscribed when running MININEC on a PC, or the errors they introduce fall within the margins of model-to-reality variances inherent in amateur antenna modeling. Indeed, the best amateur antenna situation is filled with objects and conditions whose RF properties remain largely guesswork—even if we could model them.

Some MININEC program basics

MININEC is essentially a three-part program. In the first part, the user describes the antenna within a text editor. Key words and a carefully ordered input permit the second program part to recognize both variables and their values. The second part calculates most of the output values for the model, using the well-known method of moments and its requisite matrix. For a given model, the program calcu-

lates the impedances, mutual impedances, and currents at each pulse, centered on each wire segment junction. Specifying a source position and voltage or current, plus load values and positions, permits calculation of source impedances, SWR figures relative to a standard, and the remaining unknown—either source voltage or source current. Also calculable is gain in any direction in the sphere surrounding an antenna in free space or the hemisphere surrounding the antenna over ground, so long as one uses a standard, usually dBd or dBi. Both near field and far field electrical and magnetic components, along with polarization, are derivable, as well as total field values. From these numbers, one may obtain figures for such usual parameters as front-to-back ratio, front-to-side-lobe ratio, and beam width to the −3 dB points. Grounds may be perfect or may be one of many sets of conductivity and dielectric constant combinations. Antenna material losses for common substances, from copper to steel, can modify the calculations. The results are obtainable from a run or log file that compiles the values.

The third part of the program is a plotting portion that permits the user to obtain a graphic representation of antenna performance. Azimuth and elevation patterns, in either free space or over ground, permit visual inspection of lobe maxima and minima, thus giving an overall view of performance that may escape an inspection of data tables. Possible patterns include the common (sometimes called ARRL) logarithmic polar plot, the linear polar plot, and

```
            ELNEC  ver. 3.02 / MaxP option
           (c) 1993  by Roy Lewallen, W7EL
```

TI	TITLE:	edz yagi 2M 3/4" el			
FR	FREQUENCY:	144.5 MHz. (wavelength = 6.806721 ft.)			
WI	WIRES:	2 Wires	WL	WIRE LOSS:	Aluminum
SO	SOURCES:	1 Source	UN	UNITS:	Feet
LO	LOADS:	1 Load			
GT	GROUND TYPE:	Real		LAST FILE SVD/RCLD:	
GD	GND DESCRIPTION:	1 Medium, 0 Radials		D:\ELNEC\ZEPP\EDZYG2M.EN	
PT	PLOT TYPE:	Azimuth	AR	ANAL RES:	1 Deg.
PA	ELEVATION ANGLE:	0 Deg.	RF	REFERENCE:	0 dBi
PR	PLOT/TABLE RANGE:	0 - 360 Deg. (full)	SZ	SWR Z0:	75 ohms
SS	STEP SIZE:	5 Deg.			
OR	OUTER RING OF PLOT:	Automatic scaling			
FI	FIELD(S) TO PLOT:	Total field only			

```
(BR)owse file    (DE)lete, (RE)call, (SA)ve desc    (Freq S)wp    <RET> = Plot
(AN)alyze    (CU)rrents    (Load D)ata    (OP)tions    (Print D)esc    (Src D)ata
(TA)ble    (View A)ntenna         (EX)it pgm without saving desc    (QU)it
```

Figure 3. ELNEC's main screen with its network of selections.

Figure 4. A sample of AO's 3-D "solid" antenna pattern for a dipole.

in employing key words to cue the variable values. The user must keep the manual on hand, make up a cue card, or modify an existing antenna and save it under a new name. The last process is perhaps the most useful in AO, for two reasons. First, AO provides a large collection of models already set and plotted, so most common antennas can be described by changing values. Second, once the user has described either a complex antenna geometry or a complex set of ground conditions, saving these partial model descriptions reduces work with new model elements.

AM also uses a text editor for creating new antenna models. **Figure 2** displays the start of this process by showing the initial screen with the "File" function overlaid. In creating a description, AM's editor feature is bound into the operating and calculating module and does not permit the saving of a revised model under a new name from within the program. However, AM uses a dialogue box to set material, ground, and other basic factors, followed by a template for inserting the remaining values. (If one invokes the editor independently, it lacks the dialogue and template features.) This procedure reduces the need for a reference manual or cue card during the model description phase of the work, but it does not eliminate it—especially if one requires complex sources or loads. Both AO and AM permit the addition of noncalculation-related remarks and end notes for reference.

Both AO and AM keep a running log file (called RUN. in AO and [filename].LOG in AM) for calculations. Both fill and factor the MININEC matrix to produce mutual impedance values. At this point, AM proceeds to generate pattern values. AO prints impedance, wire loss, efficiency, gain, and front-to-back values before entering the pattern generation phase to derive values for angle of maximum radiation, beam width, and side lobe data. AM displays the entire collection at once. From here, both programs are ready to generate rapid antenna plots for viewing or printing. Both contain a command to combine these steps seamlessly, although the plot numbers go into a separate file. Numeric data—including a copy of the antenna description, the segment-to-segment location of pulses, a listing of currents at each pulse, and the noted output data—accumulate in the running log file.

ELNEC has a different method of operation, one that integrates all three MININEC program facets. ELNEC operation involves constant revision of a default file or the last file used in the previous session. **Figure 3** shows the network of data presented to the user. The user can recall from the provided models, or from models saved previously, a description nearest to

a rectangular plot. In descending order, they provide a proportioned notion of maximum antenna achievement; in ascending order, they provide an increased ability to investigate fine detail of small radiation lobes.

The three parts of MININEC yield three different files for each antenna model: a description file, a running file, and a plot file. MININEC accumulates calculated data in successive stages; therefore, the running file may be subdivided into a number of more specific data files. These MININEC components have to a large measure determined the structure of commercial implementations.

AO and MN, for example, retain this structure rigorously. On AO, antenna modeling begins with the main menu. **Figure 1** is a screen print of the menu showing the "Options" screen overlaid in the normally blank space on the operating screen. To create an antenna, one must enter the text editor (TED), either using the "Edit" command or by calling up TED independently, and construct the model. The editor offers no help in ordering the elements or

Parameter	ELNEC	AM	AO-1	AO-2
Antenna 1: 10 Meter dipole, #18 copper wire, 16.78', 10 segments/wire				
Free Space				
Gain (dBi)	2.050	2.05	2.05	2.04
Source (R±X)	71.8 + 0.01	71.8 + 1.15	71.8 + 0.0	72.4 + 3.2
B/W (degrees)	78	78.7	82	79
Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'				
Gain (dBi)	7.654	7.65	7.65	7.64
TO Angle (degrees)	14	14	14	14
Source (R±X)	68.7 - 8.3	68.7 - 7.2	68.7 - 8.3	69.3 - 5.2
B/W (degrees)	80	79.9	80	80
Antenna 2: 2 Meter dipole, #18 copper wire, 3.289', 10 segments/wire				
Free Space				
Gain (dBi)	2.087	2.09	2.08	2.07
Source (R±X)	71.0 + 0.05	71.0 + 0.43	71.1 + 0.0	72.3 + 5.2
B/W (degrees)	78	78.7	79	79
Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'				
Gain (dBi)	8.021	8.02	8.02	8.00
TO Angle (degrees)	3	3	3	3
Source (R±X)	69.3 + 0.61	69.3 + 1.01	69.4 + 0.5	70.6 + 5.8
B/W (degrees)	78	78.8	79	79
Antenna 3: 2 Meter dipole, 6061-T6 Aluminum, 3.196', 10 segments/wire				
Free Space				
Gain (dBi)	2.113	2.11	2.11	2.05
Source (R±X)	71.2 + 0.02	71.2 + 0.5	71.3 + 0.0	77.4 + 10.8
B/W (degrees)	78	78.9	79	79
Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'				
Gain (dBi)	8.048	8.05	8.04	7.98
TO Angle (degrees)	3	3	3	3
Source (R±X)	69.6 + 0.83	69.6 + 0.86	69.7 + 0.8	75.7 + 11.7
B/W (degrees)	78	78.9	79	78

Table 1. Comparison of EL, AM, and AO for dipoles.

what is required. All variables appear on the main screen or one of the subscreens. Thus, reference manuals or cards are not required except when looking up specific values like the dielectric constant and conductivity for a particular type of soil. ELNEC saves nothing without orders from the user, except the last file of a session. The user specifies a file name within a subdirectory and can choose a new name for a model revision.

Subscreens in ELNEC for antenna wires, sources, and loads permit the deletion of excess items from the present record and the addition of "0" value items into which the user plugs appropriate values. Ground definition is per-

formed on a subscreen of its own. Antenna material, units of measure, and azimuth and elevation angles for plots appear on the main screen. The initiation of calculations produces matrix filling and current calculations, along with the creation of a graphics plot—unless the user specifically chooses a tabular output screen. ELNEC stores calculation results in various files, permitting screen or printer output of data subsets. Plots may be saved separately from the graphics screen. If the user saves an antenna description after calculations are completed, he or she has the option to save the matrix calculation results in a collection of files. However, one must save either all or none

Parameter	ELNEC	AM	AO-1	AO-2

Antenna 1: 3-element Yagi for 14.175 MHz (taken from AO file), 1" 6061-T6 Aluminum

Free Space

Parameter	ELNEC	AM	AO-1	AO-2
Gain (dBi)	7.988	7.99	7.97	8.04
Front-to-Back (dB)	25.115	25.24	24.91	28.14
Source (R±X)	27.3 - 1.12	27.3 - 0.98	27.5 - 1.1	26.5 + 4.6
B/W (degrees)	64	63.5	64	63
Run Time (seconds)	12	15	9	9

Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 416"

Parameter	ELNEC	AM	AO-1	AO-2
Gain (dBi)	12.167	12.16	12.15	12.18
TO Angle (degrees)	25	25	25	25
Front-to-Back (dB)	27.413	27.41	27.42	24.36
Source (R±X)	25.40 - 1.58	25.38 - 1.44	25.5 - 1.6	24.8 + 4.6
B/W (degrees)	66	65.7	66	65
Run Time (seconds)	20	40	30	30

Antenna 2: 3-element Yagi for 28.2 MHz; 1" 6061-T6 Aluminum

Free Space

Parameter	ELNEC	AM	AO-1	AO-2
Gain (dBi)	7.540	7.54	7.5	7.61
Front-to-Back (dB)	19.595	19.63	19.03	21.07
Source (R±X)	16.24 - 28.63	16.2 - 28.55	17.2 - 29.5	12.9 - 19.9
B/W (degrees)	64	64	65	64
Run Time (seconds)	20	15	15	15

Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'

Parameter	ELNEC	AM	AO-1	AO-2
Gain (dBi)	12.842	12.84	12.80	12.88
TO Angle (degrees)	14	14	14	14
Front-to-Back (dB)	24.145	24.19	23.25	25.47
Source (R±X)	15.7 - 27.8	15.7 - 27.7	16.6 - 28.7	12.7 - 19.1
B/W (degrees)	64	64.8	65	65
Run Time (seconds)	34	52	30	30

Table 2. Comparison of EL, AM, and AO for 3-element Yagis.

of these supplementary files. For complex geometries (for example, a fully tapered multi-element quad), this can run to over 100 KB of disk space per model.

These basic program descriptions do not exhaust the lists of program features, but simply give a flavor of the differences in programming style of each developer. AO remains perhaps the most traditionally structured of the MININEC implementations, but adds numerous individual calculation and graphic features. It offers the most field and graphic options, including rectangular coordinates and complete near and far field calculations and views. These even include a slow but impressive three-dimensional view of the total antenna field. See **Figure 4** for an example that plots a dipole in 3-D. K6STI has introduced modifications to the MININEC algorithms to correct for frequency-induced errors, to correct for bent-wire configurations (without the need for segment tapering), and to taper wire junction segments for adjoin-

ing linear segments that surpass the recommended maximum length ratio. Evident is the author's focus upon single source horizontal antennas, as witnessed by recommendations for the use of free space as the fundamental modeling convention, the use of default voltage sources, and the use of dBd(ipole) as the default standard of antenna gain.

In contrast, ELNEC's integrated programming provides easy access to specific data on screen and in print. Additionally, it provides the most data-intensive graphic output. At the user's choice, one may include antenna analysis figures on the graphic itself and in a succinct chart printed below the graphic. ELNEC includes segment tapering for complex antenna geometries: the user can control the maximum and minimum lengths of the segments, but the program then creates the requisite wires automatically. With dBi(sotropic) the gain standard and current sources the default, W7EL records his own focus upon phased vertical antenna

Frequency	Parameter	AM[1]	AO-2
14.000	Gain (dBi)	12.01	12.03
	Front-to-Back (dB)	19.10	23.63
	Source Impedance (R±X)	25.4 - 14.1	25.6 - 8.5
14.0025	Gain (dBi)	12.03	12.05
	Front-to-Back (dB)	20.40	25.24
	Source Impedance (R±X)	25.5 - 12.3	25.6 - 6.6
14.050	Gain (dBi)	12.06	12.07
	Front-to-Back (dB)	21.86	26.73
	Source Impedance (R±X)	25.6 - 10.6	25.5 - 4.8
14.075	Gain (dBi)	12.08	12.09
	Front-to-Back (dB)	23.37	27.72
	Source Impedance (R±X)	25.6 - 8.8	25.4 - 3.0
14.100	Gain (dBi)	12.10	12.11
	Front-to-Back (dB)	24.92	27.81[2]
	Source Impedance (R±X)	25.6 - 6.9	25.3 - 1.1
14.125	Gain (dBi)	12.12	12.14
	Front-to-Back (dB)	26.33	27.02
	Source Impedance (R±X)	25.6 - 5.1	25.1 - 0.7[3]
14.150	Gain (dBi)	12.14	12.16
	Front-to-Back (dB)	27.28	25.74
	Source Impedance (R±X)	25.5 - 3.3	25.0 + 2.6
14.175	Gain (dBi)	12.16	12.18
	Front-to-Back (dB)	27.41[2]	24.36
	Source Impedance (R±X)	25.4 - 1.4	24.8 + 4.6
14.200	Gain (dBi)	12.18	12.20
	Front-to-Back (dB)	26.70	23.03
	Source Impedance (R±X)	25.3 + 0.4[3]	24.6 + 6.5
14.225	Gain (dBi)	12.20	12.22
	Front-to-Back (dB)	25.52	21.81
	Source Impedance (R±X)	25.1 + 2.3	24.3 + 8.5
14.250	Gain (dBi)	12.22	12.24
	Front-to-Back (dB)	24.20	20.71
	Source Impedance (R±X)	24.9 + 4.2	24.1 + 10.5
14.275	Gain (dBi)	12.24	12.26
	Front-to-Back (dB)	22.92	19.70
	Source Impedance (R±X)	24.7 + 6.2	23.8 + 12.6
14.300	Gain (dBi)	12.26	12.28
	Front-to-Back (dB)	21.74	18.79
	Source Impedance (R±X)	24.5 + 8.2	23.5 + 14.6

Note 1. ELNEC and AM figures are too close to require separate listings.
Note 2. Point of maximum front-to-back ratio.
Note 3. Nearest point to self-resonance.

Table 3. Comparison of 3-element Yagi values from 14.00 to 14.30 MHz.

arrays erected over radial and ground systems requiring accurate specification.

AM shows no focus upon any particular antenna type, but rather a concern for programming itself. It incorporates structured menus in a top bar with "hot" function keys for frequently used operations. Its manual chapter on operations tracks the top bar menus, so the user will find it simple to create a reference card. The program has the operator feel of a spreadsheet with touches of GUI interfaces. Although AM lacks any of the calculation compensation features of either ELNEC or AO, it includes a special graphics program to permit plotting patterns using the HPGL feature of the HP Laserjet III's PCL5 language, thus improving the printed output quality. It also includes a DOS shell to permit various operations without terminating the program.

Like the brochures for new cars, these basic
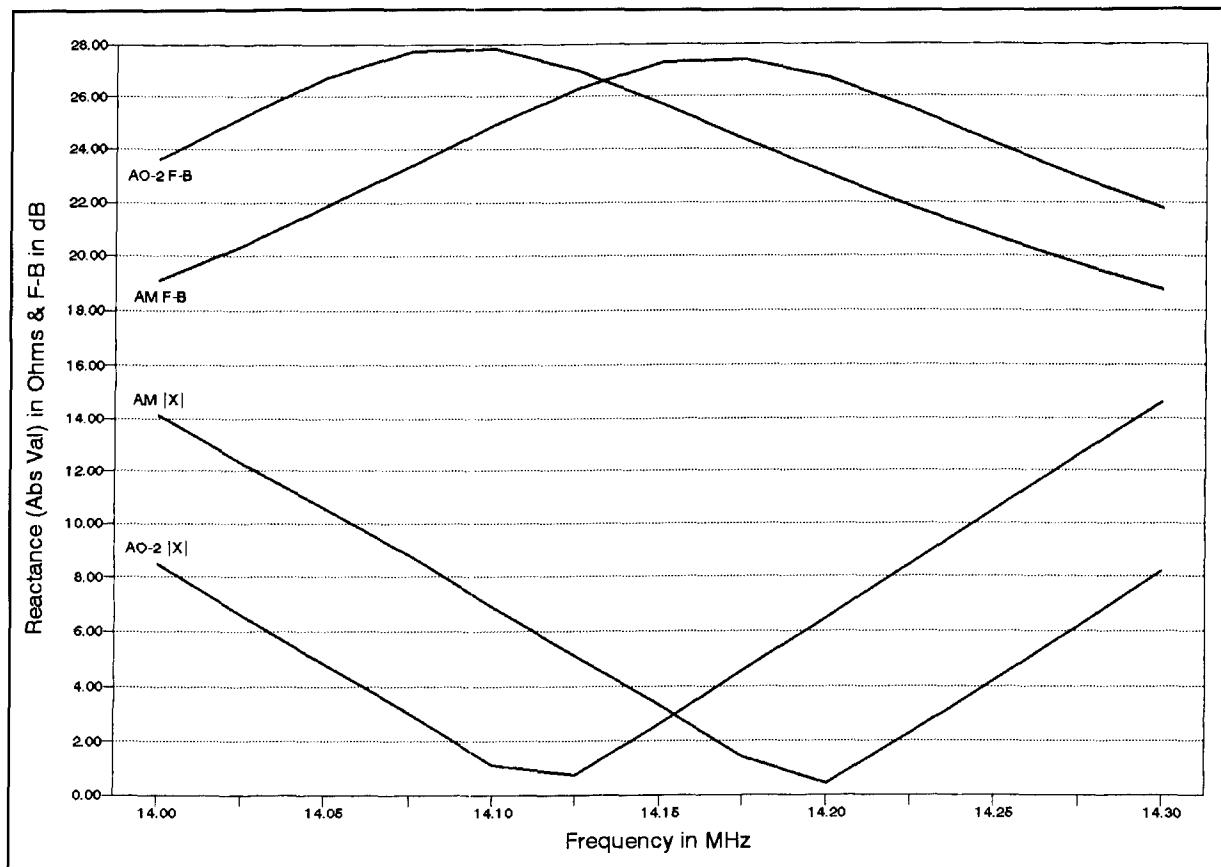
Figure 5. A graph of front-to-back ratios and absolute values of reactance for the 20-meter Yagi between 14.00 and 14.30 MHz as calculated by ELNEC and by AO with frequency correction.

descriptions of AO, ELNEC, and AM may spur some dreams about antenna modeling, but they cannot substitute for a test drive. Therefore, it's important to try out each of the programs: first to assess their technical competencies in various situations, and then to note operational advantages and limitations.

Technical competencies

There is very little the user can accomplish with one program that he or she cannot do with the others. After all, ELNEC, AO, and AM use and are subject to the limitations of the MINI-NEC calculation machine. Technical competency within each program boils down to a matter of how well each performs the basic MININEC functions, and how well each overcomes some of its limits relative to various user tasks. For casual investigation of antenna properties, each produces approximations sufficiently close to empiric reality and the NEC2 standard.

However, many hams wish to study antennas and many wish to build antennas, and a few want to do both. This increases demand for precision. One source of imprecision within MININEC relates to frequency: as frequency increas-

es and the diameters of wire elements become a more significant percentage of a wavelength, basic MININEC calculations begin to depart from the demands of the real world. AO (and MN) contain correction factors within their calculational schemes to compensate for these errors. ELNEC and AM do not. The significance of the output difference deserved a test. **Tables 1** through **6** contain the results of some initial runs, both in free space and over real ground. All runs were made on a 80386SX computer running under DOS 5.0 at 20 MHz with 4 MB of RAM under QEMM 6.02 control to clear calculation space, with 160 MB of hard disk to absorb created files. Any run times noted are relative to this test set-up.

Table 1 lists three dipoles. The first is a 10-meter wire model, the second a 2-meter wire model, and the third a 2-meter aluminum tube model. The AO1 and AO2 columns represent modeling without and with the correction factors turned on. ELNEC and AM results are consistently close; variations are the result of either the order of operations or the fixed values of certain constants, for example, pi or the speed of "light" (electromagnetic radiation in a vacuum). The wire dipoles, even at 2 meters, show

too little difference under any condition to merit distinction. Attaching the wire to end insulators will likely create more variance from the models than the amount of difference created by the frequency correction factor in AO. Differences of greater significance begin to show up with the 1-inch aluminum tube model. While the gain in dBi and the resistive component of the input impedance fall well within operational margins, the inductive reactance shown in AO2 suggests that raw MININEC may call for unduly long fat elements at VHF.

Table 2 tests more complex antennas, namely, 3-element Yagis. A 20-meter beam, taken from a file supplied by AO and slightly modified for this test, reveals that 14 MHz does not significantly display the frequency error. However, when the design is tested across a span of frequencies (refer to Table 3 and Figure 5), the effects of the difference between corrected AO and raw MININEC begin to appear even at 20 meters. Various parameters

show an approximate 75 kHz displacement, with corrected AO assigning them the lower frequency. Determining whether the displacement precision exceeds amateur construction abilities would require empirical tests beyond the capabilities of my shop.

The 10-meter Yagi also shows little specific frequency effect that goes beyond the margins imposed by building technique, surrounding terrain, and other normally unmodeled effects. However, an interesting arithmetic pattern begins to emerge. Compared to raw MININEC, models optimized initially in corrected AO typically show higher gains, lower front-to-back ratios, lower resistive source impedances, and more inductively reactive source impedances. However, while the results for ELNEC and AM are quite tightly clustered, uncorrected AO begins to depart from the cluster—usually in a direction opposite that of corrected AO.

Nonstandard antennas offer another test of the correction factor. Arrays and parasitic

Parameter	ELNEC	AM	AO-1	AO-2
Antenna 1: 2-element EDZ for 28.5 MHz, 180-degree phase-fed: #18 Copper wire				
Free Space				
Gain (dBi)	7.779	7.79	7.78	7.78
Source (R±X) (x2)	22.51 - 741	22.25 - 736	22.3 - 737	21.9 - 726
B/W (degrees)	30	29.8	30	30
Run Time (seconds)	27	15	35	35
Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'				
Gain (dBi)	12.894	12.9	12.90	12.89
TO Angle (degrees)	13	13	13	13
Source (R±X) (x2)	23.00 - 740	22.7 - 737	22.8 - 737	22.4 - 726
B/W (degrees)	30	30.4	30	30
Run Time (seconds)	62	94	122	122
Antenna 2: 2-element Yagi EDZ for 144.5 MHz, 0.75" 6061-T6 Aluminum				
Free Space				
Reflector Load (Ω)	305	305	305	280
Gain (dBi)	9.199	9.2	9.23	9.04
Front-to-Back (dB)	22.246	22.27	22.01	22.60
Source (R±X)	50.61 - 289.6	50.63 - 289.6	51.2 - 293	48.7 - 267
B/W (degrees)	32	31.5	31	31
Run Time (seconds)	43	29	45	45
Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'				
Reflector Load (Ω)	300	300	300	275
Gain (dBi)	15.196	15.19	15.23	15.05
TO Angle (degrees)	3	3	3	3
Front-to-Back (dB)	22.390	22.42	22.11	22.64
Source (R±X)	49.96 - 291.2	49.98 - 291.1	50.5 - 294	48.0 - 268
B/W (degrees)	32	31.4	31	31
Run Time (seconds)	80	90	137	137

Table 4. Comparison of EL, AM, and AO for Multi-Element Extended Double Zepps.

Parameter	ELNEC[1]	AM[2]	AO-1[3]	AO-2[4]

2-element Quad, 28.5 MHz, #18 copper (Elements: 990/f & 1037/f at 0.25 wavelength spacing)

Free Space

	ELNEC[1]	AM[2]	AO-1[3]	AO-2[4]
Gain (dBi)	7.252	7.45	7.32	7.25
Front-to-Back (dB)	19.580	14.86	18.70	20.87
Source (R±X)	88.61 + 4.68	74.5 - 13.9	85.1 - 1.9	90.1 + 3.5
B/W (degrees)	72	71.9	73	73
Run Time (minutes)	4:03	7:12	1:34	1:34

Real Ground: Cond.=5 mS/m; Dia. Const.=13; Height = 35'

	ELNEC[1]	AM[2]	AO-1[3]	AO-2[4]
Gain (dBi)	12.443	12.62	12.50	12.44
TO Angle (degrees)	13	13	13	13
Front-to-Back (dB)	17.924	14.38	17.37	18.82
Source (R±X)	91.89 + 7.69	76.7 -11.71.1	88.2 + 1.0	93.7 + 6.6
B/W (degrees)	74	72.2	73	73
Run Time (minutes)	6:53	11:57	7:31	7:31

Note 1. ELNEC model used fully tapered elements for 104 total segments.

Note 2. Antenna Model was tested at increasing numbers of segments, with an arbitrary cut-off (due to run times) of 24/side.

Note 3. AO without bent-wire correction factors used 22 segments per half wavelength.

Note 4. AO with bent-wire correction factors approached good convergence with ELNEC at 20 wires per half wavelength (10 per side). Over ground calculation time, minus the pattern generation necessary to determine beamwidth and proper angle for maximum gain, was 1:01 minutes. However, the take-off angle was known in advance, precluding the need for a second run, the first to find the proper elevation angle and the second to determine the horizontal bandwidth at that angle.

