C&E HOBBY HANDBOOKS

Electronics Theory Handbook

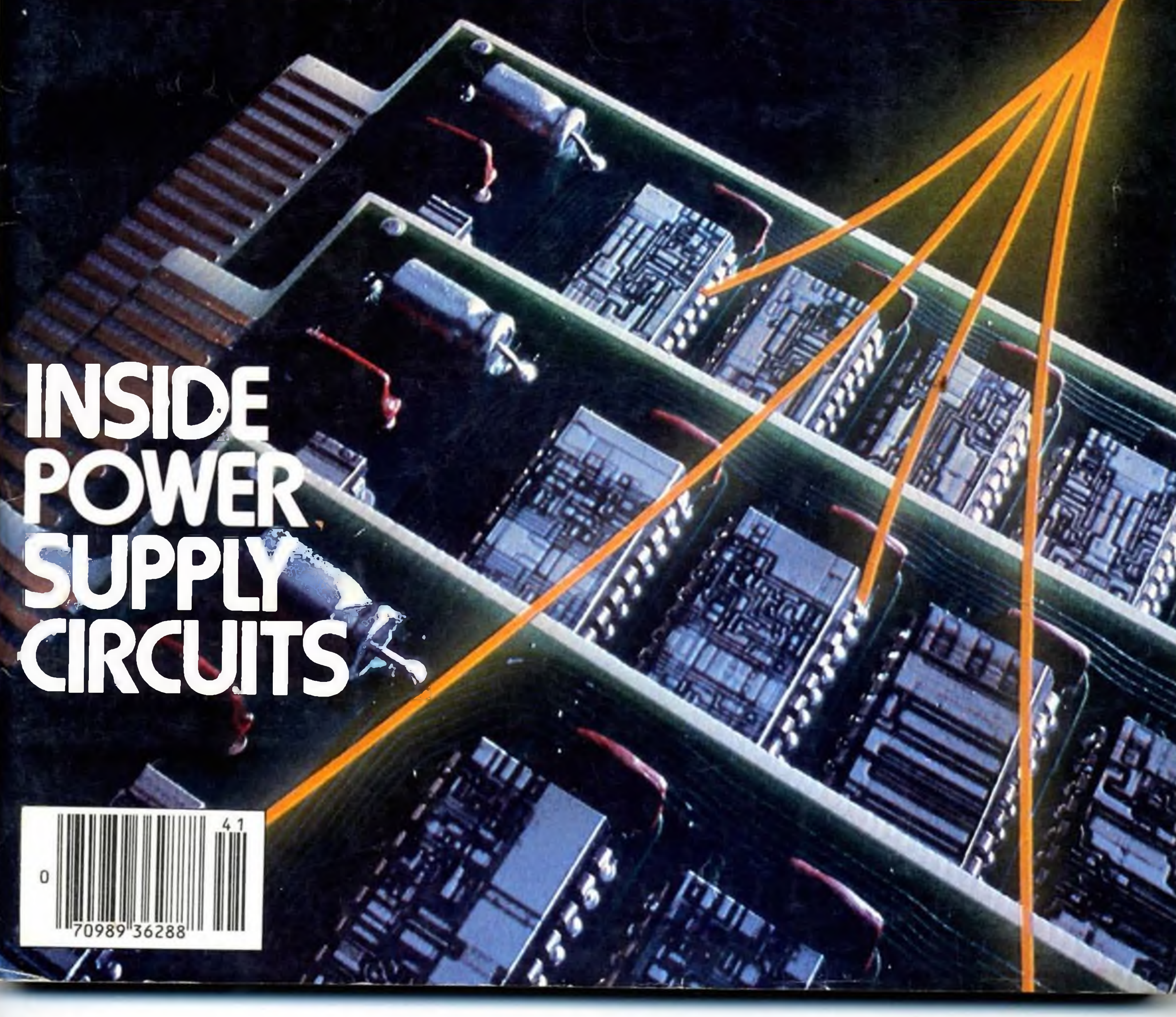
36288

SPRING 1984 $2.50

BEGINNING WITH BASICS FROM THE ELECTRON AND THE BYTE

RADIO—FROM AM TO SSB AND FM • THE ELECTRONICS OF HALLEY'S COMET

LOOKING INSIDE IC OP AMPS • ELECTRONICS OF COMPUTERS



INSIDE POWER SUPPLY CIRCUITS

Classified MARKET PLACE

C & E HOBBY HANDBOOKS — CLASSIFIED AD REQUIREMENTS

Minimum ad 15 words — Please remember to be sure to count your name and address — Count each Abbreviation (B.J. Jones = 3 words), Initial (W. Main Street = 3 words), Single Figure (P.O. Box 3333 = 3 words) or Groups of Figures (10 x 12, 35mm, F.O.B., etc = 1 word).
Your city and state are one word each, there is no charge for your ZIP. Code numbers for keying count as one word.
There will be regular headings and when no heading is requested, our judgement will decide. If you desire a special heading, there will be a $10.00 service charge.

Ads are set in uniform style with the first word in CAPS — Extra CAPS are 40¢ per word.
Proofs are not submitted on classified ads — correct insertion is guaranteed. There is no deduction for Key Number error due to illegibly written copy.
We reserve the right to refuse objectionable copy and revise copy to confrom to our standards.
We also reverse the right to request literature and merchandise from the advertiser before accepting the ad.
Advertisers using a P.O. Box must furnish complete name and address for our records.

NO CHARGE FOR ZIP CODE...NO CHARGE FOR CAPITALIZING FIRST WORD...ADDITIONAL WORDS IN CAPS — 40¢

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30

MAIL TO: C & E HOBBY HANDBOOK — CLASSIFIED,
300 W 43 Street, NY. N.Y. 10036

Your Name ____________________
Firm Address ____________________
Street Address ____________________
City ____________________ State ______ ZIP ______

Advertisers using a Post Office Box number must furnish a complete and and street address for our records.

Date Of Order ____________
Please enter the word ad, which includes ________ capatilized words under th heading:
____
No. of words rate per word ______ equals remittance.
Remittance of $__________ is Enclosed.

NO CASH DISCOUNT ALLOWED. NO CONTRACT RATE.

C&E HOBBY HANDBOOKS

300 WEST 43rd STREET NEW YORK, N.Y. 10036 (212) 397-5200 (201) 231-1518

ELECTRONIC BARGAINS. Closeouts, Surplus, Video, Audio, Cable TV, Computers, telephone, Security, Intercom, TV Parts, Books, Kits. Incredible values for hobby, industry & schools! Fascinating items unavailable in stores or catalogs anywhere. New 112 page catalog. Rush $1. ETCO, Dept 666, Box 840, Champlain, N.Y. 12919

GAINT SCREEN OSCILLOSCOPE Covered from ANY size TV. Technicians, Hams, Broadcasters, Hobbyists, Minor inexpensive changes convert TV to Super-Sensitive Testing Device. No experience needed. Follow clear-cut plans. Fully illustrated. Perfect for Workshop, Lab, School & etc. ANY size screen _ ANY size TV. Complete plans... $5. (plus $1. postage & handling)

MicroTech Dept. OP-07 Box 8561 Waco, Texas 76714

GAINT SCREEN OSCILLOSCOPE Covered from ANY size TV. Technicians, Hams, Broadcasters, Hobbyists, Minor inexpensive changes convert TV to Super-Sensitive Testing Device. No experience needed. Follow clear-cut plans. Fully illustrated. Perfect for Workshop, Lab, School & etc. ANY size screen _ ANY size TV. Complete plans _ $5. (plus $1. postage & handling)

MicroTech Dept. OP-07 Box 8561 Waco, Texas 76714

MAILING LIST, now being compiled. Simply send $1.00 cash along with your name, address, and product preference. This information will be made available to mailing list companys. Donmar Company, P.O. Box 288, Hartsmill Road, Atwood, Tennessee 38220 Attn. Mr. Don Newbill.

For as little as $500 per year we can supply your Company with a New York City—5th Ave.—address Contact: Suite 3280, 475 Fifth Ave., N.Y. 10010

Faithful, traditional, loving Oriental Ladies. Marriage only! Cherry Blossoms, Box 1021 CE, Honokaa, Hawaii 96727

GAINT SCREEN OSCILLOSCOPE Covered from ANY size TV. Technicians, Hams, Broadcasters, Hobbyists, Minor inexpensive changes convert TV to Super-Sensitive Testing Device. No experience needed. Follow clear-cut plans. Fully illustrated. Perfect for Workshop, Lab, School & etc. ANY size screen _ ANY size TV. Complete plans _ $5. (plus $1. postage & handling)

MicroTech Dept. OP-07 Box 8561 Waco, Texas 76714

EPROM PROGRAMMER KIT
Supports 2758, 2716, 2732, 2732A, 2764, 27128, 2508, 2516, 2532, 2564, 68732, 68764, 8741, 8748(H), 2749(H). On board power supply, RS232. Xon/Xoff, DTR/CTS. Kit contains P.C. Board, programmed rom & documentation. 8085 based, uses only common parts. Sent $49 + $2 shipping to B&C Microsystems, 20917 Greenleaf Dr., Cupertino, CA 95014
Calif. residents add 6.5% tax.

Electronics
Theory Handbook

PUBLISHER
Don Gabree

ASSOCIATE PUBLISHER
H.M. McQueeney

CONTROLLER
Dave Mannis

SECRETARY/TREASURER
Lillian Beck

MANAGING DIRECTOR
Steve Mydanick

EDITOR
Hank Scott

ART DIRECTOR
G.R. Design Associates

CIRCULATION DIRECTOR
Bob Martin

PRODUCTION MANAGER
Peter Zelaya

ADVERTISING DIRECTOR
John J. Abbott

C&E Hobby Handbooks, Inc.
300 W 43rd Street
New York, N.Y., 10036
(212) 397-5200

Electronic Theory Handbook, Volume 7, **Number 2**, is published quarterly by C & E Hobby Handbooks, Inc. Editorial and Business offices 300 W 43rd Street, New York, N.Y., 10036, (212) 397-5200. PERMISSIONS. Material in this publication may not be duplicated or reproduced in any form without permission. Requests for permission should be directed to—Don Gabree, C & E HOBBY HANDBOOKS, INC., 300 W 43 Street, New York, N.Y., 10036 (212) 397-5200. ISSN #0270-4994.

EDITORIAL CONTRIBUTIONS must be accompanied by return postage and will be handled with reasonable care; however, the publisher assumes no responsibility for the return or safety of manuscripts, artwork or photographs. All contributions should be addressed to: EDITOR, HOBBY COMPUTER HANDBOOK, 300 W 43rd Street, New York, N.Y., 10036.

C&E Hobby Handbook, Inc., a division of J. Q. Adams Productions, Inc.
Second Class entry applied for

What's in a word?

I can remember when it would have been disasterous to use the word "theory" on the cover of a newsstand magazine. To do so would certainly reduce the potential sale, and even cause the issue to be a financial loser. About ten years ago I thought that times have come around whereby the public, in particular, electronics hobbyists, would no longer shun the word "theory." In fact, I believed the word was an asset. So, when you believe in something, you go all the way with it, and that was how the title "Electronics Theory Handbook" came about.

Today, all electronics annuals and periodicals brag about their "theory" coverage, and use it as a selling tool. Why not? TV commercials are telling us to give expensive computers to six-year olds as birthday gifts so they will do better in school. Local school boards are scraping the bottom of barrels and asking PTA's for contributions to start up computer-equipped classrooms. I guess the word "theory" is safe to use today.

So, now you are holding the latest issue of Electronics Theory Handbook in your hands, and we are proud of its contents. The articles are selected and presented with an increasing degree of background that permits readers at all levels of electronics knowledge to find something of value—value sufficient to more than cover the cost of this issue. Dig in at the beginning and thumb away at the pages until you find a theory article for you. We know that most readers begin up front in order to refresh their previously acquired knowledge.

Let us know what you think of our annual magazine. We want to hear from you so that we can adjust to your needs. Publishing may be our business, but electronics is our first love.

Don Gabree
Publisher

On the Tail of Halley's Comet

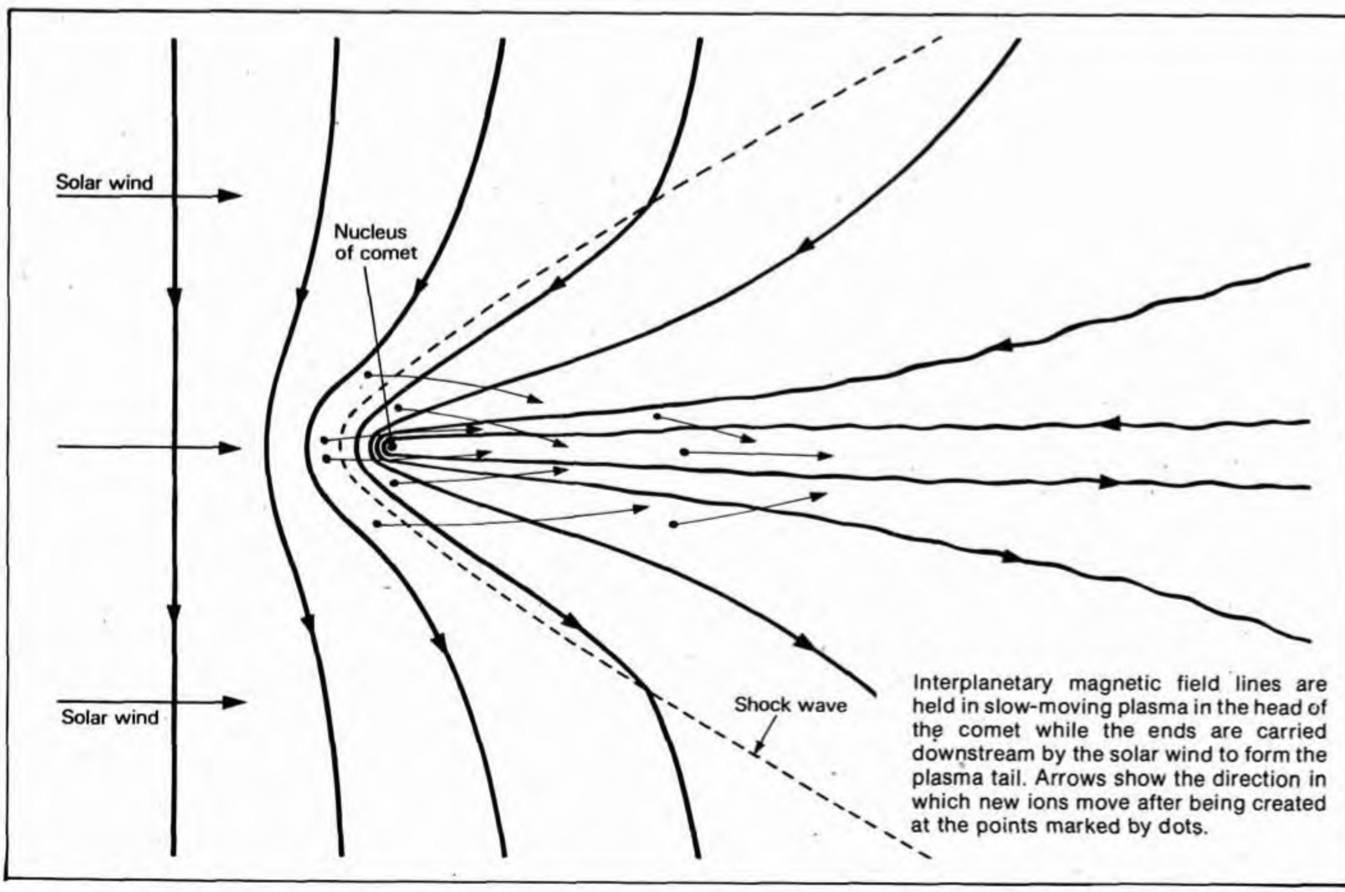
When Halley's comet returns in 1986 it will be visited by five spacecraft intent on searching for a tiny frozen nucleus at the heart of the bright gas cloud and on measuring its chemical composition to see if it is part of the material from which the solar system was formed five billion years ago.

It has recently been announced that Halley's comet has been detected for the first time in 71 years by the five-meter telescope on Mount Palomar, California. It is still outside the orbit of Saturn, but is heading towards the inner solar system along an orbit which will bring it to perihelion, its closest point to the Sun, in February 1986.

The Palomar telescope has detected a faint point of light of 24th magnitude, equivalent to an object with a cross-section of a few square kilometers. If it could resolve detail, all there would be to see now, so it is thought, would be a lump of dirty snow which has been gathered together into a rough ball two kilometers across and frozen hard. There is no coma, or flowing tail, as in the familiar picture of a comet. At the end of 1985, it will emerge from its 76-year hibernation in the deep cold of outer space and develop its coma and tail for a brief reign as queen of our skies. The warmth of the Sun's rays will bring it to life by evaporating ice from the surface.

The comet is more spectacular than a snowball melting in winter sunshine because it evaporates in the emptiness of space where the gas molecules fluoresce in sunlight, react chemically with each other and become charged and respond to the electric and magnetic fields that permeate interplanetary space. All these processes help to shape the comet into the form we know. We see the gas cloud and the dust it carries away with it; the dirty snowball at its center cannot be seen inside the bright gas cloud. Gas molecules leaving the surface react with other molecules, absorb solar photons and then emit them at their own characteristic wavelengths. The most common species identified in the cometary spectrum are the radicals OH, O, H, CN and C_2- which cannot exist in the solid state and therefore could not have been evaporated from the nucleus. They must be the daughters, produced in photochemical reactions, of unseen parent molecules released from the nucleus. The two most likely parent molecules are water and carbon dioxide. Neither has emission lines at visible or ultra-violet wavelengths in the neutral state, so they cannot be detected directly from Earth. There must also be other parent molecules to account for the presence of nitrogen and other complex compounds.

Magnetic Field. Once the molecules emerge from this chemical soup some tens of kilometers thick, there is nothing to stop them drifting away into space unless they become ionized. When they do, the loss of an electron and the acquisition of a net positive electrical charge dramatically changes their view of interplanetary space. What seemed to be a vacuum is suddenly found to contain a plasma, a gas with equal numbers of positive and negative charged particles whose behavior can be visualized by imagining that magnetic field lines are frozen into it. The field lines may be stretched or twisted, but they retain their identity and move with the



Interplanetary magnetic field lines are held in slow-moving plasma in the head of the comet while the ends are carried downstream by the solar wind to form the plasma tail. Arrows show the direction in which new ions move after being created at the points marked by dots.

plasma.

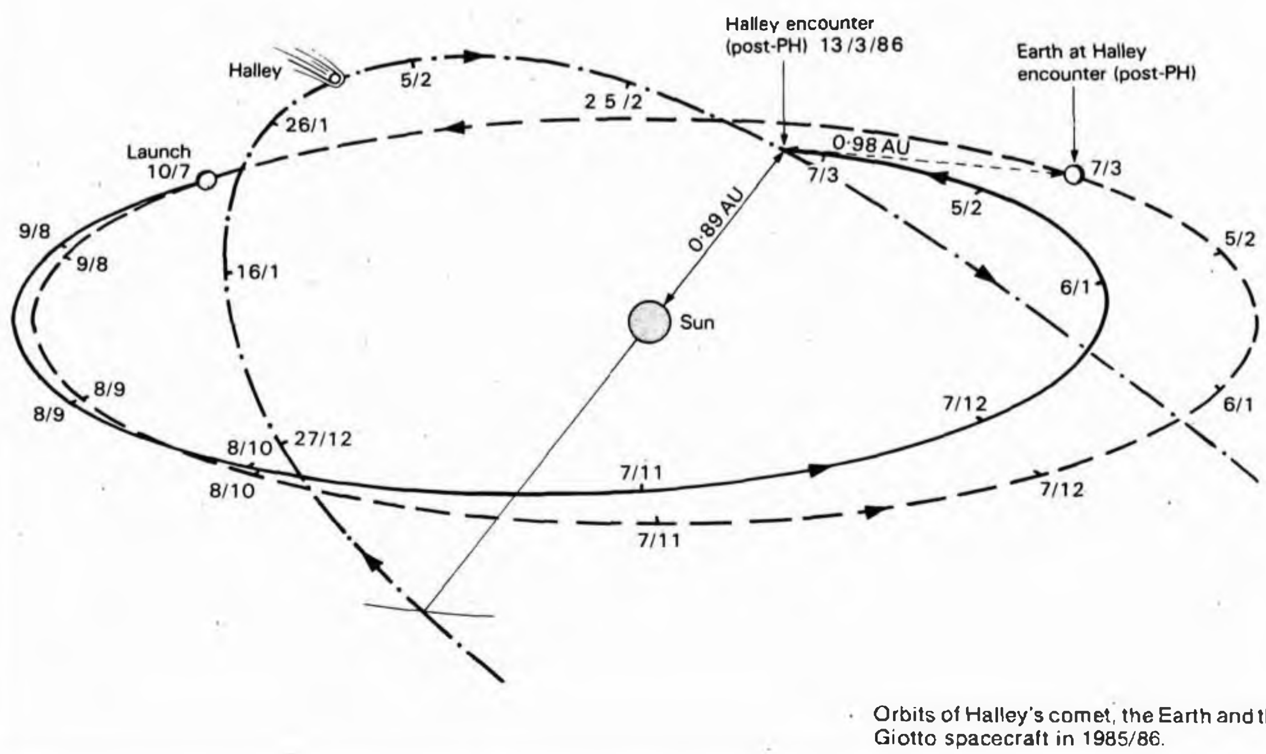
The interplanetary plasma is called the solar wind, because it blows away from the Sun at more than 400 km/s. The positive particles are protons and alpha particles, the negative ones electrons. Their density is too low to obstruct the flow of neutral molecules, but their electric and magnetic fields fill all space. Cometary molecules are ionized by exchanging charge with solar-wind protons and by sunlight. When the electric field picks up a new ion it sweeps it along with the solar wind. This additional mass in the flow has to be given momentum and energy by the solar wind, which is thereby slowed down. Now ions have an effect out of proportion to their number because they are much more massive than the solar-wind ions. The retardation is so sudden near the comet that a sonic blow shock wave is formed.

The cumulative effect of the mass-loading is greatest on the plasma in the Sun-comet line upstream from the comet, where the plasma is retarded most. Because the distant ends of field lines continue to travel at the full solar wind speed, magnetic-field lines become wrapped around the head of the comet, as shown in fig. 1; they are dragged downstream to form the filamentary plasma tail, made visible by the characteristic fluorescent emissions of those cometary ions among which H_2O^+, CO_2^+, CO^+, N_2^+ and OH^+ have been found.

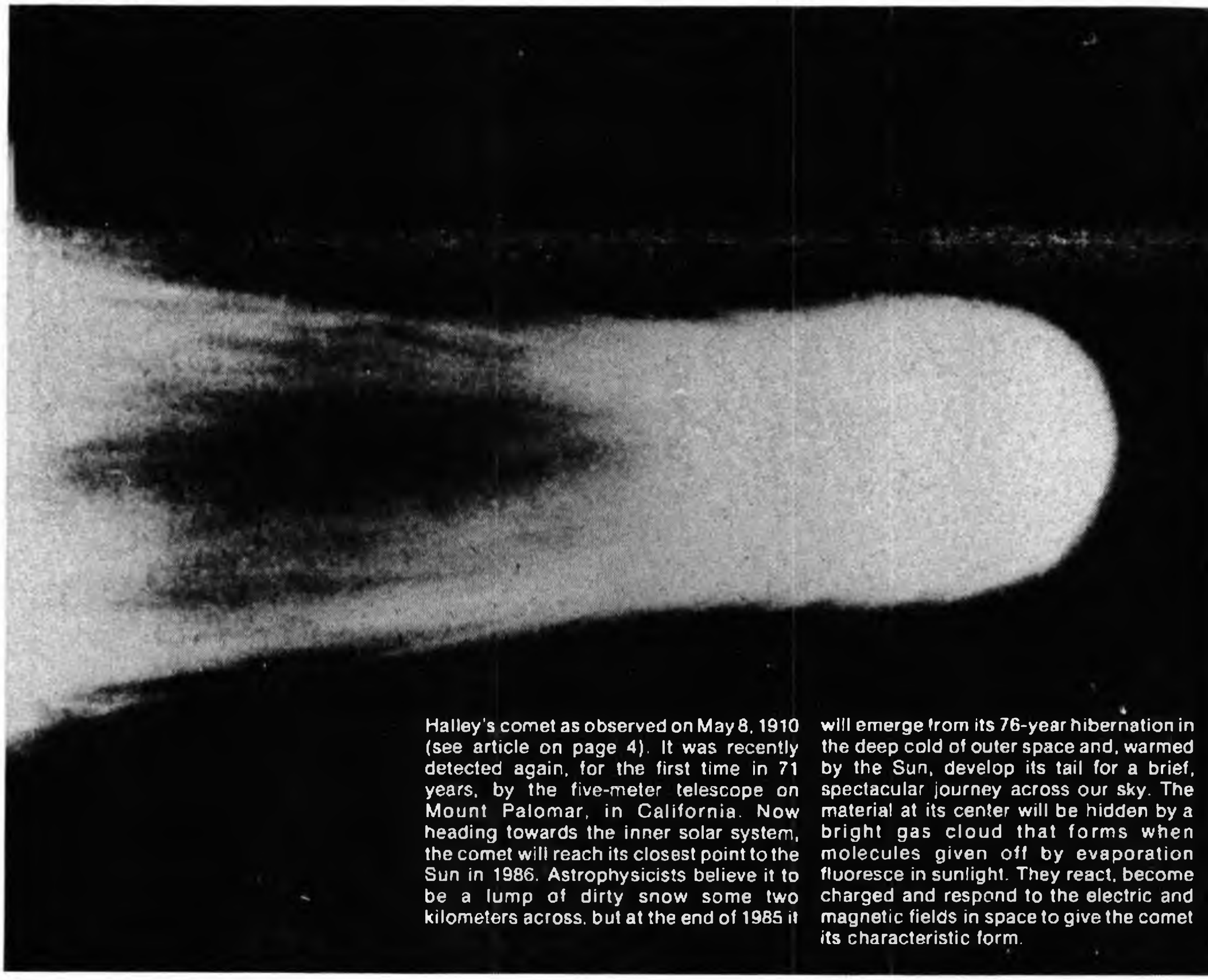
Some comets have a second tail, quite different from the first, made of dust liberated from the ice as it evaporates. Puffs of escaping gas enable the dust particles to overcome the weak gravitational attraction of the comet and escape into space. They continue to orbit as the comet does, but are pushed away from the Sun by the pressure of sunlight scattering off their surface. The dust tail is therefore directed away from the Sun, but may become curved owing to the orbital angular velocity of the particles decreasing as their distance from the Sun increases. Because the tail is illuminated by the scattered sunlight, it appears white and reveals nothing about its composition. The dust particles are believed to have a vast range of sizes from less than 10-15 grams up to, perhaps, more than one gram.

Important Clues. The origin of comets is still a mystery. Some scientists believe that there is a large cloud of comets orbiting in the outer parts of the solar system, tens of thousands of astronomical units (1 AU is the distance from Earth to Sun) from the Sun. Occasionally one is deflected into the inner solar system by the gravitational perturbation of nearby stars. The material of comets could have come recently, in terms of astronomical time, from interstellar space or be part of the material from which the solar system was formed five billion years ago. If the latter, the gas and dust released by. solar heating is especially interesting because it has been kept in a pristine state ever since it was formed, frozen hard inside a comet. Whether this is eventually found to be true or not, the chemical composition of a comet must provide important clues about its origin and about the matter in the outer reaches of the solar system.

The only way to find out if our model of the comet's structure is correct, and to identify its chemical composition, it to make scientific observations inside the coma. When Halley's comet next appears it will provide an excellent opportunity to do so, because it is by far the largest comet (measured by the rate at which it produces gas) with an accurately predictable orbit, making it possible to plan a space mission in time to



Orbits of Halley's comet, the Earth and the Giotto spacecraft in 1985/86.



Halley's comet as observed on May 8, 1910 (see article on page 4). It was recently detected again, for the first time in 71 years, by the five-meter telescope on Mount Palomar, in California. Now heading towards the inner solar system, the comet will reach its closest point to the Sun in 1986. Astrophysicists believe it to be a lump of dirty snow some two kilometers across, but at the end of 1985 it will emerge from its 76-year hibernation in the deep cold of outer space and, warmed by the Sun, develop its tail for a brief, spectacular journey across our sky. The material at its center will be hidden by a bright gas cloud that forms when molecules given off by evaporation fluoresce in sunlight. They react, become charged and respond to the electric and magnetic fields in space to give the comet its characteristic form.

meet the comet. Three agencies are now building the five spacecraft which will make the journey in 1986, namely the European Space Agency (one), and the space authorities of the USSR (two) and Japan (two). The European Space Agency has eleven member states. Its spacecraft, named Giotto after an Italian painter who included an accurate representation of Halley's comet in a painting of the Nativity of Christ, is being built by a consortium of industrial companies led by British Aerospace.

Meeting Point. Sending a spacecraft to the comet poses special problems. The orbits of the Earth and the comet are shown on page 6: the spacecraft must be placed in a third orbit which starts at the Earth and intersects the comet's orbit at the right time, using the least amount of rocket fuel so that the spacecraft can carry the largest possible collection of instruments. This requirement restricts the spacecraft to orbits lying in the ecliptic plane. The orbital plane of the comet is inclined at 165 degrees to the ecliptic, so the craft can meet the comet only at one of the two crossing points.

All the missions aim to meet the comet at the crossing after perihelion when it will be more active. Giotto must be launched with in a short period of two weeks at the beginning of July 1985 if it is to reach the comet on time. The next problem is to find the tiny, unseen nucleus—if it exists. We can only assume that it is near the center of the bright coma, which is 100,000 km across, has fuzzy edges and is not always symmetrical. This makles aiming difficult and limits the accuracy with which the comet orbit can be tracked and predicted, for it is the nucleus that is subject to the gravitational forces, not the coma.

The mission controllers will take advantage of all the observations of the comet being made from the Earth and from the four spacecraft approaching the comet, but the probable error is estimated to be greater than 500 km.

As it appraoches the nucleus, the spacecraft will run into another big problem. The relative velocity of the two bodies is extremely high, because they orbit the Sun in almost opposite directions, so tiny particles of dust drifting slowly away from the comet appear to the spacecraft to be moving directly towards it at the colossal speed of 79 km/s. A flake of dust weighing no more than a couple of postage stamps could hit the spacecraft with the same force as a bullet fired from a high-speed rifle.

Arranging to protect against such particles, which could rip out the heart of the spacecraft, is simplified by the fact that they will arrive parallel to the relative velocity vector. The spacecraft will be oriented with its spin axis pointing the same way and will carry a dust shield on the end facing the comet. Particles up to 10^{-7}

Comet Mrkos photographed in 1957. The picture shows the filamentary plasma tail and the broad featureless curve of the comet's dust tail.

gram should make a crater in the aluminum of the front shield. Larger ones may be expected to penetrate its 1 mm thickness, becoming heated to temperatures of tens of thousands of degrees and thereby vaporized by dissipation of their kinetic energy. The cloud of hot vapor spreads out and hits the rear shield, 12 mm thick, made of several layers of plastic and placed 24 cm behind the front shield.

In combination, the two shields provide more protection than a single shield of the same weight and should be able to withstand the impact of a particle with a mass of up to 100 milligrams before the rear shield is penetrated and the instruments behind it are seriously damaged. Such a blow, if not taken on the spin axis, will also tilt the spacecraft causing the Earth to lose contact with the narrow radio beam bringing a-1 the scientific information back. By the time communications could be established again the encounter would be over.

The probability of such an impact is estimated, for spacecraft passing the nucleus at a distance of 500 km, to be 90 per cent but it decreases sharply as the encounter distance increases. Estimates are based on observations of the dust tail but they are so uncertain, particularly for the large masses, that the real risk could be quite different. The mission is planned on the assumption that it will be brought to an end by such an event.

International Groups. Each of the ten scientific instruments is the responsibility of a principal investigator from a European research institution leading an international group of scientists. Two of the principal investigators are British; the others are from Federal Germany, France, Switzerland and Ireland. One of the instruments is a camera with a difficult task of taking a picture of a very rapidly-moving object from a rotating platform. It takes one picture in every rotation and selects just the portion containing the nucleus to transmit back to Earth. Only in the last few pictrues will the nucleus appear large enough for details of its surface to be resolved.

Chemical composition of the comet will be measured by a group of three different mass spectrometers. The first looks at neutral molecules and radicals and the second at positively-charged molecules and radicals. Both instruments use magnetic and electric field analysis to separate the different types of particle. The third spectrometer analyses the composition of dust particles which vaporize when they strike a gold target. To find out the mass distribution of the constituent atoms and molecules, their time of arrival at a detector after being accelerated through a constant electric field will be recorded. By comparing the measured distribution with the atomic abundance in known minerals, the mineral type of the particle can be identified.

Two instruments are designed to observe the dust tail: one will measure the amount of light scattered by dust particles in various wavelength bands; the other, from the University of Kent (see Spectrum 181) will record the size of impacts on the dust shield, its aim being to discover the size distribution of the particles.

The final group of instruments is designed to study the interaction between the solar wind and the ionized cometary particles in the plasma tail. It includes a magnetometer, analysers to find the velocity distribution of positive ions and electrons, and a detector for the high-energy charged particles. University College, London, is responsible for the positive-ion analyser.

One big scientific problem to do with all the instruments lies in making sure that their results are not affected by the spacecraft itself. For example, instruments that have to be exposed will be damaged by the impact of dust and neutral molecules, both of which produce secondary ions and electrons on the spacecraft and within instruments. Emission of such secondary particles may be expected to charge the spacecraft to a high voltage, so preventing some cometary charged particles reaching the instruments and building up a background of unwanted particles inside the instruments. Some of the material removed from the sapcecraft by impact could find its way into mass spectrometers and contaminate the cometary spectra. So, the instruments have been designed either to eliminate the known sources of contamination, or to make their effects easily recognizable. The unexpected sources must be allowed for by embodying enough cross-checks in the information from each instrument, and the collection of instruments taken together, to be able to verify the accuracy of the results.

The four spacecraft will complete another step in the dramatic exploration of the solar system which has been going on during the past ten years. Each visit to another planet, or moon, has opened ur eyes to new wonders, given us a better understanding of our own planet and provided a new per spective of its position in the cosmos. We can confidently expect that the messages sent back by the travellers to Halley's comet will do the same.

Learning While Building

By Hank Scott

There comes a time when theory has to be put to work—it's part of the learning process. What better way can you find to reinforce your newly gained knowledge than by using it in a project you assembled. It does not matter how useful the circuit may be at the time you assemble it, when matters is that your confidence is built up, and you know that larger and more difficult projects can be mastered.

To get you on the right track, Electronics Theory Handbook has assembled several interesting projects that you can build in an evening with a few parts in each project. The cost is relatively low, and the rewards weigh in your favor.

You'll need a few tools and materials besides the parts listed in the parts list for each project. Naturally, a good pair of long-nose pliers with side cutters are a must along with a few other hand tools, plus a light, low-wattage (about 25-35 watts) soldering iron, some solder and solder wick. The wick is used to sop up solder should you put too much down—a common fault with beginners. A perfboard about 3-in. by 5-in. is sufficient to mount most project parts except those with bulky transformers. Pick up some flea clips to push into the perfboard holes for use as terminal points.

You may decided to go the solderless way by using solderless (naturally!) breadboards. This technique permits you to salvage 100 percent of the parts you use, ultimately reducing the costs of future projects to a fraction of their costs. You may want to assemble some power sources at first so that you could eliminate batteries. Power supplies rated at 5, 9 and 12-volts DC regulated. Later, you may want to add some test equipment and pulse devices to your testbench.

But, enough talking! Tackle these projects, have fun, and put theory to work.

Ask Hank, He Knows!

Got a question or a problem with a project—ask Hank! Please remember that Hank's column is limited to answering specific electronic project questions that you send to him. Personal replies cannot be made. Sorry, he isn't offering a circuit design service. Write to:

Hank Scott, Editor
C & E HOBBY HANDBOOKS INC.
300 West 43rd Street
New York, N.Y. 10036

Likes Small Circuits

Hank, I like to build small gadgets and circuits just for the fun of building them. True, I do this to learn all about electronics, which is my hobby, but fun is fun. Where can I find a source of *mucho* small circuits?

—J.K., Brooklyn, MD

First, I did not know there was a Brooklyn in Maryland. So you see how much you can learn by reading this magazine. As for the source you seek, try the magazine you're reading right now *(I sent him a copy—Hank)*, and one other magazine published on an annual schedule—101 Electronic Projects. There are a few office copies still unsold which you can pick up for $4.00 (which includes postage and handling) to Worldwide Publications Inc., P.O. Box #5206, North Branch, N.J. 08876.

What the L

What is the difference between a 74LS00 IC and 7400 IC?

—T.T., Cando, ND

I just checked some mail order houses' catalogs, the difference is about 20 cents. The L's in the type number indicate that the device is a low-power Schottky device that's easy on the battery supply. Otherwise, both units are identical for most hobbyist applications. I use the Schottky devices because most of my project end up battery driven.

Mr. Clean Head

Hank, how often should I clean my cassette tape heads?

—E.R., Juneau, AK

A dirty tape head wears quicker than a clean one. I use a cassette head cleaner that makes a 10-second pass. The maker suggests cleaning after 10 hours of equipment use. I prefer 3-4 hours. Periodically I use a liquid cleaner on the heads, guides, rollers, etc. I blow dust out with a freon blower used by photographers. How long is periodically? Well, I clean the heads before each recording session or once a month, whichever comes first. To keep the heads clean, keep tape stored in the plastic boxes in which they are bought. Keep the tape door closed at all times, and dust the room and area near the cassette deck often.

Count Your Fingers

How many characters does or can the ASCII code have? I say 127. My friend says 128. Who's right?

—E.A., Salem, MO

Your friend is right. The decimal numbers for the ASCII code ends at 127, but starts at zero. So the correct amount is 127 plus 1, or 128. Remember, in digital work, you have 9 fingers. Count them starting from zero. ■

Broke a Leg

Hank, how do you solder a broken contact back onto an integrated circuit chip? I blew it on an expensive chip and don't want to go through the expense of buying a new one.

—A.N., Greenville, SC

How? Very carefully. In fact, more carefully than the technique you used to break the leg of the chip. I've tried on several occasions to do just that by using a very low wattage narrow-tip soldering iron. My results on the "fixables" were 33⅓% successful. I saved one chip and destroyed two others. Several chips were beyond saving because not enough of the leg protruded from the insulating body of the chip. If you are salvaging chips from surplus circuit board assembles as I was, take a tip from one of my readers. His idea for "one of a kind" and very expensive chips, was to cut the board using a table jigsaw all around the chip about ⅛-in. space between the chip and cut. Attach wire leads to the cutout with a low-wattage iron, then wire wrap these leads to a standoff insulator or header that is sized to insert into an IC socket. My big fingers found the job very difficult, but it worked. There must be better ways to do it, so please write and advise me. Finally, no matter if the chip is new or old, cheap or costly, always use care when inserting an IC into a socket. I use a magnifying glass to check that all the legs on the chip are seated in the IC socket before I press down. Also, an IC insertion tool *(and IC puller)* are invaluable during the hand jobs.

Soups On TV

I would like to soup up my TV to be more sensitive to distant stations as I am in a fringe area. What would be the best way to go about this? To put higher gain transitors in the IF section or to try to get higher gain in the video amplifier section? I have tried boosters to no avail, also other antennas. My TV is in good shape, just not good enough for fringe areas. My brother who lives next door gets much better reception with an older TV *(1970)* compared to my newer set *(1973)*.

—M.A., Julian, CA

Don't soup up your TV, soup up the TV signal. Install your antenna as high as you can; you may need a tower. Install a MATV (Multiple Antenna TV) system in your house with the signal booster as close to the antenna as possible. Use coax cables throughout, not

300-ohm ribbon lead-in wire. Matching transformers at the antenna and each TV set in the house is important. Finally, you may need a rotator if the TV signals come from different points on the compass.

If you are concerned about the quality or tune-up of your TV receiver, carry it over to your brothers house and see if it responds better to this TV antenna hook-up than yours. Chances are it's the antenna system or signal strenght at your brother's house that is better than your home.

7447-7448?

I see no difference between the function of the 7447 and 7448 chips. Do you agree, Hank? And, please, let's not split hairs.

—K.M., Tampa, FL

The 7447 is a BCD-to-Seven-Segment Decoder/Driver as is the 7448, except the 7447 drives a LED common-anode display like the Opcoa SLA-1, and the 7448 drives a LED common-cathode display very much like the Litronix 704. If you think they are still exacly alike, then look at their outputs for a zero decimal digit display: the 7447 delivers 0000001 and the 7448 delivers 1111110—the outputs are complementary. Admit it—there's a difference.

LED—D is for Diode

Besides giving off light, is there any difference between a light emitting diode *(LED)* and a normal diode? If so, could you tell me what the difference is?

—P.M., West Orange, NJ

All that I'd admit to is—a LED is also a diode. Unlike a silicon diode that fires at 0.7 volts, the LED fires at 1.7 volts. The average required current flow *(this is typical with most—not all)* is 20 milliamperes. I'm sure the LED is too noisy for use as a signal diode, too inefficient for rectification, or just not intended for diode use! Best bet is to forget about this question and the answer—use components designed for the job. you want done.

No Con Game

Electric power may be expensive, but the price of solar cells is out of sight. Level with me, Hank, solar cells are a bit rip off?

—W.N., Flint, MI

I recently saw a fan assembly powered by solar cells. A friend installed it in a shed, behind his home. He didn't want to string a power line because there were to many trees in the way. This fan assembly kept the heat in the shed, on sunny summers days, down to 90 degrees, when it would be over 120 degrees without the *free* fan power. Try to tell him solar cells are a rip off! You see, value is measured by the purpose, installation and person. To some, it's a godsend.

Meters Lie

On my hi-fi receiver, when turning FM, the signal strength is not at maximum position when the tuning indicator is centered. For maximum signal, the tuning indicator is about 1/32-in. to the left of zero. What should I do?

—W.D., Chatsworth, CA

You didn't say a thing about how the unit sounds! Tune for minimum distortion. I think you'll discover this occurs when the tuning meter is centered. If the sound is poor, then alignment is necessary. I know this should not happen, but it is fairly common.

Racing With the Amp

I see a lot of raceways in buildings that add wiring on the internal wall. Is this a good practice?

—A.M., Silver Springs, MD

You bet it is? It may not look good in your living room, but in the office, garage, shop, and maybe the kitchen, it's alright. The raceway is a metal conduit that is secured to the wall. It is an inexpensive way to add outlets without breaking up walls and floors. Raceways can also be used to cheaply interconnect air conditioners to the main fuse panel.

Keep Cool

Some of the larger chips on my main computer board run very warm. Can I add a heat sink to them somehow?

—W.N., Menlo Park, CA

One thing I have learned about computers is don't touch a main board, or any internal board unless you absolutely have to do so! So, keep your fingers off the chips. Also, forget about adding heat sinks, they are not required. If you are a worrier like I am, add a cooling fan if none exists. As long as a cool breeze passes over the board, no additional cooling is required. If the room temperature goes above 80 degrees, maybe you should consider air conditioning in the summer and turning the heat down in the winter.

From The Electron To Basic AC

One of the most thought-provoking discoveries of modern physics is the fact that matter and energy are interchangeable. Centuries of scientific headscratching about the nature of matter, the mystery of fire, and the once-terrifying crack of lightning have all come to focus on the smallest particle that is the building block of any given substance: the atom. An atom is necessarily matter and yet this atom of matter can undergo nuclear fission and release quantities of energy that are beyond the imagination. In the atom lies the secret of all phenomena.

By the beginning of the 19th century, the atomic theory of matter—which actually originated in 5th century Greece when the atom was named—was firmly established. It was due primarily to the efforts of 17th century scientists who—actually working in the tradition of medieval alchemy—sought the prime constituent of all matter. Mainly through the work of John Dalton, whose investigations as to how various elements combine to form chemical compounds, it came to be regarded that *an atom* was the indivisible and indestructible unit of matter.

This viable and working view of the indestructible atom served science until 1897 when the atom itself was found to be destructible! To anyone concerned with electricity or electronics, the year 1897 is a memorable one: it was the year J. J. Thomson, the English physicist, identified and experimentally revealed the existence of the first subatomic particle—the *electron!*

The First "Elecronic" Experiment. We blithely speak of electricity as the flow of electrons yet, often, we are little aware of the great body of research that went into elucidating this fundamental of basic electricity. In fact, befor the discovery of the electron, convention held that the flow of electric current was in the direction that a positive charge moved. This convention of *positive current*, being the flow of positive charges and opposite to the direction of electron flow, is still found to be useful in circuit analysis and is used even today.

Thomson's experiment established that a particle much lighter than the lightest atom did indeed exist. The electron, as it was named, was the first subatomic particle to be defined.

The experiment was conducted utilizing a rudimentary

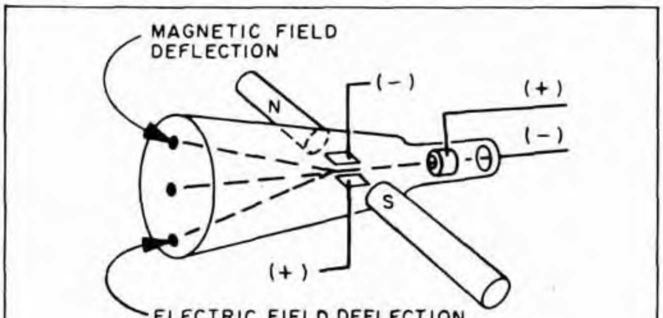


Fig. 1. Electron beam, like that in a TV picture tube (CRT), can be deflected magnetically or by an electric field. Force needed "measured" the electron.

version of a cathode ray tube—the modern version of which is in almost every home today in the form of the television picture tube. Before Thomson's experiment, it was discovered that when electric current was passed through a gas in a discharged tube, a beam of unknown nature traveled through the tube from the negative to positive terminal (opposite to the direction conventionally held as the direction of the flow of current).

This "cathode ray" beam also traveled in a straight line and was deflected by electric or magnetic forces applies perpendicular to the beam. What Thompson did was to use these facts to determine for one of the mysterious particles comprising the beam of cathode rays the relationship of its mass, *m*, to its electric charge, *e*. By deflecting the beam with a known electric force (Fig. 1) and then measuring what magnetic force applied in the opposite direction would bring the beam back to its original undeflected position, Thompson could determine the relationship of *e* to m. He established a definite value for *e/m* and thereby "discovered" the electron which as we now know, is 1,837 times smaller in mass than the lightest atom, the hydrogen atom. It also carries the smallest charge that occurs in nature; every electric charge is actually an integral multiple of the charge of the electron.

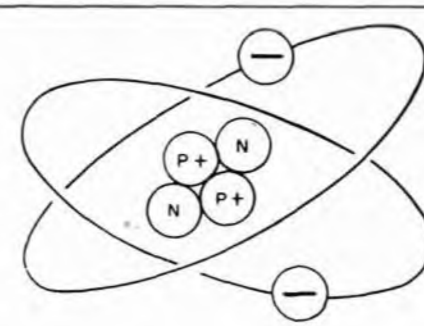


Fig. 2. Charge of each electron balances that of a proton. Other particles affect atomic mass but can be omitted in study of electronics.

From Minus to Plus. With the discovery of the electron, it was still over a dozen years into the 20th century before a graphic conception of the atom evolved. Since the atom is electrically neutral and electrons are negatively charged, the existence of positively charged particles was a necessity, and the existence of a *proton* was postulated. Eventually the nuclear model of the atom was evolved. Each atom was conceived to resemble a solar system in miniature. The nucleus—positively charged—is surrounded by a number of electrons revolving around it; the charges balance and the atom is electrically neutral (Fig. 2). Further research in the 20th century has gone on to reveal more elementary particles than you can shake a stick at: neutrons, positrons, neutrinos, mesons, quarks, and more. The number continues to grow and yet the ultimate nature of matter remains a riddle. But, in a discussion of basic electricity, only the electron and proton need concern us.

Electrons in Orbit. An atom of matter has a number of electrons orbiting around its nucleus. A hydrogen atom for example, has a single electron; carbon on the other hand has six. These electrons are arranged in rings or shells around the central nucleus—each ring having a definite maximum capacity of electrons which it can retain. For example, in the copper atom shown in Fig. 3 the maximum number of electrons that can exist in the first ring (the ring nearest the nucleus) is two. The next ring can have a maximum of eight, the third ring a maximum of 18, and the fourth ring a maximum of 32. However, the outer ring or shell of electrons for any atom cannot exceed eight electrons. However, heavier atoms may have more than four rings.

The Outer Orbit. The ring of electrons furthest from the atom's nucleus is known as the *valence ring* and the electrons orbiting in this rings are known as *valence electrons.* These valence electrons, being further from the nucleus, are not held as thightly in their orbits as electrons in the inner rings and can therefore be fairly easily dislodged by an external force such as heat, light, friction, and electrical potential. The fewer electrons in the valence ring of an atom, the less these electrons are bound to the central nucleus. As an example, the copper atom has only one electron in its valence ring. Consequently, it can be easily removed by the application of only the slightest amount of external energy. Ordinary room temperature is sufficient to dislodge large numbers of electrons from copper atoms; these electrons circulate about as free electrons. It is because of these large numbers of free electrons that copper is such a good electrical conductor. There could be no electrical or electronics industry as we know it today if it were not for the fact that electrons can fairly easily escape, or be stripped from the valence ring of certain elements we refer to as "conductors."

Electronic Charges. If an electron is stripped from an atom, the atom will assume a positive charge because the number of positively charged protons in its nucleus now exceeds the number of negatively charged orbiting electrons. If, on the other hand, the atom should gain an electron, it will become negatively charged as the number of electrons now exceeds the protons in its nucleus. The atom with the deficiency of electrons is known as a *positive ion,* while an atom with a surplus of electrons is known as a *negative ion.*

Presence of an electrical charge on a body can be illustrated by use of an electroscope (Fig. 4). Two leaves of aluminum or gold foil hang from a metal rod inside a glass case so they're free from air disturbances. When the metal rod is touched by a charged body, the leaves acquire static electricity of the same polarity and, since like charges repel, they stand apart. The greater the charge, the further apart the leaves spread.

Electron Flow. When an electrical conductor is placed between these two oppositely charged bodies, free electrong are attracted by the positive body—free electrons will move through the wire. This movement of free electrons will continue only until the excess of electrons is equally divided between the two bodies. Under these conditions, the charges on both bodies will be equal and the electron flow will end.

In Fig. 5 are a battery, lamp, and connecting leads between the battery and lamp. In this instance, the battery serves as an electric charge pump—free electrons continually ceveloped at its negative terminal by chemical action flow through the connecting leads and lamp back to the positive terminal of the battery by the attraction of oppositely charged bodies. The battery, connecting leads, and lamp form an electrical circuit which must be complete before the free electrons can flow from the battery's negative terminal to its positive terminal via the lamp. Thus, the battery serves as a source of potential difference or voltage by continually supplying a surplus of electrons at its negative terminal. Summing up, we can say a flow of electric current consists of the movement of electrons between two oppositely charged bodies.

We cannot progress very far into the study of electricity without first becoming familiar with the basic properties of electrical circuits. Just as we define distance in feet and inches, so do we define electrical properties in specific terms and units.

Potentail. Earlier, we saw that an electric charge difference has to exist between the ends of an electrical conductor in order to cause a flow of free electrons through the conductor. This flow of electrons constitutes the electric current. The electric charge difference, or potential difference, exerts a forces on the flow of free

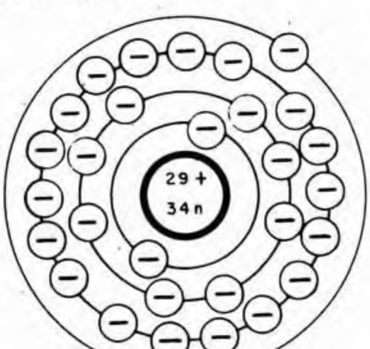


Fig. 3. The number of electrons to each ring are limited—2 in first; 8 in second; 18 for the third; and a total of 32 in fourth orbital ring. The last ring of electrons cannot have more than eight electrons.

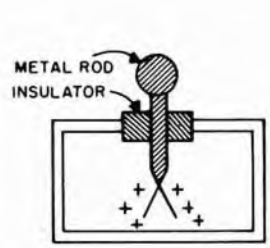


Fig. 4. Electroscope is a simple device to indicate electrical charges that are too weak to be measured with standard meters. The two leaves of gold foil pushed apart in the presents of a charged surface.

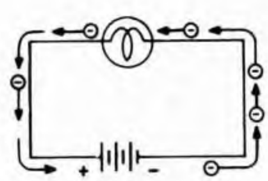


Fig. 5. Electron flow in any circuit is from negative to positive.

electrons, forcing them through the conductor. This electric force of pressure is referred to as electromotive force, abbreviated EMF.

The greater the charge or potential difference, the greater will be the movement of free electrons (current) through the conductor as there will be more "push and pull" on the free electrons. The symbol used to designate electrical potential is the letter E which stands for electromotive force. The quantity of EMF is measured by a unit called the volt. Hence, the common name most often used in place of EMF is *voltage.*

Current Intensity. We have learned that an electric current consists of a flow of charge carriers (generally free electrons) between two points of different electrical potential. The rate of flow of these charges determines the intensity or strength of this current flow. Current strength is expressed in units known as *amperes.* One ampere of current flows in a circuit when 6,240,000,000,000,000,000 electrons flow out of a negative terminal, through a conductor, and back into a positive terminal in one second. The symbol for the ampere is the letter *I* which stands for intensity.

Resistance. The flow of electric current through a conductor is caused by the movement of free electrons present in the atoms of the conductor. A bit of thought then indicates that the greater the number of free electrons present in the atoms of a particular conductor, the greater will be its electrical conductivity. Gold, silver, and copper rank as excellent electrical conductors, as their atoms readily release free electrons. On the other hand, the atoms of such elements as sulphur have almost no free electrons available and they ar ethus very poor electrical conductors. Such materials are known as electrical insulators. Between these extremes lie elements such as carbon whose atoms have a moderate number of free electrons available and thus are moderately good electrical conductors.

Even the best electrical conductors offer some opposition to the passage of free electrons. This opposition is called resistance. You might consider electrical resistance similar to mechanical friction. As in the case of mechanical friction, electrical resistance generates heat. When current flows through a given resistance, heat is generated; the greater the current flow, the greater the heat. Also, for a given current flow, the greater the resistance, the greater the heat produced.

Electrical resistance can be both beneficial and undesirable. Toasters, electric irons, etc. all make use of the heat generated by current flowing through wire coils. Resistance is also often intentionally added to an electrical circuit to limit the flow of current. This type of resistance is generally lumped together in a single unit known as a resistor.

There are also instances where resistance is undesirable. Excessive resistance in the connecting leads of an electrical circuit can cause both heating and electrical loss. The heating, if sufficient, can cause a fire hazard, particularly in house wiring, and the circuit losses are a waste of electrical power.

Electrical resistance is expressed by a unit known as the *ohm*, indicated by the letter *R*. An electrical conductor has a resistance of *one ohm* when an applied EMF of *one volt* causes a current of *one ampere* to flow through it.

Resistance Factors. There are other factors beside the composition of the material that determine its resistance. For example, temperature has an effect on the resistance of a conductor. As the temperature of copper increases, for example, its resistance increases. The increase in temperature causes the electrons in the outer ring of the atom to resist release to the free electron state. This increase in resistance is known as a *positive temperature coefficient.* Not all conductors show this increase in resistance with an increase in temperature; their resistance decreases with an increase in temperature. Such materials are said to have a *negative temperature coefficient*. Certain metallic alloys have been developed which exhibit a *zero temperature coefficient:* their resistance does not change with changes in temperature.

As you might suspect, the length of a conductor has an effect upon its resistance. Doubling the length of a conductor will double its resistance. By the same token, halving the length of a conductor will cut its resistance in half. Just remember that the resistance of a conductor is *directly proportional* to its *length.*

The cross-sectional area of a conductor also determines its resistance. As you double the cross-section of a conductor, you halve its resistance; halving its cross-section doubles its resistance. Here again, the "why" of this is pretty easy to see: there are more current carrying electrons available in a large cross-section conductor than in a small cross-section conductor of the same length. Therefore, the resistance of a conductor is *inversely proportional* to its *cross-sectional area.*

Circuit Relationship. Now that we have a basic understanding of voltage, current, and resistance, let's take a look at just how they interact under circuit conditions.

Fig. 6. In A, B, and C, the value of the resistor remains constant while the supply voltage is altered with a resulting current change.

