Proceedings of

rf technology expo 85

Disneyland Hotel, Anaheim, California
January 23-25, 1985

Return To:
Alan Victor
ET 5K2
6-11-85.

Sponsored by *RF Design* magazine

rf design
6530 South Yosemite Street / Englewood, Colorado 80111 (303) 694-1522

June 5, 1985

Dear RF TECHNOLOGY EXPO 85 Speaker:

Enclosed is your complimentary copy of the Proceedings from RF TECHNOLOGY EXPO 85. This 600-page book contains copies of the 60 papers that were presented at the EXPO in January (yours included). Printing the book turned out to be a time-consuming labor of love; but I'm sure you'll agree that it has been worth the wait.

I would like to thank you again for helping us successfully launch the first RF TECH EXPO. The response to the show has been overwhelming. Exhibitors are still commenting on the quantity and quality of the attendees; and the attendees keep telling us how invaluable the sessions were. The overall feeling seems to be--let's get together again next year!

And so we will. RF TECHNOLOGY EXPO 86 will be held January 30-February 1 at the Anaheim Hilton and Towers (just blocks from the Disneyland Hotel). 29,000 square feet of exhibit space has been set up to accommodate 150 ten-foot booths. This is twice the number of booths used at EXPO 85, but with 60 booths reserved to date, we're confident that we'll have a "full house" next year.

I will be serving as program chairman for EXPO 86. If we haven't already talked about your participation in the show, please contact me with your paper proposal (note my "Call for Papers" in the April and May issues of RF Design). Paper selections will be finalized early for EXPO 86 so that the Proceedings can be ready for distribution at the show. Proposals must be received by July 26, 1985. Speakers and papers will be announced by August 30. Speakers at EXPO 86 will receive free conference registration and a copy of the show Proceedings.

I hope I can count on your support at RF TECHNOLOGY EXPO 86. We are indebted to you for your participation at EXPO 85. Thanks again. See you at the show!

Sincerely,

James N. MacDonald

James N. MacDonald
Editor

Proceedings of rf technology expo 85

Disneyland Hotel, Anaheim, California
January 23-25, 1985

Components

Circuits

Receivers

Transmitters

Computer-Aided Design

Techniques

EMC/EMI

Fundamentals of RF Design

RF - EXPO PAPER

Hybrid Varactor-Tuned Oscillator Modules
Their Practical Applications in RF Communications

by Ron Patston,
Applications Engineer,
Avantek, Inc.
Santa Clara, CA

Introduction

Varactor-tuned oscillators (VTOs) have been used in many types of radio frequency systems over the years. Until the advent of thin film hybrid technology and its use in design at radio frequencies, however, varactor tuned oscillators, produced with conventional "discrete component" construction, have tended to be bulky and cumbersome for most system applications. Building VTOs with discrete technology also created other problems for the designer, including non-linearity, restriction to narrow bandwidths, unreliability, non-repeatability, instability over temperature, and the labor costs involved in building and tuning.

On the other hand, thin film hybrid VTOs offer wider bandwidths, reproducibility, high reliability, smaller size, lower cost, low power consumption, and extreme ease in incorporating them into new designs or retrofitting into existing designs.

What are They?

The thin film oscillators manufactured by Avantek have many features that are useful to RF system designers. The VTO-8000 series oscillators, for example, covers RF frequencies as low as 300 MHz and up to 11 GHz. They are suitable for operation over either narrow or wide frequency bands.

These oscillators are designed using a varactor diode as a voltage controlled capacitor in a thin film microstrip resonator to control the frequency of a negative resistance transistor oscillator. With thin film construction, it naturally follows that the size of these oscillators is very small in comparison with their discrete-component counterparts.

All of the VTO-8000 series oscillators are supplied in the TO-8 type package: a small, hermetic package that enhances the reliability of the product.

The basic VTO circuit is fabricated using a silicon abrupt varactor diode -- which will produce voltage-vs. frequency tuning curves that are relatively non-linear, but which are quite smooth and monotonic. As an option, it may be equipped with a silicon hyper-abrupt varactor diode -- which will provide a relatively linear tuning response across a wider bandwidth (i.e. Figure 10). These oscillators also feature good noise performance (Figure 11) plus extremely fast tuning speeds and low post-tuning drift.

Application Techniques

There are many different parameters involved in selecting and properly applying varactor tuned oscillators. System designs will generally be based upon a specific subset of the oscillator's specifications whether it be low noise, linearity, fast tuning speed, temperature stability, etc., or a combination of several of these. The primary application techniques that will be covered here are; temperature compensation, linearization, and varactor-tuned oscillator specifications as they pertain to oscillators used in phase-locked systems.

Temperature Compensation and Stabilization

The reduction of oscillator frequency change with changes in temperature may be carried out using one or more of these three basic techniques; control of the oscillator temperature, tuning voltage temperature compensation, or the use of a phase locked loop.

The temperature of a TO-packaged component is easily controlled by either a very small, low-power heater or by placing the component in a temperature-controlled chamber (oven).

There are commercially-available DC proportionally-controlled heater assemblies specifically designed for use on TO-8 type cans, or it is a relatively simple matter to fabricate a heater by mounting the component on a block (which provides thermal mass), that is temperature controlled using a proportional heater
(See Figure 9).

A self-controlling heater may be used, which employs a material with a definitive temperature vs. resistance characteristic. This material may be directly epoxied to the top of the TO-8 oscillator and then supplied with a bias voltage. The temperature vs. resistance characteristic of the material will cause it to act as a temperature controlled heater that will provide very good temperature stability at a very low cost.

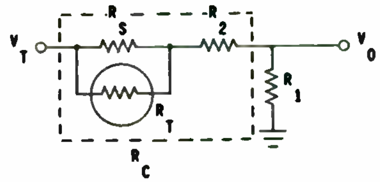
An important point to always keep in mind when using any heater approach for temperature compensation is that the temperature of the oscillator must be kept 5 to 10 degrees above the maximum expected system operating temperature. This will ensure that the oscillator will not be affected by the external temperature changes.

The primary drawback to using the heater approach is the extra power required to keep the oscillator at a higher-than-ambient temperature. The advantage is the high degree of temperature stability for a relatively low cost.

The effect of temperature on the oscillator frequency may also be reduced indirectly, by varying the tuning voltage in the proper direction to bring the oscillator back to the correct frequency. Temperature compensation of the tuning network may be done using either a Positive Temperature Coefficient (PTC) or Negative Temperature Coefficient (NTC) thermistor(or network of thermistors), depending upon the actual tuning circuitry used.

Typically, VTO-8000 Series oscillators will display a negative frequency vs. temperature drift coefficient. To compensate for this, a voltage compensation network must be put in place to offset the frequency drift. There are two simple networks which may be used: one employing a PTC thermistor, the other a NTC thermistor.

The first network is that using a PTC thermistor.

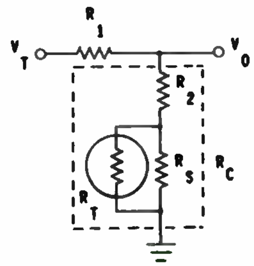


PTC NETWORK

$$V_o=(V_tR_1)/[R_1+R_2+(R_S||R_T)]$$

(Figure 1.)

The second network is that using a NTC thermistor.



NTC NETWORK

$$V_o=V_t(R_2+R_s||R_T)/[R_1+R_2+(R_s||R_T)]$$

$$R_T=R_{25}(1+A)^{(T-25)}$$

Where:

T = Temperature in C°
A = Temperature Coeff %/C°

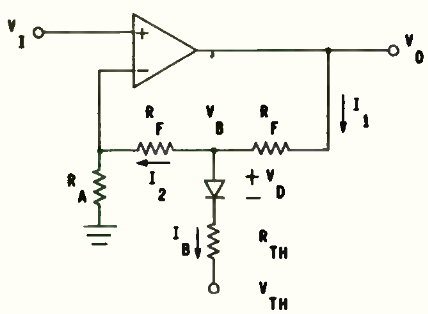
(Figure 2.)

Other types of resistance networks may be used in place of R_c depending upon personal preference. A suggestion for determining the best compensation network for the application is to hold the value of R_1 fixed and use a curve fitting routine to find the values of R_2 and R_s when the desired R_c is known for at least three different temperatures. When temperature compensation of the tuning voltage is used, the temperature sensing device should be mounted as close as possible to the oscillator itself. This will provide the shortest thermal time constant possible from the sensing device to the compensated oscillator.

Linearization

Linearizer circuits of various types have been used for years to improve the voltage-vs.-frequency curves of varactor tuned oscillators. In fact, a properly-designed and "tweaked" linearizer can provide virtually any degree of linearity required for a particular application. Linearizer circuits may also incorporate the additional function of shifting the tuning voltage provided by the system to one more appropriate for the oscillator itself.

There are typically two types of linearization schemes employed today: analog and digital. The use of an analog linearizer is desirable when the oscillator interfaces with an analog tuning voltage, or when a linear modulation spectrum at any point in the frequency range is desired. A simple analog linearizer circuit is shown in Figure 3. The primary application for a digital linearizer is in applications where the oscillator is to be tuned by a digital computer.



ANALOG LINEARIZER

(Figure 3.)

$$V_o = V_i + (I_2 R_f) + (I_1 R_f)$$

$$I_2 = V_i / R_A \quad \& \quad I_1 = I_b + I_2$$

$$I_b = (V_b - V_d - V_{TH}) / R_{TH}$$

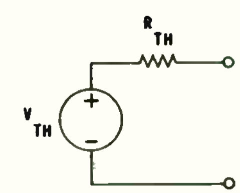
Also: $I_b = 0$ if $(V_d + V_{TH}) \leq V_b$

Therefore: $V_o = V_i(1 + 2R_f/R_A) + I_b R_A$

From this it is easily seen that V_{TH} will determine where the increase in slope will occur and R_{TH} will determine the amount of the increase in the tuning slope.

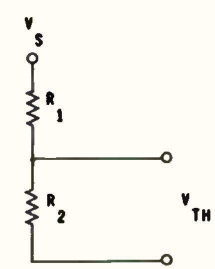
To replace R_{TH} and V_{TH} with a simple resistive divider use Thevenin's Theorum.

A circuit such as this:



(Figure 4.)

May be replaced by this:



(Figure 5.)

with some simple calculations:

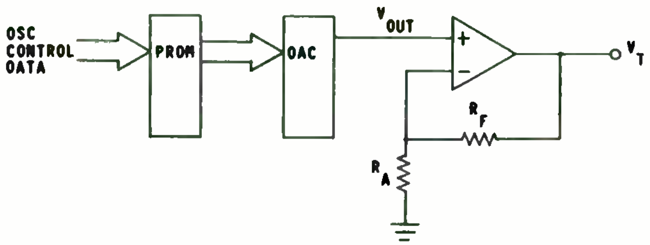
$$V=V_s/V_{TH}$$

$$R_1=R_{TH}/V$$

$$R_2=R_{TH}/(1-V)$$

In using this type of circuit one would have the capability of introducing almost any number of changes to the slope of the tuning curve, which may all be implemented in parallel depending upon the degree of linearity required. This circuit will also provide good modulation response which will only be restricted by the frequency response of the op amp itself.

One of the most efficient linearization techniques combines an analog-to-digital converter with a PROM and an op amp.



(Figure 6.)

Using this configuration and a small computer, the PROM may be programmed to provide linearity better than 0.5% across the full frequency spectrum of the VCO. The circuit will also provide extremely fast tuning response time, primarily limited by the settling time of the Op Amp.

Phase Lock

Phase-locked-loops using VCOs are becoming much more common than in the past, due to the improvements and greater availability of divider techniques and crystal multiplied sources. Some of the more important requirements for an oscillator to be suitable for a phase locked application are:

1) Phase stability (spectral purity)
2) Large electrical tuning range
3) Linearity of frequency vs control voltage
4) (frequently), the capability of accepting wideband modulation

The major concern of synthesizer designers is the phase stability or, as it is commonly termed, "Phase Noise". The phase noise generated by a VCO is primarily determined by; 1) The circuit Q (quality factor) and 2) The Q of the varactor diode. Phase noise of an oscillator will also be improved by the use of a silicon bipolar transistor rather than a gallium arsenide FET for the oscillator transistor. The VTO-8000 series features the use of silicon bipolar transistors exclusively. This means that their phase noise performance will primarily be based upon circuit Q and Diode Q.

The oscillator circuit itself is usually designed with a specific parameter in mind. In order to design a circuit with a very high Q the tuning bandwidth must invariably suffer. Therefore, in order to design an oscillator circuit for optimum phase noise performance it will ultimately end up being a fairly narrow band oscillator. With this in mind the VTO-8000 series oscillators are easily modified at the factory to provide narrow-band low noise performance.

The other governing parameter for good phase noise performance is the Q of the tuning varactor. The tuning varactor Q is primarily dependent on which of the two types of tuning varactors are used: the abrupt-tuning, or the hyperabrupt-tuning varactor.

The abrupt tuning diode will provide a very high Q along with a continuous monotonic tuning curve and it will also operate over a very large range of tuning voltages (0-50v). As a result of its very high Q, the abrupt diode offers the best-available phase noise performance. Both silicon-abrupt and GaAs-abrupt diodes are available, and both are used in VCOs. Although the GaAs abrupt diode will exhibit a higher Q (see Figure 12) than the silicon abrupt diode the phase noise performance of the oscillator will be poorer.

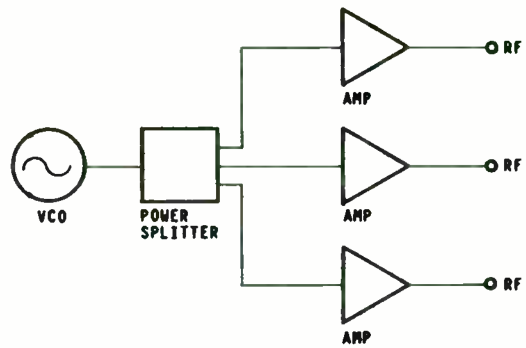
The other type of diode used extensively in the design of VCO's is the hyper-abrupt diode. The hyper-abrupt diode will provide a much more linear tuning response than the abrupt due to its linear voltage vs capacitance characteristics, this also gives the capability to cover a wider frequency range in a smaller tuning voltage (0-20v). (see Figure 10). The drawback to using the hyper-abrupt tuning diode is that it has a much lower Q than an abrupt tuning diode. From observing Figure 12, it becomes obvious that in order to achieve the maximum diode Q for a low-noise oscillator, the oscillator should be tuned at the highest possible voltage (without exceeding the breakdown voltage of the diode).

As is seen in Figure 11 the phase noise difference between using the si-abrupt and the si-hyper-abrupt diode is typically about 3 dB. The si-abrupt outperforms the si-hyper-abrupt in phase noise characteristics due to the higher Q of the si-abrupt as mentioned earlier. The general theory is that the noise performance of the GaAs varactor is degraded (even though it has a higher Q) because of the surface currents created on the diode.

From the information supplied thus far it is recommended that if a very low noise oscillator is required then the best performance will be obtained when the bandwidth is kept as low as possible (<20%) and the tuning voltage is kept high as possible (>10v). This will provide the oscillator with the criteria to obtain the optimum in low noise performance.

Integration With TO-8 Amplifiers

For applications requiring a higher power level than is available directly from the oscillator, the TO-8-packaged varactor-tuned oscillator is easily combined with readily-available TO-8 hybrid amplifier modules. Integrating such oscillators and amplifiers is a simple exercise in stripline design in a 50 ohm system. Using the TO-8 oscillator with a number of TO-8 amplifiers makes it easy to develop a single- or multiple- output system with +10 dBm output power. These types of applications are briefly outlined below.



(Figure 7.)

Applications

The modular varactor-tuned oscillator has many applications in frequency-agile systems such as digitally-controlled receivers and active jamming transmitters. In such equipment, the oscillator is usually combined with an external linearizer similar to those mentioned earlier.

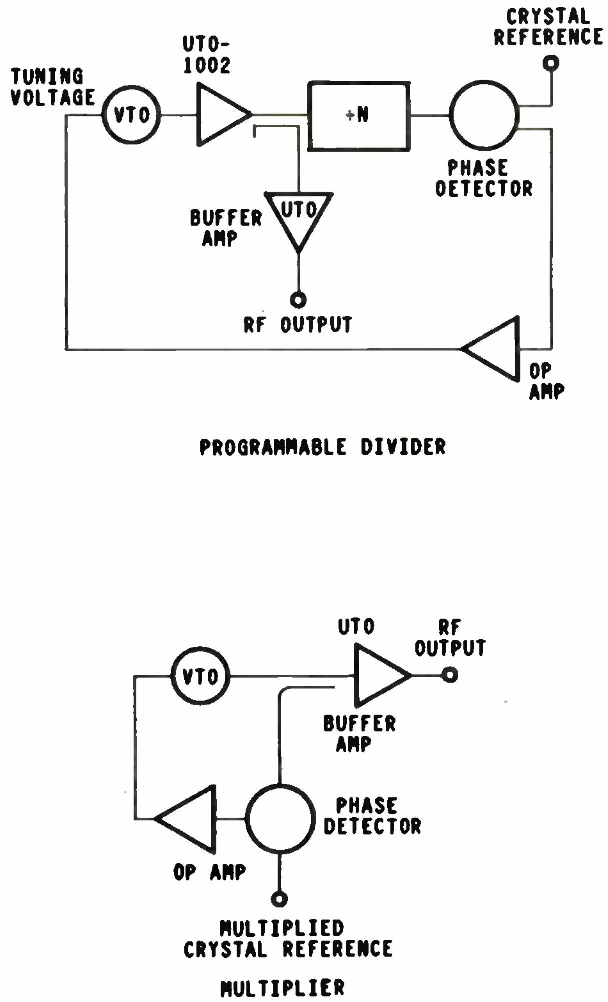
The VTO Series oscillator is also an ideal local oscillator for use in satellite earth station downconversion systems due to its small size, high reliability and the availability of the oscillators in high volume at a very low cost.

The VTO Series oscillators have been designed with a tuning bypass capacitance which is sufficient to provide the necessary RF filtering action, yet as low as possible to maximize dV/dT characteristics.

Used in a phase-locked loop circuit (Figure 8.), a VTO provides a receiver LO with stability equivalent to the reference oscillator (usually crystal controlled), yet variable in discrete steps or continuously depending on the PLL configuration. An important feature of the VTOs used in an LO application is their power vs. frequency flatness (±1.5 dB). This assures that once a receiver mixer is biased for best dynamic range that the local oscillator drive will remain constant throughout the tuning range without complex leveling circuitry.

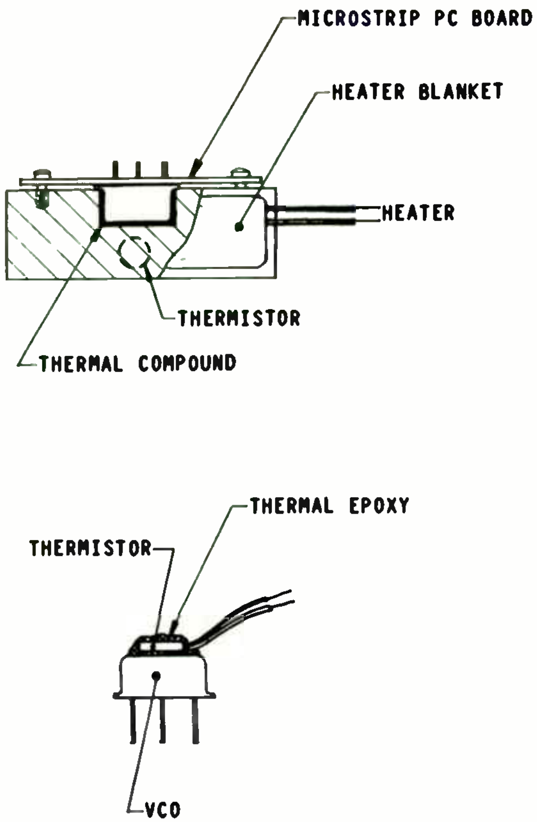
These oscillators are excellent candidates for the next generation of portable test equipment. Many designers have already found these oscillators ideal for use in frequency synthesizers, spectrum analyzers, sweep generators, and many other types of test equipment which require internal RF sources.

# # #



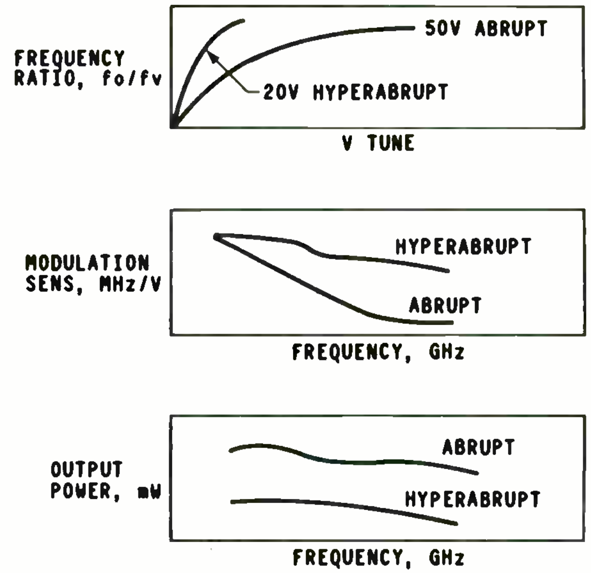
PLL

FIGURE 8



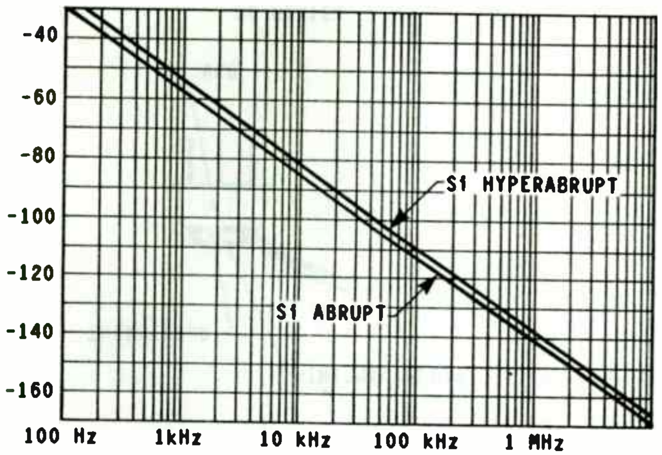
DC HEATER

FIGURE 9



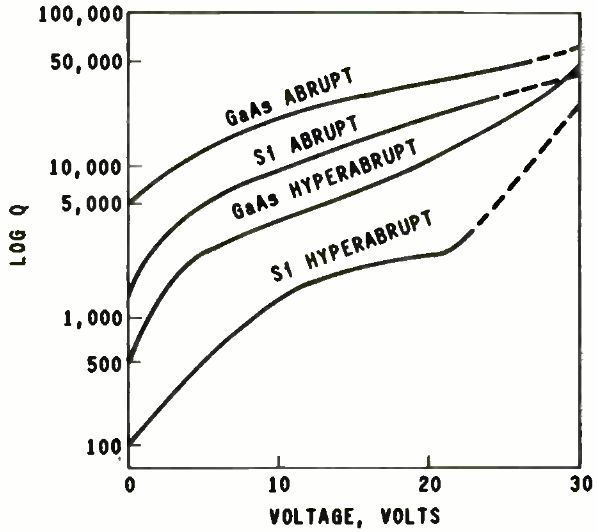
ABRUPT vs. HYPERABRUPT

FIGURE 10



PHASE NOISE

FIGURE 11



Q vs. VOLTAGE

FIGURE 12

SURFACE MOUNTED COMPONENTS

MARTIN L.BARTON
COLLINS TRANSMISSION SYSTEMS DIVISION
ROCKWELL INTERNATIONAL
DALLAS TX

INTRODUCTION

STANDARDS FOR SURFACE MOUNTED COMPONENTS (SMC'S) ARE STILL VERY MUCH IN AN EMBROYONIC STAGE AND APPLICATION DATA ON THESE NEW COMPONENTS IS SPARSE. THE DESIGNER FACES THE PROBLEM OF DESIGNING WITH SMC'S OF WIDELY VARYING CHARACTERISTICS.

THIS PAPER ADDRESSES THESE DIFFERENCES AND PROVIDES SOME DESIGN CONSIDERATIONS IN THE USE OF SUCH COMPONENTS.

CHIP COMPONENTS

CHIP COMPONENTS ARE PASSIVE, LEADLESS COMPONENTS AND INCLUDES RESISTORS, CAPAPCITORS AND INDUCTORS. THESE COMPONENTS BEHAVE BASICALLY LIKE THEIR STANDARD-SIZED COUNTERPARTS BUT ARE LIMITED IN SOME PERFORMANCE CHARACTERISITIC COMPARED TO TRADITIONAL LEADED COMPONENTS:

1) RANGE OF VALUES
2) BREAKDOWN VOLTAGE
3) POWER DISSIPATION

CHIP COMPONENTS OFFER CONSIDERABLE ADVANTAGES OVER LEADED COMPONENTS:

1) HIGHER RELIABILITY (ELIMINATION OF LEADS)
2) IMPROVED HF CHARACTERISTIC (LESS LEAD INDUCTANCE)
3) REDUCED PROPAGATION DELAYS (HIGHER DIGITAL SPEEDS)
4) LESS POWER LINE NOISE (LESS LEAD INDUCTANCE)
5) LOWER EMI (SMALLER SIZE)
6) SPACE SAVING (CAN BE MOUNTED ON BOTH SIDES OF PCB)
7) LOWER COSTS (NO LEADS TO TRIM OR FORM, NO HOLES TO DRILL & SMALLER PCB)
9) REDUCED SHIPMENT AND STORAGE COSTS
10) MORE STANDARDIZED COMPONENT SIZES

THERE ARE ALSO SOME DISADVANTAGES TO USING SMC'S:

1) LACK OF WORLD-WIDE INDUSTRY STANDARDS
2) IDENTIFICATION IS DIFFICULT (NO MARKINGS)
3) PRESENT HIGHER COST OF SEMICONDUCTORS

RESISTORS

THE PACKAGE CONFIGURATION FOR RESISTORS HAS BEEN STANDARDIZED INTO 3 BASIC STYLES:

1) FLAT RECTANGULAR "CHIP" DOUBLE-SIDED METALIZATION
2) FLAT RECTANGULAR "CHIP" SINGLE-SIDED METALIZATION
3) CYLINDRICAL "MELF" (METAL ELECTRODE FACE BONDING)

STANDARDIZATION IN SIZE, THOUGH NOT YET COMPLETE, NOW EXISTS. THE MOST COMMONLY USED SIZE IS 3.2 X 1.6 MM SIZE (0.125W). MELF TYPES ARE 6 X 2.2 MM OR 3.5 X 1.1 MM SIZE (0.25 AND 0.125W). TOLERANCES ARE 10,5,2 AND 1% WITH 5 AND 1% BEING MOST COMMON. THE TIGHTER TOLERANCES ARE OBTAINED BY ACTIVE TRIMMING TO VALUE. TYPICAL TEMPERATURE COEFFICIENT RATIOS (TCR) RANGE FROM 300 TO 100 PPM. FOR BRIDGING PRINTED CIRCUIT BOARD TRACES, I.E. CROSSOVERS, ZERO-OHM RESISTORS (LESS THAN 50 MILLI-OHMS RESISTANCE) ARE AVAILABLE.

MELF RESISTORS ARE MADE BY CROPPING THE LEADS OF AXIAL-LEADED CARBON RESISTORS AND METALIZING THE END CAPS FOR TERMINATIONS. THESE RESISTORS ARE SLIGHTLY LOWER IN COST THAN CHIP RESISTORS. COMPANIES WITHOUT THICK FILM EXPERTISE FAVOR THIS TYPE OF CONSTRUCTION. THE MAIN DRAWBACK OF THE MELF IS THE SPIRAL RESISTIVE TRACK WHICH IS INDUCTIVE AND THEREFORE LIMITS ITS USE AT HF. BELOW 500 OHMS VALUE, THE RESISTOR IS INDUCTIVE (10-15 nH) AND ABOVE 500 OHMS THE RESISTOR LOOKS CAPACITIVE (0.2 pF). IN THE US AVAILABILITY OF MELF'S IS LIMITED TO A SINGLE SUPPLIER.(TRW).

CHIP RESISTORS HAVE GOOD RF CHARACTERISTICS TO 500 MHZ (APPROX. 0.25 pF SHUNT CAPACITY) AND ARE ALSO MORE COMPATIBLE WITH AUTOMATIC PLACEMENT EQUIPMENT. HOWEVER, THEIR LACK OF MARKINGS FOR ELECTRICAL VALUE CODING NECESSITATES A SYSTEM OF COMPONENT HANDLING AND CONTROL TO AVOID MIXING OF THESE PARTS. (IF PARTS TO GET MIXED, THEY MUST BE POSITIVELY IDENTIFIED OR DISCARDED. IN MOST CASES, DISCARDING IS A LOWER COST ALTERNATIVE.)

THE DOUBLE-SIDED CHIP (MOUNTED THICK-FILM SIDE UP) AND MELF RESISTORS ARE INTENDED TO BE USED WITH A WAVE SOLDERING PROCESS. WHEN MOUNTED ON THE BACKSIDE OF A PCB THEY REQUIRE TO BE ADHESIVELY ATTACHED PRIOR TO SOLDERING WITH REFLOW SOLDERING (VAPORPHASE OR IR) IS USED THESE RESISTORS HAVE A TENDENCY TO SHIFT OFF THE SOLDER PADS (KNOWN AS "DRAWBRIDGING" OR "TOMBSTONING") DURING THE REFLOW PROCESS. THE PROBLEM IS PRIMARILY CAUSED BY THE POOR QUALITY OF THE TERMINATIONS. FOR GOOD SOLDERING YIELDS IT IS ESSENTIAL THAT THE TERMINATIONS BE CLEAN AND THERE BE A MINIMUM OF 10 MILS METALIZATION (MIL VERSION CALLS FOR THIS REQUIREMENT). ALTERNATIVELY THE RESISTORS MUST BE ADHESIVELY ATTACHED PRIOR TO REFLOW SOLDERING.

THE SINGLE-SIDED RESISTOR IS MOUNTED WITH THE THICK-FILM SIDE DOWN AND IS LESS SUSCEPTIBLE TO MOVEMENT DURING REFLOW. IT ALSO ALLOWS A HIGHER PACKAGING DENSITY DUE TO THE SMALLER FOOTPRINT. ITS MAIN DISADVANTAGE ARE THE TOTALLY HIDDEN AND UNACCESIBLE TERMINATIONS MAKING INSPECTION OF SOLDER JOINTS AND TEST ACCESS IMPRACTICAL. CURRENTLY THE ONLY MANUFACTURER OF THIS STYLE RESISTOR IS PANASONIC (JAPAN) AND IT IS NOT AVAILABLE FOR SALE. (DALE ELECTRONICS HAS A SINGLE-SIDED RESISTOR CHIP UNDER DEVELOPMENT.)

RESISTOR NETWORKS ARE PACKAGED IN 16 PIN SOIC CONFIGURATION AS

WELL AS IN FLAT PACKS. TCR'S ARE AS LOW AS 25 PPM WITH 2% TOLERANCE. THERMISTORS WITH NEGATIVE TCR'S AND 5/10% TOLERANCES ARE AVAILABLE.

VARIABLE RESISTORS - SINGLE AND MULTI-TURN, HORIZONTALLY AND VERTICALLY MOUNTING STYLES - ARE ALSO AVAILABLE. POWER DISSIPATION RANGES FROM 50mW TO 0.5W AND TCR'S ARE TYPICALLY 100 TO 250 PPM.

CERAMIC CAPACITORS

CERAMIC CAPACITORS HAVE ACHIEVED WORLD-WIDE STANDARDIZATION AND ARE AVAILABLE IN THE INTERNATIONAL 3.2 X 1.6 MM RECTANGULAR SIZE FOR VALUES FROM 1pF TO 0.1uF. FOR HIGH DENSITY DESIGNS A SMALLER VERSION 2.0 X 1.3 MM IS USED. FOR BY-PASS AND COUPLING APPLICATIONS DESIGNERS ARE ENCOURAGED TO USE THE PREFERRED VALUES OF 100pF, .001, .01, AND .1uF.THESE VALUES HAVE HI-VOLUME USEAGE AND ARE THEREFORE LOWEST IN COST AND MORE READILY AVAILABLE.

THERE ARE 4 CLASSES OF CERAMIC CAPAPCITORS:

NPO	1 TO 1000pF	1 TO 20% TOL
X7R	.001 TO .1uF	5 TO 20% TOL
Y5V	.01 TO .1uF	20 TO -80% TOL
Z5U	.01 TO .22uF	

CAPACITY AND DISSIPATION FACTOR DROP WITH INCREASE OF FREQUENCY. FOR RF APPLICATIONS NPO SHOULD BE USED DUE ITS LOW DISSIPATION FACTOR AND GOOD TCR (30 PPM). X7R IS PRIMARILY USED FOR CRITICAL BY-PASS APPLICATIONS, WHILE Y5V FOR LOWEST COST. THE MORE COMMON Z5U IS NOT SUITABLE FOR OPERATION BELOW +10 DEGREES DUE TO THE LARGE DROP IN CAPACITY (ONLY 25% OF 25 DEGREE CENTIGRADE VALUE.)

LARGE VALUES OF CAPACITANCE ARE PRONE TO MICROCRACKS WITH THERMAL SHOCK. IT IS ESSENTIAL TO SUBJECT CERAMIC CAPACITORS TO A PRE-HEAT AND POST-COOL CYCLE DURING THE SOLDERING PROCESS. MEASUREMENT OF INSULATION RESISTANCE AT 85 DEGREES C. IS A GOOD SCREENING METHOD FOR DETECTING VOIDS. BURN-IN FOR 48 HOURS AT 85 DEGREES C AND TWICE THE DC RATED VOLTAGE IS ALSO EFFECTIVE. A VOLTAGE DESTRUCT TEST (TYPICALLY 5 TIMES YHE DC RATING) IS A GOOD MEASURE OF THE QUALITY OF THE DIELECTRIC LAYERS.

A RELATIVELY UNKNOWN FACTOR IS THAT CERAMIC CAPACITORS ARE MICROPHONIC AND ARE NOT SUITABLE FOR USE IN VIBRATORY ENVIRONMNENTS. (TANTALUM OR FILM CAPACITORS SHOULD BE USED INSTEAD.)

ABOVE 100 MHz OPERATION PORCELAIN CERAMIC IS USED TO REDUCE LOSSES AND IMPROVE THE DISSIPATION FACTOR. ACHIEVABLE Q's RANGE FROM 200 TO 2000 DEPENDING UPON CAPACITANCE VALUE. GENERALLY USED IN RF POWER CIRCUITS. DUE TO THEIR LIMITED USE THESE CAPACITORS ARE COSTLY.

TANTALUM CAPACITORS

TANTALUMS ARE USED FOR APPLICATIONS REQUIRING CAPACITANCE VALUES ABOVE 0.1uF. DEPENDING ON CAPACITANCE VALUE AND VOLTAGE RATING, SEVERAL DIFFERENT SIZES ARE AVAILABLE. AT PRESENT THERE NO INDUSTRY STANDARDS FOR TANTALUM CAPACITOR SIZES MAKING INTERCHANGEABILITY A PROBLEM. SIZES RANGE FROM 3.2 X 1.6 MM TO 7.3 X 4.3 MM FOR MOLDED TYPES AND 3.4 X 1.6 MM TO 8.3 X 4.2 FOR MELF STYLE. VALUES RANGE FROM 0.1 TO 100 uF, 4 TO 50V AND 5 TO 20% TOLERANCES.

TANTALUMS ARE POLARIZED AND ARE SUSCEPTIBLE TO FAILURE WHEN EXPOSED TO REVERSE VOLTAGE. FOR NON-POLAR APPLICATIONS 2 CAPACITORS CAN BE CONNECTED IN SERIES "BACK-TO-BACK". LEAKAGE CURRENT TYPICALLY INCREASES TEN-FOLD FROM 25 TO 85 DEGREES C. GOOD QUALITY CAPACITORS HAVE LEAKAGE CURRENTS OF LESS THAN 1 uA AT 85 DEGREES C. THE LEAKAGE CURRENT CAN BE REDUCED AND STABILIZED BY BURN-IN.

ALUMINUM CAPACITORS

THESE CAPACITORS ARE USED IN LOW-COST (CONSUMER PRODUCT) APPLICATIONS. THEIR LIFE IS LIMITED DUE EVAPORATION OF THE ELECTROLYTE WITH TIME. VALUES AND VOLTAGE RATINGS ARE SIMILAR TO TANTALUM CAPACITORS. AVAILABLE SIZE IS CYLINDRICAL 6.3 X 5 MM. THERE ARE CURRENTLY NO US SOURCES.

METALIZED FILM CAPACITORS

METALIZED FILM CAPACITORS HAVE SUPERIOR CHARACTERISTICS TO CERAMIC. THERE IS NO PIEZO-ELECTRIC EFFECT AND NO FREQUENCY VERSUS VOLTAGE SENSITIVITY. THEY ARE ALSO SELF-HEALING. TYPICAL SIZE IS 5.7 X 5.7 MM. VALUES RANGE FROM .01 TO 1.0uF AND 30 TO 100 VDC.

VARIABLE CAPACITORS

VARIABLE CAPACITORS ARE AVAILABLE FROM 5 TO 30 pF IN NPO MATERIAL WITH Q MIN OF 500 AT 100 MHz. TCR's RANGE FROM 200 TO 500 PPM AND SOME HAVE NEGATIVE CHARACTERISTICS. SELF-RESONANCE OCCURS ABOVE 1 GHz. MOISTUREPROOF TYPES ARE A SO AVAILABLE. THESE HAVE THE ADVANTAGE THAT THEY ARE COMPATIBLE WITH NORMAL CLEANING PROCESSES USED IN REFLOW SOLDERING. SIZES ARE TYPICALLY 4.5 X 4.0 MM OR 3.5 MM DIAMETER.

INDUCTORS

CHIP INDUCTORS ARE WOUND ON FERRITE OR CERAMIC CORES AND THE LEADS ARE SOLDERED TO METALIZED TERMINATIONS. CONSTRUCTION RESEMBLES THAT OF STANDARD-SIZE COMPONENT. INDUCTANCE VALUES RANGE FROM 10 nH TO 10 mH WITH 10 TO 20% TOLERANCES. MAXIMUM CURRENTS ARE A FUNCTION OF WIRE SIZE AND VARY FROM 25 TO 150 mA. Q's RANGE FROM 25 TO 60 AT VHF/UHF. TCR's ARE -110 TO 300 PPM. TYPICAL SIZES ARE 2.5 X 2.0 TO 5.0 X 4.0 MM. AT PRESENT NO INDUSTRY STANDARDS EXIST FOR CHIP INDUCTORS AND INTERCHANGEABILITY IS A PROBLEM.

CURRENT DESIGNS ARE RELATIVELY FRAGILE DUE TO THE FINE WIRE SIZE USED AND THE POOR ADHESION OF THE METALIZATION TO FERRITE CORES. (CERAMIC CORES ARE MUCH BETTER IN THIS RESPECT.) CHIP INDUCTORS

ARE ALSO PRONE TO DAMAGE FROM ULTRASONIC CLEANING.

TO MINIMIZE COUPLING BETWEEN ADJACENT INDUCTORS, THE INDUCTORS SHOULD BE MOUNTED AT RIGHT ANGLES TO EACH OTHER. VALUES BELOW 200 nH ARE DIFFICULT TO MEASURE DUE TO ERRORS INTRODUCED BY THE TEST FIXTURE LEAD INDUCTANCE. THE USE OF A CORRELATION STANDARD AND COMPARISON MEASUREMENTS WILL ELIMINATE THE MEASUREMENT UNCERTAINTIES.

WHERE POSSIBLE RESISTORS SHOULD BE USED IN LIEU OF INDUCTORS AS THESE ARE MUCH MORE COST EFFECTIVE FOR DECOUPLING PURPOSES.

VARIABLE INDUCTORS ARE BECOMING AVAILALE.INDUCTANCE VALUES RANGE FROM 0.1 TO 10 uH AND Q's ARE FROM 10 TO 30. SIZE IS APPROX. 7 X 3.5 TO 10.5 X 4.5 MM.