Table 5. Comparison of EL, AM. and AO for a 2-element quad.

beams based upon extended double Zepp elements form the basis for **Table 4**. The wire array of the 180-degree phase-fed antenna again demonstrates the minimal effect of frequency upon thin wires: all the antennas are well clustered in every dimension calculated. The 2-element EDZ parasitic beam, however, tells a different story. The principles behind the antenna were first brought to my attention by Brian Egan, ZL1LE. The 3/4-inch diameter aluminum model for potential 2-meter use employs close (1/8-wavelength) spacing for mechanical reasons, and comes in at just over 8 feet from end-to-end. The reflector element requires a loading inductance, hand optimized to the nearest 5 ohms in all the models. Again, ELNEC and AM produce tightly clustered models. Frequency-corrected AO gives results that depart from the original model. Because the EDZ Yagi is purposely nonself-resonant, dimensions could be held constant, and the reflector loading varied for optimum potential performance. Again, corrected AO suggests that this "fat" element antenna is longer than raw MININEC would suggest. The requisite inductive loading to compensate for the capaci-

tive component of the reflector is lower, as is the capacitive reactance at the feed point. (As the length of an EDZ, about 1.25 wavelengths, increases, capacitive reactance at a center feed-point decreases.) Nonetheless, the uncorrected AO shows the minor arithmetical properties noted above relative to the tightly clustered ELNEC and AM results.

Table 5 shows some results for a 2-element quad design. Raw MININEC requires some extensive help to model nonlinear elements. Complex geometries with angles, especially right or acute angles, cut off parts of the antennas at the corners due to the placement of pulses. There are two ways to overcome this difficulty. One is to introduce a bent wire correction factor, as AO has done. The other is to develop a means of tapering the length of wire segments so they are very short (but within the rules) at the apex of the angles. ELNEC has adopted this technique, providing a routine that puts the minimum and maximum length of wire segments under operator control (with suggested default values), but which then performs the tapering automatically.

ELNEC quad models with fully tapered ele-

ments are large—104 total segments in the 10-meter example. The free space model required a bit over 4 minutes to run, while the real-ground model needed nearly 7 minutes. However, without either tapering or a correction factor, AM was still far from convergent with either ELNEC or AO at 24 segments per side or 192 total segments. Even with 24 segments per side, the corner segments were about twice the length of the fully tapered ELNEC corner elements. (However, AM handled 192 segments without need for program supplements.) The AM run took nearly 7 minutes for free space and 12 minutes over real ground. ELNEC's tapering produced a model that appears to agree closely with a similar antenna empirically derived long ago by Bill Orr, W6SAI.[5] The spacing of this model was increased for greater bandwidth, and the reflector circumference was increased to remove the stub in the original model.

In contrast to ELNEC's provision for tapering segment lengths, AO uses a bent wire cor-

rection factor in the calculations. The manual recommends the use of 16 segments per half wavelength for quads and 20 segments per half wavelength for deltas. However, for this quad model, AO converged most closely with ELNEC's tapered elements at 20 segments per half wavelength, suggesting the use of even more for delta beams. ELNEC's 104-segment model and AO's 80-segment model, although not as tightly clustered arithmetically as linear antenna models, likely exceed the ability of any builder to differentiate.

Whether AO performs its calculations faster or slower than ELNEC depends upon the user's situation. All free space AO quad models run faster than their ELNEC cousins by better than a 2 to 1 margin—even allowing for the pattern generation that precedes calculation of beamwidth. Real ground models, however, will be faster or slower depending upon the data the user needs. Because AO produces an azimuth pattern before an elevation pattern, and since

Parameter:	Gain dBi	F-B Ratio dB	Source 1 R ± X	Source 2 R ± X	Run Time Min:Sec

Antenna #1: 10-meter folded dipole: L = 16.5'; Dipole wire = 1" spacing; Real ground (Cond.=5 mS/m; Dia. Const.=13); Height = 20'

AM: 40 segments per 1/2 wavelength (84 total segments)

	Gain dBi	F-B Ratio dB	Source 1 R ± X	Source 2 R ± X	Run Time Min:Sec
	8.18	---	227.5 - 17.7	---	2:00

AO: 4 wires, 34 pulses autotapered

| | 8.09 | --- | 229 + 22 | --- | 2:10* |

*Run time doubled to over 4 min, due to requirement for a first run to find elevation angle of maximum radiation.

ELNEC:

No taper: 4 wires, 24 segments

a. Without parallel correction

| | 7.970 | --- | 28.0 -194.0 | | :19 |

b. With parallel correction

| | 8.104 | --- | 224.0 + 40.1 | | :19 |

Full taper: 24 wires, 38 segments

| | 8.119 | --- | 227.4 + 12.2 | | :51 |

Antenna #2: 10-meter W7EL Field Day Special: 2 folded dipoles 135-degree phase-fed; L = 16.16'; Dipole wires = 1" spacing; Elements spacing = 4.27'; Real ground (Cond.=5 mS/m; Dia. Const.=13); Height = 20'

ELNEC:

No taper: 8 wires, 48 segments

a. Without parallel correction

| | 11.182 | 4.724 | 29.9 -172.6 | - 5.2 -180.2 | 1:13 |

b. With parallel correction

| | 11.647 | 24.231 | 99.8 + 97.1 | 80.5 -168.8 | 1:13 |

Full taper: 48 wires, 76 segments

| | 11.669 | 25.328 | 108.1 + 72.2 | 69.7 -188.6 | 3:30 |

Table 6. Comparison of EL, AM, and AO for a folded dipole.

```
-------------- WIRES ---------------

Wire Conn. --- End 1 (x,y,z : ft)  Conn. --- End 2 (x,y,z : ft)  Dia(in)   Seg

  1          -5.230,  -8.670, 35.000  W2E1  -5.230,  -6.000, 35.000  8.75E-01   2
  2   W1E2   -5.230,  -6.000, 35.000  W3E1  -5.230,   6.000, 35.000  1.00E+00   8
  3   W2E2   -5.230,   6.000, 35.000        -5.230,   8.670, 35.000  8.75E-01   2
  4           0.000,  -8.310, 35.000  W5E1   0.000,  -6.000, 35.000  8.75E-01   2
  5   W4E2    0.000,  -6.000, 35.000  W6E1   0.000,   6.000, 35.000  1.00E+00   8
  6   W5E2    0.000,   6.000, 35.000         0.000,   8.310, 35.000  8.75E-01   2
  7           3.490,  -8.130, 35.000  W8E1   3.490,  -6.000, 35.000  8.75E-01   2
  8   W7E2    3.490,  -6.000, 35.000  W9E1   3.490,   6.000, 35.000  1.00E+00   8
  9   W8E2    3.490,   6.000, 35.000         3.490,   8.130, 35.000  8.75E-01   2
                                                           Tot segs:  36
       End 1 (x, y, z in ft.; conn.; or length)?




    [x],[y],[z] = Coordinates (leave coord blank for no change)
    W#E! = Conn. to end ! of wire #      RA# = Rotate azimuth # deg. (+ccw)
    L# = Change length to #              RE# = Rotate elev. # deg. (+up)
    L+#, L-# = Change length by #

Preserve Connections = OFF                    <ESC> = done (terminate entry)
```

Figure 6. A typical wires subscreen from ELNEC, with wire 4 ready for modification.
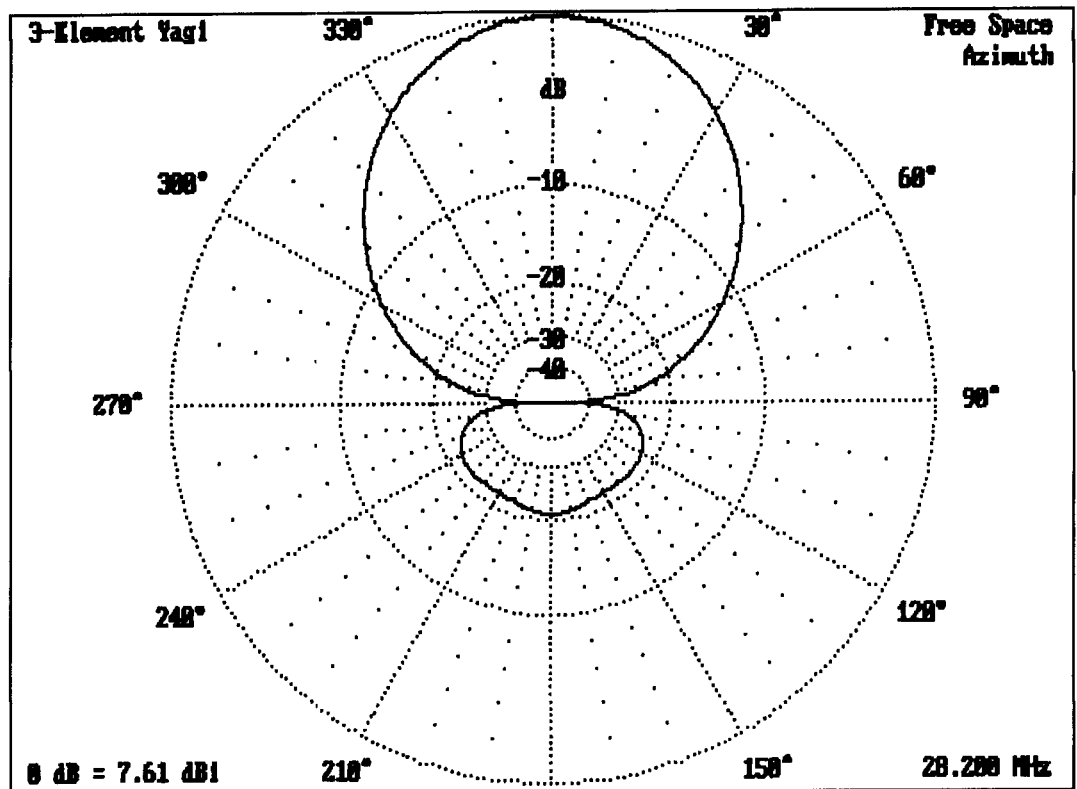
Figure 7. Pattern for a 10-meter 3-element Yagi as plotted by AO.

the latter pattern identifies the take-off angle— that is, the angle of maximum radiation—one often needs to make two pattern runs before obtaining maximum gain and front-to-back figures. The matrix fill and factor task took about a minute, but pattern generation required another five minutes to complete. Doubled, the time for full data approached the largest run time encountered for AM. (AM, incidentally, precedes the azimuth pattern generation with an elevation run, thus identifying the take-off angle and setting the azimuth pattern to this angle. In ELNEC, one selects the initial pattern, usually elevation over ground, and quickly
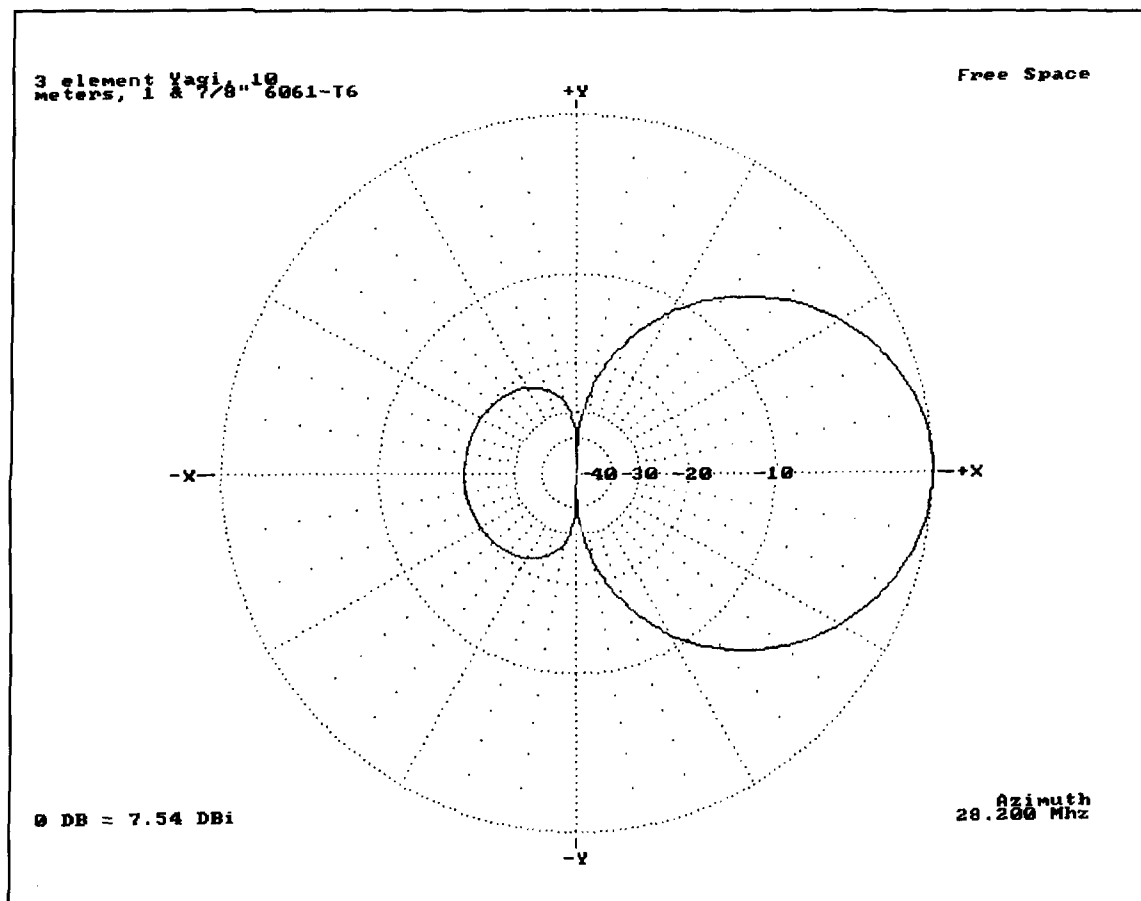
Figure 8. Pattern for a 10-meter 3-element Yagi as plotted by AM.

determines the take-off angle to use for the azimuth pattern from which is derived front-to-back and side-lobe figures.)

One complex geometry problem involves antennas that use closely spaced parallel wires. As a test of the performance of the three programs on this sort of structure, I modeled a 10-meter folded dipole. **Table 6** shows the results. To arrive at a usable result, AM employs the brute force technique of simply increasing the number of wire segments: 84 in all. AO uses an automatic segment-length tapering feature for such geometries. ELNEC provides two methods. One is full element tapering, as used with the quad. This accurate but slow technique is supplemented in the most recent version (3.02) with a close parallel wire correction algorithm that requires no increase in the number of wire segments per half wavelength. Comparative figures are shown in the table with and without the correction feature turned on, along with numbers for a fully tapered version of the same model. The dramatic speed increase allows more efficient model development, with the taper feature available for spot checks to ensure accuracy. Although the feedpoint resistance is similar in all four cases (AM, AO-taper,

ELNEC-wire correction, and ELNEC-taper), the drift in the reactance figure reflects the differences in the techniques used by each program to achieve functional results.

To further test the difference between tapering and close-wire correcting in ELNEC, I modeled a scaled version of W7EL's Field Day Special (a cousin of the ZL Special). The differences in source reactance appear again, along with a difference in the resistive component of the rear element source impedance. Nevertheless, the 3-to-1 modeling speed increase shows the benefits of implementing corrections in calculation algorithms rather than in user modeling techniques.

Calculation speed may not be the decisive factor in overall program performance. While AO calculates the fastest free space complex antennas, with or without the symmetry shortcuts, it requires one to know or guess the take-off angle to match ELNEC over real ground. Without either the tapered element or bent-wire techniques used by its MININEC relatives, AM requires an exceptionally large number of wires in antennas with complex geometries in order to converge with results from the other two programs. Nonetheless, there are a number of
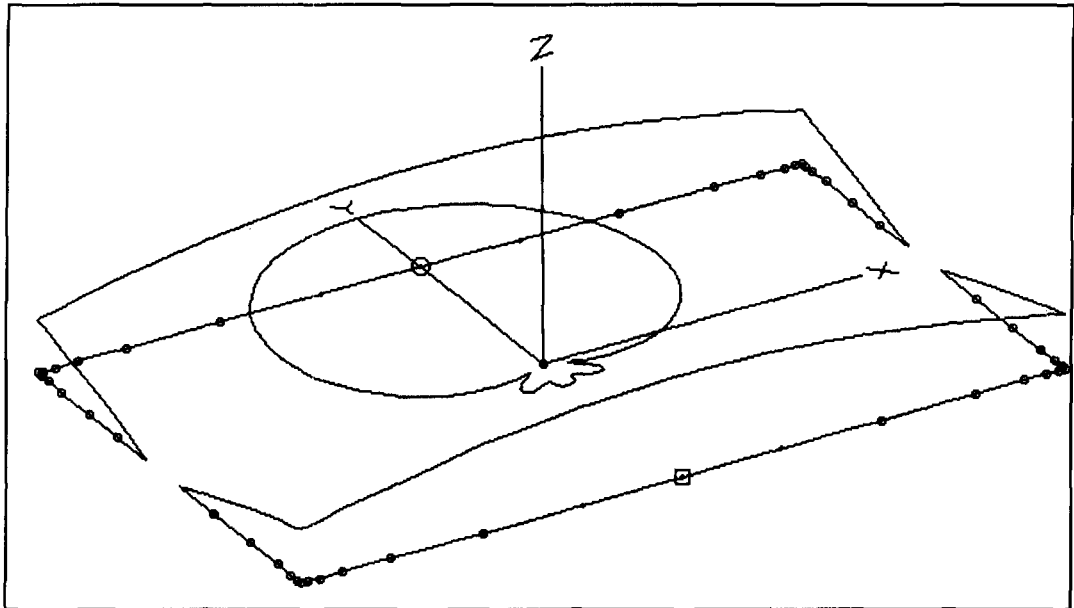
Figure 9. A sample antenna view graphic from ELNEC; AO's graphics are quite similar (without the azimuth pattern). Shown is the antenna, wire segments, source, load, current, and the azimuth pattern. AM supports a 3-D antenna view without current distribution or pattern graphics.

noncalculational factors that contribute to a program's suitability. We might think of these as resulting in an overall efficiency of performance for particular kinds of tasks rather than as raw calculation speed alone. Most of these factors belong in the category of user essentials and conveniences.

Operating conveniences

Program structure features fall into several separable categories. You can skip over those of no relevance to your proposed operation.

Antenna Set-up: There is no single best way to set up an antenna description. However, the three programs do offer some variations that the user can adapt to different situations. With these variations come some limitations.

ELNEC offers a wire subscreen for setting the X, Y, and Z coordinates of the pieces of the antenna elements. All entries are numeric, with some pleasant shortcuts. For example, for a quad, one need only specify a starting point set of coordinates. From there, using wire lengths, angles of rotation, and designators such as "Wire 1, End 2," it is possible to construct the remainder of the antenna. These mechanisms are handy for delta beams and other antennas with complex geometries, where a nonlinear intersection might result in a failure of wires to meet by an imperceptible amount. **Figure 6** shows a typical wire screen in the process of adjusting wire 4. For complex antenna geometries, ELNEC offers a "wire group" feature that permits simultaneous variation of selected para-

meters on a subset of the total number of antenna wires. This feature fulfills essentially the same function as the use of variables in AO and AM, which retain the text-editor antenna description technique of MININEC itself.

ELNEC once used a triple-plane view of the antenna to check one's work, but version 3.02 now offers a three-dimensional graphic view of the antenna, its current pattern, and its total field (if the antenna is viewed after calculations). Antennas can use either the X or Y axis for their lobe of greatest gain without slowing calculations or losing any analytical data.

To take advantage of symmetry in free space to speed calculations by simply mirror-imaging half the antenna, AM and AO require that the main lobe follow the X-axis.[6] Interestingly, AO plots place the X-axis 0-degree point at the top of the page, while AM places that point to the right, as illustrated in **Figures 7** and **8**. The use of symmetry invokes some antenna set-up restrictions, which are slightly different for each program. The feature shuts itself off over ground. Moreover, both AM and AO permit the use of symbols as variables or place holders in the wire X-Y-Z layout, with a table of equations to tell the program what numeric values to use. Changing antenna element lengths, spacing, and diameters—as well as the number of segments per element—thus becomes a simple matter of changing the value of one or more symbols. AO's symbolic power appears stronger than AM's, because AO allows the use of complex formulae. AO also permits wire rotation for constructing complex geometries. In contrast, AM offers an on-screen calculator whose values can

be spliced into the antenna description, even if the numbers have more significant digits than the display has places to show them. Like ELNEC, both AM and AO offer three-dimensional views of the antenna geometry. **Figure 9** shows a graphic antenna view from ELNEC, but AO is quite similar. AM's geometry view lacks the display of currents.

AO and AM antenna description files can be used with both programs with only minor modifications to satisfy each program's particular limitations and rules. **Figure 10** shows an AO description file set out in symbolic dimensions and fully annotated. Very little change would be needed to run it on AM. In both programs, description files are simple ASCII enumerations of variables and values. Pattern plotting files are fully transportable between AO and AM simply by changing the extensions in the filenames (.PLT vs. .PAT). ELNEC's integration in the form of a compiled BASIC program also puts its description and plot files in non-ASCII form, effectively precluding intercommunication.

Printing: Of all the packages, ELNEC offers the most convenience in producing select printed outputs from various modes of analysis. Its graphic output screen (a "Print Screen" function) prints the selected pattern in one of several forms chosen either in the options shell or on the fly. One can set the polar patterns for either linear or logarithmic scales and choose to include or exclude border data. One can also let the key analytic factors appear on the pattern for quick reference or appear only in a box below the pattern. In this way, most key data, except the detailed antenna description, appear on a single sheet, as illustrated in **Figure 11**. In addition, separate screen and printer readouts are available for the antenna description, a table of currents, the source data, and the analysis. This selectivity allows the user to commit only needed data to paper. Despite this versatility, ELNEC has one odd shortcoming: no error trapping on print requests. A printer error during a run will crash the program, forcing the user to restart and find the "LAST" file to resume operations.

With the exception of the graphic output, neither AM nor AO offer sophisticated print options. AO requires a batch file containing certain SET functions for printer selection at start-up. Not all the instruction sets work with supposedly compatible printers.[7] Print control uses no page end instruction, but relies on the operator to control the printer. All data (including a detailed antenna description and the cal-

```
2-Element Yagi for 10-meters        ;antenna title
free space symmetric                ;environment/ground description
28.20 MHz                           ;frequency
8 wires, feet                       ;# wires, units of measure
sea = 4                             ;symbol definitions
seb = 8                             ;sea/b = # segments
dia = .875"                         ;dia/b = element diameters
dib = 1"
ht = 0                              ;ht = height
sp = -4.25'                         ;sp = spacing
c = 6'                              ;c = fixed center section length
de = 7.833'                         ;de = driven element length
re = 8.750'                         ;re = reflector element length
sea    0 -de   ht    0 -c  ht   dia   ; wire chart
seb    0 -c    ht    0  0  ht   dib
seb    0  0    ht    0  c  ht   dib
sea    0  c    ht    0 de  ht   dia
sea   sp -re   ht   sp -c  ht   dia
seb   sp -c    ht   sp  0  ht   dib
seb   sp  0    ht   sp  c  ht   dib
sea   sp  c    ht   sp re  ht   dia
1 source                            ;# sources (default = voltage)
Wire 3, end1  1  0                   ;placement of source, value, phase

;from Bill Orr, "A Compact 2-Element Yagi for 10 Meters,"
    CQ, Dec., 1990, pp. 83-84.
```

Figure 10. An annotated AO antenna description with symbolic dimensions, easily modified for use in AM.

3el Yagi 10M 1"+7/8"   0 dB   ELNEC 3.02

05-24-1993 20:15:02
Freq = 28.2 MHz

-10

-20

-30

0 deg.

Gain: 7.540 dBi
Angle: 0 deg
F/B: 19.595 dB
Bmwidth: 64 deg
-3dB: 328, 32 deg
Slobe: -12.055 dBi
Angle: 180 deg
F/Slobe: 19.595 dB

Outer Ring = 7.540 dBi
Max. Gain = 7.540 dBi

Azimuth Plot
Elevation Angle = 0.0 deg.

```
| W7EL    ELNEC 3.02     3el Yagi 10M 1"+7/8"           05-24-1993  20:15:02 |

| Azimuth plot           | Frequency = 28.2 MHz   | Max Gain = 7.540 dBi      |
| Elev angle = 0.0 deg   | Peak Ht = 35.000 ft    | Source 1 impedance =      |
| Outer ring = 7.540 dBi | Wire Loss: Aluminum    |   16.237 - J 28.628       |
| Reference = 0 dBi      |   Rstvty = 4E-08       | SWR (50 ohm) = 4.174      |
|                        |   Rel Perm = 1         |     (75 ohm) = 5.321      |

| 9 Wires                | Free space             | Gain: 7.540 dBi           |
|   Wire 1: 2 segments   |                        | Angle: 0 deg              |
|     Length = 2.670 ft  |                        | F/B: 19.595 dB            |
| 1 source               |                        | Bmwidth: 64 deg           |
| No loads               |                        | -3dB: 328, 32 deg         |
|                        |                        | Slobe: -12.055 dBi        |
|                        |                        | Angle: 180 deg            |
|                        |                        | F/Slobe: 19.595 dB        |
```

Figure 11. Pattern for a 10-meter 3-element Yagi as plotted by ELNEC.

culations required) appears in the RUN. file. One can print the log screen-by-screen. To save or print the RUN. file as a unit, the user must exit AO and work from DOS. To save collected data, one must print the RUN. file or rename it for disk storage, because the program will overwrite it during the next modeling session. To edit the file to remove unneeded information, one must run the ASCII file through a text editor, such as the TED program supplied for antenna descriptions.