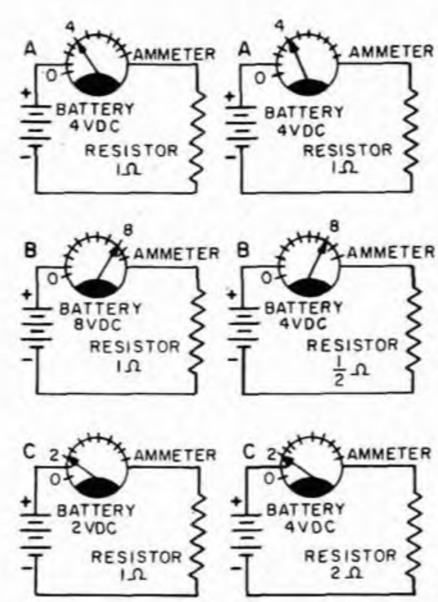


Fig. 7. Battery voltage A, B, and C is held constant while resistor is halved and doubled in value. Resulting current changes are basis for Ohm's law.

Fig. 6A shows a battery, ammeter (a device to indicate current strength), and resistor connected in series. Notice that the ammeter indicates that 4 amperes are flowing in the circuit.

Fib. 6B shows the identical setup with the exception that the battery voltage has now been doubled. The ammeter now shows that twice the original current, or 8 amperes, is now flowing in the circuit. Therefore, we can see that doubling the voltage applied to the circuit will double the current flowing in the circuit.

In Fig. 6C the same circuit appears again; this time, however, the battery voltage is one half its original value. The ammeter shows that one half of the original current, or 2 amperes, is now flowing in the circuit. This shows us that halving the voltage applied to the circuit willh alve the current flowing through the circuit.

All this boils down to the fact that, assuming the same circuit resistance in all cases, *the current flowing in a circuit will be directly proportional to the applied voltage*—increasing as the voltage is increased, and decreasing as the applied voltage is decreased.

In Fig. 7A we again see the circuit consisting of the battery, ammeter, and resistance. Notice that the ammeter indicates that 4 amperes are flowing through the circuit.

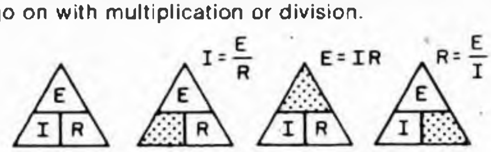
In Fig. 7B we see that the value of resistance has been cut in half and as a result, the ammeter indicates that twice the original current, or 8 amperes, is now flowing in the circuit. This leads us to the correct assumption that for a given supply voltage, halving the circuit resistance will double the current flowing in the circuit.

Fig. 7C again shows our basic circuit, but with the resistance now doubled from its original value. The ammeter indicates that the current in the circuit is now one half of its original value.

Summing things up: for a given supply voltage, *the current flowing in a circuit will be inversely proportional to the resistance in the circuit.*

Ohm's Law. From what you have seen so far, you are probably getting the idea that you can determine the current flowing in a circuit if you know the voltage and resistance present in the circuit, and the voltage if you know the current and resistance, or the resistance if the voltage and current are known.

Fig. 8. Shaded portion of triangle indicates unknown quantity in the formula. Visible factors appear in their proper mathematical relation. Just fill in the known values and go on with multiplication or division.



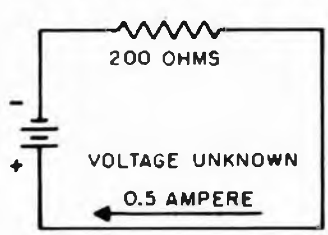


Fig. 9. Unknown quantity, voltage, found easily by applying Ohm's law.

All this is quite correct, and if sormally stated by Ohm's law as follows:

$$I = \frac{E}{R}$$

Where: E = voltage (volts)
I = current (amperes)
R = resistance (ohms)

Now, let's take a look at how this formula is used:

To find voltage:

$$E \text{ (voltage)} = I \text{ (current)} \times R \text{ (resistance)}$$

To find current:

$$I \text{ (current)} = \frac{E \text{ (voltage)}}{R \text{ (resistance)}}$$

To find resistance:

$$R \text{ (resistance)} = \frac{E \text{ (voltage)}}{I \text{ (current)}}$$

A handy way to remember Ohm's law is by means of the triangle shown in Fig. 8. Simply cover the quantity (voltage, current, or resistance) that you want to determine, and read the correct relationship of the remaining two quantities. For example, if you want to now the correct current (I), put your finger over I and read E/R. Covering E or R will yield I × R or E/I, respectively.

Ohm's Law to Determine Voltage. Let's delve a bit more deeply into Ohm's law by applying it to a few cases where we want to determine the unknown voltage in an electrical cirucit. Take a look at Fig. 9, which shows a simple series circuit consisting of a battery and resistor. The value of this resistor is given as 200 ohms, and 0.5 ampere of current is flowing through the circuit. We want to find the value of battery voltage. This is easily done by applying Ohm's law for voltage as follows:

$$E = I \times R$$

Let's go through this again, this time using a practical illustration. Fig. 10 shows a string of light bulbs, the total resistance of which is 400 ohms. You find that the bulbs draw 0.3 amperes when lighted. Let's say you would like to operate this string of bulbs from the standard 120-volt house current, but you don't know the voltage rating of the individual bulbs. By using Ohm's law for voltage, you can easily determine the voltage to light the bulbs as follows: (unknown voltage) = 0.3 (amperes) × 400 (bulb resistance) = 120 volts.

A word of caution. When cool, the resistance of a filament light bulb is much lower than when heated. As the filament is heated, the resistance increases due to positive temperature coefficient of the material as discussed earlier.

Ohm's Law to Determine Current. Now, let's take a look at a few examples of how to determine the value of unknown current in a circuit in which both the voltage and resistance are known.

Fig. 11 shows a series circuit with a battery and resistor. The battery voltage is 20 volts DC and the value of resistance is 5 ohms. How much current is flowing through the circuit?

Ohm's law for current: $I = \dfrac{E}{R}$

$$I = \frac{20 \text{ (battery voltage)}}{5 \text{ (resistance in ohms)}}$$

$$I = 4 \text{ amperes}$$

You will note that the 15 ampere fuse will not blow in this circuit.

Again to get a bit more practical, let's take a look at Fig. 12. Here we see an electric heater element connected to the 120-volt house line. We know that this particular heater element has a resistance of 20 ohms when hot. The house current line is fused with a 15-ampere fuse. We want to know whether the heter will draw sufficient current to blow the fuse. Here's how to find this out by use of Ohm's law for current.

$$I = \frac{120 \text{ (line voltage)}}{20 \text{ (Heater resistance in ohms)}}$$

$$I = 6 \text{ amperes}$$

We find from the above use of Ohm's law for current that the heater draws 6 amperes, so it can be safely used on the line fused with the 15-ampere fuse. In fact, a 10-ampere fused line can also do the job.

Ohm's Law to Determine Resistance. Ohm's law for resistance enables us to determine the unknown value of resistance in a circuit. Fit. 13 again shows a simple series circuit with the battery voltage given as 20 volts and the current flowing through the circuit as 0.5 ampere. The unknown resistance value in this circuit is found as follows:

Ohm's law for resistance:

$$R = \frac{E}{I}$$

$$R = \frac{20 \text{ (battery voltage)}}{0.5 \text{ (current in amperes)}}$$

$$R = 40 \text{ ohms}$$

Fig. 14 is a practical example of how to determine unknown resistance. Here, we want to operate a 6-volt light bulb from the 120-volt house line. What value of series dropping resistor do we need to drop the 120-volt house current down to 6 volts? The bulb draws 0.2 ampere when illuminated (hot).

We must first determine the voltage which must be dropped across the series dropping resistor. This is done by subtracting the line voltage (120) from the bulb's voltage (6). This gives us a value of 114 volts which we use in conjunction with Ohm's law for resistance as follows:

$$R = \frac{114 \text{ (voltage drop)}}{0.2 \text{ (current in amperes)}}$$

$$R = 570 \text{ ohms}$$

Resistance in Series. Many practical electrical and electronic circuits use two or more resistances connected in series. The point to remember in this case is that the total resistance is the sum of the individual resistances.

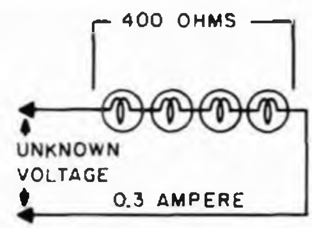


Fig. 10. Although problem looks different the basic circuit is same as that for Fig. 9.

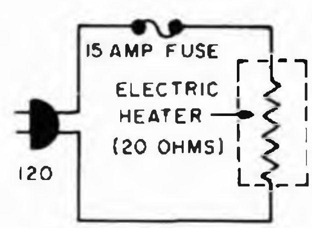


Fig. 11. Formula needed here is different since current is unknown. Just look for triangle in Fig. 8 that has I shaded.

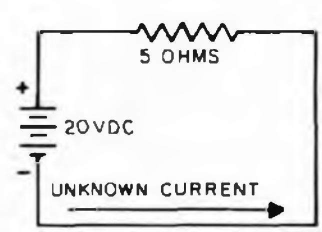


Fig. 12. Basic circuit is same as that in Fig. 11. Although three factors are given, current is unknown quantity.

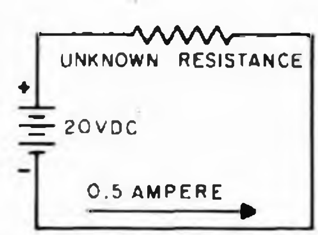


Fig. 13. Most Ohm's law problems are simple series circuits or can be reduced to simple series circuits.

This is expressed by the formula:

$$R \text{ (total resistance)} = R1 + R2 + R3 + \text{etc.}$$

where R1, R2, R3, etc. are the individual resistances. Thus, in Fig. 15 the total of the individual resistances is R (total) = 40 + 6 + 10 + 5 = 61 ohms.

Resistance in Parallel. Resistances may also be connected in parallel in a circuit as in Fig. 16. In this case the current flowing in the circuit will divide between the resistances, the greater current flowing through the lowest resistance. Also, the total resistance in the circuit will always be less than the smallest resistance since the total current is greater than the current in any of the individual resistors. The formula for determining the combined resistance of two parallel resistors is:

$$R \text{ (total)} = \frac{R1 \times R2}{R1 + R2}$$

Thus, in Fig. 16 the effective resistance of R1 and R2 is:

$$R \text{ (total)} = \frac{2 \times 4}{2 + 4} = \frac{8}{6} \text{ or } 1.33 \text{ ohms.}$$

In a circuit containing more than two parallel resistors as in Fig. 17 the easiest way to determine the total circuit resistance is as follows: first, assume that a 6-volt battery is connected across the resistor network. Pick a value that will make your computations simple. Then determine the current flowing through each of the resistors using Ohm's law:

$$I = E = 6 = 3 \text{ amperes}$$
$$I = E = 6 = 3 \text{ amperes}$$
$$I = R1 = 2 = 3 \text{ amperes}$$

$$I = E = 6 = 2 \text{ amperes}$$
$$I = E = 6 = 2 \text{ amperes}$$
$$I = R3 = 3 = 3 \text{ amperes}$$

$$I = E = 6 = 1 \text{ ampere}$$
$$I = E = 6 = 1 \text{ ampere}$$
$$I = R3 = 6 = 1 \text{ ampere}$$

Next, add the individual currents flowing through the circuit:

$$I = 2 \text{ amp.} + 3 \text{ amp.} + 2 \text{ amp.}$$

$$I = 6 \text{ amperes.}$$

Inserting this 6 amperes in Ohm's law, the total circuit resistance is found to be:

$$R = \frac{6}{6} = 1 \text{ ohm.}$$

The combined equation for determining the total resistance of *n* number of resistances in parallel would be:

$$1 = 1 + 1 + 1 + \dots 1 .$$
$$1 = 1 + 1 + 1 + \dots 1 .$$
$$R = R1 + R2 + R3 + \dots Rn .$$

Quite often an electronic circuit will contain a combination of series and parallel resistances as in Fig. 18. To solve this type of problem, first determine the combined resistance of R2 and R3:

$$R \text{ (total) } 6 \times 12 = 72 = 4 \text{ ohms.}$$
$$R \text{ (total) } 6 \times 12 = 72 = 4 \text{ ohms.}$$
$$R \text{ (total) } 6 + 12 = 18 = 4 \text{ ohms.}$$

This total value of R2 and R3 may be considered a single resistance which is in series with R1, and forms a simple series circuit. This simple series circuit is solved as follows:

$$R \text{ (total)} = 6 + 4, \text{ or a total of } 10 \text{ ohms.}$$

Power. The amount of work done by electricity is termed the *watt* and one watt is equal to one volt multiplied by one ampere. This may be expressed as: $P = E \times I$ where E = voltage in volts, I = current in amperes. Also:

$$P = E^2 \text{ and } P = I^2R$$
$$P = E^2 \text{ and } P = I^2R$$
$$P = R \text{ and } P = I^2R$$

As an example, assume that a toaster draws 5 amperes at an applied voltage of 115 volts. Its wattage would then be:

$$P = 115 \times 5 = 575 \text{ watts.}$$

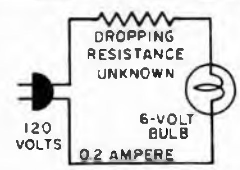


Fig. 14. This Ohm's law problem is somewhat more complex.

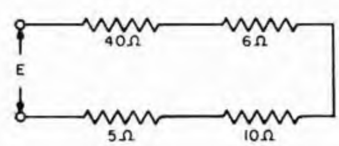


Fig. 15. Resistance in series is added. As far as voltage applied and current flow is concerned the individual resistors are only one.

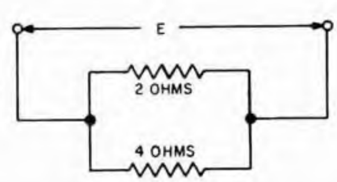


Fig. 16. Resistors in parallel are added algebraically—the result will always be a value less than that of the lowest in the circuit.

Magnetism and the Electron. The atom and a concept of its structure were a necessary preface to our discussion of basic electricity. By the same token, both are necessary to understanding basic magnetism.

As we've mentioned, electrons are in continual motion about the nucleus. The orbit is, in fact, a small loop of current and has a magnetic field that's associated with a current loop. In addition, experimental and theoretical investigation seems to indicate that the electron itself has a spin. Each electron, having its own axis, is a spinning sphere of electric charge. *Electron spin*, like the quantum and wave theories of light, is not so much a lteral interpretation of a phenomenon as a useful concept that holds water when applied to the phenomenon of magentism.

When the electron spins, the charge that is in motion produces a magnetic field. And, to briefly state the electronic explanation of magnetism, it seems that the magnetic properties of matter can be attributed to the orbital and spinning motion of the electrons comprising the atoms of the matter.

Millennia of Magnetism. Some of the basic principles and effects of magnetism have been known for centuries. The Greeks are credited as the ones who first discovered magnetism. They noted that a certain type of rock had the ability of attracting iron. Later, the Chinese noted that an elongated piece of this rock had the useful property of always pointing in a north-south direction when suspended by a string. This was the beginning our our compass.

This strange stone which intrigued people over the centuries is actually a form of iron ore known as magnetite. Not allmagnetite shows magnetic properties. Another name for the magnetic variety of magnetite is lodestone—the term lodestone being derived from two separate words, lode and stone. The term "lode" stands for guide, hence lodestone means "guide stone."

All magnets, whether natural or manmade, possess magnetic poles, which are commonly known as the magent's north and south pole. As is the case of the electrical charges (which we studied earlier) between unlike magnetic poles and repulsion between like poles, it has been found that this magnetic attraction and repulsion force varies inversely as the square of the distance from the magnetic poles.

The Magnetic Field. We all know how a magnet exerts a force of attraction on a piece of magnetic material such as iron or steel. Also, when the north poles of two magnets are brought close together, they will try to repel each other, while there will be attraction between the north and south poles of two magnets. Although it is not clearly understood just what this force of magnetic attraction and repulsion is, it is convenient to visualize magnetic lines of force which extend outward from one magnetic pole to the other as illustrated in Fig. 19.

Permeability. Magnetic lines of force can pass through various materials with varying ease. Iron and steel, for example, offer little resistance to magnetic lines of force. It is because of this that these materials are so readily attracted by magnets. On the other hand, materials such as wood, aluminum and brass do not concentrate or encourage the passage of magnetic lines of force, and as a consequence are not attracted by magnets.

The amount of attraction a material offers to magnetic lines of force is known as its permeability. Iron and steel, for example, possess high permeability since they offer little resistance to magnetic lines of force. Nonmagnetic materials have low permeability. For practical purposes, we can say that reluctance is to magnetic lines of force what resistance is to an electrical current.

Electromagnetism. Any electrical conductor through which flows an electrical current will generate a magnetic field about it which is perpendicular to its axis as shown in Fig. 20. The direction of this field is dependent upon the direction of current flow, and the magnetic field strength proportional to the current strength. If this current-carrying conductor is wound into a coil, forming a solenoid, the magnetic field will be increased by each individual turn that is added. If an iron core is inserted in this current-carrying coil, the generated field will be increased still further. This is because the lines of force are concentrated within the iron core which has considerably less reluctance than the surrounding air.

The magnetic power of a multi-turn current-carrying coil through which a core is inserted is proportional to the current flowing through the coil as well as the number of turns in the coil. The current through the coil is termed *ampere turns*. As an example, if a coil consisting of 200

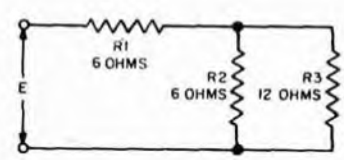


Fig. 17. Ohm's law can be used to determine the equivalent resistance of two or more resistors in parallel. Total current—then solve for ohms.

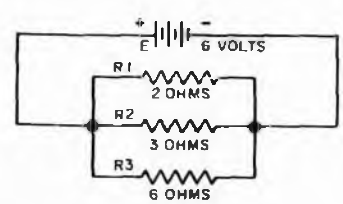


Fig. 18. Series-parallel circuit is not really difficult. Add R2 and R3 algebraically. Add effective resistance to R1 for total resistance.

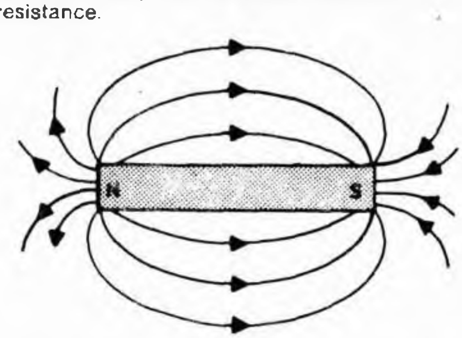


Fig. 19. Lines of force around bar magnet can be made visible by sprinkling iron filings onto white paper over magnet. Lightly tap the paper and all the filings will line up with the lines of force.

turns is carrying 2 amperes, its ampere turns equal:

Ampere turns = 200 turns × 2 amps.
Ampere turns = 400 ampere turns

Similarly a coil of 100 turns through which a current of four amperes flows also has 400 ampere turns.

Electromagnetic Induction. We saw earlier how a current-carrying conductor will generate a magnetic field which is perpendicular to the conductor's axis. Conversely, a current will be induced in a conductor when the conductor is passed through a magnetic field. The strength of this induced current is proportional to both the speed at which it passes through the field and the strength of the field. One of the basic laws pertaining to electromagnetic indcution is Lenz's law which states: "The magnetic action of an induced current is of such a direction as to resist the motion by which it is produced."

Fig. 21 illustrates two coils, A and B, which are placed in close proximity to each other. Coil A is connected in series with a switch and battery so that a current may be sent through it when the switch is closed, and coil B is connected with a current-indicating DC meter. When the switch is closed, current will flow through coil A, causing a magnetic field to be built up around it. In the brief instant that the field is building up to maximum, it will "cut" the turns of coil B, inducing a current in it, as indicated by a momentary flick of the indicating meter. When the switch is opened, breaking the current flow through coil A, the field around coil A will collapse, and in so doing will again induce a current in coil B. This time, however, the flow of current will be in the opposite direction. The meter will now flick in the direction opposite to when the switch was closed. The important thing to remember is that the conductor must be in motion with respect to the magnetic field or vice versa in order to induce a current flow. You can perform this simple experiment using two coils made of bell wire wrapped around large nails, a few dry cells in series, and a DC zero-center scale meter.

Self Induction. As mentioned a short while ago, a magnetic field is built up around a coil at the application of current through the coil. As this field is building up, its moving lines of flux will cut the turns of the coil inducing a counter-electromotive force or counter-EMF which opposes the current flowing into the coil.

The amount of counter-EMF generated depends upon the rate of change in amplitude of the applied current as well as the inductance of the coil. This value of inductance is dependent upon the number of turns in the coil; a coil with many turns will have greater inductance than a coil with few turns. Also, if an iron core is inserted into the coil, the inductance of the coil will increase sharply. The unit of inductance is known as the *Henry*.

The Transformer. One of the most important and widely used applications of magnetic induction is the transformer.

Fig. 22 shows the basic construction of a typical transformer. While two separate windings are shown here, some transformers can have as many as five or six windings.

A transformer consists of two or more separtate windings, electrically insulated from each ohter. One winding, known as the primary winding, is fed from a source of alternating current.

The alternating currents flowing through the primary induce a current in the secondary winding by virtue of magnetic induction. The transformer core is constructed from a relatively high permeability material such as iron which readily conducts magnetic flux between the primary winding and the secondary winding.

The alternating current flowing in the primary of the transformer produces a variation in the magnetic flux circulation in the transformer core which tends to oppose the current flowing in the primary winding by virtue of self-induction. The counter-EMF is just about equal to the voltage applied to the primary winding when no load is connected to the transformer's secondary winding.

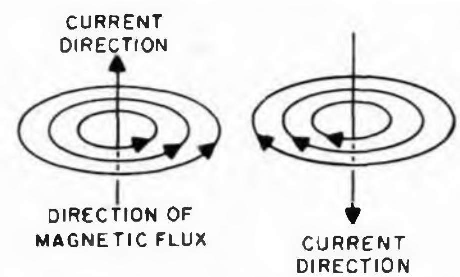


Fig. 20. Direction of flux lines is changed by direction of the current. Heavy current is needed tomake flux lines visible with sprinkled filings.

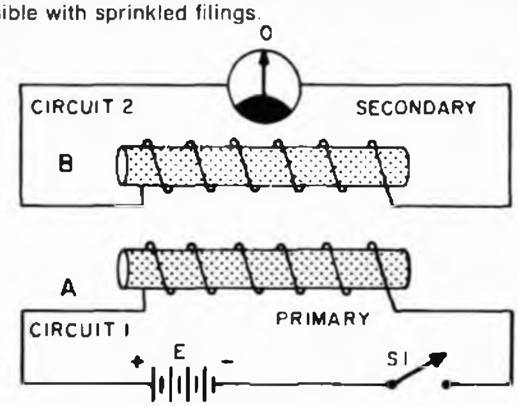


Fig. 21. Two-core transformer is inefficient since an air gap at either end does not have permeability of a ferrous metal and some flux lines do not go through core of secondary winding (B)—their effect is lost.

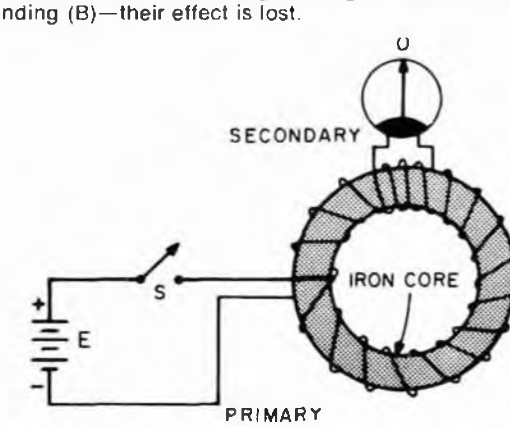


Fig. 22. Toroidal core is highly efficient but is very difficult tom anufacture. Familiar C- and E-shape core has less waste and windings are slipped over the core. Efficiency is good—about 90 percent for most designs.

This accounts for the fact that very little current flows through the primary winding when no load is connected to the secondary. The negligible current that does flow under this no-load condition is known as the transformer magnetizing current. As the current drawn from the secondary winding increases, the primary current will increase proportionately due to the reduction in the counter-EMF developed in the primary winding of the transformer.

In any transformer the ratio of the primary to secondary voltage is equal to the ratio of the number of turns in the primary and secondary windings. This is expressed mathematically as follows:

$$\frac{Ep}{Es} = \frac{Np}{Ns}$$

where Ep = primary supply voltage
Es = voltage developed across secondary
Np = number of primary turns
Ns = number of secondary turns

The above formula assumes that there are no losses in the transformer. Actually, all transformers possess some losses which must be taken into account.

Transformer Losses. No transformer can be 100 percent efficient due to losses in the magnetic flux coupling the primary and secondary windings, eddy current losses in the transformer core, and copper losses due to the resistance of the windings.

Loss of magnetic flux leakage occurs when *not all* the flux generated by current flowing in the primary reaches the secondary winding. The proper choice of core material and physical core design can reduce flux leakage to a negligible value.

Practical transformers have a certain amount of power loss which is due to power being absorbed in the resistance of the primary and secondary windings. This power loss, known as the copper loss, appears as heating of the primary and secondary windings.

There are several forms of core loss—hysteresis and eddy current losses. Hysteresis losses are the result of the energy required to continually realign the magnetic domain of the core material. Eddy current loss results from circulating currents induced in the transformer core by current flowing in the primary winding. These eddy currents cause heating of the core.

Eddy current loss can be greatly reduced by forming the core from a stack of individual sheets, known as laminations, rather than from a single solid piece of steel. Since eddy current losses are proportional to the square of core thickness, it is easy to see that the individual thin laminations will have much less eddy current loss as compared with a single thick core.

Another factor which effects eddy current loss is the operating frequency for which the transformer is designed to operate. As the operating frequency is increased, the eddy current losses increase. It is for this reason that transformers designed to operate at radio frequencies often have air cores and are void of ferrous metals.

Theory and Practice. We've come a long way from our initial discussion of the atom and its importance for an understanding of electricity and magnetism. And there's still a long way to travel to understand all about the sub-atomic nucleus and its satellites and how they are being harnessed in an ever-expanding electronics technology. But, we move ahead by mixing theory with practice—so, put your new knowledge to work in a project or two! ■

The Language Of Electronics

Anyone who has built a small project or read a beginning article on electronic theory is certain to have run across such terms as *micro*Farad, *milli*Henry, and *milli*Ampere—not to mention *mega*Hertz, *meg*Ohm, and *kilo*Hertz. The prefixes here, *micro-*, *milli-*, *mega-*, and *kilo-*, are an important part of the electronic vocabulary. It follows, then that anyone who wants to be proficient in electronics will have to develop skill in understanding and using the "language."

These prefixes are used to change the value of an electronic unit of measure. For example, if you see a resistor with the familiar brown/black/green color code, you *could* call it a 1,000,000-ohm resistor. The thing is, it's usually less awkward to call it a 1-megohm resistor. Purring the prefix *meg-* or *mega-* before the Ohm inflates the value of the unit, Ohm, by 1,000,000 times.

Similarly, one kiloVolt is rcognizable as 1,000 Volts, and one kiloHertz as 1,000 Hertz, and so on. These prefixes are usually so automatic with electronics aficionados that they will invariably refer to a millionaire as a guy who has one megabuck!

The Mini Side. At the other end of the scale, the *milli-* and *micro-* prefixes are useful for shrinking units. A Farad, for example, is too big a unit to use in everyday electronics. In dealing with the real-life capacitors (the kind you solder into circuits), we normally use a basic unit of *one-millionth* of a Farad—a *micro*Farad. The prefix *micro-* cuts up a unit into a million tiny slices, enabling us to use one such slice as a convenient-sized unit. A microAmpere, similarly, is a millionth of an Ampere; a microVolt, one millionth of a Volt.

If you need larger slices, the *milli-* prefix is available, which provides a unit only one-thousandth the size of the basic unit. A milliAmpere, for example, is a thousandth of an Ampere; that is, it takes 1000 mA (milliAmperes) to equal 1 Ampere.

To handle these tiny slices of units, it's wise to spend a few minutes learning *scientific notation*, which is designed to make it easy to handle very large and very small numbers. Once you've mastered this technique, you can manipulate all the various-sized units of electronics as easily as you can add two and two!

The Maxi Side. Take, for example, the familiar *kiloHertz* (known at one time as the *kilocycle*). A broadcasting station operating at 840 kHz (kiloHertz) in the broadcasting band is radiating 840,000 cycles of RF energy every second. To change from 840 kHz to 840,000 Hz, you can think of the "kilo-" as being replaced by "x 1000", thus:

840	kilo	Hertz
840	× 1000	Hertz
840,000		Hertz

But you can also write "1000" as "10 x 10 x 10". And you can write "10 x 10 x 10" as "10^3". (Ten to the third power, or ten cubed.) As we develop these ideas further, you will see how you can greatly simplify your future work in electronics by thinking of the prefix "kilo-" as being replaceable by "x 10^3", thus:

840 kiloHertz = 840 × 10^3 Hertz

Similarly, a 6.8-megohm resistor, measured on an ohmmeter, will indicate 6,800,000 ohms. In this case, the prefix "meg-" can be replaced by "x 1,000,000":

6.8	meg	Ohms
6.8	= 1,000,000	Ohms
6,800,000		Ohms

But you can write "1,000,000" as "10 x 10 x 10 x 10 x 10 x 10" (six of them; count 'em), which is 10^6. Thus, you should learn to mentally replace "meg-" with x 10^6, so that 6.8 megOhms becomes a 6.8 x 10^6 Ohms. The *6* is called an *exponent*, and shows how many 10s are multiplied together.

The Minus Crowd. What about the "milli-" and "micro-" prefixes? "Milli-", we've said, is one-thousandth; in a way it is the opposite of the "kilo-" prefix. Make a mental note, then, that milli- can be replaced with "10^{-3} (read as "ten to the **minus** three power"), which is 1/10 x 1/10 x 1/10 = 1/1000. Similiarly, the "micro-" prefix can be considered as the opposite of "meg-", and replaced by 10^{-6}.

The beauty of this approach appears when you are faced with a practical problem, such as, "if 1.2 milliAmperes flows through 3.3 megOhms, what voltage appears across the resistor?" From our knowledge of Ohm's law, we know that E = IR; that is, to get Volts (E) we multiply current (I) times resistance (R). Without the aid of scientific notation, the problem is to multiply 0.0012 Amperes by 3,300,000 Ohms, which is rather awkward to carry out. The same problem, however, is very easy in scientific notation, as can be seen below:

$$\begin{array}{l} 1.2 \times 10^3 \\ 3.3 \times 10^6 \\ 3.96 \times 10^3 \end{array}$$

The answer is 3.96 x 10^3 Volts, or 3.96 kiloVolts. We obtained the answer by multiplying 1.2 x 3.3 to get 3.96, and adding the -3 exponent to the 6 exponent to get 3 for the exponent of the answer. The advantage of scientific notation is that the largeness and smallness of the numbers involved is indicated by numbers like 10^6 and 10^{-3}, and the largeness or smallness of the answer is found by *adding* the 6 and the -3.

What about a division problem? For the sake of a good illustrative example, consider the unlikely problem of finding the current when 4.8 megaVolts is applied across 2 kilOhms. The problem is written as:

$$I = \frac{E}{R}$$

$$= \frac{4.8 \text{ megaVolts}}{2 \text{ kilOhms}} \cdot = \frac{4.8 \times 10^6 \text{ Volts}}{2.0 \times 10^3 \text{ Ohms}}$$

$$4.8 \div 2 = 2.4,$$

where 2.4 × 10^3 Amperes = kiloAmperes.

In division, then finding the size of the answer becomes a *subtraction* problem, in which the exponent representing the size of the divisor ("bottom" number) is subtracted from the exponent representing the size of the dividend ("top" number).

A more practical division problem answers the question, "What current flows when 5 Volts is applied across 2.5 kilOhms?"

$$I = \frac{E}{R} = \frac{5 \text{ Volts}}{2.5 \text{ kilOhms}}$$

$$= \frac{5.0 \times 10^0}{2.5 \times 10^3} = \frac{5.0 \times 10^{0-3}}{2.5}$$

$$= 2.0 \times 10^{-3} \text{ Amperes}$$

$$= 2.0 \text{ miliAmperes}$$

Note that it's perfectly legal to use 10^0 (ten to the zero power) to indicate a unit that has no prefix—in other words, one of anything.

For the Solving. Here are a few more problems:

1. The inductive reactance of a coil is given by

$$X_L = 2\pi fL$$

What is the reactance of a coil whose inductance L = 22 milliHenries, when an alternating current of frequency f = 1.5 megaHertz is applied to it?

$$\begin{aligned} X_L &= 2 \times \pi \times (1.5 \times 10^6) \times (22 \times 10^{-3}) \\ &= 207.24 \times 10^3 \text{ Ohms} \\ &= 207.24 \text{ kilOhms} \end{aligned}$$

2. An oscillator is connected to a wavelength-measuring apparatus, and the wavelength of its oscillations is determined to be 2.1 meters. What is the frequency of the oscillator?

$$F = \frac{\text{speed of light}}{\text{wavelength}}$$

$$F = \frac{3.0 \times 10^8 \text{ meters per second}}{\text{wavelength}}$$

$$F = \frac{3.0 \times 10^8}{2.1 \times 10^0} = 1.4286 \times 10^8 \text{ Hertz}$$

We wish this answer had come out with a "10^6", instead of a "10^8", because we can convert 10^6 Hertz directly to megaHertz. However, we can change the answer to 10^6, by shifting the decimal point of the 1.4286. Remember this rule: To *lower* the exponent, shift the decimal point to the *right*. (Of course, the opposite rule is also true.) Since we wish to lower the exponent by 2, we must shift the decimal point to the right by two places:

$$1442.86 \times 10^6 \text{ Hertz} = 142.86 \text{ megaHertz}$$

3. A 3.3 microfarad capacitor is being charged from a 20-volt battery through a 6.8-kilOhm resistor. It charges to half the battery voltage in a time given by

$$T = 0.69RC$$

For the particular values given in the problem, what is the time taken to charge to half the battery voltage?

$T = 0.69 \times (6.8 \times 10^3) \times (3.3 \times 10^6)$
15.4 miliseconds

4. A 365-pF variable capacitor and a 2-μH coil are found collecting dust in your junk box. You decide you might like to incorporate them into a radio but you need to know the resonant frequency of this inductive/capacative circuit. You apply the formula:

$$f = \frac{1}{2\pi \quad LC}$$

Since C = 365-pF or 365×10^{-12} Farads and L = 2-uH or 2×10^{-6} Henrys we can use these numbers, the formula and our new knowledge of exponents to determine the frequency.

$$f = \frac{1}{2\pi \quad (2 \times 10^{-6}) \times (365 \times 10^{-12})}$$

= 5,894,627.6 Hertz
= 5,895 kiloHertz
= 5.895 megaHertz

Tera to Atto. Since scientific notation is so potent, you'll probably be interested in the meaning of *all* the prefixes used in the scientific community, not just the four (micro-milli-, kilo-, and mega-)—that we've discussed so far. Very common in electronics is the *micro- microFarad*, which is $10^{-6} \times 10^{-6}$ Farad, or 10^{-12} Farad. This is more commonly known as the *picoFarad.* Similarly, a thousandth of a microAmpere is $10^{-3} \times 10^{-6}$ Ampere, or 10^{-9} Ampere. This is known as a nanoAmpere. At the other extreme, 1000 megaHertz is called a gigaHertz. See the Table of electronic prefixes and their meanings for all these prefixes, and, for a rundown of their meanings and pronunciations.

The jargon of electronics which has grown up around their prefixes is just as important as the prefixes themselves. Here are some examples of "jargonized" prefixes as they might appear in speech:

Puff—a picoFarad (from the abbreviation, PF).
Mickey-mike—a micro-microFarad (which is the same as a **puff**).
Meg—a megohm. Also, less often, a megaHertz.
Mill—a milliAmpere.
Megger—a device for measuring megohms.
dB (pronounced "dee-bee")—a deciBel, which is one-tenth of a Bel.
Mike—a microFarad. Also, to measure with a micrometer.

So, if you understand the prefixes and know their corresponding exponents, you'll have command of another set of important tools to help you do practical work in electronics. In addition, you'll be ready for the inevitable wise guy who'll ask if you can tell him the reactance of a 100-puff capacitor at 200 gigaHertz. After calculating the answer in gigaseconds, reply in femt-Ohms! ■

Equivalent Resistance Nomograph

Whether you're working at home on the final stages of a pet project or on-the-job-servicing an electronic system, nothing is quite as frustrating as discovering that the resistance value you need isn't available. And, your usual source of supply either is closed or doesn't stock the particular value. Or maybe you want a resistance within a tolerance of 1%, and just don't feel justified in paying the extra cost.

Whatever the problem, the experienced guy doesn't lose his cool, because he knows he can come up with any resistor value he needs by connecting available resistors in series and/or parallel. This combination can either be left in the circuit or replaced at some later time with a single resistor.

Making Resistors. Making resistors by *series-ing* several resistors to reach a desired value poses no problem as the resistances are additive; i.e., if you connect a 51-ohm resistor in series with a 68-ohm resistor the final resistance of the combination is 119 ohms.

However, when you parallel resistors, the resultant resistance is no longer so easy to calculate. If you connect a 51-ohm resistor in parallel with a 68-ohm resistor, the net resistance value is about 29 ohms. About the only thing you know is that the equivalent resistance of a parallel combination will be less than the value of the smallest resistor in the combination. You can't determine the equivalent resistance of a parallel combination with simple mathematics. The formula isn't complex, but it does take time to write down and solve. The easiest, fastest method for determining the values of parallel resistor combinations for the serviceman is by using an *equivalent resistance nomograph*.

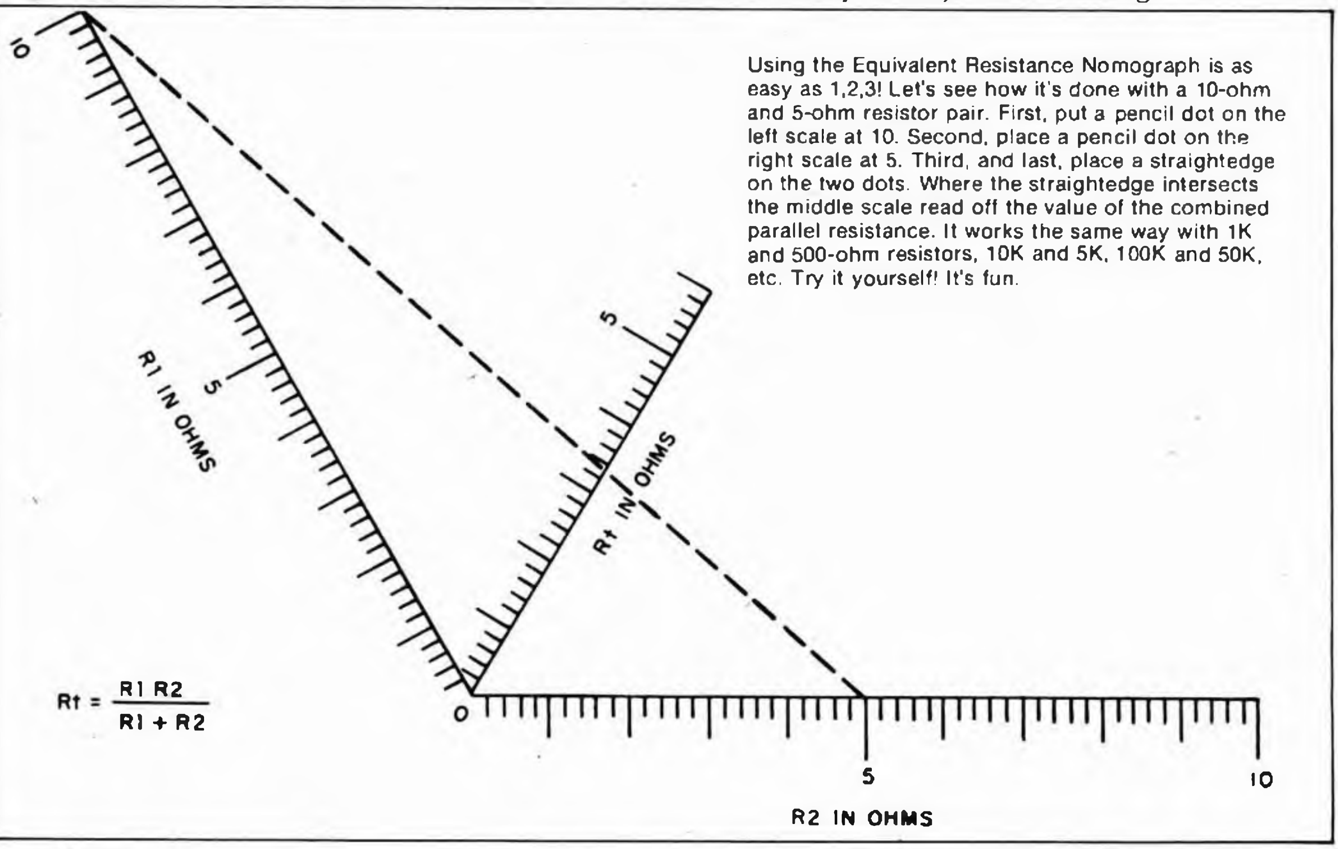
What's A Nomograph? Everyone's familiar with the old old Chinese proverb about one picture worth a thousand words. A nomograph is simply a graphic picture of a simple approach to solving a mathematical calculation. And technicians in all fields are using nomographs in ever-increasing numbers. A nomograph can be constructed to solve almost any problem, and though the actual construction may require a master's degree in math, anyone can use the final end product to solve problems which might otherwise require a college degree and bushels of valuable time when using equations.

This is one of the most appealing features of most nomographs, i.e., that you don't need theoretical knowledge of the subject to use a nomograph to solve problems in that field. All that's necessary is to lay a straightedge, or draw a line, between two known values on given scales, and read the answer where the line intersects a third scale.

Making A Nomograph. The nomograph printed in these pages is an equivalent resistance nomograph that can be cut out for use in yor work. Or, you may want to get a Xerox copy. With it you can determine the resistance of any two resistors connected in parallel in much less time than you could normally write down the mathematics required to solve the problem.

The R1 and R2 scales are equal in length, and positions at an angle of 120° with respect to one another. The Rt scale is a little more than one half the length of the other two scales, and bisects the angle between them. The scale lengths and anuglar positioning are usually by courtesy of some slaving mathematician somewhere, but, if you have the time and patience, you can construct some nomographs by trial and error. The graduations on all scales of our nomograph are of the same length and can be assigned any value that you desire as long as the same size and values are used on all scales. For example, if one major division on the R1 scale is valued 100 ohms, then one major division on the R2 scale and one major division on the Rt scale must also be valued 100 ohms. With this in mind, let's find out how to use the equivalent resistance nomograph to solve parallel resistance problems.

Using A Nomograph. The equivalent resistance nomograph can be used in either of two ways. In one application you have two resistors connected in parallel and want to know what value single resistor will be needed to replace the parallel combination. This situation often arises in breadboarding new circuits. To solve this problem you simply locate one resistance value on the R1 scale, and the other resistance value on the R2 scale. Then lay a ruler, or draw a straight line between



$$Rt = \frac{R1\,R2}{R1 + R2}$$

Using the Equivalent Resistance Nomograph is as easy as 1,2,3! Let's see how it's done with a 10-ohm and 5-ohm resistor pair. First, put a pencil dot on the left scale at 10. Second, place a pencil dot on the right scale at 5. Third, and last, place a straightedge on the two dots. Where the straightedge intersects the middle scale read off the value of the combined parallel resistance. It works the same way with 1K and 500-ohm resistors, 10K and 5K, 100K and 50K, etc. Try it yourself! It's fun.

the points located on the R1 and R2 scales. The equivalent resistance will be where the straightedge crosses the Rt scale.

In another application you know the value of one resistor and want to know what value of resistance must be connected in parallel with it to obtain a desired value. This problem may arise because your stock of resistors is depleted, or because the required resistor is not a standard value. Non-standard values of resistance cost more, of course, and at times two resistors in parallel will enable you to get the desired resistance at a much lower cost. To arrive at the value of the resistor that you need to aparllel with onre of known value to reach the odd-ball resistance you want, find the mark on either the R1 or the R2 scale for the known resistance value. Next locate the resistance of the desired resistor value on the Rt scale. Then lay a straightedge between the two points, and read the value of the required parallel resistor on the remaining scale.

Typical Problems. A typical problem will serve as an example that should bring everything into sharp focus now. Let's suppose that we have two resistors, 100,000 ohms and 47,000 ohms, connected in parallel in a project that's breadboarded and now ready for finalizing. With this parallel combination in the circuit our little jewel works fine, but the combination is bulky, unsightly, and expensive for quantity production. So boviously it's desirable to replace the bulky parallel resistor combination with a single fixed resistor.

Using the Equivalent Resistance Nomograph, locate the 100,000-ohm value on either the R1 or R2 scale (we used the R1 scale). We could have chosen any point on the scale as 100,000 ohms, but for better resolution the maximum point is the best choice. Next locate 47,000 ohms, but for better resolution the maximum point is the best choice. Next locate 47,000 ohms on the R2 scale, remembering that each major division is equal to 10,000 ohms because of thelocation of our assignment of the 100,000-ohm point on R1.

Now lay a straightedge across the nomograph so that it intersects the 100,000 and 47,000-ohm points on R1 and R2. Where the straightedge crosses the Rt scale, a line can be drawn on the nomograph, or, if you prefer, you will read the resultant resistance value, 32,000 ohms, on the Rt scale. In comparison, the correct answer, using slide rule and/or pencil and paper, of 31,950 ohms certainly will take much longer to calculate than if you use the Equivalent Resistance Nomograph. ■

Inductance And Capacitance

The lowly resistor is the solid citizen of electronics. It can be relied upon to behave with the same, predicatble performance, whether confronted with AC or DC. Unruffled by excursions into the higher frequencies, it continues to live by the guiding rule—Ohm's law—and declares that, no matter what, the current (I) it permits to flow shall depend solely on the applied voltage (E) divided by its own resistance (R).

However, the resistor's component cousins—the inductor and the capacitor—are by no means so stolid in the face of changing frequencies. The inductor, for example, grouchily shuts off more and more of the current flow as the frequency of the current passing through it gets higher and higher; while the capacitor reacts to higher frequencies in just the opposite manner—it happily allows more and more current to flow as the frequency rises.

Fortunately, the reactionary behavior of these components is just as predictable as is the single-mindedness of their resistive cousin. If we study the strange conduct of these apparently erratic citizens, we not only discover the rules which govern their odd behavior, but also perceive that ultimately, they too, are faithful to Ohm's law—in their fashion—just as are all electronic components and circuits.

Reactionaries in the Lab. To begin our study, let's set up the lab experiment shown in Fig. 1.

Here, we have an audio oscillator set to produce an output of 60 Hz, and a power amplifier to amplify that signal to the 100-volt level. The power amplifier drives a load consisting of a larte (25-watt) 390-ohm resistor connected in series with an ordinary 1½-volt (0.25-amp) flashlight bulb. An AC meter reads the output voltage from the power amplifier.

Turning on the equipment, we set the audio oscillator to 60 Hz, and gradually turn up the amplifier gain till 100 volts appears at the output. (Note: The ordinary hi-fi amplifier won't do this—you will need a public address amplifier with a high-voltage (70-volt, or 500-ohm) auxiliary output if you actually want to carry out this experiment.) At this point, the bulb will light to its normal brightness, indicating that the 100 volts is pushing about 0.25 ampere through the combined resistor/light-bulb circuit. Using Ohm's law, we can easily check this:

$$
\begin{aligned}
I &= E/R \\
&= 100 \text{ volts}/(390 \text{ ohms} + 6 \text{ ohms}) \\
&= 0.252 \text{ amperes}
\end{aligned}
$$

Note that the light bulb's resistance, R, is

$$
\begin{aligned}
R &= E/I \\
&= 1\tfrac{1}{2} \text{ volts}/\tfrac{1}{4} \text{ ampere} \\
&= 6 \text{ ohms}
\end{aligned}
$$

Please also note that it's the 390-ohm resistor, not the bulb, which is the chief authority in establishing the current. The lamp current will change very little, whether the bulb is 6 ohms, 12 ohms, or near zero ohms.

Next, let us vary the frequency of the audio oscillator;

first, down to 30 Hz, and then, up to 120 Hz. We notice that the bulb stays at the same brightness, indicating that the 390-ohm resistor is behaving in its normal, stolid fashion—that is, it is steadfastly ignoring frequency changes, and permitting its current flow to be determined solely by Ohm's law. Even if the frequency were zero (which is another way of saying DC), the bulb's brightness would remain the same, if we applied 100 volts to it.

Enter the First Reactionary. Now let's go to our parts supply bin, pick up a 1-Henry inductor, and make a few preliminary measurements on it. Using an inductance bridge, we discover that its real value is 1.05 Henry. We next connect it to an ordinary ohmmeter as shown in Fig. 2, which informs us that the inductor has a resistance of 45 ohms.

Now, let us replace the 390-ohm resistor of Fig. 1 with this 1-Henry inductor as shown in Fig. 3, and predict what will happen when we turn on the equipment. Since the ohmmeter said "45 ohms" we can predict that the current will be

$$\begin{aligned} I &= E/R \\ &= 100 \text{ volts}/45 \text{ ohms} + 6 \text{ ohms} \\ &= 1.96 \text{ amperes!} \end{aligned}$$

With this large current—nearly 8 times normal—the light bulb should burn out almost instantly! However, when we apply 100 volts of 60 Hz to the inductor-plus-bulb, we are surprised to see that the bulb lights to normal brightness! This means that the inductor is behaving like a 390-ohm resistor, and is establishing a ¼-ampere flow of current—not the nearly-2 amperes calculated from the above ohmmeter measurement.

To compound the mystery, let us now vary the frequency of the oscillator, first, up to 120 Hz—and the bulb gets dimmer!—and then gingerly, down, just a little, to 50 Hz—and the bulb gets uncomfortably brighter. Here, in the lab, is the actual behavior forecast by theory—the inductor grouchily shuts down the current flow for high frequencies, causing the bulb to go dim at 120 Hz, but it is willing to let low frequencies through, thus allowing the bulb to get brighter for 50-Hz input.

A Little Compassion for the Reactionary. To understand this "reactionary" behavior, we must understand how an inductor "feels" about an alternating current. An inductor is, after all, an electromagnet. If a steady direct current flows through it, it fills the space in its vicinity with a magnetic field. If we attempt to cut down the inductor's current, it reacts, quite understandably, by collapsing its magnetic field. But this collapsing field moves across the inductor winding in just the same way that a dynamo or generator field moves through the generator windings to make an output voltage. The inductor, then, reacts to any attempt to change its current by *acting as its own generator*, generating a new voltage of the correct polarity *to try to keep its own current from changing*. So, in its own way, the inductor is a solid, but *conservative* citizen—a citizen who tries to maintain the *status quo*.

Furthermore, the faster we try to change its current, the harder the inductor works to keep the current from changing. Therefore, the inductor sees a high frequency as an attempt at rapid changes—a threat to the *status quo*—so it works very hard to generate an opposing voltage (a 'counter-EMF' or 'counter-electro-motive-froce') to keep its current from changing. This internally-generated voltage oppposes the applied voltage more and more as the frequency rises; hence the actual current which flows drops lower and lower as the frequency rises. *This means that the apparent resistance of the inductor rises with frequency.* But this apparent resistance is not called *resistance*—since it is the inductor's reaction to the frequency applied to it, it is called *inductive reactance.*

A Resistance by Any Other Name. But whether you call it "apparent resistance" or "inductive reactance", it is still measured in ohms, and can be used as a part of the familiar Ohm's law formula. Where a simple resistive circuit answers to the expression

$$I = E/R$$

A similar circuit with resistor replaced by an inductor (coil) is described by the formula:

$$I = E/X,$$

where X is the symbol for reactance. But, since the

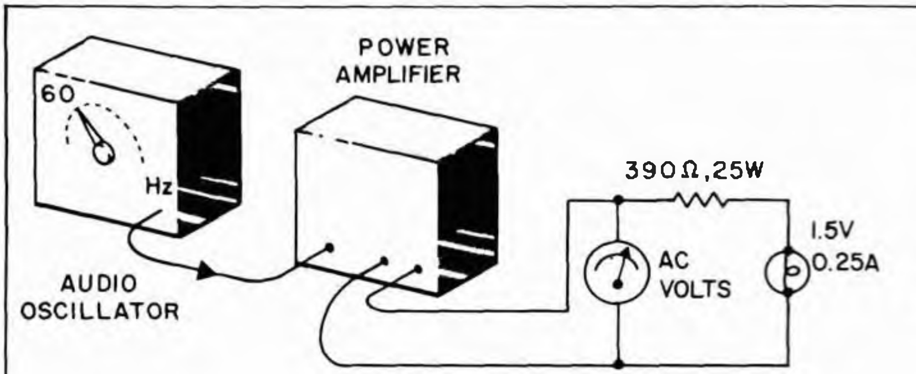


Fig. 1. Practical laboratory setup to demonstrate Ohm's law. Power audio amplifier must have a 70-volt output or a 500-ohm output transformer tap. Ordinary high-fidelity amplifiers cannot be used.

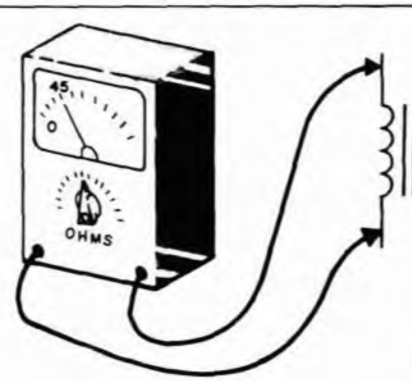


Fig. 2. Measuring the DC resistance of a one-Henry inductor (choke coil) shows it has 45 ohms of resistance in its copper winding.

amount of reactance (X) changes according to frequency, we must have a way to calculate its value at the frequency we are using. The following simple formula does that for us:

Inductive Reactance = 2 × 2 × π × frequency × inductance or, in the familiar algebraic shorthand,

$$X_L = 2\pi fL$$

where L is the inductance in Henrys, and the subscript L following the X indicates we are talking about *inductive* reactance. Therefore, the current in an inductive circuit is:

$$I = E/X_L$$
$$= E/2\pi fL$$

For Example. Let's take the 1.05-henry inductor and calculate its reactance at 60 Hz:

$$X = 2\pi fL,$$
$$= 2 \times \pi \times 60 \times 1.05$$
$$= 395.8 \text{ ohms.}$$

which, as you can see, is very close to the 390 ohms of the resistive circuit of Fig. 1. This explains why the bulb lit to about the same brightness for the inductor as for the resistor. (This also explains why we selected a 1.05 Henry inductor—we wanted you to see like values in examples.)

When we cranked the audio oscillator up to 120 Hz, the inductive reactance became

$$X_L = 2\pi fL,$$
$$= 2 \times \pi \times 120 \times 1.05$$
$$= 791.7 \text{ ohms.}$$

The new current becomes, ignoring, for the moment, the 6-ohm bulb:

$$I = E/X$$
$$= 100 \text{ volts}/791.7 \text{ ohms,}$$
$$= 0.125 \text{ ampere.}$$

which is about half the rated current of the bulb. Hence, it would be dim at this frequency. You can easily calculate that at 50 Hz, the X becomes 329.9 ohms and the current rises to the excessive value of .303 amperes; any further lowering of frequency could burn out the bulb!

Enter the Capacitor. Returning to the parts-supply bin, we now take a large, *oil-filled* (—Don't try this with an electrolytic!) capacitor, and, measuring it on a capacitance bridge, find that its true value is 6.8 μF. An ohmmeter placed across the capacitor's terminals registers an upward 'kick' of the needle as the ohmmeter's internal battery charges the capacitor, but the ohmmeter then settles down to indicate that, as far as *it* is concerned, the capacitor is—as it should be—an open circuit.

We now replace the 390-ohm resistor of Fig. 1 with the 6.8 μF capacitor, as shown in Fig. 4.

Again set the oscillator to 60 Hz, and turn on the equipment. Since the capacitor is really an open circuit—

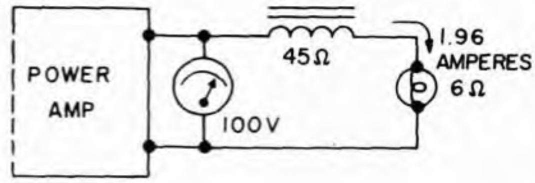


Fig. 3. Substituting the one-Henry inductor in place of the 390-ohm resistor of Fig. 1 offers the same opposition of the circuit to AC, hence the same current flows.