RF TRANSFORMERS

MINIATURE RF TRANSFORMERS ARE CONSTRUCTED WITH FERRITE CORES - TOROID AND BALUN TYPES - AND MOUNTED ON A CARRIER. THE CARRIER IS TYPICALLY CERAMIC OR GLASS-EPOXY MEASURING 4 TO 6.5 MM SQUARE. TOROIDS ARE USED FOR LOW-LOSS APPLICATIONS BUT THE BALUN CORE IS LESS SUSCEPTIBLE TO PERFORMANCE DEGRADATION WHEN CONFORMALLY COATED. IT IS ALSO BENEFICIAL TO PROTECT THE WINDINGS BY PLACING SOME HEATSHRINK TUBING OVER THE BALUN CORE AND WINDINGS. THESE TRANSFORMERS TEND TO BE LARGE AND COSTLY COMPARED TO THE NORMAL SMCs.

DIODES

ALL TYPES OF DIODES - SIGNAL, RECTIFIERS, ZENERS, PIN, SCHOTTKY, GENERAL PURPOSE - ARE CURRENTLY PACKAGED IN THE SOT-23 PLASTIC PACKAGE. POWER DISSIPATION IS LIMITED TO 0.2 W MAX. SINCE THE SOT-23 IS A 3 TERMINAL PACKAGE IT WILL HOUSE 2 DIODES. SERIES, CATHODE-TO-CATHODE AND ANODE-TO-ANODE CONNECTED CONFIGURATIONS ARE AVAILABLE. THIS NOT A COST EFFECTIVE P(ACKAGE FOR SINGLE DIODES. THE CYLINDRICAL MELF PACKAGE IS LESS COSTLY AND CAN DISSIPATE POWER UP TO 1 WATT AND IS THEREFORE MORE SUITED FOR POWER RECTIFIERS AND ZENER VOLTAGE REGULATORS. THE 0.5 W SIZE MEASURES 3.2 X 1.6 MM AND THE 1 W DEVICE XXX X XXX.

LED's USE THE DO-35 LEAD GLASS PACKAGE. THIS ALLOWS 360 DEGREE OF LIGHT TRANSMISSION. AVAILABLE COLORS ARE RED, GREEN AND YELLOW.

TRANSISTORS

SMALL SIGNAL DEVICES ARE CONTAINED IN THE SOT-23 PACKAGE. MAX DIE SIZE IS 30 X 30 MILS. POWER DISSIPATION IS LIMITED TO 0.2 W AND MAX JUNCTION TEMPERATURE IS 150 DEGREES C.(VARIES BETWEEN 125 AND 175 DEGREES C. DEPENDING ON MANUFACTURER.) AN ALTERNATIVE LO-PROFILE PACKAGE REDUCES CLEARANCE UNDER THE PACKAGE TO LESS THAN 5 MILS TO FACILITATE ADHESIVE ATTACHMENT. THE STANDARD SOT-23 HAS RAISED FORMED LEADS TO ENSURE THAT CLEANING UNDER PART IS FEASIBLE.

SOME TRANSISTORS ARE AVAILABLE WITH REVERSED BASE-EMITTER CONNECTIONS TO SIMPLIFY LAYOUTS BUT THEIR AVAILABILITY IS GENERALLY POOR. SEVERAL DEVICE MANUFACTURERES ARE ADVOCATING THE USE OF THE "J" LEADS RATHER THAN THE "GULL WING" TYPE LEADS. SINCE THE GULL WING DESIGN IS ALREADY WELL ESTABLISHED IT IS DOUBTFUL THERE WILL BE MUCH SUPPORT FOR ANOTHER PACKAGE CHANGE.

MINOR MIMENSIONAL DIFFERENCES EXIST BETWEEN THE US, EUROPEAN AND ASIAN MANUFACTURERS. PROVIDED TAPE AND REEL PACKAGING IS USED THESE VARIATIONS ARE TOLERABLE AND PRESENT NO PROBLEMS. HOWEVER, WHEN CATRIDGES ARE USED THESE DIMENSIONAL DIFFERENCES WILL RESULT IN FEED JAMS.

POWER DEVICES UTILIZE THE SOT-89 PACKAGE WHICH HOUSE A MAX DIE SIZE OF 60X60 MIL AND A MAX POWER DISSIPATION OF 1 W. THE PACKAGE IS MOUNTED FLUSH ONTO THE PCB FOR HEAT SINKING (NOTE: DEVICE HAS NO RAISED LEADS). DUAL COLLECTOR CONNECTIONS FACILITATES RF LAYOUTS. UNFORTUNATELY, THE SOT-89 HAS LIMITED SOURCING. A HIGHER POWER PACKAGE, CAPABLE OF 3 W POWER DISSIPATION, IS UNDER DEVELOPMENT AT MOTOROLA. IT IS ESSENTIALLY A MODIFIED TO-220 PACKAGE (NAMED "DPAK"). THE PACKAGE WILL BE ABLE TO HOUSE A MAX DIE SIZE OF 115X115 MILS.

FET's

FET's ARE PACKAGED IN A 4 LEADED VERSION OF THE SOLT-23 (CALLED SOT-143). GENERAL PURPOSE, SWITCHING, AUDIO AND RF FET's ARE AVAILABLE.

INTEGRATED CIRCUITS

IC's ARE AVAILABLE IN 3 DIFFERENT STYLES OF PLASTIC PACKAGES:

1) SMALL OUTLINE (SO)
2) PLASTIC LEADED CHIP CARRIER (PLCC) AND
3) FLAT PACK

THE SO PACKAGE IS 150 MIL WIDE AND IS EFFECTIVELY A "SHRUNK" DIP PACKAGE USING 50 MIL LEAD SPACING AND IS SUPPLIED WITH LEADS PREFORMED FOR SURFACE MOUNTING. THE CHANGE TO SURFACE MOUNTING MEANS THAT THE PCB INTERCONNECTION DENSITY CAN BE DRAMATICALLY INCREASED AS THE SIZE OF THE VIAS (HOLES FOR INTERCONNECTION FROM ONE SIDE TO THE OTHER SIDE) CAN BE SMALLER. IN ADDITION, INTERCONNECTS DO NOT NEED TO BE ROUTED AROUND LARGE HOLES DRILLED FOR COMPONENT LEADS.

THE SO PACKAGES ARE AVAILABLE IN 8 PIN (78X94 MIL DIE),14 PIN (78X122 MIL DIE) AND 16 PIN (78X134 MIL DIE), AS WELL AS THE WIDE BODIED VERSIONS (300 MIL) SO-16L, SO-20, SO-24 AND SO-28. THE ASIAN SEMICONDUCTOR MANUFACTURERS USE A SLIGHTLY WIDER BODY THAN THE EUROPEAN/US MANUFACTURERS. SOME MANUFACTURERS ARE ADVOCATING LEADS THAT ARE ROLLED UNDER THE DEVICE ("J" LEADS) TO REDUCE THE OCCUPIED REAL ESTATE AND ELIMINATION OF PROTRUDING LEADS FACILITATES FEEDING AUTOMATIC ASSEMBLY EQUIPMENT. HOWEVER, SOLDER JOINTS ARE DIFFICULT TO INSPECT AND TEST POINTS MUST BE PROVIDED

TO ACCESS THE LEADS. STANDARD TEMPERATURE RANGE IS 0 TO 70 DEGREE C.BUT MANY DEVICES ARE CAPABLE OF OPERATION FROM -40 TO +85 DGREE C.OPERATION. THE PACKAGE HAS NOT YET BEEN APPROVED BY JEDEC.

BEYOND 20 LEADS THE PLCC IS CONSIDERED TO BE MORE COST EFFECTIVE THAN THE SO PACKAGE. THIS SQUARE PLASTIC PACKAGE IS AVAILABLE IN JEDEC STANDARD SIZES OF 20,28,44,52,68 AND 84 LEAD OR 5,7,11,13 AND 17 PER SIDE. LEAD SPACING IS ALSO ON 50 MIL CENTERS AND "J" LEADS ARE USED.

THE PLCC PERFORMS BETTER THAN A COMPARABLE DIP - PARTICULARLY AS LEAD COUNT INCREASES - IN LARGE PART BECAUSE OF THE DIFFERENCES IN LEAD AND CONDUCTOR LENGTH. THE LONGEST TRACE ON A 64 LEAD DIP IS ALMOST 8 TIMES THAT OF THE CORRESPONMDING TRACE ON THE 64 LEAD PLCC. LONG LEADS MEANS INCREASED INDUCTANCE AND RESISTANCE. UNEQUAL TRACE LENGTH AFFECTS SYSTEM AND DEVICE PERFORMANCE BY RESTRICTING POWER AND GROUND CAPBILITIES. LONG SIDE-TO-SIDE CONDUCTOR TRACES RESULT IN SIGNIFICANT LEAD-TO-LEAD CAPACITANCES. CLOCK RATES TO 4 GHz HAVE BEEN REALIZED.

THERMAL FACTORS

WITH ALL LEADED PLASTIC PACKAGES THE MAIN HEAT PATH IS THROUGH THE LEADS. SOME MANUFACTURERS ARE PROVIDING COPPER LEAD FRAME IN PLACE OF THE CONVENTIONAL ALLOY 42 FOR SOIC's. THIS LOWER THE THERMAL RESISTANCE BY APPROX. 80 DEGREES C./W. E FREE AIR TEMPERATURE THERMAL RESISTANCE IS OF LITTLE USE TO THE DESIGNER SINCE THE THERMAL RESISTANCE IS GREATLY IMPACTED BY THE SUBSTRATE MATERIAL. APPROXIMATE THERMAL RESISTANCE VALUES OF SEMICONDUCTORS WHEN MOUNTED ON A PCB ARE:

SOT-23	420 DGREES C/W
SOT-89	160
SO-8	260
SO-14	190
SO-16	180
SO-28	140

(A CERAMIC CUBSTRATE WILL IMPROVE THESE FIGURES APPROX. 25%)

CONFORMAL COATING PROVIDES SOME IMPROVEMENT IN THERMAL CONDUCTIVITY. REDUCING THE AIR GAP OR FILLING THE VOID BETWEEN THE DEVICE AND PCB WITH THERMAL CONDUCTIVE COMPOUND WILL ALSO LOWER THE THERMAL RESISTANCE.

MOISTURE RESISTANCE

THERE HAS BEEN MUCH CONCERN THAT PLASTIC PACKAGES, BEING NON-HERMETIC, HAVE INADEQUATE RELIABILITY FOR INDUSTRIAL APPLICATIONS. THIS ISSUE IS BEING ADDRESSED BY THE SEMICONDUCTOR MANUFACTURERS AND MOST DEVICES ARE PROTECTED WITH A NITRIDE PASSIVATION LAYER OVER THE ALUMINUM DIE AND A SLICONE LAYER ACROSS THE DIE AND CONDUCTORS. IN ADDITION, SILICONE IS ADDED TO THE PLASTIC ENCAPSULANT. A CRITERIA BEING USED INDUSTRIAL/TELECOMMUNICATION APPLICATIONS IS THAT THE DEVICES MUST PASS A 2000 HOUR LIFE TEST WITH DC BIAS AT 85 DEGREES C. AND 85% HUMIDITY AND ALSO A 96 HOURS 120 DEGREE C. 15 PSI "PRESSURE COOKER" TEST. (THIS ACCELERATED TEST REPRESENTS 20 YEARS OF FIELD LIFE FOR TELECOMMUNICATION EQUIPMENT.) AN ALTERNATIVE APPROACH IS TO USE AN EXTERNAL CONFORMAL COATING OF ELASTO-PLASTIC SILICONE

FIELD EXPERIENCE INDICATES THAT WHEN THE DEVICE IS OPERATED CONTINUOUSLY AND DISSIPATING APPROXIMATELY 100 mW OF POWER, THE INTERNAL HEAT GENERATED WILL DISSIPATE ANY MOISTURE WHICH PENETRATES THE PACKAGE. POTENTIAL CORROSION IS MORE LIKELY TO OCCUR WITH CMOS (LO-CURENT/HIGH IMPEDANCE) DEVICES.

COMPONENT SOLDERABILITY

SOLDERABILITY IS PROBABLY THE MOST CRITICAL ISSUE FOR SUCCESS WITH THE SURFACE MOUNTING PROCESS. IT IS ESSENTIAL THAT THE TERMINATIONS BE COATED WITH SN60 OR SN63 OR BE PLATED WITH TIN-LEAD ALLOY ABOVE 60% TIN CONTENT AND THAT THE TERMINATIONS CONTAINING PRECIOUS METALS ARE PROTECTED FROM LEACHING BY A BARRIER LAYER.

SOLDERABILITY TESTING IS DONE BY IMMERSING THE COMPONENT FOR 20 SECONDS AT 245 DEGREES C. AND EVALUATING THE TERMINATIONS WITH AT LEAST 10X MAGNIFICATION. 95% OF THE TERMINATION AREA MUST BE COVERED WITH A NEW, CONTINUOUS AND SMOOTH SOLDER COATING. SUCH TESTS MUST BE CONDUCTED BY RECEIVING INSPECTION ON A REGULAR BASIS TO ENSURE SUPPLIER COMPLIANCE.

IN JAPAN IT IS COMMON PRACTICE TO STRESS THE COMPONENTS AND SOLDER JOINTS OF THE COMPLETED ASSEMBLY BY SUBJECTING IT TO 5 TO 10 TEMPERATURE CYCLES FROM -20 TO +85 DEGREES C. THIS IS MUCH MORE EFFECTIVE IN DETECTING WEAK SOLDER JOINTS THAN BY VISUAL INSPECTION.

RELIABILITY CONSIDERATIONS

TO ENSURE GOOD RELIABILITY THE DESIGNER MUST "PICK GOOD PARTS AND USE THEM RIGHT". SURFACE MOUNT TECHNOLOGY HAS INHERENTLY BETTER RELIABILITY THAN CONVENTIONAL INSERTED PCB TECHNOLOGY FOR THE FOLLOWING REASONS:

1) EACH LEADLESS COMPONENT ELIMINATES 2 INTERNAL CONNECTIONS

2) DELETION OF PLATED-THRU-HOLES IMPROVES RELIABILITY OF THE PCB

3) LEADS ON ACTIVE DEVICES ARE PREFORMED THEREBY ELIMINATING DAMAGED SEALS DUE TO LEAD STRESSES

4) LOWER MASS OF COMPONENT IMPROVES SHOCK AND VIBRATION CHARACTERISTICS

5) VAPORPHASE REFLOW SOLDERING LIMITS TEMPERATURE EXPOSURE TO 215 DEGREES C.(50 DEGREES COOLER THAN WAVE SOLDERING)

POWER DISSIPATION SHOULD BE LIMITED TO 70% OF MAX RATING AND JUNCTION TEMPERATURE TO A MAX.110 DEGREES C. STRESS LEVELS IN GENERAL SHOULD NOT EXCEED 50% OF MAX RATINGS. AS WITH CONVENTIONAL DEVICES THE HIGHEST FAILURE RATE COMPONENTS ARE ACTIVE DEVICES.

COMPONENT STANDARDIZATION

CONSIDERATION MUST BE GIVEN TO COMPONENT STANDARDIZATION WHEN MOVING TO AUTOMATIC ASSEMBLY. THE REDUCTION IN THE NUMBER OF PARTS BY STANDARDIZATION IS ESSENTIAL IF PARTS ARE TO BE STOCKED, PURCHASED AND MANUFACTURED IN SUFFICIENT QUANTITIES TO BE ECONOMIC.

MOST FEEDING METHODS OF AUTOMATIC PLACEMENT EQUIPMENT DO NOT UTILIZE TAPE SEQUENCING (AS IS STANDARD PRACTICE WITH AUTOMATIC INSERTION EQUIPMENT.) IT IS THEREFORE NECESSARY TO MINIMIZE THE NUMBER OF DIFFERENT COMPONENT PART NUMBERS AS THIS REDUCES THE NUMBER OF REQUIRED TAPE REELS, LOADING CHANGEOVERS OR SETUPS FOR THE MACHINE. NOT USING TAPE-SEQUENCED COMPONENT HAS THE ADVANTAGE THAT DESIGN CHNAGES ARE MORE READILY ACCOMMODATED DURING A PRODUCTION RUN.

COMPONENT PACKAGING

INTEGRAL TO THE DEVELOPMENT OF SMC's IS THE PACKAGING OF THESE DEVICES FOR USE BY AUTOMATIC PLACEMENT EQUIPMENT. THE BASIC REQUIREMENTS ARE:

1) LOCATE AND ORIENT THE DEVICE
2) IDENTIFY AND PROTECT THE DEVICE DURING SHIPMENT AND STORAGE
3) FEED THE DEVICE IN A STANDARD MANNER TO THE AUTOMATIC PLACEMENT EQUIPMENT

THE PREFERRED SYSTEM IS TO USE TAPE AND REEL WHEREVER POSSIBLE. IT HAS THE FOLLOWING ADVANTAGES:

1) SIMPLIFIES KITTING AND STORAGE
2) TOLERATES DIMENSIONAL VARIATIONS
3) ELIMINATES INADVERTENT MIXING OF DIFFERENT PARTS
4) SHORT PRODUCTION SETUP AND CHANGEOVER TIMES
5) HAS INHERENT CAPABILITY OF DELIVERING DEVICES THAT HAVE BEEN 100% ELECTRICALLY VERIFIED AT THE POINT OF PACKAGING

EIA PACKAGING SPECIFICATION RS-481 IS THE US STANDARD AND INTERNATIONAL STANDARDS ARE EVOLVING. TAPE SIZES ARE 8, 12, 16, 24 AND 32 MM WIDTH AND REELS ARE 7, 11.25 AND 13 IN DIAMETER. THE REELS WILL HOLD 4000, 9000 AND 14000 PARTS ON AN 8 MM WIDE TAPE.

COMPONENT AVAILABILITY

COMPONENT AVAILABILITY HAS BEEN THE MAJOR REASON FOR THE SLOW ADAPTATION OF SMT IN THE US. CURRENTLY MANY USERS ARE PUT ON ALLOCATION BY THE COMPONENT MANUFACTURERS DUE TO THE DEMAND OUTSTRIPPING AVAILABLE WORLD PRODUCTION CAPACITY. THIS PROBLEM IS PARTICULARLY SEVERE FOR SMALL VOLUME ORDERS. SOME SUPPLIERS ARE MAKING AVAILABLE "OFF-THE-SHELF" ENGINEERING DEVELOPMENT KITS TO ASSIST DESIGNERS IN OBTAINING LOW QUANTITY PARTS. SOME EXAMPLES ARE:

CHIP RESISTORS (PANASONIC)	120 VALUES FROM 10 OHM TO 1 MEG (24000 PARTS - 200 EACH VALUE)
CERAMIC CAPACITORS (MURATA-ERIE)	60 VALUES FROM 0.5 pF TO .22 uF (6500 PARTS) NPO, X7R, Z5U, Y5V AVAILABLE IN 1206 OR 0805 SIZE WITH OR WITHOUT NICKEL BARRIER
CHIP INDUCTORS (COILCRAFT)	64 VALUES FROM 0.01 TO 1000 uH (384 PARTS)

IT IS FREQUENTLY NECESSARY TO USE CONVENTIONAL LEADED COMPONENTS IN SURFACE MOUNT DESIGNS DUE TO THE NON-AVAILABILITY OF SUITABLE EQUIVALENT SURFACE MOUNT PARTS OR FOR COST REASONS. FOR EXAMPLE, A STANDARD DIP PACKAGE MAY NOT BE AVAILABLE IN AN SOIC. IT IS PRACTICAL AND QUITE EFFECTIVE TO MODIFY THE DIP LEADS AND "LAP-SOLDER" THE DIP. THE SPACE UNDER THE DIP CAN ALSO BE UTILIZED TO MOUNT SEVERAL CAPACITORS AND RESISTORS. MANY SMALL LEADED COMPONENTS CAN, AS AN INTERIM MEASURE, BE ADAPTED FOR SURFACE MOUNTING IN THIS MANNER. IT IS IN GENERAL MORE EFFICIENT TO BUILD ASSEMBLIES ALL SURFACE MOUNTED RATHER THAN MIX TWO PROCESSES (INSERTED WAVE SOLDERED AND SURFACEMOUNTED REFLOW SOLDERED PARTS.

ANOTHER TECHNIQUE IS TO USE LEADLESS CHIP CARRIER PACKAGES (LCC's) FOR THE ENGINEERING PROTOTYPES AND THEN REPLACE THESE WITH THE PLCC's AS THEY BECOME AVAILABLE AT A FUTURE DATE. THIS IS FEASIBLE SINCE THE PLCC AND LCC FOOT PRINTS ARE IDENTICAL (JEDEC STANDARD).

THE PRACTICE OF PLACING PARTS UNDER THE IC IS ALSO DONE IN HIGH-DENSITY DESIGNS WHEN SPACE REQUIREMENTS NECESSITATE THIS UNORTHODOX PACKING APPROACH.

FUTURE TRENDS

* DURING 1985 THE USE OF SMC,s WILL ACCELERATE AND BY 1990 THEY WILL REPLACE LEAD-IN-HOLE COMPONENTS AS THE DOMINANT PCB PACKAGING TECHNOLOGY.

* SMC COSTS WILL CONTINUE TO DECREASE AND PARITY WILL BE REACHED IN 1986 FOR MOST COMPONENTS.

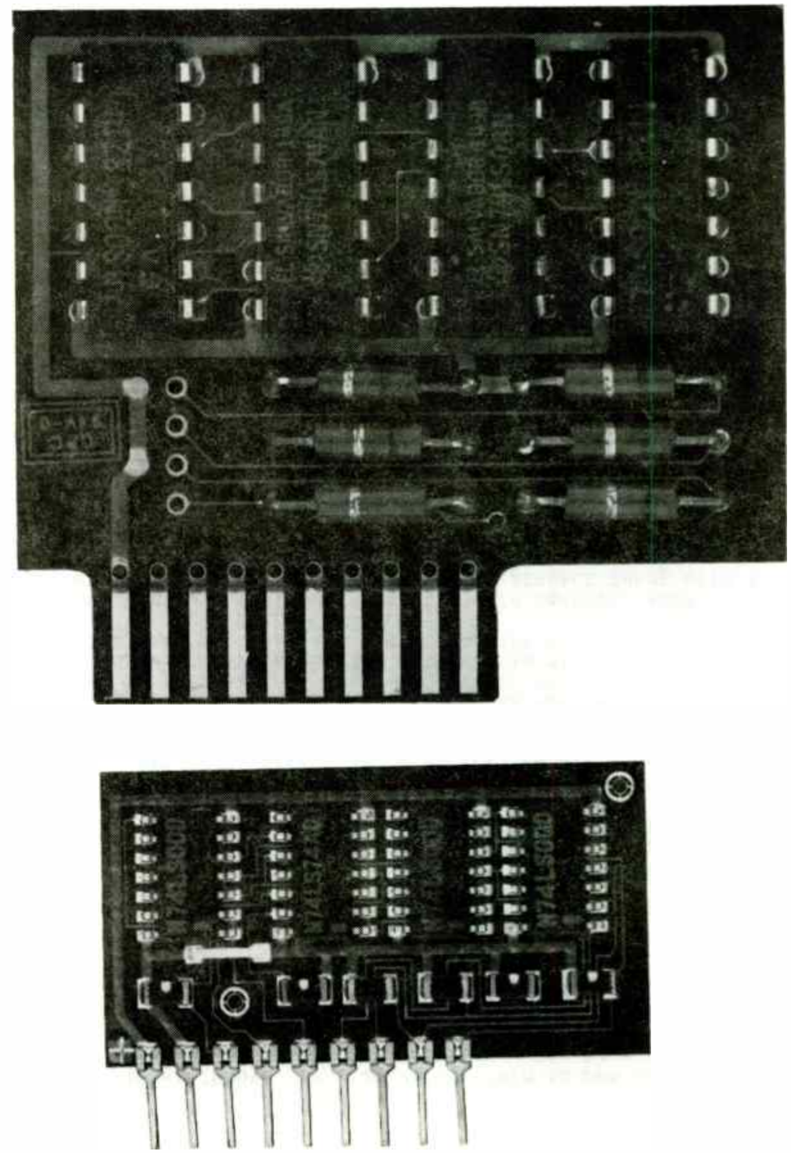
* MANUAL ASSEMBLY WILL NO LONGER BE COST EFFECTIVE.

* THE PREFERRED SMC PACKING METHOD WILL BE IN TAPE AND REEL FORMAT

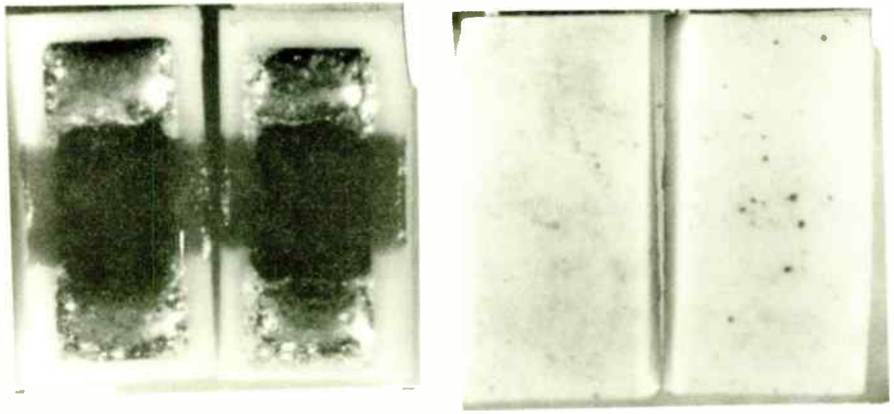
* WAVESOLDER WILL CONTINUE TO BE USED BUT FOR HIGH

DENSITY AND HIGH LEAD COUNT DEVICES VAPORPHASE REFLOW SOLDERING WILL PREDOMINATE.

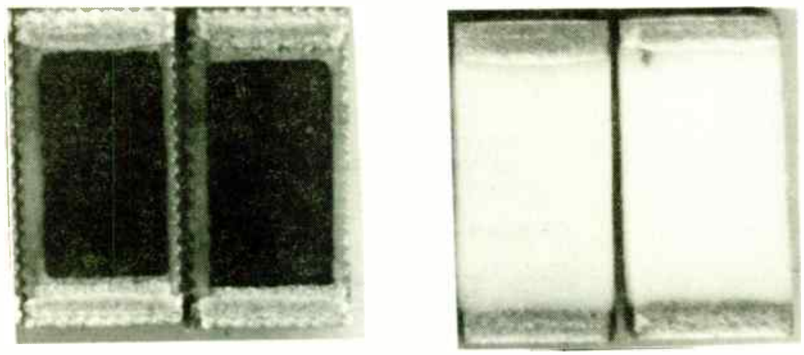
* COMPONENT QUALITY WILL REACH LEVELS SUCH THAT INCOMING INSPECTION WILL NO LONGER BE REQUIRED.



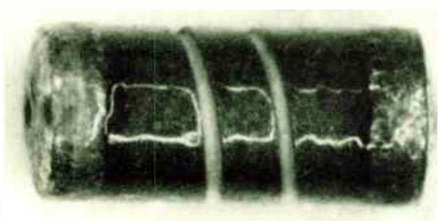
SIZE COMPARISON BETWEEN INSERTED AND SURFACE MOUNTED ASSEMBLY (AWI)



SINGLE-SIDED RESISTOR

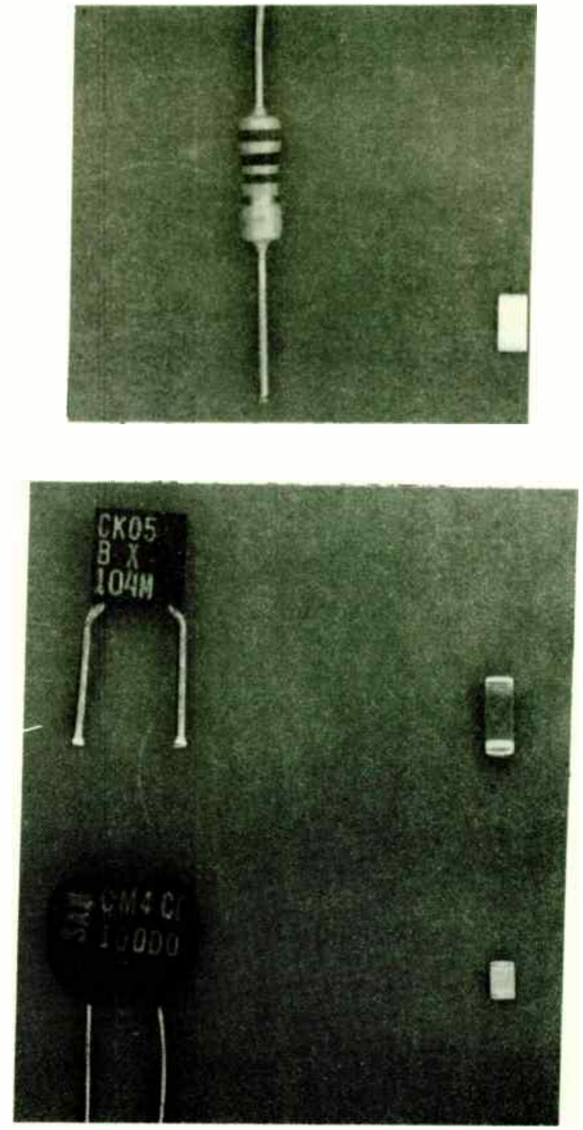


DOUBLE-SIDED RESISTOR

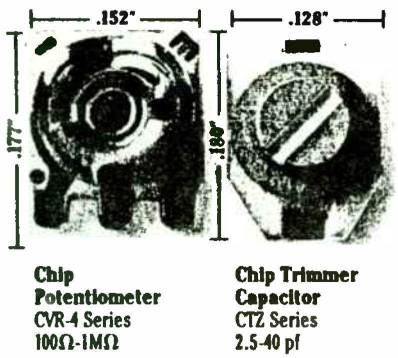


MELF (TRW)

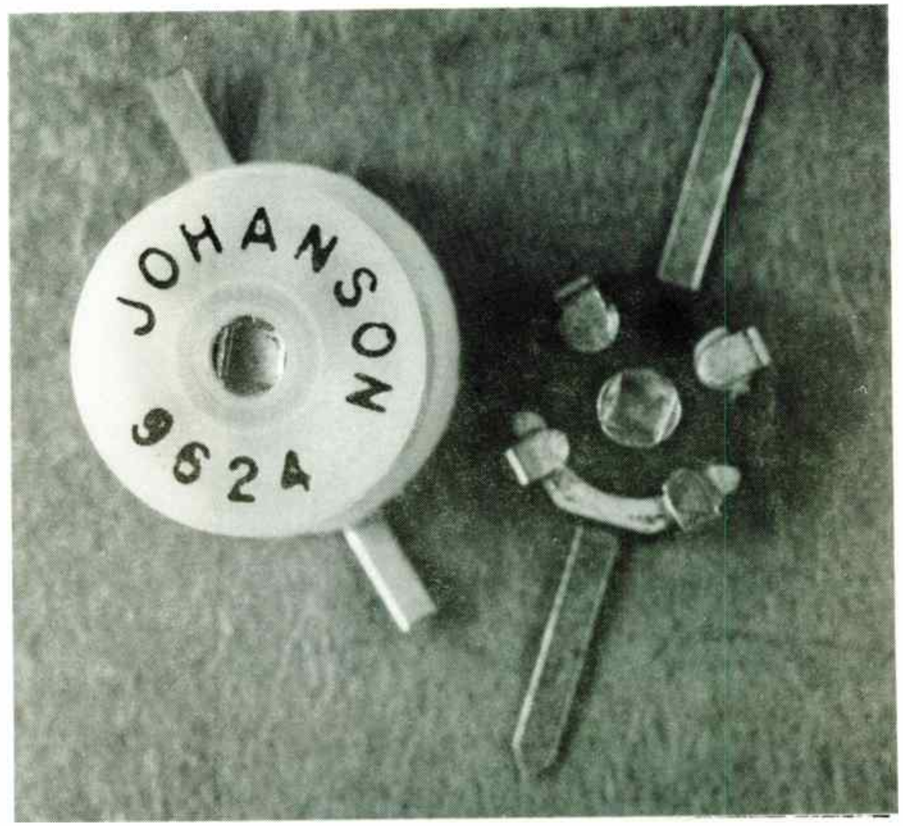
ASIAN MELF



SIZE COMPARISONS - RESISTOR AND CERAMIC CAPACITOR

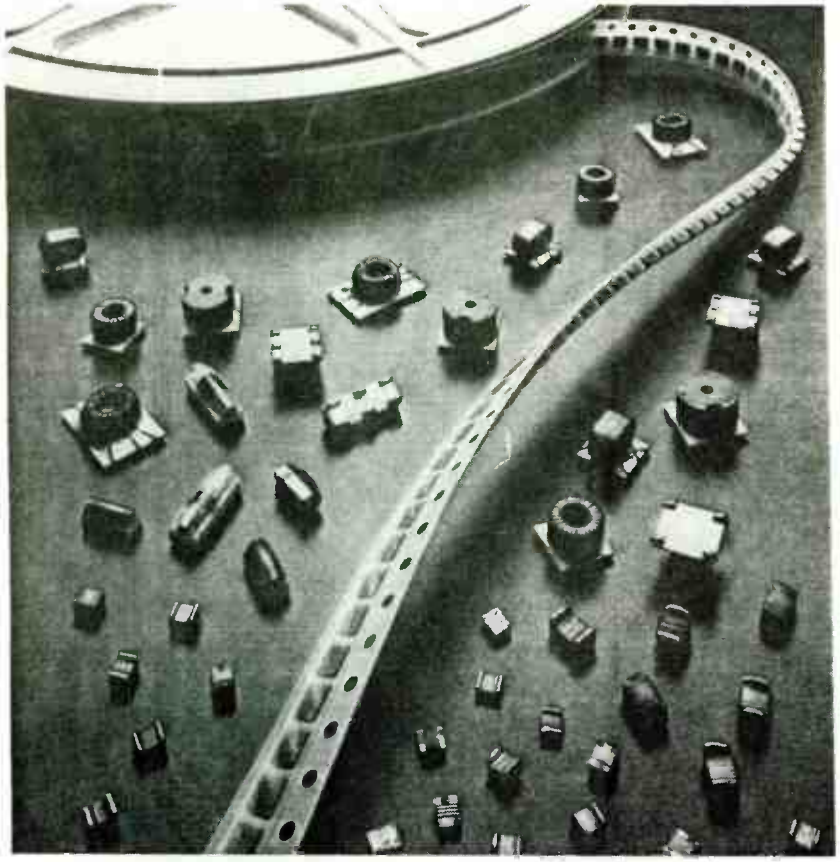


VARIABLE RESISTOR AND CAPACITOR (KYOCERA)

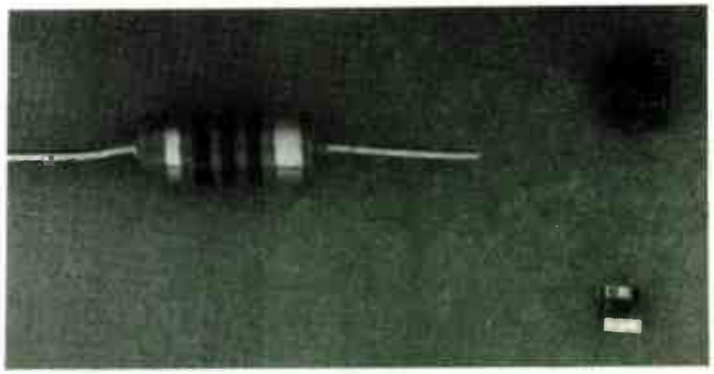


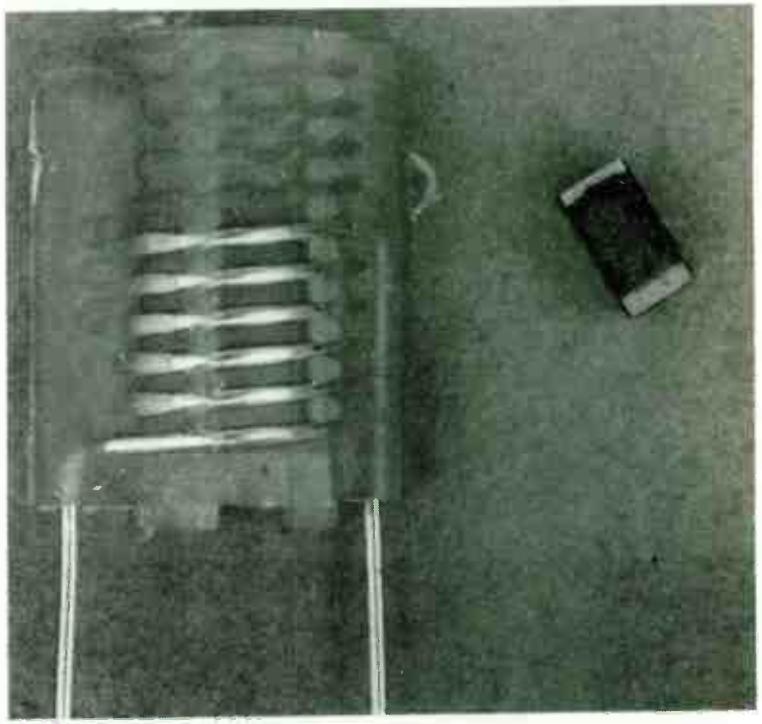
MOISTUREPROOF TYPE STRIPLINE TYPE

VARIABLE CAPACITORS



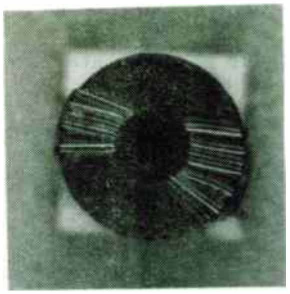
VARIOUS CHIP INDUCTORS AND TRANSFORMERS (COILCRAFT)

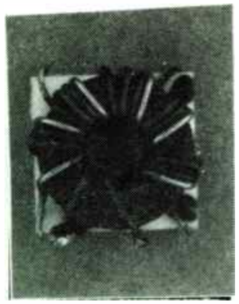




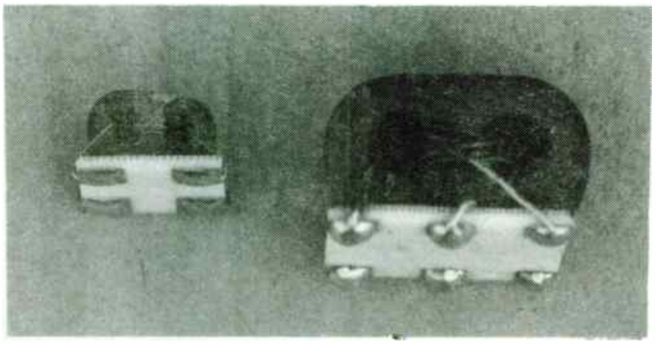
SIZE COMPARISONS - FIXED AND VARIABLE INDUCTOR