Like AO, AM collects its data in a file with the extension .LOG. This file is unique to the model and bears the same filename. It can only be overwritten if one uses the same model file in a new session. AM contains a DOS shell that permits copying and printing of the file as a unit without fully departing the program. However, for many purposes, AM log files require extensive manipulation. For example, tables of currents are useful as a check on the sensibleness of the antenna description, but lose their relevance as one makes slight adjustments in various parameters, like element length, fre-

quency, load value, etc. However, AM creates a new current table with every manipulation that requires a refiguring of the matrix, and each dutifully enters into the log. An unreviewed log gave me 17 pages of data that yielded less than 4 pages of needed data after editing within the separately run AE editor that AM uses for antenna descriptions.

If screen review of data suffices for a high percentage of your MININEC use, then the printing inconveniences of AO and AM are at most minor. However, many applications require paper data for study away from the computer screen. With the exception of files generated under frequency sweep, ELNEC offers more efficient generation of select paper than either of the other two programs.

For plot printing, AM offers two methods. Patterns may come from the on-line printing function, which appears to be quicker on a Laserjet than the screen prints of AO and ELNEC. AM also includes a separate program for employing the HPGL feature of the PCL5 HP Laserjet language to produce plotted graphics on Laserjet IIIs. Plotted outputs free the printer from dot-to-pixel correlations that produce the inexact curves shown in all the sample azimuth pattern figures. Rectangular plots require less than 1 MB of printer memory, but polar plots require at least 1 MB of added optional memory. **Figure 12** provides an example of the superb, truly camera-ready AM graphics plots. AO plots are also printable in plot form, but require a separate program from another vendor. Both AO and AM offer rectangular graphic plots as well as the more usual polar plots, and are capable of displaying and printing subfields of the total far field usually seen in antenna pattern plots. All three programs offer the potential for overlaying plots for comparison, and the resultant complex graphics are printable.

Parameter Changes: With antenna models, the user almost always wants to modify one or more variables in order to view the resultant pattern or calculated values. The programs offer different methods of performing the changes.

With AO, one may change any factor by revising the antenna description using the edit feature. This results in a permanent file revision
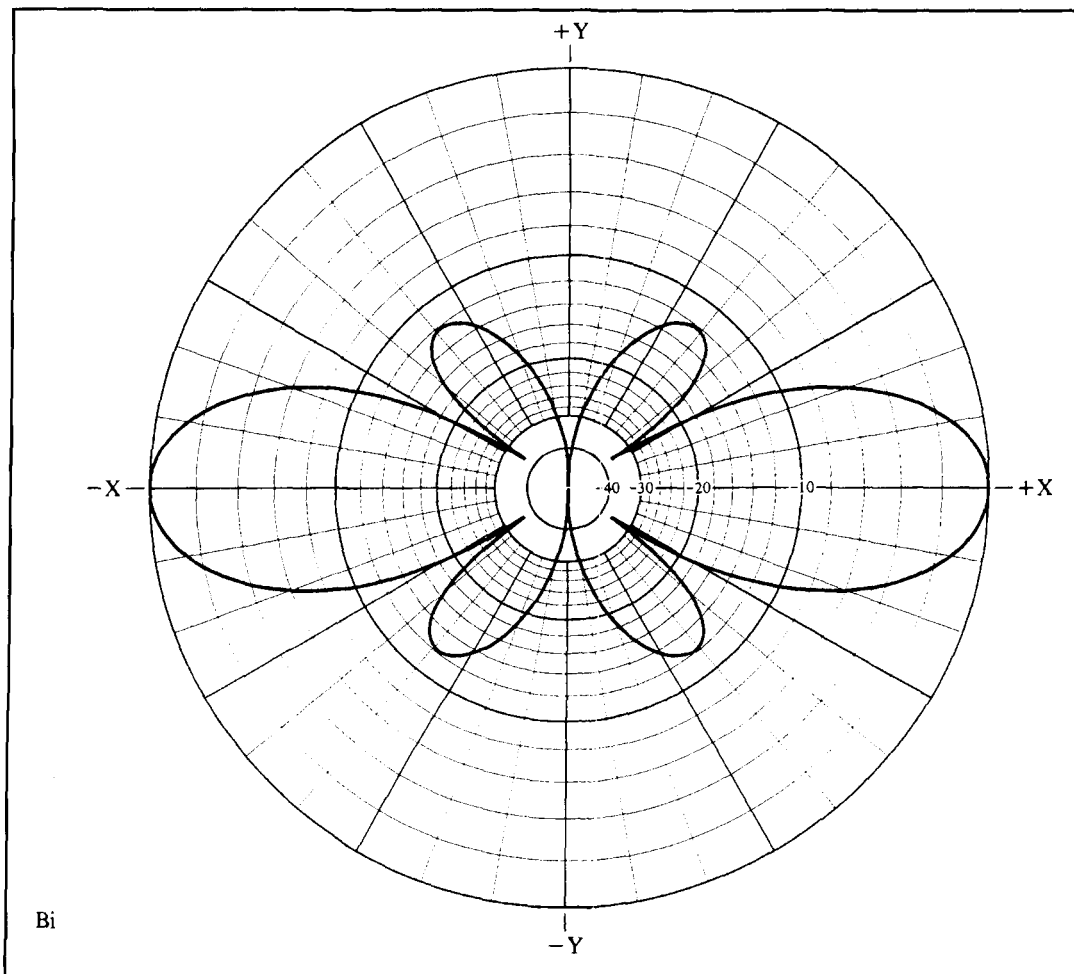
Figure 12. A sample HPGL graphics plot from AM's Antenna Draw.

and a recuing of the RUN. file. From the main menu, the user can perform a frequency sweep over almost any range with any step increment. When run with the "Gain & F/B" instruction, the program produces impedance, efficiency, gain, and front-to-back figures at the current azimuth angle—in other words, prepattern-generation data. The resultant addition to the RUN. file is a compact and easily scanned data collection. In addition, the user can change source specifications from the main menu and also enter into a versatile Options screen, shown with the main menu in **Figure 1**. Among the especially useful options to alter are the segment density to check convergence, the azimuth angle for polar plots and figures, and the toggling of the correction factors built into the program. One caution is necessary: there are 1, 2, and 4-degree options for the generation of patterns. Use the 4-degree setting only in free space. Over real ground, the program will only identify the take-off angle at one of the 4-degree points, without interpolation. Over ground, use the 1-degree selection. (In contrast, unless an elevation pattern has many lobes per pattern step unit, ELNEC will correctly identify the elevation angle, even at 5 degrees per step.)

AM offers a much larger menu of in-flight changes. Without changing the antenna description, the user can modify the sources and loads. One can also move between free space modeling and over-ground modeling by using the menu to change the environment. This can save file storage space by using one free space description per antenna and setting up ground situations, including antenna height above ground, on the go. Or, if the ground model is sufficiently fundamental, one can put it in the antenna description file and make a check run in free space as the final entry in the running log. (AM includes a soil table and calculator in its program for help in setting up both ground conditions and antenna dimensions.) The user can also change frequency, but the AO frequency-sweep provision is not present. Thus, AM creates a new current table in the log for each new frequency, inflating the running log.

Because ELNEC uses a permanent antenna description editor as its menu screen, all changes are quickly accessible. However, as noted earlier, nothing is saved without a save order, but overwritten by subsequent changes. Since select results are available on screen or in print after each run, collecting data for later use requires a different procedure. Instead of letting data accumulate in a running log, the ELNEC user looks at or prints out data on the run. With a dot matrix printer of insignificant memory, this can interrupt operations, but with a laser printer, the interruption is only momentary. For antennas with complex geometries and longer

run times, studying printed output during a subsequent calculation period allows the user to develop modification strategies within a single modeling session.

An exception to this rule is ELNEC's frequency sweep provision, new to version 3. This special subscreen permits the user to select data to be accumulated in the special run files. One can choose to save plot traces, MicroSmith-compatible files, and regular data collections. By dividing the collections into groups, including source data, load data, pattern table, currents table, and antenna analysis, ELNEC lets the user customize data collection to application needs. However, frequency sweep data files may be scanned but not printed from within ELNEC. Paper copies require a return to DOS or other ASCII-printing programs.

Repetitive Operations: Except for the new frequency-sweep feature, ELNEC is essentially a real-time program requiring user keyboard input for each desired change, whether one is moving from one model to the next or merely making changes in the present model. In contrast, AO offers a provision for batch file operation in addition to its frequency sweep feature. A DOS start-up file can specify a number of files to run in a single unmonitored session, including copying the resultant RUN. file under another name before proceeding to the next job.

More versatile are AM's batch or auto-run provisions, which can make use of menu items for each step in a repetitive process. AM's editor contains specific provisions for construction of auto-run controller files. Virtually all the function keys and the submenu selections are available. Most significant for simple but highly repetitive operations may be the "Change" options that include sources, loads, frequency, height, and environment (ground). Applications are numerous. For example, one might run an antenna through a series of height changes in small increments to check performance and source impedance between, say, a half wavelength to one and a half wavelengths in altitude.[8] The output will be a collection of data that the text editor can boil down into more tabular form. Alternatively, one might change the loading of a given antenna in small increments, seeking out the optimal load for maximum performance (however defined). Or, one might change the source position of an off-center fed antenna in small increments to find the optimal position (again, however that position might be defined). One can simulate AO's frequency sweep through the batch mode operation of AM. As with all unattended operational methods, try a small batch file first to check for smooth operation before starting an all-night run that interrupts itself on an error just as you fall asleep.

These comparisons do not cover anywhere near all the program features available in each of the commercial versions of MININEC. Details regarding the specific attributes of each system require close reading of the manuals. AM provides a 132-page printed manual with excellent illustrations. ELNEC's on-line manual (excluding added pages for the MaxP function), with an excellent starter tutorial, has grown to over 110 pages (over 216 KB on disk) of instruction and information. The 50-page AO manual (excluding additional pages for optimizer operation) also provides good information for the new modeler. If program selection requires attention to more detailed information than is presented here, borrowing manuals for close reading is the next step in the process.

Conclusions

Program efficiency is a far cry from calculation speed alone. To be efficient, a program needs to have at fingertip control all the possibilities required by particular applications. No program designer can anticipate these fully. Moreover, most enhanced versions of MININEC have been produced by amateurs out of their own curiosity and interest, and their main correspondents are likely to have been those with relatively similar interests. Thus, it is natural that the three most commercially advertised versions of MININEC examined here offer a mixture of advantage and disadvantage to the individual user.

Rather than repeat the ups and downs related to diverse tasks, let me close with an exercise every reader can reproduce from his or her own perspective. What would I include in my ideal version of MININEC (had I the ability, which I do not)? From ELNEC, I would take the menu-driven operation. Student learning curves might be steeper with this feature. Also from ELNEC, I would retain the instant and separated data print feature and the possibility for plot prints with maximum data attached. Without doubt, the versatility of element construction would be useful, as would the element tapering and the close-parallel-wire correction provisions. Incidentally, I would also retain the ability to place sources and loads as a percentage of element length, rather than having only end, center, or pulse number options.

Frequency sweep is a must in my ideal program for checking antennas across a band. ELNEC's or AO's version both work fine, although I would like to print the resultant files from within the program. From AO, I would want the calculation correction features at my command. Lest these be slighted by this compact request, they include adjacent wire taper-

---

Explanation of Parameters in Tables

1. EL = ELNEC 3.02; AM = Antenna Model 1.02; AO = Antenna Optimizer 5.0, an extension of MN. AO1 provides data on AO operation with all correction factors turned off, while AO2 gives data with all correction factors on.
2. Gain figures are in dBi to preclude confusions occasioned by showing dipoles having gain in dBd; that is, gain over a dipole. However, the "dipole" of dBd is a free-space dipole of indefinitely thin wire, whereas the models shown use finitely thick wires, whether in free space or over real ground. Real antenna improvement is always the gain in dBi of the improved model over the gain in dBi of a standard, which is often for horizontal antennas a real wire dipole over the same ground at the same height.
3. Take-off (TO) angle is the elevation angle in degrees showing peak random or maximum gain for antennas over the specified ground conditions.
4. Front-to-back (F/B) are given in dB.
5. Source impedance is given in resistance plus inductive reactance or minus capacitive reactance.
6. Bandwidth (B/W) is given in degrees between points whose gain is down 3 dB from the maximum.
7. Run-time is the total time required from the respective programs to run the MININEC matrix, factor it, and calculate all the data expressed in the tables. For AM and AO, the time does not include putting a pattern on the screen. For ELNEC, the pattern drawing time is included, because some data is most easily obtained by printing the analysis on the pattern screen. Run times are shown only where they exceed a few seconds; that is, where they exceed the time taken to read the results from the screen.

---

ing, bent-wire correction, and frequency correction. The versatility of plotted outputs is largest, but by only a bit over AM, so I shall choose it. If AO's powerful symbolic dimensioning could be made compatible with ELNEC's group element construction, it would undoubtedly make antenna modification a breeze (as well as enabling the Optimizer feature of AO, which has not been discussed here).

AM offers some powerful programming tools to strengthen my ideal MININEC program. AM's auto-run capabilities are a must (but with control programming from within the overall program and with more selectable and excludable output variables). The large spectrum of "Change" options might be included (in

case the ELNEC menu screen could not be fully preserved). Access to the calculator and to the soil definition table are convenient for empty-lap antenna construction. The use of F-keys to trigger major functions is also an excellent innovation for antenna modeling programs.

This is not everything I want. For example, variables that AM and AO require in a preoperation batch file, such as directory paths, might be placed in an accessible start-up shell. Only ELNEC currently lets you freely specify from within the program a new subdirectory that is not prewritten into a SET command, but you do have to remember where that subdirectory is located. My list could go on indefinitely. And, as noted throughout, convenient running print outputs deserve attention in one or another way by all three programs.

Each program also contains excess capabilities beyond my needs and wants. Those excesses, of course, may not be beyond your needs and wants, as measured by the applications you have in mind. On the other hand, some of my ideal wants will be irrelevant to you. Choosing a program is largely a matter of effectively listing one's needs as dictated by applications and seeing which of the available programs provides the best match. Rather than providing judgments, even about individual features, this exercise in comparative scrutiny demonstrates a method of arriving at the best individual decision. You can rightfully disagree with every evaluative particular and still get something from the exercise, even if only the clarification of your own thinking about MININEC programs.

One thing more: the surplus features of a program offer a challenge. Part of the fun of using these programs is to learn how to turn the excess capabilities into productive output. Ultimately, any tool is only as good as its user. I am now—and always will be—still learning how to use this tool called MININEC.

REFERENCE NOTES

1. ELNEC is available from Roy Lewallen, W7EL, P.O. Box 6658, Beaverton, Oregon 97007. AO and MN are available from Brian Beezley, K6STI, 507-1/2 Taylor, Vista, California 92084. Antenna Model is available from Teri Software, P.O. Box 277, Lincoln, Texas 78948. ELNEC, MN, and AM are in the $50 price range. However, a MaxP(ulse) option is available for ELNEC, and there are numerous add-on programs for MN, each at extra cost. AO is a separate program that incorporates MN and is in the $100 price range. The versions of the programs used in this study were current as of May 15, 1993. Improvements after that are not included herein.

2. For example, R.P. Haviland has had a version of MININEC available through Minilab Books, P.O. Box 21086, Daytona Beach, Florida 32019-1086. ACES (Applied Computational Electromagnetics Society) makes available BASIC and FORTRAN versions of MININEC, as well as a version of the parent program, NEC, for the PC. For information on ACES materials and membership, contact Professor Richard Adler, ACES Secretary, Naval Postgraduate School, Code EC/AB, Monterey, California 93943. One or more Macintosh versions are available, including MacMININEC 4.4 from Chris Smolinski, N3JLY, 40 South Lake Way, Reistertown, Maryland 21136. This list is not complete. Indeed, numerous engineers and hams have embellished MININEC for their own personal use, and copies of their work have been distributed. Moreover, the flux of time, economics, and effort makes the MININEC market volatile, even within the space between writing and publication.

3. The basic documentation of MININEC is contained in J.W. Rockway, J.C. Logan, D.W.S. Tam, and S.T. Li, The MININEC System: Microcomputer Analysis of Wire Antennas (Norwood, Massachusetts, Artech House, Inc., 1988). For the most recent code, see J.C. Logan and R.W. Rockway, The New MININEC (Version 3): A Mini-Numerical Electromagnetic Code, NOSC Technical Document 938, National Technical Information Service document #ADA181682, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

4. Perhaps three of the most significant recent articles in amateur journals are: Roy Lewallen, "MININEC: The Other Edge of the Sword," QST, February 1991, pages 18–22; John S. Belrose, "Modeling HF Antennas with MININEC—Guidelines and Tips from a Code User's Notebook," Antenna Compendium, Vol. 3 (Newington, Connecticut, ARRL, 1992), pages 156–164; and R.P. Haviland, "Antenna-Structure Interaction: Modeling with MININEC," Communications Quarterly, Fall 1991, pages 73–78. In general, some of the best discussions of MININEC's limitations and how to overcome them appear in the manuals for the three programs under discussion.

5. See William I. Orr, W6SAI, All About Cubical Quad Antennas (Wilton, Connecticut, Radio Publications, Inc. 1959.

6. This requirement, which may even preclude some data if not followed, is contrary to the preference I once expressed in print for putting element dimensions in the "X" column, where they are most easily inspected for errors. This possibility exists only for ELNEC, or if one accepts the loss of speed or data in the other programs. See L.B. Cebik, "A Beginner's Guide to Using Computer Antenna Modeling Programs," Antenna Compendium, Vol. 3 (Newington, Connecticut, ARRL, 1992), pages 148–155.

7. For example, the HP Laserjet options do not function with my Panasonic 4450 in Laserjet emulation mode, forcing a return to Epson emulation.

8. See L.B. Cebik, "The Effects of Antenna Height on Other Antenna Properties: A Computer Study," Communications Quarterly, Fall 1992, pages 57–79.

W2EUF wrote in response to the article by Trevor King, ZL2AKW, in the Summer 1993, issue of Communications Quarterly ("Build Your Own Direct Reading Capacitance Meter," page 103).

Dear Editor:

The article on the capacitance meter by Trevor King, ZL2AKW, in the summer 1993 issue of *Communications Quarterly* was given me by your subscriber, Al Roehm, W2OBJ, because both of us have constructed similar instruments.

Enclosed is my modification of the meter that Harry Neben, W9QB, described in the January 1983 issue of *QST*, which is mentioned by Mr. King. With a four-section, six-position range switch and 20-µA meter, there is excellent readout from 2 µF down to the pF scale. Similar results were obtained with a 50-µA meter giving a spread of 50 pF to 5 µF.

All scales are linear and 1 or 2 pF can be determined by subtracting the residual stray capacitance reading of just over 1 pF seen when the 20 pF range is employed. This has been confirmed by using several 1-pF standard capacitors in various parallel and series combinations to get 1, 2, 3, and 4 pF readings on the 20 pF scale. This stray might be less with a non-metal mounting plate, and, of course, no stray shows on other ranges.

It should be noted that only the range switch is needed after power is applied and that damping automatically is provided by the range switch on the 2-µF scale, which eliminates another control. Also, the sensitive meter is protected automatically when the meter is not being used.

The 23rd edition of *The Radio Handbook*, prepared by William I. Orr, W6SAI, and published by Howard W. Sams & Co., has a similar capacitance meter on page 26-25. This uses
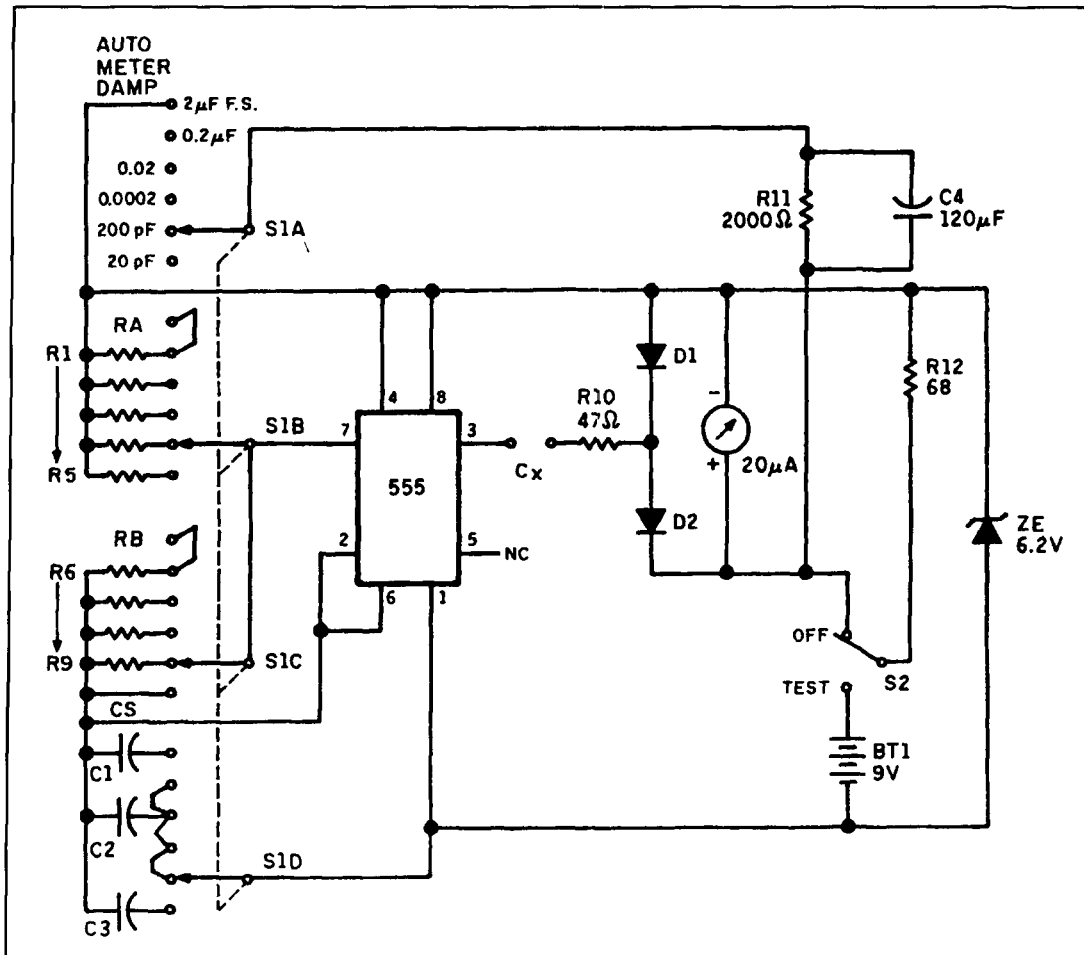
Figure 1. Capacitance meter by Harry Neben, W9QB, from *QST* of January 1983, as modified by R.E. Franklin, W2EUF.

Parts List	
R1	3.07 meg
R2	690 k
R3	69 k
R4	6.3 k
R5	4.67 k
R6	1.09 meg
R7	123 k
R8	10.1 k
R9	824 ohms
R10	47 ohms
R11	2000 ohms
R12	68 ohms
C1	30,000 pF mica = 0.03 µF
C2	10,000 pF mica = 0.01 µF
C3	1000 pF mica = 0.001 µF
C4	120 µF—24 volts DC (100 µF okay)
Ze	6.2 volts
D1,D2	1N914
S1	4 pole, 6 positions
S2	DPDT (paralleled)
BT1	9-volt transistor battery

Notes:
Subtract 1 pF on 20-pF scale.
S1 "OFF" protects the meter.
Battery current: 15 mA.
Battery volts: 9, no load; 8.6, operating
20-µA GE meter type D091 and/or Weston 301.
50-µA meter may require different values for RA,
RB, C1, C2, and C3.

a 556 chip and may be of interest, but I have not sent the circuit as it cannot be reproduced without permission from the publisher.

Richard E. Franklin, W2EUF
Mountainside, New Jersey

Mr. Franklin's modification follows.—Ed.

Simple Capacitance Meter

In the January 1983 issue of *QST*, Harry M. Nebin, W9QB, describes a capacitance meter of few parts. Results were so good that a 20 pF range was possible with a 20-µA meter, although a 50-µA meter worked fine from 5 µF down to 50 pF. No effort was made to measure larger than 2 µF. Note the damping applied to the 2-µF range. Readings are linear on all scales.

One or 2-pF capacitors were measured by subtracting 1 pF stray value on the 20-pF scale. Silver micas and high grade larger capacitors were measured on their respective ranges. After obtaining values for several capacitors, they were connected in parallel and series and the meter read very close to the calculated values of those combinations.

Construction

The circuit is shown in **Figure 1**. Stray capacitance on the 20-pF scale is decreased with short leads and lowest with parts mounted on a plastic plate in a plastic box. Silver micas are needed for C1, C2, and C3. One-watt resistors were from the junkbox and were selected by measuring with a DVM to obtain the schematic values.

C1, C2, and C3 will likely have values different from the prototype, so resistors R1 through R9 may have to be tailored to the circuit, although their values will be within reason. To calibrate, check capacitors of known values, and try resistors R1 through R9 of approximately the listed values until the meter reads the known value of those capacitors.

The test capacitors must be selected to read at the very top of each range, and then double-checked on the next higher scale. Pin 5 of the 555 chip is not needed when the range resistors are tailored to the circuit.

Component values may be determined from these equations:

$$f = I/CV$$

where:

C = pF of C1, C2, C3
V = applied voltage
I = µA
f = frequency in MHz

RA, RB, with C1, C2, and C3 determine the 555 chip frequency from the equation:

$$f = 1.44/(RA + 2RB)C$$

Note that on the 20 pF range, the frequency calculates twice that measured when operating. On other ranges, the calculated and measured frequency are close.

In the fall of 1991, we ran a reprint of a 1975 Ham Radio article by John True, W4OQ, entitled "How to Design Shunt-Feed Systems for Grounded Vertical Radiators." VE2CV sent us these comments.

I well remember John True, W4OQ's 1975 *Ham Radio* article on gamma matching grounded towers,* with particular reference to feeding radio amateur towers with quad or Yagi beams on top of the tower. I was interested therefore

*I have corresponded with John True, W4OQ, in past years. The reprint of his 1975 article indicates that he now has a new call, N4BA. But that call is not listed in the 1993 callbook, so perhaps John is a silent key? John would have enjoyed seeing the reprint of his article, and to read this correspondence.