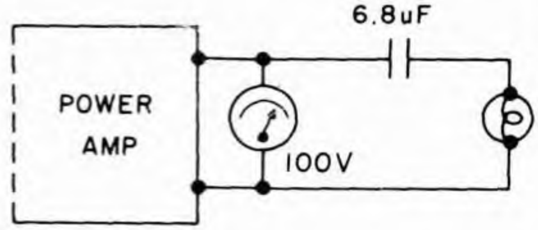


Fig. 4. A 6.8 μF capacitor substituted in the test circuit for the inductor. If the capacitor acted as an open (as it would for DC) no current would flow.

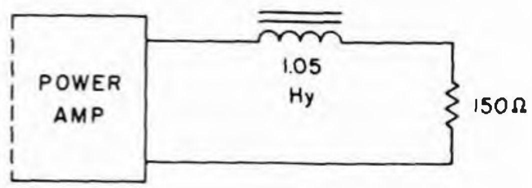


Fig. 5. Test circuit load is now the combination of the inductor (1.05 Henry) and the resistor (150 ohms). As the text explains, their combined reactance is computed by adding them at right angles!

according to the ohmmeter—one might expect no current flow at all. We are surprised, then, when the light bulb blithely ignores our oil-filled oil-filled open circuit, and proceeds to glow as serenely and confidently as it did for the inductor and resistor! When we *drop* the frequency to 30 Hz, the bulb goes dim—when we gingerly raise the frequency to 70 Hz, the bulb gets brighter. This is just the opposite of the inductive effect, where the bulb got dimmer for higher frequencies.

Capacitive Reactance. Its behavior in the face of changing frequency is exactly the opposite of that shown by inductive reactance. A capacitor's apparent resistance goes *down* as frequency goes up, and goes *up* as frequency goes down. This inverse relationship is seen in the expression for capacitive reactance, which is

$$X_C = \frac{1}{2\pi fC}$$

where X is *capacitive* reactance, and C is the capacity in *farads*, (not microfarads.)

Mathematically speaking, by placing f (frequency) in the *denominator* (bottom) of the fraction, we are saying that as f gets *larger*, the answer (X_C) gets *smaller*, which is exactly what we observed in the lab experiment above.

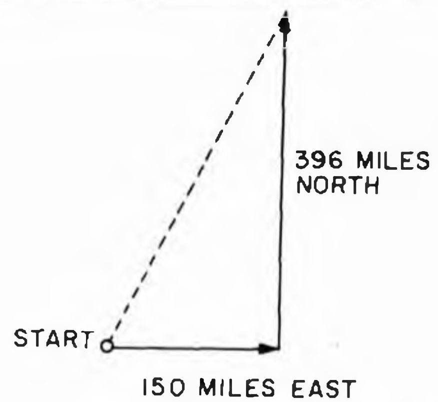


Fig. 6. Adding traveling distances in different directions yields resultant which is the diagonal (the shortest distance between the two points), also a right triangle. Back to the Pythagorean theorem!

Just as in the inductive case, we can plug X_L right into the Ohm's law formula and have a way to calculate current or voltage for a capacitive circuit:

$$I = \frac{E}{X}$$

More Examples. As examples of how the above formulas can be used, let's calculate reactances and currents of our 6.8-μF capacitor at various frequencies at 60 hz,

$$X_C = \frac{1}{2\pi fC}$$

$$= \frac{1}{2 \times \pi \times 60H \times 6.8\pi F}$$

$$= 390.1 \text{ ohms.}$$

(Note that 6.8 *micro*Farads must be expressed as 6.8×10^{-6} Farads, since the formula is always written for C in *Farads*. Alternately, you could write 6.8 microFarads as 0.000,006,8 Farads, but that's a little more awkward to handle.)

At 30 Hz, a *lower* frequency, the capacitive reactance *increases:*

$$X_C = \frac{1}{2 \times \pi \times 30\ (6.8 \times 10^{-6})}$$

$$= 780.2 \text{ ohms,}$$

and at a *higher* frequency, say 70 Hz, the capacitive reactance becomes *less:*

$$X_C = \frac{1}{2 \times \pi \times 70\ (6.8 \times 10^{-6})}$$

$$= 334.4 \text{ ohms.}$$

The current, in any case, can be found by Ohm's law. At 70 Hz, the current is

$$\begin{aligned} I &= E/X, \\ &= 100 \text{ volts}/334.4 \text{ ohms}, \\ &= 0.299 \text{ ampere,} \end{aligned}$$

while at 30 Hz the current is:

$$\begin{aligned} E/X &= 100 \text{ volts}/780.2 \text{ ohms}, \\ &= 0.128 \text{ ampere.} \end{aligned}$$

So, in their way, both inductors and capacitors are obedient to Ohm's law. Their reactances—X_L and X_C, respectively—can replace the R in the familiar Ohm's law expression, $I = E/R$ enabling us to calculate current flow for AC circuits which include inductors or capacitors.

Adding Apples and Oranges. But, what of circuits which *combine* an inductor and a resistor? Or an inductor and a capacitor? Can we simply add the two 'ohms' together? Do resistance and reactance add like apples and appels?

Unfortunately, this is not the case. Consider, for example, the circuit of Fig. 5. Here, our 1.05-Henry inductor and a 150-ohm resistor are connected in series across the 100-volt, 60-Hz source. What is the current flow?

Obviously, we need an Ohm's-law-like expression—something like: I = E, divided by the apparent "resistance" of R and L, combined.

We know from our previous work that a 1.05-Henry inductor has an X inductive reactance of about 396 ohms. We wish that we could simply add the 396 ohms of reactance to the 150 ohms of resistance to get 546 ohms—but reactance and resistance simply don't add that way. *Instead, they add at right angles!* And how, you ask, does one *add* at *right angles?*

A Journey at Right Angles. To understand how resistance and reactance can add at right angles, consider another type of problem: If I traveled 150 miles due east, and then 396 miles north, how far am I from my starting point? See Fig. 6.

You obviously cannot get the answer to this problem by adding 396 to 150, because you're certainly not 546 miles from home. But notice that the figure is a right triangle—a shape which that old Greek, Pythagoras, solved long ago. He said that if you *square* 150, and *square* 396, add the squares, and then take the square root, you will get the length (L) of the longest side:

$$L = (150 \times 150) + (396 \times 396),$$

$$L = 179{,}316,$$

$$L = 432.3 \text{ miles.}$$

This, then, is a way of adding two quantities that act at right angles to each other. Since inductive reactance is a kind of "north-bound resistance" operating at right angles to the "east bound resistance" of an ordinary resistor, we combine the two by adding a right angles, just as the two right-angle distances were added in the

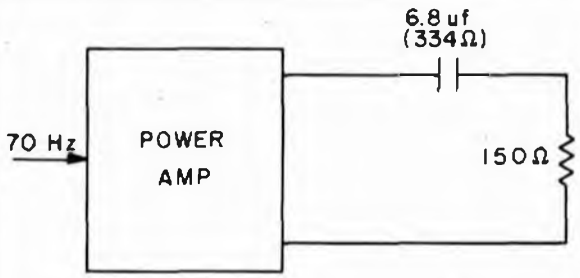


Fig. 7. The reactance of the capacitor at this frequency (70 Hz) is 334 ohms. This may be combined with the pure resistance, 150 ohms, by using the formula discussed in the text of this article.

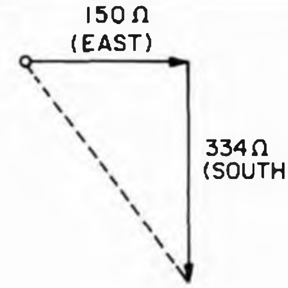


Fig. 8. This drawing shows how the two reactances are added, as also shown in Fig. 6 and, described in the text.

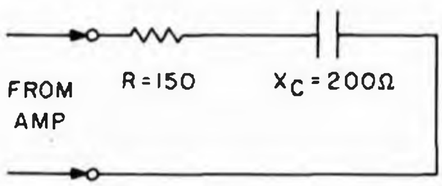


Fig. 9. Schematic showing reactance of capacitor. The two reactances add according to the formula, to yield a combined reactance (at the particular frequency used) of 250 ohms.

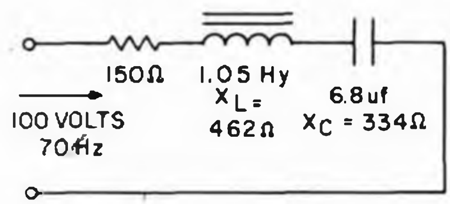


Fig. 10. Combining capacitive reactance and inductive reactance is also handled. Reactance of each at 70 Hz is computed, they are added together, and the difference is put into the formula with the pure resistance value.

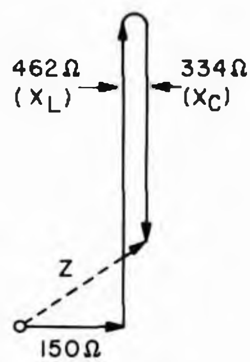


Fig. 11. The relative "directions" of capacitive and inductive reactances along with that of the pure resistance. Z (the diagonal) is the resultant of the 150-ohms resistance plus the difference between the two reactances.

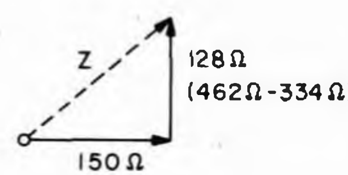


Fig. 12. Combining the reactances leaves a difference of 128 ohms and provides a directional diagram similar to that of Fig. 6. Such directional drawings are called "vector" diagrams.

above mileage problem. The effective resistance of the 396-ohm reactance and the 150-ohm resistor is therefore:

$$= (150 \times 150) + (396 \times 396),$$
$$= 432.4 \text{ ohms.}$$

What do we call this combination? It's obviously neither resistance nor reactance, but a combination of both. Since it represents a *general* way that a circuit can impede the flow of electrons, it is called *impedance,* and is represented by the symbol Z. The general formula for impedance is, therefore:

$$Z = R^2 \times X^2$$

Another Candidate for Ohm's Law. Impedance is measured in *ohms,* just as are reactance and resistance. If you know the impedance and voltage in a circuit, we can find the current I by plugging Z into the familiar Ohms Law expression:

$$I = E/Z,$$
$$= E/ \quad R^2 + X^2.$$

Using this expression we can calculate the current in the example of Fig. 5—

$$I = E/Z,$$
$$= 10 \text{ volts}/423.4 \text{ ohms},$$
$$= 0.326 \text{ ampere}.$$

Southbound Capacitors. The knowledge that inductive reactance adds at right angles to resistance leads immediately to the question, "What about *capacitive* reactance? Does it also act as a 'northbound' resistance?"

As you might except, two circuit elements as different as inductance and capacitance could never agree on the 'direction' of their reactances. If inductive reactance is 'northbound', the capacitor obstinately declares that *its* reactance is *'southbound'*, acting in direct opposition to the inductive direction, as shown in Figs. 7 and 8.

However, this makes no difference in the equation for impedance, which is still given by

$$Z = \quad R^2 + X^2,$$

where X is, in this case, X_C, capacitive reactance.

As an example, let's calculate the impedance of the circuit of Fig. 9.

$$\begin{aligned} Z &= \quad R^2 + X^2 \\ &= \quad (150)^2 + (200)^2 \\ &= \quad 62{,}500 = 250 \text{ ohms} \end{aligned}$$

As a further calculation, let us determine the current which will flow as 50 volts is applied to the terminals of the circuit in Fig. 10:

$$\begin{aligned} I &= E/Z, \\ &= 50/250 \\ &= 0.2 \text{ ampere.} \end{aligned}$$

All Together, Now. Thefinal queston is this: If a circuit includes resistance, inductance and capacitance, all at once, what is its total impedance?

An example of such a series circuit is shown in Fig. 10. Here, 100 volts at 70 Hz is applied across a series circuit of 150 ohms resistance, 1.05 Henry inductance (X_L = 462 ohms), and 6.8 μF capacitance (X_C = 334 ohms).

As a first step in determining the total impedance, we can draw the 'directions' of the resistance and the two reactances, as shown in Fig. 11.

The first thing that strikes us about Fig. 11 is that the effect of the 'northbound' (inductive) 462 ohms is partially cancelled by the 'southbound' (capacitive) 334 ohms. The *net reactance* of this circuit is:

$$\begin{aligned} X &= X_L - X_C, \\ &= 462 \text{ ohms} - 334 \text{ ohms}, \\ &= 128 \text{ ohms.} \end{aligned}$$

This combining of X_L and X_C by simple subtraction reduces the problem to a simpler one, resembling a problem we've already solved. This is graphically shown in Fig. 12.

Applying the formula for impedance

$$\begin{aligned} Z &= \quad X^2 + R^2, \\ &= \quad (128)^2 + (150)^2, \\ &= \quad 197.2 \text{ ohms.} \end{aligned}$$

The current flow is:

$$\begin{aligned} I &= E/Z, \\ &= 100 \text{ volts}/197.2 \text{ ohms}, \\ &= 0.51 \text{ ampere.} \end{aligned}$$

North vs. South. In the above problem, we saw that the 'northbound' inductive reactance (462 ohms) was very nearly cancelled by the 'southbound' capacitive reactance (334 ohms). It's pretty obvious that by changing some values, or by changing the frequency, we could make X_L *exactly equal to* X_C, and they would cancel each other completely, leaving the circuit with no net reactance whatsoever—just 150 ohms of simple resistance. This condition, where $X_L = X_C$ exactly, is called resonance. When a capacitance and inductance are in series there will be one frequency for which their net reactance will be zero. A parallel capacitance and inductance present a high impedance to a current. By using these circuits you can make filters that will pass or block certain frequencies. But that's another story altogether. ■

Frequency Fundamentals

It was 4:00 A.M. and the thermometer read 15 degrees below zero. The squad car had to deliver the package without delay! Thanks to the flashing red light and the 2-kHz note screaming from the siren, the three miles from the bus terminal to police headquarters were covered in only 2½ minutes. The chief rushed out to get the package. He likes his coffee hot.

So the siren gave out a 2-kHz note. Is that anything like say, the 60-Hz current which lights a table lamp or, maybe, the 27.155-MHz carrier a CB'er sends out from his 5-watt rig? The answer is yes—and no! No, because the siren note is actually a *disturbance of the air* which surrounds the whirling siren. The 60-Hz current which lights the lamp is actually a *disturbance of electrons* in the lamp cord. And the CB carrier is actually a *disturbance in the electro-magnetic field* which surrounds the transmitting antenna. But the answer is also yes, because, in spite of their apparent differences, the siren note, the 60-Hz current, and the 27.155-HMz signal all have something in common—the characteristic way each of these "disturbances" go through their vibrations.

Bouncing Air. If analyzed scientifically, each type of disturbance is seen to be a repetition of identical actions. When the siren gives out a constant howl, tiny air particles are forced to collide and spread apart—in steady rhythm. If the howl is a 2-kHz note, the colliding and spreading takes place 2000 times a second. Although the individual air particles move hardly any distance at all in their back and forth motion, the rhythmic collisions of these particles radiate from the siren in all directions, from particle to particle. This radiation is, of course, "sound" and travels through air at a speed of about 1100 feet per second.

Bouncing Electrons. Alternating current flows back and forth through a wire because electrons, in varying quantities, are made to push one another first in one direction and then the other—in steady rhythm. If this

AC is 60-Hz current, such back and for th motion takes place 60 times a second. The individual electrons don't move very far along the wire in either direction but their "bumping" travels through the wire at a speed of almost 186,000 miles per second. Naturally, this bumping action is made to change directions in step witht he electrons that cause it.

Bouncing Fields. Basically, a radio signal is a disturbance in which the electric and magnetic fields surrounding a transmitting antenna are distorted, first in one direction and then the other—in steady rhythm. If this signal is a 27.155-MHz CB carrier, these fields are forced to change direction 27,155,000 times a second. This rhythmic, field-reversing action radiates from the antenna at the speed of light, 186,000 miles per second.

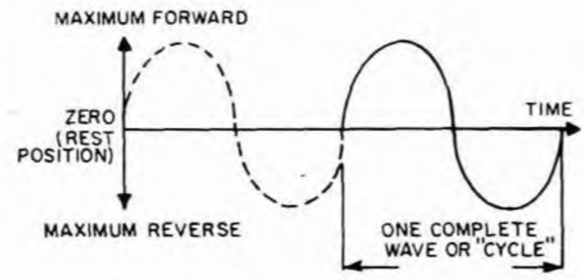
Feel the Vibrations. Now—it's the changing nature of these three types of disturbances that we are really interested in! Aside from what is actually and physically happening, all three seem to follow the same pattern of change while going through their vibrations. True, the 27.155 MHz signal does its vibrating much, much faster than the other two but its vibration pattern is much the same.

Instead of using a lot of words to describe how each disturbance goes through its vibrations, let's use a helpful mathematical tool—the graph. The ancient Chinese said that *a picture is worth a thousand words* and that's exactly what a graph is and it's worth. More important, a graph provides us with a *lasting* picture, a permanent record, of how these vibrations change their speed and direction.

Our graph is set up in the usual manner. First there are the two reference lines, the "vertical axis" and the "horizontal axis." In our graph, the horizontal axis is made to show the *passage of time*. How the vertical axis is used depends on which type of disturbance we are portraying.For the 2-kHz siren note, the vertical axis is made to measure *how far* each air particle moves during its back-and-forth motion. For the60-Hz AC current, the vertical axis measures *how many* electrons are in mass movement along the wire during their forward-and-reverse motion. As for the 27.155-MHz signal, this axis tells *how much* and in what direction the electric and magnetic fields are being distorted.

So much for the preliminaries. Next comes the most important item. It doesn't matter which disturbance we are graphing—for each *instant of time* that is represented on the horizontal axis there is a *point* above or below the axis which measures how far the vibration is displaced from its resting position. If all these individual points (and there are an unlimited number of them) are "plotted" on the graph, a continuous curve will appear. The amazing thing about this curve is that its shape looks pretty much the same for the 2-kHz note, the 60-Hz current and the CB carrier.

Study the special shape of the curve. While examining the curve (and this must be done from left to right, never from right to left!), notice that it rises abruptly, gradually tapers off, levels out for an instant, starts to fall, picks up downward speed, and falls below the horizontal axis only to repeat the process in upside-down fashion. The 2-kHz note will make 2000 of these waves in just one second while the 60-Hz current goes through only 60 of



Whether you're trying to graphically picture ripples on a sylvan pond, or the unmodulated carrier of a Citizens Band RF voltage, the familiar sine wave is all you need to draw. All sine waves have peaks culminating in "maximum forward" points and "maximum negative" points, each occurring once each cycle.

them. But the CB carrier will turn out 27,155,000 of these double swings during each second.

You Had Better Believe. This curve we have been discussing is something very special! It is fundamental to many different areas of science and engineering. It is the famous *sine curve*. The current leaving a simple AC generator varies according to a sine curve. So does the output of a vacuum-tube or transistor osciallator in an AM radio. Even the pendulum on a grandfather's clock swings back and forth in accordance with a sine-curve pattern.

But it must be understood and understood well! The sine curve is only a graph which shows the changing character of sound waves, alternating currents, and radio transmissions, etc. To be sure, a sine curve is a picture but it *is not* a picture of sound waves, currents and radio signals as they really look (if, in some way, we could actually see them).

Ironically (and students of electronics must not be misled because of this), there are sine-curve relationships in other fields of science where the physical occurence actually looks like a sine curve. If a pebble is dropped into a calm pond, for example, ever-enlarging circles (ripples) will be seen leaving the splash. However, if these ripples are viewed from the side, at water level, a train of honest-to-goodness sine curves will be observed as they leave the splash area. As another example of real-life sine curves, if two persons suspend a rope between each other and one of them whips it up and down, a beautiful sine wave will run along the rope. Nevertheless, sine curves *don't* travel through electrical circuits, they *don't* travel through air as sound waves, and they *don't* radiate from transmitting antennas! Radio signals, AC currents and sounds vibrate in unison with sine curves but they *don't* look like them! However, a check of their graphs indicate they do!

Discovering Frequency. At this point, the true meaning of "frequency" should become clear. It doesn't matter if we are talking about sound waves, AC current, or radio carriers. Frequency is the number of times these various disturbances repeat their vibrations during *one second*. From a graphical standpoint, frequency is the number of times *each* complete cycle of sine curve is repeated during *one second*. Either way you look at it, frequency can be classified by its number of "cycles per second." As everyone should now know, however, this once common designation for frequency has been replaced by, simply, "hertz" (Hz) in honor of that great German scientist of the Nineteenth Century, Heinrich Hertz.

To Sum Up. The 2000-Hz sound energy from the police-car siren, the 60-Hz AC from the wall outlet, and the 27,155,000-Hz carrier from the CB rig are three different scientific phenomena. But they have the common property of being able to be described by sine curves and, because of this, all can be measured by a common yardstick—frequency.

But hold on! Are these three so-called disturbances really so different from each other? (After all, electronics wouldn't be electronics without sound and radio signals!) They are different, but not as much as you would think. With the aid of a "transducer," one type can be *transformed* into another! Thus, a *microphone* will change 2-kHz sound waves into 2-kHz alternating current in a wire. A *loudspeaker* will change the AC back into sound waves. A *receiving antenna* will change 27.155-MHz electromagnetic energy into 27.155-MHz alternating current (AC)in a wire (the antenna feedline). A *transmitting antenna* will make the opposite change.

That's it. If the interrelation between sound, electrical, and electromagnetic frequencies now makes sense to you, you've learned a tremendously important bit of electronics theory. And don't forget the almighty sine curve. The sine curve can be used to explain theory in many fields of science, not just electronics. Nevertheless, keep clear in your mind just how the sine curve fits into electronics—what it is and what it isn't. Maybe you don't care whether your coffee is hot but you better stay *hot* on the *sine curve*. ■

Amplifiers That Oscillate

As any slightly cynical experimenter can tell you, if you want an oscillator, build an amplifier—it's sure to oscillate. Conversely, if you want an amplifier, (this same cynic will tell you), build an oscillator—it's sure to fail to oscillate, and you can then use it as an amplifier! This is well known as a corollary to Murphy's famous law, "If anything *can* go wrong—it *will!*"

Our informed cynic must have had long and unhappy experience with negative-feedback amplifiers, which are known to have at least two outstanding characteristics:

1. They function beautifully if carefully designed and built.
2. Otherwise, they oscillate!

Why do they oscillate? Or, more basically, how does a feedback amplifier differ from an oscillator?

The fundamental block diagrams of an oscillator and an amplifier with feedback bear a strong resemblance to each other, as you can see from Fig. 1. From a block diagram viewpoint, both diagrams are *very* similar. Both contain some type of amplifying device, and both have part of their output signal fed back to their input. There are only two major differences between them:

1. The amplifier with feedback contains an *inverting* amplifier; the oscillator contains a *non*-inverting amplifier.
2. The oscillator doesn't have an input.

The circuit action obtained from these two circuits is entirely different. In the amplifier with feedback, the output waveform is upside down with respect to the input, so when it is fed back to the amplifier input, it cancels a portion of the input waveform. The output is therefore less than it would be without feedback. See Fig. 2.

The feedback signals from inverting amplifiers are not "in phase" with the input signal and subtract (or reduce) the input signal level to the amplifier. When a feedback signal does this, it is called *negative feedback*.

So Why Negative Feedback? Of course, if you merely want the biggest possible gain for your money, negative feedback's not your game. However, negative feedback offers other advantages, which can be summed up by saying that the amplifier's output, though smaller, is always nearly constant for the same input signal. For example, if the amplifier weakens with age, and the output tries to drop, there is less signal to be fed back; hence there is less cancellation, and the output is restored

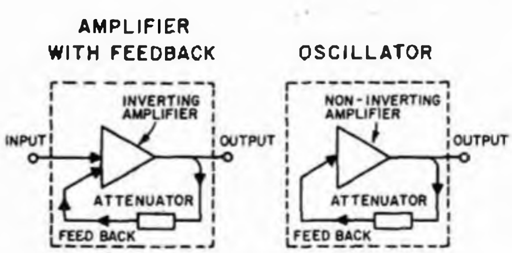


Fig. 1. Two block diagrams compare an inverting amplifier (A) to an oscillator (B).

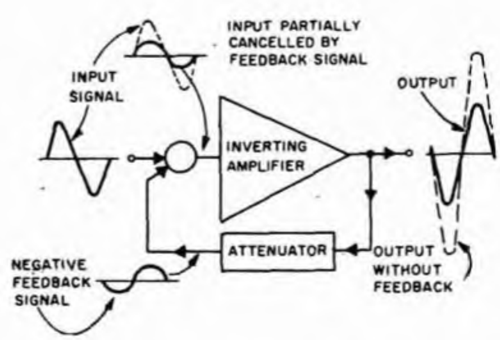


Fig. 2. Signal flow diagram illustrates how feedback in an inverting amplifier circuit reduces overall gain.

almost to its former level. Similarly, if you feed a high-frequency signal through the amplifier—so high in frequency that the amplifier can barely amplify it—the resultant drop in output reduces the fed-back voltage, produces almost no cancelling feedback signal, and keeps the output nearly the same as it was at lower frequencies. Moreover, any clipping or other distortion of the waveform inside the amplifier produces an output waveform which does not match the input; hence the non-matching part is not cancelled, and the distortion is removed, or at least greatly reduced. Without this action, hi-fi amplifiers would not exist.

So the loss in output you obtain from negative feedback repays you by providing less distortion, better long-term stability, and better frequency response—that is, the best and most uniform output in response to *all* input frequencies.

On the Flip Side. The oscillator, on the other hand, is not supposed to give the best output from all input frequencies, but is instead made to give an output at a *single* frequency—with no input at all. It's not surprising that the opposite type of internal amplifier (non-inverting) is used to obtain this opposite result. See Fig. 3.

In the oscillator, any output at all (probably the result of some random noise in the internal amplifying device) is fed back, non-inverted, to the input, where it does not cancel but instead serves as the signal at the input. This feedback signal causes an even larger output, which results in an even larger signal fed back, further reinforcing the input signal, and so on.

You guessed it—this type of feedback signal is commonly referred to as *positive feedback*. In theory, the output waveform should continue to get larger forever. In practice, the amplifier is limited in the maximum size of the signal it can deliver, so the output waveform stops growing in this amplitude. As it stops to grow, so does the positive feedback signal. Now the signal reduces rapidly and the positive feedback signal lends a hand until the signal can get no lower. This is the beginning of the first cycle of many to follow.

All well and good, you say, but if the major difference between feedback amplifiers and oscillators is the inverting or non-inverting nature of their internal amplifiers, why does an amplifier sometimes oscillate? What turns an inverting amplifier into a non-inverting one?

To answer this question, first observe that an inverting amplifier, in passing a sine-wave signal, *effectively* shifts the signal's phase by 180° as shown in Fig. 4. We say *effectively*, because it doesn't really shift the timing by delaying the signal (which is what a real phase-shifter does) but, by turning the signal upside down, the amplifier makes it look like a signal which has been delayed (phase-shifted) by 180°.

A real phase-shifter, on the other hand, is normally nothing but a fistful of judiciously connected resistors and capacitors (and sometimes inductors) which can be designed to give a 180° phase shift *at a single frequency*, such as 1,000 Hz, for example. In contrast to an inverting amplifier, it provides this phase shift by actually delaying the signal. See Fig. 5.

What happens if we combine an inverting amplifier and a 180° phase-shifter? Take a look at Fig. 6.

This combination will shift the phase of a given frequency by a total of 360° (an entire cycle) so the output is identical to the input. In effect, this combination (at 1,000 Hz) *will behave the same as a non-inverting amplifier*. See Fig. 6.

Therefore, if we build a feedback amplifier which contains the normal inverting amplifier but also (inadvertently) contains a 180° phase-shift network, the resultant circuit will oscillate at the particular frequency, (1,000 Hz in the figure) for which the phase-shifter provides 180° phase shift. See Fig. 7.

How can one "inadvertently" make a phase-shifter? It's easier than you might think. The circuit shown in Fig. 8A will provide 60° phase shift at 1,000 Hz. Three such networks connected in a "ladder" (see Fig. 8B) will proved 3 x 60° = 180° of phase shift. (But not at 1,000 Hz. Because of the way the networks load each other, the 180° shift occurs at 707 Hz. However, if an amplifier were located between each network, then the amplifier will oscillate at 1,000 Hz.) This network, if dropped into a normal feedback amplifier circuit, will convert it to an oscillator.

This circuit (Fig. 9) is known as a *phase-shift oscillator* and is widely used in electronics.

Of course, when you set out to build a phase-shift oscillator, you *deliberately* insert a phase-shifter to make the circuit oscillate. How could one ever *inadvertently* place such a circuit in a feedback amplifier, thereby producing unwanted oscillations?

Phase-shift circuits can "hide" within an amplifier, posing as other circuits. For example, vacuum-type amplifiers often have grid circuits arranged as shown in Fig. 10A. Does that resistor/capacitor circuit look familiar? In form, it's just like the phase-shift circuit above. And transistor amplifier circuits often take the

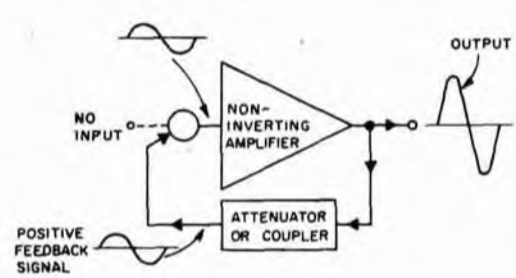


Fig. 3. Signal flow diagram shows how feedback signal from a non-inverting amplifier provides the positive feedback to produce an output signal larger than it would otherwise be without feedback.

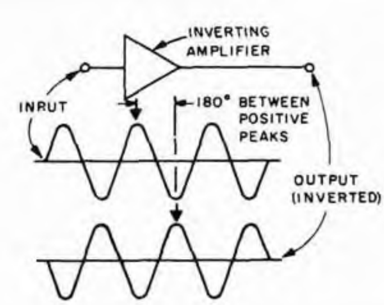


Fig. 4. The apparent phase shift of the output signal from the input signal is 180 degrees.

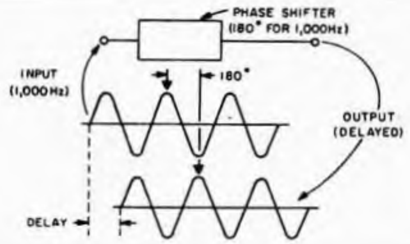


Fig. 5. The passive phase-shift circuit actually delays the output signal by a time interval measured in degrees by that portion of a sine wave so delayed. This effect appears to the apparent delay of an inverting amplifier.

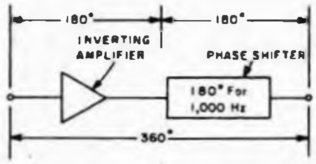


Fig. 6. Combining an inverting amplifier and 180-degree phase-shift circuit results in a 360-degree phase shift at 1000 Hertz only, which results in a non-inverting amplifier circuit.

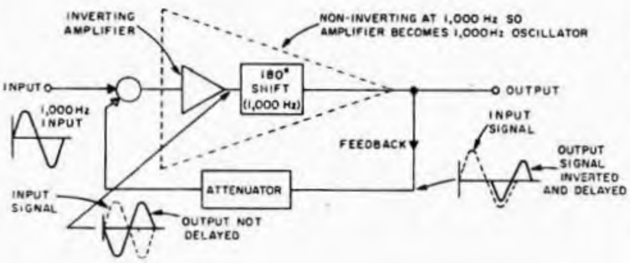


Fig. 7. Since the 180-degree phase shift is frequency selective, this non-inverting amplifier will oscillate at 1000 Hertz.

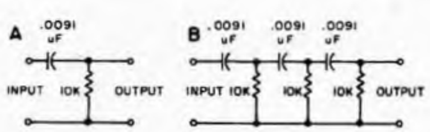


Fig. 8. In this diagram, a simple network (A) offers 60-degrees of phase shift at 1000 Hertz. Ladder three such circuits in series and the total phase shift at 1000 Hertz will be 180 degrees.

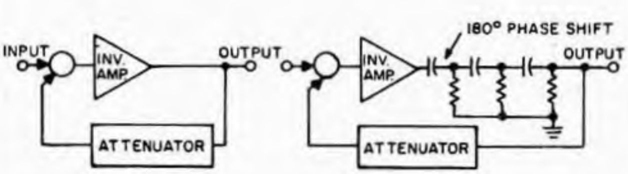


Fig. 9. Two inverter amplifiers are shown here with one having a 180-degree phase-shift network added to induce positive feedback. The RC elements limit this oscillation to be a fixed frequency.

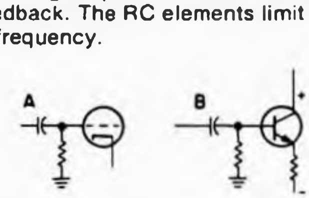


Fig. 10. External parts in this amplifier circuit have the same effect as the phase-shift circuit in Fig. 8A. However, values for the resistance and capacitance are selected to produce almost no phase-shift within the amplifiers normal frequency bandpass.

form shown in Fig. 10B.

Again, the coupling/biasing network looks just like the basic phase-shifter network. At some frequency, this network will provide 60° of phase shift. If we use three such identical networks in a three-stage amplifier we have a 180° phase-shift network "buried" inside the amplifier, masquerading as three normal coupling networks. If this three-stage amplifier is used as part of a feedback amplifier arrangement, the amplifier will oscillate at some frequency, and be quite useless for the purpose for which it was intended.

More Trouble. This is not the only way an amplifier can get into trouble. There are other types of phase-shifters that can creep into amplifiers, unrecognized, and drive the unwary experimenter up the nearest wall. This circuit (shown in Fig. 11A) can also produce a phase-shift of 60° at 1,000 Hz. Three of them, can produce the 180° phase-shift required for oscillation. See Fig. 11B. This particular network can invade amplifiers in an even more insidious fashion. The "masquerading" part of the circuit is shown heavy in Fig 11C. The dotted capacitor doesn't appear physically in the circuit, because it is the so-called "stray capacity" associated with wires, sockets, terminals, etc. Three of these circuits hiding in an amplifier, can produce an unwanted oscillation. See Fig. 12. Since the stray capacities are so small, this "oscplifier" will oscillate at a very high frequency; often so high that it is undetected as an oscillation. However, such oscillation can make an amplifier behave erratically; sometimes distorting, sometimes not; sometimes overheating, sometimes not. Fig. 7 and Fig. 12 have a lot in common.

Are feedback amplifiers the only culprits in this oscillating-amplifier business? Absolutely not! Often, so-called "straight" amplifiers—with no *intentional* feedback—will gaily oscillate away. But watch that word *intentional.* Close inspection of these misbehaving circuits usually uncovers an *unintentional* feedback path hiding within the amplifier. Consider the innocent-looking circuit in Fig. 13. This is an ordinary two-stage amplifier, obviously assigned the task of converting a small, positive-going signal into a large, positive-going signal. To help it along, the designer has even provided a decoupling network, R1 and C1. At high frequencies, C1 acts like a short circuit, effectively isolating (decoupling) the amplifier's power bus, Ecc +, from the main power bus, Ecc + +. But, at low frequencies, the capacitor acts like an open circuit—it just isn't there! A small part of the output voltage now appears across R1, and is coupled through the amplifier's power bus back to the input, arriving there with the same polarity as the normal input. True, the signal unintentionally fed back isn't very large, because the unintentional feedback path provides substantial losses for this stray signal. For example, the signal may arrive back at the input 100 times smaller than it was at the output. However, if the amplifier has a gain of 101, it makes an even larger output signal out of the fed-back signal, which then is fed back as an even larger voltage and oscillation begins.

Careless construction can get you into trouble, too. The amplifier shown in Fig. 14 is trying to convert a 10-millivolt input into a 200-ma signal needed by the load, R2. The builder has tied all ground returns to a heavy ground bus, and returned this bus to ground at only one point. Unfortunately, that single ground wire has to cary both the tiny input signal and the large output current. And, since every wire has *some* resistance, the actual circuit includes an 0.06-ohm resistor that does not appear in the original construction schematic diagram, but must be considered and is shown in Fig. 14.

Again, an uninvited, unintended feedback path has appeared, coupling the output back to the input. In the sketch, the large output current, flowing through the tiny ground-lead resistance, produces a voltage which is even larger than the original input voltage. And, since this voltage is also connected to the input (through the bias resistor R3), the fed-back voltage appears uninverted (and uninvited!) at the input, and will cause the amplifier to oscillate. What is to be done to convert these oscillators back into well-behaved amplifiers?

The general rule is *divide* and *conquer*. In the example just above, we can conquer the oscillation by dividing the ground returns, making sure that the high-current output circuits and the sensitive input circuits have their own private and individual paths to the power supply. See Fig. 15.

Short leads to the input connector are also helpful in squashing oscillations.

The misbehaving decoupling network, R1 and C2 in Fig. 13, can also be brought under control by dividing the network into *two* decoupling networks as shown in Fig. 16.

The stray capacities causing the unwelcome phase shifter to hide in the best divide-and-conquer approach to this amplifier are harder to exorcise. Your amplifier is to put feedback around only a pair of stages instead of three or more. This way, there are only two pairs of stray Rs and Cs lurking in the amplifier, and it takes at least three such pairs to make an oscillator.

The coupling capacitors, which combined to make a phase shifter in the very first example, can be prevented from ganging up on the amplifier and making it oscillate by making the product of each capacitor times its associated resistor (called the "RC product") 5 or 10 times larger or smaller than the other RC products. For example, in a three-stage feedback amplifier which has all its base resistors the same values, you could make the three coupling capacitors 2 uF, and 10 uF, and 60 uF, respectively. Again, you have divided the coupling capacitors into three widely-separated values, and conquered the oscillation.

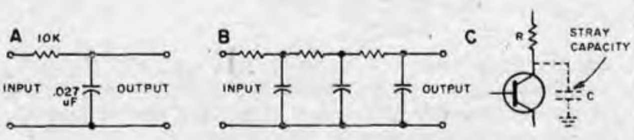


Fig. 11. Don't get confused with the circuit elements shown in (A) with those shown in Fig. 8A. The reversal of parts converts the basic circuit from a coupler to a filter with the attending phase shift. Three series-connected filters have the same effect (B) by providing 180-degree phase shift. Hidden capacitance (C) internal to the circuit element and in the external wiring will provide some phase shift of the output signal in this amplifier.

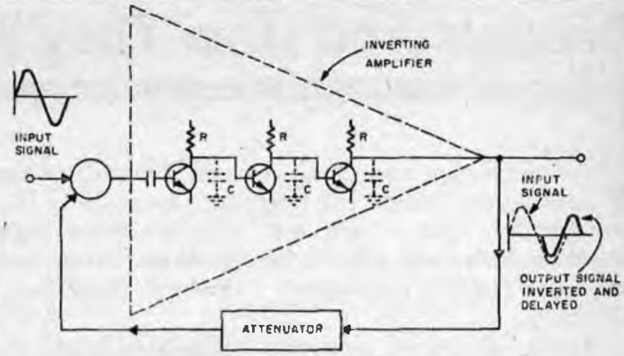


Fig. 12. An inverting amplifier may oscillate do to stray capacitance in these external circuit amplifiers providing a 180-degree phase shift at some very high frequency. Output signal is within frequency bandpass of the amplifier before phase shift causes oscillation (positive feedback).

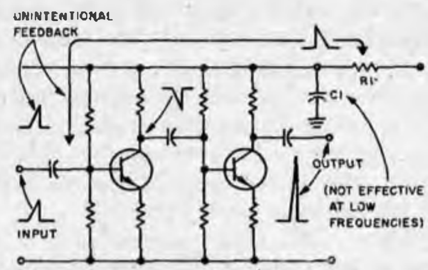


Fig. 13. Low frequency feedback occurs because the capacitor C1 is not effective at these low frequencies. Thus, any portion of the output signal that will cause a ripple in the power supply will serve as a positive feedback.

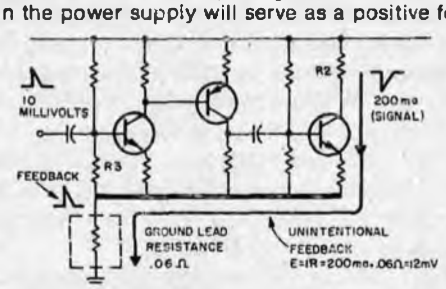


Fig. 14. A ground bus has a finite resistance, and, when high-current output signals mix with low-level signals in the same bus, positive feedback may result with the attending possibility of oscillation.

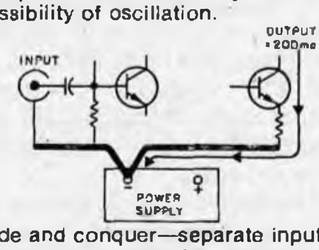


Fig. 15. Divide and conquer—separate input connections from output connections to different buses, and a good deal of the unwanted signal mixing will not occur. The resistance of the power supply must be very low otherwise all that is gained using this construction technique will be lost.

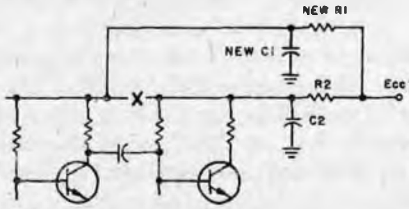


Fig. 16. A good idea is to split up the voltage distribution to many circuit points by two or more power supply decoupling networks. Compare this diagram to Fig. 13.

Crystals And How They Work

Can you imagine the chaos on the AM broadcast band if transmitters drifted off frequency as much as those inexpensive table radios do? The broadcast station engineer has a tough job. He must keep his station carrier within 20 Hertz of its assigned frequency. How does he do it?

This is, of course, done with a little help from a very basic material, the quartz crystal and some supportive circuitry. The crystal is the single component that serves to fill a basic requirement for precision frequency control. Quartz crystals not only fix the frequency of radio transmitter carriers (from CB installations to multi-kilowatt-broadcast installations), but also establish the frequency of timing pulses in many modern computers. In addition, they can provide the exceptional selectivity required to generate and receive single-sideband signals in today's crowded radio spectrum. Yet this list merely touches upon the many uses of quartz crystals. No exhaustive list has ever been compiled.

A Real Gem. This quiet controller is a substance surrounded by paradox. While quartz composes more than a third of the Earth's crust, it was one of the three most strategic minerals during World War II. And despite its plenitude, several semiprecious gems (including agate and onyx) are composes only of quartz.

Unfortunately, quartz exercises its control in only a relative manner. When it's misused, the control can easily be lost. For this reason, if you use it in any way—either in your CB rig, your ham station, or your SWL receiver—you should become acquainted with the way in which this quiet controller functions. Only then can you be sure of obtaining its maximum benefits.

What Is It? One of the best starting points for a study of quartz crystals is to examine quartz itself. The mineral, silicon dioxide (SiO^2), occurs in two broad groups of mineral forms: crystalline and non-crystalline. Only the large crystalline form of quartz is of use as a controller.

The crystalline group has many varieties, one of which is common sand. The variety which is used for control, however, is a large, single crystal, usually six-sided. The leading source of this type of quartz is Brazil. However, it also is found in Arkansas.

A property of crystalline quartz, the one which makes it of special use for control, is known as *piezoelectricity*. Many other crystals, both natural and synthetic, also have this property. However, none of them also have the hardness of quartz. To see why hardness and the piezoelectric property, when combined, make quartz so important, we must take a slight detour and briefly examine the idea of resonance and resonators.

Resonators and Resonance. As physicists developed the science of radio (the basis for modern electronics), they borrowed the acoustic notion of resonance and applied it to electrical circuits where it shapes electrical waves in a manner similar to an acoustic resonator. For instance, both coils and capacitors store energy and can be connected as a resonator (more often termed a resonant circuit). When AC of appropriate frequency is applied to the resonator, special things happen.

Pendulum Demonstrates. The principle involved is identical to that of a pendulum, which is itself a resonator closely similar in operation to our quartz crystals. To try it you can hang a pendulum of any arbitrary length (Fig. 1), start it swinging, then time its *period*—one complete swing or cycle. The number of such swings accomplished in exactly one second is the *natural* or *resonant* frequency of the pendulum in cycles per second (hertz).

You can, by experiment, prove that the frequency at which the pendulum swings or oscillates is determined by the length of the pendulum. The shorter the pendulum (Fig. 2), the faster it swings (the greater the frequency). The weight of the pendulum has no effect on

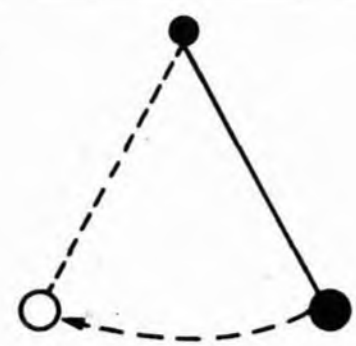


Fig. 1. A long pendulum swings at a rate determined by its length and weight at the end of the arm. To experiment at home, the arm should be made of light thread (no weight) and a small heavy mass (all the weight concentrated at one point).

Fig. 2. A short pendulum swings at a faster rate no matter what the weight is. The heavier the weight, the longer the pendulum will swing at the faster rate.

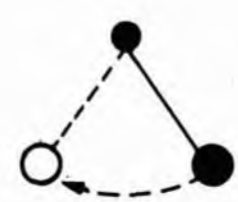


Fig. 3. The geometry of the tuning fork produces a continuous pure tone for a long time when struck gently. It is favored by piano tuners as a convenient way to carry a source of accurately-known tone or pitch.

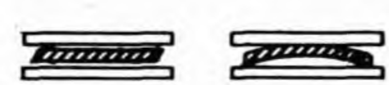


Fig. 4. Depending on how the crystal is cut or sliced from the mother crystal, some crystals warp or bend as in A and others shear as shown in B. Each of the many cuts have special characteristics.

frequency, but has a marked effect upon the length of time the pendulum will swing after a single initial push in the atmosphere: the heavier the pendulum, the greater the number of cycles back and forth before it stops.

A Real Swinger. Once the pendulum begins to swing, very little effort is required to keep it swinging. Only a tiny push is needed each cycle, provided that the push is always applied just as the pendulum begins to move *away* from the pushing point. If the push is given too soon, it will interfere with the swinging and actually cause the swing to stop sooner than it would without added energy: while if too late, added push will have virtually no effect at all. It all becomes simple when you recall your last experience "pushing" a swing with a child on it!

It is this principle—a tiny push at exactly the right time interval—which makes a resonator sustain sound or alternative-current waves. You can prove it with the pendulum by first determining the resonant frequency of a pendulum, then stopping it so that it is completely still. A series of small pushes, delivered at the natural resonant frequency, (each too tiny to have more than a minute effect) will very rapidly cause the pendulum to swing to its full arc again. Pushes of the same strength at any other frequency will have little or no effect.

The pendulum is an excellent control mechanism for regulating a clock to keep time to the second, since the resonant frequency of the pendulum can readily be adjusted to be precisely one cycle per second. However, for conrol of audio frequencies from tens of Hertz (cycles-per-second) up to tens of thousands of cycles-per-second (kilo-Hertz), or for radio frequencies ranging up to hundreds of millions of cycles per second (megaHertz), the pendulum is too cumbersome a device.

The Tuning Fork. In the audio range, the equivalent of the pendulum is the tuning fork. This is an extremely elongated U-shaped piece of metal (Fig. 3), usually with a small handle at the base. When struck, it emits a single musical tone.

The operating principle is exactly the same as the pendulum. Each of the arms or tines of the fork corresponds to a pendulum arm. But here the arms are extremely short and much heavier in proportion to their size than the pendulum. (The shorter the arms of a tuning fork, the higher the resonant frequency in the audio range.) This greatly increased mass causes the arms of the tuning fork to oscillate much longer when struck.

Not all tuning forks operate precisely like pendulums. The pendulum principle is based on a *flexing* of the arm upon its long dimension. While this is the most common operation, the fork may flex along any dimension.

It's even possible for a single, solid resonator such as a tuning fork to flex along several dimensions at once. A main part of the design of a good tuning fork is to insure that only a single dimension flexes or, in the language of resonators, only a single mode is excited.

Area Too. There's no requirement that the resonator be a completely solid substance. A mass of air, suitably enclosed, forms a resonator. This is the resonator that works on a classic guitar or violin. Here, single-mode operation is distinctly *not* desired. Instead, multiple-mode operation is encouraged so that all musical tones within the range of the instrument will be reinforced equally.

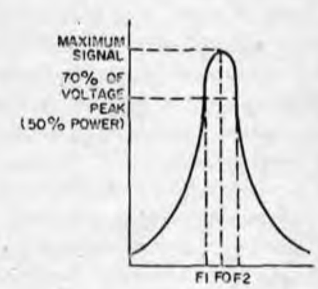
Now, with the principles of resonance firmly established, we can return to the quartz crystal and its operation.

Quartz Crystal as Resonators. Like the tuning fork or, for that matter, any sufficiently hard object, the quartz crystal is capable of oscillation when struck physically or in some other way excited.

But unlike the tuning fork, or indeed any other object except for certain extremely recent synthetic materials, the quartz crystal is not only sufficiently hard to oscillate at one or more resonant frequencies, but is piezoelectric.

Piezoelectricity. The piezoelectric property means simply that the crystal generates an electric voltage when

Fig. 5. Bandwidth characteristic of a typical tuned circuit shows the peak or maximum signal amplitude and the 70% voltage peak (50% power) points. This is the characteristic that determines overall selectivity.



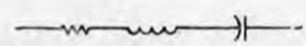


Fig. 6. Diagram is a typical equivalent circuit for a crystal. As resistance is lowered to near zero, crystal efficiency increases. In use, the crystal holder and external circuit add some capacitance across the entire circuit.

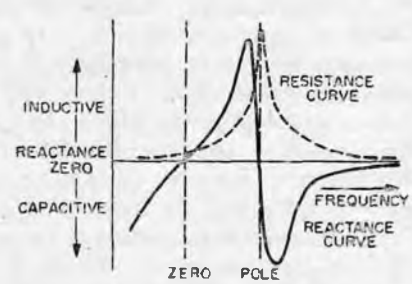
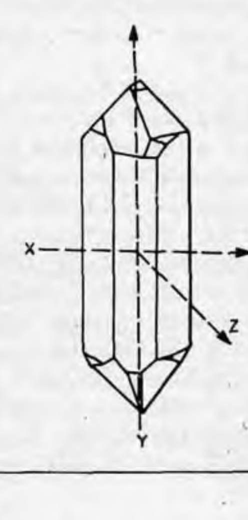


Fig. 7. Some characteristics of a quartz crystal. When slightly off resonant frequency (at the pole) a crystal exhibits inductive or capacitive reactance—just like an LC circuit.

Fig. 8. A view of a mother crystal showing X, Y, and Z axes. Crystal is sliced into blanks, ground to frequency, polished and plated (on facing sides) to make permanent electrodes. All crystals are not made perfect by Mother Nature and must be examined optically before being sliced.



physically stressed or, on the other hand, will be physically deformed when subjected to a voltage (see Fig. 4). Other familiar objects making use of piezoelectricity include crystal and ceramic microphone elements and phonograph cartridges.

This virtually unique combination of properties (sufficient hardness for oscillation and piezoelectricity) found in quartz crystals, makes it possible to provide the initial push to the crystal by impressing a voltage across it. To provide the subsequent regular pushes, a voltage can be applied at appropriate instants.

Quality Factor. Almost any discussion of resonance and resonant circuits (or for that matter, inductance) evenutally gets to a rather sticky subject labelled in the earliest days of radio as *quality factor* but now known universally as Q.

As used in radio and electronics, Q is usually defined by other means. Some of the definitions put forth at various times and places include:

- *The ratio of resistance to reactance in a coil.*
- *The ratio of capacitive reactance in a resonant circuit to the load resistance.*
- *The impedance multipication factor, and others even more confusingly worded.*

All, however, come out in the end to be identical to the definitions cited above: The Q of a resonator is the ratio of the energy stored per cycle to the energy lost per cycle.

In a resonator, high Q is desirable. Q is a measure of this energy loss. The less energy lost, the greater the Q of the circuit.

Not so obvious (and rather difficult to prove without going into mathematics) are some of the other effects of Q. A resonant circuit is never completely selective; frequencies which are near resonance but not precisely equal to the resonant frequency pass through also!

An Interesting Fraction. The greater the Q, the narrower the band of frequencies which can affect the resonator. Specifically, the so-called half-power bandwidth (Fig. 5) of a resonator (that band in which signals are passed with half or more of the power possessed by signals at the exact resonant frequency) is expressible by the fraction Fo/Q, where Fo is the resonant frequency and Q is the circuit Q. Thus a 455 kHz resonant circuit with a Q of 100 will have a half-power bandwidth of 455/100 kHz, or 4.55 kHz. This relation is an approximation valid only for single-tuned circuits; more complex circuits are beyond this basic discussion.

The Q of Quartz Crystals. When we talk of the Q of conventional resonant circuits composed of coils and capacitors, a figure of 100 is usually taken as denoting very good performance and Q values above 300 are generally considered to be very rare.

The Q of a quartz crystal, however, is much higher. Values from 25,000 to 50,000 are not unheard of.

The extremely high Q makes the crystal a much m ore selective resonator than can be achieved with L-C circuitry. At 455 kHz, for example, the band-width will be between 10 and 20 Hertz (*cycles* per second) unless measures are taken to reduce Q. Even in practice (which almost never agrees with theory), 50-Hertz bandwidths are common with 455-kHz crystal filters.

So far as external circuitry is concerned, the crystal appears to be exactly the same as an L-C resonant circuit except for its phenomenal Q value. See Fig. 6.

At series resonance, the crystal has very low impedance. You may hear this effect referred to as a *zero* of the crystal. At parallel resonance, impedance is very high; this is sometimes called a *pole*. Fig. 7 shows a plot of *pole* and *zero* for a typical crystal. The special kind of crystal filter known as a half-lattice circuit matches the *pole* of one crystal against the *zero* of another, to produce a passband capable of splitting one sideband from a radio signal. Such filters are widely used in ham, commerical and, to a lesser extent, in CB transmitters.

When a crystal is used to control the frequency of a radio signal or provide a source of accurate timing signals, either the *pole* or the *zero* may be used. Circuits making use of the *pole* allow more simple adjustment of exact frequency, while those making use of the *zero* often feature parts economy. Later we'll examine several of each type.

From Rock to Finished Crystal. To perform its control functions properly, a quartz crystal requires extensive processing. The raw quartz crystal must be sliced into plates of proper dimension, then ground to the precise size required. Each plate must be as close to precisely parallel, and as perfectly flat, as possible. The electrodes must be in proper contact with the polished plate; in many modern units, the electrodes are actually plated directly to the crystal surface, usually with gold.

The crystal plate is known as a *blank* when it is sliced from the raw crystal. The blank is cut at a precise angle with respect to the optical and electrical axes of the raw crystal, as shown in Fig. 8. Each has its own characteristics for use in specific applications. Some, notably the *X*- and *Y* -cuts, are of only ;historic interest. The *Y*- cut, one of the first types used, had a bad habit of jumping in frequency at critical temperatures. The *X*- cut did not jump, but still varied widely in frequency as temperature changed.

Today's crystals most frequently use the *AT* cut for frequencies between 500 kHz and about 6 MHz, and the BT cut for between 6 and 12 MHz. Above 12 MHz, most crystals are specially processed *BT* or *AT* cuts used in *overtone* modes. These cuts are important to crystal makers and not relevant to our layman's theory.

The blanks are cut only to approximate size. The plates are then polished to final size in optical "lapping" machines which preserve parallelism between critical surfaces. During the final stages of polishing, crystals are frequently tested against standard frequency sources to determine exact frequency of operation.

If electrodes are to be plated onto the crystal surfaces, frequency cannot be set precisely by grinding since the electrodes themselves load the crystal slightly and cause a slight decrease in operating frequencies. These crystals are ground just a trifle above their intended frequencies, and the thickness of the electrodes is varied by varying plating time to achieve precision.

Accuracy. The precision which can be attained in production of quartz crystals is astounding. Accuracy of ±0.001 percent is routine, and 10-times-better accuracy is not difficult. In absolute figures, this means an error of one cycle per megahertz. In another frame of reference, a clock with the same accuracy would require more than 11 days to gain or lose a single second.

However, such accuracy can be achieved only when certain precautions are taken. For instance, the frequency of a crystal depends upon the circuit in which it is used as well as upon its manufacture. For an accuracy of ±0.005% or greater, the crystal must be ground for a single specific oscillator. If ±0.001% (or better) circuit accuracy is required, it must be tested in that circuit only. Thus, CB transmitters are on the narrow edge of being critical. This is why all operating manuals include a caution to use only crystals made specifically for that transmitter.

When one-part-per-million accuracy is required, not only must the crystal be ground for a single specific oscillator, but most often the oscillator circuit must then be adjusted for best operation with the crystal; this round-robin adjustment must be kept up until required accuracy is achieved. Even then, crystal aging may make readjustment necessary for the first 12 to 18 months.

Frequency Variation—Causes and Cures. Possible variations in frequency stem from three major causes, while cures depend entirely upon the application.

The most obvious cause of frequency variation is temperature. Like anything else, the crystal will change in size when heated and the frequency is determined by size. Certain cuts show less change with temperature than do others, but all have at least some change.