A. Ceramic Ferrite Chip Inductor

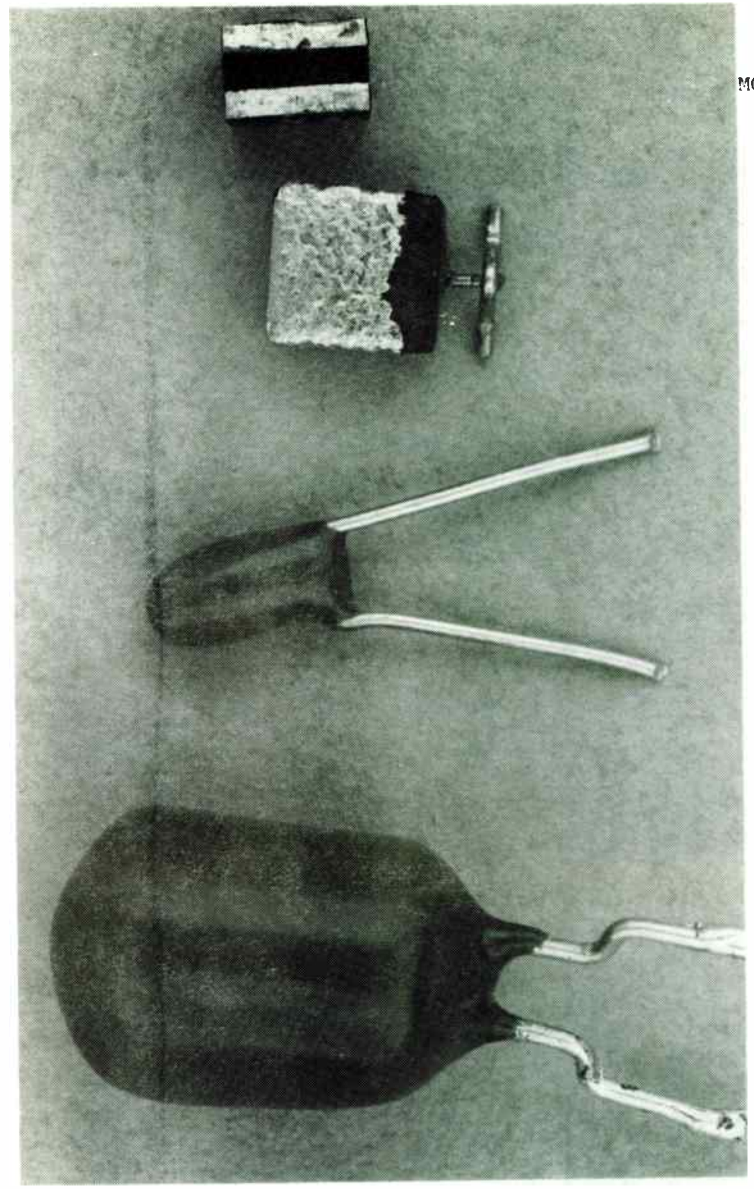




B. Toroidal Transformers (Nonpreferred)
Mounted on Ceramic Carrier



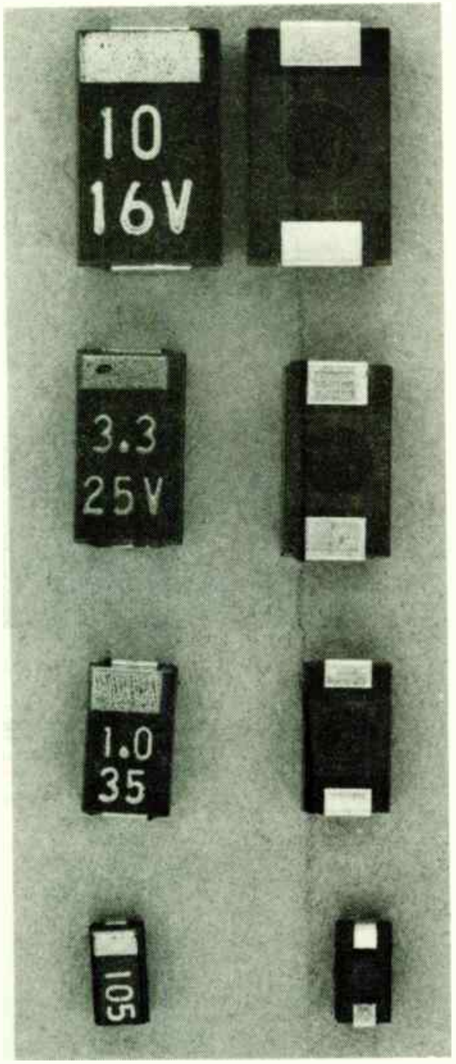
C. Balun Core Transformer (Preferred)
on 0.15- and 0.25 in. Square Ceramic Carrier



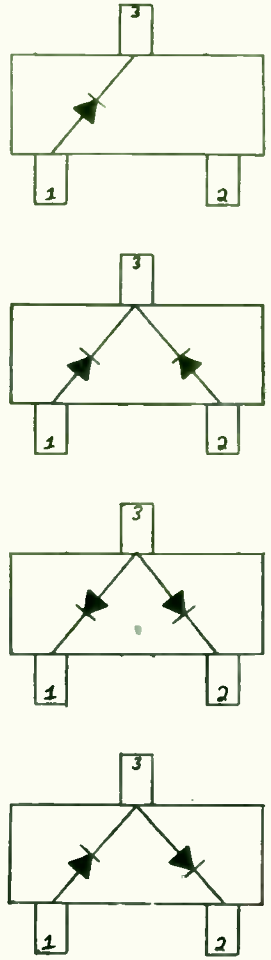
MOLDED

T-BAR

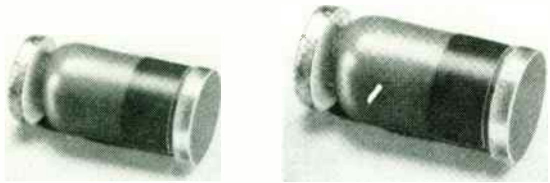
SIZE COMPARISONS - TANTALUM CAPACITORS



4 SIZES OF MOLDED TANTALUM CAPACITORS (NEC)

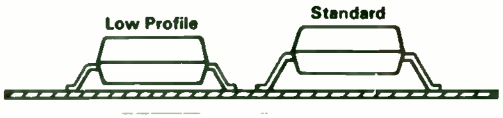


DIODE CONFIGURATIONS IN SOT-23 PACKAGE



0.2 W　　　1 W　　　(A)

MELF DIODE

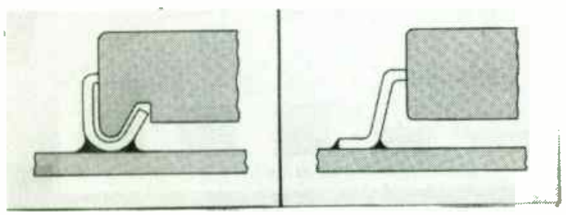


SOT-23 PACKAGE OPTION　　　(B)

SOT-89

NOTE:
Collector
has 2
possible
connections

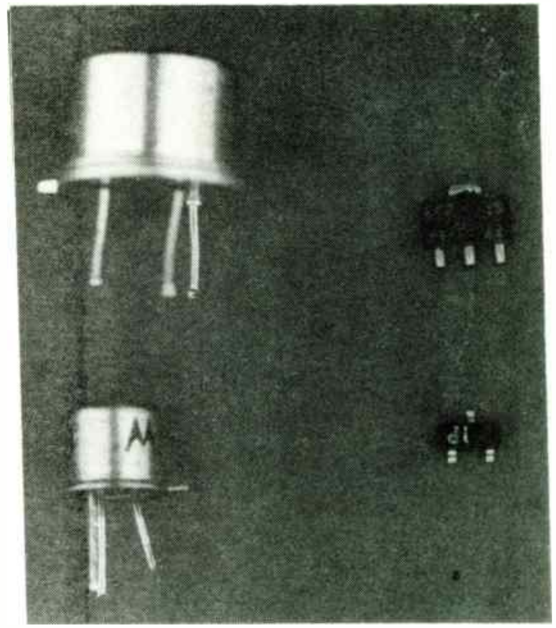
SOT-23



"J" LEAD　　　GULL WING　　　(C)

TO-5

TO-18

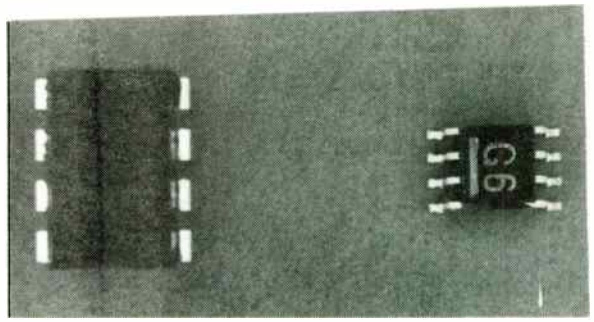


SOT-89

SOT-23

8-PIN

DIP



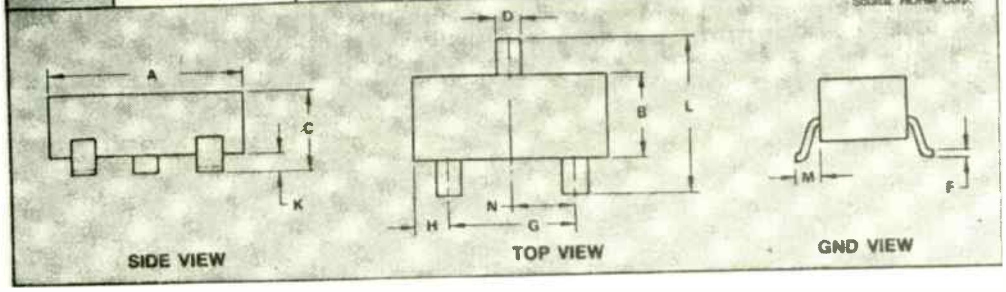
SO-8

SIZE COMPARISONS

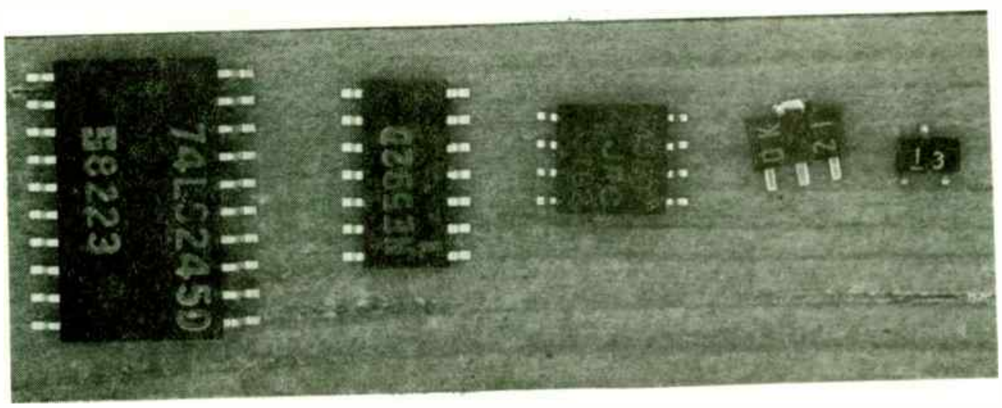
DIMENSIONAL DIFFERENCES WITH VARIOUS SOT-23 MANUFACTURERS

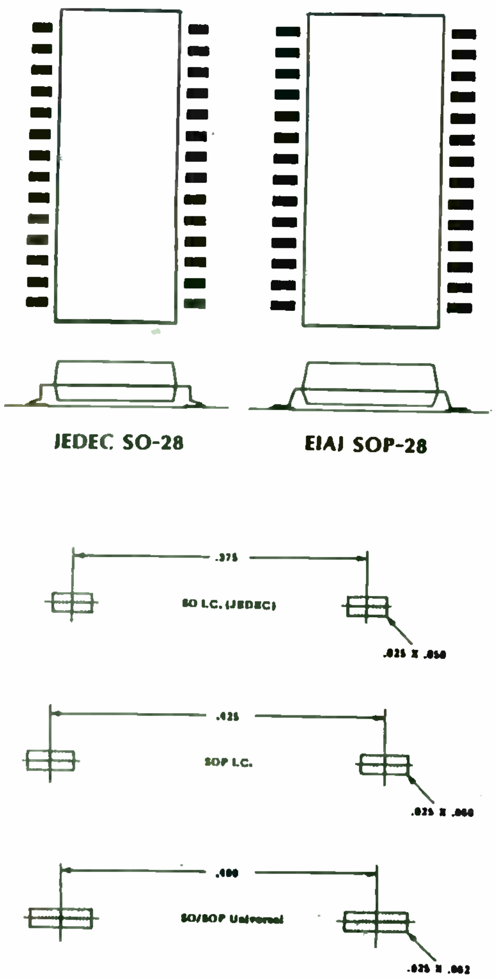
Dimension (see diagrams below)	MOTOROLA min	MOTOROLA max	ROHM min	ROHM max	FERRANTI min	FERRANTI max	AMPEREX min	AMPEREX max	NEC min	NEC max
A	2.80	3.05	2.70	3.10		2.90	2.80	2.90	2.70	3.10
B	1.20	1.40	1.50	1.80		1.40	1.20	1.30	1.50 typ	
C	0.85	1.20	1.00	1.45		1.20	0.80	1.20	1.20	1.55
D	0.37	0.43	0.35	0.50	0.40 typ		0.37	0.43	0.35	0.50
F	0.076	0.127	0.09	0.25	0.10 typ		0.085	0.115	0.10	0.26
G	1.78	2.04	1.70	2.10	1.90 typ		1.90 typ		1.90 typ	
H	0.51	0.61								
K	0.10	0.25		0.15	0.05		0.10	0.20	0.10	0.25
L	2.10	2.50	2.60	3.00		2.75	2.10	2.50	2.20	3.00
M	0.45	0.61					0.40	0.60		
N	0.89	1.02	0.94	0.96	0.95 typ		0.95 typ		0.95 typ	

Source: ROHM Corp.

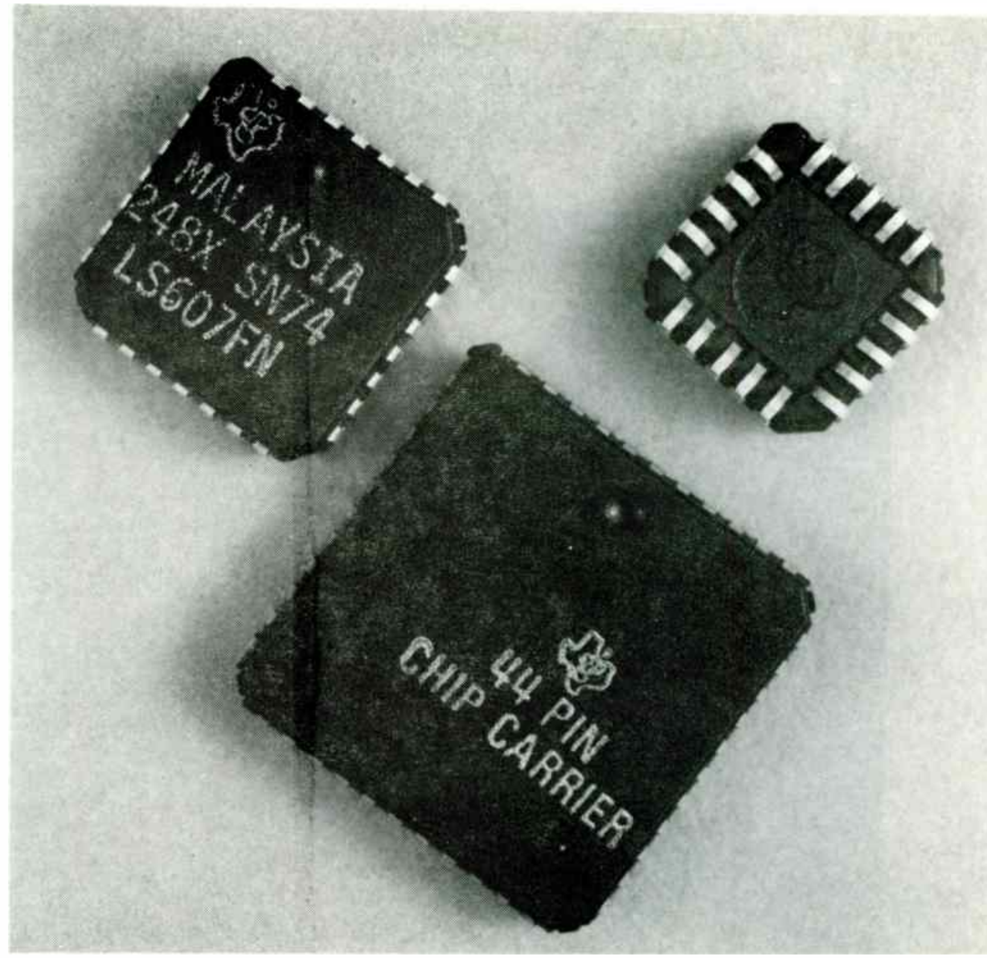


SIZE COMPARISON - SOT AND SOIC





UNIVERSAL PAD DESIGN TO ACCOMMODATE US AND ASIAN SOIC's

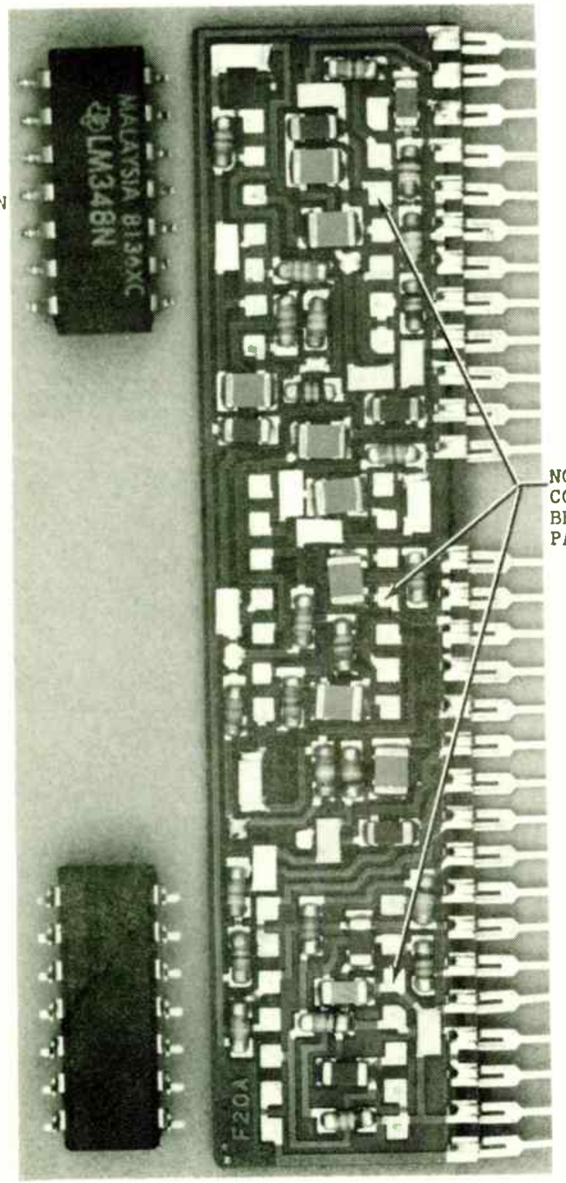


20/28/44 PIN PLCC PACKAGES

TOP

14 PIN

DIP



NOTE
COMPONENT
BETWEEN
PADS

BOTTOM

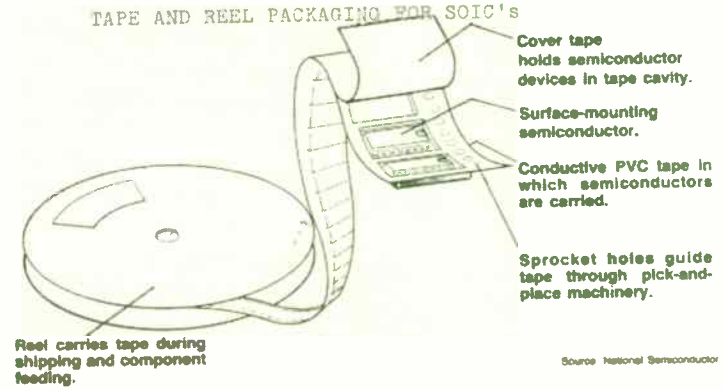
DIP

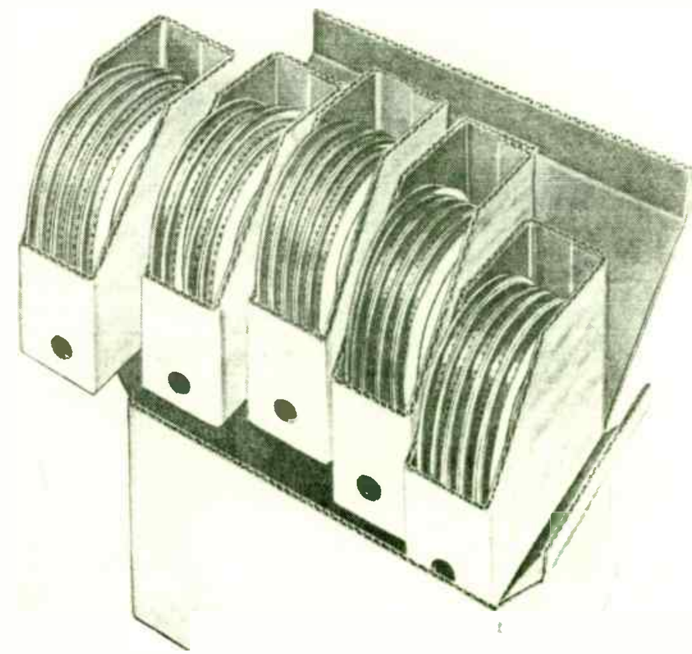
SURFACE

MOUNTED

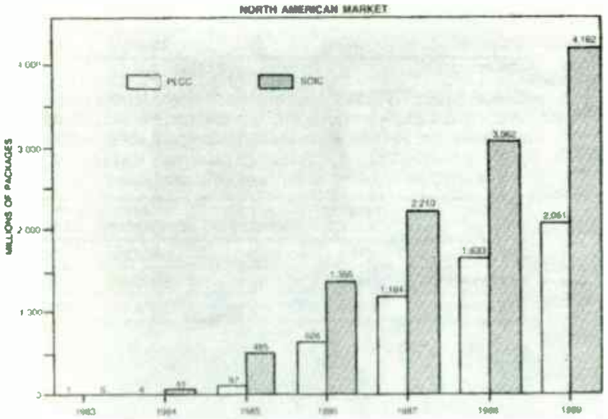
DIP's WITH PREFORMED LEADS SURFACE MOUNTED

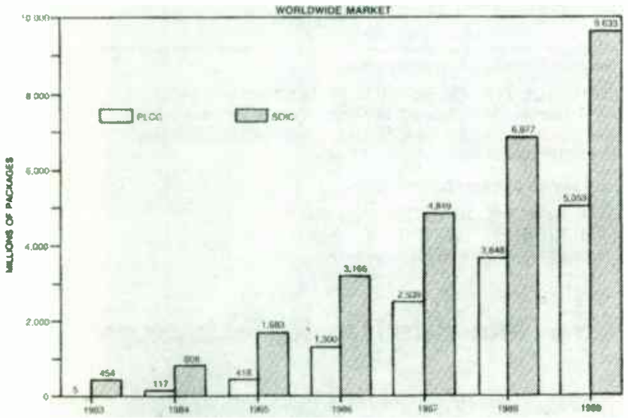
TAPE AND REEL PACKAGING FOR SOIC's





CONTAINER WITH TWENTYFIVE 7" REELS
HOLDING 100,000 CHIP COMPONENTS





Between 1984 and 1989, the consumption of surface-mounting packages is projected to increase 15-fold worldwide and 150-fold in North America. (Data source: Insight/Onsite.)

CONSUMPTION FORECAST FOR SOIC AND PLCC PACKAGES

INTEGRATED CIRCUITS FOR I.F. AMPLIFIERS AND DEMODULATORS

P. E. Chadwick
Plessey Semiconductors
Cheney Manor
Swindon
Wilts
England

Although the "single chip" radio is becoming practical in certain applications, and Dick Tracey's wrist radio is not quite so far in the future, the majority of todays receivers are conventional superhets with IF amplifiers. Because these amplifiers are operating at fixed frequency and selectivity can be provided by block filters, this area has proved to be a prime candidate for integration into monolithic circuitry. Early attempts at this in small scale integration led to a number of successful designs and the SSI circuits thus produced have, in many cases, lasted as "glue" circuits in new designs over some 12 or 14 years - something of a record in terms of linear integrated circuit lifetimes.

To a great extent, commercial pressure for I.F. circuits has come from the consumer market, although some early circuits were developed for military and professional applications. This has led to a preponderance of AM and FM circuits aimed at the consumer radio market and until recently, very little in the way of purpose designed AM/SSB demodulators. FM circuits, for communications purposes first appeared in the mid 1970's and the demands of the cordless telephone market for more than 1 million pieces per month in 1983 has boosted communication circuit sales. Circuits such as the Plessey SL6601 Phase Locked Loop demodulator and the much copied Motorola MC3357 and its derivatives are the designers standard devices, although new requirements are appearing.

A major area of compromise is in power consumption. In order to obtain operation at 10.7MHz the amplifying stages require to run at higher currents than for 455KHz operation, and some compromise is necessary. The availability of faster IC processes allows the current consumption to decrease, but the choice of IF is still related to current consumption. For example, the Plessey SL6640 running at 10.7MHz required a typical supply of 6mA at 6v. This device uses a classic limiting amplifier and quadrative detector operating at 10.7MHz. The SL6652 which includes a mixer and oscillator and runs the limiting amplifier at 455KHz typically requires 1.4mA at 2.5 volts. This reduction in power consumption is useful insofar as it allows circuits to be "stacked" in series across the supply, thus minimising total equipment power requirements. A further advantage of the use of the low frequency IF of 455KHz is the reduction in system cost. This is because the main selectivity of the equipment can now be obtained by using a ceramic filter, which is available relatively cheaply, while the "straight through" 10.7MHz approach requires the I.F. filter to be a multiple crystal unit with a much higher price. However, sufficient selectivity is required before frequency conversion to 455KHz to prevent gain compression or intermodulation in the second mixer: the provision of a suitably high gain compression point and/or third order input intercept point reduces the requirements for and thus the expense of this filter.

The Phase Locked Loop demodulator as exemplified in the Plessey SL6601 has been available for some years. It offers certain distinct advantages in some applications: for example, where extreme long term stability is required such as in receivers for remote power switching, and in military equipments operating over extended temperature ranges. The lack of a quadrature circuit (either as an L-C combination or as a ceramic resonator) is attractive, although certain parameters, such as the ultimate signal to noise ratio, are unlikely to ever reach those limits achievable with a quadrature type detector.

All circuits have certain "sensitive" parameters which are difficult to measure and/or to meet. Typical of these are AM rejection and sensitivity. AM rejection is a function of AM to PM conversion within the limiting amplifer and in order to minimise this, balanced stages are used (fig. 1). Nevertheless measurement becomes a difficulty as the following example illustrates.

A signal is applied to an FM receiver with a deviation of plus/minus 2.5KHz. Removal of the FM and substitution of AM at 30% requires the residual FM of the generator to be less than 7.5Hz if an accurate measurement is to be made of the commonly required 40dB rejections.

AM to PM conversion within an IC is generally caused by asymmetry in the output level which varies with input level - see figs 2 demonstrating the effect of offset on a limiter with varying input level. It is this AM-PM conversion which has led to the production of special low phase shift limiters for use in radar systems. The effects of operating at low collector currents such that F_t is falling does not appear to have any major effect, which is perhaps surprising, as a differential phase shift of 1 degree for a 5dB input variation (which corresponds to 30% AM) will produce an output some 43dB below a 3KHz deviation signal. Ref 1 provides some useful indications of the phase shift with input level in low phase shift limiters operating at frequencies of approximately 1/20 of F_t.

AM rejection is directly related to co-channel rejection and capture ratio, and the performance of wide band FM in this respect is well known. Narrow band FM (3KHz deviation) is much worse, because the allowable phase deviation caused by the unwanted signal is obviously decreased.

From this discussion, it may be deduced that the requirements for a limiting I.F. amplifier for FM demodulation include symmetrical limiting, achieved by the use of balanced stages, and low phase shift with input level variation, both of which are achievable in integrated form - possibly, indeed, more readily than with the use of discrete components.

Sensitivity is important for the receiver designer, and in all too many cases is defined as "3dB limiting" or some similar inexplicit term. The measurement of signal to noise ratio on a 100% basis is more meaningful, but difficult because of the noise level involved with automatic test machines and handlers. As a result, testing to levels below 5 microvolts is not very practical, while a "2 microvolt typical" sensitivity without any tolerance is of no use to the serious design engineer - especially as it may cost $1 to purchase a part and $2 to change it if the assembled equipment does not meet specification.

As previously stated, there are performance advantages in the use of an IF of 455KHz such as very low power consumption, for which the use of a high frequency I.C. process is mandatory. At frequencies below 1 or 2 MHz, the use of PNP transistors as active loads as in Fig 3 is attractive. The high frequency performance is limited by the difficulties of producing high frequency, high gain lateral PNP transistors in a monolithic I.C.

Cellular, military, amateur and even 900MHz cordless telephone equipments require signal strength indication (RSSI - Received Signal Strength Indicator), and this is a feature of new circuits. The use of a radar style successive detection logarithmic amplifier has obvious applications, although if a very linear monotonic RSSI response is required, the losses of any filter section within the amplifier strip must be carefully controlled. In addition, measurements of the RSSI response must be made with care, as it is not unknown for the logarithmic curve of the I.C. to show up hitherto unsuspected errors in the attenuators of signal generators. The use of an external precision attenuator can help, but great care must then be taken to avoid leakage. Older signal generators which use piston attenuators will give the least problems in this respect, provided that mechanical wear has not invalidated the calibration.

The use of the quadrature FM detector is almost universal, although as mentioned earlier, the Phase Locked Loop has some advantages. In narrow band FM the PLL shows little improvement in threshold extension, because the loop bandwidth approaches the IF filter bandwidth. In order to obtain adquate AM rejection, it is necessary to precede the PLL with a limiting amplifier and this means that no adjacent channel selectivity can be provided by the loop as unwanted signals will capture the limiting strip.

The quadrature detector is extremely popular for use in integrated circuit demodulators for a number of reasons, not least of which is the small number of pins required by the IC in comparison with a Foster Seeley or ratio detector - as well as the simplification and thus cost reduction of the inductive component.

There does however, seem to be some misconceptions with regard to the operation of a quadrature detector - fig 4. At resonance, the voltage across the quadrature capacitor is 90 degrees out of phase with the voltage across the tuned circuit, and this phase shift varies as the frequency is removed from resonance. In order that a linear output may be obtained, it is necessary for two parameters to be considered:

a) the linearity of the phase detector
b) the tuned circuit Q.

The variation in voltage with frequency offset follows a tangential form, while the phase detector (or analogue multiplier) will usually have a semi-sinusoidal transfer characteristic. This requires that for any given Q, frequency deviation must be small, so that distortion is minimised. However reducing Q will also reduce the available AF output and thus there is an optimum value for Q. In a tuned circuit, the phase shift is 45 degrees at the 3dB points and restriction of the phase variation to about 10 or 15 degrees gives a generally satisfactory compromise between output and distortion. This suggests the use of a Q that gives a 3dB bandwidth of about 3 times the deviation, which is generally acceptable.

The use of a high Q circuit damped by a resistor rather than a low Q coil is advisable insofar as the repeatability and stability are improved. The use of a ceramic resonator generally gives higher distortion and can also give matching problems between the filter centre frequency and that of the resonator. Additionally long term frequency stability is not always improved to the extent that might be expected.

The use of a differential output as in the Plessey SL6652 (fig 5) has some advantages. Where data is transmitted as FSK, frequency errors in the system can lead to bias distortion such that data can be lost with a single ended output. Additionally, the differential output when suitably filtered provides AFC and a convenient method of adjustment of the quadrature coil is to inject a signal at the correct centre frequency and adjust the quadrature coil for zero voltage across the output. The effects of internal offset voltages does mean however, that the differential output voltage may well not be zero at zero input. The use of a single ended output device will of course give problems of temperature drift as well.

It does not seem to be realised by many designers that the group delay characteristics of the IF filter can have a significant effect upon the distortion realised in the system, and especially where high speed data transmission is concerned, differential group delay should be minimised.

An area of some difficulty is squelch. After some years experience, the author is firmly convinced that the perfect squelch system will be designed by an engineer who can live in the same house in harmony with his/her parents, in laws, and grandparents without ever a cross word!

Squelch in the Phase Locked Loop decoder may be implemented by circuitry which looks for "cycle slipping" when the input signal is noisy. This method has two major drawbacks:

a) The number of cycles slipped at a 6dB S/N is considerably more than at 20dB S/N. Thus the change in output at low S/N ratios is adequate to drive, say, a Schmitt trigger, but at high S/N ratios, the trip point is by no means as well defined.

b) The front end noise when band limited by a filter is identical in form with that produced by a valid, noise modulated FM signal. Thus although only noise is present, the loop will attempt to demodulate it as a valid signal, and the squelch can therefore open on noise. Probably the best squelch method is to use the reduction in noise power above the AF band as an indication of S/N ratio: the noise from the front end is admittedly a valid signal, but is effectively de-emphasised, and so the HF noise is lower than would be expected. The reason that the noise is de-emphasised is that that high order sidebands of the FM signal are reduced by the filter: thus the low frequency noise components have more sidebands and effectively higher deviation than the higher frequency components. The use of carrier strength as a squelch control is again unsatisfactory because of the squelch breaking at high noise level inputs. These problems have led to the use of tone coded squelch systems (CTCSS) and one modern tactical military radio uses 5 squelch systems with majority voting!

References

1. Broadband Amplifier Applications
   Plessey Semiconductors

2. The SL6700 Versatile Radio Integrated Circuit
   Plessey Semiconductors

3. Chadwick, De Maw "Receiving with Plessey IC's"
   QST April 1981

4. Digital Techniques for Advanced Radio
   J. Masterton, P.A. Ramsdale, I.A.W. Vance
   IEE Conference on Mobile Radio Systems & Techniques, York, Sept. 1984
   P6 et seq.
   ISBN 0 85296297 5
   IEE London

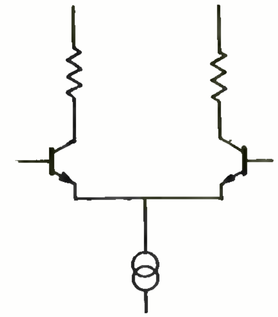


fig 1 Balanced Amplifier

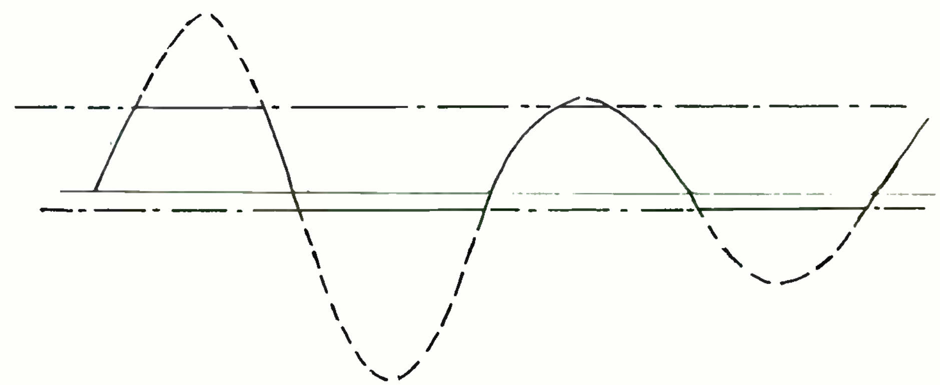


Fig 2 Assymetric Limiting leading to non equal mark - space ratio

Squelch hysteresis is required for situations where deep fading can occur. For very low signal to noise ratios hysterisis of a few dB is adequate, while at high (20dB) ratios, a larger hysterisis window gives less trouble with fading and squelch dropout. Some hysterisis is always desirable and ideally should vary with S/N ratio.

AM and SSB

Integrated circuits for AM and SSB reception are rather more rare. Leaving aside broadcast receivers, the AM requirements are now limited to CB, Aircraft, military and a few other applications. The CB requirement now tends to be multimode, while the temperature range for aircraft and military equipments are generally too exacting for broadcast type IC's. In the case of single sideband, some requirements for in channel intermodulation ratios of 60dB are extremely difficult to meet, while the market demand is currently relatively small. The growing trend to ACSB (Amplitude Compandored Single side Band) may well change this, but again, low power requirements suggest the use of a low I.F. This is not so convenient as in the FM case, because the percentage bandwidth of the IF filter is reduced. This means that the IF filter will tend to be larger, and quite possibly more expensive.

The use of RF derived, rather than AF derived, AGC is advantageous, as an RF input corresponding to a low audio frequency signal may well overload the receiver if the IF filter is fairly sharp. In addition, attack times are faster with RF derived systems because of the greater number of cycles of signal available for rectification in a given time. It is, however, important to ensure that the carrier for SSB demodulation does not block the AGC circuitry. Typical of such a device capable of very good AM/SSB performance is the Plessey SL6700 which offers a great deal of flexibility by the "addressability" of the various internal functional blocks - see fig 6. Refs 2 and 3 provide further details on the use of this device.

An area in which some difficulty can occur is in the AGC system where the ripple from the modulation remains on the AGC line. With high impedance AGC circuits used with tubes, and the high signal voltages, combined with the relatively insensitive remote cut off tube, the distortion introduced by this ripple was normally of manageable proportions. However in an IC with a 20dB/Volt AGC law, a few millivolts of ripple are capable of preventing in-channel IMD specs from being readily met. The use of an active filter in the AGC system can lead to instability unless phase compensation is used, and this is even more problematical with AF derived AGC.

Historically, the IF amplifier has provided the major part of the receiver gain, but especially in SSB systems it has become practicable to use some 50 or 60dB of gain at IF with the rest at AF. This eases layout and RF stability, but does require low noise AF amplifiers, which is a well developed technology.

The use of the RSSI output from an FM strip as an AM demodulator is possible. The logarithmic distortion introduced is only very pronounced at high modulation levels, and a two or three decade anti-log amplifier will reduce this distortion substantially. If the average RSSI output level is used to provide an offset, then acceptable results can be obtained for a wide range of input signals. It is however, unlikely that the complications involved will prove worthwhile.

IF circuits for radar applications are somewhat different, insofar as usually either a log amp, swept gain or low phase shift limiter strips are required. Although power consumption is still an important parameter, centre frequencies of 160MHz and higher militate against low power consumption. In addition, gains of 60dB at such frequencies are difficult to achieve with stability, and the production of lower gain blocks is more practical. The Plessey SL521 log amp first entered production in the 1960's and is still being designed into new systems in the 1980's. Its successors have pushed the upper frequency limit for a log strip to beyond 200MHz and this will be revised further in new generations currently undergoing development. At these frequencies, the parasitics involved is conventional dual in line and TO5 style packages are excessive, and in order to prevent performance being package limited, it has become necessary to use leadless chip carrier packages. The bonding of naked chips to hybrids can be difficult, insofar as active testing of the devices may well be impossible, or at least show little correlation with the final parameters, and the use of the chip carrier package provides both manufacturer and customer with a much higher degree of confidence. The cost of logarithmic amplifiers has now fallen to such a level that the use of a linear IF amplifier in a radar on grounds of cost can hardly be justified: even the complication of a swept gain system is unlikely to outperform the log strip, and the use of a low phase shift amplifier meets requirements for MTI, monopulse and phase encoded systems applications.

The application of direct conversion (zero I.F.) and digital techniques will doubtless lead to major charges in radio receiver design. The conventional superhet has been with us for over 60 years: technology for its complete digitisation is a long way off. Ref 4 is an example of some of the work in this field, while Collins Radio have an HF receiver which the I.F. section uses digital processing. Despite these advances, there is still a place (at the moment) for advanced analogue signal processing, using advanced IC processes to maximise performance at minimum power consumptions.

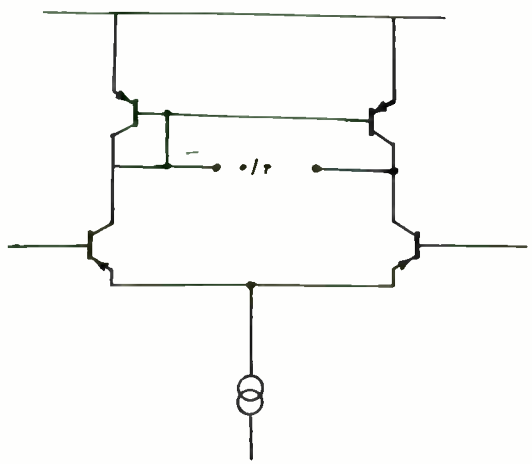


Fig 3 PNP Active Loads

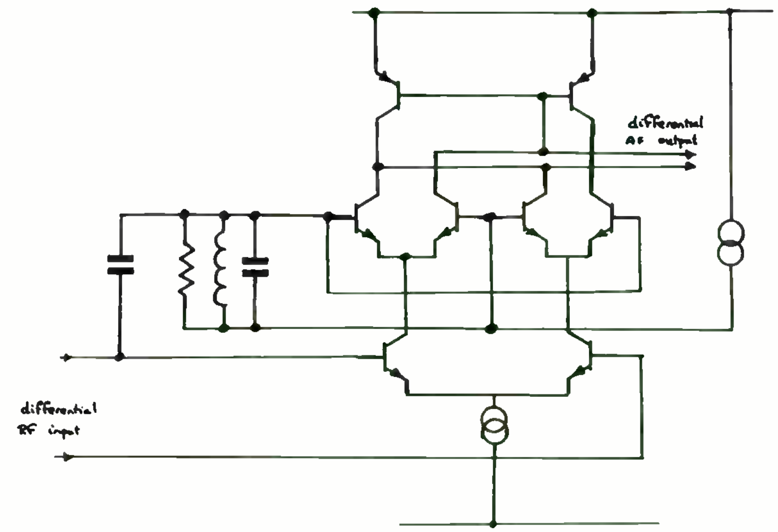


Fig 5 Balanced Quadrature Detector with Differential Output

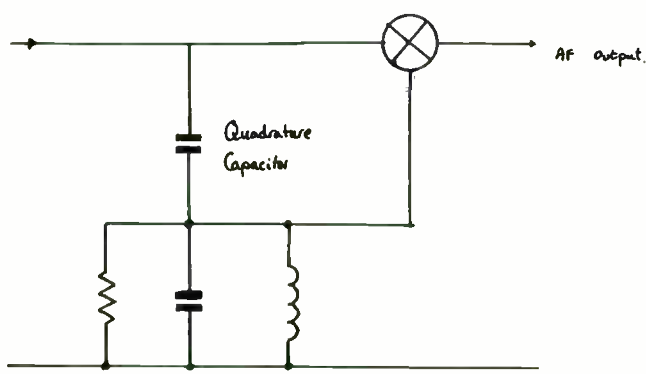


Fig 4 Quadrature Detector

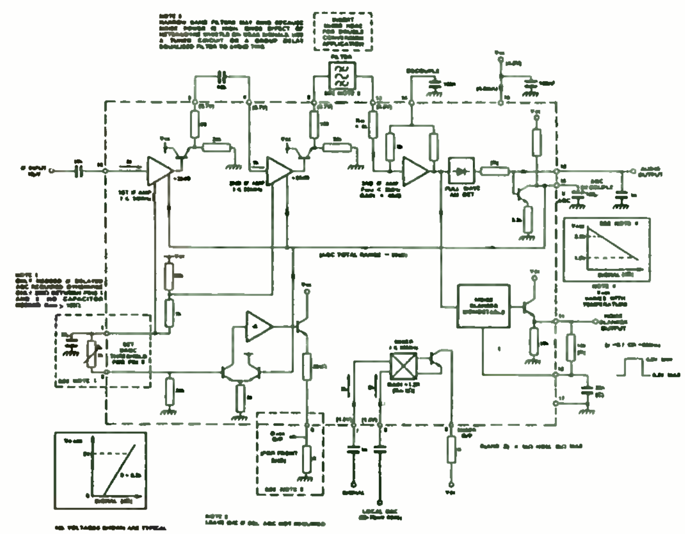


Fig 6 The SL6700 Versatile RF Integrated Circuit

MOTOROLA ADVANCED AMPLIFIER CONCEPT PACKAGE

By Alan Wood

Motorola Semiconductor
Products Sector.