W7EL   ELNEC 3.0Z   15M Tower + 20M Yagi   10-15-1993   09:51:39

↑→↓←   Rotate
<C>↑→↓←   Move
+ -   Zoom
<C>+ -   Zoom currents
A   Reset All
C   Ctr ant image
H   Highlight wire
1   Currents on/off
<C>1   1 phase:Now OFF
0   Phase indicator
M or F1   Menu on/off
O   Select cOlors
P   Print
R   Reset position
S   Seg dots on/off
<C>S   Uncon ends o/o
T   PaT o/solid/o
X,Y,Z   Move ant image+
<C>XYZ   Move ant image-
<A>XYZ   View from axis
F2   NoFlash on/off
<ESC>   Exit Ant View

• Origin
O Sources   □ Loads

Figure 1. Antenna model and current distribution for a gamma fed 1.9-MHz tower and Yagi. The heavy black line indicates currents.

to read a reprint of that article in the fall 1991 issue of *Communications Quarterly*, together with Hunter Harris, W1SI's comments on implementation.

True's graphical approach to design of gamma match systems was interesting, but not readily adaptable to applications where the spacing between the gamma rod and the tower are quite different from his model. W1SI commented on this, since he used a wider spacing for 160 meters. Nowadays, amateurs in radio are using antenna analysis programs like MININEC (or versions of it such as ELNEC or MN), which work on their PCs. Professional antenna engineers use NEC-2/NEC-3, but these programs run on a large computer, and are not user friendly. I was pleased, therefore, when MININEC was developed for the amateur in radio application, since even I can run the program on my PC; and PC versions of NEC-2 are becoming available.

But there are limitations and guidelines in the application of the MININEC code.[1] One of the guidelines is that an antenna, particularly a horizontal antenna, must not be too close to the ground (h > 0.2 wavelength), because the initial calculation of mutual and self-impedance of the segments making up the antenna, and hence the current distribution on the antenna, assumes

that the ground beneath the antenna is perfectly conducting. Thus, for antennas at low heights, the gain may not be right, but the far field pattern is correctly predicted, because the program employs Fresnel reflection coefficients for whatever user-selected ground conductivity is desired. However, the program works quite satisfactorily for well-grounded vertical antennas, so we can anticipate that we should be able to model quite successfully the present antenna system. But there is another limitation, in that two wires (or conductors) having different diameters should not be joined. A gamma rod and a tower have quite different diameters. The spacing between conductors, e.g. modeling a 2-conductor transmission line should not be close with respect to the diameter of the wires. For our gamma rod and tower, we are probably right in this respect, but we have to be careful.

The purpose of this correspondence is to give results obtained by numerically modeling W1SI's antenna system. For the 160-meter band, he used a gamma rod spacing of 5 feet from one edge of the tower. The length of the gamma rod was the height of the tower. For our model, we assumed that the Yagi was a 20-meter monoband wide-spaced 3-element Yagi on a 50-foot tower. The antenna model and current distribution on the antenna at 1.9 MHz are
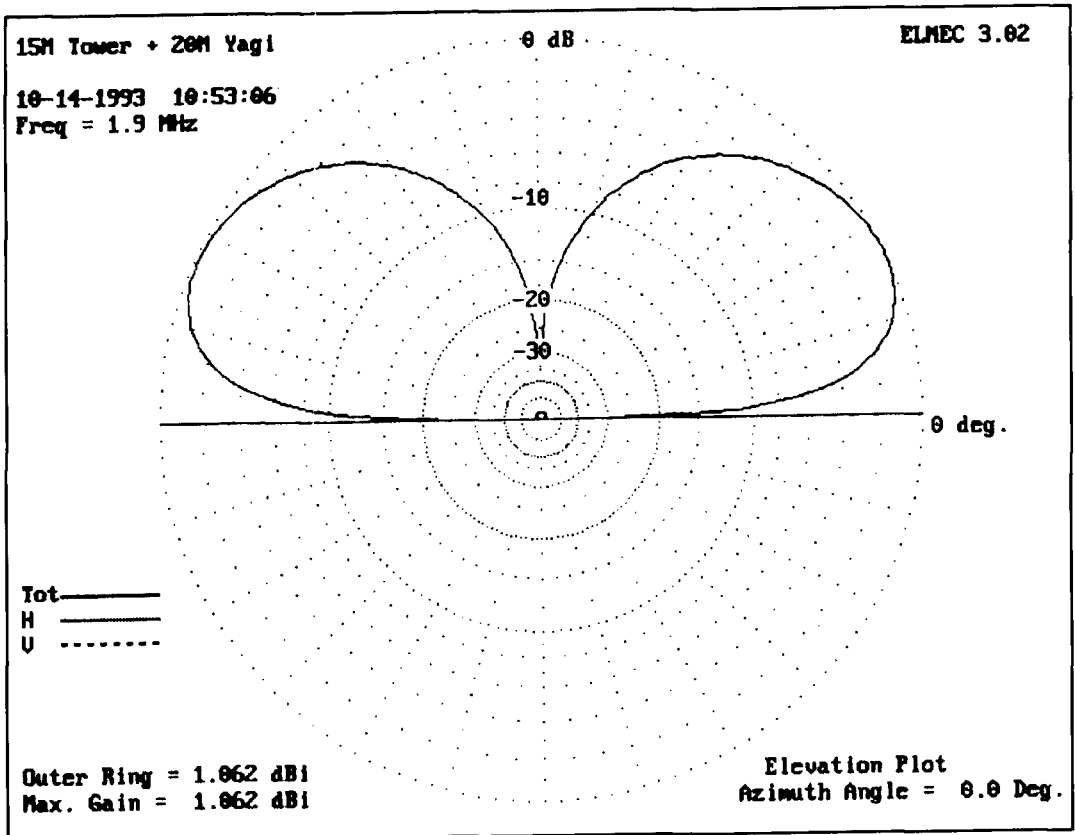
Figure 2. Elevation pattern for *Figure 1*'s antenna, frequency 1.9 MHz.
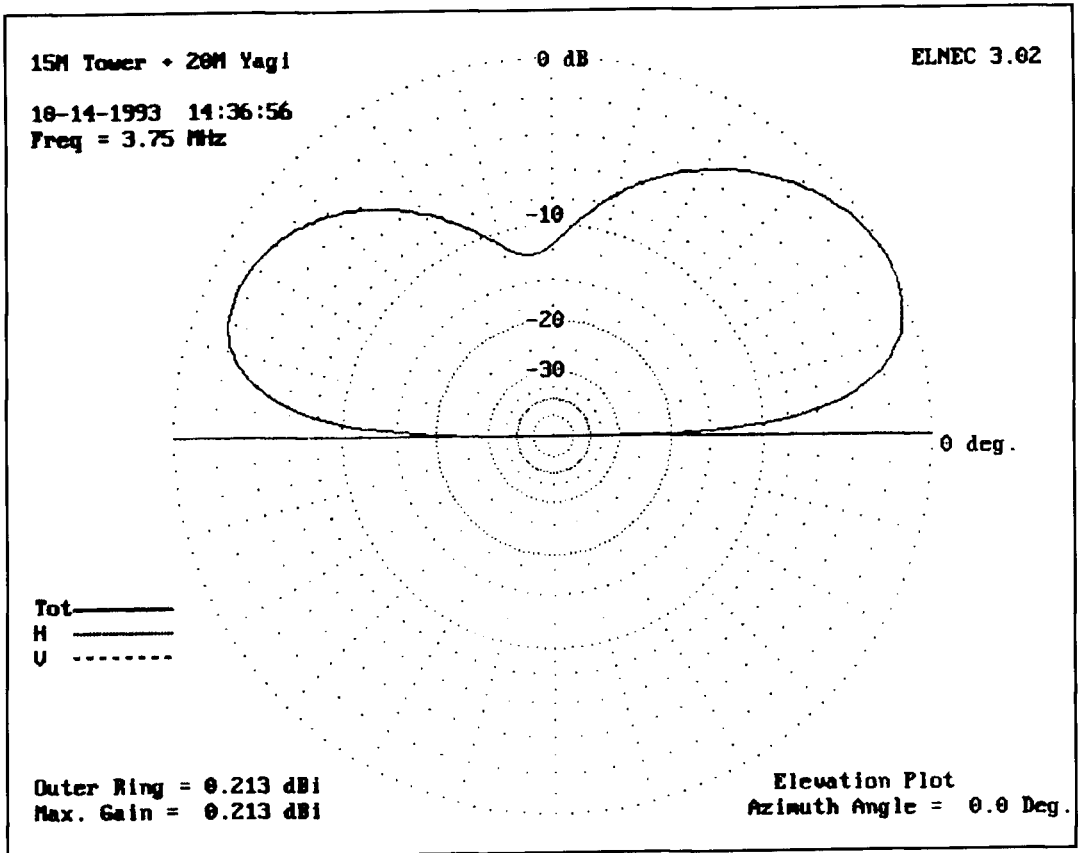
Figure 3. Elevation pattern for 3.75 MHz (see text for details on feed).
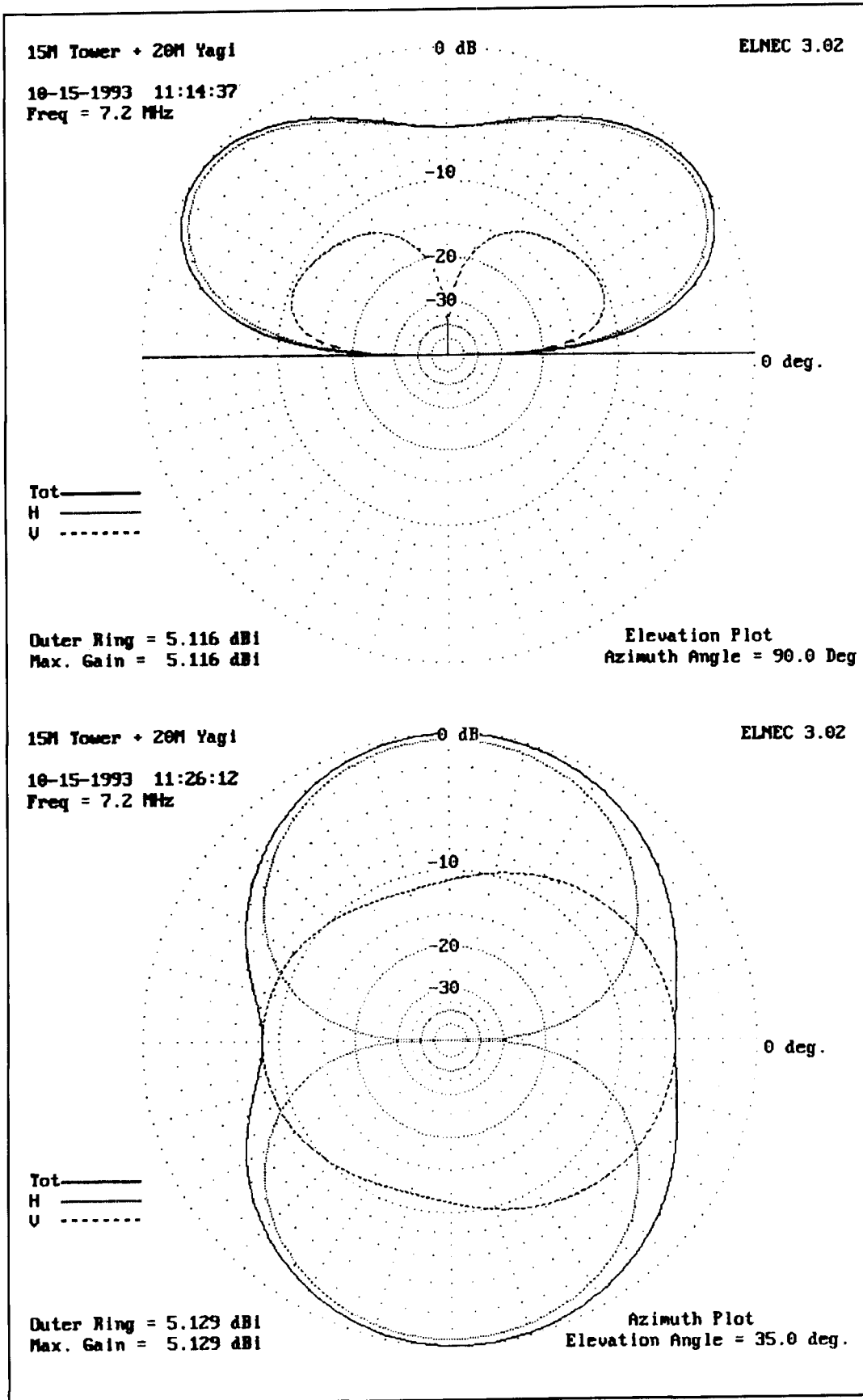
Figure 4. Radiation patterns for 7.15 MHz. A) Elevation pattern at azimuth of maximum gain; B) principal plane azimuthal pattern (at elevation angle of maximum gain).

Method of Gamma Rod Feed	Input Tower Spacing (Feet)	Impedance
Base Fed		12 − j77
Gamma Feed	2.5	26 + j405
Gamma Feed	5	67 + j648

Table 1. ELNEC's antenna system impedance 1.9 MHz.

shown in **Figure 1**, and the elevation radiation pattern in **Figure 2**, for average ground (conductivity 3 mS/m, dielectric constant 13) in front of the antenna. The antenna's impedance base fed (it is easy to lift the base of the tower off the ground for the numerical model, but not so easy at full scale), and for two gamma rod spacings are given in **Table 1**. This confirms W1SI's view that a wider spacing gives better results. As can be seen, the antenna can be tuned by a capacitor to cancel the inductive reactance at the feed point.

For the 80-meter band, W1SI used a gamma rod spacing of about a foot, and the tower was fed about 2/3 the height of the tower from the base of the tower. ELNEC's impedance at 3.75 MHz is 10 + j287 ohms. A wider spacing between the tower and the gamma rod would have been an advantage for this band also. The antenna's elevation pattern is plotted in **Figure**

3. The azimuthal pattern (not shown) is not quite circular at this frequency, but the difference is hardly more than a dB.

The antenna's gain with base feed and gamma match feed are almost the same, indicating that indeed MININEC is modeling correctly the tower and its feed.

The self-resonant frequency of the tower and the Yagi is 2.31 MHz. This is in agreement with experimental modeling I did about 15 years ago. Hence at 1.9 MHz, the electrical height of the tower and Yagi is equivalent to a 0.2 wavelength monopole (or 32.4 meters). On simple reasoning, the electrical height of the antenna system at 3.75 MHz would be 0.4 wavelength; and at 7.15 MHz, 0.77 wavelength. But the radiation pattern at 7.15 MHz is certainly not like that for a 0.77-wavelength monopole (see **Figure 4**). The antenna system exhibits directivity in the plane that contains the elements of the 20-meter beam, and surprisingly the dominant polarization in this plane is horizontal. The boom of the 20-meter beam carriers a large current (see **Figure 5**). This is a consequence of my model of the 20-meter beam. Without thinking about it, I modeled a 20-meter wide spaced beam, with reflector spacing of 0.25 wavelength, and director spacing of 0.15 wavelength. The boom connecting to the reflector carries the largest current. Other Yagi dimensions will not give such "good" pattern. By rotating my modeled 20-meter beam,


```
W7EL    ELNEC 3.02    15M Tower + 20M Yagi         10-15-1993  14:26:16

↑+↓+      Rotate
<C>↑+↓+   Move
+ -       Zoom
<C>+ -    Zoom currents
A         Reset All
C         Ctr ant image
H         Highlight wire
I         Currents on/off
<C>I      I phase:Now OFF
0         Phase indicator
M or F1   Menu on/off
O         Select cOlors
P         Print
R         Reset position
S         Seg dots on/off
<C>S      Uncon ends o/o
T         PaT o/solid/o
X,Y,Z     Move ant image+
<C>XYZ    Move ant image-
<A>XYZ    View from axis
F2        NoFlash on/off
<ESC>     Exit Ant View

          • Origin
    O Sources    □ Loads
```

Figure 5. Antenna model and current distribution for a 7.15-MHz tower and Yagi (note I have assumed a base-fed tower for this calculation). The heavy black line indicates currents. I did not bother with a gamma match for this frequency because I was concerned with pattern and currents not with feeding and matching line.

lo and behold we have a directional antenna for the 40-meter band!

In conclusion, MININEC (I used ELNEC) gives results which I believe are about right—in spite of the limitations of the code discussed above. It also, and perhaps not unexpectedly, predicts 40-meter radiation patterns that we certainly could not have guessed intuitively.

John S. (Jack) Belrose, VE2CV
Quebec, Canada

REFERENCE

1. J.S. Belrose, "Modeling HF Antennas Using MININEC: Guidelines and Tips From a Code Users Notebook," *ARRL Antenna Compendium*, Vol. 3, 1993, pages 156–164.

AF6S offers this alternative to the K1CCL propagation velocity method described by George Wilson, W1OLP, in "Tech Notes" (Communications Quarterly, spring 1993, page 83).

Dear Editor:

The K1CCL propagation velocity method that George Wilson, W1OLP, describes in "Tech Notes" (spring 1993, page 83) is clever, but tedious.

The use of separate means to measure dip-frequency accurately brought to mind my MFJ HF/VHF SWR Analyzer, Model MFJ-249, which combines a signal generator, an accurate frequency counter, and an SWR bridge—covering 1.8 to 170 MHz in a hand-held package.

A transmission line that's open at the far end appears as an open at the near end when it length is any *even* number of quarter wavelengths (i.e., *any* number of half wavelengths). A transmission line that is shorted at the far end appears as an open at the near end when the length is any *odd* number of quarter wavelengths. A clear, concise review of transmission theory appears in Caron[1] Chapter 2.

If you connect a small carbon-composition or metal-film 50-ohm resistor directly across the MFJ-249, the SWR meter reads 1:1— except well above 100 MHz, where stray inductance and capacitance enter the picture.

If you connect the end of a transmission line in parallel with the resistor, the total impedance becomes that of the resistor in parallel with the line. At frequencies where the line presents an open, the meter "sees" only the resistor, so SWR reads 1:1.

I soldered the pins of a pair of "four-hole" SO-239s together, and soldered four wires between the opposing holes and a 51-ohm, 1/4-watt resistor between the pins and one of the wires. I connected this "termination" to the MFJ-249 using a UHF double male. As expected, SWR read hardly over 1.00:1 below 150 MHz.

Next, I attached a 24-foot length of RG-8/U (Belden 9267) and put a UHF "bullet" in the PL-259 at the line's far end. **Figure 1A** shows the setup.
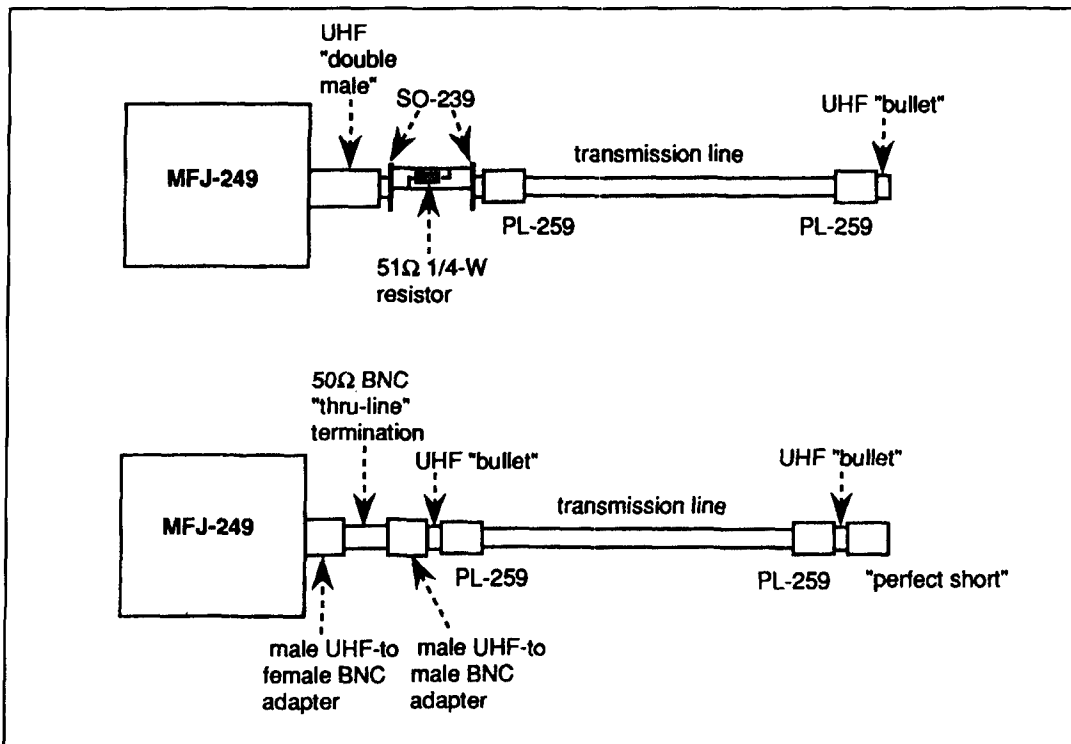
Figure 1. (A) Resistor in parallel with transmission line. (B) With thru-line termination and "perfect short" at far end of line.

	A	B	C	D	E
1	Constants:				
2	Length (meas):	24.1042	f t		
3	Len (physical)=	7.3470	meters		
4	Len (elec.)	F(open)	Vp(open)	F(short)	Vp(short)
5	wavelengths	MHz		MHz	
6	0.25	6.5319	63.99%	6.7226	65.85%
7	0.50	13.3937	65.60%	13.4210	65.74%
8	0.75	20.0468	65.46%	20.2400	66.09%
9	1.00	26.9262	65.94%	26.8830	65.84%
10	1.25	33.6678	65.96%	33.8800	66.38%
11	1.50	40.3875	65.94%	40.3940	65.95%
12	1.75	47.5613	66.56%	47.5095	66.49%
13	2.00	53.9637	66.08%	53.8858	65.98%
14	2.25	60.9219	66.31%	60.9350	66.32%
15	2.50	67.3690	65.99%	67.3350	65.96%
16	2.75	74.3725	66.23%	74.4860	66.33%
17	3.00	80.8952	66.04%	81.0450	66.16%
18	3.25	87.8790	66.22%	88.1200	66.40%
19	3.50	94.5370	66.15%	94.4710	66.10%
20	3.75	101.6320	66.37%	101.5960	66.35%
21	4.00	107.8540	66.03%	107.9500	66.09%
22	4.25	115.3420	66.46%	115.1600	66.36%
23	4.50	121.5980	66.18%	121.4660	66.10%
24	4.75	128.6500	66.33%	128.7590	66.39%
25	5.00	135.1450	66.19%	135.1680	66.21%
26	5.25	142.3170	66.39%	142.4000	66.43%
27	5.50	148.8360	66.27%	148.7200	66.22%
28	5.75	155.9190	66.41%	156.0000	66.44%
29	6.00	162.3000	66.25%	162.1500	66.18%
30		Average=	66.06%	Average=	66.18%
31		Median=	66.18%	Median=	66.19%

Table 1. Measured frequencies and resulting propagation velocities using the commercial termination.

Sure enough, the MFJ-249 showed 1:1 SWR at a series of frequencies in a 1:2:3:4:etc. frequency ratio—where the coax is 1, 2, 3, 4, etc. half-wavelengths long.

I noticed I could also read SWR peaks accurately. The peaks occur at frequencies where the transmission line looks like a short—any odd number of quarter wavelengths from an open. I recorded the frequency of every SWR null and every peak, and measured the coax length to the end of the "bullet."

Then I made a "perfect short"—a PL-259 with a copper-foil disk with a hole, soldered to the pin and to the surrounding metal body—and put it on the "bullet" at the far end.

I recorded the frequency of every SWR peak and null in the shorted cable. But when I entered all the open and shorted frequency values into an Excel spreadsheet, the variations in calculated propagation velocity seemed excessive. I substituted a commercial "thru-line" termination, with appropriate adaptors, as in **Figure 1B**, and the results improved a little. **Table 1** shows the measured frequencies and

resulting propagation velocities using the commercial termination.

Column A contains electrical length in wavelengths and column B has the corresponding frequency. The following equation for propagation velocity, Vp, in column C works for any electrical line length:

$$V_p = \frac{Len_{meters}(physical) \times F_{open(MHz)}}{300 \times Len_\lambda (electrical)}$$

Note: this equation gives the result as fraction. Excel can display fractions as percentages, so you don't need to "adjust" your equations for that purpose. But if you prefer an answer in percent, just substitute 3 for the 300 in the denominator.

Column D contains the frequency series measured with the coax shorted, and column E calculates the propagation velocity, using the same equation, but taking its frequency from column D.

The frequencies in both the F(open) and

Belden 9267, length = 24' 1.25"
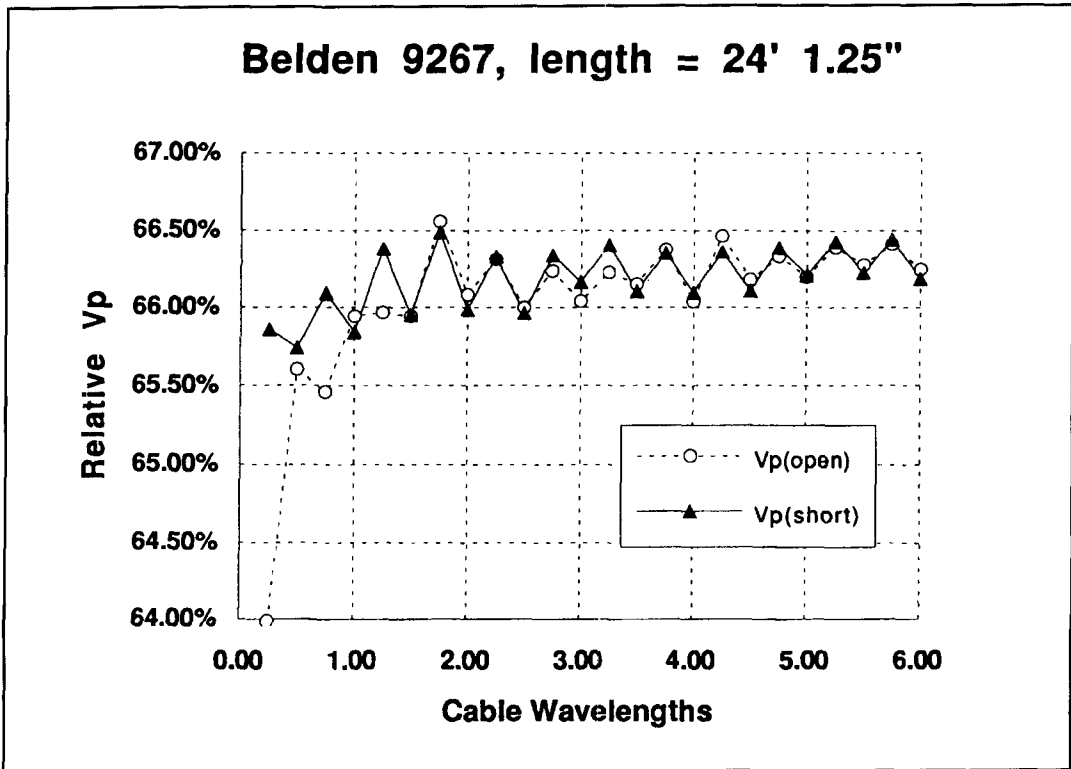
Figure 2. Plot of the results.