For most noncommercial applications, the heat-resistant cuts do well enough. For stringent broadcast station and critical time-signal requirements, the crystal may be enclosed in a small thermostatically-regulated oven. This assures that the steady temperature will cure one cause of frequency change.

The second well-known cause for variation of frequency is external capacitance. Some capacitance is always present because the crystal electrodes form the plates of a capacitor where the crystal itself is the dielectric. Most crystals intended for amateur use are designed to accommodate an external capacitance of 32 picoFarads (pF), so if external capacitance is greater than this, the marked frequency may not be correct. Crystals for commercial applications are ground to capacitance specifications for the specific equipment in which they are to be used. CB crystals also are ground for specific equipment, although many transceivers employ the 32 pF standard.

Trim a Frequency. When utmost precision is required, a small variable capacitor may be connected in parallel with the crystal and adjusted to change frequency slightly. The greater the capacitance, the lower the frequency. Changes of up to 10 kHz may be accomplished by this means, although oscillation may cease when excessive changes are attempted.

Like temperature-caused variations, frequency variations due to capacitance may be useful in special cases. Hams operating in the VHF regions obtain frequency modulation by varying load capacitance applied to the crystal in their transmitters.

The third cause for variation of frequency is a change in operating conditions in the associated circuit. This cause is more important with vacuum-tube circuits than with semiconductor equipment. As a rule, operating

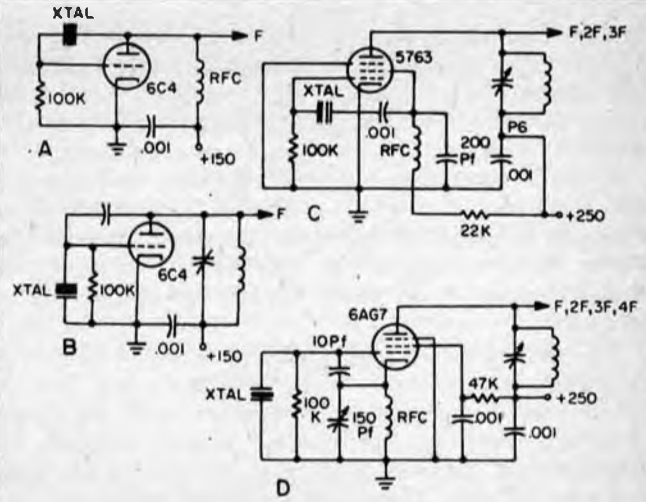
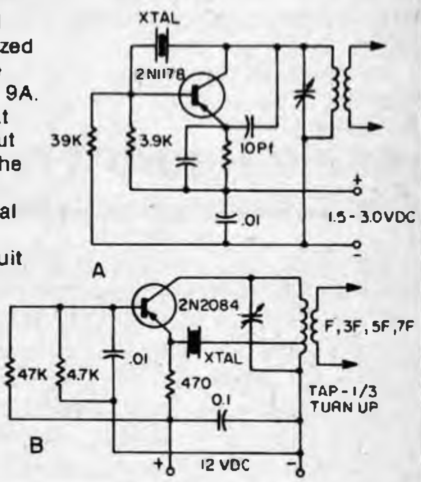


Fig. 9. The simplest crystal oscillator circuit (A) has no tuned circuits. To change frequency it is only necessary to change crystals, although a small variable capacitor across the crystal will cause some small frequency change. Miller oscillator (B) is nearly as simple as Pierce type shown in (A). Tuned circuit can pick out fundamental frequency or harmonics. Pierce electron-coupled oscillator (C) derives its feedback from the screen sircuit, eliminating need for a buffer amplifier in most cases. Colpitts oscillator (D) gets its feedback from cathode circuit. Variable capacitor in the grid-cathode circuit can trim frequency.

Fig. 10. Fundamental frequency transistorized oscillator (A) is quite similar to that in Fig. 9A. One difference is that tuned circuit in output replaces RFC unit. The overtone (harmoinic) circuit (B) uses crystal for odd harmonic feedback. Either circuit can be used for fundamental frequency operation—just tune.



voltages for any vacuum-tube oscillator providing critical signals should be regulated to prevent change.

Again, this cause can be used to provide FM by deliberately varying voltages.

Crystal Aging. A final cause of frequency variation, small enough to be negligible in all except the most hyper-sensitive applications, is crystal aging. When a crystal is first processed, microscopic bits of debris remain embedded in its structure. These bits are displaced during the first 12 months or so of use, but during that time the crystal frequency changes by a few parts per million. Extreme accuracy applications must take this change into account. For most uses, though, it may be ignored.

Using Quartz Crystals. After all the discussion of crystal theory, it's time to examine some typical circuits. While dozens of special crystal circuits have been developed for special applications, a sampling will suffice for discussion. Fig. 9 shows four typical vacuum tube crystal oscillator circuits.

The simplest of these is the Pierce circuit. Fig. 9A. While at first glance this circuit appears to employ the crystal's *zero* to feed back energy from plate to grid, the *pole* is actually used through a mathematically-complex analysis. This circuit has one unique advantage: it contains no tuned elements and, therefore, can be used at any frequency for which a crystal is available. This makes it an excellent low-cost test signal source. The major disadvantage is that excessive current may be driven through the crystal if DC plate voltage rises above 90 or so.

The Miller oscillator (Fig. 9B) is almost as simple to construct and operate as is the Pierce and has an additional advantage of operation with overtone crystals. This is the circuit recommended by *International Crystal Mfg. Co.* for use with their overtone crystals. The capacitor shown between plate and grid is usually composed of grid-plate capacitance alone. The *pole* is used here also, energy feeds back through the grid-plate capacitance, and the *pole* selects only the parallel-resonant frequency (shorting the rest to ground).

ECO. The electron-coupled Pierce oscillator (Fig. 9C) is similar to the basic Pierce. The tuned circuit in the plate offers the possibility of emphasizing a harmonic—an RF choke may be used if freedom from tuning is desired and fundamental-frequency operation will suffice.

GPO. One of the most popular oscillators of all time is the Colpitts Crystal oscillator of Fig. 9D, sometimes known as the *grid-plate* oscillator. The feedback arrangement here consists of the two capacitors in the grid circuit; feedback is adjusted by means of the 150 pF variable capacitor (the greater the capacitance, the less the feedback) until reliable oscillation is obtained. Like the other three oscillators, this circuit employs the crystal *pole* frequency.

Since all four of these oscillator circuits utilize the *pole* for frequency control, exact frequency adjustment capability may be obtained by connecting a 3-30 pF trimmer capacitor in parallel with the crystal.

Crystal oscillators may, of course, be built with transistors, too. Two typical circuits are shown in Fig. 10. Feedback mechanisms differ somewhat because of the basic differences between tubes and transistors. In general, transistorized oscillators are more stable.

As a Clock. To use a crystal as the timing element of a clock, an oscillator identical to those shown in Figs. 9 and 10 is the starting point. Crystal frequency is chosen at a low, easily-checked value such as 100 kHz. This frequency is then divided and redivided by synchronized multivibrators to produce one-cycle-per-second pulses. These may then be counted by computer counting circuits. IUC chips that divide by 10, or other programmed values, do a far better job serving as a clock because of the overall accuracy and low cost of parts.

Inside Radio Transmitters

One if by land, and two if by sea..." says the famous poem by Longfellow commemorating the midnight ride of Paul Revere in April of 1775. Revere's fellow patriot, who hung the two (if by sea) lanterns in the steeple of the Old North Church of Boston 200 years ago, was engaged in *communicating by modulation*, just as surely as today's CBer who presses the PTT switch on his microphone. For *modulation* simply means *variation*, or *change*—and it's modulation, whether you're changing the number of lanterns hanging in a church steeple, or using electronic circuitry to change the radio wave emitted by an antenna in accordance with your voice.

All communication is by modulation. For centuries, the American Indians sent messages by "modulating" a smoke stream with a wet blanket, and primitive tribes have long communicated by modulating the beat of their jungle drums. Later, semaphore flags were sued to send messages by modulating their position. Even these words you are reading can be considered modulation of the surface of a piece of paper with spots of ink.

But almost all of today's long-distance instantaneous communication is carried out by modulating radio waves. In fact, this means of communicating is now so commonplace that even the Man in the Street unknowingly refers to modulation when he speaks of "AM" and "FM". These familiar abbreviations stand for *Amplitude Modulation* and *Frequency Modulation*, respectively, and refer to the two common methods of changing a radio wave to make it broadcast words or music from one place to another.

Introducing the Carrier. A radio wave broadcast from the antenna of a transmitter is, in the absence modulation by speech or music, an unchanging, constant sine wave, as shown in Fig. 1. It is as constant and as unchanging as the steeple of the Old North Church, and conveys no more information than a steeple. It simply gives you something to monitor for thepossible later appearance of a signal.

Just as the steeple was a support or carrier on which to hang the information-giving lanterns, so the radio wave becomes the carrier upon which the speech or music is "hung". In fact, the unmodulated wave is usually referred to as the *carrier*.

The carrier may be shown as a simple sine wave, as in Fig. 1.

The height, or amplitude of the wave, indicates the strength of the signal, while the time it takes the wave to complete a certain number of cycles determines the spot on the radio dial where the signal will be received. For example, as shown in Fig. 2a if it takes only a millionth of a second for the carrier to complete seven cycles, then it will complete 7,000,000 cycles in one second, and the signal will appear on a receiver's dial at the 7,000,000-cycle-per-second (7-megahertz) point, which is on the edge of the 40-meter ham band. Such a carrier has a *frequency* of seven MHz.

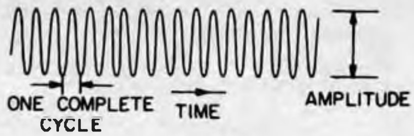


Fig. 1—Unmodulated carrier (RF) wave.

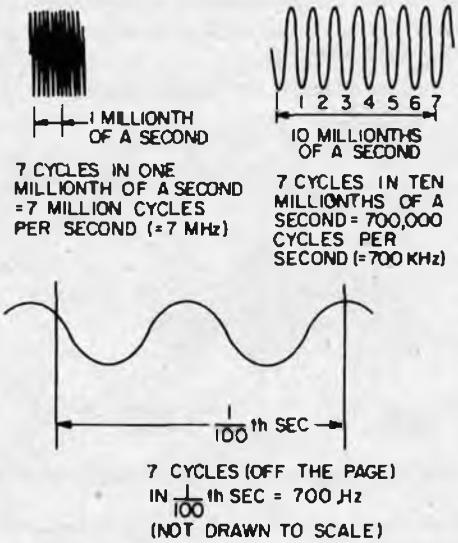


Fig. 2—Three different frequency waves. 700 Hz (bottom) is partial, not drawn to the same scale as the two higher frequencies.

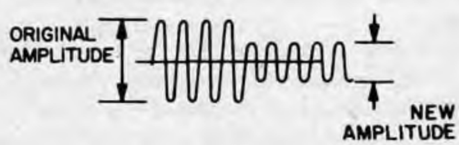


Fig. 3—Simplified amplitude modulation of the RF carrier wave.

On the other hand, a carrier taking longer to complete the same number of cycles—say, seven cycles in 10 millionths of a second (Fib. 2b)—it would complete only 700,000 cycles in one second, and would be found on the dial at 700 kHz (700 thousand Hertz), which is in the standard broadcast band.

As can be seen from the above numbers, carrier frequencies are normally very high—much higher than the speech or music (audio) requencies which we will cause the carrier to carry. For example, when a flutist plays the note F above middle C, he produces vibrations in the air which can be visualized as in Fig. 2c. Here, the time for 7 vibrations is only *one fiftieth* of a second, which is a frequency of only 350 cycles per second (350 Hertz).

But a constant (*unchanging*) carrier wave conveys no information. Something about the wave must be *varied* (modulated) to convey information to the listener. What can be changed, so that the listener can recognize that a signal has been sent to him?

AM and FM. Looking again at Fig. 1 you can see that a carrier has two obvious characteristics—its height, or *amplitude*, and its *frequency*. Changing either of these can cause a receiver to recognize that a message has been sent.

If the amplitude is changed, we call it *amplitude modulation*, or AM. If the frequency is changed, we call it *frequency modulation*, or FM.

A very simple type of AM is shown in Fig. 3. Here the amplitude of the carrier wave has been changed suddenly to half its former value.

This change in amplitude is a simple form of AM, and can convey simple messages. If Paul Revere had been a CBer, he could just as easily have prearranged a code signal which said "...one drop in carrier amplitude if by land; two drops in carrier amplitude if sea...," and served the American cause just as well (though Longfellow's poetry might have suffered).

The other obvious characteristic of the carrier wave of Fig. 1 is its *frequency*. We can also modulate this characteristic, as shown in Fig. 4. Here, instead of a sudden change in amplitude, there is a sudden change in frequency, from 7 HMz to 3.5 MHz. This is a very simple form of FM, and can also be used to convey simple messages. Since the drop in frequency represents a shift in the carrier's location on the dial, as shown in Fig. 5, two receivers, one tuned to 7 MHz and the other to 3.5 MHz, could detect this shift in frequency, and the listener could interpret it as a singla, according to a pre-arranged code.

What's PM? While the Man in the Street has made AM and FM household phrases, these modulation methods are only two of the three ways a radio frequency carrier wave may be modulated. The third method, *Phase Modulation*, or *PM*, although virtually unknown to most people, is nonetheless extremely important in such fields as data transmission and color television.

Phase modulation can be visualized as in Fig. 6. Here, neither the amplitude nor the frequency is varied, but the carrier is made to pause for a moment, and then to continue as a sine wave slightly delayed from the original. This delay is called a *phase shift*. Phase shift is usually measured in *degrees*. A phase shift equal to the time needed for an entire cycle is 360°. In the sketch, a sudden phase shift of about 70° (less than a quarter cycle) is indicated. By suitable receiver circuitry (found in every color TV receiver), this sudden change in phase can be interpreted as a signal. In color TV, it might represent a shift in hue from green to yellow.

Amplitude Modulation—A Closer Look. The sudden drop in amplitude shown in Fig. 3 is a good way to show the general scheme of AM, but it fails to tell us very much about how AM is used, every day, in our AM receivers and CB rigs. Here there are (hopefully) no sudden shifts in carrier amplitude but instead, there is a remarkable recreation of speech and musci from a distant transmitter. How is this done?

To explain, let us assume that our flutist stands before a microphone in a broadcasting studio, ready to play his 350-Hz F-above-middle-C. Let's also assume that the broadcasting station is assigned a carrier frequency of 700 kHz. Figure 7 shows how the carrier wave will appear just before the flutist plays, and just after he begins.

As you can see, the 350-Hz audio tone from the flute causes the amplitude of the carrier to rise or fall in

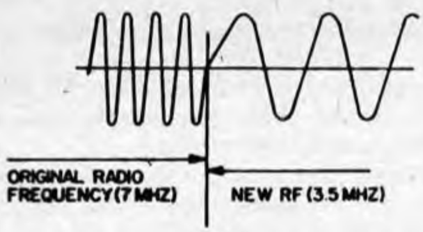


Fig. 4—Simplified frequency modulation of the RF carrier wave.

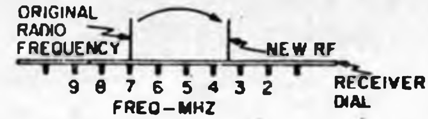


Fig. 5—If the frequency modulation were a simple change from one RF carrier frequency to another, an AM receiver could receive it if returned.

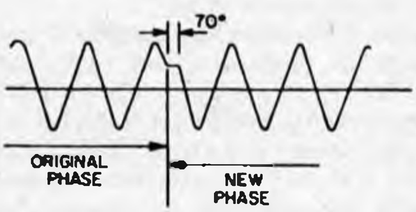


Fig. 6—Simplified phase modulation of the RF carrier wave.

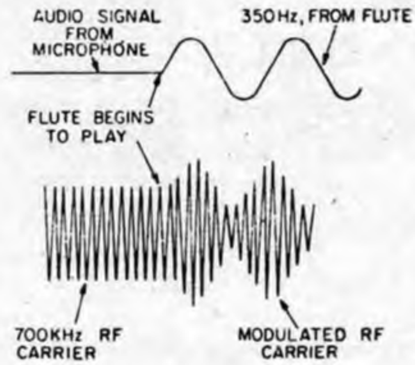


Fig. 7—The audio signal (sound) modulates the RF carrier wave.

accordace with the rise or fall of the flute wave. Note that both the top and bottom of the carrier wave are affected by the flute waveform. It is as though the carrier had been squeezed into a snug-fitting envelope, forcing it to conform to the waveform of the flute sound. The shape thus formed by the tips of the modulated carrier wave is often called the *envelope*.

The envelope of an amplitude-modulated carrier is therefore a good replica of the audio waveform coming from the studio microphone. Every shading, every change, in the sound striking the microphone will be faithfully traced out by the tips of the carrier wave. For example, if the flutist were to play more softly, the result will be as in Figure 8. If he plays more loudly, Figure 9 is the result. You will note that in Fig. 9 the amplitude modulation is so intense that, at one point, the carrier's amplitude goes to zero for an instant. This is called 100% modulation, and represents the loudest sound AM can handle. If the flutist plays even *more* loudly, the result is as shown in Fig. 10. As the figure shows, the envelope is no longer a faithful replica of the original audio waveform, so the listener will receive a distorted sound. This condition is called *overmodulation*, and is undesirable.

Hardware for AM Systems. One of the most-straightforward methods of producing AM is the method invented by the Canadian Reginald Fessenden, in 1905. In his system, a radio frequency generator produced the carrier wave, which was fed to the antenna through a carbon (variable-resistance) microphone. See Fig. 11.

Since a carbon microphone varies its resistance in accordance with the speech or music, the microphone in this primitive system acted as a valve to allow more or less of the RF carrier wave to pass to the antenna. In this way the carrier wave broadcast by the antenna was amplitude modulated by the sound waves striking the

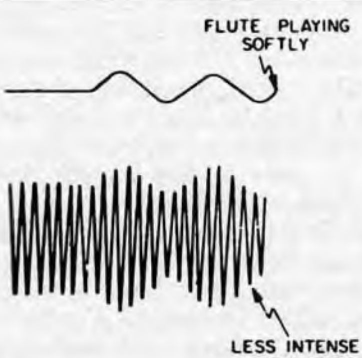


Fig. 8—Smaller audio signal modulates the RF carrier wave less.

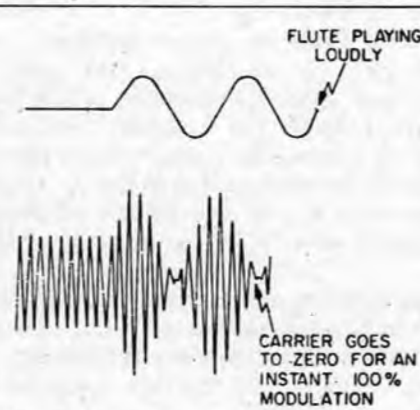


Fig. 9—Louder audio signal (sound) modulates the RF carrier wave to maximum.

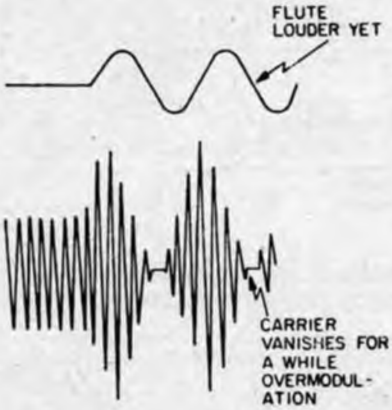


Fig. 10—Further increase in amplitude of audio causes overmodulation.

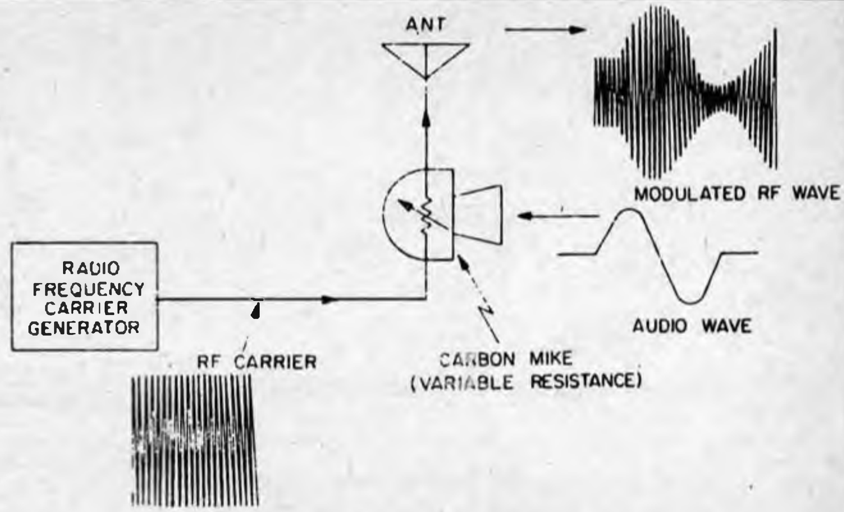


Fig. 11—The first radiotelephone braodcast transmitter, invented by Canadian Reginald Fessenden, who broadcast voice and music to a few ships at sea on Christmas Eve, 1905. It worked, but was very noisy.

microphone.

More Up-to-date Modulation Methods. Although nobody puts carbon microphones in series with antennas any more, the more modern modulation methods, such as those found in AM broadcast transmitters or CB transmitters are still rather similar to the primitive carbon-mike method. The typical modern AM system (Fig. 12), still employs an oscillator (which is crystal-controlled, to ensure that the carrier frequency is constant, and thus is found at a known spot on the dial), and a power amplifier to strengthen the carrier before feeding it to the antenna. The amount of "strengthening" is controlled by the audio signal, and the music to a distant point.

Receiving the Signal. At that distant point, we all know that the carrier power delivered to the antenna is in this way amplitude-modulated and the radiated carrier will convey the speech or wave may be intercepted by a suitable antenna and applied to a receiver. A simple receiver is shown in block diagram form in Fig. 13. Here, the weak signal from the antenna is first amplified in a carrier-wave amplifier, an RF (radio frequency) amplifier, and applied to a *detector* which can extract the original audio signal from the amplitude-modulated RF carrier. This audio wave is further amplified and then fed to a loudspeaker, which re-creates the original speech or music.

The Detector—Heart of the Receiver. By looking at Fig. 13, you can see the detector is the heart of the AM receiver which extracts the original signal from the amplitude-modulated carrier. The circuit that does this job is suprisingly simple. It resembles very closely the circuit of a typical power supply. In an AC power supply, when 117-volts AC is applied at the input plug, a DC voltage appears at the output. If, because of a brown-out or for some other reason, the 117 volts of the input drops to, say, 105 volts, the DC output will deop correspondingly. In a power supply, this output drop is undesirable—we want the DC output to remain constant even though the input AC varies. Notice that the drop in voltage of the 60-Hz power line input from 117 volts to 105 volts is *actually amplitude modulation*, of the 60-Hz input sine wave, and *the drop in the output DC has "detected" the AM that occurred at the input!* This is shown in Fig. 16.

So a simple AM detector may be thought of as a power

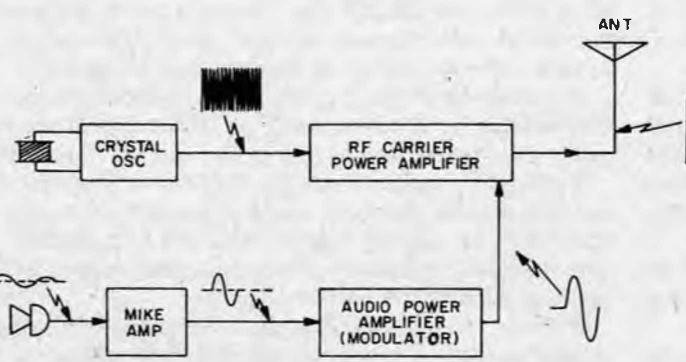


Fig. 12—Typical amplitude-modulated RF (radio frequency) transmitter.

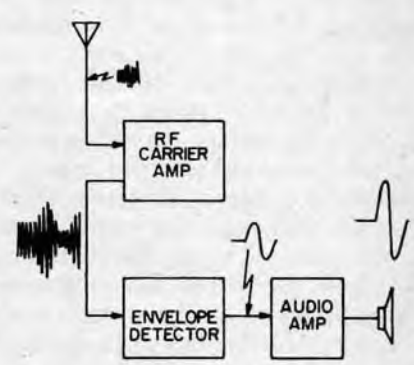


Fig. 13—Block diagram of receiver for amplitude-modulated radio frequency (RF) signals.

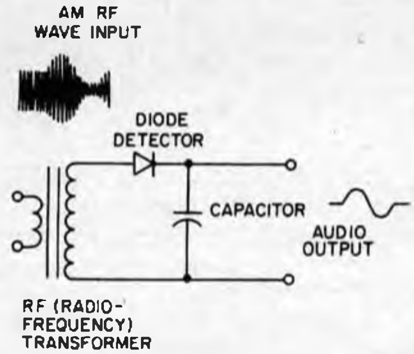


Fig. 14—Simple receiver of amplitude-modulated (AM) RF signals.

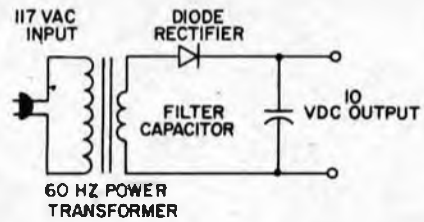


Fig. 15—DC power supply from AC house current has circuit similar to that of the simple AM receiver.

supply arranged to change its output very quickly in accordance with the amplitude changes in the AC (carrier) at its input. And, since these fluctuations in amplitude of carrier represent the original audio signals, the "power supply" (detector) output will be the same as the original audio. It's interesting to note the power supply's desirable trait, unsteadiness inoutput when the input is unsteady, is the useful operating characteristic of the same circuit when used as a detector!

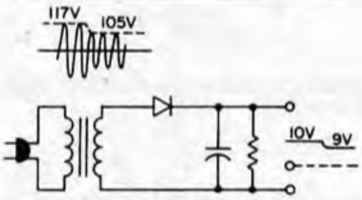


Fig. 16—Changing amplitude of voltage input to power supply provides changed (modulated) output.

Single Sideband — How It Works

All communication is by means of *change*, or modulation of some kind, whether it's electronic communication, smoke signals, or the lanterns in the Old North Church in Boston which sent Paul Revere on his famous midnight ride over 200 years ago. There are three basic ways a carrier can be changed, or *modulated*, to convey information from a radio transmitter to a radio receiver. There three ways are AM, FM, and PM (Amplitude Modulation, Frequency Modulation, and Phase Modulation). To get the most out of this theory article, you should have a good understanding of AM waveforms and what they mean.

But, examining waveforms is not the only way to gain an understanding of AM. If we look, instead, at what happens at the dial of the receiver as the carrier is modulated, we get another view of AM—one that will help us to a better grasp of the inner workings of not only AM, but FM and PM as well.

A Dial's-Eye view of AM. Imagine a radio receiver with a dial which lights up at any point where a signal appears. By moving the tuning indicator to one of the lit points, the operator could hear in his loudspeaker the speech or music being transmitted at that frequency. Such an imaginary dial would appear as in Fig. 1.

The lights indicating the presence of a signal appear in the figure as vertical lines, with strong signals appearing as tall lines and weaker signals as short lines.

Imagine further that we can use a magnifying glass to closely imspect the area around one of the bright lines, and at the same time can set the tuning indicator to that bright line, so we can hear in the loudspeaker any signal being sent (Fig. 2).

The magnifying glass shows us a single bright line indicating a carrier at 840 KHz, while the loudspeaker gives us only a low-level hiss, indicating that an unmodulated carrier is present, but no information is being sent. (To continue our Paul Revere analogy, we could say that the steeple is present, but no lanterns have been hung.)

Now, let us ask a flutist to step before the microphone and play a 700-Hz tone. Immediately, of course, the flute's tone is heard in the receiver's loudspeaker. But the most surprising thing happens at the receiver dial: Here, (see Fig. 3), two *new* frequencies appear—small bright lines on either side of the carrier. Looking closely, we see that each of the new frequencies is spaced exactly 700 Hz from the carrier—one above, at 840,700 Hz, and the other below, at 839,300 Hz. These two new frequencies are called *side frequencies*, and are the lanterns in the steeple—the indication of information being sent.

If we ask the flutist to play louder, we find that the side frequencies frow taller, until, at 100% modulation each one is exactly half the height of the carrier. (See Fig. 4).

Now, let us suppose that another flutist joins the first, and this second musician plays a tone at 1100 Hz (in this case blowing slightly louder than the first flutist). We now find a second group of side frequencies spaced 1100 Hz each side of the carrier (Fig. 5).

With the two flutists playing, we have the beginning of a *band* of side frequencies on each side of the carrier. The higher frequency is the *upper side band*, and the lower frequency is the *lower side band*.

If the two flutists are now joined in the studio by a rock group (or a symphony orchestra), the many instruments,

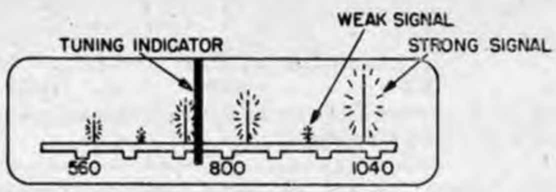


Fig. 1—Signal strenth is represented by height of vertical lines.

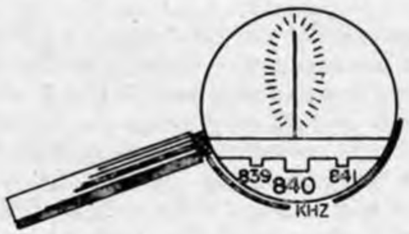


Fig. 2—Non-modulation of signal concentrates signal strength at the carrier frequency.

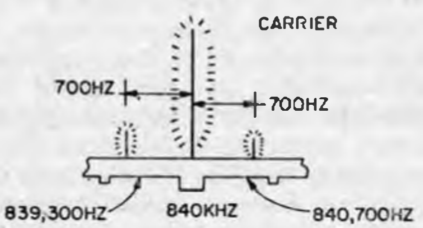


Fig. 3—Modulating the carrier with a 700-Hz audio tone generates weaker signals 700 Hz above and below carrier.

playing many different frequencies simultaneously, will cause many side frequencies to appear, and the sidebands will be composed of a very complicated group of side frequencies, constantly changing as the musicians play. Figure 6 can be thought of as a snapshot of the sidebands for a very brief period in time. This is called a *spectrum display*, and the group of frequencies shown is called the spectrum of the signal.

In this spectrum display, we observe that the upper and lower sidebands extend equally far from the carrier—about 5,000 Hz in the figure—and that they are always exact mirror images of each other. As the sound striking the microphone changes, the sidebands change in a mirror-image manner. A snapshot (spectrum display) taken a few seconds later might look like Fig. 7.

Is This Sideband Necessary? The fact that the two sidebands are mirror images of each other makes us wonder if they are both necessary. It's just as though Paul Revere's friend had hung a complicated set of lanterns on the *right* side of the steeple, and then had hung an identical set on the *left*. The second set would have been unnecessary because the first set would have given him the message, and Paul would have been on horseback and gone long before the second set could be lit!

In the same way, we can get the message through by using only a single sideband, instead of the mirror-image pair. Such a system is called a *single-sideband system*, and is abbreviated SSB. SSB offers many advantages over conventional AM systems, especially for voice communication.

Eliminating the Carrier. If only one sideband is needed to get the message through, we can also ask whether the carrier itself is absolutely necessary. Can we get the message through without the carrier?

The answer is, definitely yes! With proper modulation methods, we can eliminate the carrier completely, and still send the message. To continue our Paul Revere analogy, we are saying that it is only the lanterns that are essential to the message; the steeple could be eliminated if we could find a way to hold the lanterns in place.

Figure 8 shows the result of eliminating the carrier, but keeping *both* sidebands.

This is called a *suppressed carrier system.* The most striking result of suppressing the carrier is seen in the waveform displays of Fig. 8. First we see that, in the absence of an audio signal, there is no carrier (Fig. 8). Second, we note that the envelope representing the original audio signal is somehow tangled up inside the waveform, in such a way that it can't be extracted by a simple "power-supply"—kind of detector (envelope detector).

Sending Only the Necessities. If each sideband of a 100-watt transmitter provides only 25 watts, then eliminating the carrier cuts the total power from 100 + 25 + 25 = 150 watts, down to 25 + 25 = 50 watts. But, as we pointed out earlier, each of these two sidebands is a mirror image of the other; one of them can be eliminated withut hurting the message a bit! Fig. 8 shows the spectrum display of an SSB transmitter modulated by a 700-Hz tone. The dotted lines show where the suppressed carrier and the eliminated lower side band would have been.

Dropping one of the unnecessary sidebands reduces the required power down to 25 watts, with only a small decrease in the ability of the signal to get through. Or, if we wish, we can design the transmitter to pour all its available power into that one sideband, with a

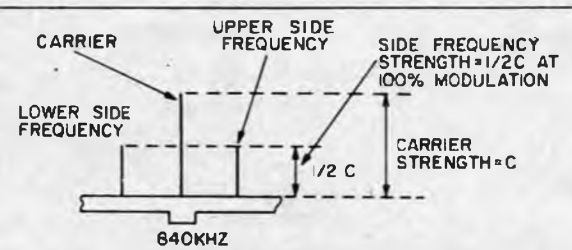


Fig. 4—Strength of upper and lower sideband frequencies is proportional to modulation.

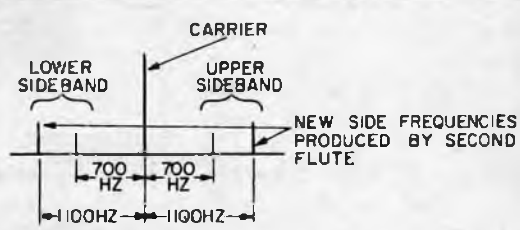


Fig. 5—Second audio modulating signal generates second set of sidebands.

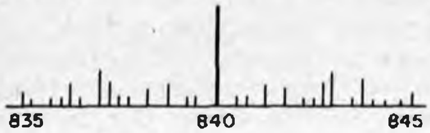


Fig. 6—Complex audio modulating signals generate complex sideband frequencies.

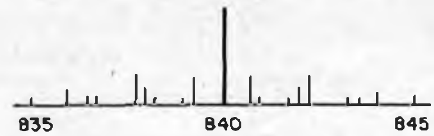


Fig. 7—A different set of audio signals generates a different set of sideband frequencies.

consequent increase in our ability to get the message through—and without changing the transmitter's size, weight, or power! The only price we pay is the requirement that we insert an artificial carrier at the receiver, to allow us to detect the information in the sideband sent.

SSB Waveforms. Now, with SSB concepts firmly in hand, if we look back to the world of waveforms we shall find a most surprising and interesting story.

Figure 9 tells this story. In Fig. 10a, we see our familiar 700-Hz flute tone, and on the right of Fig. 10b we see the SSB spectrum resulting from modulating an SSB transmitter with this tone. But on the left of Fib 10b we get our first surprise. True, we would expect no carrier, before the tone begins—we've seen this before, in Fig. 8 (the suppressed carrier case). But when our flutist begins to play, we get a single, steady sine wave from the transmitter—with no evidence of modulation of any kind! Can this be a signal? It looks just like a carrier!

The result of eliminating the carrier looks so unusual that we wish we could put it back. So that's exactly what we do. At the receiver, we insert an artificial carrier, carefully adjusted to the right frequency and strength, which restores the waveform to its more familiar appearance, (Fig. 8), and allows us to detect it with a conventional detector.

But, you might ask, why go through all the motions of first removing the carrier, and then putting it back? What advantage does this gain for us?

The most compelling reason for carrier suppression is the *saving* in transmitter power. A transmitter could pour 10,000 watts into its carrier, and only a few millionths of a watt of this would arrive at the receiver. It makes much more sense to eliminate the huge transmitter wattage, and instead build a small oscillator at the receiver to provide the few millionths of a watt needed to imitate the carrier.

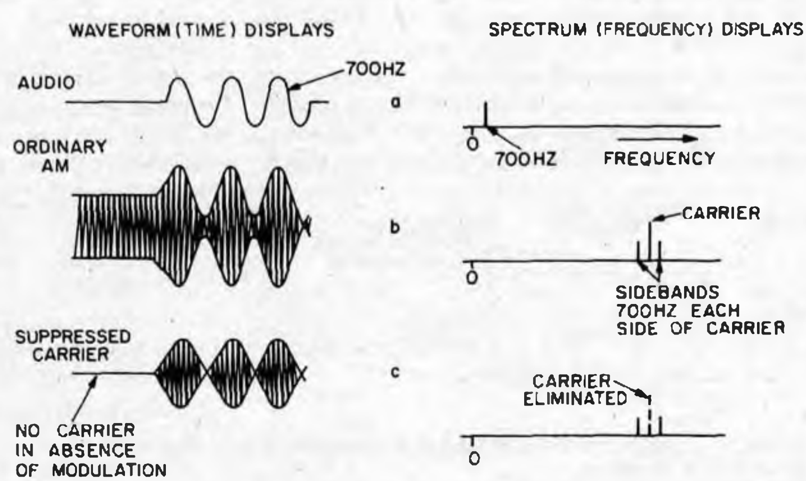
The power saving becomes even more interesting when we observe how little of the total transmitter power goes into the sidebands. At 100% modulation, a transmitter with 100 watts in the carrier will have only 25 watts in each sideband for a total of 50 watts in the pair. So, only one-third of the transmitter's 150 watts of total power is in the information-bearing sidebands.

In case you're wondering how, in Fig. 8, we could say the sidebands at 100% modulation are *half* the carrier, while here we're saying they're *one-quarter*, remember that Fig. 8 shows relative *voltages*. Since power is proportional to the *square* of the voltage, and since *one-half* squared is *one-quarter*, then each sideband's *power* at 100% modulation is *one-quarter* of the carrier power. It's the old story of apples and oranges.

Can't Tell the Lanterns from the Steeple without a Program! The steady sine wave of Fig. 10b is indeed a signal; not a carrier. In the frist place, it will get bigger (more amplitude) if the flutist plays louder, and will get smaller if he plays more softly. It is, therefore, amplitude modulated—for SSB is just a form of AM. But the *pitch* of the flute—the 700 Hz is *not* packaged neatly in an envelope, as it was in Fig. 8b (standard AM) and in Fig. 8c (suppressed carrier). The only clue to its 700-Hz pitch is the fact that this single carrier-like frequency is *700 Hz from where the carrier would have been* if we had sent it!

You can see, then, that re-inserting the carrier at precisely the right frequency is essential to recovering the proper pitch. If the carrier was supposed to be 840,000 Hz, and it is instead reinserted at 840,000 Hz—only 30 Hz high—then the flute tone will come out of the speaker 30 Hz low, at 670 Hz instead of 700 Hz. Conversely, a reinserted carrier at 839,970 Hz—30 Hz low—will obtain for us a higher pitch:—730 Hz. Therefore, we must know precisely where the carrier should have been if we are to re-create the exact sound which struck the microphone. This is the additional

Fig. 8—Elimination of the carrier leaves the sidebands, which include the audio information. Carrier is reinserted at the receiver.



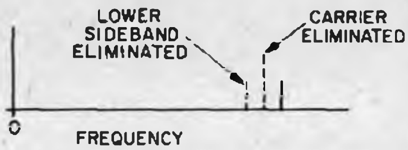


Fig. 9—Spectrum display of a transmitted signal modulated by a 700-Hz tone—lower sideband suppressed.

complexity that we endure to obtain the advantages of SSB.

Tuning by Ear. In a practical situation, the speech and music errors introduced by even a slight carrier error are so obvious that it's easy for a listener to simply readjust the inserted carrier frequency until the reproduction "sounds right." This is possible because, as soon as more than a single, simple tone is transmitted, our ears quickly detect the errors in the complex waveforms, such as speech and music, which we have encountered all our lives, and a simple twist of the knob quickly clarifies the reception.

Two-tone Waveforms for SSB. As an illustration of what can happen if more than one tone is sent via SSB, look at Fig. 10c. Here, our flute-playing pair is again intoning their 700-Hz and 1100-Hz pitches. On the right of Fig. 10d, we see the *two* resulting sidebands—one, 700 Hz from where the carrier would have been, and the other, 1100 Hz from the same location. The carrier and lower sideband are shown dotted, indicating that they have been removed. If we compare this SSB spectrum with Fig. 8c, (the suppressed-carrier, *double*-sideband case), we observe that, in *both cases, the receiver is simply looking at two frequencies.*

Therefore, the waveforms should be very much the same in the two cases, even though one arises from a single flute tone, and the other, from *two* flute's tones. In the earlier figure, (suppressed carrier), the two frequencies were 2 x 700 = 1400 Hz apart, and the resulting waveform had peaks at a 1400 Hz rate. In Fig. 10d, the two frequencies are 1100—700 = 400 Hz apart, so the peaks occur at a 400 Hz rate. Note that this frequency—400 Hz—never occurred at the microphone; it is the *difference* between the two tones generated by the flutists. But the most important fact is that the waveform of a *single* flute modulating a *double-*

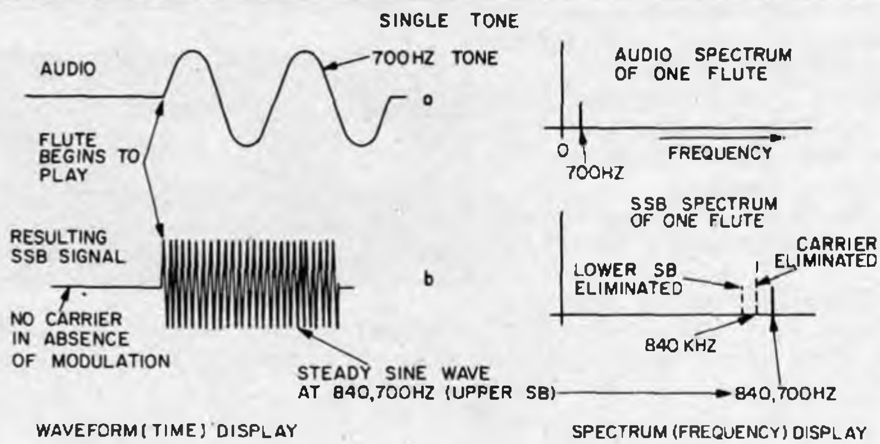


Fig. 10—At onset of 700 Hz audio modulation the carrier sideband appears. With no modulation, no sidebands are present.

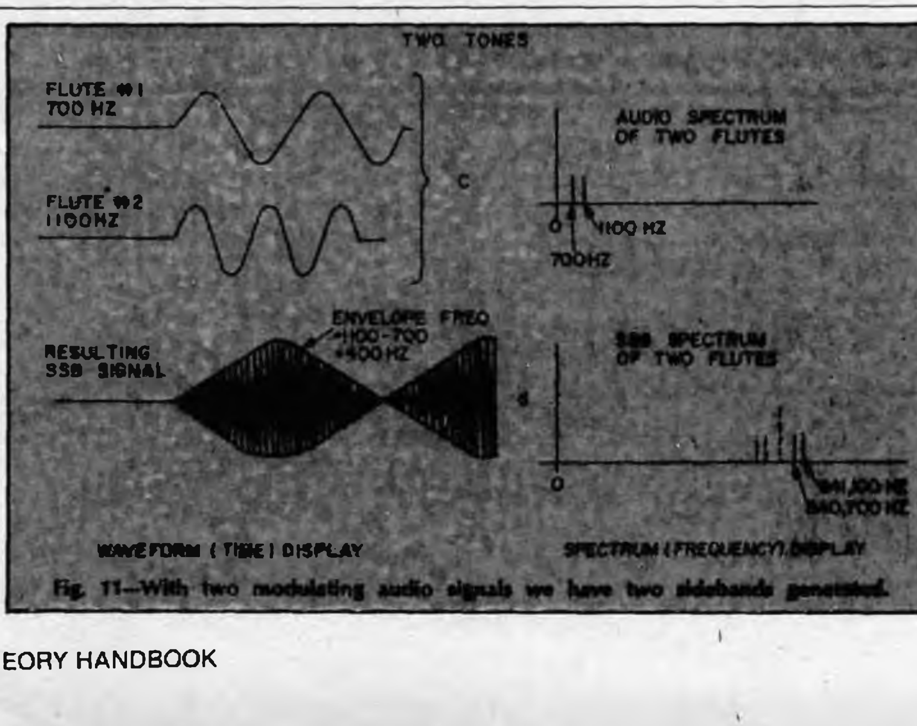


Fig. 11—With two modulating audio signals we have two sidebands generated.

sideband transmitter is the same as *two* flutes modulating a *single*-sideband transmitter.

Hardware for SSB. Since SSB is merely a very sophisticated version of ordinary AM, you might expect that you could generate a SSB signal by some kind of change to the basic AM system. In fact, a three-step change to that figure will give an understanding of how SSB is generated. In the first step, (Fig. 11), we have added another carrier power amplifier, an inverter, and an adder (the new parts are shown in bold lines).

The new amplifier provides an "upside-down" carrier which, when added to the original carrier, cancels it—provided the gains of the carrier power amplifiers are the same. The result is no carrier output at all—the carrier has been suppressed. In Fig. 12 we show the second step of the change.

Here, the *audio* signal is inverted and passed through another audio power amplifier. This inverted audio signal is used to control the output of the new inverted-carier amplifier. Since the audio supplied to the inverted-carrier amplifier is "upside-down," it will *weaken* the output of the amplifier at the same instant that the "right-side-up" audio is *strengthening* the output of the "right-side-up" carrier. The two carriers are therefore unequal, so they can no longer cancel each other; hence, an output will appear at the antenna. This output is the *double*-sideband, suppressed-carrier signal of Fig. 10c.

SSB by Sideband Filtering. The third step of our three-step change consists of merely adding a filter to eliminate one of the sidebands—either one. In Fig. 13, we have chosen to eliminate the *lower* sideband.

The signal from the antenna is now a single-sideband signal, and will have the waveforms and display of Fig. 10.

Other Roads to SSB. This SSB system was chosen as an easy-to-understand system which can be developed directly from a standard AM transmitter. It is not the most practical configuration, but, understanding it, you will easily understand the more practical systems.

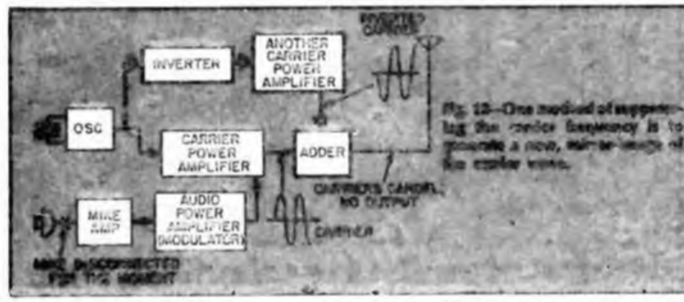


Fig. 12—One method of suppressing the carrier frequency is to generate a new, mirror-image of the carrier wave.

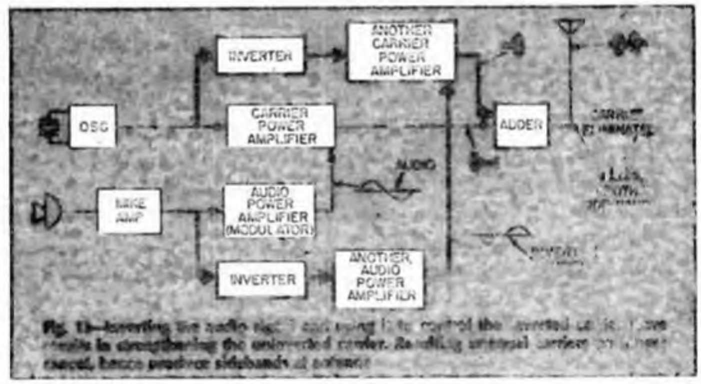


Fig. 13—Inverting the audio signal and using it to control the inverted carrier results in strengthening the uninverted carrier. Resulting unequal carriers no longer cancel, hence produce sidebands at antenna.

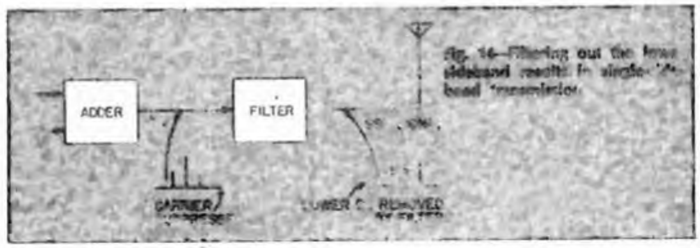


Fig. 14—Filtering out the lower sideband results in single-sideband transmission.

AM/FM Receiver Alignment

The words "receiver alignment" often conjure up a mysterious and complicated procedure which can be performed only by an expert. While it is true that receiver alignment should not be attempted by anyone who does not have the necessary skills and equipment, it is not a very difficult procedure if you have some basic guidelines. This article will explain the various procedures for aligning AM, FM, and AM/FM receivers using a minimum of equipment.

Before getting into the mechanics of receiver alignment, it is important that the service technicians understand the basic operation of the receiver, and why proper alignment is necessary. To this end, a discussion of the modern superheterodyne receiver will follow.

Receiver Fundamentals. Virtually every AM and FM receiver manufactured today is a superheterodyne receiver. They utilize a built-in RF oscillator and mixer circuit to convert the received signal to a lower frequency called an intermediate frequency (IF). In a given AM or FM receiver, the IF remains the same frequency regardless of which radio station is being received.

The basic advantage of this type of circuit is that the greatest part of the RF gain, and bandpass characteristic of the receiver, is provided by the IF stages, and is a constant for all received frequencies. Thus, the manufacturer of the receiver can precisely determine the sensitivity and selectivity of the receiver. Proper alignment ensures that the receiver performs exactly the way the manufacturer intended. Refer to Fig. 1, which is a simplified schematic diagram of the RF and IF sections of a typical modern day AM receiver.

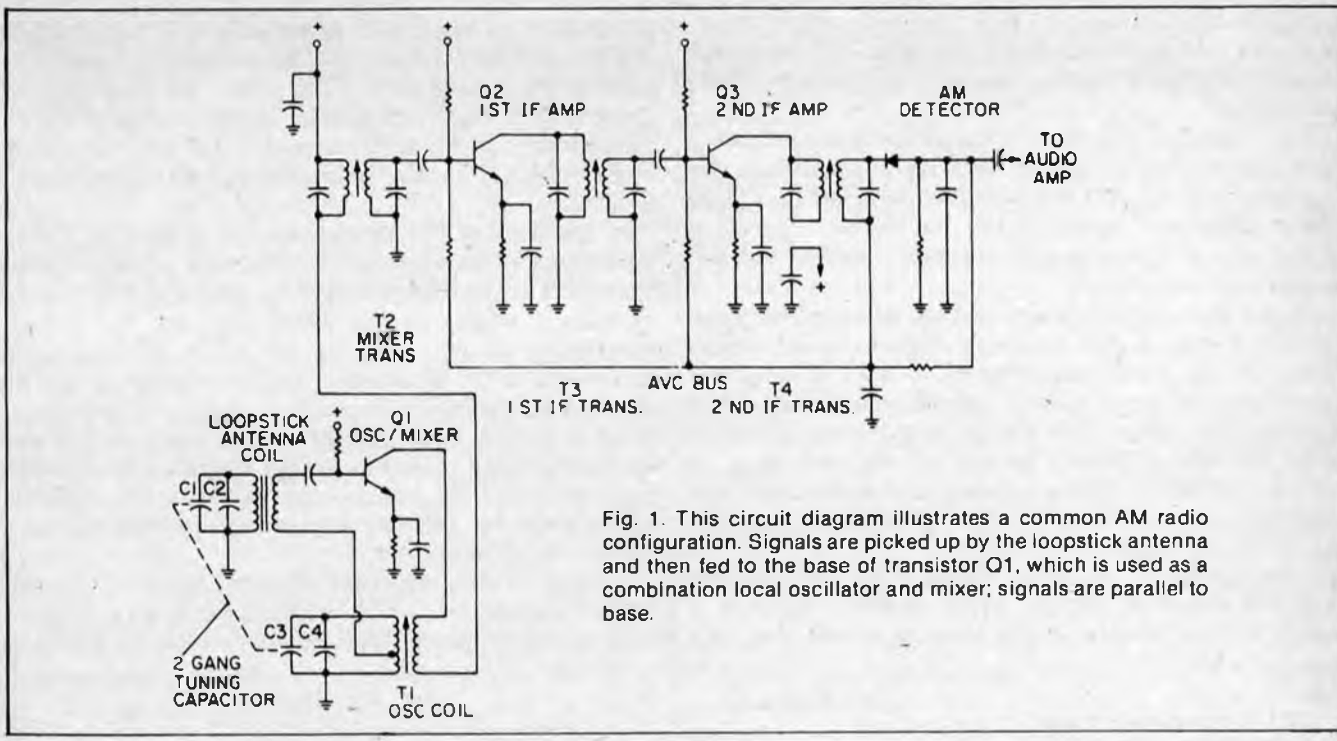


Fig. 1. This circuit diagram illustrates a common AM radio configuration. Signals are picked up by the loopstick antenna and then fed to the base of transistor Q1, which is used as a combination local oscillator and mixer; signals are parallel to base.

In Fig. 1., simplified AM radio schematic, you will note a combination mixer-oscillator stage, and two stages of IF amplification. This circuit is typical of a minimum cost AM receiver which has a two gang tuning capacitor and no FR amplifier. Such a circuit might be used in a common pocket-sized AM transistor radio. More expensive receivers will use a similiar circuit with the addition of a transistor stage for RF amplification.

In the circuit of Fig. 1, the received signals is picked up by the loopstick antenna and is fed to the base of Q1. This transistor stage is used as a combination local oscillator and mixer, and is actually a Hartly oscillator with the received signal being placed in series with the base drive of the oscillator. The frequency of oscillation is determined by C3, C4 and T1. At the same time C1, C2 and the loopstick antenna are tuned to the received signal frequency. The output of T2 contains several frequencies: the radio station frequency, the local oscillator frequency and two new frequencies equal to the sum and difference of the local oscillator and received frequency.

It is the function of the two IF stages, Q2 and Q3, to amplify the difference frequency (IF) and reject all others. This is accomplished by tuning all IF

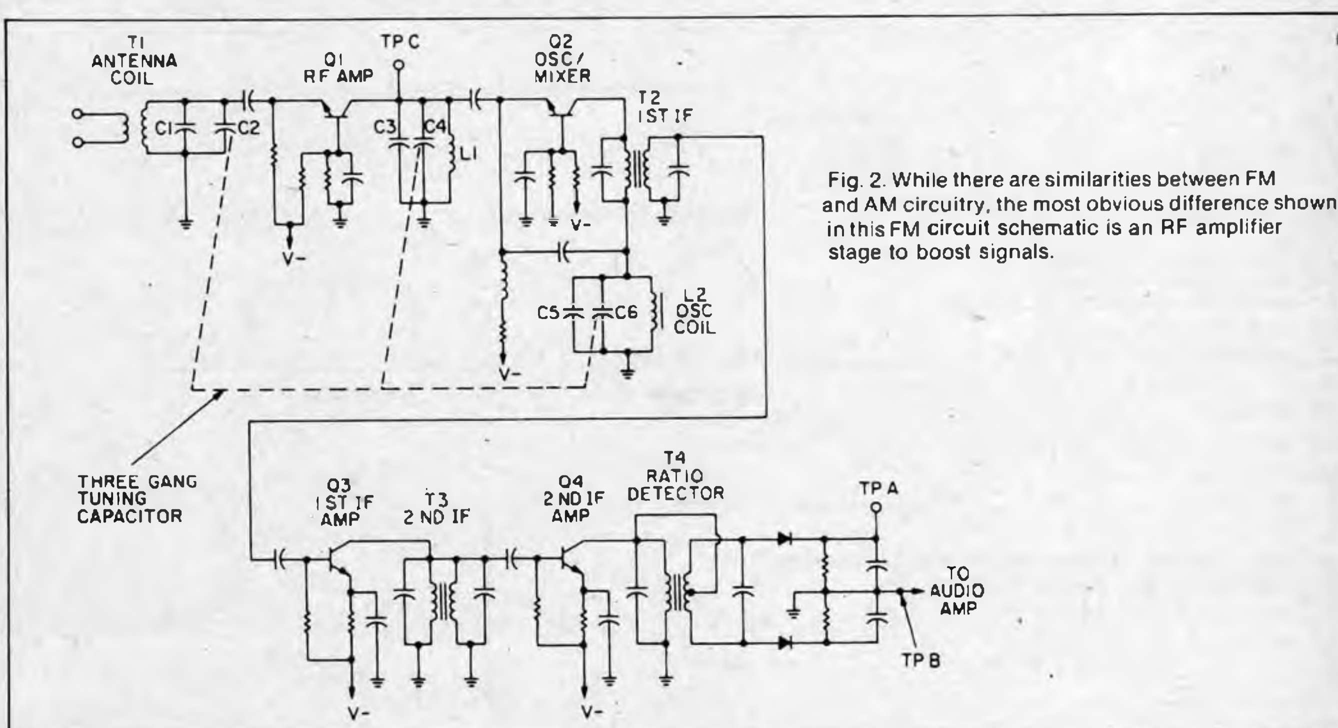


Fig. 2. While there are similarities between FM and AM circuitry, the most obvious difference shown in this FM circuit schematic is an RF amplifier stage to boost signals.

transformers, T2, T3, and T4, to the specified frequency. For most AM receivers, this is 455 KHz. The tuning of these transformers is performed during alignment of the receiver.