ABSTRACT

This paper describes the philosophy and the design of a new generation of RF power transistors which, for Land Mobile products, offer a unique design concept that will simplify the external matching requirements for high power 800MHz amplifiers. An additional benefit is the increased efficiency that can be obtained over a wider bandwidth. These improvements are brought about by the use of multiple matching sections inside the package. Presented here will be an analysis of the design of a doubly input/output matched part showing its advantages over a conventionally matched 800MHz transistor. Also described will be the performance characteristics of two RF devices, soon to be introduced by Motorola Semiconductor Product Group. They will be rated at 60Watts output power for application in mobile radio-telephones(12V operation) and base stations(24V), specifically cellular, trunked and conventional 800MHz systems.

INTRODUCTION

In a number of RF transmitter applications in the 800 to 960MHz band, e.g. paging and cellular base stations and high power mobiles it is not uncommon for the amplifier output stage to have multiple devices in parallel. RF circuit designers would prefer to replace these complex multi-device stages with a single device or at least with fewer paralleled transistors. But increasing the output power of current 800MHz transistors does present a number of problems: larger transistor die would lower the manufacturing yields, dissipating the additional heat in the existing package would limit the maximum operating temperatures, and the device impedances would be so low that broadband amplifier design would be extremely difficult.

High power RF transistors developed for these applications must therefore exhibit a number of desirable features namely:-

i. **Power Gain.** Gain should be as high as achievable using the current processing technology but not at the neglect of other important parameters, i.e. ruggedness and stability.

ii. **Power Added Efficiency.** High efficiency is of paramount importance in any high power amplifier application. Space requirements limit the volume that can be dedicated to power supplies and heatsinking structures. Invariably this results in a less efficient device operating at higher junction temperatures and consequently lower reliability.

iii. **Low Thermal Resistance.** Higher output power ratings correspond to higher concentrations of heat in a RF transistor. This is to some degree offset by a larger die size but doubling the output power, assuming similar efficiency, will double the heat dissipation in the package. Making the package larger will not necessarily decrease the thermal resistance and will certainly compromise the performance by increasing the package parasitics.

iv. **Bandwidth.** Current transmitter power amplifier designs strive to cover the full allowable operating bandwidth for their own particular application. The benefits of lower inventory and the elimination of field tuning over split band designs outweigh the added complexity in the design and the trade-offs in performance over narrow-band tuning.

v. **Stability.** An amplifier should be stable over the full operating range expected in the field.

vi. **Load-Pull Ruggedness.** A transistor should be capable of surviving an output mis-match even when operating at the design extremes. To achieve this degree of ruggedness does involve trade-offs in both gain and efficiency.

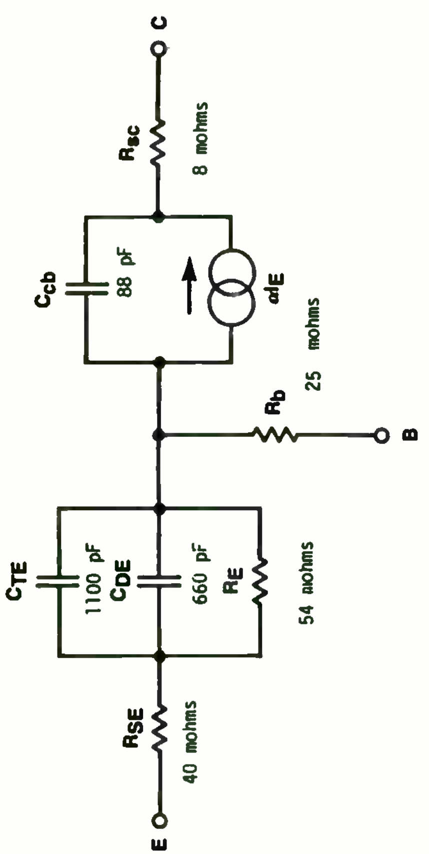
vii. **Consistent Performance.** Performance and device characteristics need to be consistent not only part to part but also batch to batch if they are to be usable by any equipment manufacturer. Inconsistency will make it difficult for the product development engineer to guarantee the final performance of his design and eventually it will lead to excessive guard-banding in the component specifications. Lower product yields when testing to a more stringent specification inevitably results in higher component cost.

Most of the above attributes are determined by the die performance and the inter-relationship of the die and the package. At higher frequencies, UHF and above, the package interface with the external circuit also becomes important. Device impedances are relatively low compared with the 50 ohm terminating impedances of an amplifier. For this reason minor variations in package position, grounding, and the values of external components can have a significant influence on the amplifier performance. In fact even minor variations in the position of the input/output shunt capacitors can easily cause an amplifier to exhibit lackluster performance.

Incorporating more of the impedance transformation network inside the package minimizes the effects of these variations and simplifies the task of the circuit designer. As an added benefit it makes the out-going RF testing by the transistor manufacturer easier, since it simplifies test fixture design and maintenance.

DIE CONSIDERATIONS

RF transistor die design is a compromise between obtaining the best performance possible in terms of power gain, saturated output power and efficiency, while still maintaining adequate ruggedness into an output mis-match, good voltage breakdowns and long term reliability. Good die yields and low production costs are also important in developing transistor die for use in Land Mobile applications.



SIMPLIFIED 'T' MODEL FOR TRANSISTOR DIE.

K4192

Figure 1

MATCHING NETWORKS

Fig. 1 is a very simplified T model of a transistor die in common base configuration. Common base is normally chosen instead of common emitter mode for class C amplifiers operating at 800MHz and above because of its higher power gain. Included in the model are the junction capacitances and the resistive losses attributed to each transistor region. The values given are typical for a 60Watt die designed for operation at 12.5Volts.

Analysis of this model at 870MHz gives us the equivalent series input impedance (Zin) and the equivalent series output impedance, (Z_{0L}), that when matched by a conjugate impedance source and load will operate at the rated output power level with minimum reflected power.

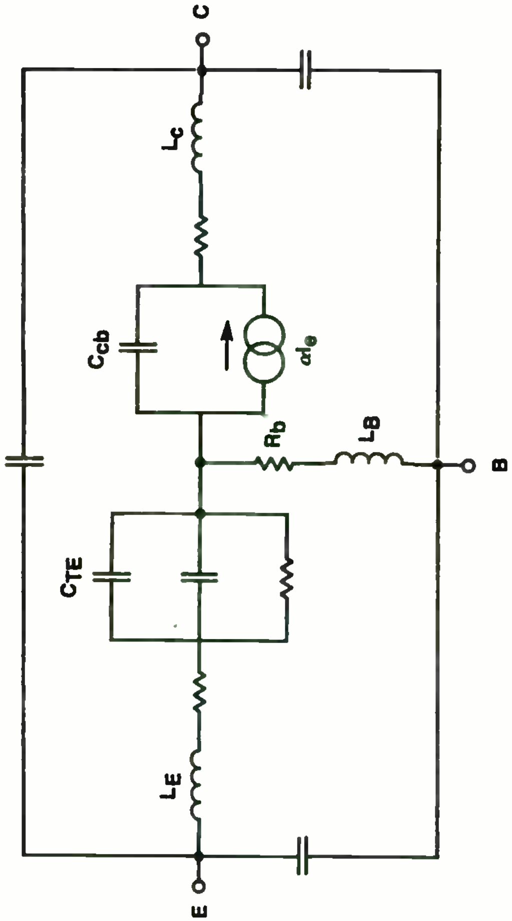
$$Z_{in} = 0.105 - j0.022 \text{ Ohms}$$

$$Z_{0L} = 0.717 - j0.38 \text{ Ohms}$$

Inspection of the series impedances given above indicates bandwidth is not inherently limited by the die below the cut-off frequency (ft). The series output impedance has the highest Q but even for this large die it is still less than 1.

RF power transistors are not generally sold in chip form but are normally assembled in packages or chip carriers before they can be usefully incorporated in discrete amplifier circuits. The package provides low resistive paths for both thermal and electrical connections. It should also provide a method of mechanically attaching the device to a heatsink. Electrical connections inside the package at high frequencies have a marked

Figure 2



Z_{IN} = 0.105 + j3.26 ohms
Z_{OL} = 0.72 + j2.90 ohms

MODEL FOR UN-MATCHED PACKAGED DIE.

K4194

effect on the performance of the transistor. Fig. 2 includes these package parasitics in the transistor model. For our purposes this model can be simplified to that given in Fig. 3 The values given are typical for a non-internally matched 800MHz package.

Analysis of this model gives the impedances at the package terminations to be :

$$Zin = 0.105 + j3.258 \text{ Ohms}$$

$$ZoL = 0.717 + j2.90 \text{ Ohms}$$

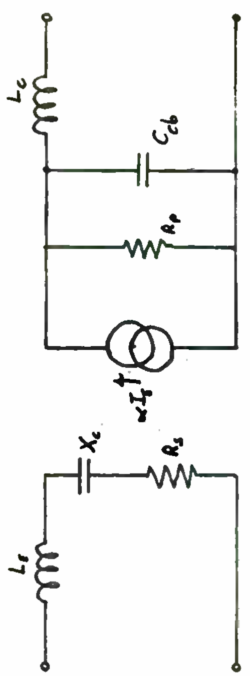
These are very low impedance levels compared to the 50ohm termination impedances of an amplifier and impedance transforming networks are essential if an amplifier is to meet its design goals. Without these networks the amplifier would exhibit :

* Poor input return loss. A large part of the drive will be reflected and thus not available for amplification by the transistor.
* Poor gain flatness and consequently limited bandwidth.
* Poor transfer of power to the load because of output mismatch.
* Instability under certain operating conditions.

Matching networks can be implemented externally but the package parasitic components will severely limit the useful bandwidth on high power devices. The inevitable losses associated with these external components and the sensitivity of the amplifier performance to component variation will also reduce the attainable bandwidth in production designs. The inherent narrow bandwidth of a packaged transistor die at high frequencies was partly solved several years ago by including part of the input network inside the package. Later further improvements were made, especially in the case of mircowave power devices, by including additional sections of input matching and output matching within the device package. The added complexity of multi-section internal matching requires the use of highly skilled labor and careful attention to detail in the assembly of these transistors. Even with these measures the product yields are relatively low compared to commercial products and consequently these parts are expensive to manufacture.

Input Network

Internal input matching performs two functions. It increases the impedances to a level that can be more readily matched by external components. Secondly, using the internal feedback inherent in the package, internal matching can be used to shape the gain- frequency response of the device. The feedback is associated with the common lead inductance and in either CE or CB configurations there will always be a small amount of common lead inductance simply due to the physical distance between the die and the grounded leads. This inductance is represented by the emitter or base die metallization, the wire-bonds from the bond pads to the lead frame and the lead-frame itself. The inductance is minimized by having several wire bonds to the die, using wide



$$R_p = \frac{(V_{cc} - V_{ce(sat)})^2}{2 P_{out}}$$

C_{cb} = 1.2 C_{obo} measured at operating volt.

FIGURE 3. DEVICE EQUIVALENT MODEL.

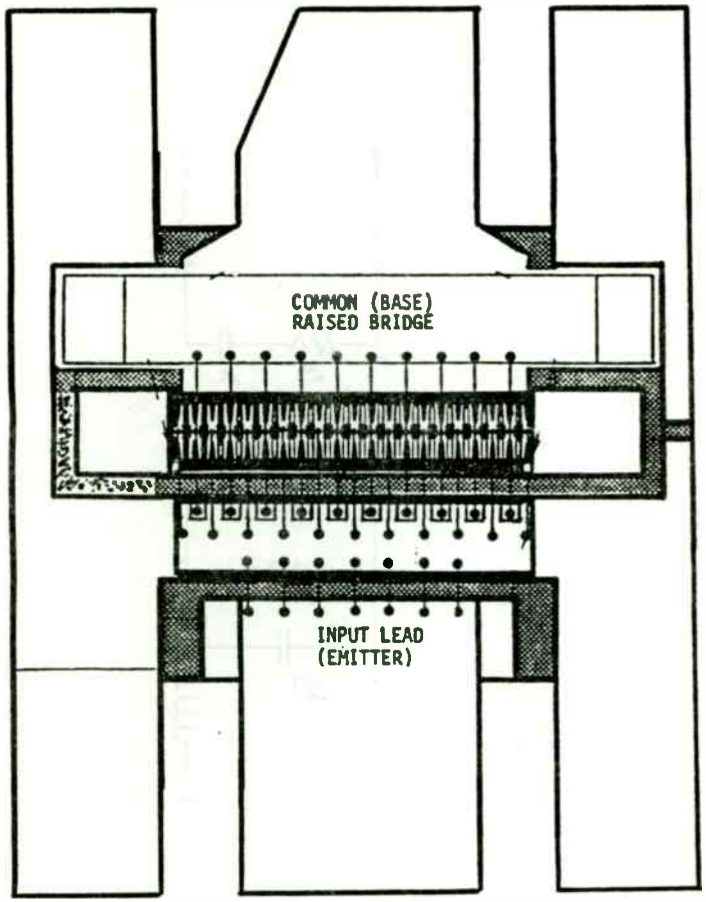


Figure 4. INTERNAL LAYOUT OF CURRENT 800 MHz PACKAGE

metallization patterns on the package and having 2 or more common leads - four is normal. This is illustrated in Fig. 4.

The self and mutual coupling that exists in a double wire bonded common base or common emitter part can be tuned to vary the gain of the transistor at a particular frequency. Using this method the 6dB gain slope for the die can be flattened over a desired frequency range.

The input impedance without matching was given earlier and is repeated here:

$$Zin = 0.105 + j3.258 \text{ Ohms}$$

Wirebond inductance and the braze area, necessary for lead attach, are responsible for the major part of the reactive component. Using current packaging techniques it would be difficult to further minimize this inductive component.

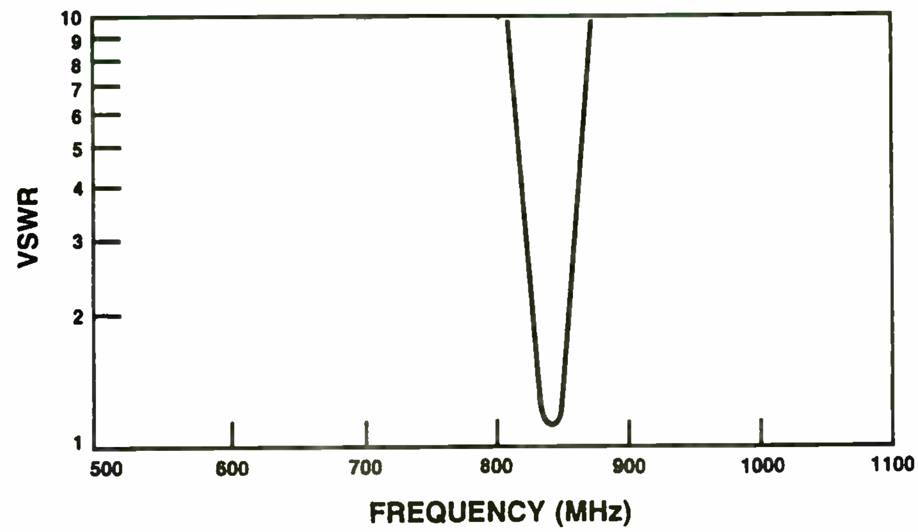
A matter of considerable importance is, however, the bandwidth over which the transistor can be operated without serious degradation in power gain and efficiency. The high Q represented by this impedance would present an insurmountable difficulty for any engineer wishing to design even a moderately broadband circuit. Additionally the high losses associated with the shunt capacitor necessary to transform the inductive reactance would severely degrade amplifier performance. This can be demonstrated using the values given in equation above. The unmatched device Q would be :

$$Q = Xs/Rs \quad \text{.................................} \quad (1)$$

$$= 3.258/0.105 = 31$$

Figure 5.

BANDWIDTH FOR NON INTERNALLY MATCHED INPUT



K4198

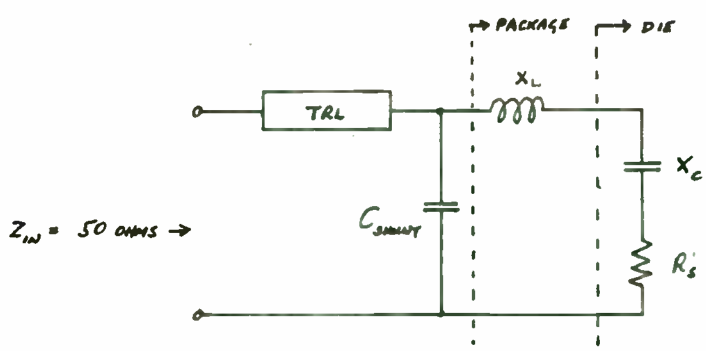


FIGURE 6. INPUT NETWORK FOR UN-MATCHED TRANSISTOR

Typical Qs for high quality chip capacitors at this frequency are in the range 100-300. This represents a gain decrease due to losses in the capacitor of between 0.9dB and 3dB. Typical gains for parts operating at 12.5Volts are 5-6dB so this does represent a significant factor in circuit performance.

The series inductance internal to the package also limits the bandwidth that can be achieved with external input matching. Fig. 5 is a plot of the frequency versus input VSWR of the input network shown in Fig. 6. This analysis assumes ideal loss-less components. The inductive reactance of the device input impedance is resonated with a single shunt capacitor at the band center. This gives the 3dB bandwidth from:

$$BW(3dB) = \frac{f_o}{Q} = \frac{f_o \, R_s}{X_s} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$= \frac{870 \times 10^6 \times 0.105}{3.258}$$

$$= 28MHz$$

The real part of the series equivalent input impedance, Rs, is inversely proportional to the area of the transistor, or more exactly the emitter periphery, which itself determines the saturated output power. This explains why low power transistors can easily be matched over several hundreds of megahertz whereas high power devices have limited bandwidth. The 3db bandwidth given in equation 2 is the theoretical maximum that can be achieved. Fano in his classic paper [1] analyzed the limitations of broadband matching a complex load. His work asserts that

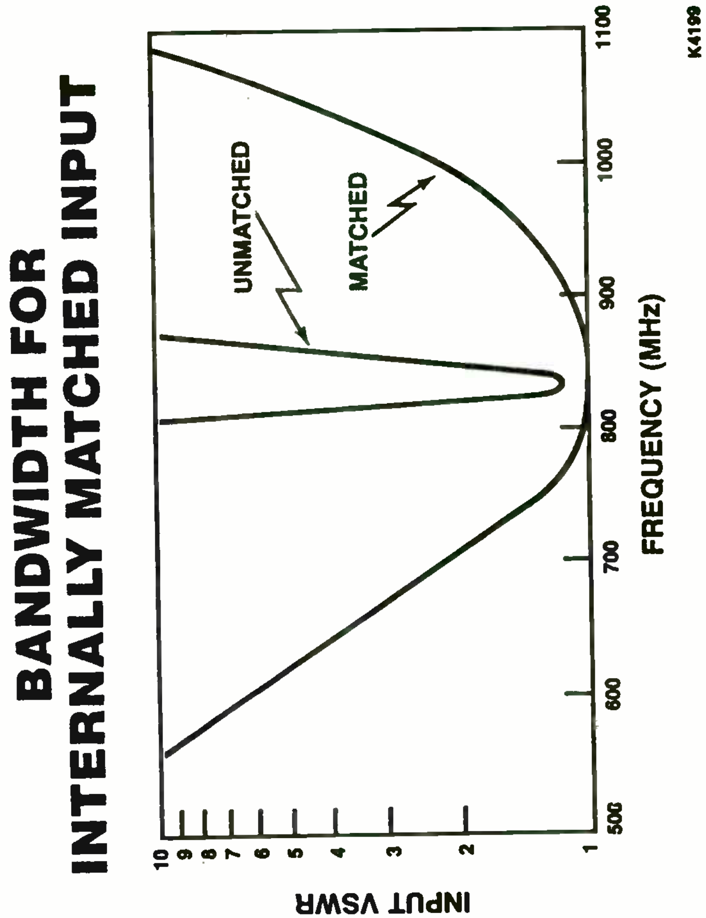


Figure 7.

increasing the number of sections does allow the 3dB bandwidth to be transformed into a nearly rectangular bandpass characteristic but no matter how complicated the network, it is never possible to match the entire available drive over a wider frequency band.

Radical improvements in bandwidth can be achieved if the series inductance is split by including a single stage of matching inside the package. Bandpass networks offer better performance than low-pass configurations using the same number of elements but low pass impedance transforming structures have a topology that can be easily integrated internally using the wirebonds for inductors and mos-capacitors for the shunt elements. Mos-capacitors can be fabricated using the same technology employed in the manufacture of transistor die and offer very low dissipation at UHF frequencies.

An alternative matching structure has been proposed [2] using a shunt-L element, inside the package, to resonate with the die input capacitance at mid-band. There are some reported advantages with this scheme namely, higher power gain, and improved collector efficiency. The shunt-L network also results in a band-pass structure with effectively zero reactance at low frequencies. This, suppresses the generation of low frequency instabilities. A major disadvantage of this matching scheme is the inability to screen the assembled device for certain DC parameters.

Fig. 7 illustrates the advantages of internal matching comparing an un-matched package with a package incorporating a single section. The input impedance measured at the device

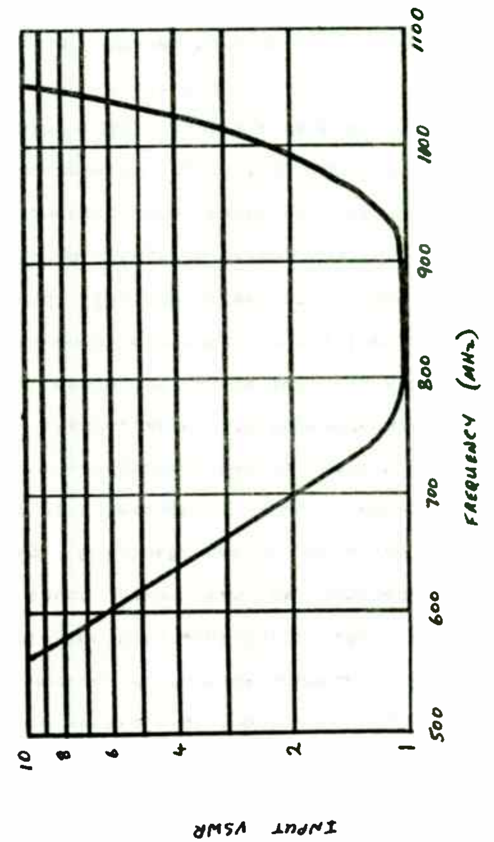
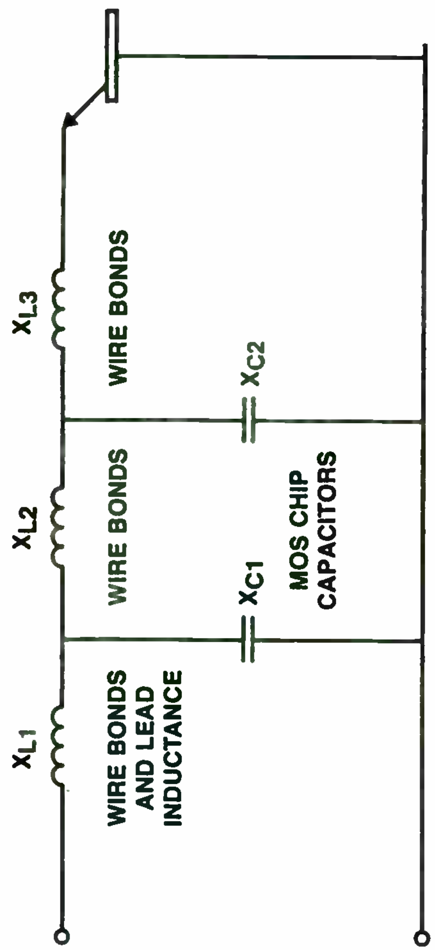


FIGURE 8. INPUT VSWR FOR DOUBLE SECTION INTERNAL MATCH.

1 LOG X LINEAR(5)

Figure 9



DOUBLE SECTION INPUT MATCHING

K4197

terminals is still relatively low but it is now practical to transform it to an higher impedance externally.

A further improvement in the input impedance can be achieved by adding a more sections of input matching. Matthaei has covered in depth the design of low-pass impedance transforming networks ideally suitable for this application [6]. With 2 sections up to 90% of the input power can be matched over the available bandwidth Fig. 8 illustrates the behavior of the double input matched device with frequency. Input impedance is now at a level where the external matching can be readily accomplished using a single section transmission line transformation Fig. 9.

Additional bandwidth can be obtained and the gain frequency response flattened by mis-matching the input at the low frequency end of the band. The 6dB/octave gain slope of the transistor die can be used to advantage to extend the low frequency response. A less than perfect input match partially reflecting the input power is compensated for by the higher device gain at lower frequencies. They do require a degree of isolation from the driver stages to prevent the low frequency mis-match affecting the stability of these earlier stages. These networks are adequately covered in the literature[3,4] and will not be further discussed here.

The input network transforms the die impedance up from 0.1 ohm to approximately the 8ohm level. The inner section conforms closely to a typical internal match seen in existing products. The outer match requires relatively high values of inductance and, because of the common base configuration, also needs to

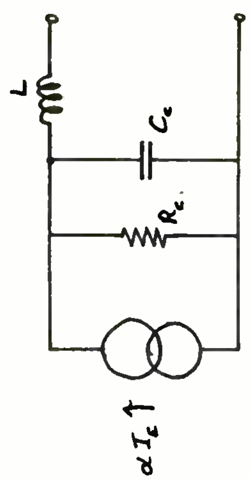


FIGURE 10. TRANSISTOR OUTPUT MODEL

carry the full emitter current with low loss. The minimum number of wires that can be used is therefore limited. The inductance is achieved by closely spacing the wires and using the mutual inductance to offset the lower self inductance of the many parallel wires.

Output Network

Reference to the transistor output model given in Fig. 10 show the collector circuit can be represented by a parallel combination of shunt capacitance(Cc) and collector resistance(Rc) and the series collector lead inductance(L). Output impedance (Zout) for this configuration is given by[8]:

$$Z = \frac{R_c}{1+(\omega_o C_c R_c)^2} + j\left[\omega L - \frac{(\omega R_c C_c)^2}{\omega C_c[1+(\omega R_c C_c)^2]}\right] \quad \ldots\ldots\ldots(3)$$

If Cc is the dominant reactive term the intrinsic Q for the output network is given by:

$$Q = \omega R_c C_c \quad \ldots\ldots\ldots(4)$$

If the inductive term dominates which it normally does for high power transistors, then:

$$Q = \frac{\omega L}{R_c}\left[1+(\omega R_c C_c)^2\right] \quad \ldots\ldots\ldots(5)$$

The maximum available output bandwidth becomes:

$$B.W. = \frac{f_o}{Q}$$

$$= \frac{1}{2\pi R_c C_c} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

if C dominates.

$$B.W. = \frac{R_c}{2\pi L[1+(\omega R_c C_c)^2]} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

if L dominates.

The value of collector resistance, Rc, can be calculated approximately, at high frequencies by:

$$Rc \approx \frac{1}{\omega_t C_c}$$

Therefore if L dominates :

$$B.W. = \frac{f_t}{L\, C_c(\omega_o^2 + \omega_t^2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

= 106MHz.

This network could be conjugately matched for maximum power transfer but half of the power would be dissipated in the collector resistance limiting the maximum efficiency to 50%. Additionally, perfect matching will not necessarily allow the transistor to reach its full output power capability because of current saturation effects. The internal collector resistance for a class-C amplifier is also highly non-linear and varies over a wide range as the transistor oscillates between saturation and cut-off during each RF cycle. In fact the shunt collector resistance is maximized during product development by the suitable selection of epitaxial resistivity and epitaxial thickness consistent with meeting the required collector breakdown voltages. High shunt collector resistance maximizes the efficiency and saturated power capability.

RF power transistors are normally operated with a collector load-line determined by assuming the maximum collector voltage swing during the device turn-off period will be twice the supply voltage. The load-line impedance can be approximated by the equation:

$$Rp = \frac{(Vcc - Vce(sat))^2}{2 \times Pout} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

This equation holds good for frequencies less than the cut-off frequency for the die (f_t). If we ignore the collector resistance (Rc) the matching problem simplifies to the collector capacitance shunted by Rp. Limitations of broadband matching for this load configuration have been analytically described by Bode[5].

We can apply Bode's resistance or attenuation integral

FIGURE 12, SERIES RESISTANCE FOR NETWORK AT THE INTERNAL COLLECTOR NODE.

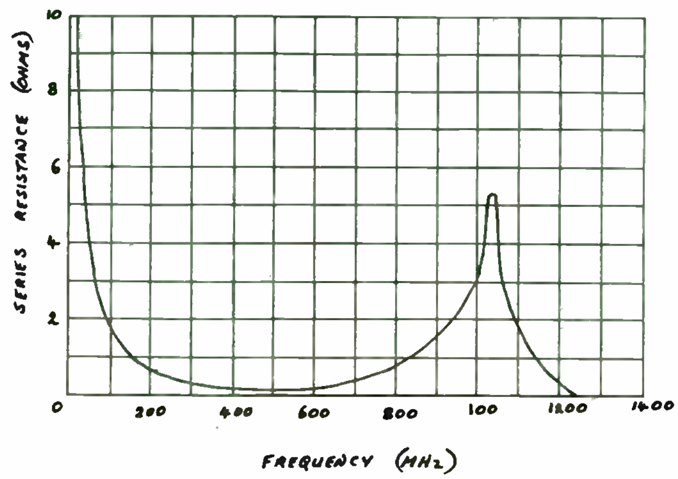
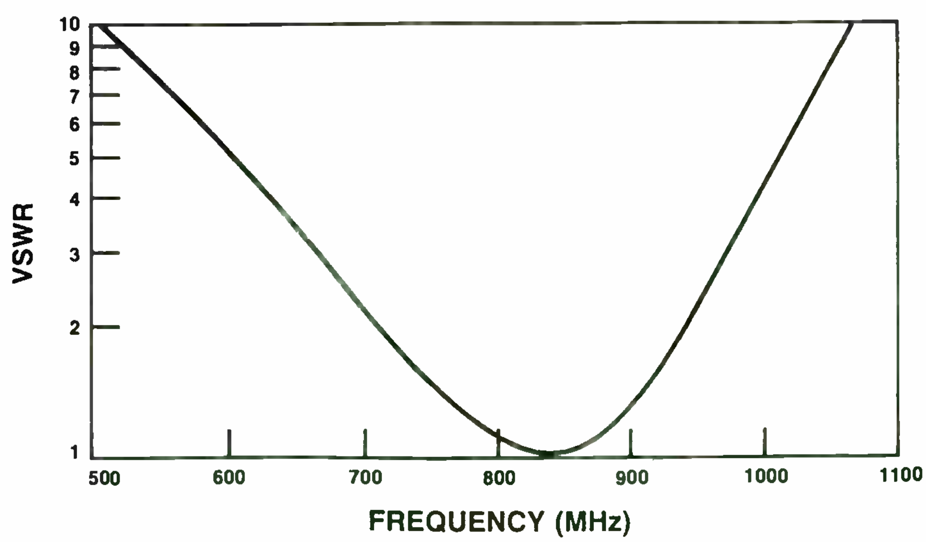


Figure 12b.

BANDWIDTH FOR NON-INTERNALLY MATCHED OUTPUT



K4200

theorem to estimate the available bandwidth for the transistor die neglecting the limitations of the package inductance:

$$\int_0^\infty R\,d\omega = \frac{\pi}{2C} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

This expression applies to any minimum reactance network including a leading parallel capacitor where the source resistance can be considered substantially infinite. Capacitance is estimated to be 1.2 times Cobo. The multiplication factor was empirically determined by comparing measured impedance data with an optimized model of the die and the package parasitic elements and has been confirmed for a number of UHF and 800MHz transistors. Using the modified capacitance value the constant resistance integral can be rewritten as :

$$\int_0^\infty R\,d\omega = 1.48 \times 10^{10} \text{ ohm.rad/s}$$

$$\text{or} \int_0^\infty R\,d\omega = 2.36 \times 10^{9} \text{ ohm.hertz}$$

Analyzing the network shown in Fig. 11 the series input resistance can be plotted for all frequencies (Fig 12). It is apparent from the graph that bandwidth is lost outside the frequency range we desire especially below 300MHz.

If the area under the curve is integrated the result should correspond to the bandwidth-resistance product calculated from the resistance integral. It can readily be seen that by restricting the area under the curve to the operating frequency

FIGURE 13, SERIES RESISTANCE AT INTERNAL COLLECTOR NODE FOR OUTPUT NETWORK INCLUDING SHUNT-L

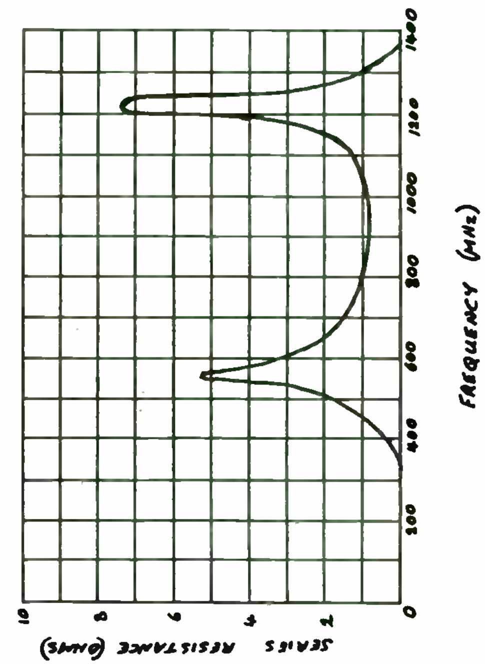
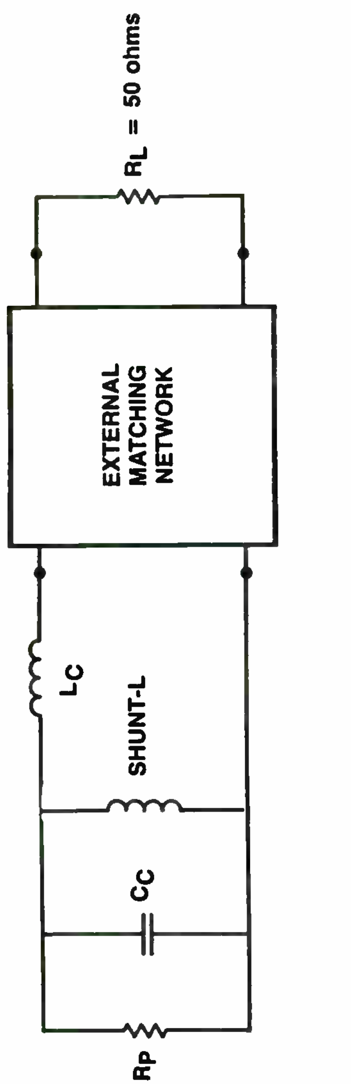


Figure 14



COLLECTOR SHUNT-L MATCHING

K4201

range and loading the internal collector node with the calculated load-line impedance the ultimate bandwidth is realized. At all frequencies outside the operating band, the series resistance seen by the internal collector terminal would need to be zero. The design of a network to match the available bandwidth would be impractical but typically only a fraction of the absolute bandwidth is normally required.

The above requirement on resistive behavior at the collector can be best met by adopting an ideal bandpass network that provides very abrupt transitions through zero resistance outside the operating range. Practical considerations, as in the case of the input network, limit the circuit topologies that can be incorporated inside the package. The un-matched case can be improved upon by some relatively simple internal changes to the package metallization which allow the die-bonding of an additional output mos-capacitor.

Tuning out the collector-base capacitance at mid-band using a shunt-L element remarkably improves the usage of the available frequency-resistance product. This is clearly illustrated in Fig 13. The series resistance has been re-plotted for the new network show in Fig. 14. Maximum broadband power transfer is enhanced by this type of network but more important the impedance match is improved over the operating bandwidth. Efficiency, which has a greater sensitivity than gain to reactive loading at the internal collector node, does not suffer the roll-off at lower frequencies that would be seen with an un-matched design. Fig. 15 is a comparative plot of normalized parallel reactance (|Xp/Rp|)

2 LOG X LINE (5)

FIGURE 15, NORMALIZED REACTANCE AT INTERNAL COLLECTOR NODE FOR OUTPUT MATCHING NETWORK

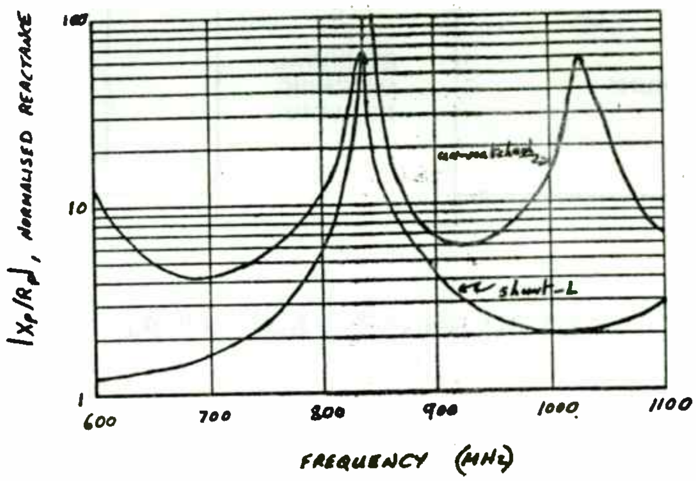
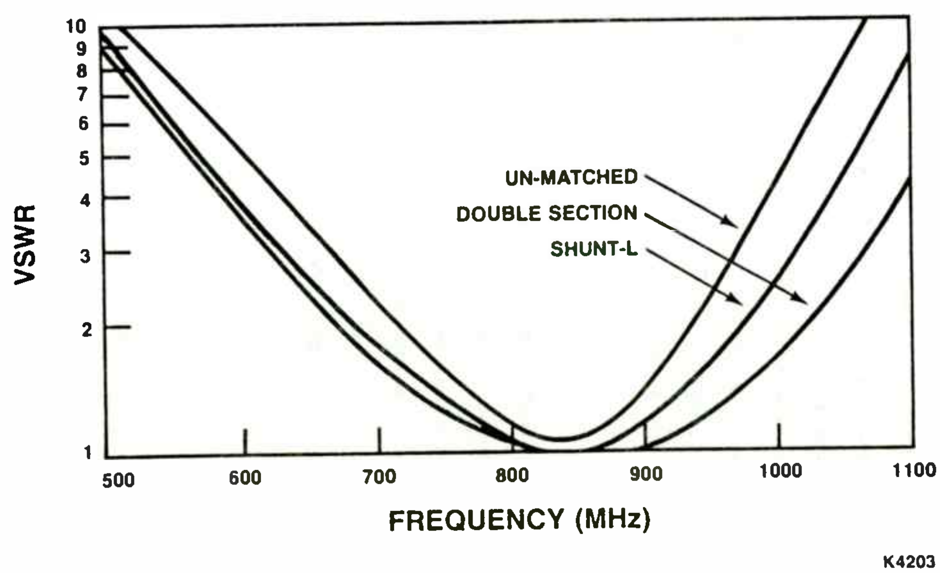


Figure 16.