F(shorted) columns correspond to alternating SWR readings near 1:1 and infinity. In the 0.25-wavelength measurement, for instance, the F(open) SWR peaks at infinity, while the F(short) SWR nulls at 1:1. In the 0.5-wavelength case, the F(open) SWR nulls and the F(short) reading peaks.

Figure 2 plots the results. Except for the open quarter-wave measurement, all fall within a 1-percent spread. If you ignore readings for lines less than 1.25 wavelengths long, the spread diminishes to 0.6 percent. The values tend to alternate, suggesting some sort of experimental error. Curiously, the higher values are null readings for the shorted line, but peak SWR readings for the open line.

If you can tolerate a half-percent error, this suggests measuring a cable one to six wavelengths long (at its end-use frequency) and considering it accurate enough. But if you are concerned about the variations, as I was, you might want to consider possible causes of error. Here are the ones that occurred to me:

1. The effective position of a simple open, like a "bullet" or a cut-off coax, varies with frequency because of electric field fringing. In their heyday, General Radio supplied calibrated opens (p/n 874-WO3) for their coaxial "trombone lines" used with their Model 1602-B UHF Admittance Meter. The outer conductor extended several of its diameters beyond the inner

conductor, producing an open that did not "move" as frequency changed.

2. An impedance discontinuity at the near end of the cable can also cause errors.

3. Non-uniformity along the length of the coax, dielectric effects, and small "skin effect" caused losses might cause measured propagation velocity to vary slightly with frequency.

What can we do to minimize these effects?

1. It's easier to make a controlled short than an accurate open, so use a "perfect short."

2. To minimize the near-end discontinuity, use a small commercial "thru-line" termination such as Tektronix p/n 011-0049-01 (50 ohms, BNC, $35)—instead of a resistor.

3. Keep your measurement frequency within a factor of two of the frequency of intended use and, as W1OLP suggested, take your measurements on the piece you intend to use. Cut a piece a bit longer than your application calls for, make the measurement, calculate the exact length needed from the resulting propagation velocity, then trim to that length.

Measure physical length from the middle of the termination to the plane of the shore at the cable's far end. (If you must measure an open-ended cable, add half the cable shield diameter

to the measured physical length and hope for the best.) If the cable is much, much longer than the total of all the adaptors, termination, etc., you need not worry about the propagation velocity in these devices; but for very short cables, the reverse is true.

This method works even if the cable impedance is not 50 ohms, but to measure a balanced transmission line, you need a balun. The Guanella type[2] is best, because it is made of transmission lines in such a way that its electrical length is constant (ideally) and measurable. Be sure to check a balun's SWR, when it is terminated in its correct output resistance. Balun SWR should be below 1.2 or so in the frequency range at which you intend to measure the balanced line. With the balun in place, proceed as before—but include the electrical length of the balun when calculating propagation velocity.

Another handy use for the MFJ-249 test setup is measuring the length of coax on a spool. An open-far-end measurement should be accurate enough—a good thing, if the far end isn't accessible. Of course, you will measure the electrical length, so you'll have to use the published propagation velocity to calculate the physical length.

On extremely long rolls, the quarter and half-wave frequencies may fall below the lowest frequency of the MFJ-249. No worries! Notice that the frequencies in **Table 1** increase in increments of the quarter-wave frequency. That means you can subtract any adjacent pair of measured frequencies, one peak SWR reading and the next null, for example, to find the quarter-wavelength frequency. Another bonus: with any long piece of cable, you can always make your velocity measurement near the desired frequency. It's handy and it eliminates one source of error.

You could just as well make MFJ-249 measurements with a terminating resistor in series rather than parallel, duplicating K1CCL's results without the coils. The difficulty lies in the impedance discontinuity of the series resistor. If you make a good series terminator using strip-line techniques and a chip resistor on double-sided pc board, it should work just as well as the shunt method. SWR nulls would occur at frequencies where the cable appears as a short instead of an open, so the nulls and peaks would reverse.

I will venture to say that the MFJ-249 is the most useful ham-priced RF instrument to hit the market since the grid-dip meter. MFJ offers it as an antenna-measuring tool, and it certainly handy for that. But the unique combination of generator, counter, and bridge, makes it useful for a far wider range of experimental work than just antennas.

David M. Barton, AF6S
San Jose, California

REFERENCES

1. Wilfred N. Caron, *Antenna Impedance Matching*, Amateur Radio Relay League, 1989.
2. Jerry Sevick, W2FMI, *Transmission Line Transformers*, 2nd Edition, Amateur Radio Relay League, 1990.

PRODUCT INFORMATION

Sipel Multi-Blade Knife From The Tool Resource

The Tool Resource has introduced the Sipel Multi-Blade Knife. Its design allows you to use one handle for many different types of hobby blades. The M34-SA knife is constructed of stainless steel and the blades can be removed or reattached with a three-step process. A scalpel is also included.

The Sipel Multi-Blade Knife (Part number M34-SA) is available for $19. Scalpel Blades (Part number M9-15C) are available for $.90 ea.

For more information contact The Tool Resource, P.O. Box 1106, W. Dundee, IL 60118. Phone/Fax: 708-468-0849.

Application Notes From Analog Devices

Analog Devices has published four new application notes.

The AN-203, *Using the AD2S80A-Series Resolver-to-Digital Converters as a Control Transformer*, for designers of position servo systems.

The AN-369, *Thermocouple Signal Conditioning Using the AD594/AD595*, practical considerations when using these one-chip thermocouple signal conditioners in temperature measurement applications.

The AN-365, -366, and -367, describe in detail evaluation boards for the AD7711, AD7712, and AD7713 24-bit sigma-delta A/D converters (respectively), to provide a complete, high-resolution channel for specific analog signal measurement applications.

And the AN-372, *MDAC Circuit Gives Freedom of Choice for Both Output Span and DC Offset Voltage*, describes use of a multiplying DAC in a configuration which allows independent setting of these two critical parameters.

For more information contact Bill Schweber. Phone 617-461-3821; Fax 617-329-1241; e-mail bill.schweber@analog.com.

Bryan Bergeron, *NU1N*
Decision Systems Group
75 Francis Street
Boston Massachusetts 02115
bergeron@hstbme.mit.edu

SURFACE ACOUSTIC WAVE TECHNOLOGY

The fundamentals of SAW technology, how SAW devices operate, and where this technology is headed

There's something pleasingly mystical about working at UHF and above. Materials normally thought of as insulators at HF are routinely used as capacitors and resistors, and component leads become inductors, tuned circuits, or antennas. Hollow metallic tubes serve as RF conduits between transmitters and antennas, the latter often consisting of nothing more than an etched PC board. Yet another apparent paradox is a class of devices that rely on *acoustic* waves for their operation.

Surface acoustic wave (SAW) devices are so named because, at the most basic level, their operation involves the excitation and detection of minute *acoustic* waves that travel over the *surface* of a substrate in a controlled manner. SAW devices command the attention of communications professionals and amateurs
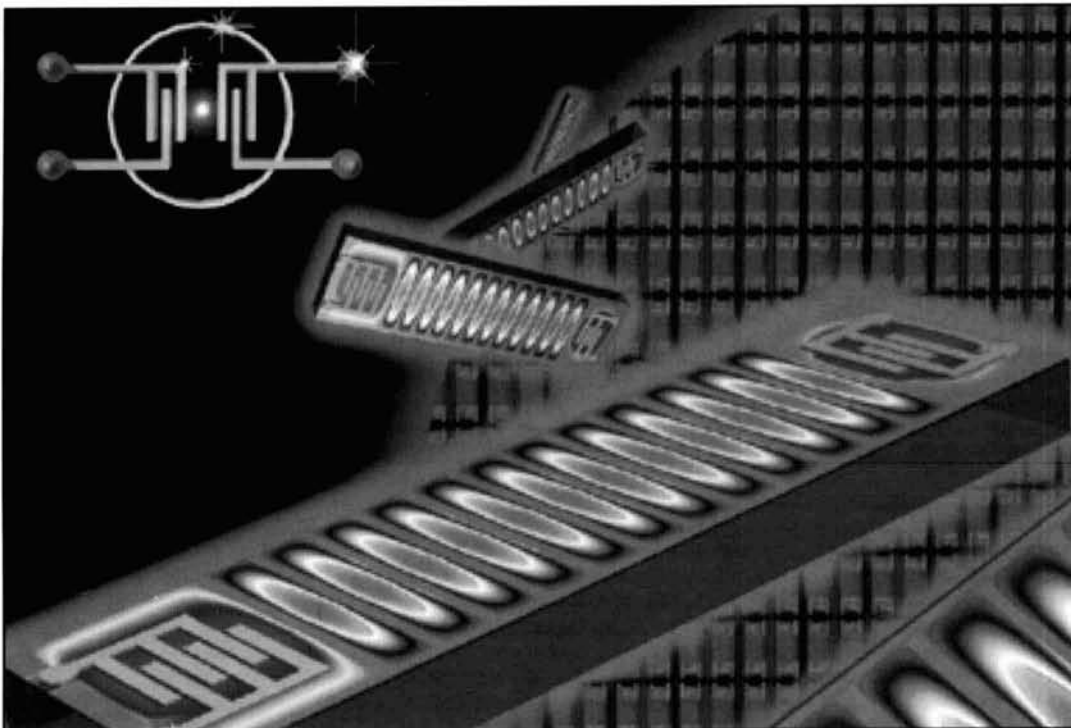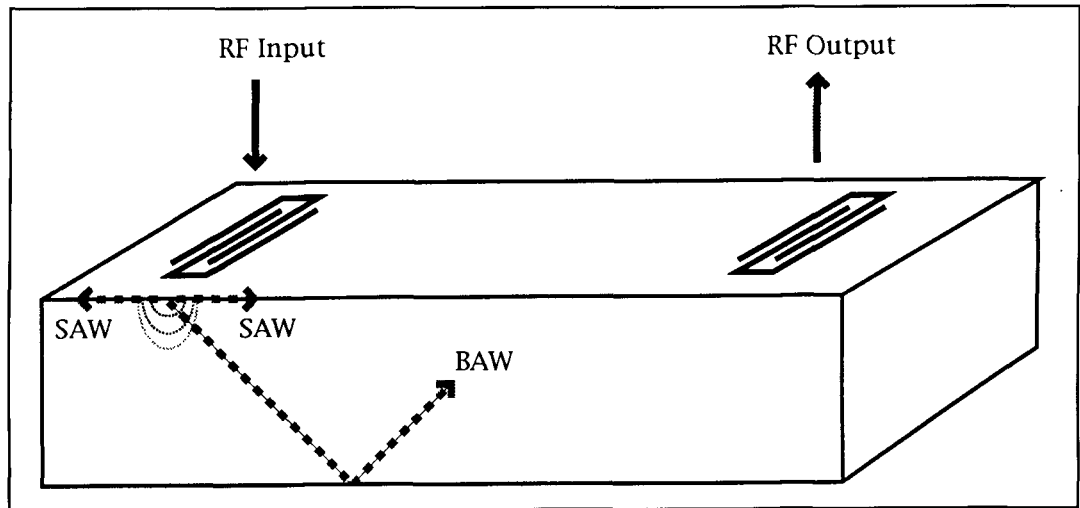
Figure 1. A simplified view of a generic SAW device showing the basic components necessary for operation—a set of RF-acoustic transducers and a substrate capable of conducting acoustic waves. The interdigital transducers produce and respond to both surface and bulk acoustic waves (SAW and BAW, respectively).

Material	Acoustic Impedance
Air	411×10^{-6}
Aluminum	17
Brass	36
Cadmium Sulfide	22
Copper	42
Glass	13
Ice	4
Lead	22
Lithium Niobate	34
Magnesium	10
Mercury	20
Mylar	3
Nickel	54
Nylon	3
Platinum	89
Polyethylene	2
Porcelain	13
Quartz	15
Rubber (Silicon)	1
Silicon	20
Stainless Steel	45
Titanium	27
Tungsten	105
Zinc	30
Zinc Oxide	36

Table 1. Acoustic impedance of both common materials and materials used for SAW device substrates, listed in 10^6 kg/m^{2}s.

because these devices provide a unique mix of excellent performance and low cost in an extremely small package. SAW components are found in a variety of applications where low weight, small form factor, ruggedness, and performance are essential—from miniature cellu-

lar phones and secure spread spectrum communications systems to color TV IF filters.

SAW devices have become increasingly important in UHF and microwave communications, where they form the basis of oscillators, amplifiers, bandpass filters, discriminators, and matched filters. SAW devices also have applications in spectrum surveillance systems, RADAR pulse compression and expansion, the generation of long pseudo-random code sequences, and in storing digital information. At the same time, SAW devices are relatively simple and elegant in both their operation and their design.

Fundamental concepts

As electronic components go, those based on acoustic waves are relatively straightforward. A prerequisite to understanding SAW devices, and acoustic wave technology in general, is a grasp of the basics of acoustics, together with a familiarity with piezoelectric transducer operation. To provide a working context for the following discussion, it may be helpful to refer to **Figure 1**. As illustrated in the figure, features common to virtually all SAW devices include metallic electrodes that have been deposited in specific configurations on a suitable piezoelectric substrate. The transducer electrodes generate and receive acoustic waves on the substrate surface in a manner defined by the configuration and relative orientation of the transducers.

Acoustics

Acoustics deals with the propagation of pressure fluctuations in fluid and, of particular importance to this discussion, solid media. The
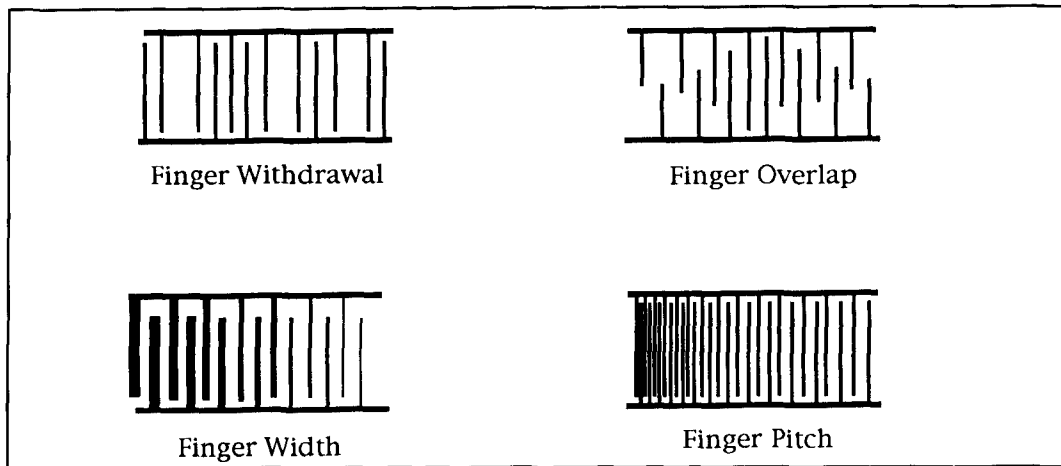
Figure 2. A sample of the more common weighting techniques used to shape SAW transducer amplitude and/or phase response. Weighting techniques allow signal amplitude to be maintained while shaping signal phase. Alternatively, phase can be held constant while shaping signal amplitude. These techniques can be combined in a single device to achieve a variety of transducer response characteristics.

properties of an acoustic wave are typically described in terms of particle displacement, velocity, pressure, and other parameters that have parallels with electronics units and concepts. For example, materials can be classified as either acoustic insulators or conductors, or something in-between. As shown in **Table 1**, materials can also be ranked in terms of acoustic impedance, a term that roughly parallels the concept of electrical impedance.

Acoustic waves propagate with variable loss and velocity in solids, as a function of the media. Furthermore, these properties can be exploited in the creation of a variety of acoustic devices. For example, the relative low velocity of acoustic waves, on the order of 5000 meters/second,[1] makes it possible to construct delay lines that provide significant delays in a minimal amount of physical space.

As illustrated in **Figure 1**, acoustic waves are generally considered as either *bulk waves* or *surface waves*. As its name implies, a *surface wave* follows a path bounded by the surface; greatest particle motion occurs within about one acoustic wavelength of the surface of the substrate.[2] *Bulk waves* travel, either longitudinally or transversely, through many wavelengths in all directions. As a result, the substrate undergoes motion and stress throughout its volume.

Although acoustic wave devices can be constructed to exploit either bulk or surface wave propagation, surface wave devices are more popular because their construction provides relatively easy access to signal paths for signal sampling. For devices based on surface wave propagation, bulk waves are viewed as undesirable. Not only do bulk waves represent energy loss, but coupling between surface and bulk waves modifies the surface wave radiation. In addition,

bulk waves carry energy to the output transducer, interfering with its response to surface waves. The influence of bulk waves on the output transducer depends on the thickness of the substrate, the parallelism of the bottom and top surfaces, the acoustic impedance match between the substrate and the supporting base, and the distance between transducers (see **Figure 1**).

SAW transducers

Transducers translate energy from one form to another. Like other *electroacoustic* transducers—such as loudspeakers, headphones, and microphones—SAW transducers convert electrical to acoustic energy and vice versa. However, instead of being driven by audio frequency signals, SAW devices operate with RF up to several GHz.

Substrates

The piezoelectric nature of the substrate supports the conversion of electromagnetic into acoustic energy at the input transducer, and acoustic energy back into electromagnetic energy at the output transducer. The coupling of electromagnetic to acoustic energy is commonly achieved through the use of interdigital transducers—metallic electrodes deposited on a piezoelectric substrate. By applying an electrical potential to opposing digits of the transducer, alternating stress patterns are set up in the substrate via the *piezoelectric effect*. As illustrated in **Figure 1**, acoustic waves subsequently created propagate normally to the electrodes on the surface (surface acoustic waves) and at various angles through the substrate (bulk acoustic waves). Acoustic waves that impinge upon the
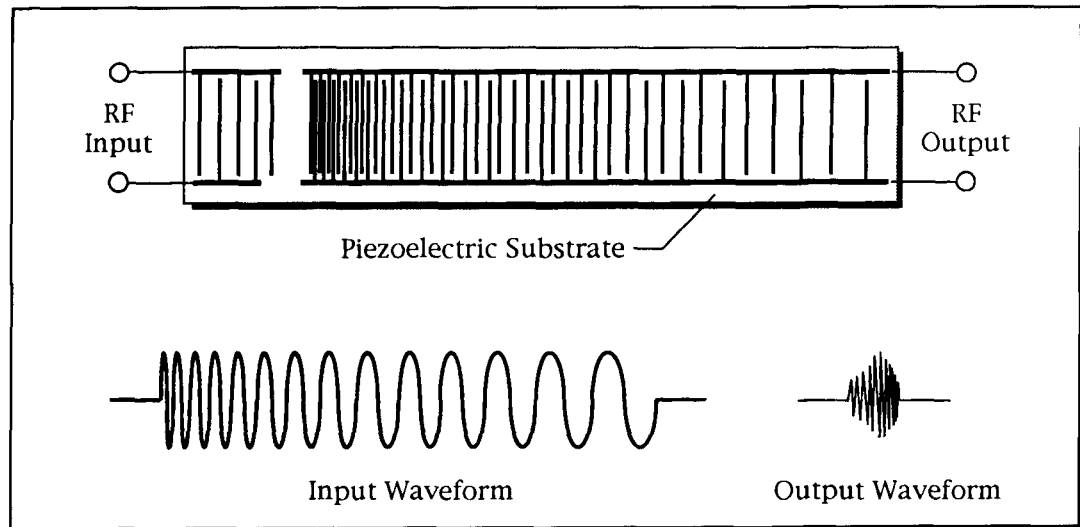
Figure 3. A SAW delay line constructed around a dispersive transducer. In this design, the input transducer (left) launches a signal toward the dispersive output transducer (right). Pulse compression is accomplished by varying the transducer finger spacing in a defined manner. This design can provide moderate compression; i.e., time-bandwidth products on the order of 500. Not shown in this figure are the thin strips of acoustic absorbing materials deposited along the left and right edges of the substrate that minimize acoustic reflections from the substrate edges.
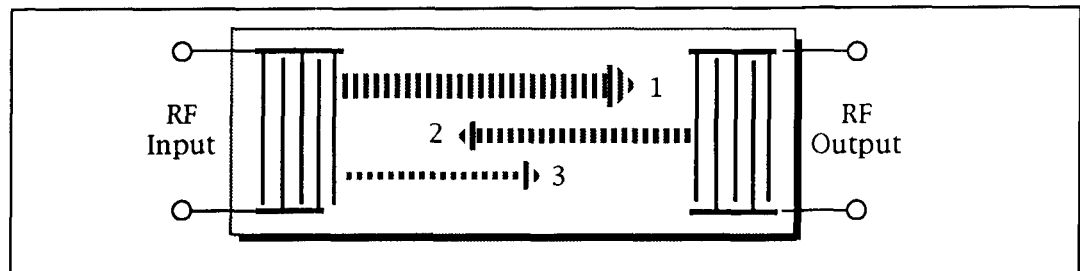
Figure 4. Triple transit echoes arise when the receive transducer reflects energy back into the transducer, thereby causing the generation of a new acoustic wave. This regenerated wave travels to the transmit transducer, where the acoustic signal is similarly regenerated, and returned to the receive transducer. Although this process continues indefinitely, the loss associated with each trip across the substrate renders the higher-order regenerative signals relatively inconsequential. Triple transit echoes limit the useful range of SAW filters.

SAW receiver electrodes establish an electric potential and RF currents in the electrodes, which can be connected to external electronics for amplification and processing.

Piezoelectric substrates are used in the fabrication of SAW devices for the sole purpose of supporting the operation of the input and output transducers. The composition of the substrate places limits on the frequency of operation, bandwidth, and time delay of the SAW device. In selecting a substrate, relevant issues include tradeoffs between high electromechanical coupling versus temperature stability, the availability of large substrates, the compatibility of fabrication techniques, spurious signal rejection requirements, the maximum allowable acoustic loss, and cost. For example, ceramic substrates are associated with severe propagation losses. Quartz, in comparison, has low propagation losses but also a low coupling coefficient. Cadmium sulfide and zinc oxide, which are

both semiconductors, have low loss and show appreciable coupling. Lithium niobate, a synthetic crystal, supports a very high degree of piezoelectric coupling.

The acoustic attenuation associated with SAW devices is a function of the substrate and the conversion loss of the input and output transducers. Acoustic losses in the substrate material can be attributed to relaxation absorption, thermoelastic loss, and scattering.[3] *Relaxation absorption,* which can be considered a form of acoustic inertia, occurs when an acoustic wave propagates through a material, but neighboring particles do not move at the same speed. Scattering losses can be attributed to imperfections in the substrate. Since scattering losses are increasingly important at higher frequencies, crystalline substances, which are relatively pure and free of imperfections, are used as substrates at microwave frequencies. In contrast, most piezoelectric ceramics are useful
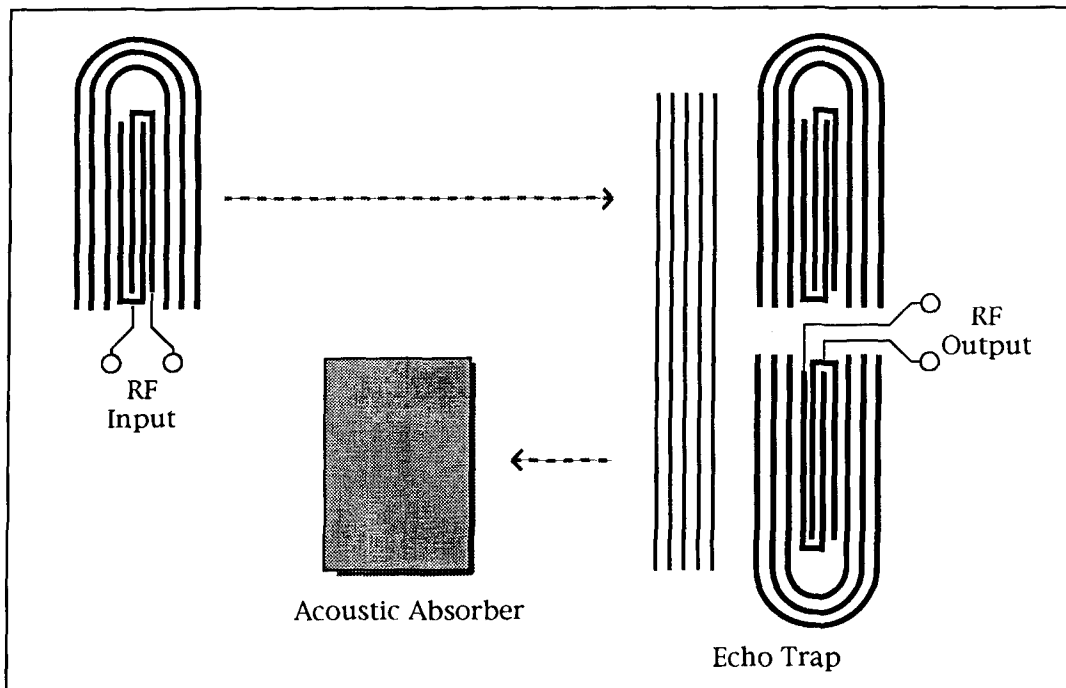
Figure 5. One of the major uses for multistrip structures is the rejection of spurious signals, such as triple transit echoes. In the above example, a delay line is constructed with an echo trap to redirect the signal normally reflected back to the input transducer into an acoustic absorber such as rubber cement.

to only about 20 MHz because of scattering losses. Thermoelastic losses can beconsidered the result of the compression and decompression of substrate material caused by the acoustic wave, with the result that acoustic energy is transformed into heat. Thermoelastic attenuation is proportional to the square of the acoustic wave frequency.