FM Circuitry. Fig. 2, is a simplified schematic of a typical FM receiver. In this diagram you will note the similarity with the AM receiver schematic of Fig. 1. This basic difference, aside from the higher operating frequency, is that the FM receiver employs an RF amplifier stage, Q1.

Most FM receivers have at least one RF amplifer, since it is the nature of FM transmission that weaker signals than AM are usually encountered. Good FM reception requires a solid signal in the IF amplifier, so that effective limiting takes place. Note also the requirement of a three gang variable capacitor instead of the two gang as appears in Fig. 1. The additional cost and size of this capacitor is one reason why most AM receivers have no RF amplifier stage.

Although the IF amplifier stages of the AM receiver and FM receiver appear to be similar, there is a substantial difference in the way in which they are designed. The bandwidth of an AM radio station is just 10 KHz, and the receiver must be designed to have an IF bandwidth no greater than this. Such a narrow bandwidth is easily controlled by the design of the IF transformers which, when turned to the same frequency such as 455 KHz, will provide the desired bandwidth of 10 KHz.

In the case of FM reception, and especially Stereo Multiplex FM, good reception requires a receiver with at least 150 KHz bandwidth. It is the nature of FM to have significant sidebands on either side of the carrier frequency, as far away as 75 KHz. Such a wide bandwidth in the IF stages of an FM receiver cannot be attained by tuning each stage to the same frequency. What is done is to stagger tune each stage, so that the resultant overall response has the required bandwidth. Note that each IF transformer in Fig. 2 has separate tuning slugs for primary and secondary which permit stagger tuned alignment.

Because of the problem of attaining proper alignment in the IF stages of an FM receiver, the more expensive designs utilize special filter circuits which are tuned to

Fig. 3. A matching circuit must be used when attaching sweep generators to FM sets.

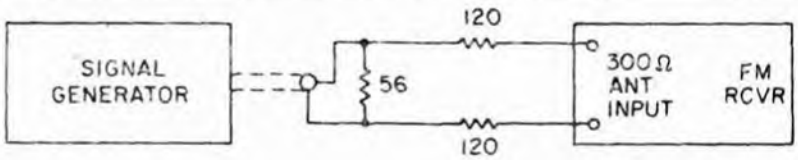


Fig. 4. When using both a sweep generator and oscilloscope for aligning an FM set, they must be connected in the manner shown in the diagram on the right side.

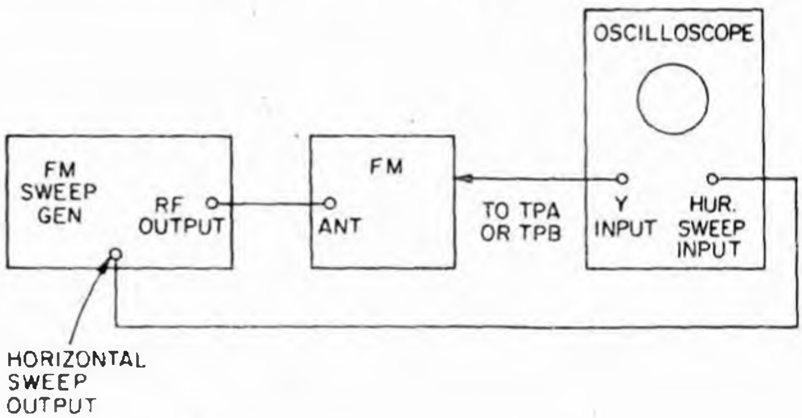


Fig. 5. The lower-quality FM receivers will have a waveform resembling A; while better sets will show something more like B.

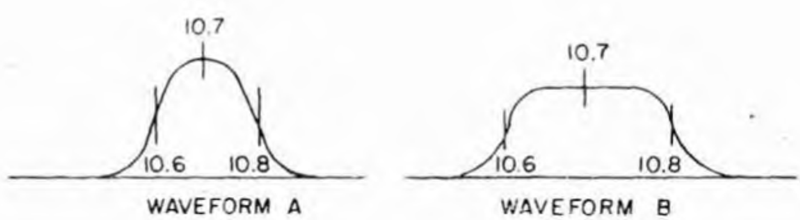
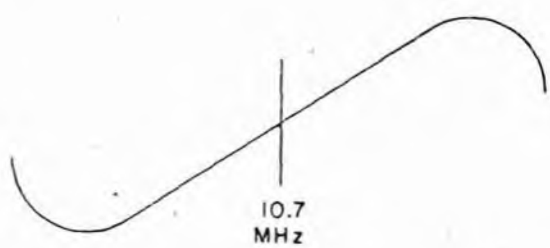


Fig. 6. Adjust the secondary of the discriminator transformer 'til it centers 10.7 MHz.



the factory and require no adjustments in the field. You will find that many modern stereo receivers are designed this way. However, these types of receivers still require alignment of the RF and local oscillator sections of the unit as you would expect.

AM Receiver Alignment. Alignment of an AM receiver is a relatively simple procedure, and can actually be performed by using existing radio stations as a signal source instead of a signal generator. However, if a signal generator is available, it is always best to use it as described in the following paragraphs.

When performing the alignment of an AM receiver, connect a few loops of wire across the output cable of the signal generator and loosely couple the loops to the AM antenna coil. Use the smallest RF output from the generator that will produce a reliable meter reading. For a tuning indicator you can connect a VTVM or VOM set to measure AC volts across the voice coil of the speaker.

An alternate method is to measure the DC voltage level on the AVA (Automatic Volume Control) line of the receiver. This requires a high impedance DC voltmeter such as a VTVM.

The alignment procedure will, of course, depend upon the number and type of adjustments in the specific receiver under test. If possible, obtain the manufacturer's alignment procedure. In the absence of such information you can use the following procedure.

IF Alignment. The first adjustment to be performed is the IF alignment. Set the signal generator at 455 KHz. with about 30% to 90% amplitude modulation. Loosely couple the output of the generator to the antenna coil of the receiver and listen for the audio modulation in the receiver speaker as the RF generator is varied about 455 KHz. If you do not obtain any audio around this frequency, the IF of the receiver may be another frequency, such as 262 KHz.

Once you have determined the proper frequency, set the generator at 455 KHz. or whatever the specified intermediate frequency may be. Connect the VOM across the voice coil of the receiver and adjust the IF transformers (T2, T3, and T4 in Fig. 1) for the maximum reading of the meter.

As you progress with the IF alignment, you may find that you can lower the RF output of the signal generator to prevent overload of the receiver. When you are satisfied that all IF transformers are tuned so that no further increase in meter reading can be attained, the IF alignment is complete.

The next set of adjustments will align the upper and lower settings of the tuning capacitor. This is accomplished in two steps, which will have to be repeated several times, since the adjustments have some interaction with each other.

The initial adjustments are made at the high end of the broadcast band. Set the tuning dial to a silent point (where no radio station is received) near 1600 KHz. Set the RF signal generator to the same frequency. Adjust the two timer capacitor (C2 and C4 in Fig. 1) for maximum meter reading.

Then set the tuning dial of the receiver to a silent point around 600 KHz. and set the signal generator to the same frequency. Adjust the oscillator tuning coil (T1 in Fig. 1) for maximum meter reading.

Repeat the adjustments for 1600 KHz and 600 KHz as described above until no further improvement can be made. Once you have done this, the AM alignment will be completed.

FM Receiver Alignment. Proper FM alignment requires the use of a sweep generator and oscilloscope, unless the receiver under test has a fixed tuned IF section which is permanently aligned at the factory. The generator may be coupled to the antenna terminals of the receiver using the circuit of Fig. 3, which matches the 50-ohm impedance of the generator to the 300-ohm input impedance of the FM receiver.

Some receivers may have highly selective front ends which make injection of the 10.7 MHz IF through the RF stages difficult. In these receivers you may couple the output of the signal generator to the input of the first IF amplifier stage through a 10-pf capacitor. This point is shown in Fig. 2, as test point C. Use only sufficient signal strength to achieve a usable display on the oscilloscope. Fig. 4, is a typical connection diagram of the receiver, generator and oscilloscope.

The first section of the FM receiver to be aligned is the IF amplifier. Set the signal generator to a center frequency of 10.7 MHz. and a sweep width of about 450 KHz. An FM sweep generator should have a marker system which allows you to identify the important frequencies of interest, such as 10.6, 10.7, and 10.8 MHz. These markers indicate the desired bandwidth of the IF amplifier.

Connect the Y input of the oscilloscope to test point A of Fig. 2, and adjust the scope controls to obtain a display similar to Fig. 5. Adjust all IF transformers (T2, T3, & T4 in Fig. 2) so that the amplitude of the display is maximum, while preserving the symmetry of the waveshape at about 10.7 MHz. Bear in mind that low cost FM receivers will have a waveshape similar to that of waveform A, while high quality receivers will have the preferred waveshape B of Fig. 5.

The next step is the adjustment of the Ratio Detector or Discriminator. Connect the Y input of the oscilloscope to test point B of Fig. 2, and adjust the secondary of the Ratio Detector or Discriminator transformer (T4 of Fig. 2) to place the 10.7 MHz marker at the center of the S curve, as shown in Fig. 6. Readjust the primary of the same transformer for maximum amplitude and straightness of the line between the positive and negative peaks. This completes the alignment of the IF section of the receiver.

Trimmer Adjustments. The final adjustments to the FM receiver are made to the trimmer capacitors, which are connected in parallel with the main FM variable capacitors. These are shown in Fig. 2 as C1, C3 and C5. Note that some FM receivers, especially the higher quality ones, may have four such trimmer capacitors. These adjustments are made with the FM tuning dial set to near the high end of the range, (108 MHz).

Set the FM dial to a silent point near 108 MHz. Set the signal generator to the same frequency as the dial setting, and couple the output of the generator to the antenna terminals of the receiver as shown in Fig. 3. Connect a DC voltmeter to test point A as shown in Fig. 2, and use only enough signal output from the generator to attain a reliable meter reading. Adjust C1, C3, and C5 as shown in Fig. 2, for maximum meter reading.

As a final check, set the receiver dial to a silent point near the low end of the range, 88 MHz, and set the generator frequency to the same frequency. Now, by varying the generator about 88 MHz while watching the DC voltmeter, you can determine how accurate the FM dial is. If necessary, you may adjust the oscillator coil L2, or slip the dial pointer so that the dial reading agrees with the signal generator frequency

Keep in mind that if you do make this adjustment you will have to go back to 108 MHz and readjust the trimmer capacitors again. In addition, if you are working on an AM/FM receiver slipping the dial pointer will also require a realignment of the RF section of the AM receiver. For these reasons it is not recommended that the dial be slipped unless it already has been placed far out of calibration through some previous servicing procedure.■

Horizontal Dipole Antenna

The Horizontal Dipole is a simple effective antenna. Antenna wire and accessories required for a horizontal dipole antenna construction are inexpensive and readily available. In fact make-do replacements from surplus sources cost even less, or may be available for free.

It is the length of the dipole antenna that makes for efficient use as a receptor of an incoming radio wave. When used as a transmitting antenna, it radiates efficiently and, at the same time, displays a proper impedance match to the output of the transmitter.

Dipoles are of critical length when used as transmitting antennae because of the need for proper match. Length and match insure ideal characteristics for reception as well. It is true that length and match are less critical if the antenna is to be used for receive only, because of the high sensitivity of modern receivers. For example, an antenna cut to length on one of the lower-frequency short-wave broadcast bands, will also receive well on ohter international broadcast bands. However, a dipole cut to length and properly oriented may be of considerable help in receiving some favorite, but very weak station.

Dipole Length. The physical length of the antenna is related to the wave-length of the signal frequency to be received or transmitted. Frequency in megaHertz, and wavelength in meters are related as follows:

$$\frac{\textbf{Wavelength}}{\textbf{(meters)}} = \frac{\textbf{300}}{\textbf{Freq. (MHz)}}$$

For example, the wavelength of a 3.75 mHz signal frequency would be:

$$\textbf{Wavelength} = \frac{\textbf{300}}{\textbf{3.75}} = \textbf{80 meters}$$

A dipole is a half-wavelength antenna and, therefore, its theoretical length would be one-half of this value, or 40-meters long. In practice, however, there are capacitive end-effects which cause a dipole that is cut to exactly the so-called "free-space" wave-length to be resonant on a lower frequency than the calculated value. In fact, to make the antenna an exact "electrical" half-wavelength long, it is necessary to shorten the physical length by 5-percent. Hence the dipole length for 3.75 MHz resonance would be:

$$\textbf{Dipole Half-Wavelength} = \textbf{0.95} \times \textbf{40} = \textbf{38 meters}$$

Since the dipole antenna is fed at the center and separated into two quarter-wavelength segments as shown in Fig. 1, each side of the antenna would be 19 (38 ÷ 2) meters long.

Physical antenna length for each quarter-wave segment of the dipole can be obtained by multiplying the 19 meters times the meters-to-feet conversion constant of 3.2808, obtaining a value of 62.34 feet.

A conversion from metric to linear length results in a very simple equation that can be used to determine the length of the quarter-wavelength segment of a dipole:

$$\textbf{Length (feet)} = \textbf{234/f(mHz)}$$

A hand calculator is an aid if you wish to make your own antenna calculations.

Dipole Dimension Charts. Quarter-wave segment

FREQUENCY IN MHz	DIPOLE λ/4 IN FEET
160 METERS	
1.81	129.28
1.83	127.86
1.85	126.49
1.87	125.13
1.89	123.80
1.91	122.91
1.93	121.24
1.95	120.00
1.97	118.78
1.99	117.59
80 METERS	
3.525	66.38
3.550	65.92
3.575	65.45
3.600	65.00
3.625	64.55
3.650	64.11
3.675	63.67
3.700	63.24
3.725	62.82
3.750	62.40
3.775	61.99
3.800	61.58
3.825	61.18
3.850	60.78
3.875	60.39
3.900	60.00
3.925	59.62
3.950	59.24
3.975	58.87
40 METERS	
7.025	33.31
7.050	33.19
7.075	33.07
7.100	32.96
7.125	32.84
7.150	32.73
7.175	32.61
7.200	32.50
7.225	32.39
7.250	32.28
7.275	32.17
20 METERS	
14.025	16.69
14.050	16.66
14.075	16.63
14.100	16.60
14.125	16.57
14.150	16.54
14.175	16.51
14.200	16.48
14.225	16.45
14.250	16.42
14.275	16.39
14.300	16.36
14.325	16.33
15 METERS	
21.05	11.12
21.10	11.09
21.15	11.06
21.20	11.04
21.25	11.01
21.30	11.98
21.35	10.96
21.40	10.94
10 METERS	
28.2	8.30
28.4	8.24
28.6	8.18
28.8	8.13
29.0	8.07
29.2	8.01
29.4	7.96
29.6	7.91

Table 1—Here are the dimensions for cutting half-wave dipoles for the various Amateur frequencies below 30 MHz. The number given represents one-half the total antenna length.

DECIMAL PART OF ONE FOOT	INCHES (APPROXIMATE)
0-.05	0
.05-.10	1
.10-.15	2
.15-.25	3
.25-.35	4
.35-.45	5
.45-.55	6
.55-.65	7
.65-.75	8
.75-.85	9
.85-.90	10
.90-.95	11
.95-1.0	12

Table 2—Use this table to convert the decimal portion of the feet measurement given in Table 1, to inches. Always cut your antenna a bit longer at first. It's easier to trim it down than to add length later.

lengths for each of the Amateur bands, 10 through 160 Meters, are given in Table 1. For example, each quarter-wave segment of a dipole antenna cut to 14.2 MHz in the 20-Meter band should have a length of 16.48 feet. Table 4 gives dimensions for dipole quarter-wave segments for reception on the various shortwave broadcast bands. Dipole lengths for the various WWV frequencies are given in Table 3.

Lengths are given to a decimal part of one foot in the tables. In addition, Table 2 permits you to make an approximation in inches. In fact, when erecting an antenna for use with a transmitter, there are other variables, such as proximity to ground and metallic surfaces, as well as antenna element diameter, that influence the exact resonant frequency. Therefore, cutting an antenna within an inch or two of indicated value is satisfactory. For example, in cutting the dipole for 14.2 MHz use, a practical value is 16-feet, 6-inches. Note from Table 2 that the six-inch figure is appropriate for a decimal part falling between 0.55 and 0.65.

It has been my experience in cutting antennas, that it is preferable to cut elements somewhat on the long side, permitting you to trim down the antenna to a specific resonant frequency after initial tests have been made. Of course, antennas for receive-only use are not nearly so critical as to their length. Consequently, in Table 4, one dimension is given for operation over the specific shortwave broadcast band.

Of course, it should be stressed that an inch or two of length has a much more decided effect on the resonant frequency on a higher frequency Amateur band as compared to a lower one. Thus, you should be more careful in cutting the dipole for 10 or 15 Meters, as compared to the cut for 80 to 160 Meters. For example, a differential of 3-inches on 10 Meters might result in a frequency change of approximately 1 megaHertz, while a similar differential on 80 Meters corresponds to a frequency shift of only 20-25 kiloHertz.

Dipole Directivity. The horizontal dipole is directional. As a transmitting antenna, it sends out maximum radio energy (radiation) in the two directions broadside (perpendicular) to the antenna wires (Fig. 2). As a receiving antenna, it displays maximum sensitivity to radio signals arriving from the same two directions. RAdiation and sensitivity taper off at angles away from the perpendicular, declining to a minimum in the direction along the line (parallel) of the antenna wire. The response pattern of Fig. 2 is a theoretical one. The antenna does radiate energy at other angles and is sensitive to incoming signals as well. The extent ot the differential depends upon a number of variables including type of antenna, proximity of ground, nearby metallic structures, propagation conditions, transmission line system, etc. It is a fact though, that maximum radiation and sensitivity occur perpendicular to the antenna wire and minimum in the direction of the antenna wire. The figure-eight pattern is itself rather broad, and it is only at angles near to the angle of the antenna wire that the response is sharply down.

In practice then, it is a good idea, if possible, to erect the dipole antenna with an orientation that places it broadside to the direction toward which you wish to radiate maximum signal or display maximum sensitivity. If your intent is to minimize the pickup of an interfering signal, you should point the dipole antenna wire in that direction.

Dipole Antenna Components. Essential components

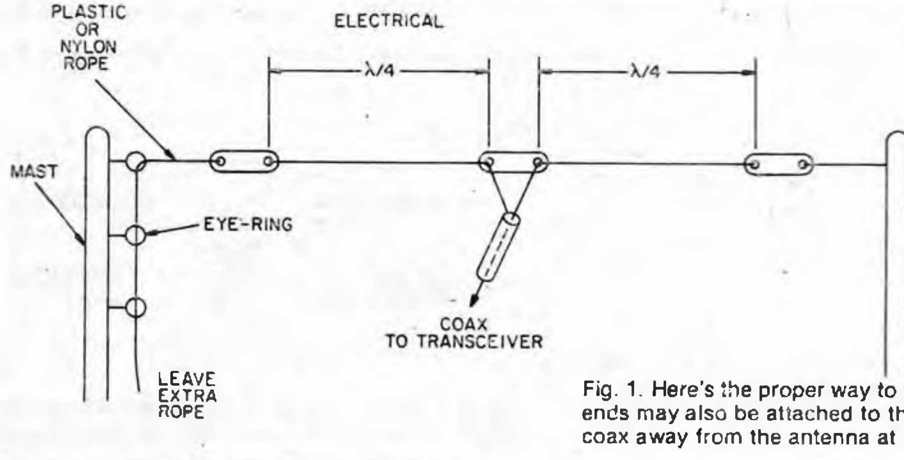


Fig. 1. Here's the proper way to erect a dipole. Either or both ends may also be attached to the house or garage. Lead the coax away from the antenna at right angles.

Table 3—For receive-only operation, the dipole is still a very good choice. Here are the optimum lengths for broadband operation. Remember to orient the antenna for the area you wish to listen to specifically.

BAND METERS	FREQUENCY IN MHz	DIPOLE λ/4 IN FEET
120	2.3-2.495	97.5
90	3.2-3.4	70.9
75	3.8-4.0	60.0
60	4.75-5.06	46.8
49	5.95-6.2	38.36
41	7.1-7.3	32.5
31	9.5-9.775	24.4
25	11.7-11.975	19.8
19	15.1-15.45	15.3
16	17.7-17.9	13.15
13	21.45-21.75	10.8

of the dipole antenna are: antenna wire, dipole center connector, end insulators, support rope, transmission line, and other accessories as needed. The antenna wire can be the popular 7-strand, #22 type, which is common and inexpensive. When it can be found at low cost, our personal preference is for #14 or #16 solid, insulated wire. A good-quality, insulated wire gives you added safety and weather protection. Insulation in no way interfers with the radiation or pick-up of signal.

Available end insulators are usually made of porcelain and are 1.75 to 3-inches long. They are oval-shaped or rectangular, some having a ribbed construction. Two holes are provided, one for the antenna wire itself and the other for the support line. Support line can be nylon rope or strong plastic clothes-line with a nonmetallic core. To make it easy to lower the antenna, for cleaning or experimentation, the support line at one end can be fed down through eye-bolts to ground level, as shown in Fig. 1.

A coax-to-dipole connector, Fig. 3, is the ideal method of linking the dipole antenna to the coaxial transmission line. This connector provides a durable and reasonably weather-proofed connection, providing for convenient connection and detachment of transmission line. An alternative plan is to use an end insulator at the center. The two condustors of the transmission line can be attached firmly, soldered and taped to the antenna wire on each side of the center insulator.

There are various support means for horizontal dipole antennas. The variety of TV-antenna hardware such as chimney, side-wall and roof mounts, permit easy attachment to a house or garage. Support itself can be a 5 or 10-foot section of TV mast. Free-standing and guyed masts are available for ground-mounted supports. A telescoping TV mast is versatile because of its ease of erection and let-down. Guying is required. Guy rings are spaced approximately every 10-feet along such a telescoping mast.

Use good quality coaxial line, either 50-ohm or 70-ohm. Preferred types are RG-58A/U (50 ohms) or RG-59A/U (70 ohms) for low power applications. RG-8A/U is recommended for higher-powered applications, and installations where a long feed line, from antenna to transmitter, is necessary.

Erection of Dipole. Plan your installation according to length, height, and directional orientation. You must consider the space required by the antenna, and where the line must be brought into the house.

Safety and performance are important criteria. For safety reasons, keep the antenna clear of power lines. Be certain that if the antenna falls when erected, or while under erection, it cannot fall across electrical wires. Make certain that under no circumstances, can mast or wire come in contact with power lines if you lose control of the mast or antenna. Keeping clear of power lines also improves the antenna performance. You will pick up less power line noise on receive. On transmit, you will radiate the least signal into the power lines, minimizing loss and possible interference with home entertainment units such as television receivers and high-fidelity amplifiers.

Orient the antenna to best meet your needs. If you wish to radiate maximum signal east and west, orient

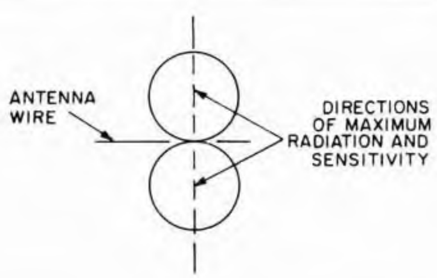


Fig. 2. If you're interested in working a particular area with yur dipole, then be sure to orient it properly. The maximum amount of RF is radiated perpendicular to the line of the antenna wire itself.

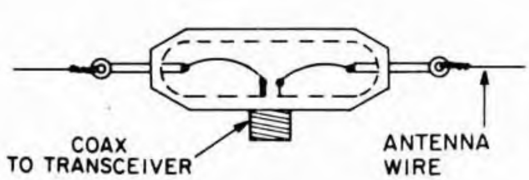


Fig. 3. The best method for mating the antenna to the transmission line is a coax-to-dipole insulator/connector.

FREQUENCY IN MHz	DIPOLE λ/4 IN FEET
2.5	93.6
5	46.8
10	23.4
15	15.6
20	11.7
25	9.36

Table 4—Here are the dimensions needed to cut a dipole for WWV time station frequencies. WWV is an excellent source for receiver frequency calibration as well as the correct time. WWV's transmitters can be heard world-wide, 24-hours every day.

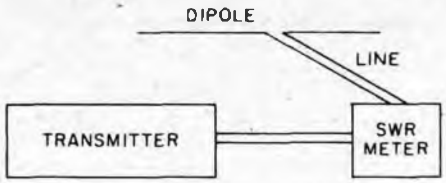


Fig. 4. An SWR meter should be inserted in the transmission line between the transmitter and antenna for accuracy.

the antenna wire north and south. In a built-up area, it is not always possible to find an ideal mounting situation. However, within reason, it is not necessary that the two antenna ends be the same exact height above ground. Neither must the two quarter-wave segments of the dipole be in an exact line. Stay as close as you can to the idealized situation, but don't worry if you must make limited departures. The antenna will still perform well if you are reasonable in the changes you make.

Receive Only. The same general considerations must be made in the erection of a receive-only antenna, with the exception that power handling capability and transmitter matching are no longer factors of concern. Thus the antenna need not be cut as precisely. The two-wire transmission line can be made of lamp cord or, preferably, a good quality 300-ohm TV ribbon line. A combination of dipole antenna and TV line makes a good combination for short-wave listening on the international broadcast and radio amateur bands. A receiving dipole cut for 35-feet on each segment is a reasonable antenna for all-band listening. However, if you want peak performance on some particular band, you can then cut a separate receiving antenna for that particular band. Orient this antenna with its figure-eight reaching out in the favored direction.

Tuning With An SWR Meter. An SWR meter connected between transmitter and transmission line, Fig. 4, can be used to measure the resonant frequency of a dipole. To go a step further, the antenna can now be trimmed or extended if it does not resonate to the desired frequency. The results can be observed by the SWR meter, as the antenna resonant frequency is moved up or down the band. Since it is easier to trim off rather than to add on wire length, cut the initial antenna wire longer than the specified value for the particular frequency, in order to catch up with any variables that might influence resonance. A practical example will demonstrate an acceptable procedure.

Assume an antenna is to be cut for 7150 kHz in the 40 Meter amateur band. Table 1 indicates a dipole length of 32.73 feet. This suggests a dipole length of 32-feet, 9-inches. Cut each dipole element to 33-feet, which would be for a resonant frequency of 7100 kHz. Erect the antenna on a temporary basis.

Measure the SWR every 25 kHz between 7025 and 7225. Set the readings down in a table form of frequency vs. SWR. Determine the precise frequency at which the SWR reading is minimum. This would be the resonant frequency of the dipole. In our example, it was exactly 7050 kHz.

Inasmuch as the resonant frequency reading is low, you can now trim the antenna to attain the desired resonance. Be careful not to trim off too much. According to Table 1, each trimmed inch corresponds to a frequency change of approximately 20-25 kHz. In our example, we trimmed of six-inches, and increased the resonant frequency to 7141 kHz. If your SWR reading is low, and the resonant indication falls near to 7150, let well enough alone.

The plot of our experimental antenna is shown in Fig. 5. Note that even at the band edges, the SWR reading is reasonable. In the range between 7050 and 7250 kiloHertz, the SWR meter indicated almost ideal performance.

Antenna Tuners at Work. The primary function of an antenna tuner, Fig. 6, is to provide a proper match between your antenna system and transmitter. In so doing, your transmitter sees a proper load and is able to operate at the optimum conditions of its design. The tuner does not alter the performance of the antenna or the SWR on its transmission line. Rather, it makes certain that an improper SWR does not result in unfavorable operation or possible damage to your transmitter. Primarily it is a transmitter protector.

However, a tuner has a number of secondary benefits that enhance antenna system experimentation and permits the use of antenna systems that are not, in themselves, ideal for matching the standardized 50-ohm output of modern ham radio equipment. Again, it must be stressed that the tuner does not influence the

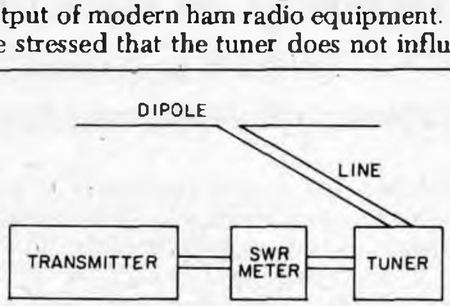


Fig. 6. The antenna tuner allows the transmitter to work into an optimum load even if the dipole itself is not cut for the correct frequency. A tuner won't correct an inefficient antenna installation.

f(mHz)	SWR
7.005	1.65:1
7.025	1.5:1
7.050	1.35:1
7.075	1.21:1
7.100	1.1:1
7.125	1.08:1
7.150	1.06:1
7.175	1.1:1
7.200	1.2:1
7.225	1.36:1
7.250	1.51:1
7.275	1.75:1
7.295	2.2:1

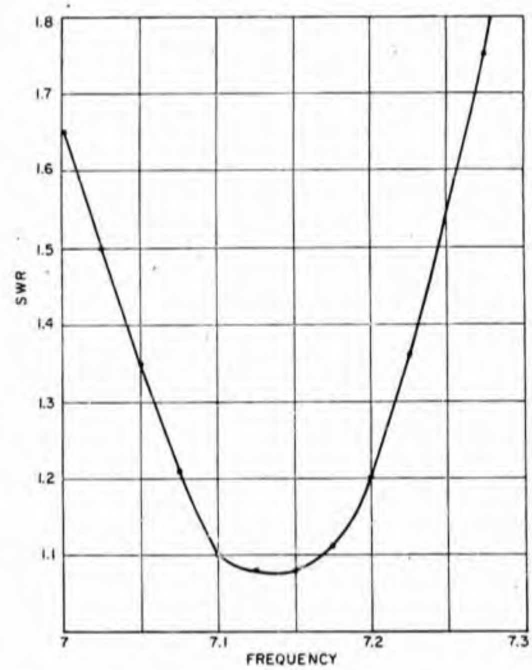


Fig. 5. Plotting the points recorded in Table 5 of measurements made of our experimental dipole for 40 meters pinpoints the antenna's resonant frequency. Even though the curve is rather steep, we manage to achieve an SWR under 2.0:1 for just about the entire band.

Table 5—Measuring SWR vs. frequency and plotting the results will help you determine the exact resonant frequency by converting these plot points into a sample graph as in Fig. 5.

performance of an antenna system. Rather, it acts as an interface between an antenna system and transmitter.

An additional secondary benefit is that it reduces harmonic and spurious signal radiation because it blocks the path between any such signals generated by your transmitter, and the radiating antenna. The tuner also rejects incoming signals that are on frequencies removed from the desired operating band. In effect, it reduces the sensitivity of the receiving system to image and other spurious frequency components.

A tuner makes the dimensioning of a horizontal dipole antenna less critical. It extends the frequency range of operation of the antenna that will provide an ideal match to the transmitter. For example, an 80-Meter dipole cut to 3750 kHz, will be made operable over the entire 80 Meter band from 3500-4000 kHz. The electrical performance of the antenna will not differ greatly from an antenna cut precisely to some specific frequency on the band. Even though the SWR on the transmission line might be rather high at the band extremes, the transmitter itself will look into an optimum load.

Conclusion. The horizontal dipole is indeed a versatile antenna, giving good performance at low cost. It should be dimensioned properly, and should be used with an SWR meter to evaluate its performance. A tuner insures proper match between dipole antenna system and transmitter, and also extends the operating bandwidth of the dipole in terms of proper matching to the transmitter. Let the dipole start you off in your first experimental activities with antenna systems.

Propagation — Its Effect On Shortwave Communications

The propagation of all radio waves may be defined as: "The traveling of wave energy through the atmosphere." How do signals get from one side of the world to the other? The ionosphere, a region of the earth's atmosphere extending upwards from 50 to 350 miles (80-560 kikometers), has ionized layers with the ability to reflect electromagnetic energy of certain frequencies.

The purpose of this article is to review, briefly, the components of the ionosphere, some propagation characteristics, and some ways that you can tell whether or not the amateur bands are in good condition for DX communications.

The Ionosphere. In 1901, Marconi succeed in transmitting radio signals from England to America, over the bulge of the earth. Since it was known that electromagnetic waves travel in straight lines, Heaviside and Kennelly proposed in 1902 that the radio waves had been reflected from an atmospheric layer consisting of free electric charges. Heaviside called it the "electrified layer", and others called it the "Heaviside layer." Later, Appleton discovered between the two layers, and to make room for other possible layers, he called the

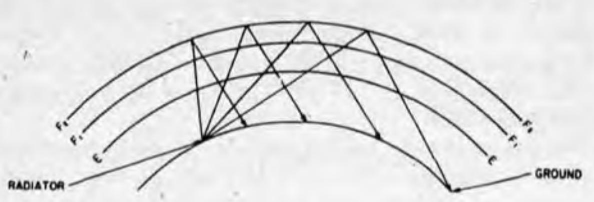


Fig. 1. The diagram above shows the effects of interaction between three layers on the reflection of radio waves back to Earth receivers.

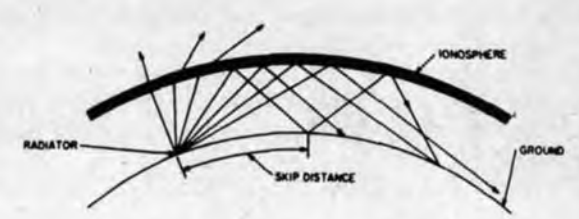


Fig. 2. This diagram demonstrates the principle of skip. You can see effects of ionospheric reflection of radio waves.

Heaviside layer the E layer, and the new layer the F layer. Subsequently a lower layer was discovered, called the D layer. It was also found that the F layer consisted of two layers which separated during the day and merged at night, and the concentration of electrons decreased in all layers at night.

The D-layer (50 miles or 80 km up) assembles during the day and disappears at night. It primarily affects 160 to 180 meters, limiting daytime coverage to groundwave by absorbing skywave signals.

The E-layer (75 to 120 km up) provodes stable skywave coverage during the day on 10 through 40 meters, and on 80 to 160 at night. The nominal skip distance is around 1,000 miles (1,600 km), but double hops can and do occur. During primarily the summer months, sporadic-E—extra-thick patchy clouds of ionization—opens 6 and 10 meters for short duration openings during the day or night.

The F-layer at night is about 200 miles (320 km) high, and during the day, it separates into the F_1-layer (125 miles or 200 km up), and the F_2-layer (250 miles or 400 km up). These layers are the mainstays of the DX'ers looking for worldwide coverage. A single hop covers 2,000 to 3,000 miles (3,200-4,000 km), and multiple hops cover up to 12,000 miles (19,300 km). What gets reflected by the F_1 and the F_2-layer depends upon the degree of ionization and the angle of the arriving signal at the ion layer.

The Maximum Useable Frequency. The maximum useable frequency (MUF) is simply defined as the highest frequency that can be used over a particular transmission path.

The MUF is affected by three variables: the 11-year sunspot cycle, the annual cycle, and the daily cycle.

Mid-1980 is generally predicted as the next peak in the sunspot cycle. Although the upcoming peak is not expected to equal the all-time historical cycle centering on 1958 when the MUF cleared 100 MHz on many occasions, 10 and 15 meters should be superlative.

The annual cycle varies from summer to winter. Due to the amount of solar radiation received by the several ion layers, the MUF is lower during the winter than summer (although atmospheric noise during the summer may hide those benefits). The overall height of the F_2-layer during the winter day is lower too, by some 200 miles (320 km) than during the summer day. The F_1-layer is also lowered somewhat. The effect is that more hops are required to cover the same distance, thus introducing more variables into the degradation of the signal over the transmission path. Winter or summer, the height of the nighttime F layer remains about the same.

The daily cycle is the most variable. The MUF is higher during the day than at night. During the short winter day, the "day" effect may seem brief. The twilight hours of DX'ing during the winter may be longer, and very interesting.

Twilight DX'ing. The rate of change of ionization is greatest at sunrise and sunset. (Listen to the broadcast band (540-1600 KhZ) for a vivid example of this.)

80 and 40 meters provide very interesting DX'ing conditions worldwide during the twilight hours. The specialty of greyline DX'ing, sending signals around the world following the day-night demarcation, is a science in itself. The key is in timing, knowing when both ends of the path will be in twilight. Take a look at *80 Meter DX'ing,* by ON4UN, John Devoldere (from the Ham Radio Publishing Group, Greenville, NH 03048), for a further discussion of this type of DX'ing.

On the higher bands, DX'ing the twilight hours, by pointing your antenna into the daylight, can maximize the coverage of your signal. In the daylight, the ionization at the reflection points of the transmission path are near maximum, and any D-layer absorption is rapidly disappearing.

Band Conditions. Aside from the three-fold variables affecting the MUF, there are other short-term, solar-induced factors that affect propagation.

Sunspots, storms on the surface of the sun, can affect conditions by several types of radiation. Sudden ionospheric disturbances (SIDs) caused by solar flares can cause instant blackouts of worldwide radio communications on the earth.

OTher, less severe, disturbances caused by intense radiation from the sunspot (or group of sunspots) tend to be longer lasting. As the radiation is sprayed outwards (just like pointing a garden hose away from you and turning quickly in a circle), sweeping across the earth, the effects on the earth may last from four to five days at a time. The sun rotates on an average 27 period, and the long-lived sunspot eruptions may have a recurring effect on the earth approximately 27 days, and multiples thereafter, until the particular sunspot (s) disappears.

The Auroral Zones. The magnetic poles of the earth draw most of the particles toward the north and south polar regions. These auroral zones are centered roughly 23 degrees from the magnetic poles.

The size of the auroral zones changes with the intensity of the solar radiation. Flares and sunspots will increase the geomagnetic activity in the polar regions.

Signals traversing the auroral zones will have a characteristic polar flutter induced on them. If the geomagnetic storm is severe enough, signals will be absorbed.

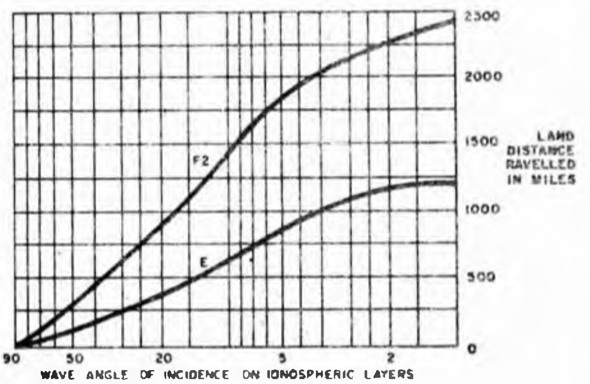


Fig. 3. The graph avoce shows the relationship between the angle of incidence of radio waves striking the "E" and "F2" layers of the ionosphere and the distance which they will travel over land. Obviously, the lower the angle, the greater the distance travelled by the waves.

As signals travel a Great Circle path, the high latitude paths will be the first to disappear as these reflection points are absorbed. As the storm continues, the lower latitude signal paths will disappear into the noise and, eventually, only signals from the south (in the northern hemisphere) will be heard. Those reflections points are outside the enlarged zone.

Sources of Information. How do you know if band conditions are good enough to warrant your time and effort to chase some DX? One time-consuming way is to tune across the bands, stopping at each signal encountered to see who it is.

An alternate method is to do some homework. A number of the amateur magazines offer long-range forecasts which are based upon extrapolations of current data. The newsletters offer more up-to-date forecasts, projecting just a week or two into the future.

Such newsletters include: the weekly *hr report* (Greenville, NH 03048), the weekly *West Coast DX Bulletin*(77 Coleman Drive, San Rafael, CA 94901), the bi-weekly *Long Island DX Association Bulletin* (Box 173, Huntington, NY 11743), and the bi-weekly *Mail-A-Prop* (11307 Clara St., Silver Spring, MD 20902). Except for *hr report*, send a #10 SASE for subscription information.

A telephone call to Dial-A-Prop (516-883-6223) will deliver a message tape 24 hours a day, seven days a week. Tapes are normally updated on Tuesdays.

Also, W1AW runs a weekly propagation bulletin amongst its daily airing of news and OSCAR bulletins. Check a recent copy of *QST* for the schedule, or ask the ARRL (Newington, CT 06111) for a W1AW schedule (send as SASE).

The most timely information is provided by WWV. Unlike the other sources cited above, the National Bureau of Standard's propagation information is oriented for the engineer and scientist, and not for just the radio amateur.

WWV's Propagation Bulletin. At 18 minutes past each hour, WWV (on 2.5, 5, 10 and 15 MHz) broadcasts a tape supplied by the National Oceanic and Atmopsheric Administration's Environmental Research Laboratories.

The message contains number counts of the solar flux, the A-index, and the K-index.

The solar flux is a measure of solar radiation: units range from 60 to 400. The higher the flux, the higher the MUF. In late 1978, the flux was averaging 160.

The A-index is a measure of geomagnetic activity ranging from 0 (extremely quiet) to 400 (very disturbed) units. An A-index of more than 20 usually wipes out polar path signals. Only on a very few occasions in the past five years have I seen this index exceed 100.

The K-index is mathematically related to the A-index and is a measure of geomagnetic activity. Its range is 0-9.

As the flux and A-index counts are updated daily at 0000 GMT, and the K-index is updated every six hours at 0000, 0600, 1200, and 1800 GMT, it is possible to track short-run variations from day-to-day and from six-hour period to six-hour period.

In practical terms, you can build a set of references for yourself by noting the WWV counts under normal and abnormal conditions. The set of references may differ for different operators.. The better the equipment and the antennas, the longer you have to work DX under worsening or marginal conditions.

Conclusion. I have only touched on some the principles and highlights of propagations and determining band conditions. One of the better books for those seeking further readings is a classic by Stanley Leinwoll, *Shortwave Propagation*, published by John F. Rider Publishers, New York, 1959. It would be worth hunting for on library shelves. For a detailed explanation of the H+18 WWV bulletin, pick up a copy of *Introduction to Solarterrestial Phenomena and the Space Environment Services Center* Stephen J. Mangis, NOAAS Technical Report ERL 315—SEL 32, put out by the U.S. Government Printing Office, Washington, 1975.

Amateur band DX'ing doesn't *have* to be hit or miss. Do a little reading and a little studying, and watch your DX totals go up.■

Fiber Optics

A fiber no thicker than a human hair is slimming down the bulky steel and copper clad wires that are traditionally used in the communications industry. This silicon based fiber, which can relay 1,000 messages at a time, is part of the new world of fiber optics, a light duct for communication.

Transmitting messages that one can hear by light is nothing new; Alexander Graham Bell invented a photophone over 100 years ago that transmitted speech on a beam of light. But it is only now that the potential of light beam communication has been realized.

short circuits or the associated dangers. Fibers are smaller, lighter and more rugged than cables. In addition, they are secure from outside entry since they are nearly impossible to tap.

The optical fiber works on the principle of complete internal reflection. To understand this requires a short disgression about light.

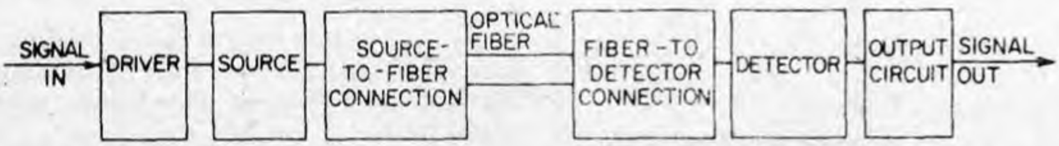


Fig. 1. Fiber optic links are being used as the transmission medium for a large number of applications. Signal in & out conectors are an important circuit part.

A Link of Light. Fiber Optics is becoming a widely used transmission link (Fig. 1). Fiber optic links were used to provide television coverage of the 1980 Winter Olympics in Lake Placid. Also, computer networks are being linked with fibers.

A major concern in communications and data transmission, is how to send larger amounts of information with greater accuracy over a smaller medium. As speeds increase and more information is handled, the difficulties of preserving signal accuracy in copper cable increases.

Problems caused by electromagnetic interference (EMI), signal radiation, inductance and attenuation increases as frequencies get larger (Fig. 2).

Many experts predict that fiber optics will be the transmission medium of the 21st century. Advantages of fiber optics links include great signal accuracy, large information carrying capacity, and EMI immunity.

Optic fibers can be used in electrically noisy environments without fear of signal disruption. This is perhaps their greatest asset. At the same time, the fiber will not radiate energy that could interfere with other equipment.

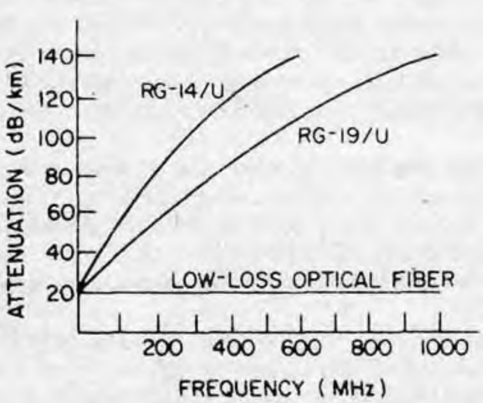


Fig. 2. Attenuation isn't related to signal frequency as it is with coaxial cable.

Secure and Isolated. Since its dielectric properties offer electrical isolation, the fiber is not prone to sparks,

What is commonly called the speed of light (3×10^8 meters or 186,284 miles per second) is actually the speed of light in a vacuum. Light travels slower in other media. When light goes from one medium to another, it changes speed and is bent.

White light is composed of all the colors in the visible spectrum. The colors, which travel at different speeds outside of a vacuum, will refract at different angles. Therefore the light emerging from a prism is dispersed into a rainbow of colors.

This dispersion is due to the index of refraction, of the prism, which is the ratio of the speed of light in a vacuum onto its speed through a material. Glass has a refractive index of around 1.5, although it varies depending on compositon. A vacuum has an index of refraction of one.

Since the speed of light in a medium is a function of its frequency, the value of n changes in the prism along with the frequency. Because of this property and the shape of a prism, white light is dispersed into its color components.

Before looking closer at light, we must define two terms; normal is perpendicular to the interface of two materials; angle of incidence is the angle between the light ray and the normal in the first material.

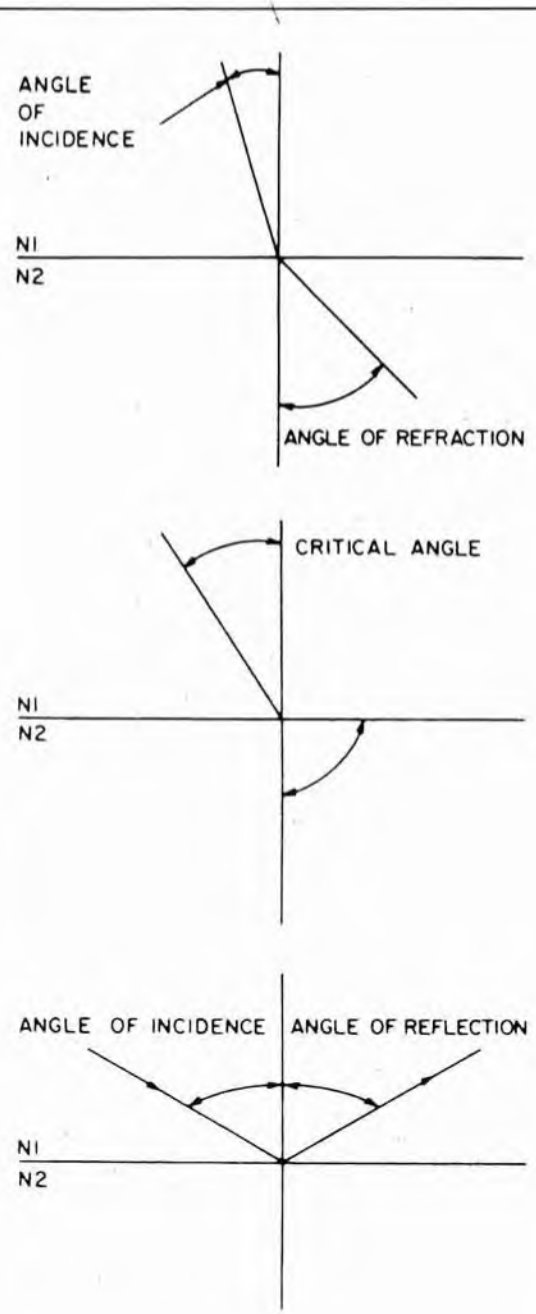
Reflection And Refraction. As light passes from a material with a high refractive index to a material with a low one, light is refracted away from the normal (Fig. 3A-C). As the angle of incidence increases, the angle of refraction approaches 90 degrees.

At 90 degrees, the angle of incidence is called the critical angle. If the angle is increased, the light will be reflected and not enter the second material. The angle of reflection is always equal to the angle of incidence.

Optical fibers work on this principle. A beam of light is partially reflected within the fibers.

Once light reflects down the optical fiber, it will continue to do so. The fiber is two layered with (glass or plastic) one layer surrounding the other. The core has a higher refractive index than the cladding.

Light injected into the core and striking the core-to-cladding interface at greater than the critical angle will be reflected bakc into the core. Since the angles of incidence and reflection are equal, the light will continue to rebound down the fiber.



Figs. 3A-C (top to bottom). Any ray of light that propagates from one medium to another will be bent. When the angle of incidence is equal to the critical angle, rays are no longer refracted within the material. All the rays are reflected from material. Note: For this case, N1 is larger than N2.

Light striking the intrface at less than the critical angle will pass into the cladding and be absorbed.

Optical fibers are constructed either as a bundle or a single fiber. A bundle, similar to stranded wire, is a number of fibers in a single jacket. Each fiber carries light independently.

Because each fiber may have slightly different characteristics, bundles are not highly efficient. Single fibers, which are today's trend, are a single core-cladding arrangement.

Glass Is The Best. Glass (or silica) is the superior optical fiber and more expensive material. Plastic is popular, especially for short distances and lower speeds. A compromise is plastic-clad silica (PCS), which uses glass for the core and plastic for the cladding.

Perhaps the most important fiber characteristics are mode and the index of refraction of the core (Fig. 4). Mode is a complex physical property describint the propagation of electro-magnetic waves. For our purposes, the possible paths of light in an optical fiber defines mode.

The most simple fiber is the multi-mode step-index fiber, which has a diameter from 125 um to over 400 um. The large core diameter allows many modes of light propagation.

Since light reflects at different angles for every mode, some light rays will propagate more than others. A ray going straight down the centr of the core arrives first at the other end. The more a ray is reflected, the later it arrives: the light has spread out (Fig. 4).

Modal Dispersion. This spreading is calle dmodal dispersion. For multimode step-index fibers, typical modal dispersion figures are 15 and 30 ns/km. For light entering the fiber at the same time, the straightest beam will trvel a kilometer 15 or 30 ns faster than the most reflected beam.

Billionths of a second may seem insignificant; often it is. However, dispersion is a primary limiting factor of a fiber's operating speed, especially over long distan ces. Dispersion is sometimes expressed as a frequency—200 MHz/km—indicating the highest operating frequency possible before dispersion affects the system.

Dispersion, and other factors, makes the multimode step-index fiber the least efficient type, although it is suited for many applications. Its large size and simple construction offer low cost, good light acceptance and working ease.

One way to reduce modal dispersion is to reduce the core's diameter until the fiber will propagate only one mode. This is done in the single-mode fiber, which has a core diameter of 2-8 um (size of a human hair).

These fibers are the most efficient, however their small size makes them unattractive for all but the long-distance, high-speed (above 50 MHz) applications that might interest cable television or telephone companies. They are difficult to manipulate and require a laser light source.

Only 5ns Per Km! Another way to reduce modal dispersion is to sue multi-mode graded-index fibers. The core's refactive index is not uniform: it is highest in the center and gradually tapers off to that of the cladding. The further the light is from the center axis, the faster it travels.

Instead of being sharply reflected, the light is gently bent. This causes the beams of light to reach the same point in the fiber closer in time than in a multimode step-index fiber. Modal dispersio can be reduced to 5ns/km.

Multimode graded-index fibers are small, typically 125-150 um. While the complexity of the core makes them a bit more expensive than multimode step-index fibers, their efficiency makes them popular.

An important feature of optical fibers is light gathering ability. The fiber should accept and transmit as much light from the source as possible. Light gathering ability

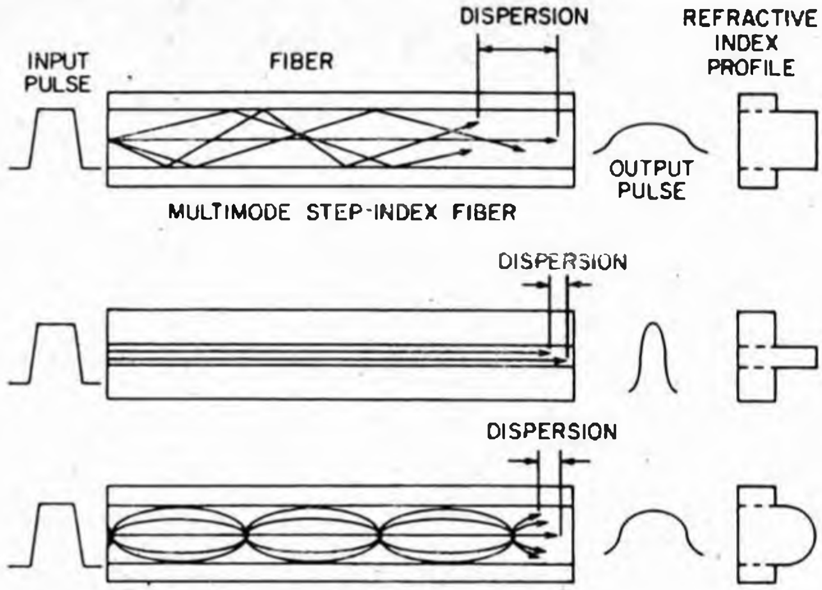


Fig. 4. The core of optical fibers have indexes of refraction that are slightly higher than those of the cladding. Most rays reflected within the core. Top: A fiber optic core allowing a wide dispersal pattern. Below this is a single mode fiber. It has minimal dispersion. Bottom: A multi-index fiber: It causes the rays to bend, thereby lowering dispersion.

depends on two things; core size and the numerical aperature of the fiber.

The Acceptance Core. A large core accepts more light than a small one. The numeric aperature (NA) is a measurement of the amount of light that can be accepted by a fiber. Because light is reflected only if it strikes the cladding at greater than the critical angle, we can form a cone: the acceptance cone (Fig. 5).

The NA is the sine of the acceptance cone's half angle O. For step-index fibers, the numerical aperature can be dervied from the following formula:

$$NA + n_1^2 \bullet n_2^2$$

where n_1 is the refractive index of the core and n^2 is the refractive index of the cladding. A large NA means the fiber will accept a large amount of light from the source.

Along with the spreading of light through dispersion, attenuation is a main consideration in describing fiber-optic link efficiency. As late as 1968, fibers with losses of 1000 dB/km were common. In 1970, the first fiber was developed with a medium loss of 20 dB/km. Presently losses are as low as 1 dB/km. Fiber losses of 10 dB/km are typical for a quality, low-loss fiber.

However, a low-loss fiber is not always needed. A high-loss fiber of 400 dB/km can be used for certain applications. For a 10-meter run, the loss is only 4 dB. Besides, medium- and high-loss fibers are usually larger and easier to manipulate.

One cause of fiber loss is through spectral attenuation (Fig. 6). Some wavelengths travel through a fiber better than others. For example, an 820 nm wavelength travels fairly well, while a 950 nm wave is sharply attenuated. Attenuation is specified by wavelength.

Sensitive Detector. Spectral attenuation raises two issues: component matching and spectral purity. If a fiber propagates an 820 nm wave well, then the source should provide a strong output at this wavelength and the detector should be sensitive to it.

In addition, the source should emit a limited range of wavelengths. LEDs have spectral widths of around 25 to 40 nm; for a laser, the width is 5nm.

Another cause of loss in a link is only a portion of source light is coupled to the fiber. Fiber-to-fiber connections introduce loss.

Dispersion effects signalling speeds; consider a pulse train. If the pulse rate is too high, dispersion will cause the pulses to overlap until the detector is unable to discern a single pulse.

Besides modal dispersion, there is spectral dispersio caused by different wavelengths traveling at different speeds. Spectral purity is important.

Connectors are an important consideration. In connecting wires, you ahve several choices: soldering, crimping, plugging, wrapping or various combinations. The dica is to achieve a connection that is strong both mechanically and electrically.