IMPROVEMENT IN BANDWIDTH WITH OUTPUT MATCHING



for a shunt-L network and a conventional un-matched transistor. It can be seen that the shunt reactive component for the shunt-L match is higher at the low end of the band than in the case of the un-matched device. For a good match the reactance should be at least twice the parallel resistive component ($|Xp/Rp| > 2$) within the operating band [7].

Again as in the case of the input network an additional section of low-pass transformation can be included to further increase the impedances to a level which eliminates the need for an external shunt-C. Fig. 16 shows the bandwidth attainable with the network shown in Fig. 17.

External Matching Requirements

The high impedance levels present at the terminations of this package do greatly simplify the external matching requirements. The device can be matched to 50ohms with a single section transmission line with a characteristic impedance in a range that can be readily fabricated. The photograph (Fig. 19) and the circuit schematic (Fig. 20) of the broadband fixture used for device evaluation illustrate the simplicity of the external matching. The elimination of the troublesome shunt capacitors close to the transistor package does simplify the production and enhance the consistency of the amplifier performance. A paper to be given in a latter session of this seminar will describe the design of an amplifier using 4 MRF898 parts paralleled to produce 220Watts output power over 850-900MHz bandwidth. This design was

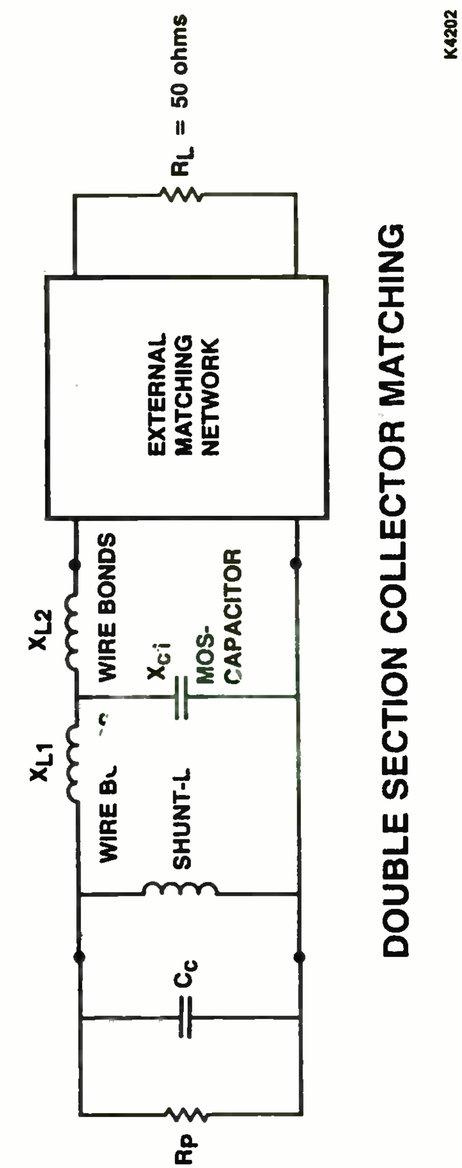
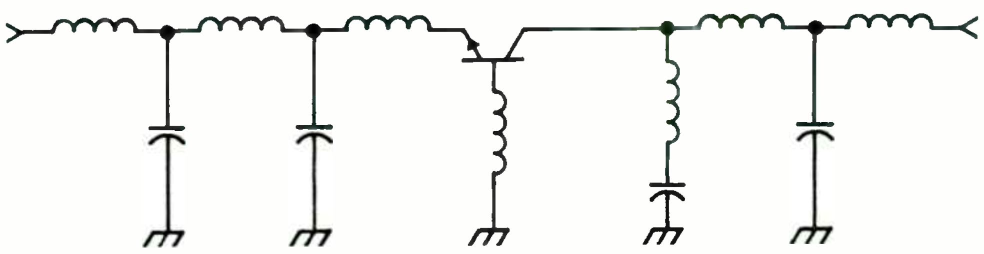


Figure 17

Figure 18b.

SIMPLIFIED MRF898 SCHEMATIC DIAGRAM



K4217A

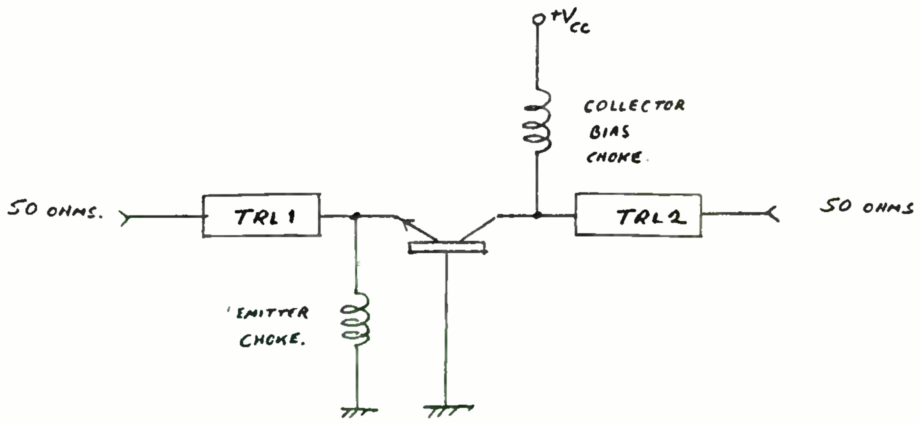
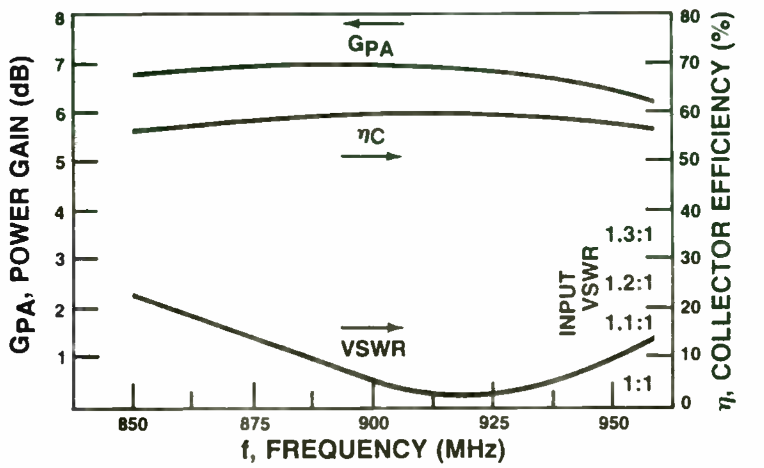


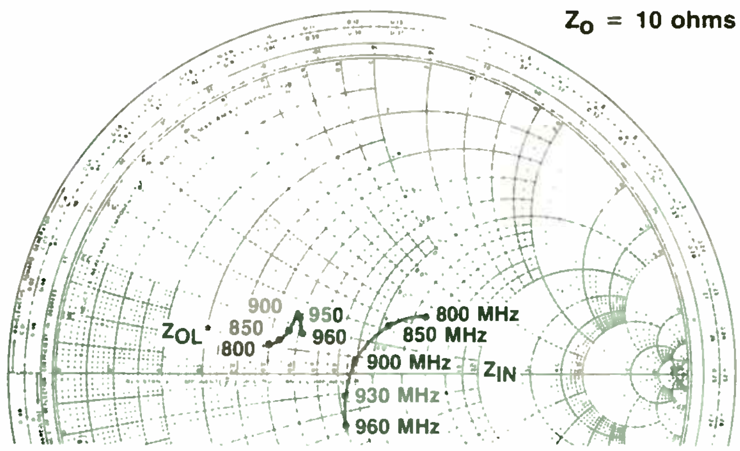
FIGURE 20. EXTERNAL MATCHING AND DC BIASING.

MRF 898
TYPICAL BROADBAND CIRCUIT PERFORMANCE



K4208

IMPEDANCE OR ADMITTANCE COORDINATES



FREQ. (MHz)	Z_{IN}	Z_{OL}*
800	12.9 + j5.1	5.0 + j1.0
850	10.4 + j3.2	5.3 + j1.2
900	8.9 + j0.7	5.6 + j1.6
930	8.3 − j1.2	5.8 + j2.4
960	8.0 − j2.8	6.1 + j1.7

P_o = 60 WATTS,
V_{CC} = 24 V

*Z_{OL} = Conjugate of optimum load impedance into which the device operates at a given output power, voltage and frequency.

K4210

CONCLUSION

The design of RF power devices for high power, high frequency operation involves a number of compromises. Most of which have been outlined above. The important points are that the added integration of additional matching inside the package can have the multiple benefits of easier usage, improved performance and better testability.

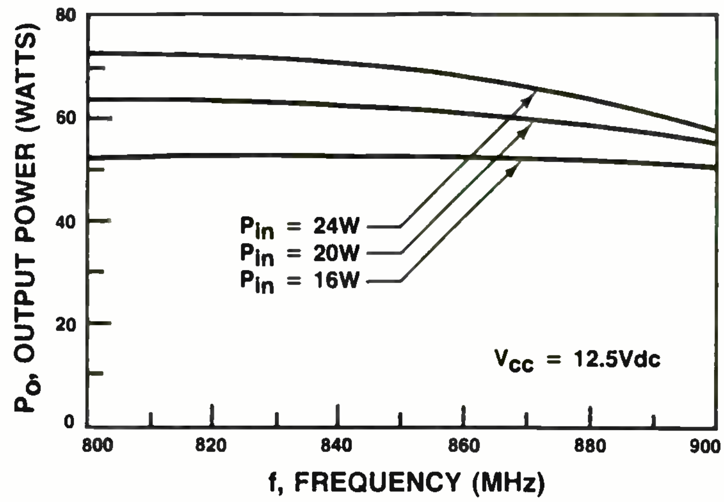
Conventional single input-matched parts will continue to be used at lower power levels but at higher power and higher frequency innovated product design is needed if devices are to be of practical value.

The package design outlined here offers several advantages over conventional 800MHz packaging:-

* **SIMPLER EXTERNAL MATCHING** - Higher device impedances eliminate the need for critical shunt-C capacitors and allow single section transmission line matching.
* **HIGHER EFFICIENCY** - High performance die and the use of shunt-L collector matching enable high efficiency (>60%) to be maintained over a greater bandwidth.
* **BETTER THERMAL PERFORMANCE** - Larger package and higher operating efficiencies result in lower thermal resistance.
* **WIDER BANDWIDTH** - Internal matching minimizes the effects of package parasitics allowing broader bandwidth and a minimum of variation in gain and efficiency across the operating band.

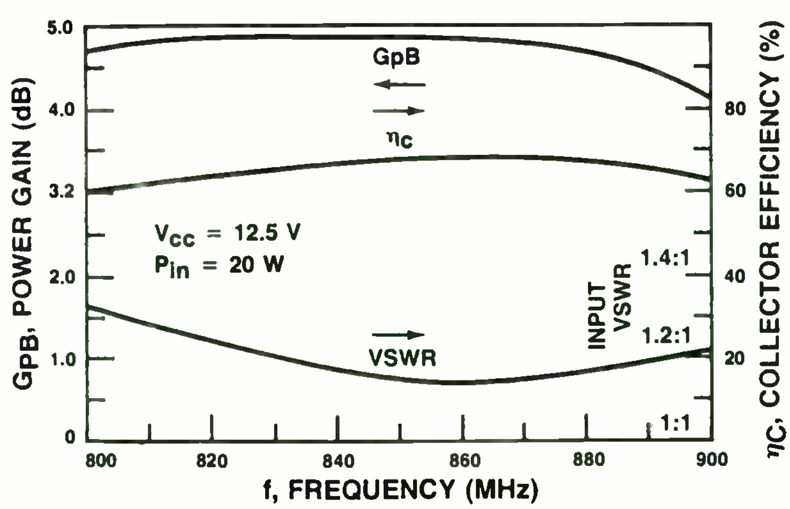
What of the future? Operation in excess of 100Watts output power at 900MHz has already been demonstrated with no changes required in the external matching. In fact this package concept can be extended to products operating at both higher and lower output powers than the examples given and the design is also feasible for products in the 400-512MHz land mobile band.

MRF 848
OUTPUT POWER versus FREQUENCY



K4205

MRF 848
PERFORMANCE IN BROADBAND CIRCUIT



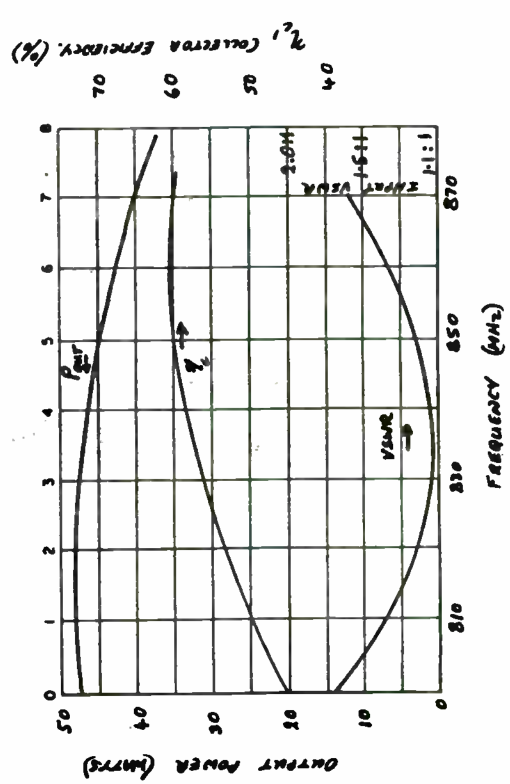
K4206

rapidly executed because the simple external circuitry operated first time with performance close to the design goals and the amplifier required a minimum of further tweaking.

PERFORMANCE

The accompanying graphs illustrate the performance of both the 12Volt and the 24Volt versions of these devices. Noteworthy is the flatness of the gain and efficiency response across the design bandwidth and the extension of this outside the normal frequency range of interest. For comparison Fig. 21 shows the broadband performance for the MRF846 40Watt device assembled in a conventional 800MHz package. Efficiency at the low end of the band is greatly reduced and the input vswr at the band-edges is much higher than the MRF848. The package size and higher efficiency result in both new transistors having a thermal resistance less than 1 Watt/°C.

Figure 21, Typical Broadband Performance - MRF846



REFERENCES

[1] R. M. Fano, "Theoretical limitations on broadband matching of arbitrary impedances", J. Franklin Inst., vol 249, pp 57-83, 139-155.

[2] E. J. Colussi, "Internally Matched RF Power Transistors", Microwave Journal, April 1978, pp 81-84.

[3] O. Pitzalis Jr and R. A. Gilson, "Tables of Impedance Matching Networks Which Approximate Prescribed Attenuation versus Frequency Slopes", IEEE Trans. MTT-19, pp 381-386, Apr. 71.

[4] W. H. Ku and W. C. Petersen, "Optimum Gain-Bandwidth Limitations of Transistor Amplifiers as Reactively Constrained Active Two-Port Networks", IEEE Trans. on Circuits and Systems, vol. CAS-22, No. 6, June 75, pp 523-533.

[5] H. W. Bode, "Network Analysis and Feedback Amplifier Design", New York : Van Norstrand, 1945

[6] G. L. Matthaei, "Tables of Chebyshev Impedance-Transforming Networks of Low-Pass Filter Form", Proc. IEEE., Aug. 1964, pp. 939-963.

[7] O. Pitzalis, Jr and R. A. Gilson, "Broad-band Microwave Class-C Transistor Amplifiers", IEEE Trans. MTT-21, no. 11, pp 660-668, Nov. '73

[8] R. E. Hejmanoski, et al "Broadband Combinable Power Amplification 1350-1850MHz", ECOM-0042-1, Aug. '72

THE HYBRID POWER AMPLIFIER MODULE FOR CELLULAR RADIO TELEPHONE DESIGNS

By Norman E. Dye and James P. Oakland
Motorola, Inc.
Phoenix, Arizona

Introduction

In 1977, Motorola Semiconductor began design efforts on what was to be a family of 800 MHz hybrid power amplifier modules for use in the land-mobile-telephone industry. Conventional and trunked commercial FM radio designs at 800 MHz were seen as major targets, but cellular telephone was the bright star in the future - the area in which to concentrate development efforts. This paper highlights the design, construction, performance and reliability of the module used in many of the 800 MHz cellular mobile telephones being manufactured today.

In the last 14 years Motorola has designed, built and sold a variety of power amplifier modules for use in the VHF (136-174 MHz), UHF (400-512 MHz) and the 806 to 950 MHz frequency bands. The majority of these modules have gone into VHF and UHF portable and mobile commercial FM radios manufactured both domestically and abroad. As shown in Figure 1 however, the list of demonstrable applications for the hybrid P.A. module is not limited. The most recent applications are, of course, the 800 MHz radio designs and in this area the use of modules in radio P.A. sections is becoming the industry standard. Many radio manufacturers have abandoned conventional discrete designs in favor of the hybrid module approach for the reasons summarized in Figure 2.

The U.S. cellular mobile application in the frequency range from 825 to 845 MHz has received the most attention due to the enormous business potential represented. The MHW808A1 module shown in Figure 3 was designed specifically for this application and is compatible with the unique cellular system requirements as they relate to the P.A. section of the radio telephone. Shown in Figure 4 is a block diagram of the module incorporated into a typical cellular radio P.A. section.

The module has three stages of gain with the overall gain adjustable over a full 30 dB range by controlling the supply voltage to the first stage. A complete summary of the electrical specifications and typical RF performance curves are illustrated in Figures 5 through 8. An important feature to note is the excess bandwidth capability of the module as witnessed by the 806 to 870 MHz bandwidth specification and the even more broadband characteristics shown in the typical RF performance curves. In fact, the MHW808A2, which is identical to the MHW808A1 except for a 806 to 890 MHz bandwidth specification, is sourced from the same product line. The relatively broad bandwidth of the module is the result of design - not chance. Experience has shown narrowband designs to be more prone to instability under source and load mismatch conditions. Additionally, the output power versus input power curves shown in Figure 7 are for general interest and are not included to suggest gain control via RF drive adjustment. As stated above, the module is designed for gain control via supply voltage adjustment of the first gain stage. Most significantly, however, the module is designed to maintain stability (i.e., no spurious outputs) over the complete range of definable radio operating conditions which impact on the module in the form of varied levels of RF drive, supply voltage, gain control voltage, source and load mismatch, temperature and, especially true for the cellular application, RF output.

A circuit schematic for the module is shown in Figure 9. The three active devices feature high figure-of-merit geometries to assure maximum stage performance and gold top metal for high reliability. To enhance overall module stability, all three stages operate in the common-emitter configuration. The first stage, the gain control stage, is biased for large signal class-A operation and the last two stages for class-C operation featuring threshold base bias with temperature compensating Schottky diodes. The threshold bias sets the average emitter-base junction potential at 0.35 to 0.4 volts and is used primarily to eliminate stability problems at low levels of RF drive when the transistors are just beginning to turn on. Schottky diodes with the inherently low forward voltage characteristic are compatible with the threshold bias voltage range and provide satisfactory temperature compensation for power degradation at reduced temperatures. In Figure 10, the benefits of threshold bias are illustrated along with a graph depicting the forward voltage versus temperature characteristic for the 1N5817 Schottky diode at 50 mA forward current.

The virtual ground concept has been employed at the emitter of the output transistor to minimize the deleterious effects of common lead impedance on RF gain associated with common-emitter operation. The emitter wirebonds are not returned directly to ground, but first to the bottom side of the first shunt capacitor at the base and at the collector, and then to ground through relatively large isolation inductances. The net effect is to isolate the first transformation loop at the base and collector from ground. As configured, a virtual ground is established at the emitter wirebond pads on the transistor chip which, in effect, eliminates common lead inductance and allows a stage gain equaling that achievable with common-base operation without sacrificing module stability.

In addition to the input and output impedance levels being 50 ohms, each interstage is also designed to be at the 50 ohm level to facilitate testing of individual stages and to provide a cascadable, gain block option. In general, low-pass Chebyshev impedance matching networks as shown in Figure 11 are used to transform the low base and collector impedance levels to the terminal or interstage level of 50 ohms. Listed in Figure 12 are the considerations for determining the number of matching network sections required. Consider for example, the real part of the base impedance for the third stage transistor is approximately 0.35 ohm - a transformation ratio of nearly 140 when transformed to 50 ohms. In Figure 13 the circuit element values and corresponding passband transducer loss curves are shown for 1, 2, and 3 sections of matching for transforming the 0.35 ohm base impedance to 50 ohms over the 800 to 900 MHz operating frequency band. In this design example, both bandwidth and transformation ratio are known, fixed values. If minimum transducer loss in the passband was the primary criterion for selecting a circuit configuration, then the choice of two sections with a maximum of 0.007 dB loss would seem most reasonable and, in fact, two sections are used in the actual design.

In Figures 14 and 15 are a listing of the various types of inductors used in the circuit realization and a photograph illustrating specific examples of each type. The primary factors for determining which inductor type to be used are the inductance value required and the circuit area available. At 800 MHz with high transformation ratios it is not uncommon to encounter inductance values less than 1 nH. In the design example above, the first series inductor at the base of the third stage transistor ranged in value from 0.12 nH to 0.78 nH depending on the number of matching sections used in the design. Inductance values in this range are necessarily realized by accurately controlling the height and length of a wirebond or wirebond array and, in this particular case, the wirebond array used to access the base wirebond pads on the transistor. Array heights are controlled by using glass rod forms with diameters specified to within ± 1 mil and, more recently, with the use of sophisticated automated wirebond equipment.

Larger values of inductance are realized with lengths of electrically short microstrip transmission line formed on the ceramic circuit substrate. Shown in

Figure 16 is a sketch of the microstrip configuration and a graph illustrating characteristic impedance and inductance per unit length as a function of line width for a common ceramic material - aluminum oxide. The non-linearity of the relationship between line width and inductance per unit length can often be used to the designer's advantage in selecting an optimum physical realization of a specific inductance value.

Although not used in the MHW808A1 design, airwound coils are also used to realize the larger inductance values. Many of the VHF and UHF designs incorporate the airwound coil.

The capacitor types used in the module realization are listed in Figure 17. In addition, a photograph illustrating examples of each type is shown in Figure 18. In general, all capacitors used for RF impedance matching purposes are MOS. The cost effectiveness and piece part control afforded by internal manufacturing and the relatively high Q of the MOS structure are the primary reasons for this choice. MOS capacitors are currently manufactured with capacitance values in the range from 0.5 to 750 pF using silicon dioxide as the dielectric insulator. Silicon nitride has also been used where higher values of capacitance per unit area are required to minimize capacitor plate area.

In situations where even higher capacitance values are required ceramic chip capacitors are utilized. NPO ceramic chip capacitors available in the value range indicated for MOS capacitors can be used for RF impedance matching elements but are physically larger and, in general, exhibit a lower Q. However, for supply bypassing and dc blocking, which typically dictate larger capacitance values to accommodate both a RF and a low frequency function, high dielectric constant ceramic chip capacitors are the ideal choice. For some applications, most notably the dc blocking application, high capacitance values can be simulated at a specific frequency by selecting a lower value chip capacitor that is self-resonant at the frequency of interest. Finally, in those situations where high-pass impedance matching networks are used, series capacitors used for RF impedance matching elements serve a dual function and also provide dc blocking. This technique is employed at the collector of the first stage transistor in the MHW808A1 module where, not only does the series capacitor serve a dual function, but the shunt inductor is used as a RF matching network element and the means for bias insertion.

Shown in Figure 19 are the two most common approaches to bias insertion - one of which is the technique discussed in the preceding paragraph regarding high-pass matching networks and the dual function of the circuit elements. The other approach, the approach used for the second and third stages in the MHW808A1 design, utilizes a separate section of high impedance transmission line appropriately bypassed at the end closest to the supply voltage access terminal. The line length is generally selected to yield a high shunt impedance at the point of insertion in the circuit, but can be chosen to present a shunt inductance value capable of parallel resonating the output capacitance of the transistor involved.

Resistive elements are formed using the evaporated nichrome adhesion layer located between the ceramic substrate and the metal conductor. Nichrome is exposed by selectively etching the metal conductor and is laser trimmed to within 1% of the desired resistance value.

To bridge the gap between design concepts and reality, a "thumb-nail" sketch of the module construction is in order. Tooled ceramic substrates purchased from an outside vendor are metallized using thin-film techniques. Topside metal conductor patterns are defined using photolithographic procedures and plated to final thickness. Nichrome resistors are laser trimmed to value. The substrates are laser scribed and broken into individual circuit boards and at this point are ready for subsequent assembly operations. All active devices and MOS capacitors are attached to the circuit boards using gold-silicon eutectic bonding. Other components including leads, bridges, ceramic chip capacitors and diodes are soldered to the circuit board using gold-tin solder preforms reflowed in a hydrogen furnace. The circuit board is wirebonded and at this point is a completed circuit board subassembly. The appropriate combination of circuit board subassemblies is soldered to a plated copper flange using a low temperature indium based solder and the nearly complete module is ready for initial RF testing using internally designed, fully automated RF test equipment. If the module passes the initial RF test, it is conformally coated and a plastic cover is attached. The module is than RF tested once again, packaged and shipped to the warehouse

The MHW808A1 module has recently undergone several major cost reduction steps. As originally designed, the first stage circuit was assembled on a single piece of aluminium oxide ceramic measuring 0.3" x 0.65" and the second and third stages were build on a single piece of beryllium oxide ceramic measuring 0.6" X 0.65". The information shown in Figure 20 highlights the major differences between the two ceramics that impact on module design and cost. The choice of BeO as a substrate material for stages two and three was based solely on thermal considerations with the thermal conductivity of BeO being five to six times that of Al_2O_3. However, the limited availability and the extremely high cost of BeO forced a redesign effort aimed at utilizing alternate assembly procedures which minimize the use of the BeO ceramic. The photograph shown in Figure 21 features the original or conventional design contrasted with the newer cost reduced design which resulted from this effort. In the cost reduced design BeO usage is limited to two small carriers measuring 0.08" x 0.15" to which the active die are mounted. The remaining circuitry for stages two and three is constructed on two pieces of Al_2O_3 ceramic with the overall assembly dimensions being the same as the conventional design. In summary, BeO usage was reduced from 0.39 square inches to 0.024 square inches - a factor of 16.25.

The second cost reduction feature is closely linked to the first and involves the ceramic thin-film metalization scheme. Historically, best results in metalizing BeO using thin-film processes have been achieved with an evaporated nichrome

adhesion layer followed by evaporated gold and then plated gold. For Al_2O_3, a less expensive metal scheme, the components of which are evaporated nichrome, evaporated copper, plated copper, plated nickel and a gold flash plating yields very satisfactory results and allows the use of copper in place of gold as the primary conductive metal. For the microstrip structure the metal closest to the ceramic surface is the most critical. Shown in Figure 22 is a sketch of the microstrip cross-section and a corresponding graph illustrating the normalized current density in the top and bottom side metal as a function of penetration depth measured in units of skin depth. From the graph a metal thickness of four or five skin depths closely approximates the maximum penetration depth into the conductor and can be used as a specification for minimum metal thickness. The chart presented in Figure 23 lists skin depth information in microinches for copper and gold at several frequencies. For the 800 MHz module designs, four to five skin depths represents 350 to 500 microinches of thickness dependent upon the metal chosen. This is a significant metal usage and further validates the cost effectiveness of using Al_2O_3 in place of BeO.

The third, and last, cost reduction feature is the implementation of automated wirebonding. Prior to automated wirebonding, manual wirebonding using glass rod forms on critical wirebond arrays was the single most labor intensive operation in the module construction. Commercially available equipment was purchased and specially modifed to accommodate the complex wirebond scheme and is fully operational at this time. It is estimated the wirebond time per module has been reduced from eight minutes to less than forty seconds.

All of these cost reduction features combined with offshore assembly in Motorola's Seremban, Malaysia facility have resulted in module selling prices that are extremely competitive with the pricing of discrete transistor lineups required to build the equivalent amplifier.

As always, reliability is a key issue with radio manufacturers. The radio designer must feel confident the component selections he makes will not result in unexpected reliability issues and adversely affect the salability of the end product, in this case the radio telephone. Listed in Figures 24 and 25 are the reliability tests completed for both the conventional and cost reduced designs. With exception to the thermal shock testing, which is a destruct test designed to detect such problems as ceramic fracturing under extreme temperature stress, sample groups of modules were subjected to each test with before and after data recorded to identify failures. These tests were performed under the supervision of Motorola's Reliability and Quality Assurance organization and copies of the test conditions and verified test results are available upon request.

FIGURE 1

APPLICATIONS FOR HYBRID POWER AMPLIFIER MODULES

(136 MHz TO 174 MHz, 400 MHz TO 512 MHz, 806 MHz TO 950 MHz)

1. MOBILE AND PORTABLE CELLULAR RADIO TELEPHONE
2. CONVENTIONAL/TRUNKED FM TWO-WAY COMMERCIAL RADIO
3. PAGING SYSTEMS
4. UTILITY USAGE MANAGEMENT
5. ALARM SYSTEMS
6. BIO-MEDICAL ELECTRONICS
7. WILDLIFE TRACKING SYSTEMS
8. RF DATA TRANSMISSION LINKS
9. MARINE RADIO SYSTEMS
10. AMATEUR RADIO
11. REMOTE SURVEILLANCE LINKS

K4096-1

MOTOROLA INC.

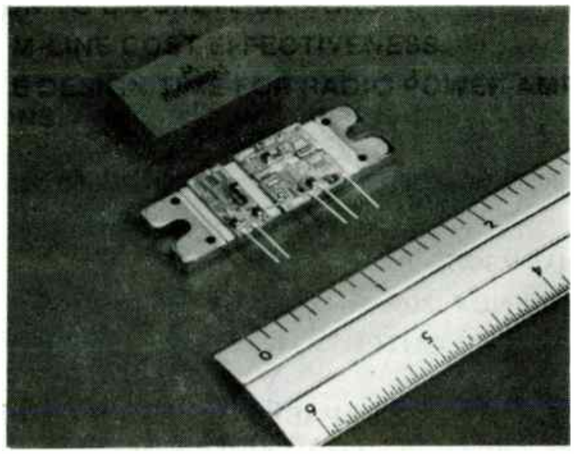


FIGURE 3 — MHW808A1

FIGURE 2

CONSIDERATIONS FOR CHOOSING HYBRID APPROACH OVER CONVENTIONAL DISCRETE DESIGN APPROACH

1. INHERENT SIZE ADVANTAGE.
2. GUARANTEED "SUB-SYSTEM" PERFORMANCE (i.e., BANDWIDTH, INPUT RETURN LOSS, GAIN, EFFICIENCY RUGGEDNESS, STABILITY, HARMONIC SUPPRESSION, DYNAMIC RANGE).
3. HIGH LEVEL OF PROVEN RELIABILITY.
4. MINIMIZE CLOSE TOLERANCE ASSEMBLY PROCEDURES COMMON TO DISCRETE DESIGNS.
5. BOTTOM-LINE COST EFFECTIVENESS.
6. REDUCE DESIGN TIME FOR RADIO POWER AMPLIFIER SECTIONS.

K4095

FIGURE 5

ELECTRICAL SPECIFICATIONS FOR MHW808A1

(CONDITIONS: P_{IN} = 30mW, P_{OUT} = 7.5W,
V_{S1} = V_{S2} = 12.5V, $V_{CONT} \leq$ 12.5V)

BANDWIDTH: 806 TO 870 MHz

GAIN: 24dB

MAXIMUM INPUT VSWR: 2:1

MINIMUM EFFICIENCY: 30%

MAXIMUM HARMONIC OUTPUT: -42dBc @ $2f_o$, -60dBc @ $3f_o$ AND HIGHER

LOAD MISMATCH STRESS: CAPABLE OF WITHSTANDING P_{OUT} = 9W, V_{S1} = V_{S2} = 16V, LOAD VSWR = 30:1

POWER DEGRADATION WITH TEMPERATURE: LESS THAN 1.7dB FROM -30° C TO +80° C
(REFERENCE: 7.5W @ +25° C)

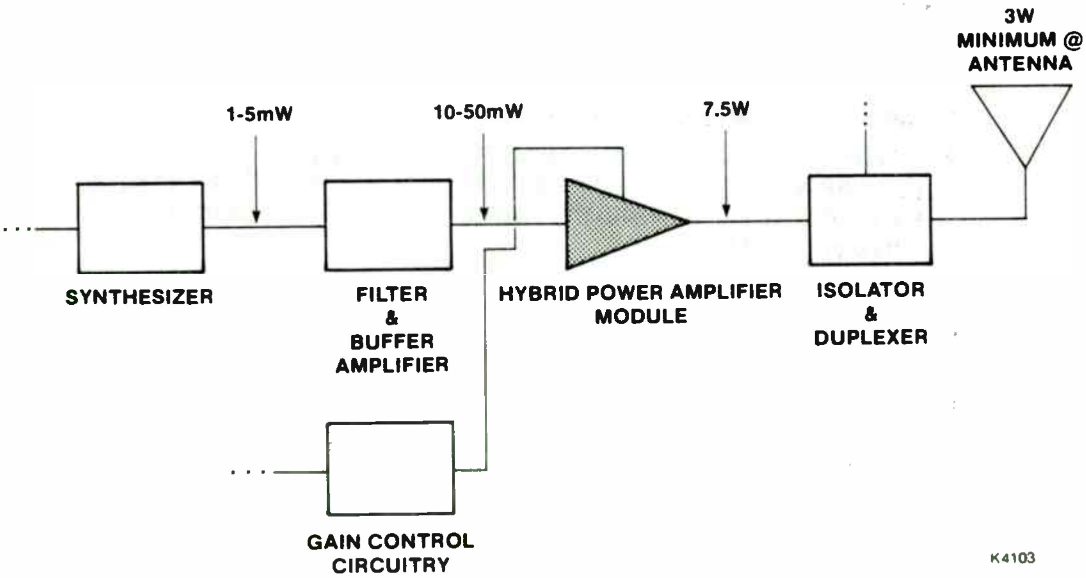
STABILITY: SPURIOUS OUTPUTS $\leq$ 70dBc FOR P_{IN} = 0 TO 30mW, V_{S1} = V_{S2} = 10 TO 16V, V_{CONT} = 0 TO 12.5V AND LOAD VSWR = 4:1

K4103E

MOTOROLA INC.

FIGURE 4

TYPICAL APPLICATION FOR CELLULAR MOBILE RADIO

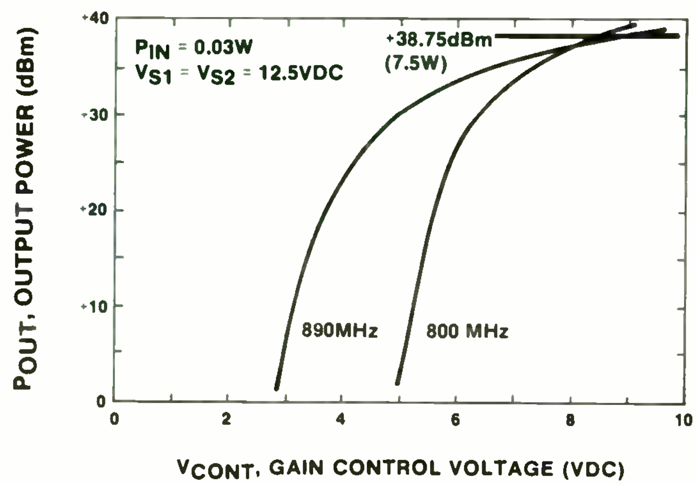


K4103

MOTOROLA INC.

FIGURE 7

OUTPUT POWER VERSUS GAIN CONTROL VOLTAGE FOR MHW808A1

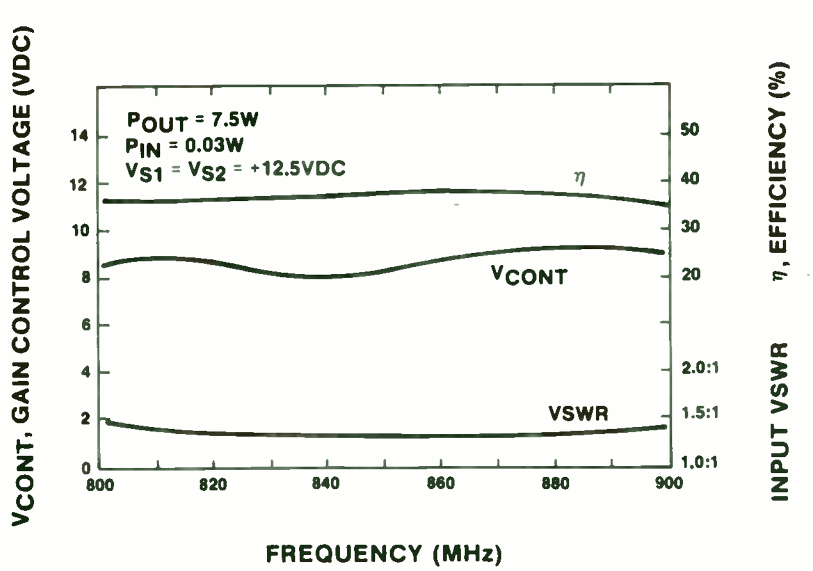


K4103D

MOTOROLA INC.

FIGURE 6

GAIN CONTROL VOLTAGE, INPUT SWR, EFFICIENCY VERSUS FREQUENCY FOR MHW808A1

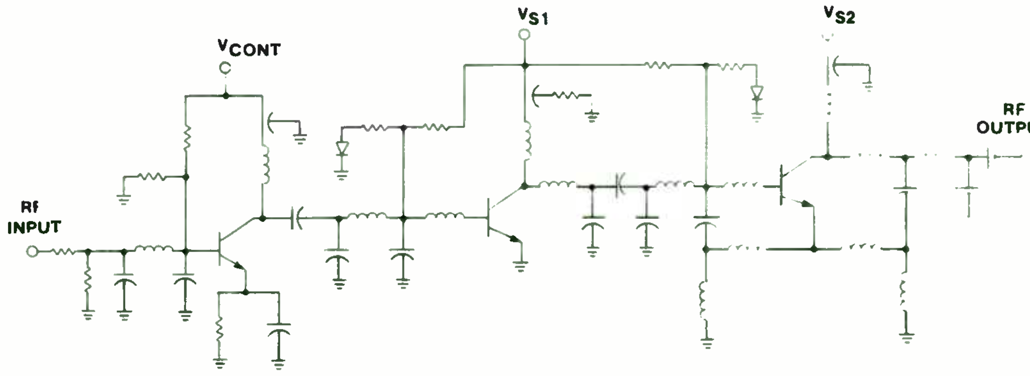


K4103B-1

MOTOROLA INC.

FIGURE 9

CIRCUIT SCHEMATIC FOR MHW808A1

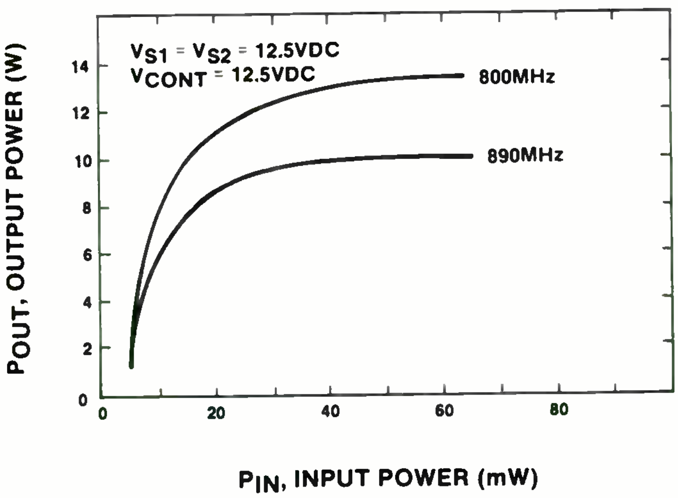


K4103F

MOTOROLA INC.

FIGURE 8

OUTPUT POWER VERSUS INPUT POWER FOR MHW808A1



K4103C

MOTOROLA INC.