Transducers

Surface wave transducers typically consist of two sets of interleaved metallic electrodes or fingers deposited on a piezoelectric substrate, such as lithium niobate. The fingers of these interdigital transducers are separated by one half of the acoustic wavelength at the operating frequency. Typical finger spacing for a transducer designed to be driven by a 100-MHz RF signal (or to generate a 100-MHz signal when acting as a receiver) is less than 10 μm.

Whereas the lowest operating frequency supported by a SAW device is limited by the maximum allowable size of the substrate, the upper frequency limits are set by resolution in the photolithographic processes used to form the transducer fingers. Transducers up to about 600 MHz can be constructed with conventional photolithography techniques. Transducers capable of higher frequency operation require scanning electron microscopes to expose the photoresist. With modern fabrication tech-

niques, SAW operating frequency ranges from a few MHz to over 3.5 GHz.

It's often useful to think of acoustic transducers as antennas because the coupling of the transducer to the piezoelectric substrate requires consideration of directivity, acoustic impedance, and radiation resistance. For example, both finger geometry and spacing are dependent on intended operating frequency. Interdigital surface wave transducers have a fractional bandwidth defined by 1/N, where N is the number of finger pairs. Electrically, surface wave transducers can be thought of as a capacitance shunted by a radiation resistance that depends on finger length as well as the composition of the underlying substrate.

Transducer geometry is often modified to shape the amplitude and/or phase response of the SAW device. These geometric modifications or *weighting* schemes include alterations in finger width, pitch, withdrawal, and overlap (see **Figures 2** and **3**). Finger *width weighting* involves varying the width of successive fingers in an array, while maintaining a constant separation between finger centers. Finger *pitch weighting* varies the separation between finger centers, whereas finger *width weighting* involves modifying the overlap between opposing fingers. Finger *withdrawal weighting* uses selective withdrawal or absence of electrodes, so the average electrode density approximates the desired weighting.

Used alone, each weighting scheme makes it
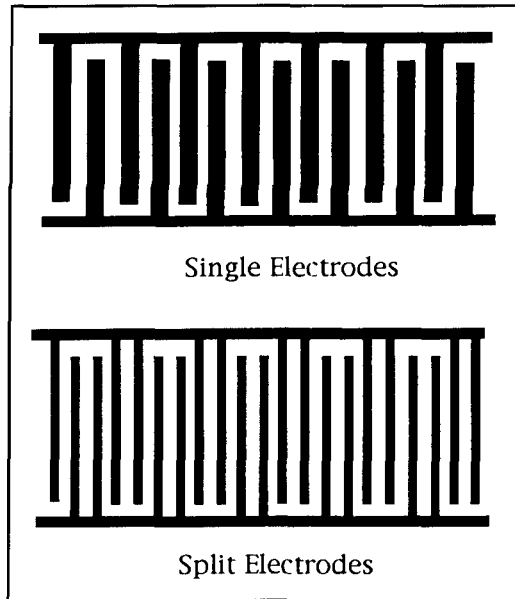
Figure 6. Intra-transducer reflections can be reduced by substituting double or split electrodes (bottom) for single electrodes (top). Transducers with single electrodes have one electrode per half-wavelength, whereas transducers with double electrodes have two electrodes per half-wavelength.

possible to independently define the amplitude and phase responses of a SAW device. Alternatively, these and other weighting techniques can be combined to achieve specific transducer amplitude and phase response characteristics. Variations in electrode width can also be used

to affect amplitude and frequency response.

Transducers have their limitations. Transducer performance is degraded by tight coupling effects, unwanted acoustic waves, electrode resistance, electromagnetic crosstalk, acoustic reflections from terminations and substrate edges, diffraction of the acoustic wave, and electrode interaction. All these factors contribute to distortion, and therefore limit the useful range of SAW devices. *Tight coupling effects* arise because transducer input and output impedances are not fixed, but are a function of frequency, so the generator and load cannot be maintained at the same impedance.

Unwanted acoustic waves include *regenerated waves* (triple transit echoes), *internal reflections* (intra-transducer reflections), and *bulk waves*. Triple transit echoes are due to signals that arrive at the receiving transducer, are reflected back to the input transducer, and then back to the receiving transducer (see **Figure 4**). The triple transit echo or signal results in part because as an acoustic wave travels under a transducer, it induces a voltage across the bus and external load, which in turn generates another acoustic wave within the transducer. The resulting signal has three times the delay of the original signal. In addition, if the original signal path has a loss of X dB associated with it, the triple transit signal will experience a further loss of 2X + 6 dB.[4] Triple transit echoes can be reduced by making one of the transducers unidirectional and tuning it so no energy is available for reflection, or by building a reflect-
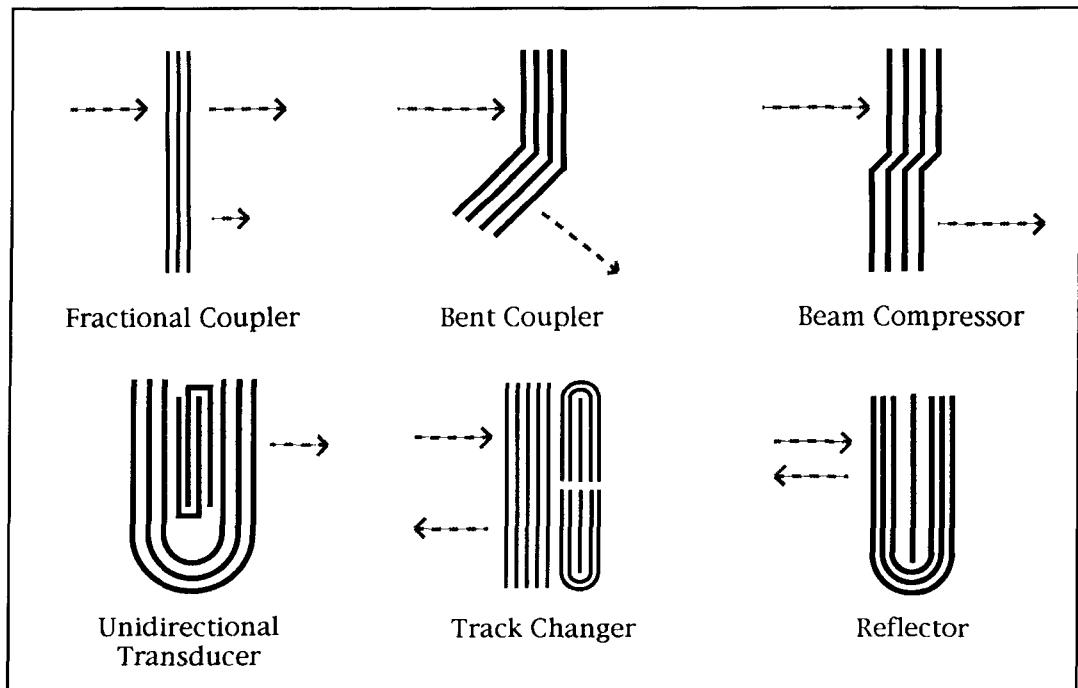
Figure 7. Common multistrip coupling configurations. Multistrip geometry defines how acoustic waves are moved or steered from one region of a piezoelectric substrate to another.
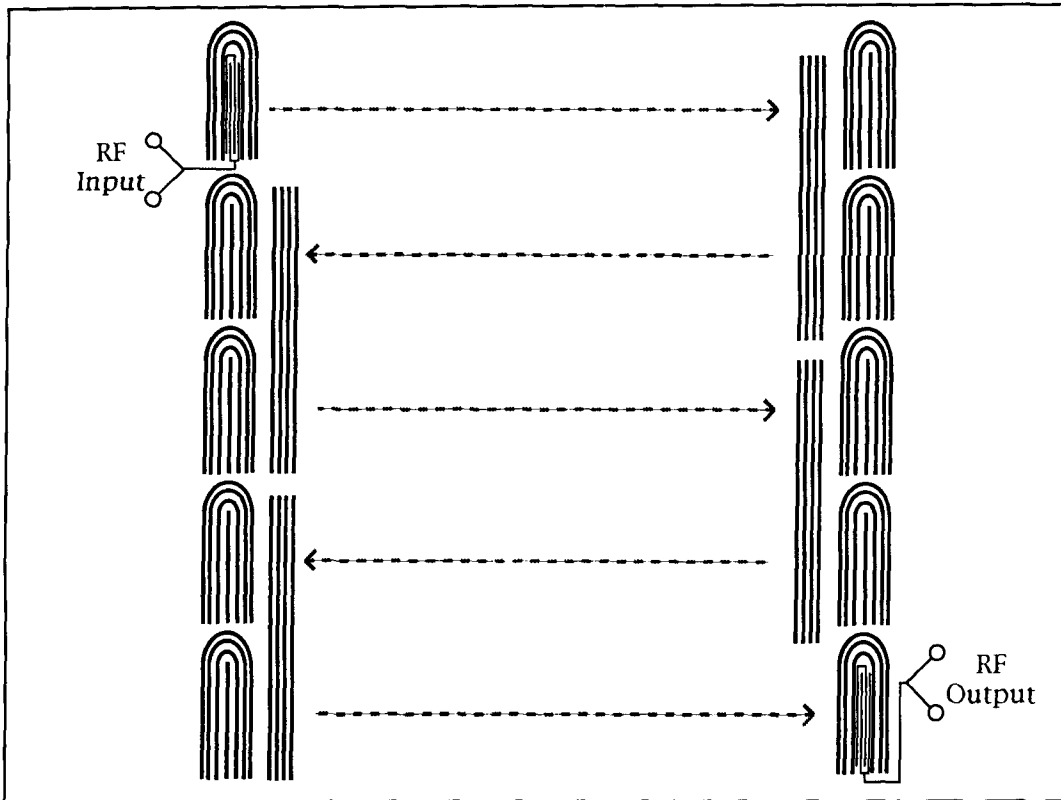
Figure 8. Extended SAW delay times can be achieved on a relatively small substrate by using multistrip structures to fold the acoustic signal, much like the prisms used to extend the light path in binoculars.

ing element that redirects and absorbs the reflection (see **Figure 5**). However, the latter approach has a finite loss.

Intra-transducer reflections, a problem with transducers having many electrodes, can be reduced by using split electrodes, each 1/8 wavelength instead of 1/4 wavelength wide (see **Figure 6**). This approach is limited by the resolution of the method used to deposit the transducer patterns on the substrate.

Because bulk waves travel at higher velocities than surface waves, they may reach the output transducer before the desired surface wave signals. As shown in **Figure 1**, bulk waves must pass through the volume of the substrate and then reflect off the back side of the substrate before reaching the output transducer. The intensity of bulk waves reaching the output transducer can therefore be minimized by roughening the back side of the substrate, which has the effect of dispersing the waves throughout the substrate.

Acoustic wave *diffraction* depends on the nature and length of the substrate, the geometry of the transducers, and the wavelength of the acoustic wave. On long paths, diffraction causes beam spreading, and therefore amplitude and phase distortion. Amplitude and phase distortion is especially significant in transducers that use a large number of electrodes or weighting

schemes involving variability in electrode overlap. The presence of metallic electrodes on the substrate surface slows the acoustic wave velocity due to shorting of the piezoelectric field. In addition, the mass loading due to the weight of the electrodes results in changes in the local acoustic impedance, which results in acoustic reflections.

In addition to the constraints that areimposed by the acoustic environment, transducers have several purely electrical limitations as well. For example, transducers have a finite electrical resistance, which results in signal loss. Electromagnetic coupling between input and output transducers can also introduce signal distortions. This form of coupling can usually be minimized by proper packaging and shielding, including metalizing and grounding the bottom of the piezoelectric substrate.

Multistrip couplers

Multistrip couplers, in the form of shallow grooves etched in the substrate or strips of deposited metal, are used to extend and enhance the basic performance of SAW devices. Once an acoustic wave has been generated by a transducer, the resulting wave can be modified and redirected in a variety of ways
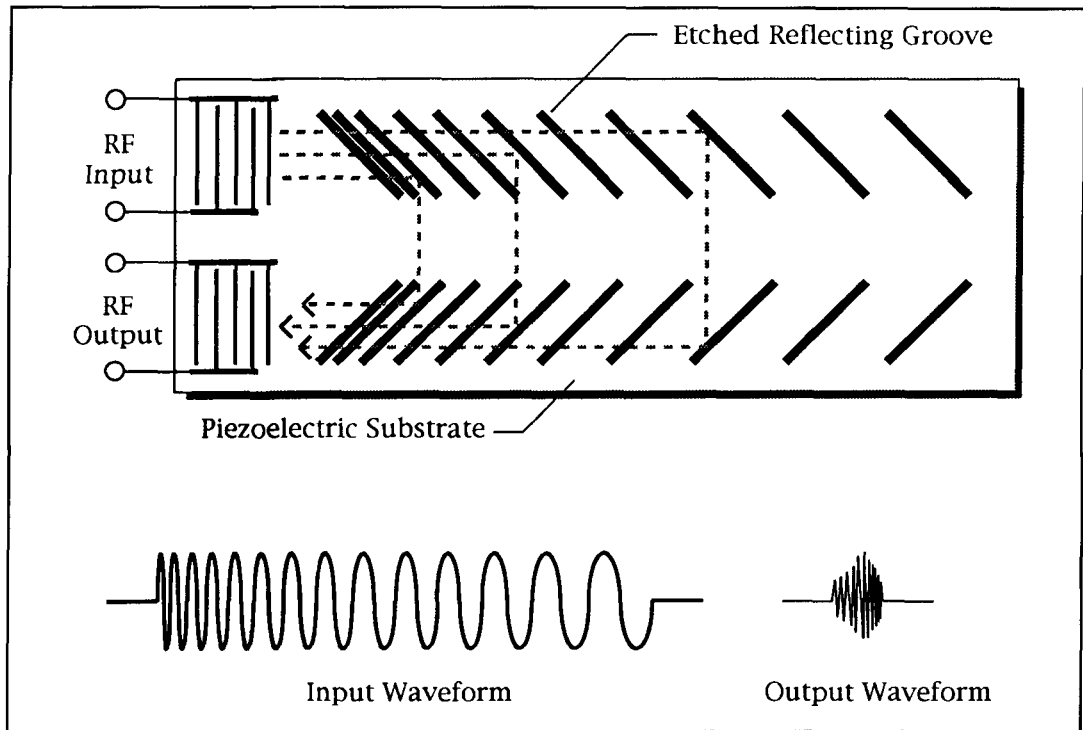
Figure 9. A SAW delay line using reflective array compression (RAC). Signals generated by the input transducer are reflected by shallow grooves etched in the substrate and reflected to the output transducer. This design supports large time-bandwidth applications; i.e., time-bandwidth products up to about 10,000.

on its trip to the receiving transducer (see **Figures 5, 7, 8,** and **9**). The applications of multistrip couplers range from rejection of spurious signals, such as triple transit echoes and suppression of bulk acoustic waves, to the creation of filters with delays virtually impossible through alternative approaches.

Grooves are generally considered to be superior to metallic overlays in the construction of multistrip couplers. Not only are grooves more easily reproduced, but metallic overlays are dispersive and introduce phase distortions in the filter response. In addition, the reflectivity of conducting overlays on high coupling piezoelectric substrates is usually excessive. Multistrip grooves, by comparison, are resistant to flaws; a multistrip coupler constructed of grooves can contain a large number of defects and still operate. A small number of flaws in a multistrip coupler constructed of metallic overlays will render the coupler useless. Although multistrip couplers are almost always fabricated on a piezoelectric substrate, the piezoelectric properties of the substrate are required only for the operation of the input and output transducers.

SAW devices

Amplifiers were the earliest acoustic wave devices (see **Figure 10**), and some were capable of providing gain up to 100 dB with band-

widths of up to 50 percent at frequencies of several hundred MHz. Although work in developing acoustic wave amplifiers continues, practical applications of surface wave technology have been restricted to passive delay lines and signal processing components. The most notable acoustic wave components—filters, delay lines, and oscillators—are described in more detail below.

SAW filters

Traditional filter design, normally considered from a frequency-domain perspective, is concerned with the lumped inductance and capacitance values, resonant frequencies, passbands (frequencies that are passed through the filter), and stopbands (frequencies that are not allowed through). When considering SAW filters, it's more intuitive to work from the perspective of a spectrum analyzer; that is, in the time domain. If the incoming signal is considered as a propagating wave, filtering can be viewed as the process by which signals are passed through a number of delay paths, after which the delayed signals are added. Under some conditions, the delayed signals add constructively (the passband), and under other conditions, they add destructively (the stopband). **Figure 11** shows the schematic symbol for a SAW filter device with interdigital electrodes.

SAW *bandpass filters* are by far the most common form of SAW technology in use today, especially as replacements for the conventional LC filters used in television receiver IF circuits and cable systems. SAW bandpass filters are capable of providing stop-band rejection of at least 60 dB, with a flat in-band response to 0.1 dB, and phase deviation of only a few degrees (see **Figure 12**). Other advantages of *SAW bandpass filters include: excellent dynamic range, stability, and linearity in a package that's both rugged and lightweight.

SAW filters are commonly used in applications where it's necessary to pass a very narrow band of frequencies relatively unattenuated, with a sharp dropoff outside the limits of the band. For example, it's not uncommon for the SAW filters used in television receivers and cable systems to have a shape factor (–3 dB to –40 dB bandwidth) as small as 1.1.[5] In addition to providing excellent performance, SAW filters are extremely space efficient, especially when compared to the multiple resonant LC sections of a conventional filter. Piezoelectric and mechanical resonators, while also providing excellent performance, are relatively expensive.

SAW filters cannot be easily detuned because their properties are established when they are fabricated. However, variations in temperature can affect transducer and substrate geometry through thermal expansion, and therefore affect filter performance. Thermal expansion of the substrate can be minimized by regulating the environmental temperature and avoiding substrates that are sensitive to temperature fluctuations.

Perhaps the major drawbacks associated with SAW bandpass filters are a relatively high insertion loss (about 20 dB)[6] and a bandwidth limited at the lower end by the size of the substrate (about 100 kHz). Loss can be reduced through the use of special transducer geometries, but not to the point where SAW bandpass filters can be used for the front-end of the IF

where the signal-to-noise ratio would be affected. SAW bandpass filters are better suited for later stages in the IF where loss would not be detrimental to the signal-to-noise performance. Bandwidth can be extended by connecting several SAW transducers in series, each operating at a different center frequency.[7]

Delay lines

Delay lines, traditionally constructed of transmission lines or equivalent components, are used to delay a signal for a predetermined length of time. Delay lines have a variety of applications, from delaying the luminance signal in the luminance channel of a color TV receiver, to compressing a received RADAR signal.

The problem with using transmission lines to delay or compress signals is that the lines must be relatively long to realize an appreciable time delay. A number of conventional solutions have been advanced to address the transmission line length problem. One approach has been the development of a special transmission line in which the center conductor is wound in a helix. Another has been to create the equivalent of a transmission line with lumped capacitors and inductors. However, both of these approaches are limited in the amount of delay that can be practically achieved. For time delays longer than a few pulse periods, SAW delay lines are used.

SAW delay lines, in the form of dispersive filters, are used primarily for pulse compression in applications where significant delays (or pulse compression) must be introduced in an environment where space is at a premium. One SAW approach to achieving pulse delays is to monotonically decrease transducer finger spacing (see **Figure 3**). For greater compression (e.g., time-bandwidth products up to 10,000), a *reflective array compressor* (RAC) design, illustrated in **Figure 9**, can be used. Discretely variable delay lines can be created by electroni-
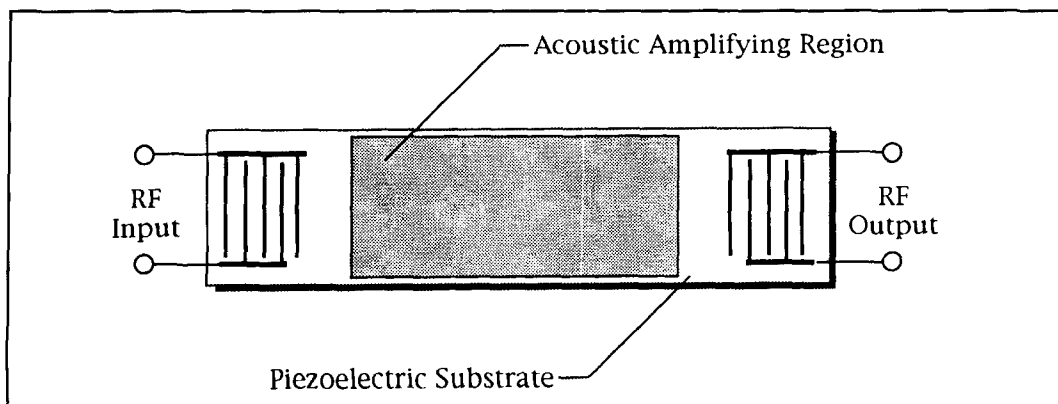
Figure 10. The basic structure of a SAW acoustic amplifier. SAW acoustic amplifiers have yet to reach the popularity of SAW filters and delay lines.
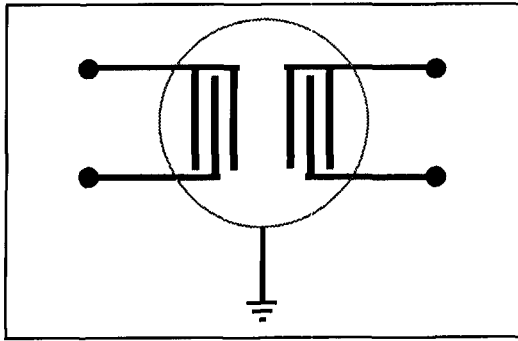
Figure 11. The schematic symbol for a SAW filter device with interdigital electrodes.

cally selecting from a set of transducers spaced along the path of the acoustic beam.

Oscillators

Of the common methods used to generate signals, including those based on phase locked loops (PLLs) and direct digital synthesis (DDS), a quartz crystal oscillator is often used to stabilize or control the frequency of oscillator circuits. However, fundamental frequency crystals are only practical up to about 25 MHz. At higher frequencies, overtone crystals or some sort of frequency multiplier must be used. In addition, quartz crystals are relatively fragile, and often require considerable mounting space.

SAW oscillators, which can be constructed by simply feeding a SAW delay line's output back to its input via an amplifier with excess

gain (see **Figure 13**), provide a number of advantages over quartz crystal and LC oscillators. Because SAW oscillators are planar devices, they are extremely space efficient and can be made mechanically rugged. Secondly, SAW oscillators can operate from about 20 MHz up to over 1 GHz. Qs on the order of 10^5 at 100 MHz and 10^4 at 1 GHz are typical. The frequency of oscillation is a function of the transducer patterns and the relative positions of the input and output transducers, and not the overall dimensions of the substrate. The substrate can therefore be constructed as thick as needed, and bonded to other objects—providing mechanical stability and a thermal sink. Because the oscillation frequency is a function of the relative separation of the interdigital transducers, SAW oscillators can be used as strain gauges, with the frequency output varying in accordance with changes in the SAW delay substrate.

Although SAW resonators can be both smaller and less expensive than traditional quartz or LC resonators, the *long-term stability* of SAW oscillators is less than that of quartz crystal oscillators. The aging rates of SAW oscillators, which can result in slight changes in input and output transducer geometry as well as relative transducer distances, are about an order of magnitude greater than crystal oscillators.