The aim of the fiber-optics connector is alignment, aligning the fiber with the source, detector and other fibers. Any misalignment causes loss; the smaller the fiber, the more difficult the alignment.

House The Fiber. One way to connect the fiber is to house it in a bushing and attach a screw-on connector holding the fiber to the bushing. With a LED, a drawback is the fiber is separated from the emitting area of LED. The emitted light spreads out, coupling only a portion to the fiber.

If attenuation does not affect signaling speeds, dispersion does. Consider a pulse train. If the pulse rate is too high, dispersion will cause the pulses to overlap until the dectector can no longer distinguish different pulses.

Besides modal dispersion, there is also spectral dispersion caused by the different wavelengths travelling at different speeds. Spectral purity is important. An LED with a spectral widt4h of 25 nm allows less dispersion than one with a 50 nm width.

Fiber Link. Besides the optical fiber, the link includes a transmitter and receiver. The transitter (LED or laser) has a drive to operate the source; the receiver (photodiode or transistor) has an output circuit to rebuild

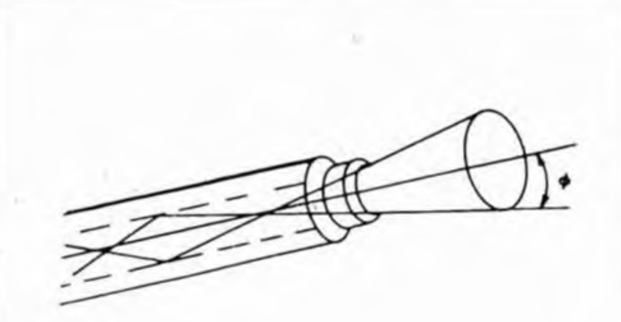


Fig. 5. The acceptance cone defines the fiber's total light gathering ability.

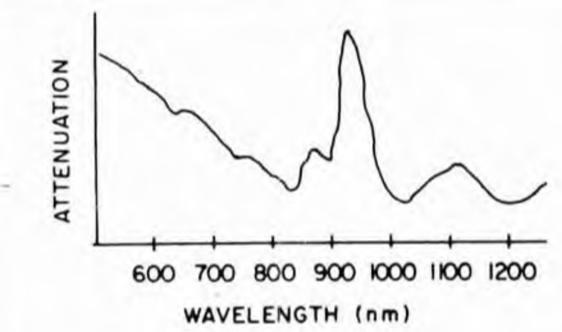


Fig. 6. Some wavelengths travel through fibers ore efficiently than others do.

the signal back into its original form. One other important part of the link is the fiber-optic connector for joining the fiber to the source, detector and other fibers.

For the hobbyist, a Fiber-Optic Experimenter's Kit is offered by AMP Incorporated through Edmund Scientific. The kit allows six TTL- and CMOS-compatible links to be built. Two links each wit speeds of one megabit, one or 100 kilobits and 10 kilobits are possible.

The kit contains four meters of fiber, a liberal assortment of AMP fiber-optic connectors and bushings, and over 70 active devices from Motorola.

Fast, reliable and now in kit form, fiber optics has a bright future.

Power Supply Primer

The catalog of a leading electronics supplier contained this glowing description: A superhet, shortwave receiver covering the standard broadcast band through 20 Meters. Its floor-standing cabinet was luxurious walnut, its audio output push-pull into a high-quality speaker. The set boasted low-current drain and the latest circuitry. The price?—a mere $49.75.

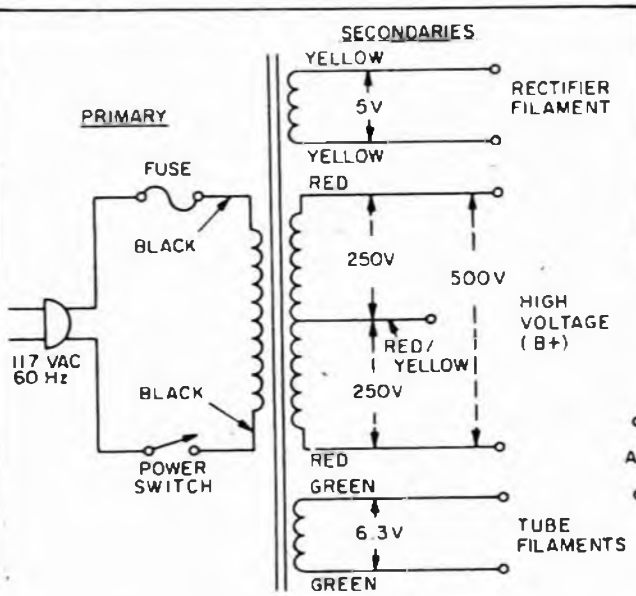


Fig. 1. Schematic diagram of typical power transformer, showing voltages and EIA color-coding of leads from the various windings.

The catalog was Allied Radio's and the date was 1932. The radio was a console meant for the living room, and it no doubt pulled in the A&P Gypsies with reasonable fidelity. The thing is, it required batteries for power.

Here's the battery complement for the handsome, but hungry, *Knight* 8 vintage receiver: three 45-volt "B" batteries for tube plates; one 2-volt "A" cell for lighting filaments; one 22.5-volt "C" battery for biasing tube grids. This mountain of Evereadys cost $9.00, a rather steep tab even in the good old days. And they could have pooped out right in the middle of a Herbert Hoover speech.

Super Supplies. That danger is gone, thanks to power supplies. Now, a receiver takes raw electricity from the utility company and converts it to filament, plate, or bias voltages. It does the same for transistorized circuits. Or it

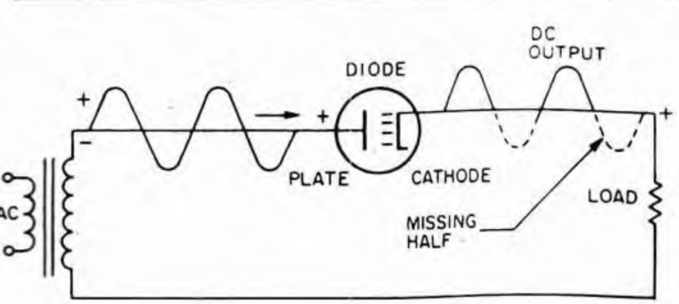


Fig. 2. Half-wave rectifier, showing how diode produces pulsating DC output by repelling electrons when plate is driven negative.

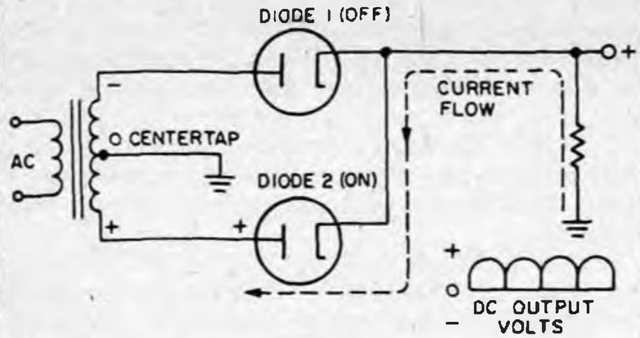


Fig. 3. Full-wave rectifier is a pair of diodes driven alternatively to take advantage of both halves of AC input voltage.

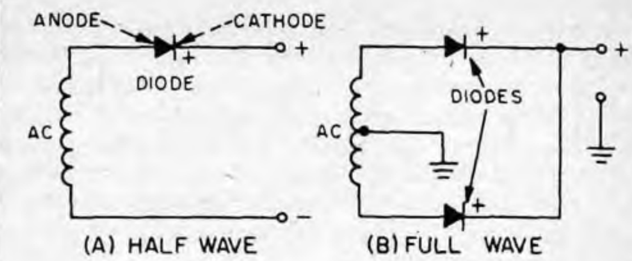


Fig. 4. Solid-state silicon diodes in similar half- and full-wave rectifier circuits are better than tubes since they don't need filaments or heaters for rectifying action.

perhaps participates in the growing trend to 3-way operation, where you use the same device at home, in a car, or carry it as a portable. The supply not only powers the equipment in the home, it also recharges the portable batteries. Cost is low because AC power is priced about 6¢ per kilowatt hour—which means you can operate a plugged-in table radio for about 50 hours on a penny.

Though power supplies operate circuits of vastly different voltage and current requirements, the basic principles are the same. In most instances a supply accepts house current—usually 117 volts AC alternating at 60 Hz (cycles)—and performs the following steps.

- *Transforming Voltage.* The power company provides 117-volts AC for home outlets, but it's hardly the value that many electronic devices demand. The plates of receiving tubes require about 100 to 250 volts for operation, while transmitting tubes may need a "B+" several hundred volts higher. Transistors, on the other hand, usually function at less than 30 volts. So the first task of the supply is to transform voltage to the desired value. In many CB sets, for example, there's plate-voltage requirement of 250 and filament-voltage requirement of 12.6 VAC. The power transformer delivers these levels.
- *Changing AC to DC.* Furnishing correct voltage is not enough. Those voltages must often be DC—and the power company provides alternating current. So the second function of a supply is to *rectify*, or convert AC to DC. If a rectifier malfunctions in your radio you'll soon learn its function. The symptom is annoying hum in the speaker (caused by 60-Hz alternations in the audio). In a TV set, suffering rectifiers can put a thick, dark, "hum" bar across the screen.
- *Filtering.* Though rectifiers change AC to DC the product is far from suitable because it contains objectionable ripple. This will be attacked y the filter, which smooths the pulsations to pure DC.

The final step of the supply depends on the designer. He can add a *bleeder*, choose a *regulator*, or insert a *divider* at the output. We'll look at these extras, but first consider how the supply's basic parts operate.

The Transformer. In Fig. 1 is a typical power transformer that's been produced by the millions with only slight variations for countless vacuum tube circuits. As we'll see, the transformer acts to create a voltage change between its primary and various secondary windings. The trick's based on the turns-ratio between the various windings. If turns in the secondary winding number twice those of the primary winding, then output voltage doubles; if turns in the secondary winding are a fraction of those in the primary winding, then a stepdown in voltage occurs.

Thus, in Fig. 1, the rectifier filament, which operates at 5 volts, however, has about five times as many turns as the primary winding. The colors shown for the windings, incidentally, are standard and observed by many transformer manufacturers throughout the world.

The *centertap* connection of a winding splits the voltage in half. In our example, the high-voltage secondary winding is capable of 500 volts across the full winding (red to red), but only 250 volts between the centertap (red/yellow) and either end. The most important job for a centertap occurs in a *full-wave* supply, as we'll see in a moment. Note that a protective fuse and a power switch are located in one primary winding lead (black) of the transformer.

Rectification. The two filament voltages from our transformer (5.0 for the rectifier and 6.3 for other tubes) will need no further processing. AC can be applied directly for filament heating (or for lighting pilot lamps on the front panel). High voltage, however, must be converted to DC before powering tube plates or transistor collectors and drains.

A circuit for changing AC to DC is a *half-wave* rectifier, shown in Fig. 2. It's based on a diode's ability to conduct current in only one direction. The rectifier cathode boils off electrons (negative) which are attracted to the plate when the plate is driven positive by incoming AC.

When the next half-cycle of the AC appears, the plate is driven negative, so electrons are repelled at this time. The net result is shown in the output: a series of positive voltage pulses appearing at the load. (The dotted line shows where the negative side occurred.)

In practical circuits the half-wave rectifier is usually reserved for light-duty power supplies. It's inefficient because it fails to make use of AC voltage half the time (during the negative pulses). Secondly, those wide spaces between pulses are difficult to filter because of low ripple frequency. In a half-wave rectifier, the pulsations occur at 60 Hz, the same frequency as the applied line voltage. But don't underrate the half-wave supply because it's widely used where the power requirements are low and the circuit is inexpensive to manufacture.

Full-Wave Supplies. Transmitters and high-power equipment overcome the half-wave's shortcomings with the full-wave system. It's nothing more than a pair of diodes that are driven alternately so they consume every bit of AC input voltage. The key to full-wave operation is the centertap on the transformer's secondary winding. As applied AC appears across the complete winding, it makes the top-end negative (as shown in Fig. 3) and the bottom-end positive.

The centertap at this time establishes the zero voltage point because it's at the common, or grounded, side of the circuit. During the time the lower diode (No. 2) has a positive plate, it does the conducting. Next, the applied AC voltage reverses and makes the top diode plate (No. 1) positive so this tube now conducts.

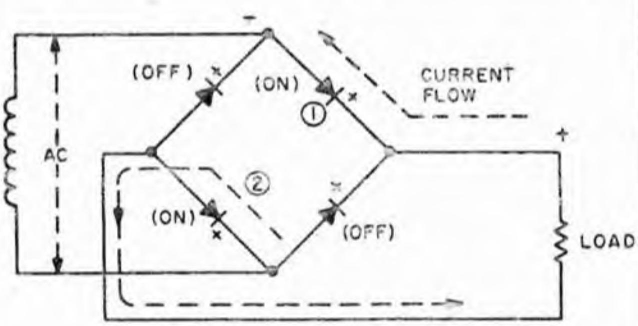


Fig. 5. Solid-state bridge rectifier provides full-wave rectification with centertapped transformer by using four diodes. Very often, the four diodes are packaged in plastic and only four connecting leads show.

This load-sharing combination of two diodes and a centertapped power transformer not only improves efficiency, but doubles the ripple frequency. An input of 60 Hz emerges as 120 Hz in a full-wave arrangement because every half-cycle appears in the output. This reduces the pulsating effect (cycles are closer together) and the DC becomes easier to filter.

If you purchase a transformer, watch out for one pitfall. It may be rated, say, "250 volts CT" and appear to be suitable for a rig with a 250-volt plate supply. In a full-wave supply, however, the transformer voltage output would be only 125, since a centertap reduces the voltage of a winding by one half. This can be avoided by specifying a transformer that has 250 volts *each* side of centertap or, stated another way, "500 volts CT."

Solid-State Rectifiers. Tube rectifiers are still widely found in electronic equipment, but they're destined for the Smithsonian Institution. Solid-state equivalents are superior because they don't need filaments or heaters to accomplish the same rectifying action. They're several hundred times smaller and much cooler in operation.

Circuits using these semiconductors, though, are similar to those of vacuum tubes. As shown in Fig. 4, diodes can be used in equivalent half- and full-wave arrangements.

Unlike tubes, though, solid-state diodes rectify AC and DC by a semi-conductor effect at the diode junction (a region between the anode and cathode). The action, in simplified fashion, occurs when "current carriers" in the material flow toward and away from the junction under the influence of applied AC. When few carriers appear at the junction, little current gets through the diode; conversely, when many carriers are in the area, they reduce the junction's opposition to current flow. Depending on the way the diode is connected in the circuit, it can recover either the positive or negative half of the AC.

Bridge Rectifiers. Another common arrangement is the full-wave bridge (Fig. 5). Though it uses four diodes, it offsets this disadvantage by an ability to produce the same output as a regular full-wave supply without a centertapped transformer. It accomplishes the feat by operating one pair of diodes during each half cycle. And as one diode pulls current out of the load, its partner pushes current into it.

The net effect is a total voltage across the load which is about equal to the applied AC. We've shown how it occurs vfor diodes 1 and 2 in the diagram (Fig. 5) but a comparable action occurs in the other diodes when the AC switches polarity.

Filtering. The next major section of the supply is the filter, which smooths out the ripple. Its two major components are often a *capacitor* and a *choke* which eliminate pulsations by dumping a small amount of current from the peak of each ripple into the "valleys" between them. The result, as shown in Fig. 6, is pure DC fit for a tube or transistor.

In operation, pulsating DC arrives at the filter choke, a coil of wire wound on a soft iron core. As the name implies, the choke attempts to oppose any change in current flow. The rippling part of the wave, therefore, encounters high reactance in the choke and fails to get through. This is aided by the filter capacitor which is charged by ripple voltage.

As the ripple falls (between pulses), the capacitor discharges part of its stored current into the "valley." Thus the combined effect of choke and capacitor results in smooth DC which can have ripple as low as a few percent of the total voltage.

You won't find the choke in some power supplies because it's an expensive item. Many designers eliminate it (especially in mass-produced equipment) by using a resistor instead, as shown in Fig. 7. The resistor does the

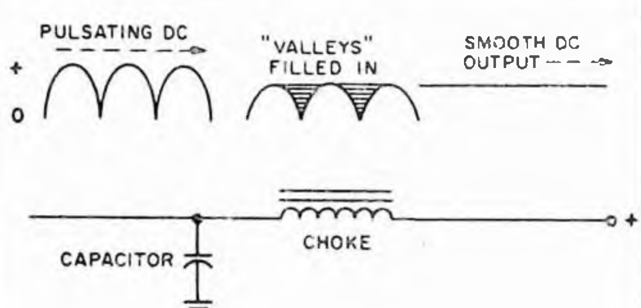


Fig. 6. Curves depicts how reactance of choke and discharge of capacitor eliminates ripple and fills in valleys for smooth DC.

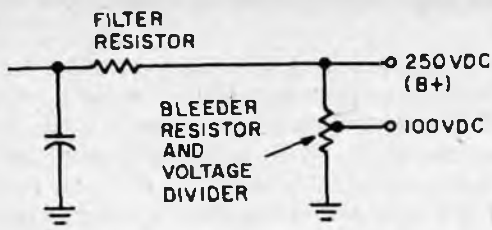


Fig. 7. Choke can be replaced by filter resistor. Also shown is bleeder resistor that serves both as output regulator as well as voltage divider.

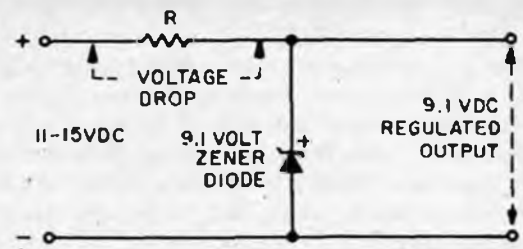


Fig. 8. For more efficient regulation than produced by bleeder resistor, Zener diode shunted across DC output will hold voltage to specified limit.

job of filtering, but with one penalty: it reduces the amount of available voltage at the output. Yet, the loss can be tolerated in many circuits and filter resistors are common.

Another use for resistors in a supply is to serve as a *bleeder*, also shown in Fig. 7. In this function, it protects parts in the supply rom possible damage due to sudden voltage surges when the supply is first turned *on*. Also, a bleeder helps stabilize voltage output when the load changes (as in a keyed ham transmitter) by always drawing some small degree of load current. Bleders, too, are found in dangerous high-voltage circuits where they bleed-off the stored charge of filter capacitors that could deliver a lethal shock to a repairman (even after the equipment has been turned *off*.)

Note that a tap can be added to the bleeder to provide a second output voltage from the supply. Now the bleeder becomes a *voltage divider*. As such, it can supply the designer with multiple output voltages for operating various devices in a circuit. Bleeder resistors are not used much in today's circuits since most solid-state components operate at only a few voltages.

Voltage Regulation. A ham who's received a "pink ticket" from the FCC for chirpy signals, a color TV that's gone fuzzy, a shortwave receiver that won't stay on frequency—all may suffer from a problem in voltage regulation. Line-voltage fluctuations or other electrical swings can cause poor, unstable operation. So the engineers have come up with methods for "stiffening" a power supply.

If, say, line voltage changes from 105 to 130, they design the circuit to operate at 100 volts. Whatever voltage arrives over the line is reduced to 100, and the surplus is dumped (usually in the form of heat). To perform this task, the regulator establishes a reference point, then regulates around it.

A common example is the Zener diode found in the power supply of many amateur transceivers. Since these rigs can operate from a car's battery or generator, supply voltage can swing from 11 to 15 volts. This could happen if you're standing for a traffic light, then pull away, causing a shift between car battery and generator. If the transciever is on at this time, receiver tuning could be thrown off because of large changes in local oscillator voltage.

A Zener diode can compensate for the shift, as shown in Fig. 8. At first glance it appears as an ordinary diode connected backward. Since the cathode (upper) terminal is connected to the positive side of the supply, there's a "reverse bias" condition. A Zener diode, however, "breaks down" (or "avalanches") whenever it is rated (Zener) voltage is exceeded. In our example, the Zener is a 9.1-volt unit, so the diode conducts current as the supply voltage shifts from 11 to 15 VDC.

Yet we see 9.1 volts indicated at the output. Secret of the Zener's ability to hold at 9.1 is that it detours part of the supply current as the voltage increases. Since a resistor is in series with that current flow, a voltage drop (as shown) appears across the resistor. Thus, any increase in supply voltage is dissipated across the resistor and effectively subtracted from the output. This automatic and continuous action occurs for any voltage above 9.1—the Zener's nominal rating—so the output is said to be regulated.

Fusing Power Supplies. You can protect your power supply by installing a fuse or circuit breaker on the primary or secondary side of the transformer. A fused primary gives good overall protection but does not have the sensitivity needed for some circuits. By installing a fuse between the transformer's secondary center tap and ground you can improve fusing sensitivity. In solid state circuits the value of the fuse is all-important due to the low voltages involved, but since fuses do not react fast enough to save a transistor or IC chip you might have to resort to semiconductor protection such as using a Zener diode.

More and Merrier. This barely brushes the subject of power supplies, since the variations are nearly endless. More than 20,000 volts for the picture tube of a color TV are derived from a special "flyback" transformer. It captures voltage from rapidly moving magnetic fields in the set's horizontal scanning section. An oscilloscope power supply contains strings of adjustable voltage dividers to move the pattern of light on the screen in any direction.

There are also high-current supplies with massive rectifiers for battery charging and super-smooth lab supplies for circuit design. But behind most of them are the sample principles which transform, rectify, filter, and regulate a voltage so it can do the job at hand. ■

Power Supplier For IC Circuits

Recently I ordered a single-board computer that is used as a development system and trainer. It is a good microcomputer trainer and is used by many schools.

I was ready to begin fingerboning 6502 commands into the microcomputer when I noticed that the machine had no power supply. In my hast to get into this type of microcomputing, I had overlooked the power supply. The company price list revealed a power supply, but it cost $50. Since it was only a 5-VDC, 4-amp, with ±12-VDC, 500-mA on the side, it was easy to beat that $50 price.

There are several popular circuits used in power supplies for typical low to medium current range microcomputers. For really big computers, which require 25-ampere supplies, we will defer to the makers of commercial supplies.

The S-100 is probably the most popular mainframe microcomputer on hobbyist market. Also, there are plenty of industrial, scientific and engineering computers on-line and many of them are S-100 machines.

The S-100 system utilizes distributed voltage regulation. All digital devices, especially, TTL, require regulated DC power supplies. One approach is centralized regulation. In these machines, there is a 5-volt DC high current power supply located somewhere in the machine. It will supply the same power to all printed circuit cards plugged into the microcomputer-s motherboard.

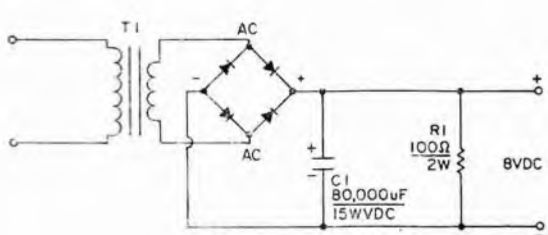


Figure 1. This circuit is designed for the simple s-100 power supply. The transformer is a heavy duty Triad F-28U. It has dual secondary voltage ratings of 6.3 and 7.5 volts AC at 25-amperes.

Lots of Current. It seems that a problem with a centralized system is that the voltage can be 5-volts only at one system location. Since these supplies draw large amounts of current there will be a sufficient voltage drop in the printed wiring tracks to reduce the voltage below the viable level for TTL devices (4.5 to 4.75 volts).

Some power supplies, those that use feedback type regulators, have a sense line sampling the output voltage. These power supplies are a little more useful because the sense line can be connected to the point on the motherboard where the voltage level is critical. The power supply will then attempt to keep the voltage at 5 volts DC at that point.

But at all other points, the voltage will be higher or lower, depending uponwhere the5-volt line is connected to the PC board and how far that point is from the point of measurement. In my own Digital Group, Inc. Z-80 system, I measured the voltage at 5.25 at one point,and 4.8-volts at the far end of the system. If it had been any lower at te extreme end, some of the memory devices may have trouble functioning properly and if any higher it would have burned out TTL devices right and left.

Another problem with the centralized regulator system is the power supply cost. A 5-volt DC, 15-ampere power supplycost. A 5-volt DC, 15-ampere power supply can cost a lot more than building an unregulted 8-volt S-100 power supply and then using buck apiece 1-ampere, 5-volt three-terminal regulators on each card.

Disaster. One final problem with the central method is the parts replacement cost if something should go wrong. If the series passelement of the power supply shorts or if the voltage reference element goes open, both events will cause the full unregulated power suply input voltage (8 volts on the S-100) onto the 5-volt bus—with disaster following for the TTL devices on the board.

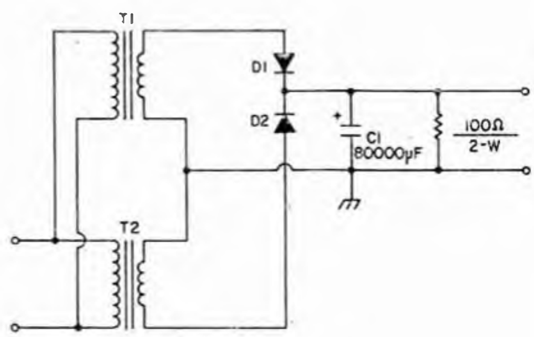


Figure 2. This is a viable but more expensive alternative, which is to use two Triad F-22A transformers in a regular full-wave rectified situation. This will deliver the 20-amperes which is needed.

When one of these big supplies opens up a shotgun blast of high voltage, there might be significant damage before the overvoltage protection circuit works. The higher the input voltage, with respect to the output (5-VDC) line, the worse the damage. On the S-100 type of power supply, only the devices on the offending regulator are affected. While the other scheme could wipe out the entire system, the S-100 will burn up only the devices on the single board.

The problem is that we need to supply the 8-volts DC required by the array of three terminal 5-volt DC regulators used on S-100 boards. Furthermore, this voltage must be backed up with not less than 10 amperes of current for even moderate systems. For some systems,

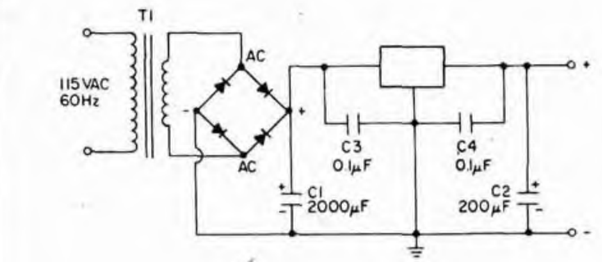


Figure 3. A substantial power supply circuit which forms the basis for the more elaborate circuit which is presented in Fig. 4.

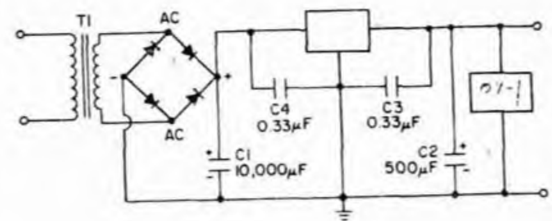


Figure 4. Probably the most commonly needed power supply for small, single-board computers, this 5-volt DC, 5 ampere circuit is the smallest to be built by the typical microcomputer hobbyist. The transformer is rated at 6.3-VAC at 8-amperes.

even more current is required.

Fig. 1 shows the circuit for the simple S-100 power supply. I have built this circuit into a Vector cabinet containing the Vector S-100 motherboard and there was plenty of room. Keep in mind when you use this cabinet, however, that the aluminum used in its construction will not support the transformer that I specify here. Place a piece of ⅜-inch plywood between the main support struts to support the transformer. Alternatively, cut up a 5- to 7-inch high piece of 19-inch rack panel to make the support.

The transformer specified here is the Triad F-28U. This heavy duty device is rated at 6.3/7/5 VAC, 25-amperes. The dual secondary voltage rating is one factor making this transformer more useful to microcomputer builders. The regulators used in S-100 require at least 8 volts DC at their input to operate.

Under fully loaded maximum current conditions, this level is sometimes hard to maintain using just a 6.3-VAC transformer, as we have only an 800-millivolt margin allowing for voltage drop. The 7.5 volts AC, however, will, when rectified, keep us in the gallgame. The voltage is selected using the primary winding.

There are three lines to the primary: a common and two voltage taps. Select the primary tap that yields 7.5-VAC across the secondary.

Bridge Rectifier. Since this transformer is use in a bridge rectified situation, we cannot legally draw more than one-half of the 25 amperes it is rated to deliver. This 12.5 amperes will supply many small S-100 systems, but is insufficient for larger ones. We can stretch the point by properlyf cooling the transformer. Overheating is due to exceeding the transformer's primary rating.

I have used this same transformer at 20-amperes in a bridge rectified situation and it ran cool, because I placed a 100 c.f.m. muffin fan a few centimeters from the transformer. Place the fan so that it will blow along the same line as the laminations to maximize cooling. This will allow the air to blow into the windings, not into the fish papr covering on th outside of the winding.

The rectifier is any of the 25-ampere, 50 volt PIV (or more) stacs that are available. I used the General Electric GEBR-425 device. Be a little wary of some companies' ratings in this type of device; they tend to be optimistic to your detriment. I heat-sinked the rectifier using one of the larger finned aluminum sinks designed to TO-3 transistors (but undrilled to I could customize it for the rectifier).

The rectifier/heatsink combination was positioned so that the air from the fan would also blow over the heatsink, improving heat transfer. Silicone heat transfer grease was used on the surface of the mounted rectifier.

The capacitor is probably the single most important part of this power supply. The nominal rating is given at 2000-uF to 50,00-uF range, depending upon whether or not you attempt to get the entire 25 amperes out of the transformer. But this is a minimum and in the case of the S-100 system, I doubt the wisdom of the standard because of the distributed regulation system used.

I prefer to use as much capacitance as I can obtain. For my own power supply, I found an 80,000-uF at 15-WVDC computer grade capacitor at a local distributer. My recommendation is some value of capacitance between 75-000-uF and 200,000-uF even if two or more units have to be wired in parallel to achieve the value.

Small Load. The 100-ohm resistor is used to place a smallminimum load on the power supply. It was found that output voltage was over +12 volts under no load conditions and that seemed a little too close to the working voltage of the filter capacitor. As a result, I tried various values of load, and found that 100-ohm, 2-watt seems the best solution.

When wiring this power supply, remember that when you are dealing with a high current situation. Use heavy duty wire and wire connectors in this supply. I recommend at least 8-gauge wire, or, two 12-gauge wires in parallel, for each run. Such wire sizes are not usually available in radio-TV parts distributers, but auto parts stores almost always carry them.

There are several alternatives to the transformer selected for this project. You could, find a surplus transformer rated at 6.3 or 7.5-VAC at bunches of current. These transformers are popular with the designers of high power radio transmitters as filament transformers and are occasionally seen in electronic surplus stroes or in the hamfest markets. Or you could use the Triad F-22A transformers.

This is a 6.3-VAC, 20 ampere model, so is rated slightly lower than the F-28U that was specified. It will deliver 10 to 20 amperes in the bridge configuration, depending upon how brave you are. A viable, if expensive, alternative, is to use two Triad F-22A transformers in a regular full-wave rectified situation (see Fig. 2). This will deliver the 20 amperes required. In this circuit, the primary windings are connected in parallel, while the secondaries are connected in series.

If you accidentally connect the secondaries backwards, then the output voltage will be zero. In that case, reverse the connections of either the primary or the secondary of one of the transformers. The problem is phasing the AC sine wave properly. Diodes D1 and D2 are stud-mounted 50 volt PIV, 25 ampere (or more) types. Again, the rectifier should be sturdily heatsinked.

Low Grade Supply. This power supply project is probably the lowest grade power supply that will be useful to microcomputerists. It is useful for small digital projects of any type using TTL (or CMOS operated at 5-VDC) and some of the smaller single board computers. The transformer can be any 6.3-VAC filament transformer rated at 2 amperes or more.

The rectifier is a bridge stack rated at 50 volts PIV, 1 ampere or more. The capacitor used in the filter fills the 2000-uF/ampere rule, although it would not hurt to used more capacitance. The output capacitor is strictly optional, although recommended because it improves the transient power supply response. It has a value that follows the rule 200-uF/ampere. It will be; happy with any value from 100 to 500-uF.

The small value capacitors are used to improve the immunity of the regulator to noise on the input voltage line and from within the computer. These capacitors should be mounted as physically close to the regulator as possible. Most builders place them directly on the regulator terminals.

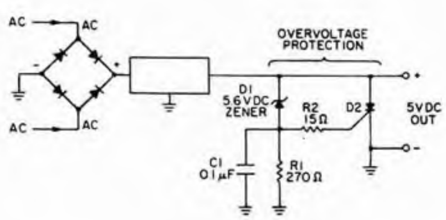


Figure 5. Circuit for adding overvoltage protection to your power supply. The circuit is called an SCR crowbar, because it represents a brute force method of doing the job of over-voltage protection.

The regulator itself can be any of the standard three-terminal 5 volt IC regulators: LM-309, 7805, LM-340-5, etc. Avoid the plastic TO-220 package types, as they should be limited to 750-mA unless well heatsinked.

Common Power Supply. The 5-volt, 5-ampere power supply (see Fig. 4) is probably the most commonly needed on small single-board computers.

The transformer for this power supply is rated at 6.3 VAC/8-amps.

There are also several regulators now on the market that will hack at 5 amperes. The regular that was selected was the Lambda Electronics, Inc. (515 Broad Hollow Rd., Melville, N.Y., 11746) type LAS-1905.

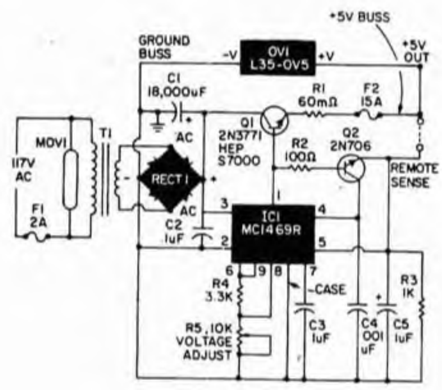


Figure 6. Circuit for a 5-volt DC, 10-ampere power supply. The high current lines must be 12 gauge wire or larger. The circuit features output current limiting to a constant 10-amps.

The rectifier used is the same 25 ampere rectifier as used in the S-100 power supply project.

Follow The Rule. The filter capacitor (C1) follows the 2000-uF/ampere rule, so a 10,000-uF unit was selected. Note that 8000-uF/25-WV-DC is commonly available.

The device labeled OV-1 in Fig. 4 is an over voltage protection module made by Lambda. These devices are designed to trip at 6.6-volts and will short the power supply line to ground if the voltage reaches this point.

If you wish to build your own overvoltage protection circuit then try something like the circuit in Fig. 5. The circuit is called an SCR crowbar, probably after the fact that it represents a brute force method of doing the job. Diode D2 is a silicon controlled rectifier (SCR) and is rated to handle at least the current of the power supply. It is connected in parallel across the +5-volt DC output lines, but remains inert until a voltage appears at the gate terminal.

This triggering voltage is supplied by Zener diode D1. At potentials less than 5.6-volts, the normal TTL operating range, the zener will not conduct current. But, at potentials greater than 5.6-volts DC the Zener conducts and creates a voltage drop across R1 that will fire the SCR. When the SCR turns on, the output lines of the power supply are shorted by ground. This will blow the primary fuse or burn out the tansformer if there is no primary fuse.

10 Amp Supply. Fig. 6 shows the circuit for a 5 volt, 10-ampere power supply. This is the circuit that I used on my first Digital Group, Inc. computer. The regulator is a Motorola MC1469R (HEP C6049R also works). There are two suffixes available for the MC1460; buy the one with the R otherwise, the rating is too low. This device is a 500-mA regulator but is insufficient to power the computer. We can boost the output current by using the

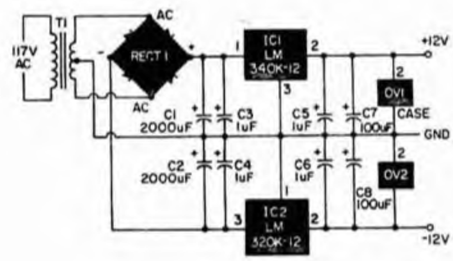


Figure 7. A plus/minus 12-volts DC power supply, this is a circuit that is needed in the S-100 microcomputer, and most others as well. This power supply is also useful for opamps, and other linear IC applications.

IC regulator to control the base circuit of a series-pass transistor (2N3771). The heavy lines in Fig. 6 are the high current lines and there must be 12 gauge wire or larger.

This power supply features output current limiting, which means that the maximum output current, even into a short circuit, will be 10-amperes. The active element in the current limiter is transistor Q2. The base-emitter bias of this transistor is the voltage drop across the 60-milliohm (R1) resistor in series with the emitter of Q1. When the current reaches 10 amperes, the voltage drop is 0.6 volts, which is sufficient to turn on Q2 and shut down the supply. You can make a 60 milliohm resistor by connective five 0.33 ohm auto radio fusistors in parallel, or by winding one of fine gauge wire yourself.

The exact output voltage can be adjusted using potentiometer R5. Set the output voltage as close to 5.0 volts as you can.

The last power supply that we will consider here is the ±12 VDC, 1 ampere circuit shown in Fig. 7. This circuit is also needed in the S-100 microcomputer and most others as well. The transformer is a Triad F-56X (or equivalent), rated at 25.2 VAC at 2.8 amperes. The extra 800-mA gives us a margin of safety. Nowt: when buying the transformer for this supply, make absolutely sure that the transformer that you buy has a centertapped secondary!

The rectifier is a 2-amp or greater bridge stack rated at 100 WVDC or more. Noté that we are not using the rectifier as a fullwave bridge, but as a pair of half wave bridges, with the transformer center tap as the zero-volt reference. The output of one half-bridge will be negative to ground, while the other is positive to ground.

The filter capacitors are selected for the 200-uF/ampere rule, but should be higher because this circuit uses halfwave rectifiers. I recommend that you use 2000-uF for light-duty loads, but 3000-uF or 5000-uF or more for the full 1-ampere per side. This will compensate for the halfwave voltage rectification.

This power supply is also useful for operational amplifiers, and other linear IC applications. ■

Putting OP Amps To Work

Few linear, (non-digital) integrated circuits (ICs) have achieved the wide popularity of the op amp in hobbyists projects of every description. Combined with few other circuit components, the op amp circuit vastly outperforms the multiple transistor circuit of yesteryear. Because you will encounter the op amp time and time again in many circuit applications and may desire to work up your own op amp circuits, it is essential to become familiar with the device.

This will introduce you to the basics of op amp circuits and applications. Four simple rules of operation are logically applied to trace and deduce the operation of the basic linear op amp circuits. Test circuits are included for "hands on" familiarity. Only a few mathematical relationships are listed to effect a comparison of the several circuits.

Some Basics. As shown in Fig. 1a, the op amp has two input terminals and one output terminal. A plus sign at one input identifies the *non-inverting* or "follower" input. When this input goes positive (with respect to the other) the output voltage, VOUT, also goes positive, thus "following" the polarity of the input. The minus sign at the other input identifies the *inverting* input. When this input goes positive (with respect to the other), VOUT goes negative, thus "inverting" the polarity of the input. In technical terms, the input stage of the op amp is a balanced "differential input" amplifier and is one which responds to the *difference* between the voltages.

Output voltage VOUT equals VIN times AVOL (AVOL equals the open-loop voltage gain as listed on spec sheets for the op amp used). Input impedance (AC resistance) *Zi* and output impedance *Soi* are shown in Fig. 1b. These are the primary characteristics of the op amp. If a perfect op amp could be constructed, AVOL and *Ai* would be infinitely large and *Zoi* would be zero. Actually, for a general purpose 741 op amp, AVOL equals 200,000, *Zi* equal 2 megohms, and *Zoi* equals 74 ohms.

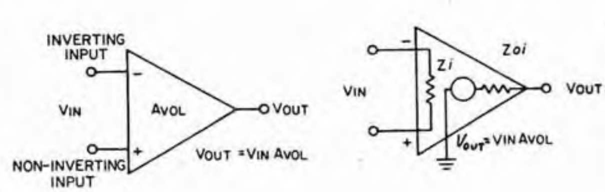


Fig. 1. Two engineering views of a basic op amp. In A, the circuit symbol also provides some insight to what the op amp will do. The circuit model in B should be referred to when investigating circuit design.

Important Concept. In linear applications such as an AC or DC amplifier, the op amp is operated closed-loop with negative feedback. To do this, the loop is "closed" by connecting a feedback circuit from the output terminal to the *inverting* input; this results in negative feedback. That porton of the output voltage fed back to the input tends to negate, or oppose, the applied input signal voltage. As will be shown later, the resulting closed-loop gain, AVCL, is very much smaller than the open-loop gain. Also, the closed-loop gain now depends on the particular feedback circuit itself and not on the actual value of the openloop gain. This makes it possible to build amplifiers with precise closed-loop gains. Among other beneficial effects, negative feedback imparts high linearity and stability to the amplifier.

Rules Of Operation. A few basic facts of op amp operation will be stated as rules and applied to the operation of several circuits. The implications and meanings of these rules will become clear as you apply them to the circuits.

1. The *difference* in voltage between the + and — input terminals is always small and can be assumed to be zero. (This fact is a direct result of very high AVOL.)

2. The current entering the + and — input terminals is small and can be assumed to be zero. (This rule is a direct result of a very high *Zi*.)

3. If the + input terminal is at ground voltage or zero, the — input terminal can be assumed to be virtually at ground voltage or zero. (This rule is also a direct result of high AVOL.)

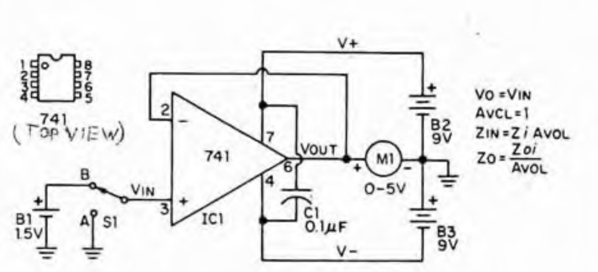


Fig. 2. Here an op amp is hooked up to provide unity gain. The output voltage equals the input voltage due to the feedback of the non-inverted output of the op amp to the inverting input effectively reducing the gain of the stage to 1.

4. When an op amp is connected in a negative feedback configuration, a voltage change at the + input must result in an equal voltage change at the — input terminal. (This is a description of rule 1 in operation.)

Setting Up. Breadboard the op amp circuit shown in Fig. 2. You can use perforated board and flea clips to assemble the circuit. Better still, use a solderless breadboard kit. Use only the 741 op amp for IC1. These are short-circuit proof and are internally frequency compensated to prevent oscillations. Switch S1 may be simulated by a clip lead.

Install disc capacitor C1 as close as possible to the IC. Use either a common tie point for all ground connections or a heavy ground bus. Keep the input lead wires well separated from the output lead wires. The prototype breadboard uses a 50 uA DC meter connected in series with a 100,000-ohm, 1% resistor for meter M1. Alternately, you can use your VOM (1000 ohms/volt or better) to measure output voltage. Use two fresh nine-volt transistor batteries for B2 and B3 and 1½ volts (an AA cell) for B1.

• **Unity Gain Follower.** Simplest of the op amp circuits, the unity-gain follower shown in Fig. 2 has a direct connection from output to the inverting input. This provides one-hundred percent negative feedback. With switch S1 at position A, the + input is grounded and meter M1 indicates zero. By rule 3, the — input is at virtual ground. Therefore, VOUT is also at virtual ground. With S1 set to position B, the meter now indicates the voltage of battery B1, near 1.5 volts. By rule 4, the 1.5 volt increase at the + input must be accompanied by a 1.5 volt increase at the — input. Hence, VOUT must rise to 1.5 volts.

The resulting closed-loop gain, or AVCL, is unity (one unit out for one unit in). However, the high AVOL inside the IC itself is still present, enforcing close compliance with the several rules. As noted on Fig. 2, actual input and output resistances ZIN and Zo are much improved due to feedback. The resultant input resistance now equals *Zi* times AVOL or 400,000 megohms for the 741 op amp! The resultant output resistance now equals *Zoi* divided by AVOL, or .0035 ohms! Consequently, the unity gain

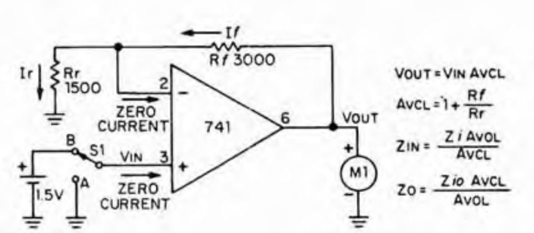


Fig. 3. With the feedback signal reduced, the gain of this op amp non-inverting voltage amplifier is 3. The meter should read 4.5 volts with 1.5-volt input.

follower can duplicate the input voltage at its output without loading down the input voltage source due to the high input resistance and with high accuracy due to the low output resistance. Actually, input and output resistances are degraded somewhat by secondary factors. Nevertheless, this unity gain follower offers the highest input resistance and lowest output resistance of the several basic circuits.

• **Non-Inverting Voltage Amplifier.** Stable op amp voltage amplification is obtained by feeding back only a portion of the output voltage. Alter your breadboard circuit to include feedback voltage divider resistors R*f* *and* Rr, as shown in Fig. 3. With Rf equal to 2Rr, only one-third of the output voltage is fed back to the inverting input.

With S1 at position A, the + input is at ground voltage and the meter indicates zero. By rule 3, the — input is at virtual ground or zero. With zero voltage across Rr, current Ir is zero. In view of rule 2, If always equals Ir and is zero in this case. With zero current in Rf, the — input and the output voltage must be equal and zero in this instance.

Compared with the unity gain non-inverting amplifier, actual output resistance Zo is greater by a factor of (1 + AVCL) or three times as much, still acceptable small at this (and even much higher) gain.

By connecting additional input resistors to the — input and upon applying several input voltages, the amplifier will sum the several input voltages at the output. For this reason, the amplifier is often termed a summing amplifier and the — input is termed the summing node or input.

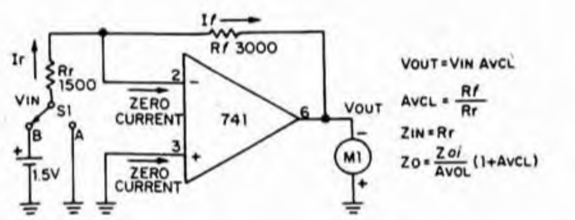


Fig. 4. The circuit arrangement of the op amp provides an inverting output with a gain of 3. Compare it to fig. 3.

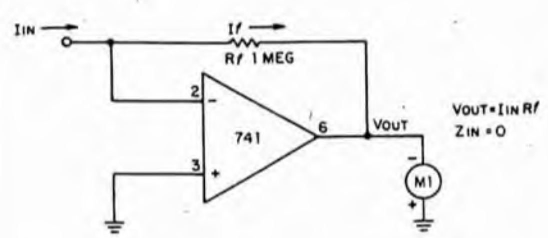


Fig. 5. Using only one external circuit element, this op amp functions as a current-to-voltage converter.

With S1 at position B, the + input is raised 1.5 volts and the meter indicates 4.5 volts. By rule 4, the — input must rise to 1.5 volts matching that at the + input. The op amp does this by forcing a current into the feedback voltage divider as shown. With 1.5 volts across Rr, current Ir equals 1.5 volts divided by 1500 ohms, or 1 mA. Also, the voltage across Rf equals 1 mA times 3000 ohms or 3 volts. Thus, VOUT equals 1.5 plus 3 or 4.5 volts.

Closed-loop voltage gain AVCL equals 1 + (Rf/Rr) or 3 in this case. Compared with the unity gain circuit, actual output resistance Zo is three times grater and input resistance ZIN is one third that of the unity gain circuit. This reflects the effect of feeding back 1/3 of the output voltage. To obtain a closed-loop gain of ten, resistor Rf must equal 9Rr, and so forth.

• **Inverting Voltage Amplifier.** Alter the breadboard circuit to that of Fig. 4 including reversal of the meter. With S1 at position A, the meter indicates zero. The proof of this result is identical to that of the non-inverting voltage amplifier with switch at position A. With S1 at position B, the meter indicates 3 volts (actually, minus 3 volts since the meter is now reversed).

In this case, the — input does not rise to 1.5 volts. By rule 3, with + input grounded, the — input must remain at virtual ground. Therefore, and quite importantly, the voltage across Rr equals the input voltage, or 1.5 volts divided by 1500 ohms, or 1 mA, flowing in the direction shown. Since If equals 1 mA times 3000 ohms, or 3 volts. With the — input at virtual ground, the otuput voltage must be minus 3 volts as indicated on the reversed meter.

The closed loop voltage again, AVCL, is simply Rf/Rr, or 2 in this case. Quite unlike the previous cases, actual input resistance ZIN equals Rr, the input resistor.

• *Current to Voltage Converter.* A variation of the inverting amplifier, the current to voltage converter shown in Fig. 5, omits input resistor Rr. Because the — input must remain at virtual ground for linear operation, this circuit cannot accept an input voltage. Instead, it accepts an input *current* and is used to measure very small currents. If IIN were 1 microA, the output voltage would be 1 microA times 1 megohm, or 1 volt. By making Rf very large, the circuit can measure extremely small currents. The input resistance of this circuit is zero. The output voltage, VOUT, equals IIN times Rf.

If you breadboard this circuit, you may observe a small output voltage at zero input current. This output "offset" voltage is caused by the flow of a small bias current from output to input through the large feedback resistor, Rf. Unless special op amps having very low bias currents are used, it is necessary to include a nulling circuit to reset.

• **Input Bias and Offset.** Although rule 2 assumed zero input currents, an op amp does require a small input current Ib to bias the input stage into linear operation. For the 741, Ib may range up to .5 uA. The difference between the input bias currents at the two inputs is the input offest bias current ziio. This current is usually much smaller than Ib. Both Ib and Iio cause an objectionable output offset voltage with Rf is very large. To restore the output voltage to zero, add the nulling circuit potentiometer R1 and resistor R2 as shown in Fig. 6. With S1 open, adjust the control until meter indicates zero.

If Rr is small, and upon closing S1, you may observe that the meter again loses its zero. This is caused by the input offset voltage Vio resulting from slight mismatches of the input transistors. Input offset voltage Vio is defined as that input voltage required to restore VOUT to zero. It is measured under open-loop conditions with very low value resistors at the input. For the 741, *Vio* may range up to 6 millivolts. Conveniently, the 741 includes terminals allowing compensation for input offset voltage. Add potentiometer R3 and adjust the control with S1 closed, until the meter indicates zero. If both circuits are included, adjust the controls several times in succession.

Conclusion. Having become acquainted with the basic operation of the op amp, and with some knowledge of the six primary op amp specifications, you will now be able to experiment with op amp circuits with some degree of confidence rather than apprehension. With some appreciation of how and why the circuit functions as it does, how the performance of the several circuits compare with each other, and how negative feedback plays its part, you will find that op amp literature and circuits are more easily understood. ■

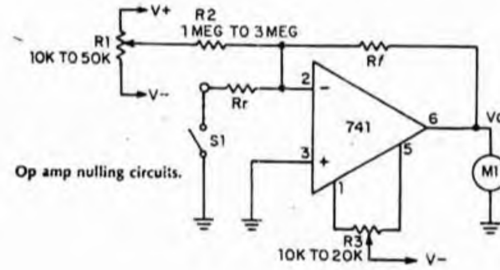


Fig. 6 When required an offset voltage can be added to the op amp circuit.

Inside Synthesizers

There is no doubt that electomic synthesizers have made a major impact on the world of music, and are here to stay. No longer are they limited to experimental avant-garde works. There are dozens of albums of traditional music (both classical and popular) played on synthesizers, and they are widely used in radio and television commercials, and popular music of all kinds; rock, jazz, and even western.

A good synthesizer can take the place of literally dozens of other instruments, and can easily produce sounds that would be difficult, it not impossible, to duplicate on acoustic instruments. Synthesizers are fascinating to experiment and practice with.

Many people, however, are reluctant to try them, fearing they are too complex for the beginner.

Actually, synthesizers can be quitre easy to understand, if you just take them one step at a time. Synthesizers are really nothing more than packages of modules, that can be connected in various ways to produce different sounds. Once you understand the basics it's not hard to design an experimental synthesizer of your own.

Basically, there are four types of modules in snythesizers; oscillators, which produce the basic signal; filters, which modify the harmonic content, or tone of the signal; amplifiers, which modify the dynamics or volume of the signal; and, finally, some method of controlling all the modules.

You could use potentiometers to control all these devices manufally (in fact, that is what the early electronic music composers did). That method is very awkward and tedious. Most modern snythesizers use some form of voltage control in their electronic circuit.

The voltage to a voltage-controlled oscillator (VCO) would control the frequency of the signal, or the pitch. The voltage to a voltage-controlled filter (VCF) will determine how much of the original signal will be attenuated. The voltage to a voltage-controlled amplifier (VCA) determines the strength, or level of the signal.

Keyboards. There are a number of voltage sources that are used in snythesizers. The most common is the keyboard, which is similar to those found on pianos and organs. Figure 2 shows a simplified voltage control keyboard. Each switch is a key. If S1 is closed, none of the resistors are in the circuit, so the full five volts is available at the output. If S2 is closed, R1 is brought into the circuit, so the output voltage is dropped to a lower value For this example, let's assume all four resistors are rual, and each has a one volt drop across it. This means that closing S2 produces a four volt output. S3 would bring both R1 and R2 into the circuit (in series, so their values add), so the output would now be three volts. S4 would bring 3 resistors into the circuit, giving a two volt output, and S5 is connected to all four resistors, giving an output of only one volt.

One important thing to notice with this kind of arrangement, is that only one voltage can be produced at one time. That is, the keyboad is monophonic. Polyphonic (multiple note) keyboards are used in comercially available synthesizers, but they are outside the scope of this article.

You can build a monophonic keyboard like Figure 1 from a toy organ keyboard. The switches marked B are for a trigger signal (explained later) which simply tells the synthesizer when a key has been depressed. There are two switches for each key.

The resistor values would be determined experimentally. The exact voltage doesn't matter so much as the pitch of the VCO, so you'd hook the keyboard to the VCO and try various resistors until it sounds right. You can use fixed resistors connected with spring clips, or potentiometers, which are ore expensive, but give you precise control of the sound.

Sequencers. Another common voltage source is the

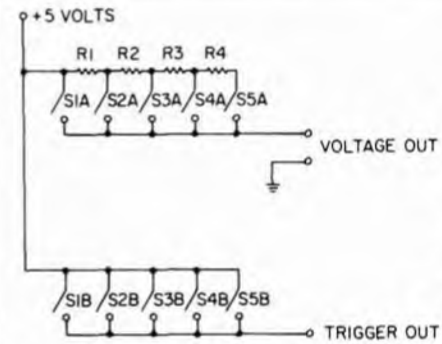


Fig 1. In this version of a simplified voltage keyboard, each switch is a key.

Complex in appearance but simple to understand, electronic synthesizers are not an avant-guard musical instrument. Today they are a permanent part of music and will no doubt remain there for quite a long time.

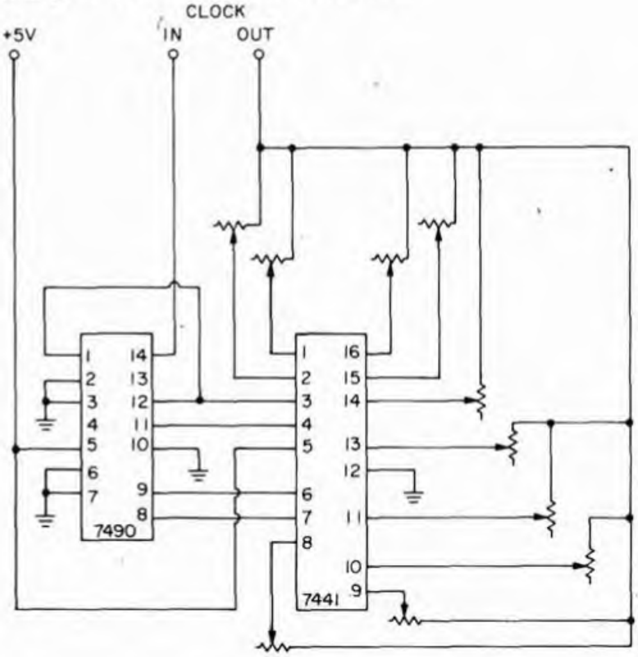


Fig. 2. This simplified schematic of a sequencer utilizes two integrated circuits to order various voltages that are used in producing music. A clock controls the sequence.