FIGURE 11

LOW-PASS CHEBYSHEV IMPEDANCE MATCHING NETWORK CONFIGURATION

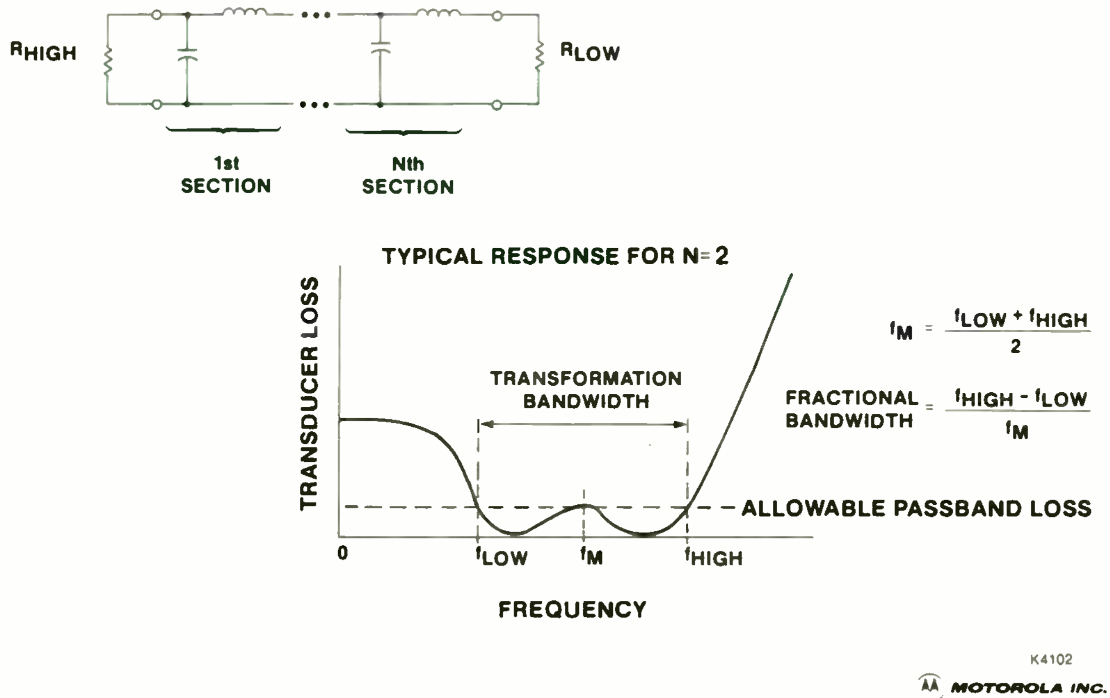


FIGURE 10

THRESHOLD BIAS FOR CLASS-C OPERATION

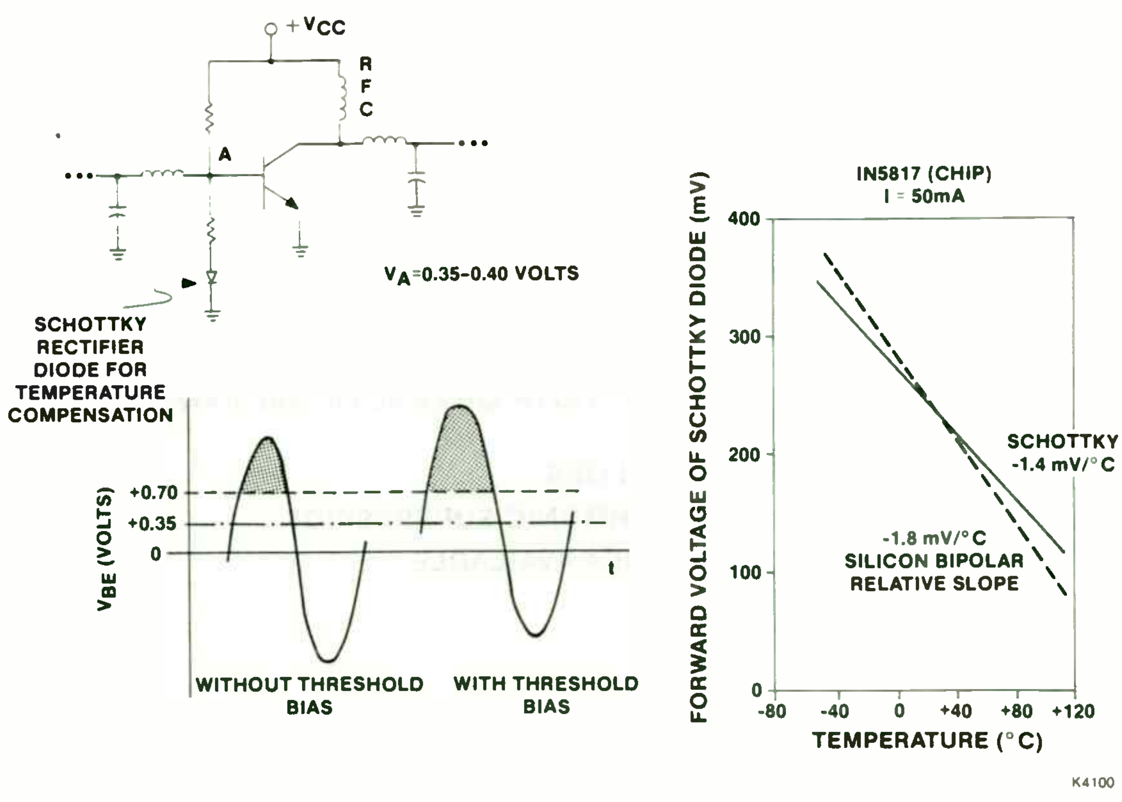
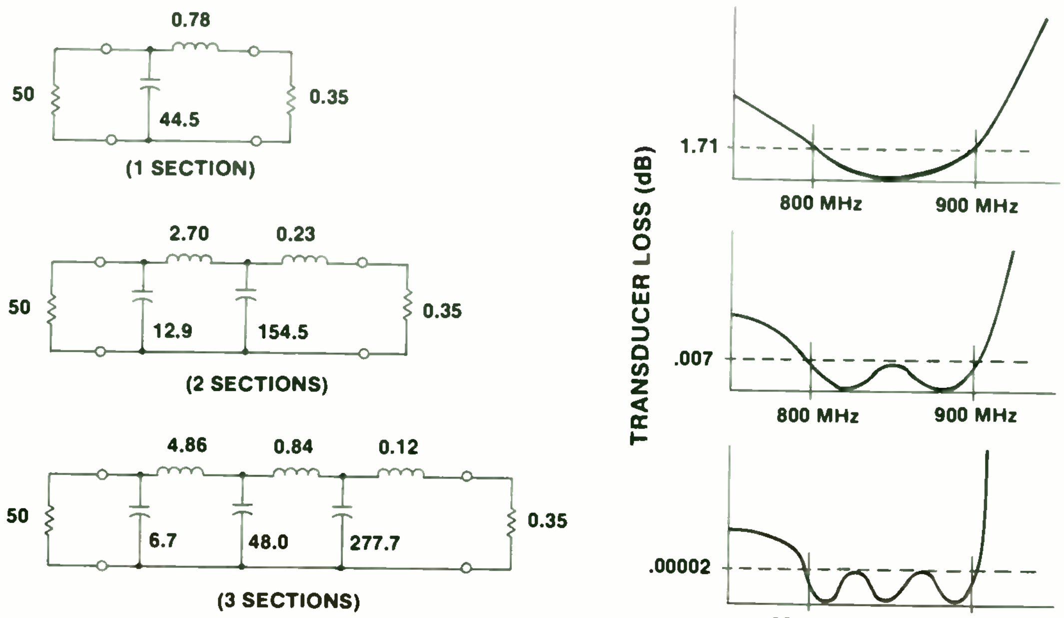


FIGURE 13

DESIGN EXAMPLE ILLUSTRATING PASSBAND TRANSDUCER LOSS VS. NUMBER OF MATCHING SECTIONS



UNITS FOR RESISTANCE: OHMS
UNITS FOR INDUCTANCE: nH
UNITS FOR CAPACITANCE: pF

K4099

MOTOROLA INC.

FIGURE 12

CONSIDERATIONS FOR DETERMINING NUMBER OF MATCHING SECTIONS NEEDED

1. TRANSFORMATION RATIO (i.e., R_{HIGH} / R_{LOW}).
2. TRANSFORMATION BANDWIDTH OR FRACTIONAL BANDWIDTH REQUIRED.
3. ALLOWABLE PASSBAND LOSS.
4. LEVEL OF REQUIRED HARMONIC SUPPRESSION.
5. AMOUNT OF CIRCUIT AREA AVAILABLE.

K4092

MOTOROLA INC.

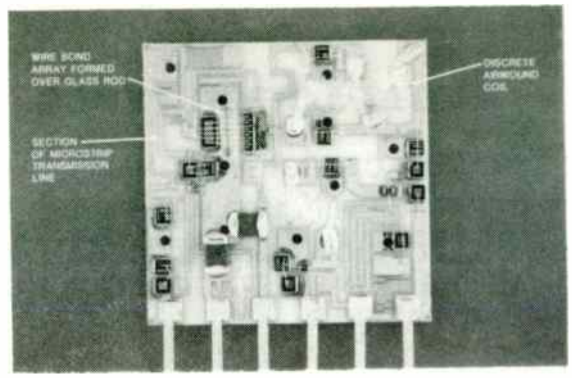


FIGURE 15 — EXAMPLES OF INDUCTORS USED IN MODULE DESIGN

FIGURE 14

INDUCTOR TYPES

1. PRECISELY FORMED WIRE BOND ARRAYS (0.1 TO 1 nH).
2. LENGTHS OF ELECTRICALLY SHORT MICROSTRIP TRANSMISSION LINE (1 TO 30 nH).
3. DISCRETE AIRWOUND COILS (1 TO 50 nH).

K4094

MOTOROLA INC.

FIGURE 17

CAPACITOR TYPES

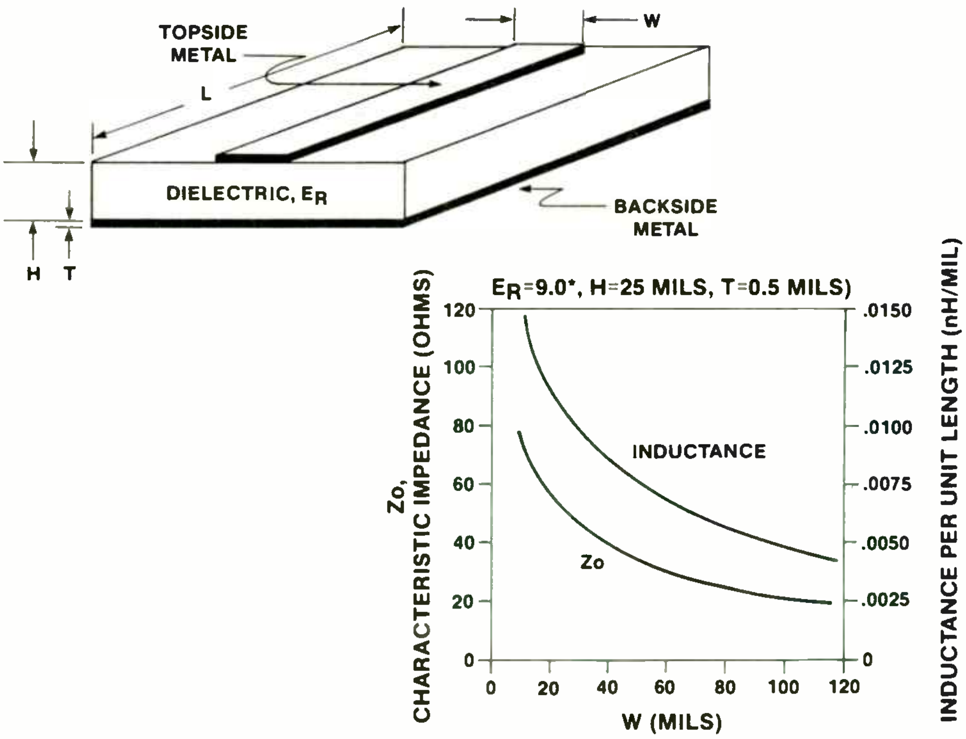
1. MOS (MANUFACTURED INTERNALLY). 0.5 TO 500 pF
2. NPO CERAMIC CHIP. 1 TO 500 pF
3. HIGH DIELECTRIC CONSTANT CERAMIC CHIP (USED PRIMARILY FOR INTERSTAGE BLOCKING AND SUPPLY BYPASSING).
   18,000 pF

K4093

MOTOROLA INC.

FIGURE 16

PROPERTIES OF TYPICAL MICROSTRIP CONFIGURATION



*E_R FOR ALUMINUM OXIDE

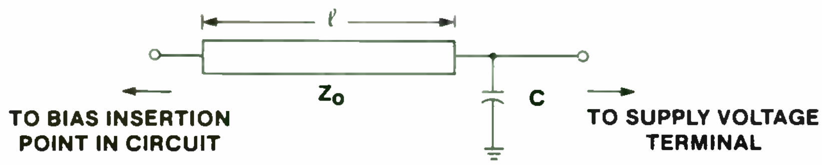
K4101

MOTOROLA INC.

FIGURE 19

BIAS INSERTION TECHNIQUES

1. CHOKES FORMED WITH LENGTHS OF HIGH IMPEDANCE MICROSTRIP TRANSMISSION LINE:



   a. $\lambda/8 < \ell < \lambda/4$

   b. $55 < Z_o < 90$ OHMS

   c. CAPACITOR, C, PROVIDES LOW IMPEDANCE TO GROUND OVER OPERATING FREQUENCY RANGE.

2. IN SITUATIONS WHERE HIGH-PASS IMPEDANCE MATCHING NETWORKS ARE USED, SHUNT INDUCTORS ALSO PROVIDE A MEANS OF BIAS INSERTION.

K4098

MOTOROLA INC.

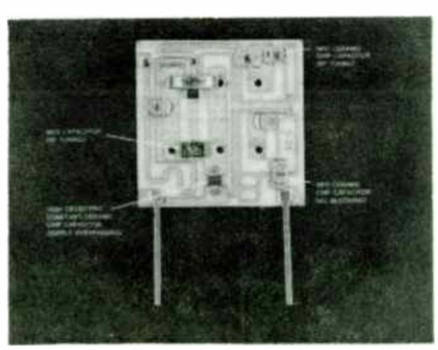


FIGURE 18 — EXAMPLES OF CAPACITORS USED IN MODULE DESIGN

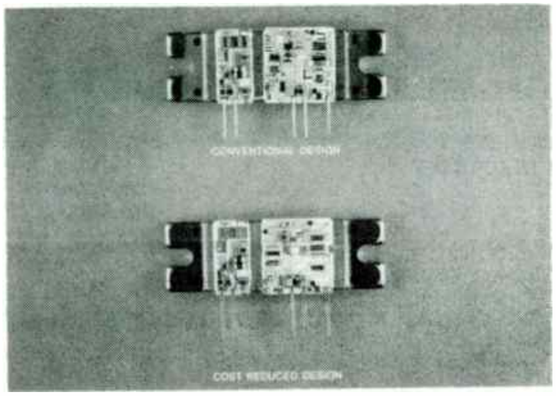


FIGURE 21 — CONVENTIONAL AND COST-REDUCED DESIGNS FOR MHW808A1

70

FIGURE 20

ALUMINUM OXIDE VERSUS BERYLLIUM OXIDE

	ALUMINUM OXIDE	BERYLLIUM OXIDE	
RELATIVE DIELECTRIC CONSTANT	9.0	6.7	
THERMAL CONDUCTIVITY	0.34	1.84	W/CM ·°C
METAL ADHESION STRENGTH	SATISFACTORY	POOR	
AVAILABILITY	SATISFACTORY	POOR	
COST	$0.19	$2.70	PER SQ. INCH

K4090

FIGURE 23

SKIN DEPTH VERSUS FREQUENCY FOR COPPER AND GOLD

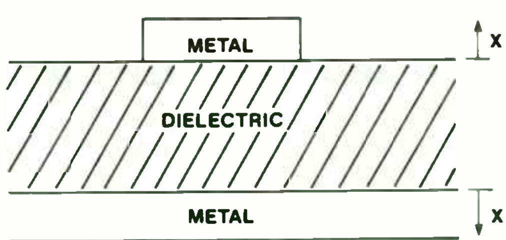
	COPPER	GOLD	
150 MHz	212	236	μ INCHES
450 MHz	123	136	μ INCHES
850 MHz	89	99	μ INCHES

K4089

MOTOROLA INC.

FIGURE 22

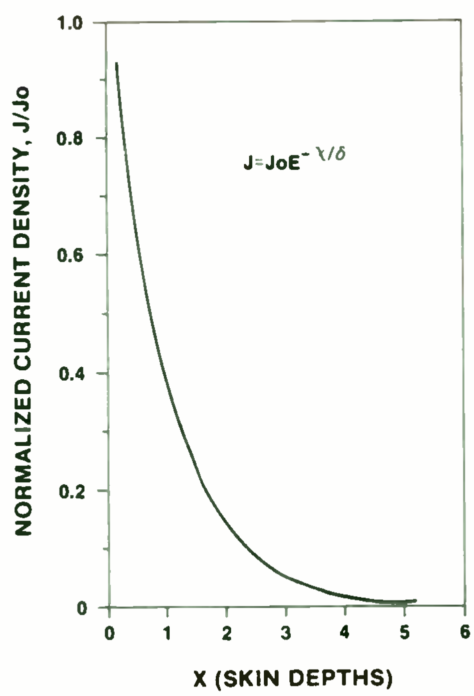
CURRENT DENSITY WITHIN CROSS-SECTION OF MICROSTRIP CONDUCTOR



J ≡ CURRENT DENSITY IN THE METAL CONDUCTOR EXPRESSED AS A FUNCTION OF X.

Jo ≡ CURRENT DENSITY IN THE METAL CONDUCTOR AT THE METAL/DIELECTRIC INTERFACE (X=0).

δ ≡ SKIN DEPTH = $\sqrt{\dfrac{2}{\omega\mu\sigma}}$



K4097

MOTOROLA INC.

FIGURE 25

RELIABILITY TEST RESULTS

	SAMPLE SIZE	FAILURES
1. HIGH TEMPERATURE STORAGE LIFE	35	0
2. MOISTURE RESISTANCE	84	0
3. STEADY-STATE OPERATING LIFE	40	0
4. CYCLED OPERATING LIFE	16	0
5. TEMPERATURE CYCLING	80	0
6. THERMAL SHOCK	10	0
7. VIBRATION	10	0
8. MECHANICAL SHOCK	10	0
9. SOLDERABILITY	10	0
10. LEAD BEND	10	0
11. SMOG ATMOSPHERE, H_2S, NO_2, SO_2	8 of each	0

K4103G

MOTOROLA INC.

FIGURE 24

RELIABILITY TESTING SUMMARY FOR HYBRIDS

1. HIGH TEMPERATURE STORAGE LIFE (125° C, 1000 HOURS)
2. MOISTURE RESISTANCE (85° C, 85% RELATIVE HUMIDITY, WITH DC BIAS APPLIED)
3. STEADY-STATE OPERATING LIFE (100° C, 1000 HOURS)
4. CYCLED OPERATING LIFE (50° C TO 110° C, 5000 CYCLES, 6 MINUTES PER CYCLE)
5. TEMPERATURE CYCLING, AIR TO AIR (-55° C TO +125° C, 10 MINUTES AT EXTREMES)
6. THERMAL SHOCK, LIQUID TO LIQUID (0° C TO +100° C)
7. VIBRATION, VARIABLE FREQUENCY (10, 55, 10HZ ON X, Y AND Z AXIS)
8. MECHANICAL SHOCK (500G, 1 MSEC, 3 PLANES)
9. SOLDERABILITY (260° C, 10 SECONDS)
10. LEAD BEND (8 OZ., 90° BEND AND RETURN)
11. SMOG ATMOSPHERE, H_2S, NO_2, SO_2(75% RELATIVE HUMIDITY, 96 HOURS)

K4103A

MOTOROLA INC.

HIGH POWER CLASS A AND CLASS AB TRANSISTORS

Prepared by:

MICHAEL J. MALLINGER
Vice President - Marketing

ACRIAN

HIGH POWER CLASS A AND CLASS AB TRANSISTORS

For UHF TV and Cellular Base Station Applications.

OUTLINE OF PRESENTATION:

A) UHF TELEVISION

I. Overview of Requirement — System Needs

II. UHF TV Class A Performance

III. Transistor Performance Characteristics and Design Criteria

IV. Transistor Performance Achieved

V. Circuit Design Concepts and Performance Achieved

B) CELLULAR BASE STATION

I. Overview of Requirement — System Needs

II. Base Station Class AB Performance

III. Transistor Performance Characteristics and Design Criteria

IV. Transistor Performance Achieved

V. Circuit Design Concept and Performance Achieved

A) UHF TELEVISION

I. OVERVIEW OF REQUIREMENT -- SYSTEM NEEDS

The high power UHF TV transmitter has the following system needs:

Operating Frequency Range: 470-860MHz

Input Signal Levels	Frequency
Visual at - 8dB	FO
Aural - 7dB	FO + 4.5MHz
Color Sub Carrier - 16dB	FO + 3.5MHz

All intermodulation products are measured in dB below the peak sync pulse and are specified at 60dB down. Since this spec includes preemphasis the transistor is specified at -50dB.

This specification set is usually referred to as the European test method and is considered to be the most stringent of the specifications for this system type.

The solid state amplifier power level desired is 100 watts. This can be used to drive a high power travelling - wave tube amplifier to 1KW or higher. To achieve a 100 watt amplifier will require the following combination in the final stage: (Using binary combination)

2 - 70 watt
4 - 40 watt
8 - 20 watt
16 - 12 watt
32 - 8 watt

It should be noted that there are a number of systems which have combined 32 or more of the lower power transistors to achieve the needed power. The combining of a large number of transistors is expensive to build and to maintain. The use of a pair of transistors could be costly if it meant downtime due to failure. A reasonable compromise is to use an 8-way combined unit utilizing 20-watt transistors. This improves the initial cost and allows for "graceful degradation" should a single transistor fail.

A typical lineup would then look as follows:

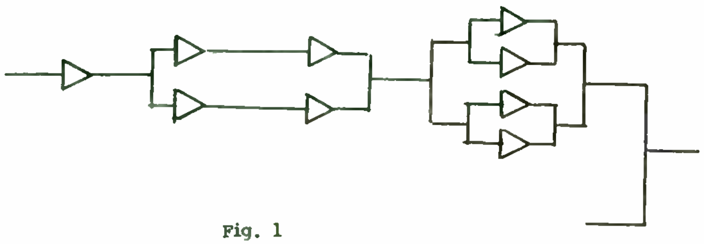


Fig. 1

The use of combined transistors in the final stages improves the intermodulation characteristics and reduces the impact of interstage mismatch. It also allows for manufacture of subassemblies consisting of a transistor pair which are then integrated into the entire power amplifier.

The key transistor is the 20 watt power device used in the final stage. This presentation will detail the design and characteristics of the 20 watt power transistor.

II. UHF TV CLASS A PERFORMANCE

Class A transistors are used to achieve the required linearity in line with the system specs for multitone intermodulation distortion. The device is usually biased Class A at the supply voltage of 24 volts and at the current necessary for full Class A operation — assuming 25% efficiency the 20 watt device would be biased at the DC power of 80 watts which is 24 volts, 3.3 amps. The device must therefore be capable of a high dissipation at this bias condition and also withstand load mismatch under full operation.

The device must also have a high degree of linearity at the operation point and therefore must have a 1dB compression point approx 25% above the rated power therefore the 20 watt power device has power output in excess of 25 watts.

The high power Class A Transistor therefore incorporates all of the state-of-the-art technology in order to produce a product which has the high power output characteristics and still retains the low capacitance needed to operate at 860MHz.

III. TRANSISTOR PERFORMANCE CHARACTERISTICS AND DESIGN CRITERIA

The operating specifications for the 20-watt Class A transistor are as follows:

Frequency Range: 470-860 MHz

Power Output: 20 watts peak sync at IMD of -50dB

Power Gain: 8.5dB

Load Mismatch Capability 3:1 under full operation

Operating Conditions: VCC = 24 volts, IC = 2.7 amps

D.C. Safe Operation Range (SOAR) 24V, 3 amps

Utilizing silicon based technology with NPN microwave power interdigitated designs which include diffused ballasting and gold topside metalization the generic design conditions, for the total device, become:

Emitter Periphery (EP) : 6700 Mils

Base Area (BA): 1200 square Mils

Base Periphery (BO): 900 Mils

EP/BA = 5.8

Chip Design:

Acrian has designed a cell structure which when combined in a push pull package and consists of eight (8) cells on each side will provide the total required active area. This geometry is interdigitated and incorporates diffused ballast resistors, gold topside metal and silicon nitride surface passivation. The photo (PHOTO 1) is a closeup of the cell structure.

The final device is built into a push pull package with eight cells on each side. The product incorporates two steps of low pass match (series L shunt C) on the input. This provides a transformation of the input impedance from the chip to the package terminals and nets an input impedance on each side of 3 + j9 OHMS °. The final device is shown in the photo (PHOTO 2).

IV. TRANSISTOR PERFORMANCE ACHIEVED

The transistors constructed and evaluated have provided the following results:

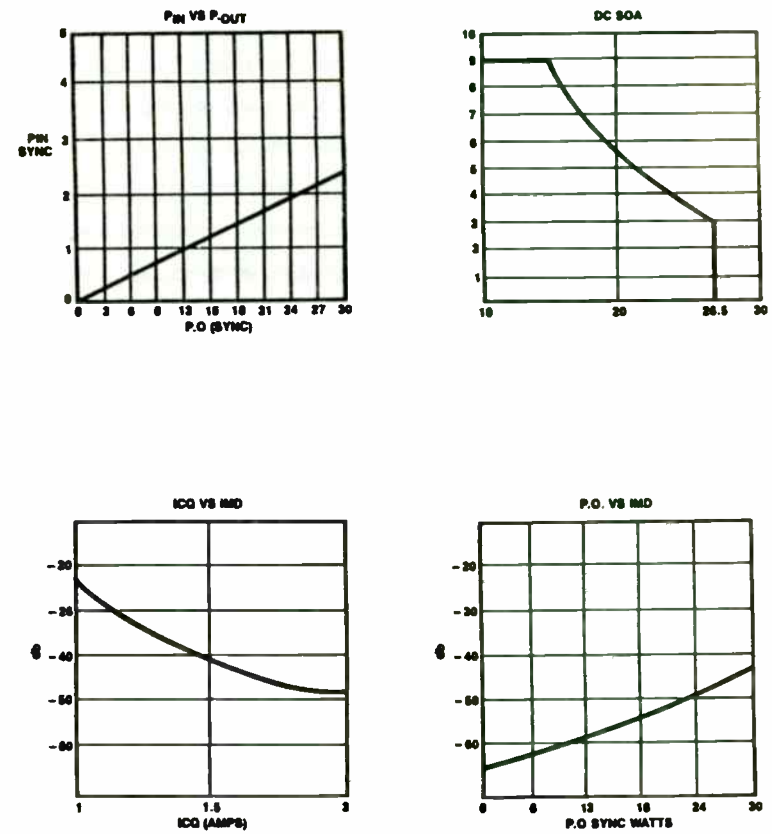


FIG. 2

V. CIRCUIT DESIGN CONCEPTS AND PERFORMANCE ACHIEVED

The transistor is a push pull device therefore the circuit employs the push pull design concept incorporating the balanced to unbalanced 1:1 transformer on the input and output. The device impedances are transformed from the package levels to 25 ohms utilizing microstrip transformers; lumped elements on fiberglass circuit board material (teflon) with a dielectric constant of 2.5 (See Photo 3). This was selected since it is readily available and is quite commonly used in the broadcast industry. If size were critical the design could be accomplished using ceramic alumina in about 1/4 the total area.

A full schematic of the final circuit is as follows:

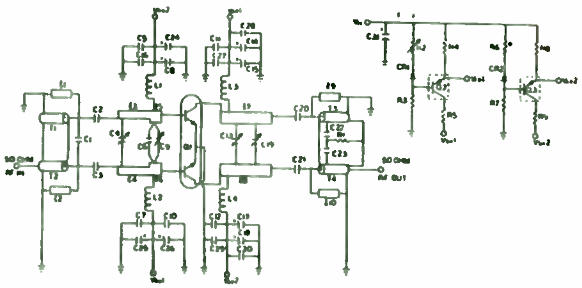


FIG. 3

This circuit is tuneable for a given channel so that it can be peaked for best IMD performance. It will cover perhaps 5 channel quite well without any returning.

Details on the circuit schematic and bill of materials are included in the Appendix.

B) CELLULAR BASE STATION

I. OVERVIEW OF REQUIREMENT -- SYSTEM NEEDS

The cellular base station requires a power amplifier capable of achieving:

Frequency Range: 850-960 MHz

Supply Voltage: 24 volts

Power Output: 45 watts (45 + losses)

Power Gain: 28dB

Dynamic Range: -28dB from Full Power

Load Mismatch: to 1 - after circulator

Stable into a 2 to 1 load mismatch

Projected MTTF 15,000 hrs.

The cell site will consist of a large number (24 to 96) of these transmit amplifiers and therefore the total unit size and power consumption are key points to consider.

II. BASE STATION CLASS AB PERFORMANCE

The base station is required to service a number of units within the cell and must be able to automatically adjust the output power depending on the distance to the mobile unit. Therefore, the power output of the base unit must be adjustable over a wide range. The Bell specifications call for adjustment of power over a range of 28dB down from the full system spec output. Also as the density of cells increases the power output of operation will be reduced. The ability to service a wide dynamic range dictates that the transmit power AMP be designed with transistors working Class AB. This therefore prefers the transistor to be designed common emitter.

III. TRANSISTOR PERFORMANCE CHARACTERISTICS AND DESIGN CRITERIA

The transistor specs for power output of 60 watts, Class AB common emitter with 6.3dB power gain and rugged into a 5 to 1 load mismatch indicate the complexity of the task -- to this point there has been no such product. The current systems use a pair of common base power transistors each providing 35 watts of power output.

Chip Design:

To achieve this performance using NPN silicon bipolar transistors it was necessary to design a product with total active silicon characteristics as follows:

Emitter Periphery (EP) 5,000 Mils

Base Area (BA) 900 Square Mils

Base Periphery (BP) 700 Mils

EP/BA 5.8

An interdigitated structure was selected due to the excellent history of this configuration for designs of this type as used in UHF TV and the EW band of 500-1000MHz.

The chip incorporates diffused emitter ballasting, gold topside metalization and silicon nitride passivation.

The final device consists of 2 groups of six cells on each side of the push pull device.

Package:

It was also decided to use a push pull flange mount package to allow for a simplified circuit versus using a large single ended structure which would have lower terminal impedances. The package is designed with a short series path (input to output) and therefore only required a single step on input match on the final device. Package outline drawing, Figure 7. The transistor flange is the emitter lead inductance -- improving power gain. A side benefit is the much improved input return path -- a capacitor across the two inputs which is much easier to implement versus chip caps to ground in the more conventional approach. The package is sealed be adhering a ceramic lid in place with an epoxy preform -- a commonly used technique in the mobile/cellular product area.

IV. TRANSISTOR PERFORMANCE ACHIEVED

The final transistor performance over the frequency range is as follows:

POWER OUTPUT VERSUS FREQUENCY

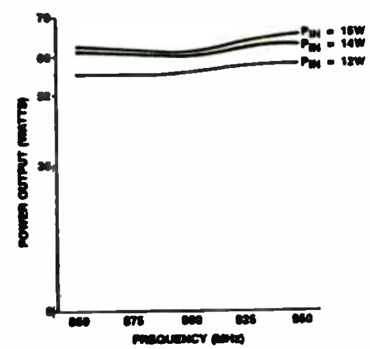


FIG. 4

V. CIRCUIT DESIGN CONCEPT AND PERFORMANCE ACHIEVED

The circuit design is a conventional layout which has been successfully used to cover octave bands in this general range. The concept utilizes an input and output balanced to unbalanced transformer and microstrip matching networks. The bandwidth is quite easily covered with the full spec performance achieved without any tuning. The trimmers are to allow for minor variations in both the circuit and the transistors. The board material is teflon fiberglass with a dielectric constant of 2.5.

The characteristic impedance of each side of the device is 5 + j 13 OHMS on each half of the input. A low pass/high pass match is used to transform the impedance up to the level desired prior to the bal/unbal unit.

The final configuration performance as follows:

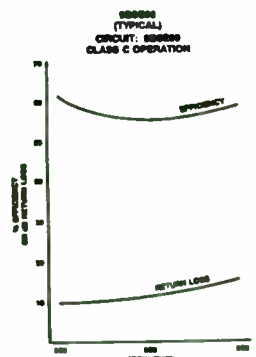
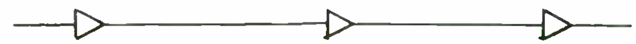


FIG. 5

Studies were conducted to verify the junction temperature during full operation with the results showing a junction temp rise of less than 90 deg. cent when fully stressed.

VI. NEXT GENERATION SYSTEM PERFORMANCE

The high power transistor is used in conjunction with the lower power devices as follows:



9BSE2E 9BSE14E 9BSE60E

FIG. 6

This unit has passed the Acrian qualification procedures as set up by Acrian to simulate the bell procedure and will go thru the bell inspection shortly. It is presently undergoing tests at the labs.

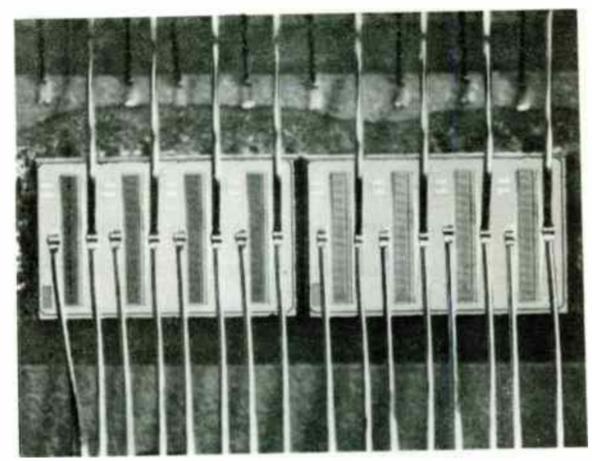


PHOTO 1

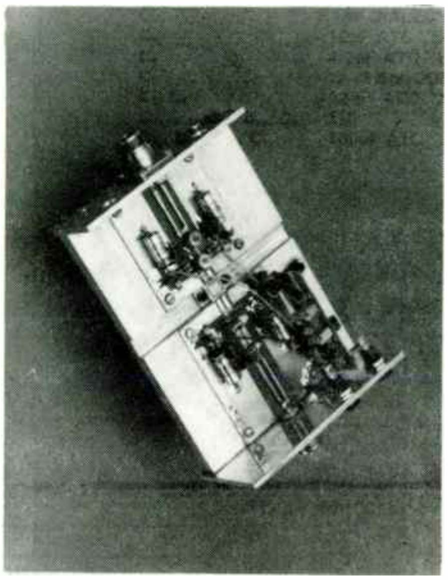
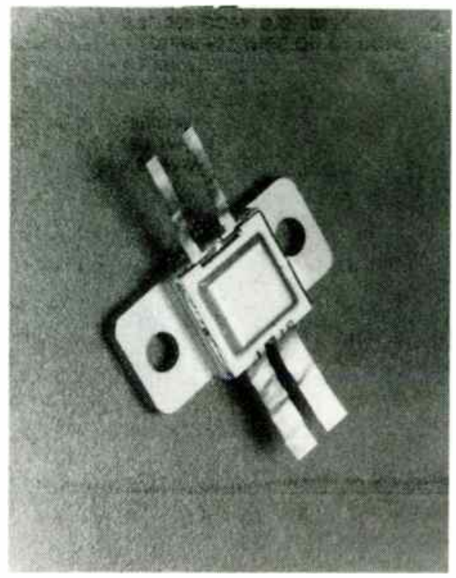
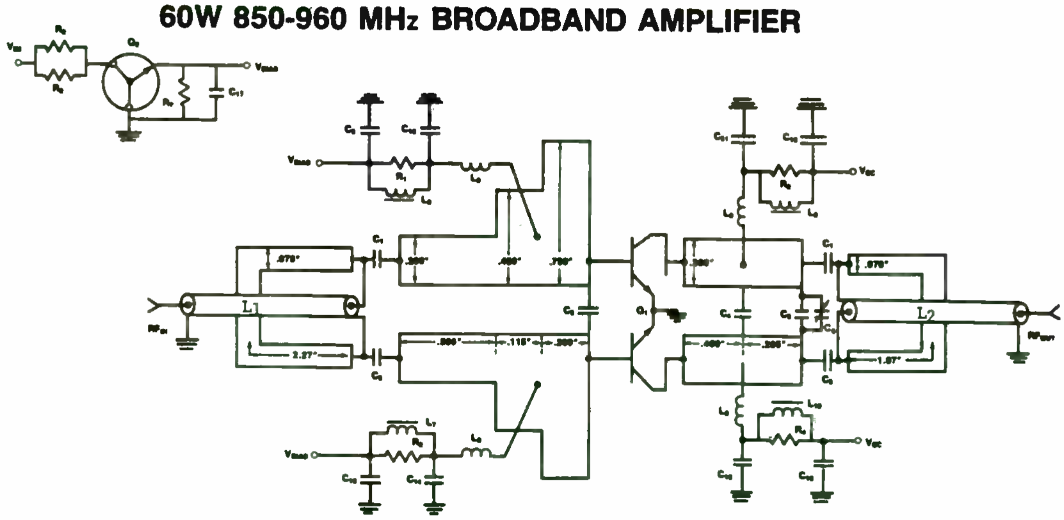


PHOTO 2

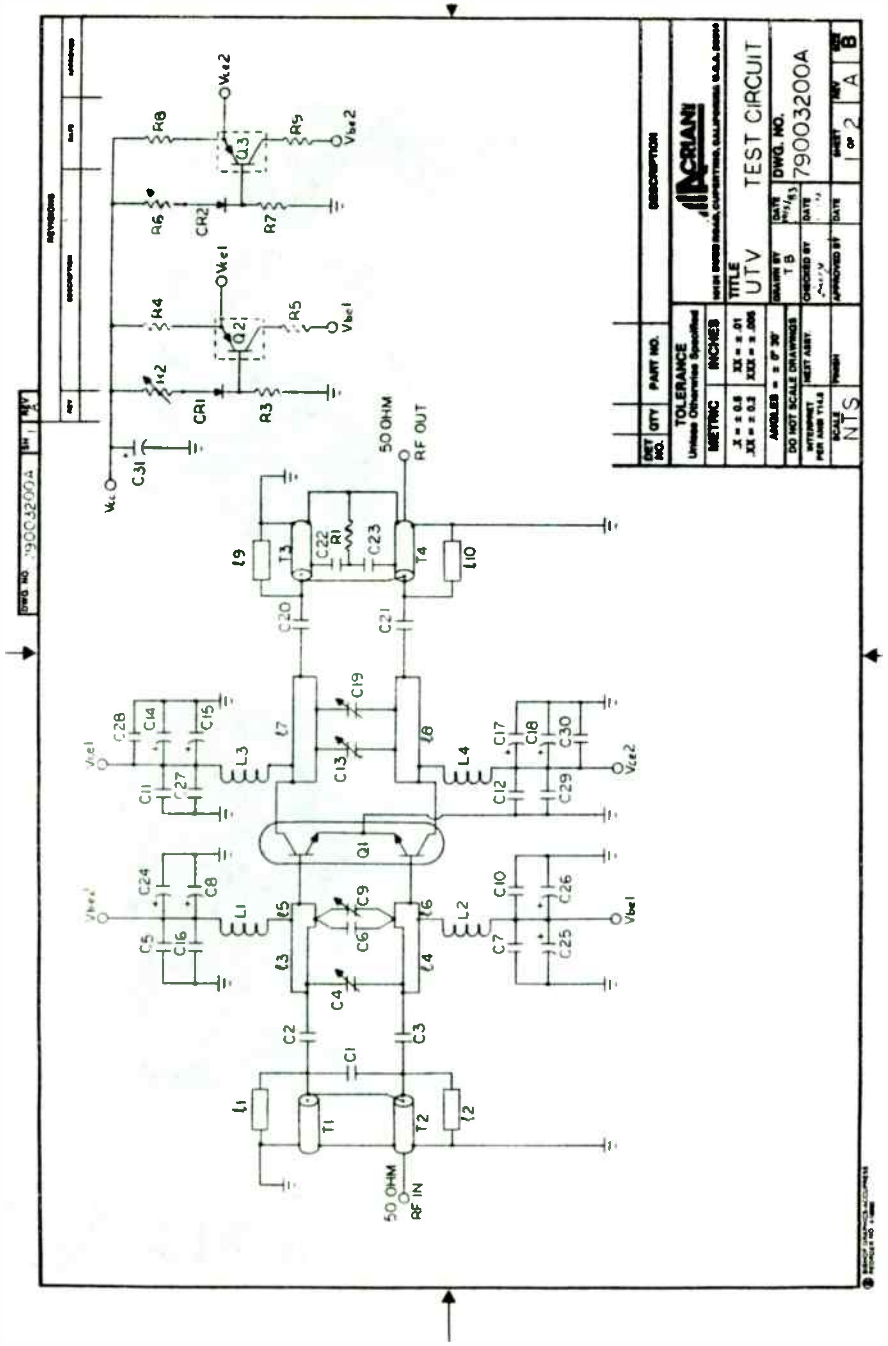


PRELIMINARY
9BSE60
SCHEMATIC DIAGRAM
60W 850-960 MHz BROADBAND AMPLIFIER



BOARD MATERIAL: 1/32" TEFLON - FIBERGLASS

C_1, C_2	- 11.5 pF ATC "B"	L_1	- 2.5" 50Ω COAX O.D. .032"	Q_1	- 0510-50
C_3	5.1pF DIALECTRIC LABS	L_2	- 2.5" 25Ω COAX O.D. .089"	Q_2	- BYI-1
C_4	- 10pF ATC "B"	L_3, L_6, L_7, L_{10}	- 5 TURNS #22 WIRE ON Q1 CORE		
C_5	- 4.7pF ATC "B"	L_4, L_9	- 4.7 μH		
C_6	- .3 – 3.5pF JOHANSON	L_5, L_8	- 0.1 μH		
C_7, C_8	- 32 pF ATC "B"				
C_9, C_{12}, C_{13}, C_{16}, C_{17}	- 1μF	R_1, R_2, R_3, R_4	- 15Ω ¼W		
C_{10}, C_{11}, C_{14}, C_{15}	- 100pF ATC "B"	R_5, R_6	- 15Ω 1W		
		R_7	- 8.2Ω ½W		



APPENDIX

TEST CIRCUIT PARTS LIST

CAPACITORS

C_1, C_6 - 4.7 pF ATC Series A
C_2, C_3, C_{20}, C_{21} - 33 pF ATC Series A
C_4, C_9 - 1.2-3.5 pF Film Dielectric Trimmer
C_5, C_7, C_{11}, C_{12} - 0.01 Microfarad, 50 Volt Disc Ceramic
C_8, C_{15}, C_{17}, C_{25} - 1 Microfarad, 50 Volt Tantalum
C_{10}, C_{16}, C_{27}, C_{12} - 0.1 Microfarad, 50 Volt Disc Ceramic
C_{13} - 0.6-6 pF Piston Trimmer
C_{19} - 0.35-3.5 Piston Trimmer
C_{18}, C_{24}, C_{14}, C_{26} - 10 Microfarad, 50 Volt Electrolytic
C_{28}, C_{30} - 0.001 Microfarad, 50 Volt Disc Ceramic
C_{31} - 100 Microfarad, 50 Volt Electrolytic

RESISTORS

R_1 - 10 OHM, ½ Watt Carbon Composition
R_2, R_6 - 500 OHM Potentometer
R_3, R_7 - 4.7 KOHM, ¼ Watt Carbon Film
R_4, R_8 - 1 OHM, 3 Watt, 1% Carbon Film
R_5, R_9 - 47 OHM, ¼ Watt, Carbon Film

DIODES

CR_1, CR_2 - 1N4148

TRANSISTORS

Q_1 - ACRIAN UTV200
Q_2, Q_3 - MJE 172

INDUCTORS

L_1, L_2 - 0.47 MicroHenry Molded Inductor
L_3, L_4 - One turn #18 gauge wire on a 0.325 inch form

MICROSTRIPLINES

ℓ_3, ℓ_4 - 0.075 in. X 0.65 in. $\ell 1$, $\ell 2$ .035 in. x 1.1 microstrip on board
ℓ_5, ℓ_6 - 0.120 in. X 0.310 in.
ℓ_7, ℓ_8 - 0.120 in. X 1.33 in. $\ell 9$ $\ell 10$ .035 in. x 1.1 microtrip on board

TRANSFORMERS

T_1, T_2, T_3, T_4 - 50 OHM semi-rigid coaxial cable (0.056 in. X 1.1 in.) soldered to microstripline measuring 0.035 in. X 1.1 in.)