The future of SAW technology

SAW devices are becoming increasingly important in UHF and microwave communica-
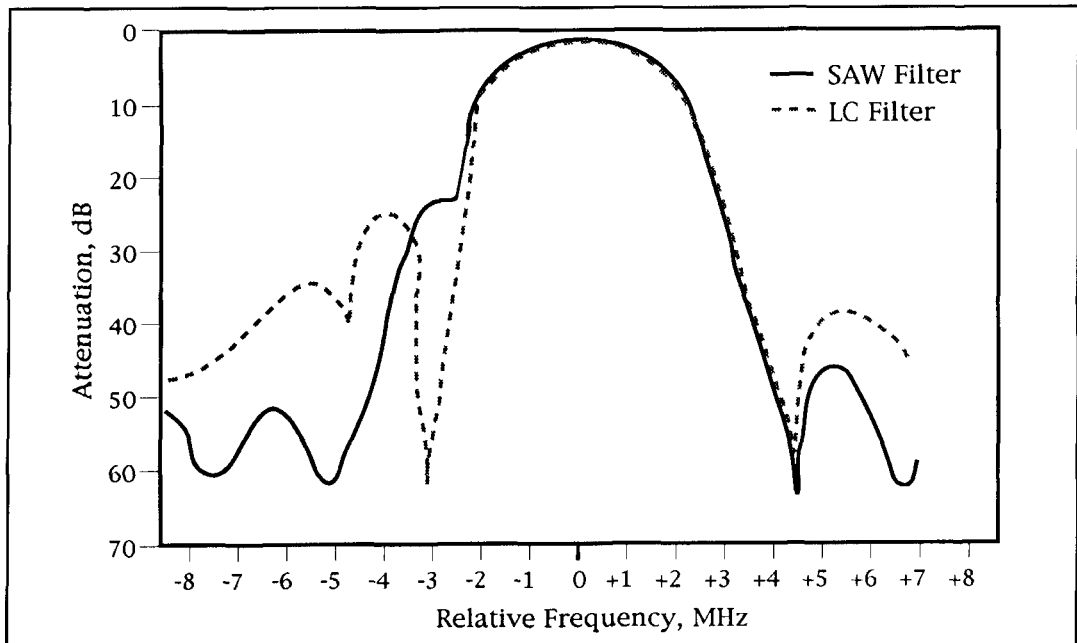
Figure 12. SAW versus conventional LC bandpass filter response. In addition to providing a relatively flat in-band response, properly designed SAW bandpass filters introduce negligible phase deviation.
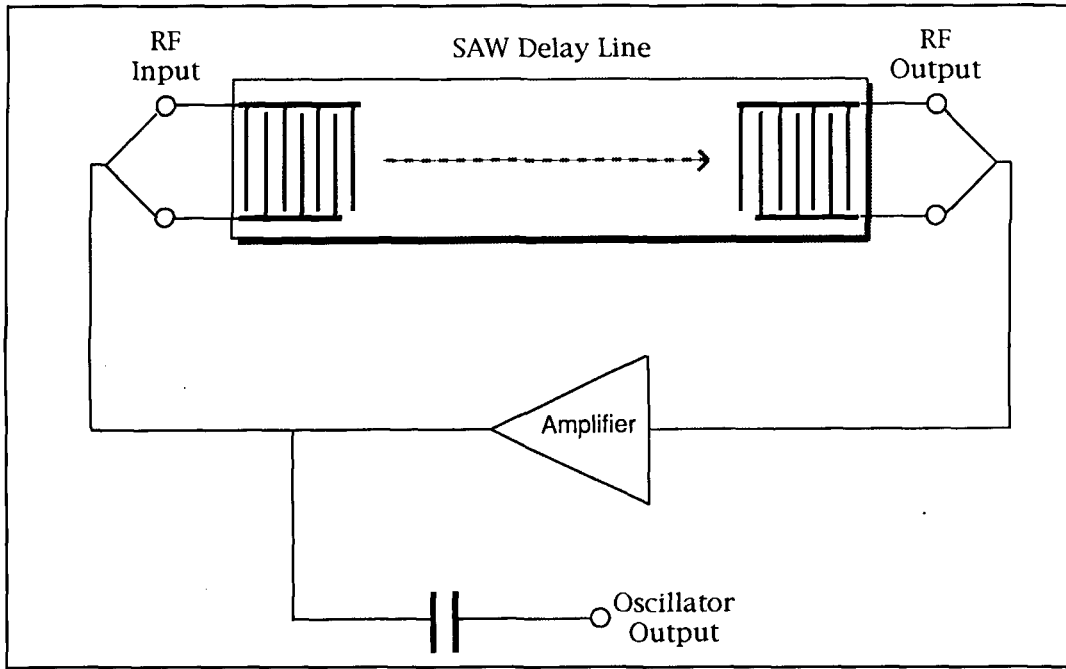
Figure 13. A basic SAW oscillator can be constructed by feeding the output of a SAW delay line back to its input via an amplifier.

tions. They can be used to reduce cost and size, while simultaneously improving reliability of communications systems. SAW technology shows particular promise in spread spectrum communications, owing to the wide bandwidths and high operating frequencies supported by SAW devices.

Although SAW devices are not obviously superior to some competing technologies, like direct digital synthesis, they have the edge in applications requiring ruggedness and reliability over other considerations. In addition, SAW devices have the added advantage of simplicity. Like DDS, one advantage of using SAW filters is that they can be designed and replicated with ease; amplitude and phase specifications can be met by a single filter. Similarly, SAW devices are stable, do not require adjustment, and can be interfaced with standard IC amplifiers. As is the case with all modern electronic components, further size reductions in SAW devices can be expected. Current SAW filters designed for cellular phones are about the size of SMT capacitors and resistors.

SAW devices are finding a variety of applications in electronic warfare. For example, SAW devices are used to provide simple time delay, frequency dependent time delay, discretely variable time delay, bandpass filtering, and generation and recognition of various coded waveforms. SAW devices, in the form of tapped surface wave delay lines, can be used to generate both frequency hopped and pseudo noise phase modulated signals for spread spectrum communications. There are also applica-

tions of acoustic digital memories, similar in operation to charge-coupled devices, but operating at hundreds of MHz.

For those who want to learn more about the SAW technology, there are a number of excellent texts available. The work by Mario[3] provides a good introduction to the basics of electromechanical coupling involved in electroacoustic transducers. For more in-depth information on the physics of electroacoustic transducers see the texts by Zdenek,[8] Merhaut,[9] and Rosenbaum.[10] The text edited by Matthews[11] provides an outstanding description of the physics relevant to the design and construction of SAW filters. ∎

REFERENCES
1. J. Feinstein, R. Nelson, G. Farney, D. Preist, A. Scott, and P. Walchli, "UHF and Microwave Devices," *Electrical Engineering Handbook*, Fink and Christiansen editors, 1986, McGraw-Hill, New York.
2. G. Weiss, "Electrical Transducers," *Handbook for Modern Electronics and Electrical Engineering*, Belove, Hopkins, Nelson, Rosenstein, and Shinners editors, 1986, John Wiley & Sons, New York.
3. R. Mario, *Acoustics and Electroacoustics*, 1988, Artech House, Inc., Norwood, Massachusetts.
4. R. Mitchell, "Surface Acoustic Wave Transversal Filters: Their Use and Limitations," *Component Performance and Systems Applications of Surface Acoustic Wave Devices*, 130-9, 1973.
5. P. Horowitz and W. Hill, "High-frequency and High-speed Techniques," *The Art of Electronics*, 1980, Cambridge University Press, Cambridge, Massachusetts.
6. L. Hoke, "Television Broadcast Receivers," *Electrical Engineering Handbook*, Fink and Christiansen editors, 1986, McGraw-Hill, New York.
7. R. Velimir, *Principles of Acoustic Devices*, 1983, John Wiley & Sons, Inc., New York.
8. S. Zdenek, *Vibrating Systems and Their Equivalent Circuits*, 1990, Elsevier Science Publishing Company, Inc., New York.
9. J. Merhaut, *Theory of Electroacoustics*, 1981, McGraw-Hill, Inc., New York.
10. J. Rosenbaum, *Bulk Acoustic Wave Theory and Devices*, 1988, Artech House, Inc., Norwood, Massachusetts.
11. H. Matthews, *Surface Wave Filters: Design, Construction, and Use*, 1977.

Rick Littlefield, *K1BQT*
109A McDaniel Shore Drive
Barrington, New Hampshire 03825

QUARTERLY DEVICES
The Harris Semiconductor HFA3600 Low-noise Amplifier/Mixer

Building a good UHF receiver has always been a challenge. However, the 900-MHz consumer-electronics revolution continues to generate a windfall of special-purpose chips that makes the job easier. In this edition of "Quarterly Devices," we'll look at a new offering from Harris Semiconductors—the HFA3600 Low-noise Amplifier/Mixer.

General description

The HFA3600 is a sensitive 800 to 1000-MHz down-converter on a chip that features a 50-ohm input impedance, 2.3-dB noise figure LNA, and active mixer. By combining these

critical receiver elements into a single 14-pin SOIC package, Harris engineers have lowered the external parts count and removed much of the "touchiness" inherent in UHF front-end design. Combined LNA/mixer intercept performance is a respectable –16.7 dBm (referred to input), and the LO drive requirement is relatively low—around –3dBm. Like many newer consumer-application chips, the HFA3600 runs on 4 to 5 volts and provides an integrated power-down function to conserve battery life in hand-held products. Typical power-on state consumption is just over 11 mA.

Figure 1 shows the HFA3600's internal block diagram. The LNA is internally optimized for direct 50-ohm input. A collector spiral inductor improves stability at lower frequencies while optimizing gain in the 900-MHz range (see **Photo A** for an inside look). In addition to a 2.3-dB noise figure, the LNA section provides nearly 13 dB power gain with a third-order output intercept of +12.8 dBm.

The HFA3600's single-balanced active mixer features open-collector output. This gives designers the flexibility to set conversion gain and match the device into a wide range of IF filters. At –3 dBm LO injection, mixer conversion gain is around 7 dB with a third-order output intercept point of +3.2 dBm.

To ensure stability and consistent performance, LNA and mixer bias currents are laser-trimmed to hold within 1 percent over the device's temperature range. The bias network is gated to provide a fast-switched power-down mode for battery-powered or TDM (time-division multiplex) applications. Standard TTL-level input controls the gate, reducing current consumption to 250 µA when logic zero is applied. Although characterized as a 5-volt device, the HFA3600 operates comfortably at 4 volts with minimal degradation in performance.
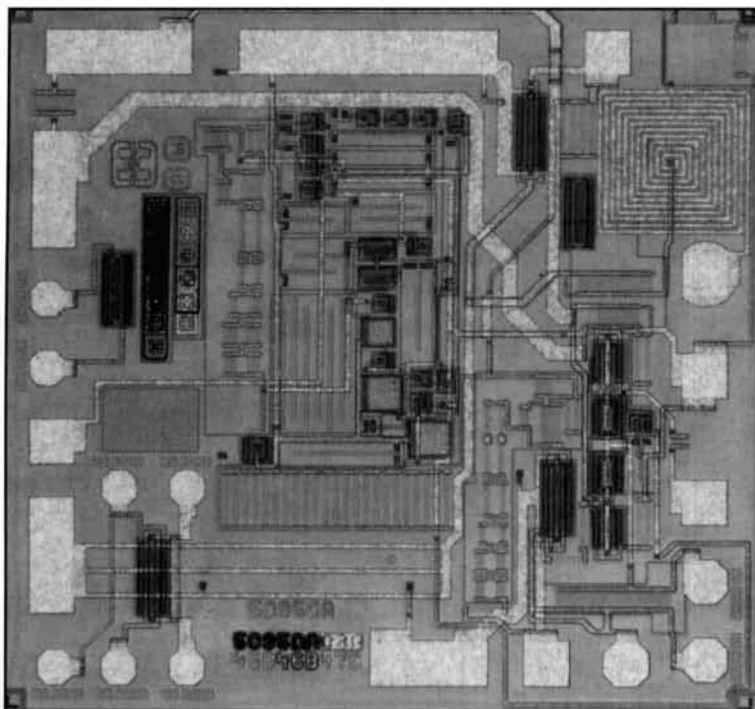
Photo A. Inside the HFA3600.

Detailed operating parameters are presented in **Table 1**.

A typical circuit

Figure 2 illustrates a typical down-converter application circuit. The device's input frequency range, which is established by the LNA's internal network, covers from 800 to 1000 MHz. IF output can range anywhere from 45 to 200 MHz—depending on the frequency of the 1st LO signal. Synthesized or continuously tunable receivers operating in the IF region may be used as tunable IFs.

Supply Vcc for the LNA section goes to pin 1. Appropriate RF and low-frequency bypassing is used to decouple the Vcc line.

Pins 2, 4, 5, 10, and 12 are bonded to the pc groundplane. Ground pins are strategically placed to help isolate the various input and output ports on the device.

Pin 3, the LNA's input port, must be AC-coupled to the 50-ohm antenna line. Note that all AC-coupled inputs and outputs require a minimum coupling capacitance of 100 pF; a value of 1000 pF is suggested in the application note. SMD chips are a must at these frequencies.

Pins 6 and 7 make up the LO port. The bypass capacitor on pin 6 must be positioned as close to the device as possible. LO input goes to pin 7 through a coupling capacitor.

The device's power-down function is controlled at pin 8. This can be driven by TTL or
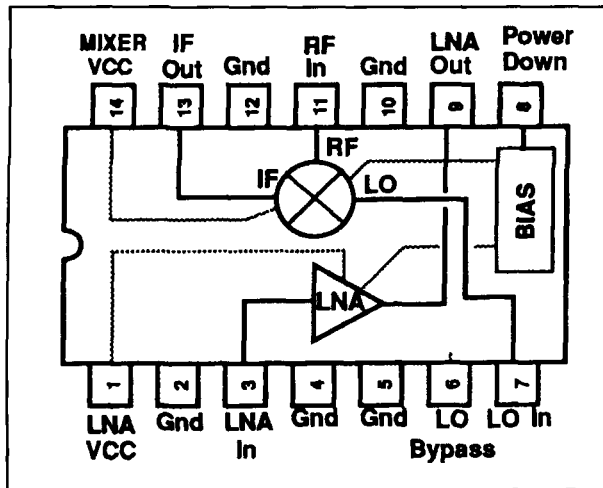
Figure 1. Internal block diagram of the HFA3600.

CMOS-level logic. Logic-low input invokes the power-down mode. An internal pull-up is provided to enable the chip when pin 8 is left open.

Pin 9 is the LNA's output port. In most applications, this will be looped to the mixer's RF-input port at pin 11 through a 50-ohm image filter. The image filter rejects unwanted UHF signals occurring on the image side of the LO, preventing them from entering the mixer.(Note: The image filter is typically of a bandpass configuration.)

Pin 13 is the mixer's IF output port. Because this is an open-collector circuit, Vcc is supplied externally through a resistor. Resistor value

DC Electrical Specifications

SYMBOL	PARAMETER	CONDITION	TEST LEVEL	TEMP	ALL GRADES			UNITS
					MIN	TYP	MAX	
I_{CC}	Total Supply Current @ 5V	Normal PD= 2V	A	25°C		11.3	12.5	mA
		Shutdown PD =0.8V	A	25°C		250	375	µA
V_{IH}	Shutdown Logic High	Normal Mode	A	25°C	2		V_{CC}	V
V_{IL}	Shutdown Logic Low	Shutdown Mode	A	25°C	-0.3		0.8	V
I_{IL}	Shutdown Input Current	PD = 0.4V	A	25°C	-200	-150	-100	µA
I_{IH}	Shutdown Input Current	PD = 2.4V	A	25°C	-45	-24	-3	µA
V_{LNA-IN}	LNA Input DC Level	Normal Mode	A	25°C		0.79		V
		Shutdown Mode	A	25°C		0.37		V
$V_{LNA-OUT}$	LNA Output DC Level	Normal Mode	A	25°C		4.9		V
		Shutdown Mode	A	25°C		5.0		V
V_{MX-RF}	Mixer RF$_{in}$ DC Level	Normal Mode	A	25°C		0.79		V
		Shutdown Mode	A	25°C		tbd		V
V_{MX-LO}	Mixer LO$_{in}$ DC Level	Normal Mode	A	25°C		2.1		V
		Shutdown Mode	A	25°C		tbd		V
t_{ON-OFF}	Shutdown On-Off time		B	25°C		10		µs

Table 1. HFA3600 operating specs. (Continued on next page)

AC Electrical Specifications

All characterization results have been obtained with the use of a standard evaluation board.

SYMBOL	PARAMETER	TEST LEVEL	TEMP	ALL GRADES			UNITS
				MIN	TYP	MAX	
LNA (V_{CC} = +5V, T_A = 25°C, Test Figure 1 and f = 900MHz unless otherwise noted in characterization curves)							
S_{21LNA}	LNA Gain	B	+25°C	11.8	12.8	13.8	dB
S_{12LNA}	LNA Reverse Isolation	B	+25°C		23		dB
S_{11LNA}	LNA Input Return Loss	B	25°C	6.0	7.3		dB
S_{22LNA}	LNA Output Return Loss	B	25°C	10.0	13.0		dB
$P_{-1dBLNA}$	LNA Output 1-dB Gain Compression Point	B	25°C		-2.0		dBm
IP_{3LNA}	LNA Output 3rd-order Intercept	B	25°C	+11.2	+12.8		dBm
NF_{LNA}	LNA Noise Figure	B	25°C		2.30	2.60	dB
MIXER (V_{CC} = 5V, T_A = 25°C, f_{LO} = 825 MHz@-3 dBm, f_{RF} = 900 MHz, f_{IF} = 75 MHz and Test Figure 1 unless otherwise noted)							
PG_C	MIXER Power Conversion Gain	B	25°C	5.9	7.0	8.1	dB
S_{11RF}	MIXER RF Input Return Loss	B	25°C	8.0	11.0		
S_{11LO}	MIXER LO Input Return Loss	B	25°C	18.0	26		dB
NF_{MIXER}	MIXER SSB Noise Figure	B	25°C		12.1	13.9	dB
$P_{-1dBMIX}$	MIXER Output 1-dB Gain Compression	B	25°C		-7.5		dBm
IP_{3MIX}	MIXER Output 3rd-order Intercept	B	25°C	+1.0	+3.2		dBm
C_{OUTMIX}	MIXER IF Output Capacitance	B	25°C		2.3		pF
G_{RF-IF}	MIXER RF-IF Isolation (includes matching network)	B	25°C		25		dB
G_{LO-IF}	MIXER LO-IF Isolation (includes matching network)	B	25°C		16		dB
G_{LO-RF}	MIXER LO-RF Isolation	B	25°C	16	21		dB
$G_{LO-LNAin}$	Mixer LO-LNA$_{in}$ Isolation	B	25°C	42	50		dB
$G_{LNAout-RF}$	LNA$_{out}$-Mixer RF$_{in}$ Isolation	B	25°C	35	40		dB
(LNA + MIXER) V_{CC} = 5V, T_A = 25°C, f_{LO} = 825 MHz @-3 dBm, f_{RF} = 900 MHz, f_{IF} = 75 MHz and idealized lossless external filters							
CPG_C	Power Conversion Gain	B	25°C		19.8		dB
CNF	Noise Figure	B	25°C		3.97		dB
CIP_3	Input 3rd-order Intercept	B	25°C		-16.7		dBm

Table 1. HFA3600 operating specs.
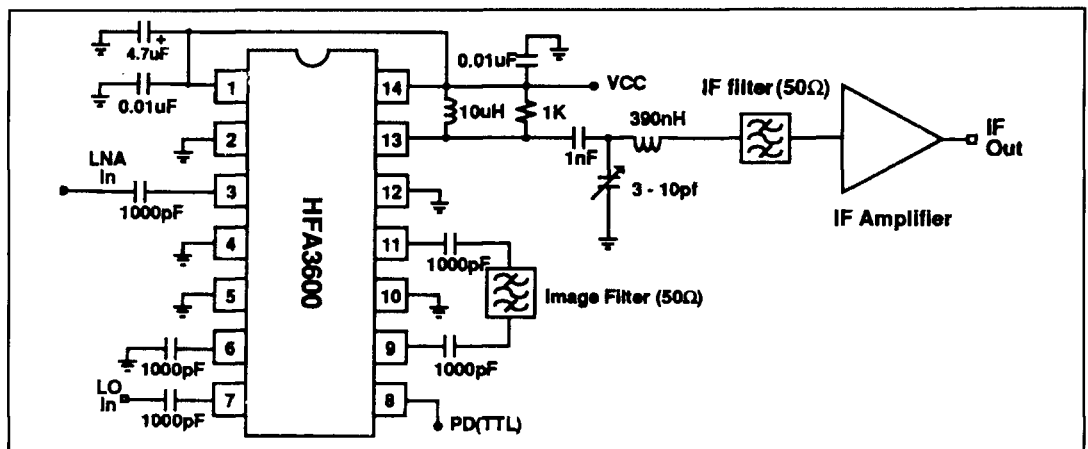
Figure 2. Typical circuit.

Evaluation Board layout. Scale X1 Top View
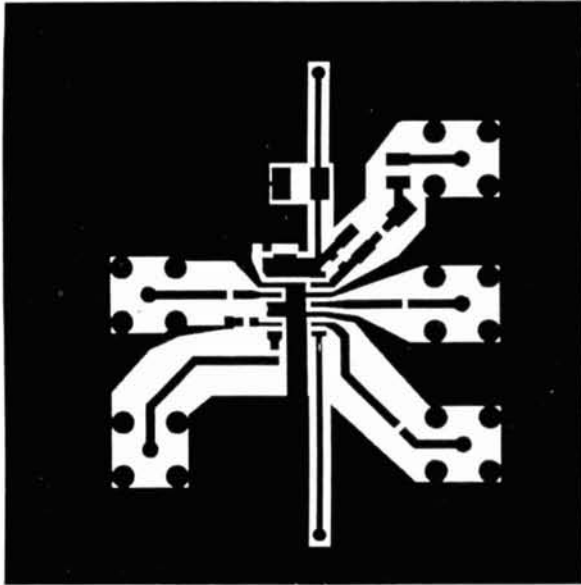
Component List.

R1 Res, fixed 1KW

L1 Ind., fixed 10uH

L2 Ind., fixed 390nH

C3,C4,C5,C7,C10,C11 Cap, fixed 1nF

C1,C6 Cap, fixed.01uF

C2 Cap, fixed Tantalum. 4.7uF

C8 Cap, var. 3 to 10pF

Cr1 Diode DL4001

Evaluation board Component Placement
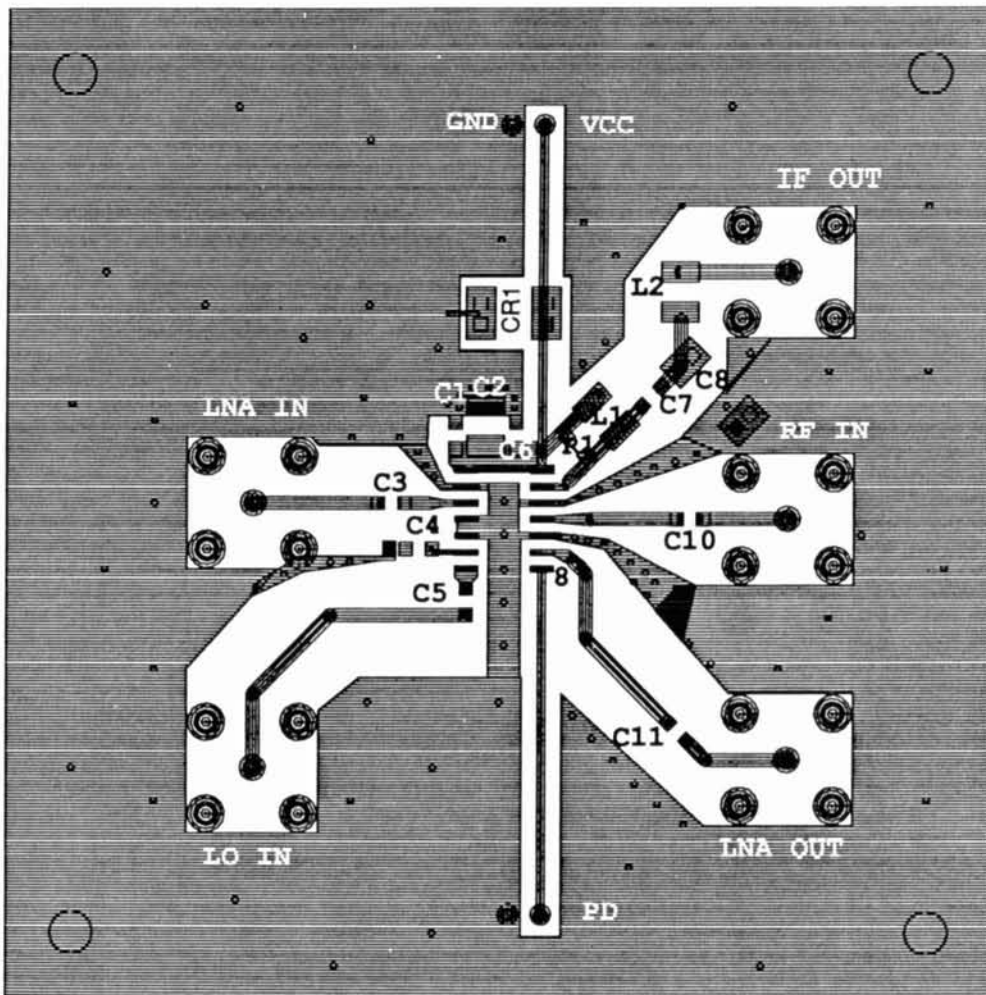
Figure 3. Board layout for experimentation.

will determine mixer conversion gain and IP3. (Acceptable values range from 300 to 1500 ohms, with smaller values providing less gain and a lower output impedance.) An optional 10-µH RF choke may be installed in parallel with the resistor to lower DC resistance and increase output swing. Some form of impedance transformation is recommended when feeding a 50-ohm IF system.

Vcc to pin 14 powers both the RF mixer and bias network portion of the chip. Note that 0.01 bypasses are shown at *both* Vcc inputs (1 and 14). These must be positioned as close as possible to their respective pins and provide a direct low-impedance path to pc ground.

Although an integrated device like the HFA3600 can greatly simplify UHF construction, you must still adhere to sound 900-MHz RF-design practices. For example, non-stripline traces must be short and direct—with angles or pads shaped or chamfered for minimum parasitic contribution. Chip capacitors and resistors are a must, and pc-board dielectric should be suitable for 1-GHz circuitry. To achieve optimum gain and stability, use two-sided board with abundant feedthroughs to solidify the groundplane. **Figure 3** provides a board layout suitable for general experimentation. When the HFA3600 is used in a single channel application or in conjunction with a tunable IF, a crystal-controlled UHF-transmitter IC like the Motorola MC13176 can provide a simple and stable 700-MHz LO source (see "Quarterly Devices," Fall 1993).

Applications

The HFA3600 is primarily designed for use in portable cellular phones, wireless data transceivers, UHF mobile radios, cordless phones, and wireless telemetry. However, the HFA3600 should work equally well in any specialized communication project that requires a UHF receiver front end. Here are a few possibilities:

- Transverters for 900-MHz SSB service.
- ATV down-converters for TV receivers.
- Repeater control-channel and voice-uplink receivers.
- Packet backbone data receivers.
- Spread-spectrum equipment.
- TDM-multiplexed digital-audio transceivers.
- Block down-converters for VHF-range panadapters or analyzers.
- Simple 800-MHz scanner converters.

HFA3600 price and availability

The HFA3600 is presently available from Harris Semiconductor for $3.75 each in 100 lots and $3.56 each in 1000 lots. Application notes, samples, evaluation boards, and referrals for technical assistance are handled through the local Harris sales office in your area (to obtain the number of your local representative, call 1-800-442-7747). Small quantity purchases of the HFA3600 may be made through any Harris authorized distributors, including Newark Electronics (1-800-367-3575).