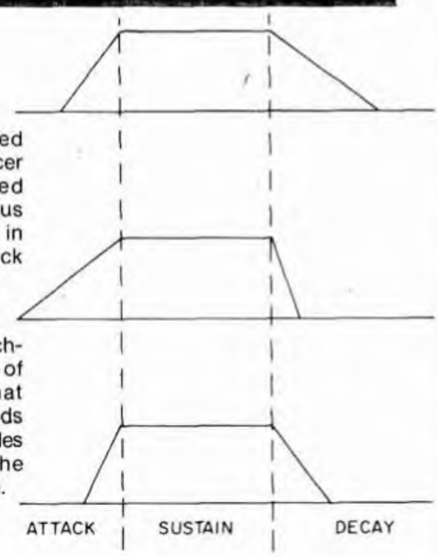


Fig. 3. These three sketches illustrate the range of envelope voltages that can be made. The speeds of attack and decay modes are controlled thru the envelope voltage mode.

sequencer. As the name suggests, this device produces a sequence of voltages. For examp,e a sequencer's output might be set to 5V, 4V, 3V, 3.5V, 3V, 2.5V, 3V, 4V. When the sequence reaches its last position, it loops back around and starts over.

A schematic for a simple sequencer is shown in Fig. 2. Note that the sequencer requires as oscillator to control it. This is not the VOC that produces the signal you hear. It is a separate oscillator, or clock, that controls the speed of the sequence, consisting of a low frequency squre-wave.

As a matter of fact, a spare oscillator can also be used as a varying voltage source. If the controlling oscillator is a very low frequency you can hear its wave-shape (discussed below), but it is in the audible range, very complex tones can be produced.

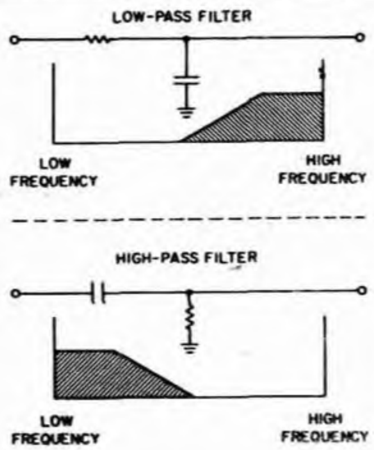
Function Generators. Another varying voltage source is the function generator or envelope generator. When a function generator receives a trigger signal (such as from a keyboard's second set of switch contacts) its output voltage rises from zero to some specific maximum level (attack). The maximum level is usually held for as long as the trigger voltage is present (sustain). When the key is released, the voltage drops back down to zero (decay). Figure 3 shows some typical evenlopes. Figure 4 is the schematic of a simple function generator. Any of these voltages can control a VCO, a VCF, or a VCA, in any bomcination that you wish to try.

Synthesizing a Flute. For discussing the rest of the synthesizer, let's assume we want to synthesize a flute sound.

We start out with a VCO to produce the original signal. VCO's may produce any of a number of wave-shapes. The most common are shown in Figure 5. A sawtooth-wave has all of the harmonics. For example, if the frequency of the sawtooth is 100 Hz., then it will also contain tones at 200 Hz. (the second harmonic + 2 x 100), 300 Hz. (3rd harmonic + 3 x 100), 400 Haz. fourth harmonic), 500 Hz. (fifth harmonic), and so forth. The ear automatically combines all of these tones into one raspy sound with an apparant pitch of 100 Hz.

A square-wave, on the other hand, only has odd harmonics. Again, assuming a fundamental frequency of 100 Hz., there will also be tones at 300 Hz. (third harmonic), 500 Hz. (fifth harmonic), 700 Hz. (seventh harmonic), and so on. The resulting sound is rather reedy, vaguely similar to an oboe.

A triangle-wave also contains all the odd harmonic,



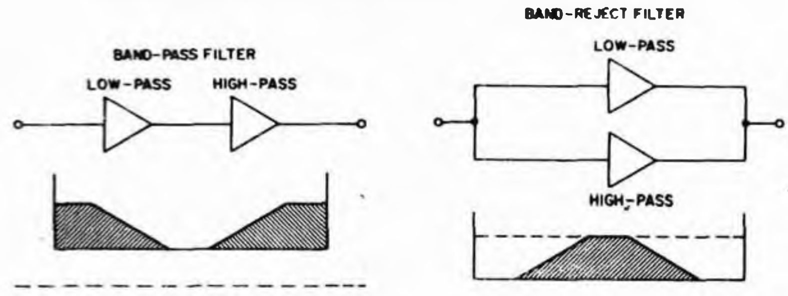


Fig. 8. Combining two filters in parallel or series creates a simple frequency selector.

Fig. 7. The arrangement of the capacitor and the resistor determines cut-off points.

but they are somewhat weaker than in a square-wave, so the sound is brighter.

A sine-wave has no harmonics, and the sound is very piercing. Actually, it is somewhat unpleasant to listen to by itself, so this waveform is sually only used as a fluctuating voltage source fo control other modules for such effects as tremolo and vibrato.

Filtering. Since the sound of a flute is somewhere in between that of a triangle-wave and a sine-wave, we start out with a triangle-wave, and filter out some of the upper harmonics. (A squre-wave could also be used, but more filtering would be required). For this purpose we would use a low-pass filter, which lets low frequencies pass through it, but blocks higher frequencies (i.e., harmonics). For most synthesizer uses low-pass filter is the handiest, but its mirror image; a high-pass filter can also be useful. If you combine these two basic filters as in Figure 8, you have a band-pass or a band-reject filter, so there are a number of effects that can be achieved.

The filters in Figure 7 are fixed filters. That is, their cut-off frequencies are determined by the components values, and remain constant unless the components are changed. This can be troublesome in a snythesizer. Let's say you want a sound with the fundamental, the third and the fifth harmonics only. If we start out with a 1,000 Hz. signal, the third haronic is 3,000 Hz., and the fifth is 500 Hz. So we use a low pass filter to lobck everything about about 5,500 Hz. Now, suppose we double the frequency and play a 2,000 Hz. note. The third harmonic would be unaffected at 4,000 Hz., but the fifth harmonic would be 6,000 Hz., which means it would be blocked by the filter. Or, if we reduce the frequency to 500 Hz., the filter will pass the fundamental (500 Hz.), the third harmonic (1,500 Hz.), the fifth (2,500 Hz.), and also the seventh (3,500 Hz.), and the ninth (4,500 Hz.). Obviously the sound can change quite a bit as the pitch varies.

The solution is to use a VCF, and control it along with the VCO. The patch is shown in Figure 9.

We now have a flute-like sound, but unfortunately, it won't sound very realistic, because each note will instantly be at its maximum level as soon as the key is depressed, and will instantly cut off when the key is released. Real-world sounds aren't like that, and the effect is terribly unnatural. Here is where the function generator and VCA come in. The new patch is shown in Figure 12.

Since the flute is a wind instrument, the sound will take some time to built up to maximum, as the player blows air through the tube. We can simulate this with a fairly slow attack. The sound will also take some time to die out when the player stops blowing, but not as long as it took to build. We'd use a moderate decay time. The evenlope voltage applied to the VCA would look like Figure 13. When no key is depressed, the VCA is cut off. No sound is heard. When a key is depressed, the function generator is triggered. A changing voltage is applied to the VCA. The volume of the signal will vary in step with the voltage.

The noise generating circuit shown in Fig. 14 is a simple but important addition to the synthesizer. It makes for a great deal of flexibility and provides background noise for effect.

At this point we have a fiar impression of a flute. Just how good it will sound will depend on the quality of the circuit. Those shown in this article are suitable only for experimentation.

We can do a little more to improve realism. For instance, we could add a little noise, as in Figure 15, to simulate the sound of the musician's breath. This noise should have a somewhat faster attack and decay, than the VCO signal, at a much lower level.

We could also add a tremolo (or fluctuating effect) by controlling the VCA with a very low frequency sine wave oscillator (about 5 or 6 Hz.). The bias is a manually adjusted negative voltage that is equal to the peak voltage of the sine wave. This is to cancel out the sine wave when the function generator is not triggered. Thus, the VCA will be cut off between notes.

Synthesizing a Drum. Now that we've synthesized a flute, lets' go back to the patch in Figure 15. Would you believe this is the same patch you would use to produce a drum sound? In this case the level of the noise should be much higher than that of the VCO, and the envelopes would be changed to those of Figure 18A.

Figures 18B and 18C show additional envelope settings for unique sounds, not heard in nature. Even with this simple patch, many variations are possible. Besides changing the envelope, you could change the waveshape of the VCO, or substitute a second FCO for the noise genrator, or use high-pass or band-pass filters instead of the low-pass module. You could also use a function generator (or oscillator) to control the filters, or

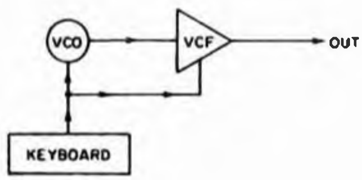


Fig. 9. Using a VCF with a VCO, the proper harmonics of an instrument can be created.

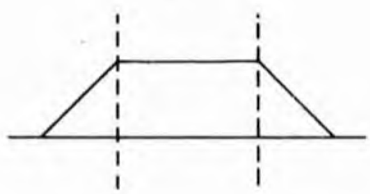


Fig. 10. A low-pass VCF is used to control frequency. Enclose photocell and lamp together.

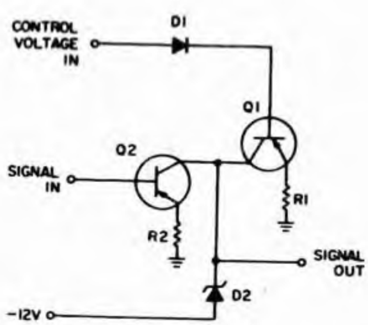


Fig. 11. A simple device, a VCA controls the volume generated so snythesized sound seems real.

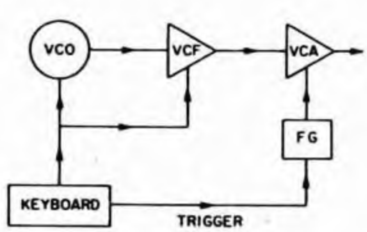


Fig 12. This block diagram illustrates how the VCA control fits into the synthesizer.

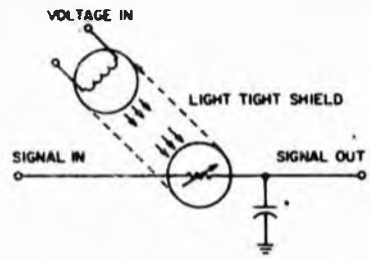


Fig. 13. The VCA has an envelope voltage, which is shaped like this, applied to it.

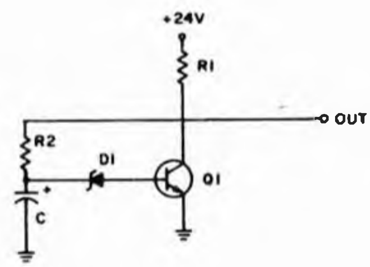


Fig. 14. The noise generator of the synthesizer provides background noise for effect.

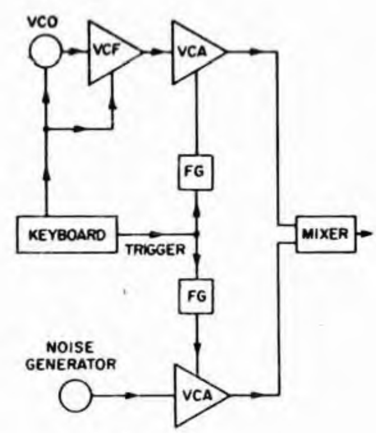


Fig. 15. Adding a little noise to the circuit adds strange qualities to the sound.

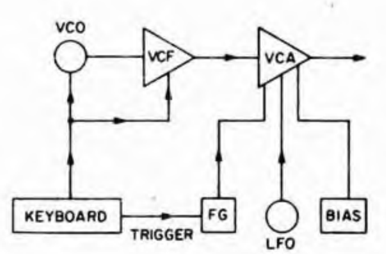


Fig. 16. A variation of the basic patch, which is given in Fig. 15, for the tremolo effect.

use you imagination.

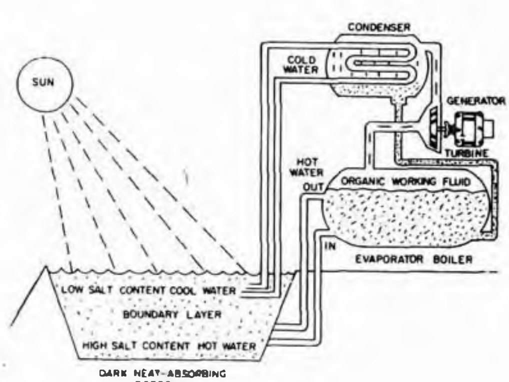
As you can see, synthesizers really aren't as complex as they seem, yet you can get hours of pleasure out of the simplest collection of modules. Each variation in the way the modules are wired totether, will give you a different audio effect. Think of yourself as the conductor who orchestrates the sounds of many musicians. With a synthesizer, you have many more possibilities than with acoustic instruments. The latest trend in music is called New Wave. Its sounds are generated strictly from synthesizers, without any acoustic instruments being used. Eerie sounds seem to echo through distant galaxies, producing strange & hypnotic effects.

Watts From Sunlight

By the early 1900's, Israel expects to meet much of its electricity needs with solor generated power. An all solar powered 150 kilowatt generating plant was put into service in late 1979, and more ambitious projects are slated for the future.

The Ein Bokek project, located on the shores of the Dead Sea, makes use of the concept of the solor pond—a body of water whose salt content is such that the water in its depths rises to high temperatures—and a turbogenerator powered by this heat energy. The combination of these two relatively simple and low cost technologies has made possible an innovative approach to electricity production. This large scale application of solar technology is the first of its type.

Solar Priorities. The Israeli government was understandably interested in giving a high priority to the development of solar energy. Continuing hostility from the oil-producing Arab countries dictated energy conservation long before it became necessary for the rest of the world. Israel was one of the first countries to take advantage of the abundant and free energy from the sun; nearly ever rooftop sports a solar water heater, with its distinctive panels and collecting tanks.



This diagram explains the solar pond generating system and shows the transformation of sunlight into electrical energy by taking the heat generated and powering a turbine.

A 15,000 square foot solar pond was constructed at Ein Boket by excavating an area of the Dead Sea shore, damming the seaward side, and lining the excavation with a black, heat-absorbing rubber substance. Water was let in, and its salinity level was constantly monitored to ensure that the proper gradient level would be reached. The high salt and mineral content prvents normal convective cooling, and the salinity increases with depth. The water at the bottom of the pond rose nearly to boiling temperature after a few hours under the desert sun.

The hot water was then circulated through a heat exchanger, where it heated an inorganic working fluid to the gaseous state. This gas powers a turbine attached to an electric generator.

The turbogenerator system that converts the heat of the solar pond into electricity is a unit known as the Ormat Energy Converter (OEC). This is a low-temperature, low-pressure generating system that was originally designed to produce electricity from waste heat. It uses a closed system wherein the working fluid is heated up, used to run a turbine, then condensed and used over again. The OEC is ideal for use in applications where a constant 175-200°F heat flow is available.

Encouraged by the success of the pilot plant at Ein Bokek, the Israeli government has the project's two contractors, Ormat Turbines and Solmat Systems, at work on a 5,000 kilowatt power station. This generating station should be operational by 1981, and will be the first module of a system with an eventual capacity of 2,000 *megawatts*. Meanwhile, the Rin Bokek plant has been constantly producing 150 kilowatts, day and night, winter and summer. The Dead Sea is a body of water approximately 50 miles long and 11 miles wide: an enormous potential for electricity generation!

A Practical Energy Source. Solar ponds occur naturally in many parts of the world, including the U.S., and they can be man-made as well. While the Ein Bokek project was built in an area where temperatures in excess of 100°F are common, the solar pond is equally viable in more temperate climates.

This ambitious Israeli project is just one example of how the dependence on expensive and potentially hazardous energy sources can be offset by solutions developed from Nature. ■

A Peek At Computer Circuits

It seems that the day of the home computer has at last arrived. Everywhere, we read about these new machines that are changing our lives, running things better than we ever could ourselves. Everybody tells us that the solid state computer is the newest and most radically advanced thing to come along since the sun. But, you know, in spite of all these claims, the personal computer is not all that new. In fact, you've been using one all your life; it's called a brain.

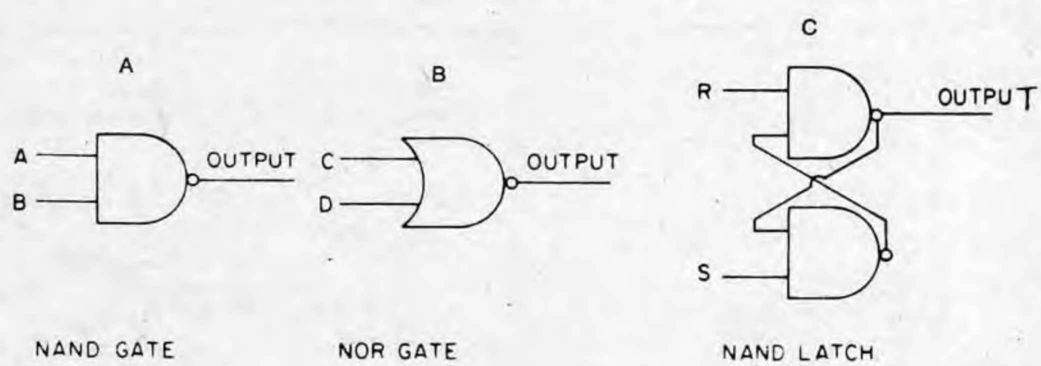
Many people, electronics hobbyists in particular, have long suspected some fundamental similarities between brains and electronic computers. In this article we're going to compare these two types of information-processing systems.

Let's begin on familiar ground. In Figure 1 we have the various logic gates; the important building blocks of any electronic computer, big or small. The NAND gate's output state is a function of its two inputs, and it remains at a logical 1 so long as the input state is not 11. An input of 11 sends the output to a logical 0. On the other hand, the NOR gate's output is usually a logical 0 except when the input is 00, in which case the output assumes a logic 1 value. Using combinations of just these two gates we can synthesize any desired logic function, no matter how complex.

As you know, the 0 and 1 symbols can denote voltages, with 1 usually representing a high voltage and 0 a low voltage. The ones and zeros may also be interpreted in a philosophical sense as True and False, respectively. Why should we use numbers to represent qualities such as truth and falsity? It's mainly a convenience which allows logic to be expressed inmathematical form. If logic can be handled algebraically, then solution of logical problems becomes routine, and can be handled by machines—computers.

But a computer needs more than decision-making apparatus (NAND and NOR gates); it needs memory. We store the computer's instructions (program) in memory, and also use memory to retain the solutions generated by the computing process. One basic form of memory is the NAND latch (Figure 1c). When a little control gating is added to this basic circuit, we can get various flip-flops, which are probably the most familiar storage elements. Other forms of stroage include magnetic cores, magnetic tape, paper tape and the exotic magnetic bubble device. Regardless of the storage form, however, memory simply holds data until it's called up for processing or readout.

Neurons. Contrast the human information-processing system to the electronic computer. Like a computer, the human brain is a complex entity built up from enormous numbers of simple fundamental units. These basic units are called nerve cells or neurons. Several classes of neurons exist; nevertheless, all share the same generic traits. Figure 2 shows one typical neuron. The tentacle-like elements are called dendrites, and they function as the input leads to the cell. The single long filament leaving the cell is an axon, which is the cell's output.



A	B	OUTPUT
0	0	1
1	0	1
0	1	1
1	1	0

C	D	OUTPUT
0	0	1
1	0	0
0	1	0
1	1	0

R	S	OUTPUT
0	0	NOT ALLOWED
1	0	0
0	1	1
1	1	NO CHANGE

Fig. 1. Some Basic logic circuits: a) A NAND gate gives a 1 output for any input except 11. b) the NOR gate only gives a 1 output when both inputs are 0. The NAND latch (c) is a memory device which is the basis of the common flip-flop circuit.

Nerve signals travel along the thin membrane that encloses the cell's protoplasm. Adjacent cells communicate with one another across synapses—gaps between axons and dendrites of different cells.

Logically the neuron is more complex than a NAND or NOR gate. It functions as a majority gate, which may be described as follows: Suppose, for example, that we construct a gate with five inputs and one output, and that the output will be a logical 1 whenever a majority (3, 4, or 5) of the inputs is in the 1 state. If there are less than a majority (0,1, or 2 that is) of logical 1 inputs, the output is a logical 0. Such majority gates are not encountered too frequently in electronics, one notable exception, however, being Motorola's CMOS MC14530 dual 5-input majority gate. As a rule a neuron has more than five inputs—typically, up into the hundreds. Regardless of its complexity, however, a majority gate can be expressed in equivalent form as a combination of NAND and NOR gates.

From the foregoing you can see that computers can be logically equated to neuron networks; nevertheless, certain physical differences exist between biological and electronic systems. To begin with, electronic signals are viewed as differences in potential, which govern the flow of free electrons. Signals in nerve nets, however, consist of waves of polarity reversal on the surface of polarized neuronal membranes. In Figure 3 you can see that a nerve's membrane normally possesses excess positive charge on its outer surface, and excess negative charge on the inner surface. Excitation of the nerve brings about a polarity reversal at some point on the membrane, which then induces polarity reversal at adjacent points on the membrane, and the signal spreads. The reversed-polarity condition at any given point reverts to normal after several miliseconds, while the wave of polarity reversal propagates to new points on the membrane. Charge distribution on the membrane's surface is controlled by movements of sodium and potassium ions, and the rate of ion flow through the membrane limits the speed at which nerve signals travel to about 100 meters/second. Contrast this with the speed of an electronic signal traveling down a transmission line with a polyethylene dielectric: about 2×10^8 meters/second. Biological systems are thus inherently slower than electronic ones.

Memory. Two physical characteristics distinguish the human nervous system from the electronic computer. Communication between the functional units in a computer is established by connecting wires. Interneural communication involves diffusion of certain chemicals across synapses. Memory is likewise a chemical phenomenon, involving chemical changes in nerve cells and associated structures. Regardless of these physical differences, however, the two systems possess a fundamental logical similarity, as we've seen. Is it now possible to map the information flow in the human system in the same way that we are accustomed to in electronics?

Yes, to a certain extent. First of all, no one alive and in good health would consent to have men in white lab coats probing about inside his or her head. So direct information about the living, functioning nervous system is hard to come by. But a lot of data can be obtained concerning brain structure is an indirect way, and this is the domain of cognitive psychology. Research along these lines can be frustrating; it's very much like trying to figure out what's inside a computer by seeing how the machine responds to different sequences of input signals. Nevertheless, psychological data, together with whatever biologists can provide, has allowed many deductions about the brain's organizations.

Let's look at Figure 4, which is a simplified human

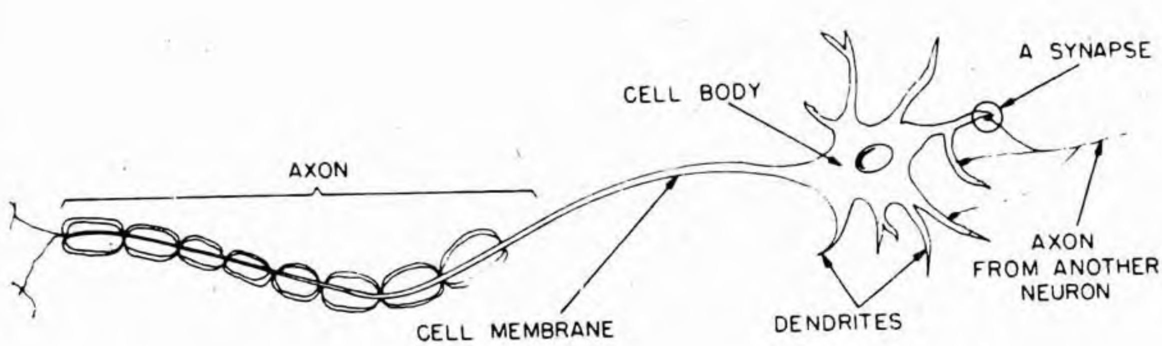


Fig. 2. A motor neuron, one of the basic building blocks of the biological information processing system, is a single cell and functions as a multi-input majority gate. The inputs to the dendrites and the cell body come from other neurons. These inputs, often numbering in the hundreds, are averaged and the majority input determines the output along the axon—sort of biological democracy. The axon then leads to another neuron.

cognitive map. Memory elements are represented by squares, while circles stand for control processes. This is an arbitrary distinction because both functions, memory and control, are apparently performed by cells. Keeping that fact in mind, let's analyze the major features of Figure 4, saving the details for later. First, visual and auditory information from theo utside world are converted by the eyes and ears respectively into nervous impulses. These signals are routed to buffer memories, the sensory registers, which act to preserve fleeting data long enough for the rest of the system to process it. By convention the visual and auditory sensory registers are called, respectively, the "icon" and the "echo."

The block labeled LTM is long-term memory, a presumably permanent storehouse filled with all the important facts accumulated during a lifetime. For example, LTM contains enormous quantities of information such as 1 + 1 = 2, V = IR, the rules of English grammar, telephone numbers, and so on. In contrast, STM or short-term memory has a very limited capacity. Furthermore, information fades away or gets dislodged very easily from STM. It is possible to consciously retain STM information, however, by the process of rehearsal, repeating things mentally to one's self. As the arrows of Figure 4 indicate, rehearsal takes information out of STM and then re-enters it. At the same time, rehearsal can facilitate entry of information into LTM, as we'll see later.

Recognition. The final circle represents two independent processes, chunking and pattern recognition. Because these processes seem to have the same inputs and output, they coincide on the map. Chunking is an organizational process wherby sensory register information gets grouped by means of LTM rules before entering STM. This chunking does two important things: 1) it allows more information to be crammed into STM, and 2) it makes the information easier for LTM to assimilate. Note that information cannot enter LTM directly; instead, it must be worked on in STM first. For this reason, STM is commonly called working memory by psychologists. Apparently, STM is the site of what we experience as consciousness. The last process, pattern recognition, is the identification of visual or auditory inputs, and is a very important part of conscious experience.

As we stated earlier, the map in Figure 4 is far from complete. Missing entirely are the autonomic processes, such as breathing, which the brain controls. Gone too are the outputs which, for example, would control muscular action. It would also appear from Figure 4 that all inputs are processed equally, yet we know this is not true. As you read these lines countless sounds, sights, and pressure sensations are available to your senses, yet your attention is fixed on the letters that you scan. This phenomenon of selective attention is extremely important, but a little too complex to cover here.

Keeping in mind the fact that our cognitive map is greatly simplified, let's consider its features in more detail, starting with the sensory registers. Experiments with the icon have revealed that visual information is stored for less than a second before fading. Moreover, new visual stimuli erase the old iconic information. If this did not happen, our apprehension of an image would lag about a second behind its appearance. By contrast, echoic storage lasts longer than that of the icon (the exact duration is subject to dispute, however), and erasure per se doesn't seem to occur. The necessity of echoic storage is readily seen when you consider that lanugage is a serial phenomenon, made up of sounds which follow each other in time. To make sense out of language, what is and what was must both be available. For example, a single word like *computer* is made up of many basic sounds called phonemes. All must be available in order for the pattern recognition process to identify the word.

STM. Let's consider STM. One piece of evidence suggesting an STM that is distinct from LTM is Milner's Syndrome, a mental impairment caused by damage to the hippocampal area of the brain. Victims of this syndrome are quite normal in most respects, but are unable to permanently memorize new information. For example, given a list of words, they can retain the list in memory for any desired length of time by continually rehearsing. When instructed to stop rehearsing, they lose the list in a matter of a minute or so. One patient was able to read the same magazine again and again without getting bored. The patient's doctor had to introduce himself daily, even though he had been treating the man for years. Yet, the patient could recall the events in his life before the accident. According to our cognitive map, we

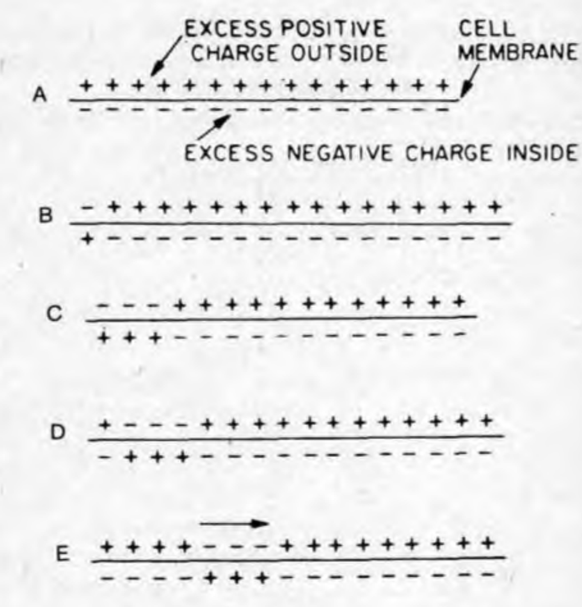


Fig. 3. Nerve signals are waves of temporary polarity reversal along the cell membrane. In the first drawing above (A) the nerve membrane is in a quiescent state. in (B) the polarity reversal begins and progresses along the membrane (C & D). The polarity returns to normal (E) and this return travels behind the reversal at an equal speed (about 100 mps).

can theorize that the damage broke the link between LTM and STM. The victim could use STM and recall old items from LTM, but he was unable to put new information into long-term storage.

Not only is STM storage temporary but it also possesses only limited capacity—typically five to nine items. This figure varies according to the nature of the items being memorized, but the average person's short-term capacity is about seven letters, numbers, or words. This seven-item limit is commonly referred to as the memory span. Let's clarify the memory span concept now. If a list of items is presented one at a time to a subject, that person can recall the list in perfect order, after seeing the items just once, if the list length is less than or equal to about seven items. Mistakes will appear with longer lists. Of course, we can memorize more than seven items, but to do so will require several presentations of the list plus some rehearsing. Since rehearsal both refreshes STM and transfers information to LTM, learning of lists longer than the memory span evidently requires storage space in LTM.

Suppose that the following list of letters is presented to you. Could you memorize it at one pass?

VTVMSCRFETUJT

You could if you treated the items like this:

VTVM SCR FET UJT

This type of organization is called chunking, which means ordering the list into meaningful groups. Since meanings are governed by the pattern of storage in LTM, previously acquired LTM information is being used here to facilitate STM storage. The chunked list has four items, well within the memory span, and can be quickly memorized by anyone familiar with electronics. But to an uninitiated person the list contains thirteen items; a few rehearsals would be necessary to enable perfect recall.

Let's try another experiment. Can you memorize the following list quickly and without a lot of rehearsing?

110001100101000011

If you can count in binary, it's a cinch. Look at the list in this way:

110	001	100	101	000	011
6	1	4	5	0	3

Below each triplet is its decimal equivalent. If you can mentally do such a conversion, you're left with only six decimal digits to memorize. When the time comes for recall, simply change the decimal numbers you've memorized back to binary. With some practice, you should be able to handle lists of about twenty-one ones and zeros, much to the amazement of your friends. Note that you're using old LTM information—the rules of binary-to-decimal and decimal-to-binary conversion—to assist in memorizing new material.

RAM. Leave STM now and consider long-term storage. First of all, think of all the different kinds of information contained in LTM. Sights, sounds, tastes, smells, and words are all stored in some codified form. The information can be reached quickly without first searching through a lot of extraneous material. In electronics we call such a memory structure random-access. This random-access characteristic is one of LTM's most important features, and current theories hold that a close relationship exists between the structure of LTM and the grammar of language.

Certainly one of the most important characteristics of LTM is its relative permanence of storage. However, we all know that recall is often marred by forgetting. Is this a memory failure? It might be, but some say no. Consider the experiences of the neurosurgeon Wilder Penfield. While performing surgery he electrically stimulated areas of patients' brains. Since the patients were conscious during brain surgery (there are no pain receptors in the brain), they were able to report their sensations. Many reported long-since forgotten, seemingly trivial incidents from the past. These sensations were vivid, filled with minute visual and auditory details. In fact, they were often described as being like movies. This type of evidence is frequently

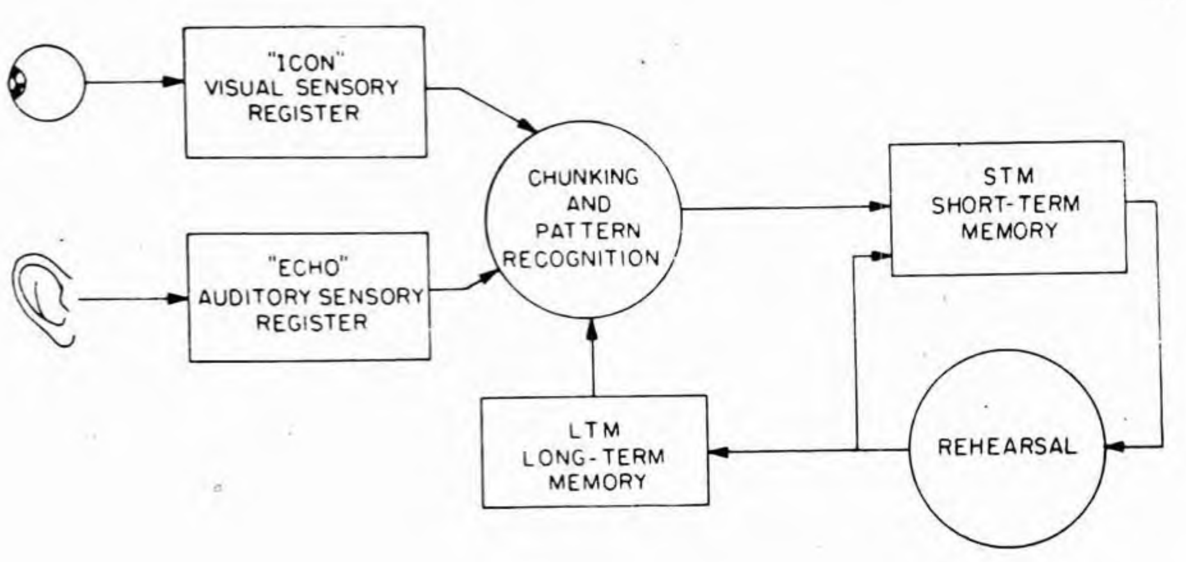


Fig. 4. A simplified model of the information processing system of humans and higher animals. In addition to the "ECHO" and "ICON" sensory registers there are others for tasts, smell and touch. In the lower box, the LTM, is stored everything you know except for the most recent inputs which are being processed as you read this page. These immediate inputs are going into the short term memory and rehersal loop; some of what you have read will be recorded in the LTM—there is, however, no guarantee your brain will be able to recall it. Something may be in memory but the brain can't always find it.

cited in support of the notion that everything stored remains stored, and that forgetting is the inability to find stored information. However, we don't know that the patients' sensations were true memories. They might contain some factual information, but the great bulk of detail could be manufactured, just as in a dream. The question of the permanence of LTM storage remains open.

There is no question, however, about the role of the chunking process in LTM. Chunking is the most effective method of guaranteeing long-term storage. Tests on mnemonists, people with exceptional powers of memory, show conclusively that structuring information in some imaginative way is the key to successful memorization. For instance, given a list of words to memorize, the mnemonist might weave them together into a fanciful story. Even if the process sounds silly, it works remarkably well.

So far we've only scratched the surface of the topic of human information-processing, and in an article of this size we can't do much more than that. Many questions remain unresolved. For instance, are STM and LTM two physically separate blocks of memory in the brain, as they would be in a calculator or computer? Probably not. For more fascinating unanswered questions try *Human Memory* by Roberta Klatzky ($6.95, W. H. Freeman and Co., 660 Market St., San Francisco, Calif., 94104). Much work remains to be done in this field, and if you like computers, you may just become hooked on cognitive psychology—the psychology of comprehension. ■

The UART — It Makes I/O Possible

No, UART does not mean United Authority for Radio Tubes! It stands for *Universal Asynchronous Receiver Transmitter.* The UART is a special LSI integrated circuit that allows us to make a serial communications or I/O port for our microcomputers. With a single UART chip and a few other devices mounted on a blank S-100 prototyping card (such as the Vector® products), we can make a serial I/O port rather easily. So who is the designer of a serial I/O for *your* microcomputer? Well, with the UART, *thou* art!

Serial Data Transmission. Transmission of data from one unit to another might take place over a distance of a few feet, a few yards, several miles or across a continent. Except in the shortest runs, e.g., from one equipment cabinet to another, it is usually more economical to use *serial* data transmission, in which the bits of the world are sent over the communications channel one by one. If we wanted to use parallel transmission, which is a lot faster, it would require not less than one wire or radio or telephone channel for each bit transmitted—and that can be expensive! In serial transmission, only one channel is needed, so that it is the lowest cost way of accomplishing the job!

But microcomputers tend to have parallel organization. The eight-bit data bus is a parallel structure. Before we can transmit the data over a serial channel, therefore, we have to assemble the data in serial form. This was once a bit difficult, and required the interaction of both hardware and software, but today, we can make the job a lot easier because of the UART.

This LSI device will accept parallel format data on the transmitter inputs, send it out in serial form, and then reassemble it in parallel form on the receiver end. The typical UART contains both receiver and transmitter sections, so it may be used as either receiver, transmitter, or both.

The UART will also add certain bits needed in the parallel transmission. Figure 1 shows the general format for serial data transmission. The data line will sit HIGH (i.e. logical 1) when there is no transmission taking place. The data word transmission begins with a start bit, which is a LOW condition. The transmission equipment, and the receiver section of the UART, sees the HIGH-to-LOW transition of the start bit as the signal that the data word is beginning. Following the start bit, are five, six, seven or eight bits of the data word. The typical UART chip is programmable to permit any of these data word bit lengths. The ninth bit could be a parity bit.

The concept of parity is used to check the data transmission for errors. The parity can be odd, or even, and this refers to the number of HIGH bits in the data word. the UART will set the parity bit HIGH or LOW to make the parity of the total word equal to that programmed into the system. Finally, there are 1, 1.5 or 2 stop bits. For most of our applications, we will use 2 stop bits, but some applications (like teletypewriters, for example) require 1.5 stop bits.

Figure 2 shows a typical data word transmission (without parity bit). In this case, we are transmitting the word 101100101, and adding two stop bits.

Clocking In. Data transmission means that the operations of the transmission are independent of a master system clock. But this does not mean that the clock must be accurate, and relatively stable. In general, RC-timed clocks are not acceptable, so we must use crystal control of the clock.

Figure 3 shows the circuit for a suitable clock, while Fig. 4 shows the pinouts for various division ratios of the internal counter of the 4060 chip used in the oscillator. The 4060 chip is a multistage binary counter that includes its own oscillator stage.

The 4020 device, incidentally, is similar, except that it does not have the oscillator. Crystal Y1 will have a frequency selected by the desired baud rate of the data transmission, and the division ratio selected. The clock frequency applied to the input of the UART must be 16 times greater than the desired baud rate. If, for example, we want to drive a 300 baud printer from the UART, then clock frequency must be 16 × 300, or 4800 Hz. If we select the 2^8 output of the 4060, the division ratio is 256, so the frequency of the oscillator should have a crystal frequency of 256 × 4800 Hz, or, 1228.8 kHz. A frequency

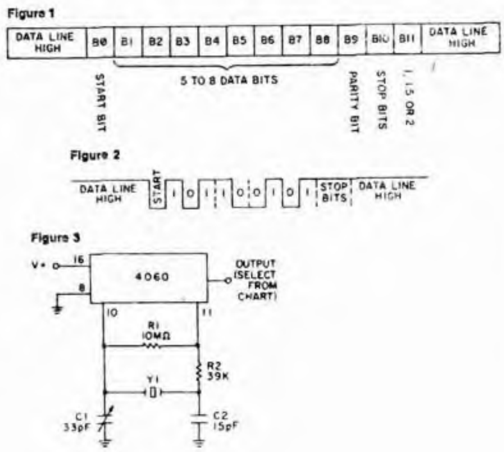


Fig. 1. The general format for serial data transmission. Fig. 2 shows a typical data word transmission without parity bit. In this case, we are transmitting the word 101100101, and adding two stop bits. Fig. 3. shows the circuit design for a suitable clock, utilizing a 4060 chip.

of 1.229 mHz is a standard crystal frequency, so that could used. Trimmer capacitor C1 could then be used to trim the frequency to exactly 1.2288 mHz.

The UART. Figure 5 shows the block diagram of a typical UART. This particular device is the 1602, but it is pin-for-pin compatible with devices such as the AY1013 and others. The two sections, receiver and transmitter, are completely independent except for the common connection to the power supply and control section. We can use either the receiver, the transmitter, or both sections at once.

The transmitter section contains an eight-bit *transmitter hold register*, and a *transmitter register*. The hold register is a buffer, or storage area. When the THRL (transmitter hold register load) control line is brought LOW data on the eight input lines is loaded into register. This data is transferred into the transmitter reggister for assembling into serial format and then for transmission.

The transmitter section also has certain other control lines, including the following: THRL, WLS1, WLS2, EPE, THRE, TRE, TRC, CRL and PI. These are defined as follows:

THRL	Transmit Hold Register Load
WLS1,2	Word Length Select
EPE	Even Parity Enable
THRE	Transmit Hold Register Empty
TRE	Transmit Register Empty
TRC	Transmit clock
CRL	Control Register Load
PI	Parity Inhibit

The receiver section is a mirror image of the transmitter section. The serial data is disassembled from serial to parallel format, and the eight parallel bits are stored in a receiver register. The receiver ouput is the ***hold register*** which is connected to the eight-bit output lines. These output lines are tri-state types, so will float at a high impedance when not enabled. This feature allows us to connect the UART directly to a data bus, an application that will be covered later.

The receiver section also has its own signals: PE, FE, OE, RRD, DRR, RRC and SFD; which are defined as follows:

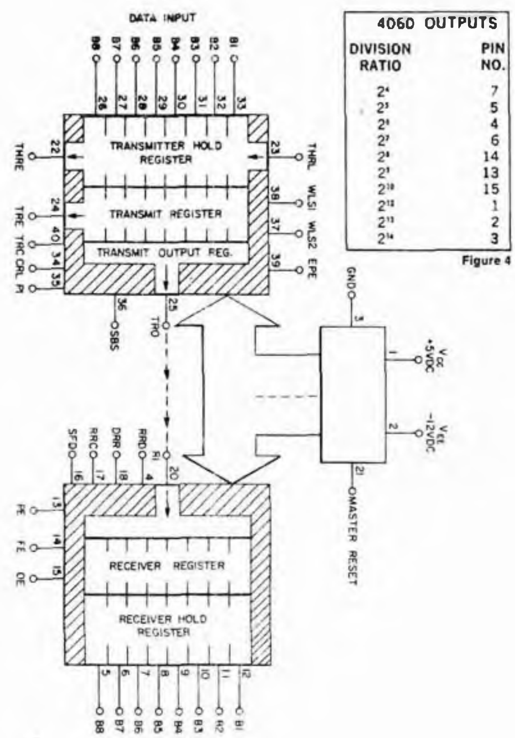


Fig. 5. This is the block diagram of a typical UART. This particular device is the 1602, but it is pin-for-pin compatible with other devices such as the AY1013. The two sections, receiver and transmitter, are completely independent.

Figure 8

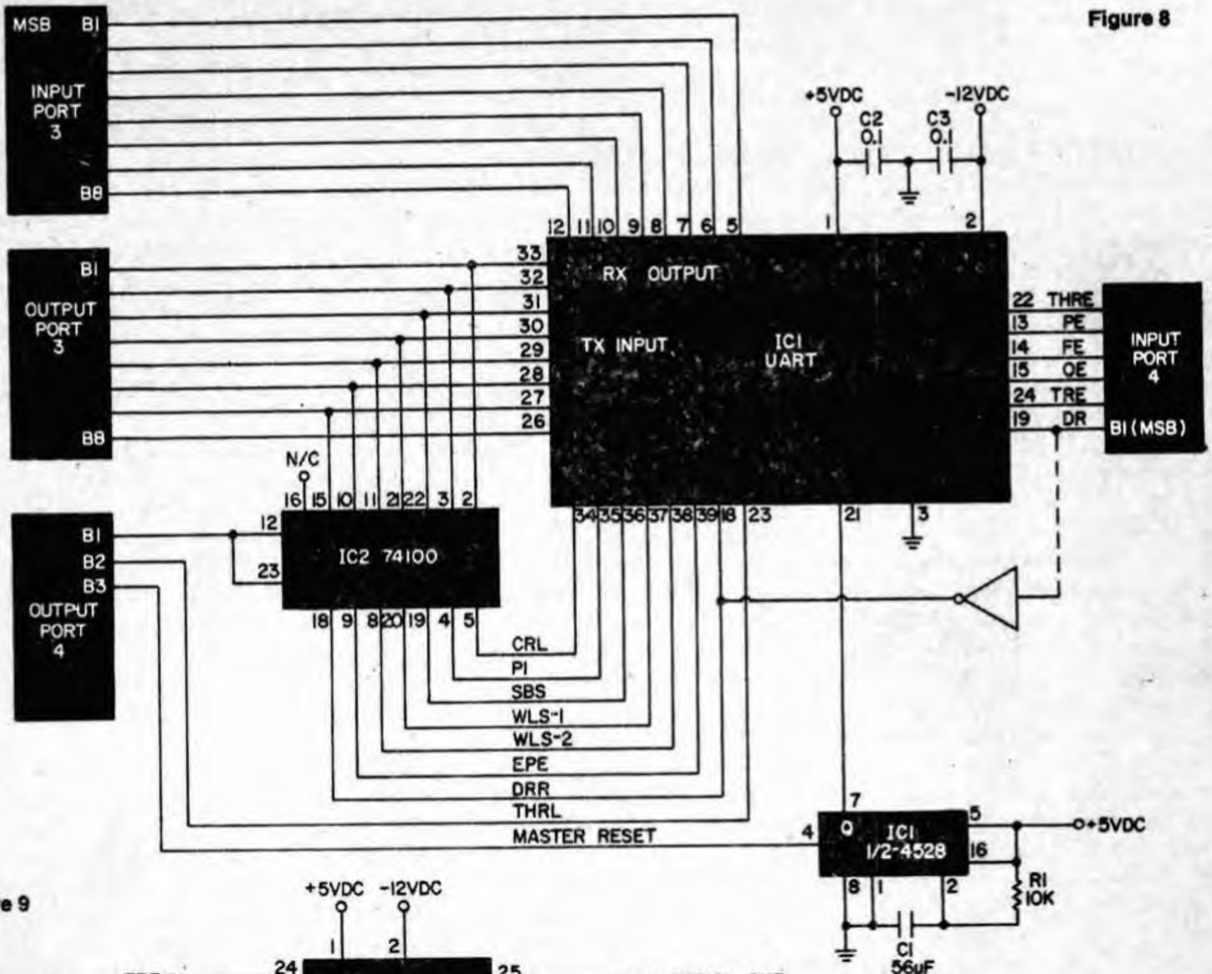
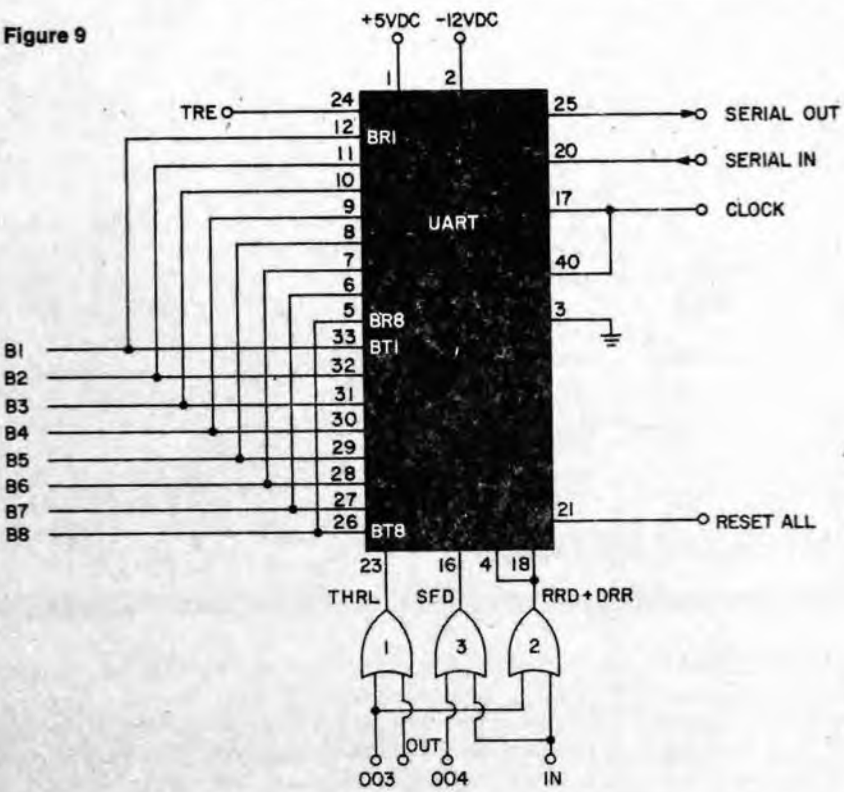


Figure 9



The transmitter section circuit of the UART is shown in Figure 6. The receiver section of UART is shown in Figure 7. The transmitter requires eight input bits, a clock at 16X the baud rate, and certain programming commands. The receiver section obeys the control signals wired into the transmitter section. Figure 8 shows the connection of the UART to a data bus. All the outputs of the UART are tri-state. The methods shown in Figure 8 are basically for the 8080A device, but similar connections will allow use of the Z-80 and others. Figure 9 shows a method for connecting the UART to I/O ports of an existing microcomputer device.

PE	Parity Error
FE	Framing Error
OE	Overrun Error
RRD	Receiver Register Disconnect
DRR	Data Receive Reset
RRC	Receiver Clock
SFD	Status Falg Disconnect

Basic UART Connections. One of the principal attractions of the UART device is the simplicity of its connection, allowing almost anybody to implement circuits based on the UART. The transmitter circuit is shown in Fig. 6, while the receiver circuit is shown in Fig. 7.

The transmitter requires eight input bits, a clock at 16× the baud rate, and certain programming commands. It also has a reset (pin no. 21 resets both the receiver and transmitter, as well as the status flags) control terminal. The control flags must be set either HIGH or LOW, depending upon the requirements of the situation. We apply +5 volts to each line through a 1000 ohm, ¼-watt resistor, which sets the terminal permanently HIGH. A grounding switch will set the pin LOW if desired. In most of our applications, I suspect that these pins will be HIGH. In that case, they are shorted together, and then connected to the +5 volt line through a single 1000 ohm resistor. In an application later in this article, we will allow setting of the control lines under program control.

The receiver section is shown in Fig. 7, and it is even simpler than the transmitter section if the transmitter section is also included in the circuit. The reason for this is that the receiver section obeys the control signals wired into the transmitter section. If the receiver section is to be used alone, however, we will also have to include the appropriate control signals. Otherwise the circuit is the same as in Fig. 7. We have a 16× clock, serial input, eight output data lines, and certain flags: DR, PE, FE, and OE. We also have the DRR signal input. In the circuit shown the *data received* (DR) line is inverted and applied to the *data receive reset* (DRR) line. The inverter allows us to rest the DR flag.

UART Interfacing. The UART can be interfaced to the computer in several different ways. We can connect it directly to the data bus, either as a memory-mapped device or a dedicated I/O port, or, we can connect it directly to an I/O port of the computer. One of the applications to follow will allow you to interface the UART to existing parallel I/O ports of a microcomputer.

Figure 8 shows connection of the UART to a data bus. This particular method of connection is possible because all of the outputs of the UART are tri-state. This means that, unless the output is enabled, the line floats at a high impedance to both ground and +5 volts. But, when the line is enabled, it will have a low impedance to +5 volts when HIGH and a low impedance to ground when LOW. In Fig. 8, we take advantage of this feature and connect the receiver output lines in parallel with the transmitter input lines and the data bus of the microprocessor. We can connect the *master reset* (pin number 21), to either the system power/on reset, or, to some point that can be used to reset the UART.

The methods shown in Fig. 8 are basically for the 8080A device, but similar connections will allow use of the Z-80, 6502 or 6800 devices. Also, programming for each microprocessor will be different, so no examples will be given here.

A method for connecting the UART to I/O ports of an existing microcomputer is shown in Fig. 9.

The circuit in Fig. 9 uses two I/O ports to control the UART. Input port number 3 is connected to the receiver output lines of the UART. The receiver status flags are connected to input port number 4, although some of these will not be used by some programmers. There are two output ports used (3 and 4). Port number 3 is used to output the eight data bits, plus the control signals to the control inputs of the UART. Port number 4 is used to control the circuitry.

The 74100 device is used to store the control signals output from port number 3 so that the same port can also be used for the data bits to follow. But this IC is not strictly needed because the UART contains its own register to input the control signals. We could connect PI, SBS, WLS1, WLS2, EPE, and DRR to the output lines in parallel with the transmitter input lines. The CRL (control register load) line can then be brought HIGH momentarily to strobe the data on the output port into the control register. If the 74100 is used, then bit B1 of port number 4 is used for this purpose.

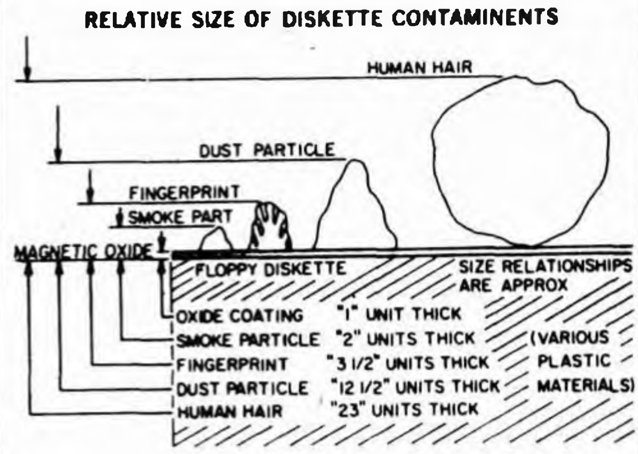
Bit B2 of port number 4 is used to strobe the transmit hold register load (THRL) with a LOW whenever we want to input the data on port number 3 into the transmit hold register of the UART. Bit B3 of port number 4 is used to reset the UART. This job is actually done by a positive-edge triggered monostable multi-vibrator (IC-3), which is triggered when bit B3 of port number 4 goes HIGH. ■

How Floppy Disk Systems Work

Sooner or later just about every serious computer hobbyist reaches their computer's memory limit. There are some applications that require too much I/O action to be practical with a cassette based system or require too much memory space to fit in a RAM. It's this need for high-volumn storage and rapid access time that makes a hobbyist disk system desirable. These systems don't come close to commerical hard disk systems in terms of performance but they can easily fill the computer hobbyist's memory bill.

Floppies have become the missing link to a midrange of random access memory systems. The floppy offers higher performance at lower cost than cassette and similar types of Input/Output (I/O) devices.

Well Packaged. The present standard floppy is an 8-inch flexible disk of a plastic material coated with a



The metal oxide surface of a floppy disk is very sensitive to contamination. Each one of these contaminants (right) could cause a dropped bit. Since the disk is always kept in it's protective envelope, this is usually not too much of a problem in daily usage.

magnetic oxide. Looking a little like a popular 45 RPM record, it is sealed inside a jacket and there are no grooves on the surface.

The disk cannot be removed from the jacket which is designed to protect the recording surface. The disk is visable at a slot, a spot, and a hole in the center of the jacket. Users are told by the instructions that we must not touch exposed areas of the disk or write on it with anything firmer than a Q tip. Finally the user is admonished to return the jacketed disk to its outer envelope after they have finished using it.

The natural skin oil of fingerprints can damage the quality of music in needle and groove recordings, and in the super-miniature world of floppies, a fingerprint can destroy an entire segment of data. A dust particle can waste a dozen sectors and a human hair can reduce the effectiveness of the *Read/Write/Erase (R/W/E)* heads.

It is understandable why the media, as the flexible disk is sometimes called, is permanently sealed with all of those implicit instructions. There is, however, an internal jacket wiper that continually cleans the rotating disk and removes contaminents, and floppies are reasonably rugged.

Hardware. The disk-drive hardware is add-on equipment to the main frame of the computer. Inside the drive are motors, driving mechanisms, and interfacing electronics that enable the drive unit to "talk" with the controller.