Note: All microstriplines were calculated for 1/32 in. dielectric teflon glass 2 oz. copper clad substrate (Er=2.575).

HIGH-VOLTAGE UHF POWER STATIC INDUCTION TRANSISTORS

by

Robert J. Regan
Scott J. Butler

GTE Laboratories Incorporated
Waltham, MA 02254

Abstract

This paper describes a new family of UHF power static induction transistors (SITs). These new transistors have significant advantages with respect to output power, breakdown voltage, efficiency, and terminal impedances, compared to more conventional bipolar transistors and MOSFETs. One of these devices, a new single-ended multicell SIT, has demonstrated 180W cw output power with 6 dB gain at 225 MHz with >70% drain efficiency while operating at 60 Vdc. Small signal measurements on single-cell SITs indicate 10 dB gain in the UHF range and a unity power gain frequency in X-band.

Introduction

Present-day high-frequency electronic system designers and manufacturers are required to deal with a number of issues which are determined by the characteristics of the transistors used in their designs. One of the most pressing problems is the conversion of line voltage power to the lower voltages required by presently used power transistors. This conversion generally requires the use of large, heavy, and costly magnetic components. Another is in the power circuit design where very low transistor terminal impedances must be accommodated by high transformation ratio impedance-matching networks. In addition, thermal considerations are very important, and transistors with the highest power conversion efficiency are desired. GTE Laboratories has addressed these problems in the ongoing development of a family of SITs. The SIT embodies the best combination of power, frequency, gain, efficiency, and breakdown voltage of any semiconductor device yet reported.[1] Thus, it is possible, using SITs, to design and implement efficient, high-gain, high- frequency power amplifiers and oscillators which are capable of operating at high dc supply voltage levels.

Background

High-voltage SITs, which operate at relatively low frequencies (100 MHz), have been reported by Kotani et al.[3] High-frequency SIT performance has been reported by Kane and Frey,[4] but their devices operated at low voltage and relatively low power levels. Thus, high-frequency, high-power SITs have been limited to operation at low voltages (<50V),[2] until the recent development of high-voltage UHF power surface-gate SITs.[1,7,8] Advanced microwave MOS power FETs[5] have shown good performance, but these devices also operate at low voltage levels (28V) and bipolar transistors, the most commonly used UHF power transistors, all operate below 40V.

In order to achieve high-frequency and high operating voltage concurrently, the surface-gate SIT has been optimized in various designs for operating voltage levels between 60V and 100V with good power gain at UHF and L-band frequencies.

SIT Operation

SITs are a special class of junction field-effect transistors (JFETs), in which the current flowing "vertically" between the source and drain is controlled by the height of an electric potential energy barrier under the source.[6] Such a barrier will develop when the channel is depleted of mobile charge carriers by reverse biasing the gate junction. The height of the barrier is influenced by both the gate and the drain bias potentials. The channel current in an SIT is primarily due to electron transfer. Thus, the SIT is a majority carrier device, free from the deleterious effects of minority carriers. Since the electrons have high mobility and travel at saturated velocity through the depleted channel, the SIT is a high-frequency device which achieves high-voltage operation by virtue of the high intrinsic breakdown fields of bulk material.

SIT Electrical Performance

SITs have been fabricated and characterized at GTE Laboratories with pitch ranging from 15 μm to 7 μm.[1,7,8] This paper presents the most recent performance obtained from 7 μm pitch devices. In order to increase the gain and reduce the interelectrode capacitance in these devices, local oxidation (LOCOS) is employed to separate the gate and source in the vertical direction. A simplified cross section of this LOCOS surface gate SIT (SGSIT) is shown in Figure 1, while Figure 2 illustrates typical dc I-V characteristics for a multicell LOCOS SGSIT. Figure 3 shows a set of typical electrical performance data, normalized to device size, where appropriate, by using the width of the active gate.

Small signal S-parameter measurements, taken using a Hewlett-Packard HP8409 network analyzer system, have been used in conjunction with SUPER-COMPACT to establish an accurate equivalent circuit model [Figure 4(a)] for this SIT. Figure 5 compares the measured and modeled small-signal gain data. The gain calculated using the measured S-parameters from 0.5 GHz to 4 GHz is identified as (A) on this figure. The model element values which provided a "best fit" to the four measured S-parameters were determined by an optimization procedure performed using SUPER-COMPACT. The plot identified as (B) on Figure 5 indicates that the gain calculated from the optimized model correlates very well with that calculated from the measured data. Once determined to be a fairly accurate electrical representation of the device, the model was used to determine the performance down to 100 MHz and also to evaluate the influence of package parasitics on the microwave performance of the SIT chip.

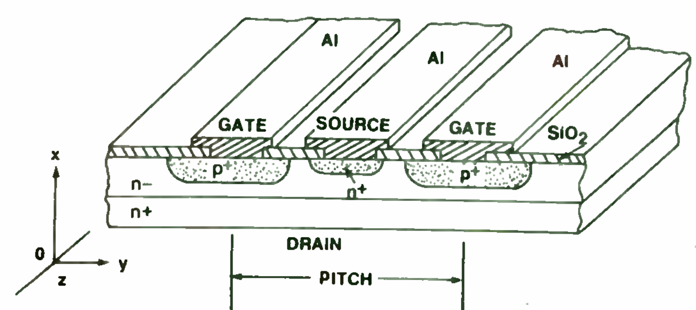


Figure 1. Typical cross section of a LOCOS surface gate SIT (SGSIT)

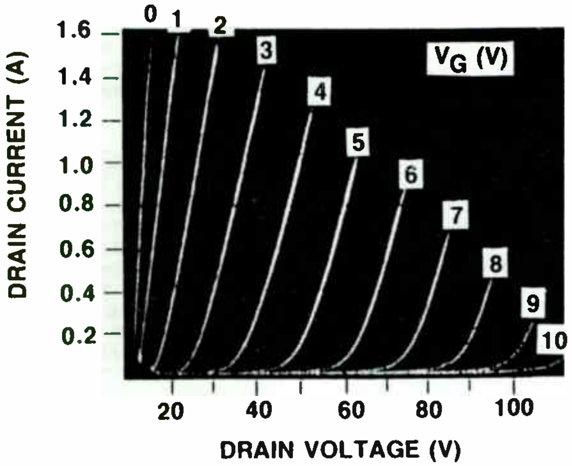


Figure 2. SGSIT DC I-V characteristic (W_g = 24 cm)

Blocking Voltage	BV_{DG}	135V
Voltage Gain	μ	10
Transconductance	g_m	50 mS/cm
ON-Resistance	R_{ON}	100Ω • cm
Input Capacitance	C_{SGO}	3 pF/cm
Output Capacitance	C_{DGO}	2 pF/cm
Unity Power Gain Frequency		>6 GHz

Figure 3. Electrical performance Data for 7 μm pitch LOCOS SGSIT

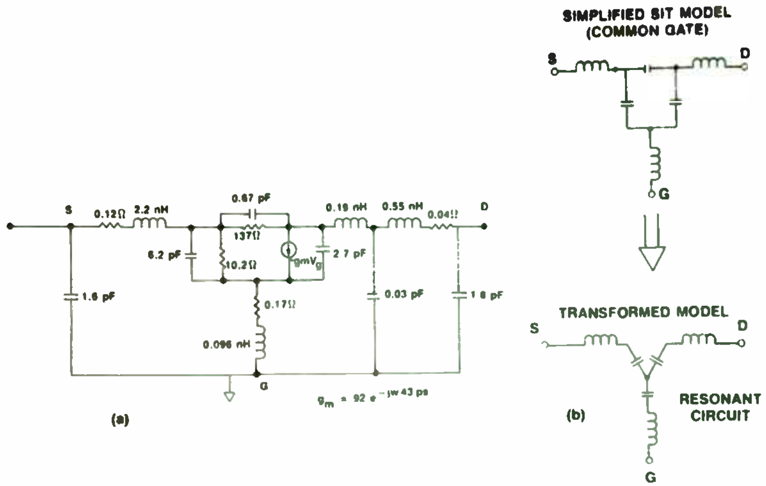


Figure 4. SGSIT equivalent circuit model

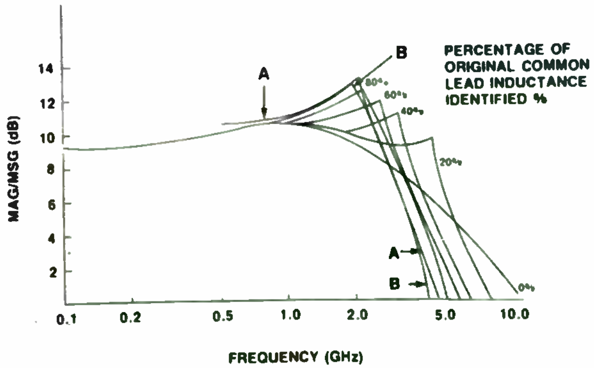
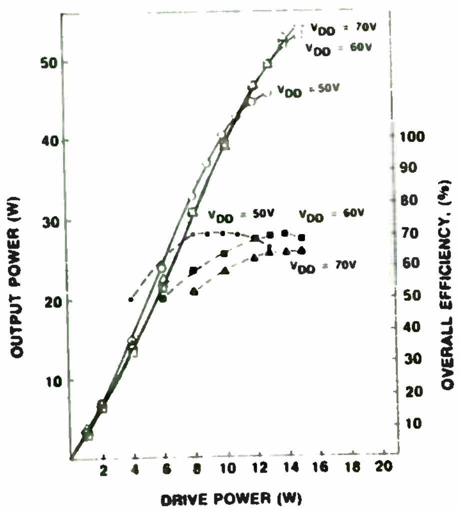


Figure 5. Small signal gain vs frequency of a 7 μm pitch SGSIT

As expected, common lead inductance is the most influential package parasitic component limiting the frequency response. Although the common lead inductance in the packaged SGSIT is less than 0.1 nH, there is considerable influence on the high-frequency performance. The reactive portion of the intrinsic SGSIT equivalent circuit may be converted to a common-node configuration using a delta-Y transformation as shown in Figure 4(b). When this is done, it is clear that a series resonant circuit is formed in the common lead. This results in a resonance peak in the small-signal gain at about 2.5 GHz.[9] At this frequency, unilateral gain is approached; however, the gain decreases rapidly above this frequency. As shown on Figure 5, the common lead parasitic impedance was reduced in steps between the optimized equivalent circuit value and zero. The resultant circuit at each step was analyzed from 100 MHz to 10 GHz using SUPER-COMPACT. In each case the low frequency gain remained about the same, while the unity power gain frequency increased. The resonance peak frequency was determined by the common lead reactive element values. The computed unity power gain frequency without common-lead parasitics was found to be 10 GHz, and the gain-frequency characteristics contained no resonances, as expected. Additional analysis of the equivalent circuit with all of the package parasitics removed resulted in minor improvements, indicating the relative importance of the common lead inductance in limiting device performance.

Large-signal tests with 7 μm pitch SGSITs have been conducted at UHF frequencies for three different size devices. For example, three SGSIT cells were combined in a single stripline-type package, GTE 02-140-50 EXP, and tested under "Class B" operating conditions. As indicated on Fig. 6, operating this device at lower drain voltage results in more linear performance with slightly higher gain but lower P_{SAT}. In addition, overall efficiency is also defined and plotted on this graph. It is interesting to note that the drain efficiency ($\eta_D = P_o/P_{dc}$), for example, at $P_o = 50W$ with $V_{DD} = 60V$ and $P_{DRIVE} = 14W$, is $\cong$ 85%. This is very high compared to any other UHF power transistor.



$$\text{OVERALL EFFICIENCY} = \frac{P_{out}}{P_{dc} + P_{drive}} \times 100\%$$

Figure 6. 7 μm pitch SGSIT 225 MHz cw performance (W_g = 6 cm)

Figure 7 illustrates the performance of a packaged six-cell SGSIT (W_g = 12 cm) at various operating conditions (class "A" and "B"). As expected, much higher gain ($\cong$ 8.5 dB) and more linear operation is possible when the device is biased in a class "A" mode of operation (curve 3), albeit with lower saturated output power ($P_{SAT} \cong 75W$). Compared to the three-cell class "B" data shown earlier, this six-cell device performs as expected, providing approximately twice the output power (> 100W) with the same power gain. The unprecedented 180-W performance of the single-ended packaged 12-cell SGSIT (W_g = 24 cm), GTE 02-140-180 EXP, is shown in Figure 8. This device consists of six chips eutectically mounted on three specially configured metallized BeO package inserts which are mounted in place within a specially designed power package using Au/Tin eutectic preforms. Quartz glass rods are epoxied in place within the package to provide support for the gate and source aluminum bond wires. Up to 180W cw was demonstrated with this device. Drain efficiency was also very high, peaking at 76% at $P_o \sim 175W$. The surface temperature of all of the SGSIT cells in this device was monitored during this test using infrared techniques and found to be approximately 95°C ± 5, indicating excellent die bond and insert bond integrity and uniformity.

Conclusion

Power SGSITs fabricated with small pitch have been shown to be capable of unusually high bias voltage operation at microwave frequencies while exhibiting high efficiency and gain as well as high input and output impedances. As a new class of high-voltage microwave power transistor, the SIT is presently under consideration for applications such as phased array radar systems, broadcast transmitters and high scanning rate electron beam systems.

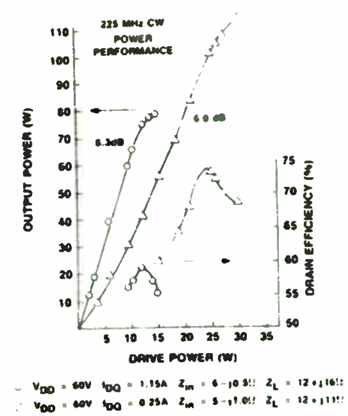


Figure 7. 7 μm SGSIT 225 MHz cw performance (W_g = 12 cm)

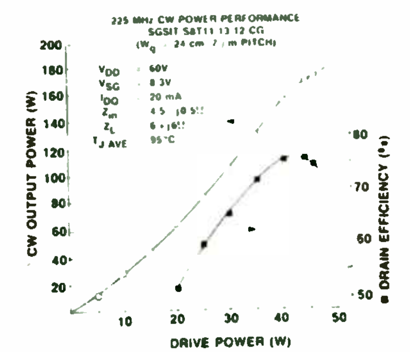


Figure 8. 7 μm pitch SGSIT 225 MHz cw performance (W_g = 24 cm)

Acknowledgment

The authors wish to acknowledge the support of Paul Haugsjaa; the important contributions to this work by Adrian Cogan, Izak Bencuya, and Fred Rock in the areas of device design and process development; the technical assistance of Anthony Varallo, Marguerite Delaney, Charles Herrick, Theresa Rubico, Steve Rose, and Maureen Sullivan; and the process engineering contributions of Emel Bulat.

This work was supported, in part, by the USAF Systems Command, WPAFB, under Contract No. F33615-82-C-1702.

References

1. A. Cogan, R. Regan, I. Bencuya, S. Butler, and F. Rock, "High Performance Microwave Static Induction Transistors," IEDM, Washington, DC, December 1983, Paper 9.5.
2. Y. Kajiwara *et al.*, "A 100 W Static Induction Transistor Operating at 1 GHz," Proceedings of 11th Conference on Solid State Devices, Tokyo (1979).
3. M. Kotani *et al.*, "Characteristics of High-Power and High-Breakdown-Voltage SIT with the High Maximum Frequency of Oscillation," IEEE Trans. ED, Vol. 29, p. 194 (February 1982).
4. M. Kane and R. Frey, MSN, September 1984, p. 46.
5. D. Moss, "ISOFET Evolution Yields High-Power Devices," *Microwaves and RF*, February 1983.
6. J. Nishizawa, Editor, *Semiconductor Technologies*, 1982, pp. 220-240, North-Holland Publishing Co.
7. A. I. Cogan and P. O. Haugsjaa, "Two-Dimensional Field and Potential Analysis in SIT," Proceedings of 38th Device Research Conference, Ithaca, NY, paper #3A-8 (1980).
8. R. Regan, A. Cogan, I. Bencuya, S. Butler, and P. Haugsjaa, "Improved Performance of High-Voltage Microwave Power Static Induction Transistors," 14th European Microwave Conference, Liege, Belgium, September 1984.
9. J. Nishizawa, ed. Semiconductor Technologies, North Holland Publishing Co., 1982, p.228.

"USE OF SAW TECHNOLOGY IN THE RF SYSTEMS OF THE 1980's"
by

Carl A. Erikson, Jr.
Director of Processing Operations

Andersen Laboratories, Inc.
1280 Blue Hills Avenue
Bloomfield, CT 06002

ABSTRACT

Today's Surface Acoustic Wave (SAW) devices are no longer a research curiosity but are derived from a mature technology, providing low cost, high volume production benefits over wide performance ranges. SAW devices include delay lines, bandpass filters, resonators, convolvers, matched filters, and dispersive delay lines.

As RF engineers understand the advantages of this technology, more and more of them are using SAW devices in such systems as channelized filter banks, compressive receivers, chirp radars, spread spectrum communication systems, and ECM equipment.

This paper will summarize the current usage of SAW devices in RF systems, both commercial and military, and also provide some insight in future trends and applications.

INTRODUCTION

SAW devices have been demonstrated to be extremely successful in many state-of-the-art RF systems to include commercial[1], military[2], and space[3] applications. These systems' demands for high performance, highly reliable components have enhanced SAW devices' reputation, allowing them to offer viable alternatives to other technologies such as digital processing, charge coupled devices, and acousto-optics. As RF engineers begin to understand the advantages of SAW technology, more of them are using SAW devices in their designs of such systems as channelized filter banks, compressive receivers, chirp radars, spread spectrum communication systems and ECM equipment.

WHAT IS A SAW DEVICE?

A SAW device can be defined as a passive, electro-acoustic device that allows acoustic energy to be generated, manipulated, and detected on a piezoelectric substrate. There are three basic parts to any SAW device (See Figure 1). First, a highly polished piezoelectric substrate such as quartz or lithium niobate is used. The property of piezoelectricity allows the surface acoustic wave to be generated and detected. Table I lists several parameters for typical substrates used in fabricating SAW devices. Second, thin metallized or grooved structures need to be fabricated on the substrate's surface by standard metallization and photolithographic techniques. These structures which include interdigital transducers (IDT), multistrip couplers (MSC), reflective gratings, and waveguides are designed to perform the basic function of the SAW device; e.g., delay line, filter, resonator, convolver, etc. Third, the patterned substrate needs to be packaged and connected to the package's terminals. The package provides the mechanical support and hermetic environment for the SAW substrate which can be extremely sensitive to surface contamination.

Several major advantages of SAW devices which RF engineers should be aware of are:

1. Compactness - Because surface acoustic waves travel on the order of 10^5 times slower than electromagnetic waves, significant signal delays are achieved in a short length of substrate, typically 8 microseconds per inch. For comparison, if coax cable were utilized for the same 8 microsecond delay, over one mile of cable would be needed!

2. Ease and Versatility of Design - Many types of devices can be readily designed with standard CAD techniques because there is a direct correspondence between the finger placement and weighting techniques of the transducers (representing time-domain) and the desired output (usually frequency domain); namely, the Fourier transformation. See Figure 2.

3. Economical - The planar processing techniques used to fabricate SAW devices and the great reproducibility in manufacturing have made them very economical for many aplications. Costs may be as low as $.50 for high volume TV IF SAW filters or may run at $5000 to $10,000 per unit for a high performance, ovenized SAW dispersive delay line. In addition, the ability to easily integrate SAW substrates with hybrid substrates and components has been a key in providing low cost SAW oscillators, programmable matched filters, and internally-tuned SAW devices.

APPLICATIONS OF SAW DEVICES

Table II lists a few of the known applications of SAW devices based on type of function of the device. For each major function a few comments and examples will be mentioned.

SAW DELAY LINES

The amount of separation between the input and output transducers relates directly to the time delay of the SAW device, namely T_{DELAY} = Separation/Velocity$_{SAW}$. The amount of metal versus free surface between the transducers, the substrate's surface anomalies, and SAW material parameter changes with temperature will affect the time delay. Figure 3 is an example of a wideband SAW delay line centered at 700 MHz.

Because of its temperature stability, quartz is often utilized for delay lines. Quartz substrates as long as 12 to 15 inches have been used, giving time delays approaching 100 microseconds. To increase time delays above this, techniques such as rounding the edges of the substrate, multistrip couplers, and cascading substrates have been tried.

SAW FILTERS

The filter functioning is one of the larger volume applications for SAW devices. Because SAW filters can be designed with optimized amplitude and phase responses for most of the world's television standards, millions of devices are fabricated every year for the intermediate frequency stages of monochrome and color television receivers. Other IF filtering applications include both basic and addressable CATV converters and decoders, TV tuners and Data Modems. Filtering for video and sound modulator outputs for CATV and Satellite receivers are other high volume applications. Figure 4 is a typical frequency response of a SAW vestigial sideband (VSB) filter.

SAW OSCILLATORS AND RESONATORS

SAW oscillators have been developed in response to both military and commercial needs to provide compact, stable and high performance sources in the high frequency range (100 MHz to 1100 MHz). SAW delay lines or resonators are integrated with standard hybrid circuitry into highly reliable and hermetically sealed packages. Figure 5 is a photograph of an actual device. Because the SAW oscillators operated at fundamental frequencies in the VHF and UHF range, the need for frequency multipliers and post multiplier filtering is reduced or eliminated. Both fixed frequency and voltage controlled SAW oscillators can be fabricated.

In addition to sources, one major use of SAW oscillators is in sensors. Parameters such as temperature, force, pressure, vapor density, and magnetic fields have been "sensed". For example, under stress (tension or compression) a SAW delay line or resonator will change length. When used as the feedback element in an oscillator, this change in length relates to a phase change around the oscillator loop causing a shift in output frequency. When the outputs of two oscillators (one where the SAW is stressed and the other is used as a reference) are mixed together to provide an IF output, a fairly linear stress sensor can be obtained (See Figure 6). The advantages of this technique are wide dynamic range, good signal to noise ratio, and an IF ouptut not dependent on temperature as long as the temperatures of both SAW devices track.

SAW TAPPED DELAY LINES

SAW delay lines can be used to provide coding schemes. Figure 7 demonstrates how PN type codes can be formulated by switching the polarity of the transducer fingers in each bit. Fixed codes containing more than 1000 bits have been fabricated. Figure 8 is a plot of the correlated output of a 13 bit Barker code SAW device. Programmable matched filters have been made by interfacing with hybrid switching circuits.

SAW CONVOLVERS

Programmable SAW correlators use monolithic SAW convolvers to provide the maximum flexibility for waveform programming, being optimized for PN phase coded waveforms used in spread spectrum communications and phase coded radar. Figure 9 is a schematic of a programmable SAW correlator. The input signal, s(t), at the IF frequency (fo) is applied to one input port of a monolithic SAW convolver while a locally generated reference signal, r(t), is applied to the second input port. The output signal, c(t), is the convolution of the reference with the input signal provided that the two signals coexist within the device processing time window of ΔT. ΔT relates to the length of the pickup plate as seen in Figure 10 of an actual SAW convolver. Because of the nature of the SAW convolver, the output signal is at twice the frequency of the input with twice the input bandwidth and is compressed in time by a factor of two. If the reference signal is the time reversed replica of the input signal, the output signal is then the desired autocorrelation of the input. Figure 11 is an example of a correlated peak of a 63 chip Bi-Phase code. Normally the PN code is a maximal sequence and a large number of codes and time reversed references can be generated from simple shift registers. Codes to 1000 or more chips in length at chip rates from a few megaHertz to over 100 MHz can be correlated with essentially instantaneous programmability.

SAW DISPERSIVE DELAY LINES

Analog pulse compression using SAW dispersive delay lines (DDL) is a common technique for optimizing the range, resolution, and signal to noise performance of pulsed radar. Subsystems can be configured with both expansion and compression channels. Figure 12 is a schematic of such a subsystem. An unweighted DDL is used in the expansion channel to generate the linear FM (chirp) signal to be transmitted. The conjugate (opposite slope) DDL is used in the compression channel to perform matched filter signal processing. Often, a weighting function to reduce sidelobes is designed into the compression DDL. Figures 13 and 14 show typical frequency responses of an expander/compressor pair while Figure 15 is the systems' compressed pulse.

SAW DDLs are suited for dispersions under 100 μsec, wide bandwidths up to 500 MHz and center frequencies up to 1 GHz.

FUTURE TRENDS

Several major trends are extending SAW technology into new markets. High frequency (>400 MHz) SAW components are being developed to meet the challenges of cellular radio[4], higher IFs and microwave links. Successful and economical manufacture of these high frequency devices will require a totally integrated approach to include every part of the process, its equipment and environment.

New materials are being developed to optimize parameters such as temperature stability and better piezoelectric coupling for specific applications. Zinc oxide[5] can be sputtered on low cost substrates at high production rates. Lithium tetraborate[6] has a high SAW coupling and low temperature coefficient of delay. This material should be useful in oscillators. Berlinite[7] is another material useful for oscillators because it is temperature compensated and has a piezoelectric coupling higher than quartz. In addition, the integration of SAW structures on gallium arsenide[8] has allowed the basic acoustic functions of delay, tapping and filtering and the basic electronic functions of amplifying, summing, weighting and memory to be implemented in a single substrate.

The economical manufacture of SAW devices with grooved structures (e.g., reflective array compressors, resonators and buried IDT filters) in large quantities will become a great challenge for many SAW component houses because of the complex fabrication process and in-process testing requirements in order to obtain the high performance necessary in state-of-the-art RF systems.

CONCLUSION

SAW components offer the RF engineer a viable product in the 10 MHz to 1100 MHz frequency range. Understanding the many advantages and functional applications of this technology allows the RF engineer great design alternatives in his attempt to meet the requirements of state-of-the-art communication and radar systems of the 80's.

BIBLIOGRAPHY

[1] "High Volume TV-IF Filter Design Fabrication and Applications", G. Toboika et.al., 1984 IEEE Ultrasonics Symposium Proceedings, to be published.

[2] "Wideband Radar Signal Processor Based on SAW Convolvers", I. Yao et.al., 1984 IEEE Ultrasonics Symposium Proceedings, to be published.

[3] SAW Filters for Military and Spacecraft Applications", W. J. Tanski, F. C. Meyer, and L. P. Solie, Microwave Journal, February 1982, pp 53-66.

[4] "High Performance SAW Filters with Several New Technologies for Cellular Radio", M. Hikita, et.al., 1984 IEEE Ultrasonics Symposium Proceedings, to be published.

[5] "Surface Acoustic Waves Bandpass Filters on ZnO/Pyrex Substrate with Zero Temperature Coefficient", Ph. Defranould, 1983 Ultrasonics Symposium, pp 341-344.

[6] "Lithium Tetraborate - A New Temperature Compensated Piezoelectric Substrate Material for Surface Acoustic Wave Devices", N. M. Shorrocks, et.al., 1981 IEEE Ultrasonics Proceedings, pp 337-340.

[7] "Temperature Dependence of the Material and Acoustic Wave Parameters of Berlinite", D. S. Bailey et.al., 1983 IEEE Ultrasonics Symposium Proceedings, pp 335-340.

[8] "GaAs SAW/MESFET Programmable Tapped Delay Line", S. W. Merritt et.al., 1984 IEEE Ultrasonics Symposium Proceedings, to be published.

THREE PARTS OF A SURFACE ACOUSTIC WAVE DEVICE

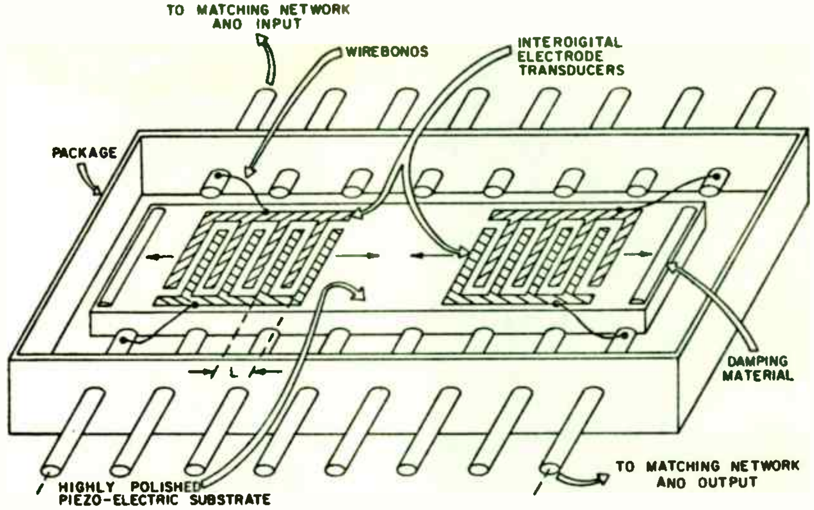


FIGURE 1

TABLE I

Surface Wave Characteristics of Various Substrates

Material	Cut	Propagation Direction	Coupling Coefficient k^2 (%)	Temperature Coefficient (ppm/°C)	Velocity (m/s)
Quartz (HC)	−20° Rotated Y	X	0.25	−32	3209
Quartz (ST)	+42.75° Rotated Y	X	0.16	(Note 1)	3157
$LiTaO_3$	Y	Z	.74	−35	3230
$LiNbO_3$	Y	Z	4.5	−94	3488
$LiNbO_3$	−128° Rotated Y	X	5.2	−72	3990
$Bi_{12}GeO_{20}$	110	001	0.85	−140	1620

Notes:
1. TC = .03 ppm/(Δ°C)^2

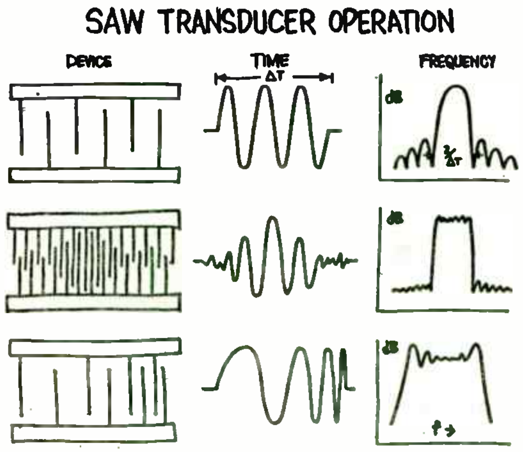


FIGURE 2

The Impulse Response, ΔT, is related to the Frequency Response by the Fourier Transform

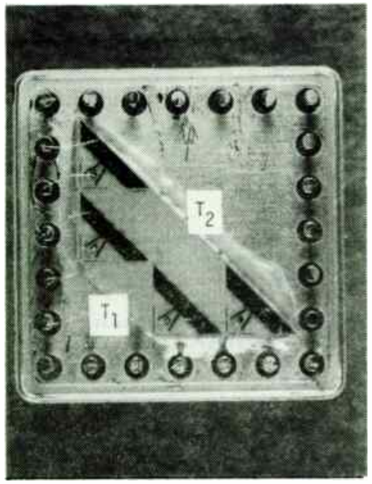


FIGURE 3

Dual Wideband SAW Delay Line
T_1 = 1.5 μsec, T_2 = 3.0 μsec
fo = 700 MHz, $BW_{@1dB}$ = 180 MHz

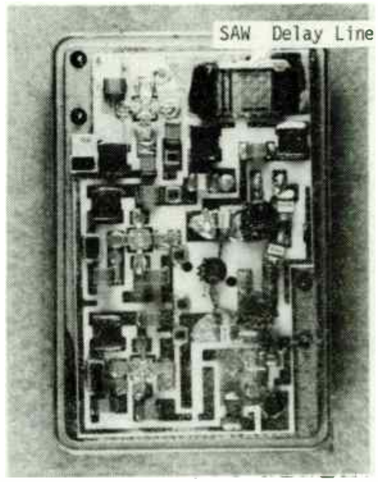


FIGURE 5

SAW Hybrid Oscillator

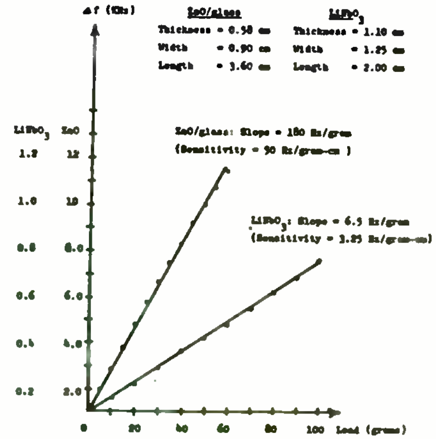


FIGURE 6 *

Effects of strain on dual-channel SAW oscillators on $LiNbO_3$ and ZnO-on-glass.

*Taken from "Bulk and Surface Acoustic Wave Sensors", C. Chuang, PhD. Thesis, Univ of California, Berkeley, 1982, pg 94.

TABLE II
APPLICATIONS OF SAW DEVICES

DELAY LINES

Fusing, MTI Radar, communications path length equalizer, altimetry, time ordering, target simulation.

FILTERS

Color TV, Radar, CATV, Repeaters, Transponders, ECM.

OSCILLATORS, RESONATORS

Stable sources (VHF to microwave), local oscillators for communications and coherent radar, sensors.

TAPPED DELAY LINES

Fourier transformation, acoustic image scanning, clutter reference radar, ECM deception.

CONVOLVERS

Synchronizer for spread spectrum communicators, Fourier transformation.

DISPERSIVE DELAY LINES

Radar pulse expansion/compression, variable delay systems, spectrum analysis, compressive receiver, chirp z transform (CZT), adaptive bandpass filters.