Conclusion

The 800 to 1000-MHz radio spectrum continues to be a hot area for new RF devices—with plenty of room for exploration and new-product development. If you come up with a unique HFA3600 application or construction project you'd like to share with our readers, please write to *Communications Quarterly*, P.O. Box 465, Barrington, New Hampshire 03825—or call our editorial offices at 603-664-2515. We'd love to put your name in print! ∎

PRODUCT INFORMATION

New High Voltage Relay & Contactor Application And Product Guide

Kilovac Corporation has published an expanded 75 page Application and Product Guide for industrial and aerospace applications.

The new guide begins with an introduction to high voltage relays and contactors. It explains the physics of vacuum and gas-filled relays, defines relay and contactor terms, and shows, with cut away type drawings, common high voltage relay designs. It also includes information on various high voltage and power switching relay applications and the criteria that should be used to select the proper relay to ensure trouble-free performance. Special sec-

tions include RF deratings, capacitive discharge loads, power switching loads and high insulation resistance relays for test equipment.

Kilovac's Application and Product Guide introduces some of Kilovac's new 270 Vdc relays and contactors. These include the AP5 and AP10 relays, rated at 5 amps and 10 amps at 270 Vdc, respectively. Also new are the AP90 and the AP350 with 90 amp and 350 amp load switch ratings at 270 Vdc.

If you have any questions about your specific relay application, or would like to obtain Kilovac's Application and Product Guide, call 1-800-253-4560 or write Kilovac Corporation, P.O. Box 4422, Santa Barbara, CA 93140.

A.E. Popodi, *AA3K/OE2APM*
Moosstrasse 7/6, A-5020, Salzburg
Austria

A NEW METHOD FOR MEASURING CABLE LOSS

Try this simple technique

In an existing antenna system, it is often necessary to measure the attenuation of the coaxial cable that connects the antenna to the transceiver—especially after many years of outdoor operation. The result can then be compared with data sheet values to see whether any deterioration has occurred. Likewise, if you install a new cable of unknown origin, you undoubtedly would like to know its true condition. However, measuring cable loss can be a frustrating endeavor, especially when the loss is small, the cable is short, or when the frequency is low.

The simplest, though not the best, way to determine the cable loss seems to be to measure the difference between input and output power of the properly matched cable with an RF wattmeter. A more accurate method is to measure the input and output voltages with an RF voltmeter. Both methods require a power source, proper impedance matching, and instruments that, in most cases, the amateur does not have. Ground currents in the test setup and mismatch errors can cause false readings if you consider that the difference between input and output voltage is only 0.89 percent for an RG-213 cable, 2 meters long, at a frequency of 28 MHz. It is not unusual to measure a larger output power than input power.

There is a useful method that is especially suitable for short cables (see **Reference 1**). It is based on the fact that the ratio of input to output voltage of an unterminated, quarter-wavelength cable driven from a low impedance source, is nearly equal to the cable loss in nepers (1 neper = 8.69 dB). However, for this test you need a power source, a little test fixture consisting of an amplifier and cable driver, and access to both ends of the cable at this fixture.

It is therefore desirable to have a method that does not require costly instruments, cable matching, terminations, and that is simple, fast, and accurate.

The method that meets these requirements is based on the fact that the maximum of the real part of the input impedance of an unterminated or shorted cable is directly proportional to the cable loss in nepers.

Figure 1 is a plot of the real and imaginary part of the cable input impedance $Z_{IN} = R_{IN} + jX_{IN}$ for an unterminated cable, plotted versus frequency. It shows that the input resistance R_{IN} has a maximum value at frequencies that are multiples of the frequency at which the cable is half a wavelength long. The reactive part, X_{IN}, is zero at these points. Calculations show that R_{max} is directly proportional to the cable loss in nepers. The value of R_{max} decreases with successive peaks whereas R_{min} increases with increasing frequency.

How to measure R_{max}

The ideal instrument for measuring R_{max} is the noise bridge. Because the reactive part is zero at R_{max}, the reactance dial must be set to zero. Because R_{max} in most cases exceeds the R range of the bridge, you must shunt the cable input with a resistor to bring it into range. If, for instance, the R range of the bridge is 250 ohms, you can use a noninductive 240-ohm resistor in parallel with the cable. If the resistance reading
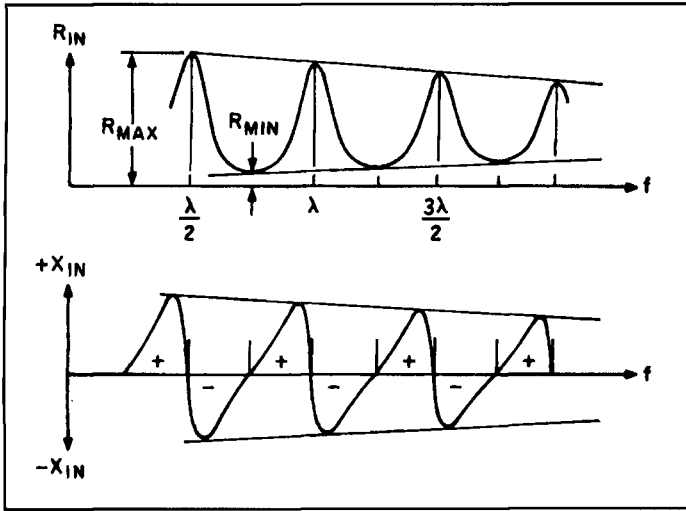
Figure 1. Plot of the real and imaginary part of the cable input impedance $Z_{IN} = R_{IN} + jX_{IN}$ for an unterminated cable, plotted versus frequency.

of the bridge for null then is, for example, 200 ohms, the R_{max} is actually 1200 ohms.

Calculating cable attenuation

Now we must prove that we can calculate the cable loss from R_{max}. The basic transmission line formula for the input impedance is given below and illustrated in **Figure 2**:

$$Z_{IN} = \frac{Z_L + Z_0 \tanh(a + jb)}{1 + \frac{Z_L}{Z_0} \tanh(a + jb)} \quad (1)$$

where Z_L is the load impedance, Z_0 is the characteristic cable impedance, a is the cable loss in nepers, and b is the cable length, expressed in electrical degrees. If we assume $Z_L = \infty$ (open cable output), the formula simplifies to:

$$Z_{IN} = \frac{Z_0}{\tanh(a + jb)} \quad (2)$$

We proceed by using the identity:

$$\tanh(a + jb) = \frac{\sinh 2a + j \sin 2b}{\cosh 2a + \cos 2b}$$

Because the electrical cable length at the half-wavelength frequency is b = 180 degrees, $\sin(2b) = 0$, and $\cos(2b) = +1$, the formula simplifies to:

$$\tanh(a + jb) = \frac{\sinh 2a}{\cosh 2a + 1}$$

The formula for Z_{IN} now becomes:

$$Z_{IN} = R_{MAX} = Z_0 \frac{\cosh 2a + 1}{\sinh 2a} \quad (3)$$

This shows that the reactive part of Z_{IN} is zero at the point where $Z_{IN} = R_{max}$ and that the input impedance is a pure resistance that can be measured easily with the noise bridge Using the known identities:

$$\sinh 2a = \frac{e^{2a} - e^{-2a}}{2}$$

$$\cosh 2a = \frac{e^{2a} + e^{-2a}}{2}$$

Equation 3 now becomes:

$$\frac{R_{MAX}}{Z_0} = \frac{e^{2a} + e^{-2a} + 2}{e^{2a} - e^{-2a}}$$

Multiplying numerator and denominator by e^{2a}, we have:

$$e^{4a} - e^{2a} \frac{2Z_0}{R_{MAX} - Z_0} - \frac{R_{MAX} + Z_0}{R_{MAX} - Z_0} = 0$$

Because e^{4a} is the square of e^{2a}, we have a quadratic equation for e^{2a} that allows us to calculate the cable loss. Its solution is:

$$e^{2a} = \frac{Z_0}{R_{MAX} - Z_0} + \sqrt{\left(\frac{Z_0}{R_{MAX} - Z_0}\right)^2 + \frac{R_{MAX} + Z_0}{R_{MAX} - Z_0}}$$

or

$$a = \frac{1}{2} \ln \left\{ \frac{Z_0}{R_{MAX} - Z_0} + \sqrt{\left(\frac{Z_0}{R_{MAX} - Z_0}\right)^2 + \frac{R_{MAX} + Z_0}{R_{MAX} - Z_0}} \right\} \text{ nepers} \quad (4)$$

where Z_0 is the characteristic cable impedance and R_{max} is the input resistance as measured with the noise bridge.

Calculating cable loss for the shorted cable

The equation above was derived for the

unterminated cable. However, the principle can also be applied to the shorted cable, as shown in **Figure 3**.

The input impedance of the shorted transmission line is:

$$Z_{IN} = Z_O \tanh(a + jb) \qquad (5)$$

The input resistance maxima now occur at odd multiples of the quarter-wave frequency. In this case, the value of b is 90 degrees and $\sin(2b) = 0$, but $\cos(2b) = -1$. The quadratic equation for e^{2a} now becomes:

$$e^{4a} - e^{2a} \frac{2R_{MAX}}{R_{MAX} - Z_O} + \frac{R_{MAX} + Z_O}{R_{MAX} - Z_O} = 0$$

and the cable attenuation for the shorted cable is:

$$a = \frac{1}{2} \ln \left\{ \frac{R_{MAX}}{R_{MAX} - Z_O} + \qquad (6) \right.$$

$$\left. \sqrt{\left(\frac{R_{MAX}}{R_{MAX} - Z_O}\right)^2 - \frac{R_{MAX} + Z_O}{R_{MAX} - Z_O}} \right\} \text{ neper}$$

Although **Equation 6** looks quite different from **Equation 4**, it delivers exactly the same results. You can use either formula, regardless of whether the cable is open or shorted.

Practical example

Let us examine an RG-213 cable 33.238 meters long, unterminated, with a nominal velocity constant of 0.66. This cable will be one wavelength long at 5.957 MHz:

$$f_0 = \frac{300 \times 0.66}{33.238} = 5.957 \text{ MHz}$$

Because the velocity factor may not be exactly 0.66, the measurement frequency may need to be varied to find the exact value that gives R_{max}. (The purpose of this calculation is to aid in finding R_{max}; the actual frequency is not needed in the formula.) The noise bridge is used to measure R_{max}. Because R_{max} may exceed the range of the noise bridge, a noninductive resistor is placed in parallel with the cable input. I recommend use of the highest possible resistor value that can *just* be measured with the bridge. If, for instance, the bridge potentiometer has 250 ohms, a resistor between 240 and 250 ohms would be best. Here is the measurement procedure:

First, set the reactance dial of the bridge to the zero point, because the reactive part of Z_{IN}
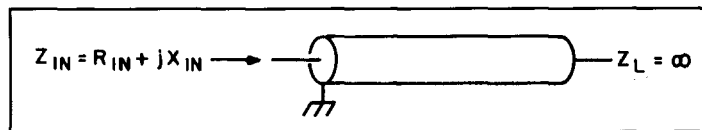
Figure 2. The basic transmission line formula for the input impedance.

is zero for $R_{IN} = R_{max}$. Do this by connecting a 240-ohm resistor at the X terminal of the bridge. Assuming a measuring frequency of 4 x 5.957 MHz, tune the receiver approximately to this frequency and null the bridge. The reactance dial is now set to zero—a setting that does not necessarily agree with the dial.

Second, measure the value of the variable bridge resistor with a digital multimeter. The reading should be 240 ohms, but it will deviate in most cases. If the reading is, for instance, 247.4 ohms, the readings are too high and must be reduced by a factor of 240/247.4 = 0.97. I recommend this resistance check because most noise bridges are rather inaccurate at the higher resistance values.

Third, connect the coax cable in parallel with the 240-ohm resistor at the X terminal and null the bridge without touching the reactance dial. Do this by adjusting R and, if required, the frequency until you obtain a deep noise minimum. Measure the variable resistor of the noise bridge with the DVM once again. In our example, the reading is 154.6 ohms. The corrected value is then 154.6 x 0.97 = 150 ohms.

Fourth, calculate R_{max}. Because R_{max} and the 240-ohm resistor are in parallel, the formula

$$\frac{R_{MAX} \times 240}{R_{MAX} + 240} = 150 \text{ ohm} \quad \text{gives} \quad R_{MAX} = 400 \text{ ohm}$$

yields $R_{max} = 400$ ohms

Thus, **Equation 1** becomes:

$$a = \frac{1}{2} \ln \left(\frac{50}{400 - 50} + \right.$$

$$\left. \sqrt{\left(\frac{50}{400 - 50}\right)^2 + \frac{400 + 50}{400 - 50}} \right) = 0.1256 \text{ neper}$$

Because 1 dB = 0.115 neper, the cable loss in dB is 0.1256/0.115 = 1.093 dB or (1.093 x 30.48)/33.238 = 1 dB per 100 feet. The data sheet value for this frequency is 1.1 dB.

If you want to find the cable loss at another frequency, disconnect the cable briefly and null the bridge with the 240-ohm resistor alone at the new frequency because the zero reactance point may be different. Then proceed as before.

If your receiver does not have continuous tuning, pick one or more frequencies that pertain either to multiples of the half-wavelength frequency or to odd multiples of the quarter-
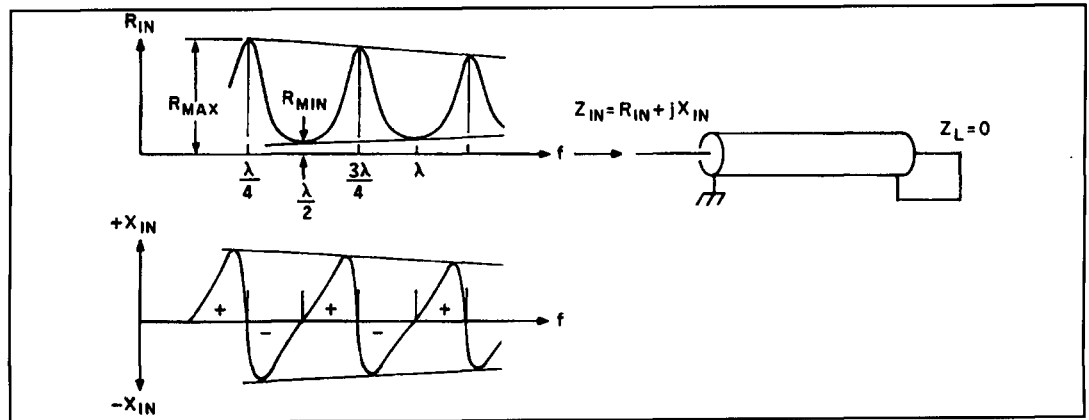
Figure 3. The equation derived for the unterminated cable can also be applied to the shorted cable.

wavelength frequency that will then be covered by the receiver.

Calculating loss at other frequencies

So far we have calculated the loss for only one frequency. But if you make another measurement at another frequency, you will have two sets of attenuations, a_1 and a_2, and two frequencies, f_1 and f_2. Using the interpolation method described by K2BT,[2] you can calculate the loss at any other frequency using the following procedure.

First, calculate two constants m and n:

$$m = \frac{a_1 f_2 - a_2 f_1}{f_2 \sqrt{f_1} - f_1 \sqrt{f_2}} \quad n = \frac{a_2 \sqrt{f_1} - a_1 \sqrt{f_2}}{f_2 \sqrt{f_1} - f_1 \sqrt{f_2}}$$

The cable loss at a_3 at the frequency f_3 is then:

$$a_3 = m\sqrt{f_3} + nf_3$$

Measurement accuracy

Measurement accuracy depends mainly on the accuracy of the R dial readings for large resistance at the operating frequency. Because your noise bridge may not be accurate enough, I recommend measuring the actual value of the noise bridge variable resistor after balancing with a DVM. Because there are several measuring frequencies available for a given cable, I recommend using the highest possible frequencies because the R_{max} value is lower and therefore easier to measure with the noise bridge.

Calculations and measurements indicate that a 5-ohm error in the noise bridge reading for an 18-meter long RG-213 cable at a frequency of 30 MHz causes an 11 percent error in the loss calculation. However, the resistance error is generally much smaller than 5 ohms if you use a digital multimeter to measure the actual noise bridge resistance.

If you add the parallel resistor with a T-con-

nector, the added parasitic capacitance has a negligible effect on R_{max} and the frequency *if* the reactance dial zero setting is determined with the T-connector in place. Inspection of **Equation 1** shows that an error in the assumed characteristic impedance of 10 percent results in an approximately equal error in a.

Summary

This method of measuring the cable loss is based on the fact that the maximum value of the input resistance is directly dependent on the cable loss and can be used to calculate the latter accurately. Measurements can be made at multiples of the half-wavelength frequency when the cable is unterminated, or at odd multiples of the quarter-wavelength frequency if the cable output is shorted. Both formulas give the same results. The method is simple, delivers excellent results, and doesn't require any instruments on the antenna side of the cable. Measurements can be made from the shack! No impedance matching or terminations are required and only a noise bridge is needed. Even noise bridges with questionable accuracy can give good results if you measure the noise bridge resistance with a DVM after nulling, and if you determine the zero-reactance point with the parallel resistor before the test. Thus, calibration errors of the reactance dial are eliminated.

For very short cables, the accuracy suffers because the noise bridge readings are less reliable when the cable input resistance is very high. In this case, the method described in **Reference 1** is most useful. In general, R_{max} varies between a few hundred and several thousand ohms.

Practical cable tests, together with the interpolation method, produced excellent agreement with measurements using power meters on both cable ends. The new method is also a very useful application of the noise bridge.

REFERENCES
1. A.E. Popodi, AA3K/OE2APM, "A Novel Method for Measuring Cable Attenuation, *Ham Radio*, May 1989, pages 65–73.
2. Forrest Gehrke, K2BT, "Real Coax, Impedance, and Phase Relationships," *Ham Radio*, April 1987, pages 8–14.

TECH NOTES *(from page 8)*

of electrolyte may be present. Sea water in the sea is a particularly good electrolyte.

The HF radio referred to on the boat is a marine HF radio. It is not an amateur radio, as I don't hold the appropriate license. The radio is used for normal marine communications particularly during long ocean races and return voyages. For the interest of amateurs, the antenna system consists of a ground connection to the ocean, a manual tuner, and an antenna which consists of the back stay which holds the mast up. The length of antenna from the tuner to the top insulator is about 14 or 15 meters, with a further meter or so from the tuner down to the ground connection. The backstay acts as a random radiator and is tuned by the manual tuner. The system is useful up to about 6 to 8 MHz. Above about 6 MHz, the capacitance of the top insulator begins to look like a short circuit to the radio signal, and the system loses efficiency. For 8 MHz and above, I normally use a supplementary whip antenna in amongst the forest of other (amateur VHF and UHF) antennas on the stern rail.

I hope this brief article may be of use to someone, and hopefully it will assist others in the removal of similar problems and save some wastage in materials which would otherwise be destroyed by electrolysis. ∎

PRODUCT INFORMATION

New MULTIMEC Switches

SoLiCo/mec introduces a new 1K switch option available in both illuminated and non-illuminated versions.

The square cap measures 14.3 mm and may be complimented by the addition of the 2k bezel which snaps into the designer's front panel. Depending on the position of the pc board behind the front panel, the 1k cap can be extended, flush-mounted, or be recessed within the 2k bezel.

The illuminated 1K option comes complete with high intensity LED illumination, and offers the designer the option of using interchangeable legends or standard printed graphics. The MULTIMEC switch module is totally sealed to IP-67M standards and has a mechanical life rating to 10,000,000 cycles minimum and an electrical rating of 50 mA at 24Vdc. Standard operating temperature is –40°C to +65°C. A high temperature option is available with an operating temperature of –40°C to +160°C.

For more information and/or to receive a MULTIMEC Engineering catalog, contact SoLiCo/mec, 75 Locust Street, P.O. Box 14756, Hartford, CT 06114; or call 1-800-275-7089, Fax 1-800-275-7092.

ADVERTISER'S INDEX

We'd like to see your company listed here too. Contact Arnie Sposato, N2IQO, at (516)681-2922 or FAX at (516)681-2926 to work out an advertising program to suit your needs.

ASTRON CORPORATION

9 Autry
Irvine, CA 92718
(714) 458-7277 • FAX (714) 458-0826

ASTRON POWER SUPPLIES
• HEAVY DUTY • HIGH QUALITY • RUGGED • RELIABLE •

SPECIAL FEATURES
- SOLID STATE ELECTRONICALLY REGULATED
- FOLD-BACK CURRENT LIMITING Protects Power Supply from excessive current & continuous shorted output
- CROWBAR OVER VOLTAGE PROTECTION on all Models except RS-3A, RS-4A, RS-5A, RS-4L, RS-5L
- MAINTAIN REGULATION & LOW RIPPLE at low line input Voltage
- HEAVY DUTY HEAT SINK • CHASSIS MOUNT FUSE
- THREE CONDUCTOR POWER CORD except for RS-3A
- ONE YEAR WARRANTY • MADE IN U.S.A.

PERFORMANCE SPECIFICATIONS
- INPUT VOLTAGE: 105-125 VAC
- OUTPUT VOLTAGE: 13.8 VDC ± 0.05 volts (Internally Adjustable: 11-15 VDC)
- RIPPLE Less than 5mv peak to peak (full load & low line)
- All units available in 220 VAC input voltage (except for SL-11A)

MODEL VS-50M

SL SERIES
• LOW PROFILE POWER SUPPLY

MODEL	Colors Gray	Black	Continuous Duty (Amps)	ICS* (Amps)	Size (IN) H × W × D	Shipping Wt. (lbs.)
SL-11A	•		7	11	2⅝ × 7⅝ × 9¾	12
SL-11R	•	•	7	11	2⅝ × 7 × 9¾	12
SL-11S	•	•	7	11	2⅝ × 7⅝ × 9¾	12
SL-11R-RA		•	7	11	4¾ × 7 × 9¾	13

RS-L SERIES
• POWER SUPPLIES WITH BUILT IN CIGARETTE LIGHTER RECEPTACLE

MODEL	Continuous Duty (Amps)	ICS* (Amps)	Size (IN) H × W × D	Shipping Wt. (lbs.)
RS-4L	3	4	3½ × 6⅛ × 7¼	6
RS-5L	4	5	3½ × 6⅛ × 7¼	7

RM SERIES
• 19" RACK MOUNT POWER SUPPLIES

MODEL	Continuous Duty (Amps)	ICS* (Amps)	Size (IN) H × W × D	Shipping Wt. (lbs.)
RM-12A	9	12	5¼ × 19 × 8¼	16
RM-35A	25	35	5¼ × 19 × 12½	38
RM-50A	37	50	5¼ × 19 × 12½	50
RM-60A	50	55	7 × 19 × 12½	60
• Separate Volt and Amp Meters				
RM-12M	9	12	5¼ × 19 × 8¼	16
RM-35M	25	35	5¼ × 19 × 12½	38
RM-50M	37	50	5¼ × 19 × 12½	50
RM-60M	50	55	7 × 19 × 12½	60

MODEL RM-35M

RS-A SERIES

MODEL	Colors Gray	Black	Continuous Duty (Amps)	ICS* (Amps)	Size (IN) H × W × D	Shipping Wt. (lbs.)
RS-3A		•	2.5	3	3 × 4¾ × 5¾	4
RS-4A	•	•	3	4	3¾ × 6½ × 9	5
RS-5A		•	4	5	3½ × 6⅛ × 7¼	7
RS-7A	•	•	5	7	3¾ × 6½ × 9	9
RS-7B	•	•	5	7	4 × 7½ × 10¾	10
RS-10A	•	•	7.5	10	4 × 7½ × 10¾	11
RS-12A	•	•	9	12	4½ × 8 × 9	13
RS-12B		•	9	12	4 × 7½ × 10¾	13
RS-20A	•	•	16	20	5 × 9 × 10½	18
RS-35A	•	•	25	35	5 × 11 × 11	27
RS-50A	•		37	50	6 × 13¾ × 11	46
RS-70A	•		57	70	6 × 13¾ × 12⅛	48

MODEL RS-7A

RS-M SERIES

MODEL	Continuous Duty (Amps)	ICS* (Amps)	Size (IN) H × W × D	Shipping Wt. (lbs.)
• Switchable volt and Amp meter				
RS-12M	9	12	4½ × 8 × 9	13
• Separate volt and Amp meters				
RS-20M	16	20	5 × 9 × 10½	18
RS-35M	25	35	5 × 11 × 11	27
RS-50M	37	50	6 × 13¾ × 11	46
RS-70M	57	70	6 × 13¾ × 12⅛	48

MODEL RS-35M

VS-M AND VRM-M SERIES
• Separate Volt and Amp Meters • Output Voltage adjustable from 2-15 volts • Current limit adjustable from 1.5 amps to Full Load

MODEL	Continuous Duty (Amps) @13.8VDC	@10VDC	@5VDC	ICS* (Amps) @13.8V	Size (IN) H × W × D	Shipping Wt. (lbs.)
VS-12M	9	5	2	12	4½ × 8 × 9	13
VS-20M	16	9	4	20	5 × 9 × 10½	20
VS-35M	25	15	7	35	5 × 11 × 11	29
VS-50M	37	22	10	50	6 × 13¾ × 11	46
• Variable rack mount power supplies						
VRM-35M	25	15	7	35	5¼ × 19 × 12½	38
VRM-50M	37	22	10	50	5¼ × 19 × 12½	50

MODEL VS-35M

RS-S SERIES
• Built in speaker

MODEL	Colors Gray	Black	Continuous Duty (Amps)	ICS* Amps	Size (IN) H × W × D	Shipping Wt. (lbs.)
RS-7S	•	•	5	7	4 × 7½ × 10¾	10
RS-10S	•	•	7.5	10	4 × 7½ × 10¾	12
RS-12S	•	•	9	12	4½ × 8 × 9	13
RS-20S	•	•	16	20	5 × 9 × 10½	18
SL-11S	•	•	7	11	2¾ × 7⅝ × 9¾	12

MODEL RS-12S

*ICS—Intermittent Communication Service (50% Duty Cycle 5min. on 5 min. off)