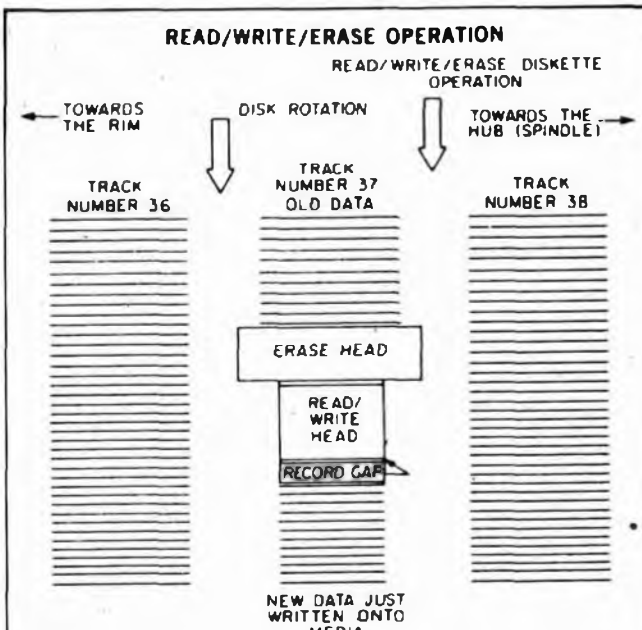
The diskette is inserted into the drive unit through a small door in the front. Once the door is shut, it is locked by the drive unit logic until the door release button is pushed to disable the drive assembly. The drive spindle centers and grasps the center of the diskette firmly as the motor comes up to speed. During power-up the diskette reaches a speed of 360 RPM, and *R/W/E* heads are stepped out to track 00 and a mechanical index hole provides the first location pulse for disk timing. In the IBM format this is the only reference to a physical location on the disk.

The floppy is firmly held against the recording surface, and the heads are positioned by a precise stepping motor. While the heads are positioned over the desired track, they ride above the spinning diskette. Once the correct track is located, in what is called a "seek operation," a head loading coil pulls the heads down onto the magnetic surface. This operation is called loading the heads and it is controlled by a computer program.

When the heads are to be moved again, in another seek operation, to a new track, they are first unloaded, stepped to a new track, and then reloaded to a new track. This requires about three milliseconds per track of movement. It will further require about 50 milliseconds, per track, to move the heads and about 15 milliseconds of settling time.

Part of System. Floppies should always be thought of as a part of a sophisticated cluster of mechanical drivers, computer electronics, and software called a computer. Fundamentally this computer consists of a central processor unit (CPU), some memory, some interfacing devices, called controllers here, and an entry device such as a keyboard CRT, and some software.

Computer operation is made possible by a written program entered into the computer's memory and operated on by the internal microprocessor. The programs are fed by any one of a number of techniques: a paper tape punch, keyboard, cassette tape, or teletype. All of these are slow and time consuming. New ways are continually being introduced to feed the voracious appetite of computer memories. Floppies are the most



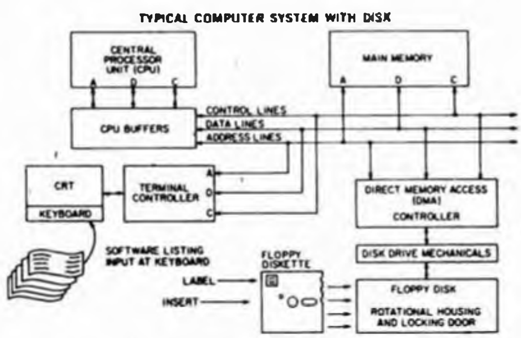
There is no physical track on the disk as with an LP record. The tracks are formed by the passage of the record (write) head over the surface of the disk. The disk itself appears to be a solid sheet of recording tape material.

versatile of the program instruction loading techniques.

In a computer's time frame things go on a million times faster than in our brain's time frame. In such a whirlwind existence, telling computers what to do was a difficult problem. The answer was in the development of software which the computer could store in memory for reference each time it needed a new program instruction.

Software. The computer—the IC's, printed circuit boards, filter capacitors, chassis, and power supply—are all part of the hardware. All of this state-of-the-art electronics is just so much junk without a program, a way to make the computer compute. The instructions, written in computer languages, are called software.

We talk with the floppy (human to floppy) through a series of interpreters comprised of software and hardware. First we place our instructions, using perhaps, Fortran in a machine assembly language. Located within this language and acting as a general interpreter, is a section of software called a file manager. Its job is to take general instructions, count words to be put on the floppy, caluclate the number of sectors, tracks and floppies that will be needed to store your file. All we do is to tell the file manager how much and where; it will do the rest. It will even put the data on the floppy then check to see if it got there, and if not correct its own errors. The special language of the computer is based on the numbers zero and one.



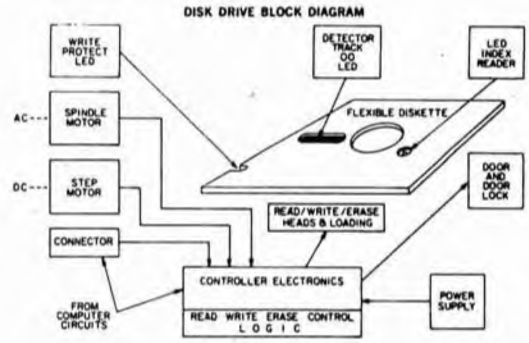
This block diagram of a typical computer system utilizing CRT readout and disk drive shows the parallel connections of the memory, drive controller and terminal controller.

These two values comprise one "bit" of information. At this level there are no grey areas, no informational maybes. Facts are either a zero or they are a one. In this language a word has only one length of the microcomputer. It is 16-bits long. That is 16-bits having two states or thirty two pieces of information. However most peripherial devices, such as the floppy controller, that hunk of electronics that interfaces (talks) between the floppy and the CPU, is designed to speak in half words. This half word is called *Byte*, pronounced "bite." A byte has 8-bits so there are two bytes to a full 16-bit word computer.

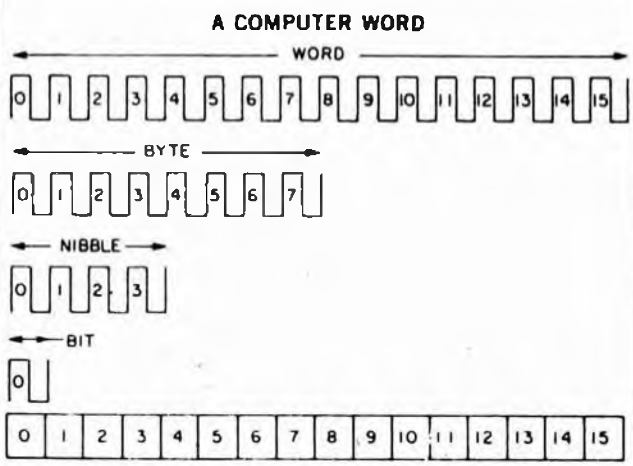
The story doesn't stop here. A new term is emerging in the industry as they learn to manipulate the byte. It is the half byte and is called a nibble.

Disk Mechanics. The R/W/E slot in the floppy jacket is two inches long by ½-inch wide. That little hole on the opposite side (8 inch diskette) is the mechanical index hole, which is a physical starting place (read by a LED sensor) for recording information on the oxide surface.

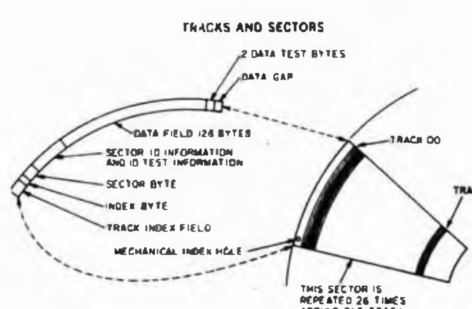
On a tape cassette you can break off a plastic tab thereby preventing further over recording on that



You can easily see from this diagram how the physical configuration of the disk is related the the various electronic systems needed to control information storage and retrieval.

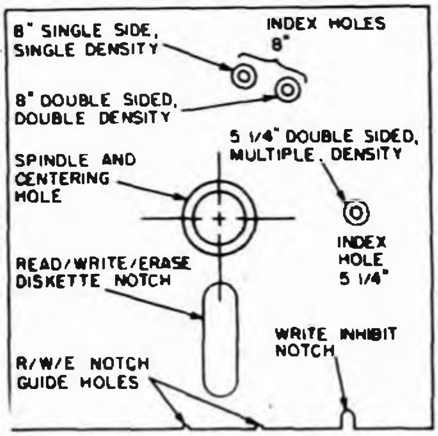


Software is the artful manipulation of the WORDS, bytes, nibbles, and bits utilized in transferring something meaningful between humans and machines without resulting in the total disorientation of the two parties involved.



This diagram shows the relationship between track information and sector information as they are physically located upon the disk when utilizing the IBM-type disk formatting.

HOLE AND NOTCH LOCATIONS



The disk's guide holes serve the same purpose as the arm on a record changer which positions the tonearm according to LP size.

cassette. An industry term for this is 'record protect.' Floppy disk jackets have this same feature, in a slightly different form. On some jackets there is a write protect slot cut in the paper of the jacket. When not needed the slot is covered over. This slot permits a photo-optical system to shut off the *write* electronics when the system detects a slot. In this way valuable programs already recorded are not destroyed by writing over the program. With some drive units there is a switch to override this system, with other drives we must tape over the write protect slot.

The erase head most widely used may be either a tunnel-erase or a stradle-erase head. The tunnel-erase head design minimizes the influence of noise from data in adjacent tracks. This more clearly defines the erase band and improves the signal to noise ratio. Present usage seems to favor the tunnel-erase head design.

To place information on the diskette (to *write*) and to retrieve information (to *read*) a software plan called a *format* is employed to pre-organize diskette data fields. Where only a single reference is made to the mechanical index the resulting formatting is called soft sectoring. Most of the information presented here is for a single density, soft-sectored, IBM formatted diskette. After receiving the pulse that represents the index hole, the rest of the floppy is formatted from the software, or computer program.

IBM Format. The IBM 3470 formatted diskette, one of the more popular industry standards, has 77 tracks, with 26 sectors, (data spaces) formatted per track. There are 128 bytes per sector, 256 nibbles, or 1024 bits. The total numbers of sectors per single density diskette is 2002. Of the tracks, 74 are for data storage, two are set aside as alternate "bad tracks reserves," and another track is reserved for maintenance purposed. A typical data transfer rate from floppy to controller is 250 kilobytes per second.

The tracks are numbered from the outside in with number 00 on the rim. Number 76 is the last track and is nearest the hub. Remember that there are no real tracks that you can see. They are the products of a software format, in this case IBM format 3470. These single sided, FM coded (more about that in a moment) disks can have a recording density of 3408 bits per inch.

A second type of sectoring is called hard-sectoring. Thirty two holes are cut in the diskette. These become the index marks for each sectored area. There is a 23% increase in data packing in hard-sectoring but the industry seems to prefer the soft-sectored format.

The sectors each contain a data field, with data gaps to guard this information, sector and track identification, again with gaps to protect this information, and guard bytes to further isolate sectors of information. All activity is in byte-length half-word groups.

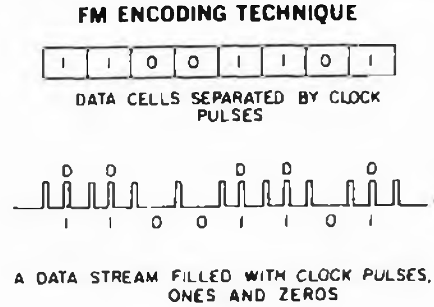
Frequency Mod. The technique used to place data on the diskette is called Frequency Modulation (FM). Clock Pulses, 4 micro-seconds apart, from a 250-KHz clock generator, are placed on the floppy sectors forming data cells. This is the time from one clock pulse to the next. A magnetic transition within a data cell is read from the data stream as a one; no transition as a zero. The data bit will fall into the center of a 4 microsecond data cell.

When the controller performs a read operation, the data stream going back to the drive electronics and on to the controller where the data is separated from the clock pulses, a known frequency, and also the index mark, sector data, and addressing information is removed before sending the date on to the Central processing unit.

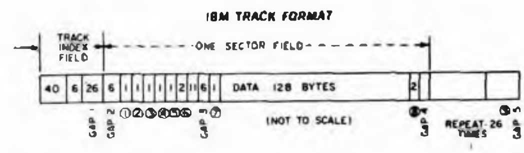
As you might imagine there are sometimes errors in either a read or a write operation. A software test program is designed to search out these errors. They will be classed as either hard or soft errors. That is an error traceable to a piece of hardware such as a faulty motor is a hard error. An error traceable to a poorly formatted sector will be a soft error. One possible soft error would be a bad sector such that data could not be written into the sector. The controller would store the address of that sector in a *Bad-Sector* file in memory and search out another sector. Later if the computer addressed that bad-sector, the file manager would discover that it was bad and immediately go to the address of the new sector used in place of the damaged one. At some future date the user might want to replace the floppy if there are too many bad sectors.

As the industry gets better at the techniques to pack and crunch data onto a small recording surface, floppy usage will increase. Shugart and Perteck, among others, are offering the double density floppy employing a recording code permitting double data packing using the MFM or Modified Frequency Modulation code. Two

FM ENCODING TECHNIQUE



The binary data pulse stream (lower line) is translated into and stored in the data cells (top line). 1's and 0's are shown.



IBM definitions: byte 1—address mark; 2—track #; 3—side #; 4—sector #; 5—sector size; 6—ID test bytes; 7—data mark; 8—data test bytes; 9—trailing gap.

other codes now in use are the Modified MFM or the Group Code Recording (GCR) technique. Double sided recording heads have also been introduced allowing recording on both sides of the floppy with transfer rates of 500 Kilobytes per second and 256 bytes per disk sector.

Mini-Floppy. Hardly had this double density, double sided floppy been introduced when the mini-floppy popped on the stage. Only 5¼ inches in diameter, it uses a coding scheme called the Modified MFM, with the index hole shifted 90 degrees to the right. The Radio Shack add-on TRS Mini-Disk System, uses a mini floppy. Cost is about $500 for the "first Disk-Drive." As with all disk systems, a certain number of sectors are devotes to housekeeping, sometimes called labeling. Those sectors usually consists of a directory, test programs, and other sector and track information so that it will talk smoothly with its computer system.

The flexible diskette is a very useful peripheral device used with a I/O controller and can be used to store special diagnostics (hardware test programs) and debugging routines. ■

Computer Operating Systems

Eventually, every enterprising home computerist, if he has been using cassette storage for any length of time, will look at the system he started out with and begin dreaming of upgrading it to do even more wondrous things. Visions of printers, X-Y plotters, color graphics, and controlling external devices will dance in the heads of those who have become relatively competent at programming and operating their basic systems.

Disk Drives. One of the first and most popular upgrades to a personal computer is, of course, the addition of a floppy disk drive. It is also one of the most dramatic. A disk system, Fig. 1, really makes a computer come alive and seem more "magical" after those interminable waits for a program to load from cassette. It also makes the system a serious contender for use as a business tool.

The big advantage of a disk system is the increased storage obtained and the speed with which stored programs and data can be read and utilized. We hear much about single and double density, access time,

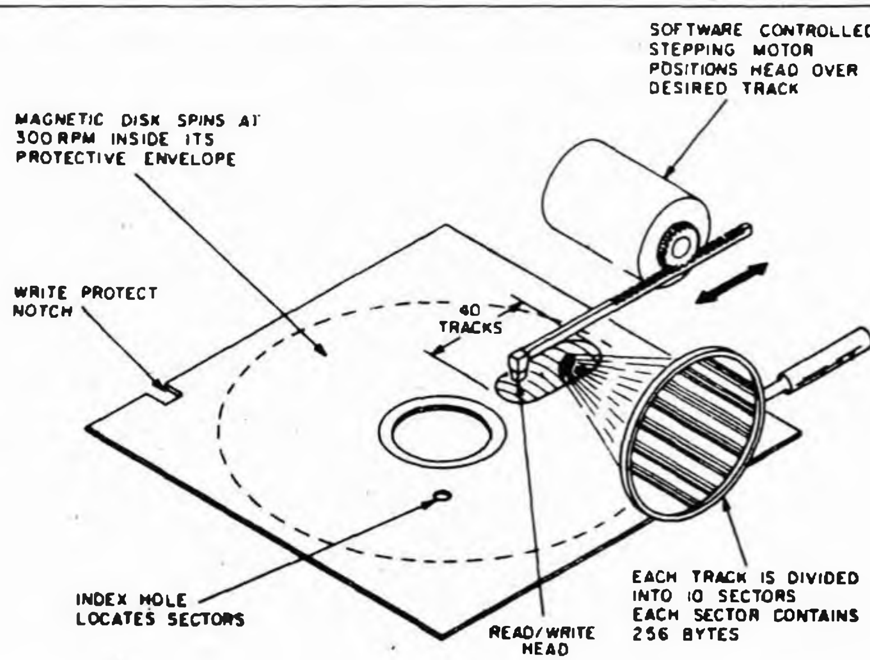


Fig. 1. There are many different sizes and types of floppy disks but they all look much like the one in this diagram. The disk tracks are not visible lines but rather sectors of a single magnetic coating. The read/write head is positioned by a stepping motor. Some systems have read/write heads for both sides of the disk. Index hole marks the first sector.

formats, etc., that is, the physical capacity of the disk itself. But one major consideration is often neglected. It takes a very sophisticated computer program, that is, software, to even begin to use the capabilities of a floppy disk system.

Operating Systems. Such a program, called an *operating system*, is far beyond the abilities of most home computer users to write and is therefore usually supplied by the manufacturer or by an independent software house. Anyone who is contemplating upgrading his system to include floppy disks should look well to the type of operating system that will be available to him. Also, by examining what an operating system is and what it does, we will be a better idea of some of the true potentials of today's microcomputers.

Essentially, the operating system is a program which supervises and coordinates the operation of all the various programs and peripheral devices (printers, disk drives, terminals, etc.) operating on the computer system. It keeps track of the order in which things are done, the names of different programs in memory, their locations, and it allows the user to tailor the system's operation to optimally suit his needs.

To understand how the operating system does this, we will need to use the concepts of 'devices" and "files". A device can be any peripheral device connected to the computer: a CRT terminal, a p rinter, a disk drive, and X-Y plotter, etc. The operating system communicates with and uses these devices by means of special sub-programs known as *device drivers*.

Different peripherals reqauire different software routines to drive them and an operating system may not have drivers for all possible devices. Nonetheless, a set of such routines is essential if the computer is a to communicate with external contraptions, be they teletypewriters or robot arms. The ambitious programmer who wishes to run some exotic gizmo with his computer usually has to deal himself with writing the software to run it as well as building the hardware. A good operating system, though, will provide an easy method for the user to add device drivers of his own making.

Files. The concept of a "file" is a little more complicated. To explain it, w will use the example of the floppy disk. The Floppy disk is both a "device" which requires its own device driver, and a gigantic "file" itself. It is the place where all the other blocks of data and programs, or "files" are kept.

In fact, you can think of the disk drive, the device, as the filing cabinet, the physical storage unit. All the pieces of paper and records in the filing cabinet, then, correspond to the data blocks or files, recorded on the disk. The way the date is stored on the disk is illustrated in the accompanying diagram.

The floppy disk is a circular piece of material much like magnetic recording tapes. It rotates inside a protective envelope under the read/write head of the disk drive. The disk is divided into a number of concentric tracks, and each track into a number of sectors A sector is the smallest unit of information the disk can access at one time. In a typical home computer diskette, each sector contains 256 bytes and each track contains ten sectors. On a diskette with 40 tracks, that

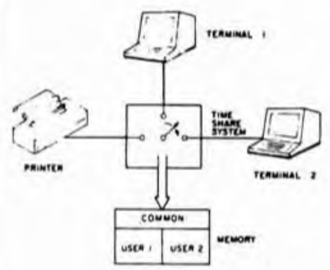


Fig. 2. In a time sharing system various terminals, printers and other "slow" peripherals share the use of a single computer to enable more efficient use of the machine. This requires a special operating system to keep the different inputs and programs sorted.

addes up to over 102,000 bytes of storage!

With the disk rotating continuously under it, the read/write head moves quickly in and out until it has found the track requested by the computer. It then waits over that track until the proper sector comes around. At that point the head "ta's" the disk and either reads data from it into the computer or writes data onto the disk.

File Finder. The job of the operating system in all of this is to keep track of where all the things are on the disk so they can be found when requested by a program. The operating system acts much the same as a poor, harrassed file clerk who is constantly asked to fetch and file records. To ease its task, it generally uses part of the disk to establish a *directory* for itself. The directory is a little refeence file the operating system uses to tell itself where everything is stored on the disk.

The name of each file (a file can consist of as many sectors as necessary, all grouped under one name) is chosen by the user and generally means something in English, such as STARTRK or BIOTHM. The operating system reads the file name when it is enterd on the terminal, then looks it up on the disk directory. From the directory it finds out what track and sector it must go to to find the beginning of the file. We might also liken this process to finding a book in the library. First we go to the card catalog (the disk directory) and find out what aisle and shelf (track and sector) the book is on, then we go get the book.

So far, we have seen the actions of the operating system with the main means of mass storage, the disk drive. Let us now expolore some of the things that are made possibly by a typical microcomputer operating system. Not all the methods described here work for all operating systems because the approach of different authors varies. The general concepts, however, apply to all operating systems.

One major thing we can do with an operating system is take a file and transfer it from one device to another, or copy it from one disk to another. Many operating systems contain a program called PIP (Peripheral Interchange Program), which simply copies a file from one device to another. Using this program, a businessman can record the status of his inventory or reciepts on a disk kept outside his shop and update that backup disk, say, every week. Thus, if anything should happen to his daily working disk, he only has to go back one week in his paper records to get things straightened out. A better idea is to update the backup disk daily—with PIP it only takes a few seconds.

An operating system increases in versatility if it is able

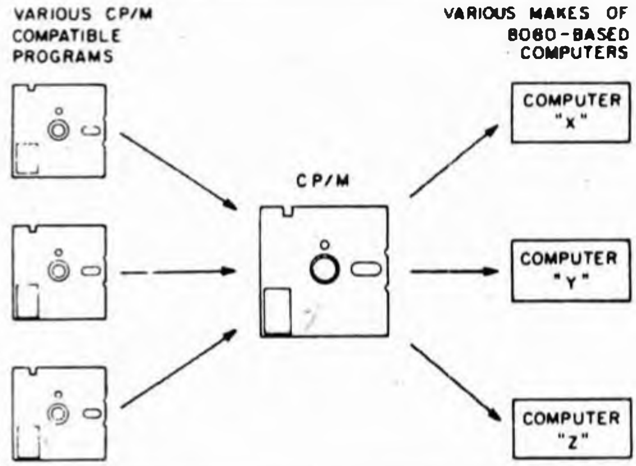


Fig 3. The CP/M operating system can be modified slightly to run on any 8080-based micro-computer. Applicatons programs can then be written to conform to CP/M's requirements rather than those of the computer. Thus, any computer running CP/M can use any of the thousands of independently produced programs written to run under that operating system.

to treat certain peripheral *devices* as if they were *files*. In that way it can use them as soruces of data to be copies to a corresponding file on the disk and later used by a program. For instance, if a program is monitoring a temperature at some peripheral, that information can be read as a file by the operating system and stored as a file on the disk. It can later be read as a file from the disk and copies as a file to the printr or used as a data file by another program.

With various peripheral devices connnected to the computer (CRT, printer, disk, etc.) the operating system must always be aware of which one is really running the show, that is, which terminal is allowed to give commands to the operating system, tell it to shut up, power down, or fetch a file. This terminal is known as the *console terminal*.

In most cases, the console terminal is the CRT terminal since many printers don't even have keyboards. It is, however possible to have more than one console terminal and more than one user program working on the same computer system at the same time. This is done by means of a method called *time sharing*, and it is one of the most powerful features available on today's computers.

Thumb Twiddling. Peripheral devices, especially those with mechanical parts, are terribly slow when compared to the speed of a microprocessor CPU. While the printer is busy printing the letter "G," the CPU is spending most of its time twiddling its electronic thumbs waiting until the printer is ready for it to send the next character to be printed.

A time sharing system makes use of these idle moments by paying attention to other console terminals. To do this, it "polls" them one after another in much the same way an automobile distributer services spark plus in rapid succession. The operating system also has to assign certain portions of memory to each user and keep track of who is where and when.

In addition, the system must keep track of priorities: if two terminals are ready for its attention at the same time, it must decide "who's on first." Thus, in a time sharing computer, our operating system not only plays the role of file clerk and scheduler, but, it seems, that of diplimat. See fig. 2.

Time shared systems on microcomputers seem to be *the* coming thing for small to medium sized businesses. Studies have shown that what a small businessman needs is a system with four terminals or less which sells for under $10,000. And manufacturers are currently falling all over each other trying to meet that need.

A typical such system might be found in a parts store where inventory control is critical. One terminal might be out at the counter or in the shipping area. Sales and shipments are entered here and the computer than automatically subtracts the parts from the inventory file and adds the sale amount.

Back in the warehouse, there could be a terminal where incoming parts would be entered and added to the inventory file. Meanwhile, in the office, inventory reporting and bookkeeping could be done using both files with a third terminal along with a printer. And all this could be (and *is* being) done by a single computer with the necessdary memory, disk space, and an efficient operating system.

One thing must be noted, though. There is a distinct difference between the operating system and the various programs and computer languages which do the actual programs and calculations that the user finds useful. The operating system actually just does the housekeeping. But it represents someting more, something very important.

The operating system provides a software *interface* between a computer language like BASIC or a large application program like "Geneal Ledger" and the piece of hardware that is the actual computer. So we often speak of a language or a program that will run *under* a certain operating system. That means, no matter what sort of funny piece of hardware it may be, if a certain operating system has been adapted to run on it, any piece of software written to run on that machine also.

D.O.S. A number of manufacturers have produced Disk Operating Systems to run on their machines. Processor Technology ahs PTDOS. Heathkit has HDOS, Cromemco has CDOS, etc. There is also a very popular operating system which has been modified to fit a wide varieyt of computers. It is called CP/M (Fig. 3) and was produced by Digital Research (P.O. Box 579, Pacific Grove, CA 93950). The reason CP/M (for Control Program/Micro) is so popular is that it has been adapted to run on a number of different machines.

Programs written by many indenendent software companies need then only be written in such a way that they run under CP/M and they can then be used on any computer that has a CP/M operating system. This is a big incentive for independent producers of software to get busy and write programs that are CP/M compatible because they know they will find a large market.

Having more software available in turn makes the computer more versatile for the user, and increases its value to him. Expanding a personal computer from a hobby to a serious tool for business or career still represents a sizeable investment. The add-ons are themselves expense, but the user should look beyond that to the features of the operating system he will be used—both at what he can do with the system itself and how much more of the world of micro-computer software it will open up to him.

Computer Communications

Most people who have begun investigating the world of computers are probably familiar with the most common means of communicating with the machine—the keyboard terminal. This type of input output, or I O, device reads the different keystrokes typed by the operator and converts each into a distinct pattern of I's and O's which are then transmitted to the computer through one of its I O *ports*. A port is simply a location (sometimes a memory address) where the CPU looks for incoming information from the terminal and to which it transfers data to be read by the terminal.

For the computer to receive meaningful data from a keyboard terminal, the operator must first formulate it in terms of the alphanumeric characters available on the keyboard and the computer must be programmed to interpret the symbols sent to it. This works quite well for a great many applications, but what if we want the computer to do something that requires so much constant input that the task of formulating it and typing it in all the time would be too dull and repetitive for a human to put up with all the time?

Computer Control. Let us take for an example, the task of reading—and later controlling—the temperature in some device. While that job could be done by a simple thermostat, we will see that using a computer will enable us to expand incredibly the possibilities of using the temperature data it reads.

The first problem is to get the information into the machine. We could have an operator read a thermometer and type the readings in via the keyboard, but he would soon get frustrated, especially if we wanted readings every half second or so. With the computer, we can do it automatically. This whole subject of automatic data gathering and external control is called *process* control. It is simply obtaining information from the real world, processing that information according to a program in the computer's memory and performing operations based on that information.

That means that an incredible variety of specialized I O devices and programs must be devised and tailored to the specific application. These can range from sensors of all kinds to stepper motors, valves and electric trains. We will now look at some of the general methods that are used to interface a computer to the outside world.

Analog to Digital. Since the real world does not normally present convenient digital data, some method must be found to translate *analog* signals into digital data. The general device for doing this is called an analog to digital, or A/D convberter. An analog signal such as an audio wave, has constantly changing values so the best a digital device can hope to do is to *sample* positions of the wave and to approximate its shape in the form of digital information. Given the speeds at which computers operate, this can be done with surprising accuracy.

Let us look at the wave form shown here. It is applied as a continuously varying voltage to the input of the A/D converter which is at one of the computer's input ports. The computer than reads the instantaneous value at regular intervals and accepts the digital value through its port. It can thus form an approximation of the wave form. The greater the frequency of this *sampling rate*, the more accurate the digital approximation will be.

Each sample that is read by the computer represents the value of the analog signal at a given point of time. For an eight-bit A/D converter, this value can be resolved into 256 increments ($2^8 + 256$). For example, it would be possible for such a device to measure a range of temperatures from) degrees to 100 degrees Celsius to an accuracy of better than 0.4 degrees Celcius! This accuracy can be enhanced by more complex (and more expensive) 10- or 16-bit converters.

The opposite of A/D conversion, digital to analog conversion, is also done, often on the same circuit board known, naturally, as an ADAC. Here, the digital information is output at a specific rate to a network of resistors, filters and op amps to produce an analog signal. The most sophisticated D/A converters are used in computer music applications where the highest quality complex waveforms are essential.

One inexpensive example of these principles, now available to hobbyists, is a device that allows the user to digitally record his voice in the computer's memory and then play it back by having the computer output those memory contents through the D/A section of the device to a speaker. In all this there is a tradeoff the user must be aware of. The faster the sampling rate and hence the more accurate the representation of the wave form, the more information must be stored—and the more memory must be taken up with stored data.

Of course, most real world applications are concerned with more than just inputting and outputting the same information. The real power of the computer, after all, is to process the information and make decisions based on it. And the true purpose of real-world connections is to enable the computer to act independently on those decisions. In this context, then, they are merely the things which implement the end results of the processing which is the computer's main task.

Control. By outputting analog signals, the machine could be used to control conventional analog device. Let us take a look at our example of temperature control. The A/D converter, under software control, samples and presents temperature readings to the computer which reads and compares this data with its program information. If the temperature is outside the specified parameters, the computer outputs a digital value to the D/A converter which applies the proper analog voltage to the motorized temperature control until the readings again satisfy the limits.

Well, we certainly don't have to invest in a computer simply to sue it as a thermostat! As I said and will emphasize again, the power of the computer lies in its speed and its ability to analyze large amounts of data and make decisions based upon its. So let us set our machine a somewhat more challenging task.

Multiple Data. We have a chemical process in which the computer must simultaneously mointor several thermometers, flow meters and pH meters, control various valves and temperature controls, and, in addition, prepare a human-report on materials used, temperature variations and the like. In addition, it must be alert for critical conditios (or combinations of conditions) and sound an alarm, and shut down operations if necessary. It should be apparent that the main problem here is not so much in the various on-off or analog device at the machine's ports, but more in the considerations put into the software program.

Having the computer monitor many different inputs simultaneously is accomplished with an analog multiplexer. A typical such device can monitor 16 channels by rapidly scanning the inputs. In much the same way, an automobile distributor can service 8 different spark plugs seemingly at the same time.

Parts of the computer's progrtam can then store the different values in memory locations alloted to the different input channels. This information can then be manipulated analysed and output in whatever manner the programmer thinks best.

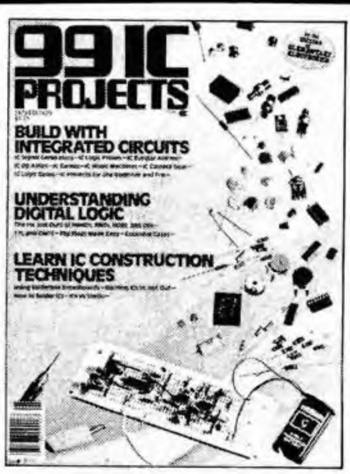
There are two more devices that are very important in process control applications. The first is the *real time clock*. Since all computer functions are controlled by various timing circuits in "clocked logic" there is a central timing reference in the form of a crystal oscillator known as the system clock. It's typical frequencies in home computers are between 2 and 4 MHZ. Since a crystal controlled clock is extremely accurate, divider circuits can be added to give the system an internal source of time reference calibrated in hours, minutes, seconds and fractions of seconds. The computer can then be programmed to refer to this time standard in order to time and delay the external processes it is controlling.

The other important device is one which enables the computer to turn on and off high voltage or high power devices safely. Obviously, the low voltage levels present at the computer's output port are not capable of controlling large devices by themselves. One solution is to use them to turn relays on and off which in turn controls the devices that do the actual work. The problem with relays and transistor switches is that there is danger of high voltage getting back into the computer and destroying components. The solution for simple on-off functions is to isolate the voltage with an *opto-isolator*. As the diagram shows, this is a device consisting of a photo transistor and an LED. The LED is connected to the computer output and is totally electrically isolated from the phototransistor. The phototransistor is turned on by the light emitted when the LED is on. It can thus be used to turn on any type of relay or other device while protecting the computer circuits from dangerous voltage.

This is the type of device we would use to sound alarms and activate sprinkler systems of our computer detected dangerous conditions, in our chemical process. Of course, with the computer in charge of things that should not happen.

If it did, the operators could use the computer to analyze the recorded data and the records of its control functions to generate a report on the entire process.

And that brings us back to the point I have been emphasizing. The program control and analysis of the data are the crucial aspects. The external devices must sense accurately and carry out the mahcine's orders exactly. But the whole system is still merely a tool which is used to extend the pwoer of the human mind.



Learn Through Building

You have by now absorbed many pages of basic theory and the chances are you will not retain much of it unless you put your newly gained experience to work. One way to do this is by building projects. Should you select a project that is enjoyable to build and use afterwards, the experienced gained and the application of your recently obtained information will be reinforced in your memory. Furthermore, the larger your basic theory base be, the more theory you will be able to understand and retain at a later date. The building blocks you gain today will become structures of knowledge tomorrow.

You should find the building of projects a fun thing to do. At first, it may be expensive, but later, as you acquire a junkbox filled with spare and salvaged parts, the cost will diminish. Eventually, you may be able to assemble a project without any cost at all. Build, it makes for fun and retained knowledge.

PRECISION VOM CALIBRATOR

Until now, most of the calibrator circuits appearing in hobby magazines could not be considered as primary reference standards. Instead, they were **transfer standards,** since the builder would be instructed to align his calibrator using a voltage reference of known accuracy. The obvious reaction of most readers was: "If I had access to an accurate voltage reference to begin with, why would I want to build this calibrator?"

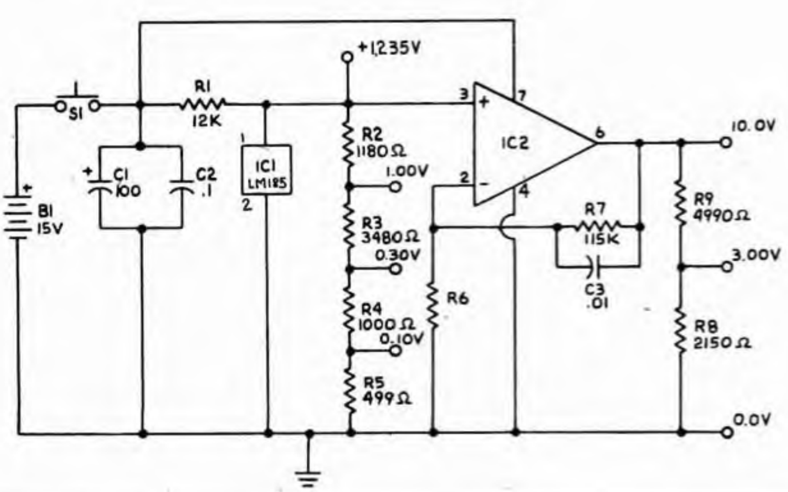
Our sentiments exactly. Now National Semiconductor comes to the rescue with a voltage reference IC, the LM185, having an output of 1.235 bolts 1%. What's more, this voltage remains stable in the face of changing ambient temperature and supply current.

The circuit diagrammed here produces six useful reference voltages from .100 V to 10.0 V. As noted above, the 1.235-volt output is accurate to within 1%. All of the other outputs are accurate to within 2% except for the 3-volt output, which has a tolerance of 4%. Reduced accuracy on all derived outputs is the result of errors introduced by the 1% resistor tolerances. Bear in mind, however, that worst-case accuracies are quoted here.

Be certain that the input resistance of the instrument being calibrated greatly exceeds the resistance at the circuit node being read. Most of you who worry about calibration have high-impedance (10-megohm) FET voltmeters, the loading effects of which are negligible here.

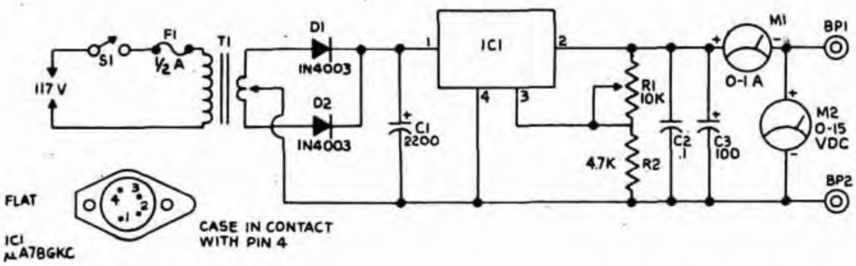
PARTS LIST FOR
PRECISION VOM CALIBRATOR

B1—ten AA cells in series to yield 15 volts
C1—100 uF, 25 V electrolytic capacitor
C2—.1 uF ceramic disc capacitor
C3—.01 uF polystyrene or mylar capacitor
IC1—LM185 1.235-volt reference IC (National Semiconductor)
IC2—3140A FET-input op amp (RCA)
All Resistors ½w, 1% precision unless noted otherwise
R1—12,000-ohm, ½-watt 10% resistor
R2—1,180-ohm, ½-watt resistor
R3—3,480-ohm, ½-watt resistor
R4—1,000-ohm, ½-watt resistor
R5—499-ohm, ½-watt resistor
R6—162,000-ohm, ½-watt resistor
R7—115,000-ohm, ½-watt resistor
R8—2,150-ohm, ½-watt resistor
R9—4,990-ohm, ½-watt resistor
S1—SPST normally open pushbutton switch



VARI-REG POWER SUPPLY

There are lots of good power supplies on the market, but why not build your own and save a bundle? This circuit can provide voltages between 5 and 15-volts DC at currents up to one ampere. Be sure to heat-sink the uA78GKC regulator by bolting it to either a commercial aluminum heat sink or to your supply's cabinet (if it's made of aluminum). Mount C2 and C3 as close as possible to pins 2 and 4 of IC1. If you cannot locate a 28 VCT transformer, go to something slightly higher, say 32 VCT. The same goes for the transformer's current rating; for example, you could use a 2-amp device.



PARTS LIST FOR VARI-REG POWER SUPPLY

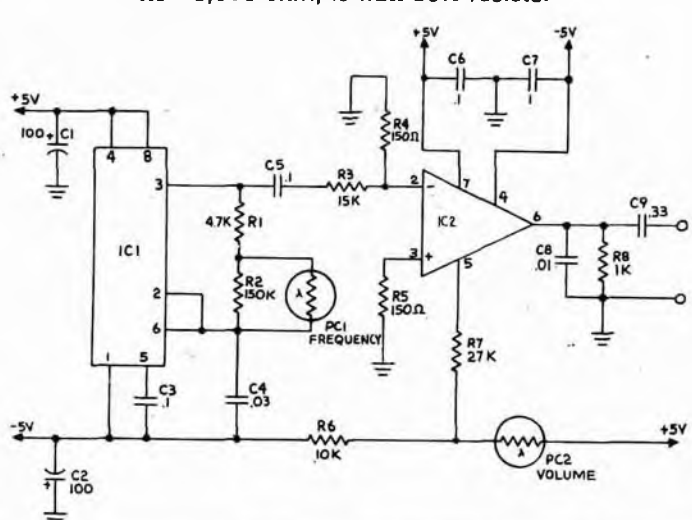
BP1, BP2—binding post
C1—2200-uF electrolytic capacitor, 40-WVDC
C2—0.1-uF ceramic disc capacitor, 35-WVDC
C3—100-uF electrolytic capacitor, 25-WVDC
D1, D2—1N4003 (1A, 200 PIV) rectifier diode
F1—0.5-Ampere slow-blow fuse
IC1—uA78GKC adjustable voltage regulator
M1—0-to-1 Amp DC meter
M2—0-to-15 Volt DC meter
R1—10K-ohm linear-taper potentiometer
R2—4700-ohm, ½-watt 5% resistor
S1—SPST toggle switch
T1—28VCT, 1.2-Amp power transformer (see text)

THEREMIN JUNIOR

Let's return now to prehistoric times, at least as far as electronic music is concerned. Way back then, nearly forty years ago, an odd-looking and equally odd-sounding instrument known as the Theremin was born.

PARTS LIST FOR THEREMIN JUNIOR

C1, C2—100 uF, 16-WVDC electrolytic capacitor
C3, C5, C6, C7—.1 uF ceramic disc capacitor
C4—.03 uF mylar capacitor
C8—.01 uF mylar capacitor
C9—.33 uF mylar capacitor
IC1—555 timer
IC2—RCA 3080 transconductance op-amp
PC1, PC2—cadmium sulfide photocell (Radio Shack 276-116 or equiv.)
R1—4,700-ohm, ½-watt 10% resistor
R2—150,000-ohm, ½-watt 10% resistor
R3—15,000-ohm, ½-watt 10% resistor
R4, R5—150-ohm, ½-watt 10% resistor
R6—10,000-ohm, ½-watt 10% resistor
R7—27,000-ohm, ½-watt 10% resistor
R8—1,000-ohm, ½-watt 10% resistor



Playing the Theremin entailed waving one's arms spastically between two sets of antennas. The purpose of all this was to modulate the RF fields in the vicinity of these antennas, thereby producing accompanying changes in the frequency and volume of the sound emitted by the instrument.

Controlling the sound was both difficult and inexact. As a result, the Theremin never gained widespread popularity, but was instead relegated to the domain of avant-garde composers and science-fiction-movie soundtracks.

Despite the shortcomings, the Theremin is great fun to play, so we decided to create a simple solid-state circuit, Theremin Junior, for those of you too young to have experienced the real thing. In this instance, photocells replace the Theremin's antennas. To play, you move your hands to cast shadows on two photocells, one of which controlls pitch—the other, volume. PC1, the pitch-control photocell, varies in resistance as the intensity of the light shining on its surface varies. This causes a change in the frequency of square-wave oscillator IC1.

Similarly, modulating PC2's resistance with light changes the voltage at pin 5 of IC2, which controls the gain of the circuit. High light intensity results in high frequency and high volume. Frequencies between 150 and 4800 Hz, approximately, can be produced at a maximum amplitude of about 0.5 volt peak-to-peak.

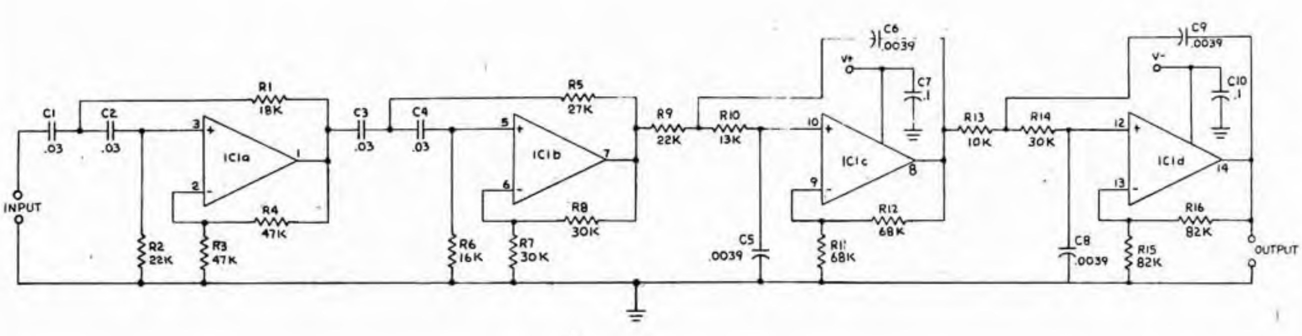
AUDIO BANDPASS FILTER

There are two different approaches to bandpass-filter design. The first involves use of a high-Q resonant network. You'll find this type of device sold as a CW filter, an application in which it excels. However, the selectivity of a resonant bandpass filter is such as to favor a very few frequencies to the exclusion of all others, and this makes it useless in voice reception. To filter the garbage out of an SSB transmission, you need a filter that freely passes the band of frequencies between about 300 and 2500 Hz but drastically attenuates frequencies outside the passband. An audio filter of this type is constructed by cascading (i.e., hooking in series) very sharp high- and low-pass fitlers.

That's what we've done here U1a and U1b comprise a sharp, 4-pole Butterworth high-pass filter with a 300-Hz cut-off. The two remaining stages function as a low-pass 4-pole Butterworth filter having a 2500-Hz cut-off frequency. Overall circuit gain is 16. Insert the filter into your receiver's audio chain at a point where the input signal level will be less than 100mV peak-to-peak. If the filter's extra gain causes problems, chop its output down with a resistive divider. A dual supply furnishing anywhere between 2.5 V and 15V can be used to power the circuit.

PARTS LIST FOR AUDIO BANDPASS FILTER

C1-C4—0.03-uF polystyrene capacitor
C5, C6, C8, C9—.0039-uF polystyrene capacitor
C7, C10—0.1-uF ceramic disc capacitor
IC1—LM324 quad op amp integrated circuit
R1—18,000-ohm, ½-watt 5% resistor
R2, R9—22,000-ohm, ½-watt 5% resistor
R3, R4—47,000-ohm, ½-watt 5% resistor
R5—27,000-ohm, ½-watt 5% resistor
R6—16,000-ohm, ½-watt 5% resistor
R7, R8, R14—30,000-ohm, ½-watt 5% resistor
R10—13,000-ohm, ½-watt 5% resistor
R11, R12—68,000-ohm, ½-watt 5% resistor
R13—10,000-ohm, ½-watt 5% resistor
R15, R16—82,000-ohm, ½-watt 5% resistor

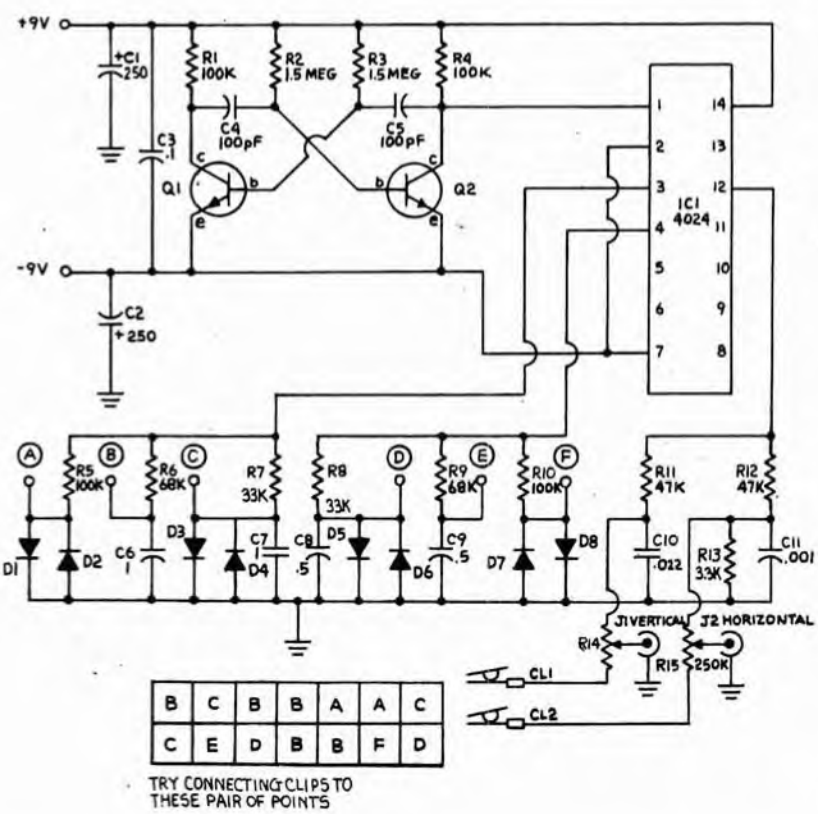


VIDEO PATTERN GENERATOR

Those of you with oscilloscopes might enjoy bread boarding this pattern generator. Feed the signal at J1 to your scope's vertical input, and connect the horizontal input to J2. Attach the clips to the selected pairs of test points, then adjust potentiometers R14 and R15 to create complex images. Output signals are about 1-volt, peak-to-peak.

PARTS LIST FOR VIDEO PATTERN GENERATOR

C1, C2—250-uF electrolytic capacitor, 25 VDC
C3—0.1-uF ceramic disc capacitor, 35 VDC
C4, C5—100-pF polystyrene capacitor, 35 VDC
C6, C7—1.0-uF mylar capactitor (non-polarized), 35 VDC
C8, C9—0.5-uF mylar capactitor 35 VDC
C10—0.022-uF mylar capacitor, 35 VDC
C11—0.001-uF mylar capacitor, 35 VDC
CL1, CL2—alligator clip
D1 thru D8—1N914 diode
IC1—4024BE CMOS ripple divider
J1, J2—phono jack
Q1, Q2—2N3904 NPN transistor
R1, R4, R5, R10—100K-ohm ½-watt 10% resistor
R2, R3—1.5-Megohm ½ watt 10% resistor
R6, R9—68K-ohm, ½-watt 10% resistor
R7, R8—33K-ohm, ½-watt 10% resistor
R11, R12—47K-ohm, ½-watt 10% resistor
R13—3300-ohm, ½-watt 10% resistor
R14, R15—250K linear-taper potentiometer

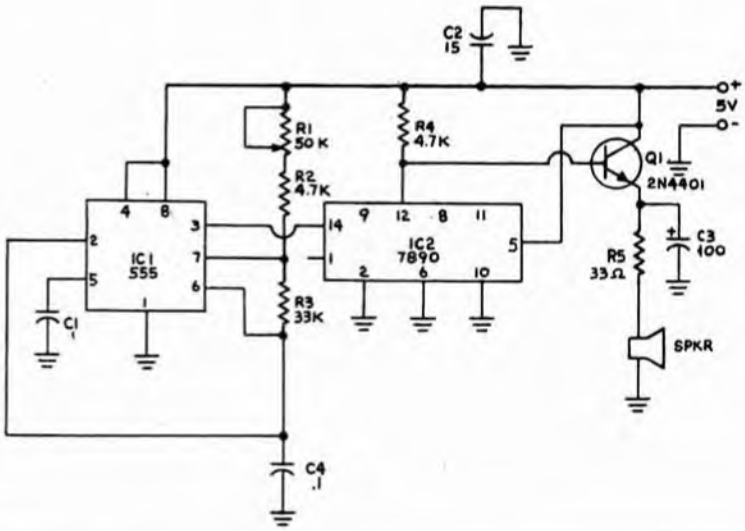


GUITAR TUNER

By taking advantage of the frequency stability of the 555 timer IC operating in an astable mode, an oscillator can be constructed which can be used as a tuning aid for the guitar. The first string of the guitar, E, produces a note with a frequency of 82.4 Hertz. The frequency of the oscillator is set to twice this value, 164.8 Hertz, and then followed by a divide-by-two stage to produce the desired frequency. The purpose of the divide-by-two stage is to guarantee that the waveform produced has a duty cycle of exactly 50%. This produces a note with no second harmonic distortion. The frequency of oscillation of the circuit is set by adjustment of R1, R2, and C2 also determine the frequency of oscillation but these components are fixed values and need no adjustment. The output of IC2 is fed to an emitter follower to provide current gain to drive a loudspeaker. C3 acts as a low-pass natural sounding note. The circuit is powered by a 5-volt supply, and this voltage **must** fall within the range of 4.75 to 5.25 volts for IC2 to operate properly.

PARTS LIST FOR GUITAR TUNER

C1, C4—0.1-uF ceramic capacitor, 15-WVDC
C2—15-uF electrolytic capacitor, 15-WVDC
C3—100-uF electrolytic capacitor, 15-WVDC
IC1—555 timer
IC2—7490 decade counter
Q1—2N4401
R1—50,000-ohm linear-taper potentiometer
R2, R4—4,700-ohm, ½-watt 10% resistor
R3—33,000-ohm, ½-watt 10% resistor
R5—33-ohm, ½-watt 10% resistor
SPKR—8-ohm PM type speaker

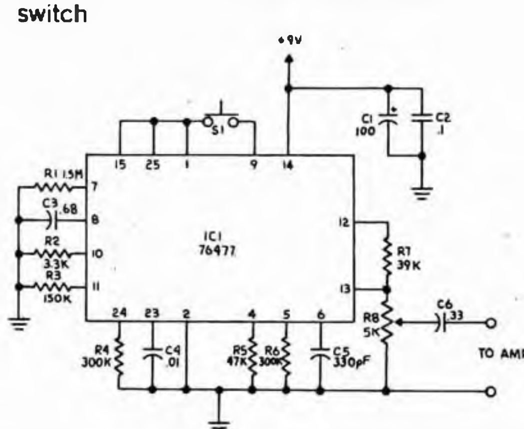


KABOOM CHIP

No, IC1 does not disintegrate in a fiery blast when S1 is pressed, but it does feed a mighty impressive burst of explosion-like noise to your amplifier. The more powerful your amplifier is, the more realistic the effect becomes. Just be sure that your speaker can handle the power safely. Maximum output from this circuit is about one volt peak-to-peak, which you can feed to the high-level input of any amp. One final note of caution: Don't overdo it, or you may find your home surrounded by the local SWAT team.

PARTS LIST FOR KABOOM CHIP

C1—100-uF, 25-WVDC electrolytic capacitor
C2—.1uF ceramic disc capacitor
C3—.68-uF mylar capacitor
C4—.01 uF mylar capacitor
C5—330-pF. polystyrene capacitor
C6—.33 uF mylar capacitor
IC1—SN76477 sound-effect generator
R1—1.5 Megohm-ohm, ½-watt 10% resistor
R2—3,300-ohm, ½-watt 10% resistor
R3—150,000-ohm, ½-watt 10% resistor
R4, R6—300,000-ohm, ½-watt 10% resistor
R5—47,000-ohm, ½-watt 10% resistor
R7—39,000-ohm, ½-watt 10% resistor
R8—5,000-ohm audio-taper potentiometer
S1—SPST normally open pushbutton switch



Now NRI takes you inside the new TRS-80 Model 4 microcomputer with disk drive to train you at home as the new breed of computer specialist!

NRI teams up with Radio Shack advanced technology to teach you how to use, program and service state-of-the-art microcomputers...

It's no longer enough to be just a programmer or a technician. With microcomputers moving into the fabric of our lives (over 1 million of the TRS-80™ alone have been sold), interdisciplinary skills are demanded. And NRI can prepare you with the first course of its kind, covering the complete world of the microcomputer.

Learn At Home In Your Spare Time

With NRI training, the programmer gains practical knowledge of hardware, enabling him to design simpler, more effective programs. And, with advanced programming skills, the technician can test and debug systems quickly and easily.

Only NRI gives you both kinds of training with the convenience of learning at home. No classroom pressures, no night school, no gasoline wasted. You learn at your convenience, at your own pace. Yet you're always backed by the NRI staff and your instructor, answering questions and giving you guidance.

Now training includes either the TRS-80 Model 4 Microcomputer with Disk Drive or TRS-80 Color Computer with Computer Access Card; professional LCD multimeter; the NRI Discovery Lab; and hundreds of demonstrations and experiments.

(TRS-80 is a trademark of the Radio Shack division of Tandy Corp.)

TRS-80 Model 4 Plus Disk Drive To Learn On and Keep

NRI gives you hands-on training on the latest model in the most popular line of Radio Shack microcomputers: the TRS-80 Model 4.

You not only learn to program your Model 4, you learn all about it... how circuits interact... interface with other systems... gain a real insight into its nature. Under NRI's expert guidance, you even install a disk drive, verifying its operation at each step.

You also work with a professional 4-function digital multimeter and the NRI Discovery Lab,® performing over 60 separate experiments.

The powerful, versatile TRS-80 Model 4 microcomputer disk drive and other equipment you train with are yours to use and keep!

Same Training Available With Color Computer

NRI offers you the opportunity to train with the TRS-80 Color Computer as an alternative to the Model 4. The same technique for getting inside is enhanced by using the new NRI-developed Computer Access Card. Only NRI offers you a choice to fit your specific training needs.

The Catalog is Free. The Training is Priceless.

Get all the details on this exciting course in NRI's free, 104 page catalog. It shows all equipment, lesson outlines, and facts on other electronics courses such as Electronic Design, Industrial Electronics, Video/Audio Servicing... 12 different career opportunities in all. Keep up with the latest technology as you learn on the latest model of the world's most popular computer. If card has been used, write to:

NRI McGraw-Hill

NRI Schools
McGraw-Hill Continuing Education Center
330 Progress Avenue
Scarborough, Ontario M1P 2Z5
or telephone 416-293-1911

We'll give you tomorrow.