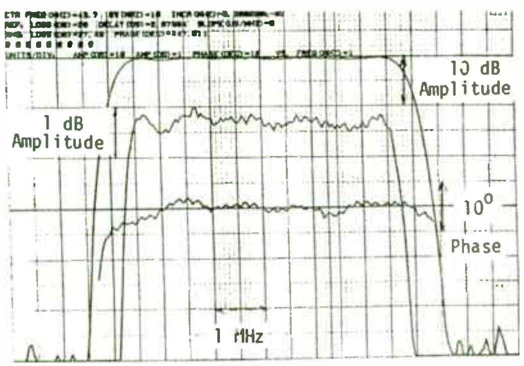
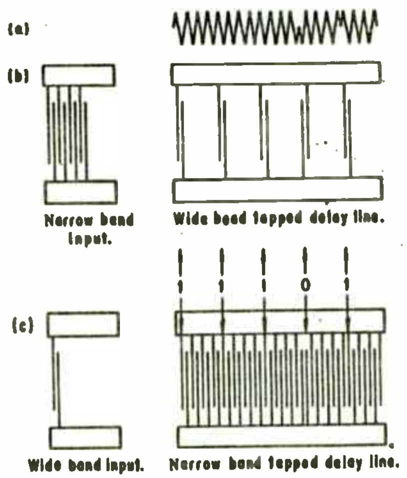


FIGURE 4

Frequency Response of a SAW VSB Filter



Phase Coding Schematic

FIGURE 7

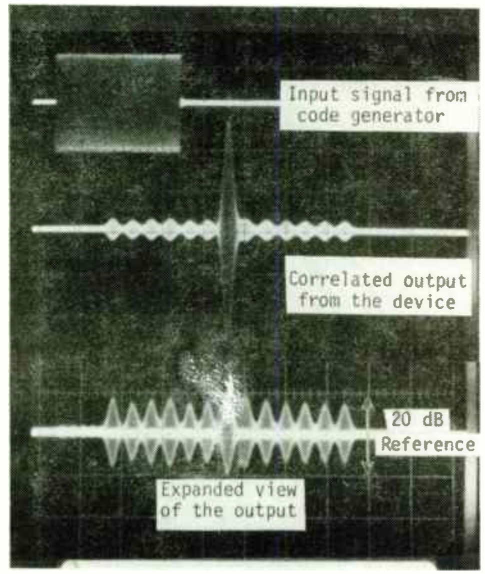
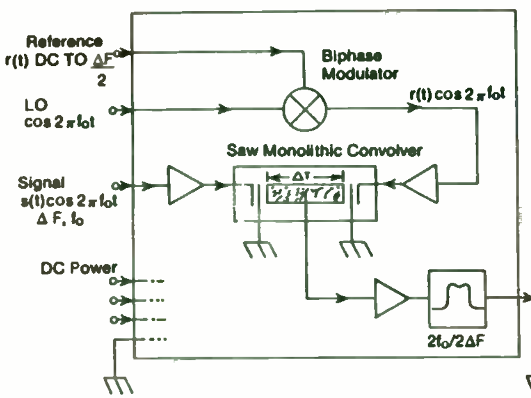


FIGURE 8

Correlated output of 13-Bit Barker coded SAW device



$$c(2t) = \cos 4\pi f_0 t \int_{\Delta T} s(\tau)\, r(2t-\tau)\, d\tau$$

If $r(t) = s(-t)$ and is bounded by ΔT
$c(2t) = \Psi_{ss}(2t) \cos 4\pi f_0 t$
Where $\Psi_{ss}(2t)$ is the auto correlation function of $s(t)$ compressed in time by a factor of two

FIGURE 9

Schematic of SAW Programmable Correlator

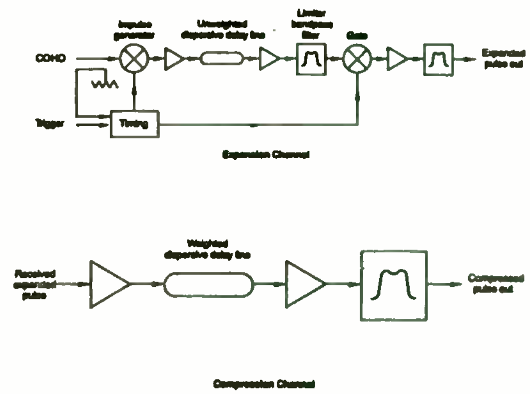


FIGURE 12

Typical Configuration for a Pulse Expansion/Compression Subsystem

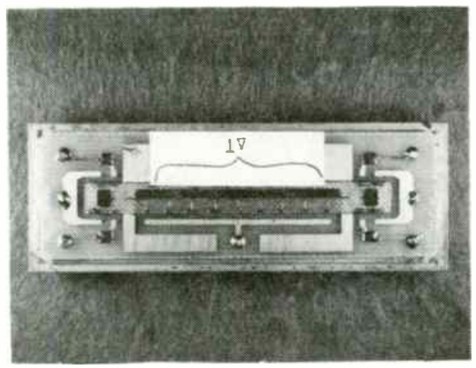


FIGURE 10

Monolithic SAW Convolver showing pick-up plate, ΔT

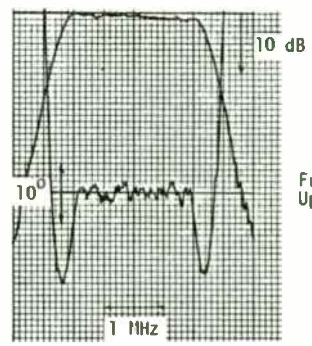


FIGURE 13

Frequency Response of Unweighted Upchirp SAW ODL (Expander)
fo = 31.5 MHz

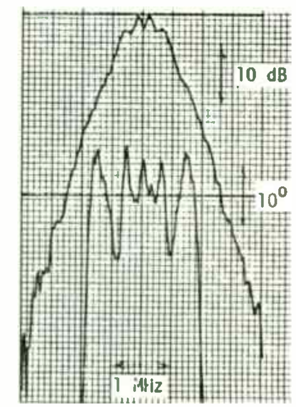


FIGURE 14

Frequency Response of Weighted Downchirp SAW ODL (Compressor)
fo = 31.5 MHz

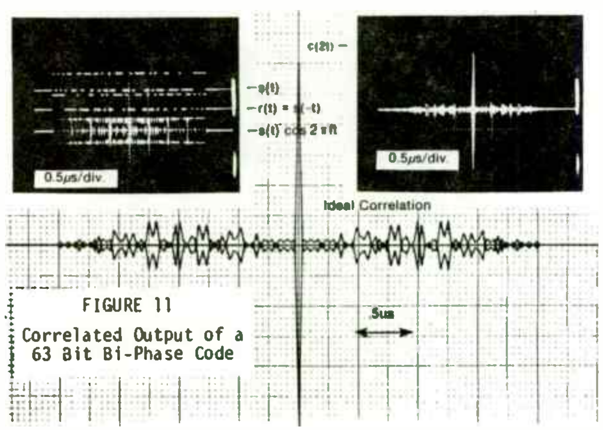


FIGURE 11

Correlated Output of a 63 Bit Bi-Phase Code

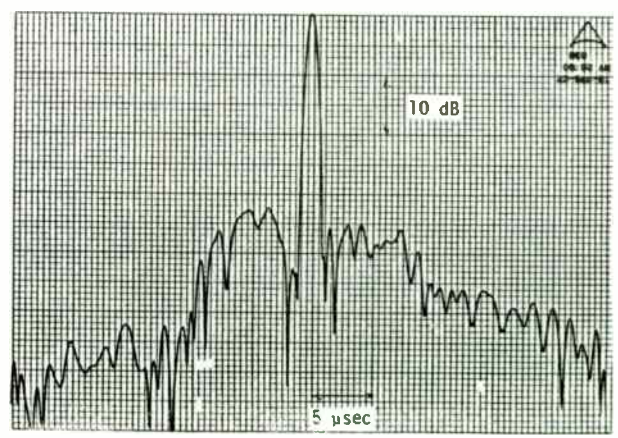


FIGURE 15

Compressed Pulse of the Subsystem

PROGRAMMABLE RF SIGNAL PROCESSORS FOR SPREAD

SPECTRUM COMMUNICATION SYSTEMS

J. LATTANZA, F.G. HERRING, P.M. KRENCIK, A.F. CLERIHEW

Hazeltine Corporation

Research Laboratories

Greenlawn, N.Y. 11740

(516) 261-7000

ABSTRACT

This paper describes the results of applying surface acoustic wave (SAW) technology and custom LSI RF control circuits to the problem of implementing small size, low cost RF signal processors for tactical communication equipment. Physical characteristics and electrical performance are described for two representative processors comprising a programmable RF correlator for the matched filter detection of spread spectrum signals and a bandpass transversal filter having a programmable center frequency. Additionally, it is shown how the concept can be extended to the design of programmable HF/VHF notch filters and transversal equalizers.

INTRODUCTION

Programmable matched and transversal filters are key processors in tactical spread spectrum systems. These filters provide the processing gain necessary for jam resistant operation and the electronic programmability essential for a secure system. Small size, low cost, and operation in a military environment are additional requirements imposed by tactical systems.

This paper describes the design and performance of both a matched and a transversal filter that meets these requirements by using a custom-designed LSI chip to provide the programmability for a surface acoustic wave (SAW) tapped line. This combination of bipolar LSI and SAW technology results in programmable filters with near theoretical electrical performance, substantially reduced size and cost, and excellent reproducibility.

The matched filters described will meet the requirements of most airborne, vehicular, or manpack systems. The volume, power dissipation, cost, and electrical performance are superior to either their digital correlator counterpart or to SAW analog convolvers. Additionally, the matched filter is designed to operate at carrier frequencies up to 300 MHz and is therefore capable of processing code rates up to 100 Mb/second. These performance levels cannot, presently, be achieved by digital or CCD correlators.

Historically, programmable matched filters implemented with surface acoustic wave technology have realized the peripheral electronic circuits as thick film hybrids, using multiple transistor chips and standard SSI integrated circuits. In large time-bandwidth systems, this configuration is both large and costly since both RF switching and logic control must be provided at each tap, i.e., one tap per unit of time-bandwidth. Parasitic elements, inherent in this implementation, also limit the useful frequency range to lower IF frequencies. These constraints have been surmounted with the design of a custom LSI chip. This chip provides programmable switching for 32 contiguous taps and it can be cascaded directly to permit matched filtering for PN coded words having time bandwidth products of over 1000.

The performance capability provided by this design is illustrated by description of a 128-bit device operating at 80 MHz with a 12.8 Mb/s code rate and a 256-bit device centered at 240 MHz with a code rate of 64 Mb/s. Additional parameters, for the matched filters, are summarized in Table I. The experimental results which are presented demonstrate the near-theoretical performance obtained.

The implementation and preliminary electrical performance of a programmable transversal filter, operating as an agile bandpass filter in the 55 to 65 MHz range, is described. Description of the custom-designed LSI chip used in these filters is also presented.

PROGRAMMABLE LSI/SAW MATCHED FILTERS

Implementation

The basic implementation of the correlator is that of an electronically programmable SAW tapped delay line having one (switchable) tap for each chip of the input-coded word. The total length (delay) of the line is made equal to the length (time duration) of the phase-coded input signal and the tap spacing is made equal to the length of one chip in the coded word. The phase at each tap (i.e., through each tap to a summation node) matches the phase code of the signal and the device operates to compress the time duration of the input signal. Electronic programmability is provided by active circuits contained on a custom designed LSI chip. For a constant amplitude input signal, all taps are weighted equally and the only programmable function required is the phase.

Table I. PMF Parameters

Parameter	Device 1	Device 2
Number of Taps	128	256
Tap Spacing	78.125 ns	15.625 ns
Code Rate	12.8 Mb/s	64 Mb/s
Code Length	10 μs	4 μs
Center Frequency	80 MHz	240 MHz
Programming Rate (max)	25 Mb/s	25 Mb/s
Modulation Waveform	MSK	MSK
Power Dissipation	2.0 watts	4.0 watts
Form Factor	1.75" x 0.75" x 0.140"	2.9" x 2.3" x 0.75"
Volume	0.18 cu in.	5.0 cu in.

A block diagram illustrating this scheme is shown in figure 1. The SAW line is an ST-cut quartz crystal whose metallization pattern consists of an input transducer and an array of equally spaced taps. The input transducer is apodized to provide the matched filter spectral response for an MSK waveform. Both the launcher and taps employ a split-finger configuration to minimize reflections. The metallization pattern is obtained by a photolithographic process after a 1500-Angstrom aluminum thin film has been E-beam evaporated on the active surface of the crystal.

The electrical signal is coupled from the SAW line by wire bonding each tap

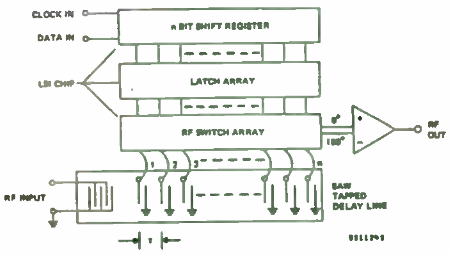


Figure 1. Programmable SAW Matched Filter, Simplified Block Diagram

to the control circuits contained on a custom-designed LSI chip. Binary phase control (0° or 180°) is obtained by feeding each tap output through an RF switch to either one or two differentially connected summing buses. Each RF switch, in turn, consists of two UHF bipolar transistors and many (32) of these switches are contained on a single LSI chip. The switch condition, and thus the binary phase of the tap, is controlled by the parallel output of a shift register (SR) that has been serially loaded with the proper binary code. In practice, the parallel output of the SR is actually transferred into a holding register that controls the individual RF switches. This action permits a code change to be loaded into the SR while the SAW line is the matched filter to the previous (latched) code.

Custom Switch Chip

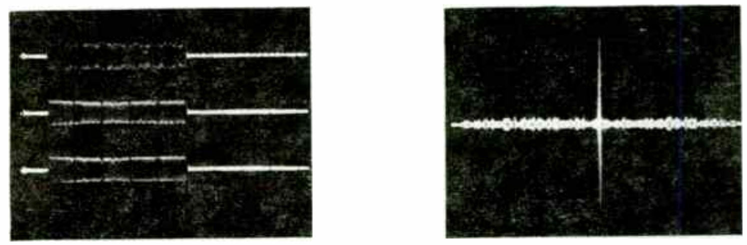
The excellent electrical performance, low volume, and low cost result from the development of a fully customized LSI chip for these filters. This integrated circuit, designed using 6 micron, washed emitter, junction-isolated bipolar technology, contains 32 logic-controlled RF switches and on-chip summing amplifiers. The chip layout is specifically designed to permit the chips to be cascaded easily and facilitate wire bond connection to the SAW line.

The chip measures 224 mils long and 130 mils wide with 48 I/O pads. The physical form factor and cell layout was dictated by RF considerations associated with the interconnection of the RF Switch and the SAW line. The interconnection lead length should be a minimum to reduce the parasitic inductance, and the interconnections should be of uniform length to reduce phase variations. The best way to achieve this is to have the RF pads along one edge of the chip facing the SAW line. Other I/O pads should be arranged that the chips can easily be cascaded along the length (propagation axis) of the SAW crystal to provide switching for any number of taps.

All of these features were incorporated in the design as illustrated by the chip topography shown in the photograph of figure 2. The RF input pads, seen along the bottom edge of the chip, are on 6.6 mil centers and all the RF and logic circuits required to switch a single tap are contained within a cell 5.8 mils wide.

To simplify the interface with external circuits, both the clock and transfer inputs were designed to be TTL compatible. Data inputs are selectable as either TTL or voltage translated ECL. The data output from the shift register is only supplied as voltage translated ECL. The advantages of voltage translated ECL are higher clock speeds and lower power dissipation.

The transfer function of the RF switch has a 3-dB degradation at 304 MHz and an ON/OFF isolation of 20 dB at 250 MHz. The chip was designed to operate over a full temperature range of -55° to +125°C. In addition to the 32 logic-controlled RF switches, latch



(a) Expanded Pulse Waveform (Impulse Response) for PN Code (top trace), All Zeros (middle trace), and All Ones (bottom trace) Scale: 2 μs/cm.

(b) Correlated Output for 128-Bit MSK Input Signal (Note 20 dB sidelobe level.) Scale 2 μs/cm.

Figure 4. Waveforms for 128-Bit MSK Programmable SAW Matched Filter

evidenced by the near equal response to these codes and the flatness (better than 1 dB) of the impulse responses.

Although not evident from the photograph, the baseline spurious is down by over 30 dB. Figure 4(b) is the correlated output for a 128-bit PN MSK input waveform. The sidelobe level is 21 dB down and this is within 1 dB of the theoretical response. The measured 6 dB pulse width is 138 ns, compared to the theoretical value of 123 ns.

Figure 5 is a plot of the theoretical loss to the correlated peak of the compressed pulse over the temperature range for the quartz SAW line. The shape of this curve is a result of the parabolic temperature dependence of an ST-cut line having a zero slope temperature coefficient at 27.5°C. The loss at the temperature extremes can be reduced, at the expense of a slightly increased mid-temperature loss, by designing the metallization mask to match an offset carrier frequency. This offset frequency results in a match (zero loss) at two temperatures symmetrically displaced from 27.5°C. This tradeoff is shown in figure 5 for

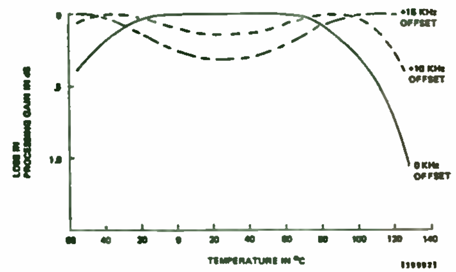
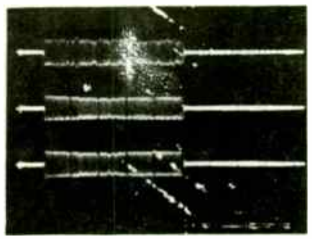


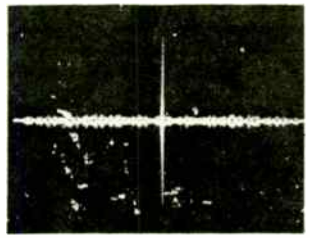
Figure 5. Processing Gain Loss vs Temperature

offsets of +10 and +15 kHz. The oscilloscope photographs of figure 6 demonstrate the performance achieved for a +10 kHz offset frequency design and the loss of the correlated peak at -55° and +125°C is seen to approximate the curve of figure 5.

One advantage of the LSI/SAW implementation is the high dynamic range that is possible as a result of the linear property of the system. Table II, below, shows the dynamic range to internal (transistor) noise and to



(a) Expanded Pulse Waveform (Impulse Response) for PN Code (top trace), All Zeros (middle trace), and All Ones (bottom trace) Scale: 2 μs/cm.



(b) Correlated Output for 128-Bit MSK Input Signal (Note 20 dB sidelobe level.) Scale 2 μs/cm.

Figure 4. Waveforms for 128-Bit MSK Programmable SAW Matched Filter

evidenced by the near equal response to these codes and the flatness (better than 1 dB) of the impulse responses.

Although not evident from the photograph, the baseline spurious is down by over 30 dB. Figure 4(b) is the correlated output for a 128-bit PN MSK input waveform. The sidelobe level is 21 dB down and this is within 1 dB of the theoretical response. The measured 6 dB pulse width is 138 ns, compared to the theoretical value of 123 ns.

Figure 5 is a plot of the theoretical loss to the correlated peak of the compressed pulse over the temperature range for the quartz SAW line. The shape of this curve is a result of the parabolic temperature dependence of an ST-cut line having a zero slope temperature coefficient at 27.5°C. The loss at the temperature extremes can be reduced, at the expense of a slightly increased mid-temperature loss, by designing the metallization mask to match an offset carrier frequency. This offset frequency results in a match (zero loss) at two temperatures symmetrically displaced from 27.5°C. This tradeoff is shown in figure 5 for

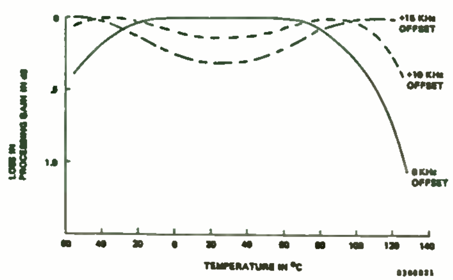


Figure 5. Processing Gain Loss vs Temperature

offsets of +10 and +15 kHz. The oscilloscope photographs of figure 6 demonstrate the performance achieved for a +10 kHz offset frequency design and the loss of the correlated peak at -55° and +125°C is seen to approximate the curve of figure 5.

One advantage of the LSI/SAW implementation is the high dynamic range that is possible as a result of the linear property of the system. Table II, below, shows the dynamic range to internal (transistor) noise and to

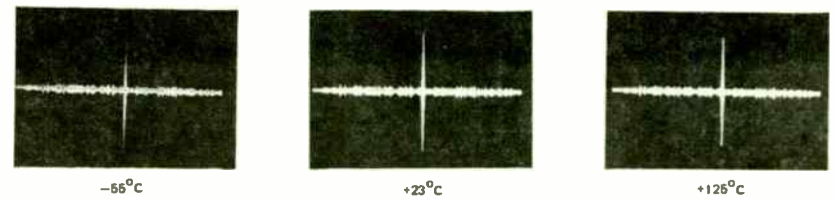


Figure 6. Temperature Performance of Matched Filter

Table II. Measured Input-Output Characteristics

80 MHz PMF	
Output Noise Level	-79 dBm
Maximum Output Signal (1 dB Compression)	-6 dBm
Dynamic Range to Noise	73 dB
Output Clock-Noise Level	-56 dBm
Dynamic Range to Clock-Noise	50 dB
Insertion Loss to Correlated Peak	26 dB

parasitically coupled clock noise to be 73 and 50 dB, respectively.

The dynamic range to clock noise is important in those applications which require a new code to be loaded during the time of an anticipated epoch (e.g., continuous code change). The measured insertion loss (expanded pulse input to compressed pulse output) of the PMF is 26 dB.

Another advantage of the linear properties of the device is the absence of any processing gain (PG) loss due to quantitization or sampling. For an MSK waveform, the theoretical PG is:

$$PG\ (dB) = 10 \log \left[\left(\left(\frac{\pi^2}{16}\right) \cdot \frac{1}{T_c}\right)(128\ T_c)\right] \quad (1)$$

where T_c is the chip duration and $(\pi^2/16)(1/T_c)$ is the noise bandwidth. For a chip rate of 12.8 Mb/s:

$$PG\ (dB) = 18.9\ db \quad (2)$$

The measured processing gain at room temperature was 18.6 dB or 0.3 dB less than theoretical. The processing gain at the temperature extremes was not actually measured. However, since both the sidelobe profile and the shape of the main lobe remained unchanged, it is reasonable to assume that the amplitude response of the system is the same, and that any loss in processing gain is bounded by the reduction in amplitude of the correlated peak at the temperature limits (0.1 and 0.3 dB).

Performance of 240 MHz PMF

The 240 MHz PMF uses 8 LSI chips to correlate a 256-bit, PN-coded, MSK input waveform. The code rate and tap spacing are 64 Mb/s and 15.625 ns respectively. Figure 7 is a photograph of the matched filter. Note that the LSI circuitry and other discrete components are resident on a thick film hybrid that supports both the digital and analog RF signals. The SAW crystal is located in the center of the package. The input to the SAW tapped delay line consists of a center-fed, cosine weighted, 7.5 λ transducer that

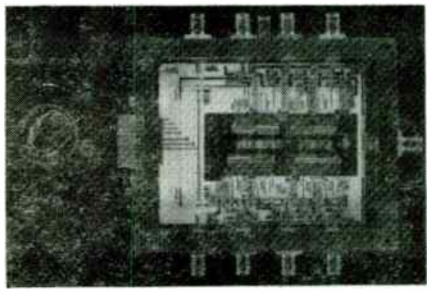


Figure 7. Programmable SAW Matched Filter, 256 Bit

launches an acoustic wave in each direction along the propagating axis of the crystal. The transducer employs a split electrode configuration and the radiating aperture measures 2.54 mm.

The SAW artwork was designed and layed out using a Calma Interactive Graphics System and is pictured in figure 8. Since this device has a high code rate which requires very close tap spacing (49.3 microns), the resulting tap density makes it impractical to cascade LSI chips in the same manner as done in the 80 MHz device. To alleviate this problem, the bidirectional property of the input transducer is used to advantage. An array of 128 active taps is located on each side of the input transducer. These taps are spaced correctly in the time domain to realize 256 evenly spaced chips. Dummy taps are inserted between active tap sections to ensure timing accuracy between the two sides of the device. Fanout structures are also used to conveniently interconnect each SAW tap to the RF input pad of the LSI chip.

The electrical performance of the matched filter is summarized in the oscillograph of figure 9. Figure 9 (a) top trace illustrates the impulse response of the PMF programmed for a direct M-sequence PN code. The middle and bottom traces show the PMF impulse response for an all ones and all zeroes, 256 bit code respectively. The flatness of the response illustrates the balance and uniformity of the LSI chip. Note that there are less than 5 inoperative taps in the 256 tap array. Any discontinuities appear at the chip-to-chip interface, and improvements could be made by trimming the output load resistors in each of the eight sections. Figure 9 (b) top trace illustrates the input 4 μs 256 bit MSK code. The middle trace shows the compressed pulse and near sidelobe level.

The sidelobe level is 19 dB down and this is within 0.5 dB of the theoretical correlated response (see figure 10). The excellent symmetry of the

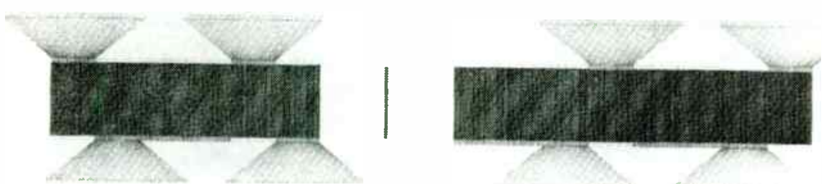
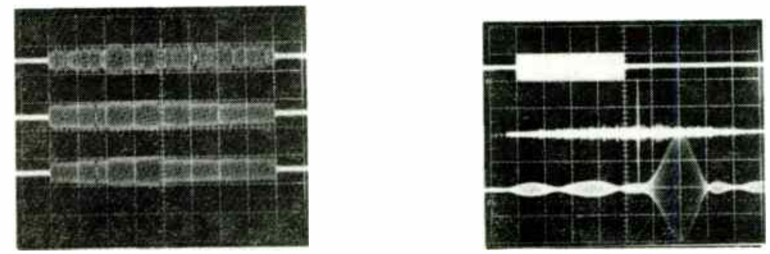


Figure 8. Programmable Matched Filter, SAW Mask Layout



(a) Expended Pulse Waveforms (Impulse Response) for PN Code (top trace), All Ones (middle trace), and All Zeros (bottom trace). Scale: 0.5 μs/cm.

(b) Top Trace: Input 256 MSK Code Scale: 1 μs/cm. Middle Trace: Correlated Output for 256 Chip MSK Input Scale: 1 μs/cm Bottom Trace: Correlated Output-Mainlobe Scale: 20 ns/c.

Figure 9. Waveforms for 256-Chip MSK Programmable SAW/LSI Matched Filter

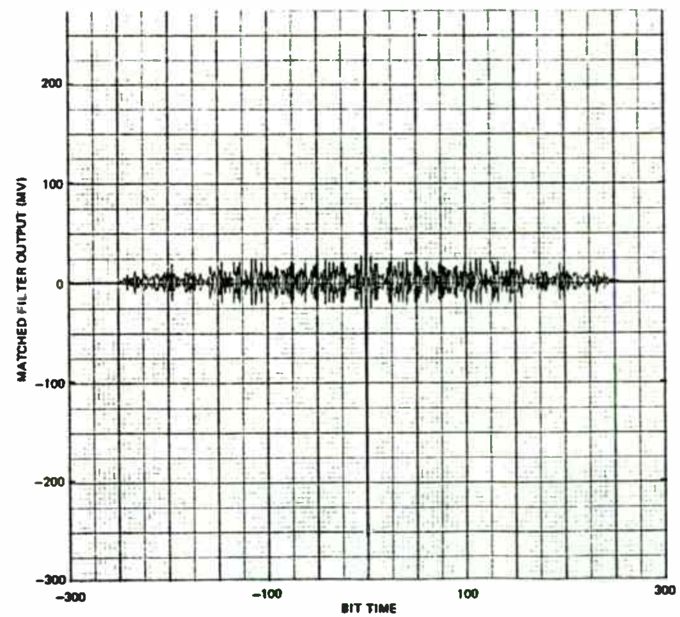


Figure 10. Theoretical 256 Chip Response

correlated main lobe is also shown in the bottom trace of figure 9. The 6 dB pulse width measures 30 ns which compares favorably to the theoretical value of 25 ns.

The measured insertion loss (expanded pulse input level to compressed pulse output level) of the PMF is 30 dB. The measured processing gain at room temperature was 21.5 dB. This compares favorably to the 21.98 dB theoretical value for a 256 bit MSK line. Table III summarizes additional test results for the programmable filter.

PROGRAMMABLE TRANSVERSAL FILTER

As a result of the success on the design of the two PMFs, work was undertaken to extend the basic concept to the design of programmable transversal filters (PTF). Three types of transversal filters are possible and these include bandpass and bandreject

Table III. Test Results of 240 MHz PMF

Characteristic	Measurement
Insertion loss	30 dB
Sidelobe level	-19 dB
Processing gain degradation	0.48 dB
Compressed Pulse Width (6 dB)	30 ns
Direct RF Feedthrough (below peak)	-35 dB
Clock Feedthrough	-40 dBm
Spurious level (below peak)	-47 db
Input/Output Impedance	50 ohms

filters, having center frequencies under program control, and transversal equalizers which are capable of synthesizing an arbitrary amplitude and phase characteristic. The programmable feature permits open loop (logic) control or closed loop control as part of an adaptive processor.

Theory

Figure 11(a) shows the basic transversal configuration. It consists of a (SAW) delay line having 2n+1 weighted taps separated by a delay, T. Following each delay interval, the signal is sampled, weighted in both amplitude and phase (A_k), and summed to form the output. The general input-output relationship is

$$y(t) = \sum_{n=0}^{N} A_n\, x\,(t - nt) \qquad (3)$$

and the impulse response is

$$h(t) = \sum_{n=0}^{N} A_n\, \partial\,(t - nt) \qquad (4)$$

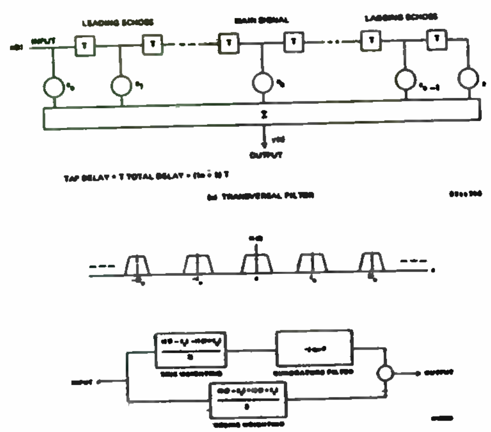


Figure 11. Transversal Filter Configuration

To obtain the frequency response of the filter, it is only necessary to take the Fourier transform of h(t). Thus,

$$H(f) = \sum_{n=0}^{N} A_n e^{-j2\pi f n T} \qquad (5)$$

H(f) is clearly periodic with period $f_o = 1/T$. If H(f) is designed for a low pass response at f = 0, then this window will be repeated at $f = \pm f_o$, $\pm 2 f_o$, $\pm 3 f_o$ . . . as shown in figure 11(b). If it were desired, for example, to have the nominal center of the filter at 60 MHz, then the possible choices for f_o are

$$f_o = \frac{60\ \text{MHz}}{n}\ ;\ n = 1,2,3\ .\ .\ . \qquad (6)$$

which correspond to the following values of T,

$$T = \frac{n}{60\ \text{MHz}}\ ;\ n = 1,2\ .\ .\ . \qquad (7)$$

The value of n is then selected as the largest integer which results in only one bandpass window within the programmable range of the filter. This value is given by

$$n = \text{INT}\ \frac{f_c}{\Delta f + (B/2)} \qquad (8)$$

where INT is the integer part of the expression

f_c is the center frequency

Δ_f is the programmable range

B is the filter bandwidth

and for typical parameters is.

$$n = \text{INT}\ \frac{60\ \text{MHz}}{10\ \text{MHz} + 1\ \text{MHz}/2} = 5 \qquad (9)$$

Smaller values of n can be used but these result in more taps than necessary. Excessive taps must be avoided since electronic weight control is required for each tap used.

Given a particular bandpass characteristic, the actual tap weights can now be found using a computer optimization technique known as the Remez Exchange Algorithm. It is now necessary to define the configuration of the PTF that will result in a programmable center frequency.

Let H(f) denote the entire frequency response of the tapped transversal filter (including all repetitions of the window). If the impulse response of the filter is multiplied by $\cos 2\pi f_1 t$, the resulting frequency response will be

$$1/2\ H\,(f - f_1) + 1/2\ H\,(f + f_1) \qquad (10)$$

Note that from equation (4), this is equivalent to multiplying A_n by $\cos 2\pi f_1 t$.. Similarly, if h(t) is multiplied by $\sin 2\pi f_1 t$, the resulting spectrum is:

$$1/2j\ H\,(f - f_1) - 1/2j\ H(f + f_1) \qquad (11)$$

Now consider the implementation shown in figure 11(c). Clearly, the frequency response of this filter is given by

$$W(f) = 1/2H\,(f-f_1) + 1/2H\,(f+f_1) + \text{sgn}f\left[1/2\ H(f+f_1) - 1/2H\,(f+f_1)\right] \qquad (12)$$

or

$$W(f) = \begin{cases} H\,(f - f_1)\ ;\ f > 0 \\ H\,(f + f_1)\ ;\ f < 0 \end{cases} \qquad (13)$$

For positive values of f_1, all positive frequency windows are shifted to the left by f_1 and all negative frequency windows are shifted to the right. For negative values of f_1, the reverse will be true. Thus, the capability of shifting the window up or down in frequency is obtained by using two identical transversal filters in the two branches of the composite filter. One branch multiplies the weights by a cosine function; the second branch multiplies the weights by a sine function, and additionally, adds a quadrature phase shift. The frequency f_1 is the programmable aspect of the filter.

Programmable Bandpass Filter

A SAW transversal, implemented as a weighted-tap tapped delay line, can be used to synthesize a bandpass filter characteristic having a programmable center frequency. The SAW taps are apodized to produce the desired center frequency bandpass characteristic using the Remez-Exchange algorithm. It is usually desirable to have a large number of taps (i.e., a long impulse response) in order to obtain a generally rectangular bandpass characteristic and a low (frequency-domain) sidelobe level. Using 63 taps, a filter having the characteristics shown in table IV was synthesized. The practicality of a design having such a large number of taps is a result of the availability of the 32-tap LSI chip described earlier.

The physical configuration of the PTF is illustrated in figure 12. There are four parallel tracks on a single SAW crystal and each track contains 63 apodized taps. The taps on each track are binary phase weighted and summed by the two 32-tap logic controlled RF switch chips. Each track is then assigned a binary weight and the four tracks are then summed. This configuration effectively provides for a 4-bit (3 bits plus sign) quantized signal at each tap. The tap weights are then cosine-modulated in an I channel and sine-modulated in a Q channel to produce the desired frequency shift.

The computer-plotted frequency response for the design selected is shown in figure 13. This is the center frequency response and includes the effects of the 4-bit precision tap weights. The effect of the finite tap weight precision (4 bits) is to degrade the out-of-band sidelobe level as the center frequency is shifted. The worse case degradation is an out-of-band response that is 30 dB down.

Table IV. Programmable Center Frequency Transversal Filter

Center Frequency	55 to 65 MHz
Bandwidth	1 MHz
Shape factor	1.5:1
In-band ripple	0.5 dB
Stop-band rejection	40 dB
Number of taps	63

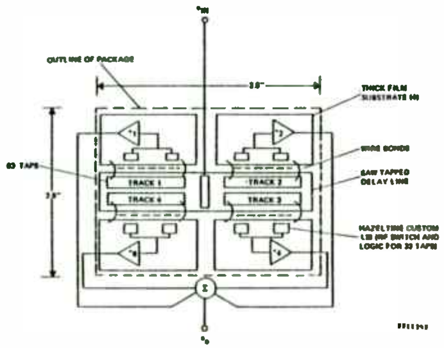


Figure 12. Configuration for 63-Tap Programmable SAW Transversal Filter

The PTF layout was performed on a Calma GDS-II interactive graphics system, and the configuration is shown in the photograph of the final unit (figure 14). In this implementation, two SAW crystals were employed and each crystal contained two tracks of 63 taps. The center fed input transducer is five wavelengths and each tap consists of 2 finger pairs. The aperture length is 200 mils, all electrodes are split (6.5 micron width) to minimize reflections, and the tap spacing is 83 nanoseconds or 10.3 mils. The input matching networks, SAW crystals, (eight) LSI chips and the interconnections are contained on a 2-inch by 2-inch thick film hybrid substrate.

Electrical Performance

Electrical testing performed to date is not complete but shows promising results. The major problem is electrical leakage that results because of the high insertion loss from the input transducer to the output of the tap. This loss is on the order of 55 dB and can be reduced by increasing the number of finger pairs in a tap or by using a lower loss piezoelectric material such as lithium niobate ($LiNbO_3$).

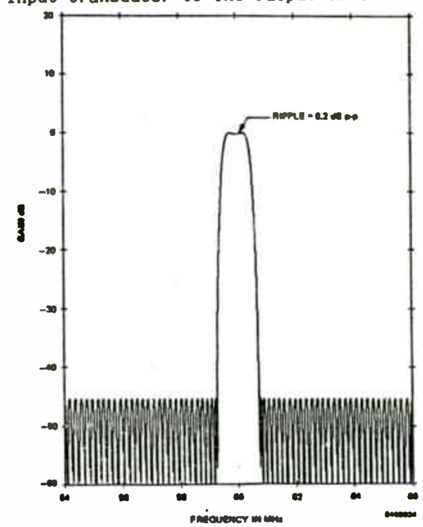


Figure 13. 63-Tap Bandpass Filter Response

The impulse response of the filter is shown in figure 15 and is in close agreement with the theoretical response. The (center) frequency response is shown in figure 16. The high sidelobe level in figure 16 is a result of leakage, and it is planned to correct the problem before proceeding with further testing.

Conclusion

This paper has shown that a hybrid configuration based on SAW and full-custom LSI technology is a viable method of producing miniature, low cost, high performance matched filters for use in tactical spread spectrum equipments. Near theoretical electrical performance can be obtained and was reported for two designs having 128 and 256 taps and 12.8 and 64 Mb/s code rates, respectively.

The design concept for implementing a matched filter can be extended to the general case of a programmable transversal filter. These devices are

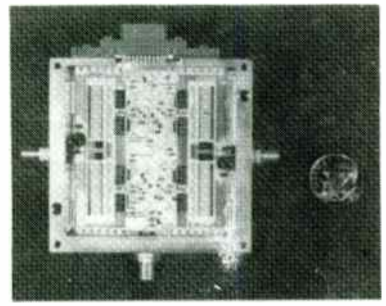


Figure 14. Photograph of Programmable Transversal Filter

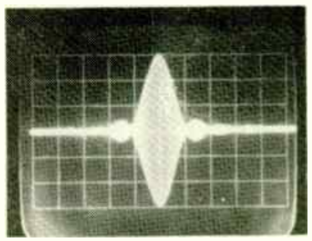


Figure 15. PTF Impulse Response

useful as an agile bandpass filter, a rejection or notch filter for interference removal, and for the purpose of amplitude and phase equalization as required to improve intersymbol interference or effect multipath cancellation in communication channels. The design approach was described and theoretical and preliminary experimental data presented for a 63-tap programmable center frequency bandpass filter.

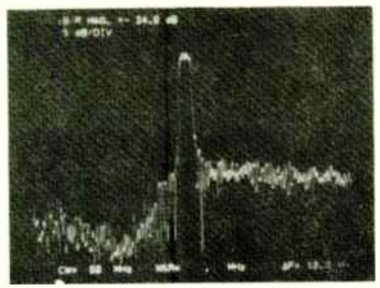


Figure 16. PTF Frequency Response

SAW Filter Applications.

Not too many engineers are deeply interested in the physics of the SAW technology, more interested in the applications and technology benefits for the job which they have in hand for their own engineering problem. This paper is intended to present applications of the SAW technology rather than how the devices are designed and fabricated

SAW filters are used as resonators, filters and oscillators for a wide range of different applications. The main features being small size, non tuneable and fully characterised. Each of the following applications will be covered -

1. IF Filters
2. Front end filters.
3. Filter banks.
4. Modems.
5. P.C.M.
6. Television
7. Satellite Receivers
8. Spread spectrum communications
9. Radar applications

As with many other things in engineering, compromises and tradeoffs must be made. Whilst SAW filters have many advantages, they also have problems associated with them. This paper explains the inherent problems and some ways of overcoming them, for example the spurious signals and insertion loss.

The Author

This paper is presented by Mr. Ron Towns who is the SAW Sales Director at CRYSTAL TECHNOLOGY. Ron Towns has been working in Surface Waves since 1974 and has seen engineering activity in Plessey, Signal Technology, Siemens and Crystal Technology.

SAW Filter Applications.

INTRODUCTION Not too many engineers are interested in the Physics of the SAW technology, more interested in the applications and technology benefits for the job which they have in hand for their own engineering problem. This paper is intended to present applications of the SAW technology rather than how the devices are designed and manufactured.

BACKGROUND SAW components handle key functions in entertainment electronics and professional telecommunicationu engineering.

Besides their widespread use as frequency stabilizing devices and as bandpass filters in the VHF and UHF ranges, SAW commonents are also employed to implement complex signal-processing functions like fast correlation of fixed or programmable signal forms. SAW components come as resonators, transversal filters, dispersive delay lines and as convolvers.

They are used in television transmitters and receiving stations, CATV and satellite TV installations, as well as in both analog and digital telecommunications networks, radar equipment and highly sophisticated electronic surveillance receivers.

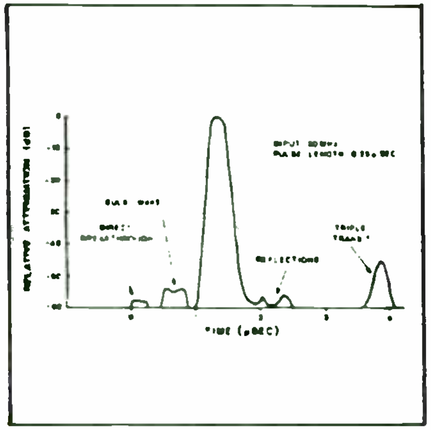
For the user the benefits are as follows :

* small size
* unbeatable precision
* extremely high reproducibility
* superior shape factors
* linear phase
* no tuning

These components are produced on crystal or lithium-niobate substrates and are notable for their excellent long term stability and very low temperature drift.

As with many things in engineering, comprimises and trad-offs must be made. Whilst SAW filters have many advantages, they also have problems associated with them. The engineer must be familiar with the inherent problems and how to overcome them.

IF FILTERS Saw filters are ideal solutions to some IF filter requirements. In most developments, to date, filter insertion losses in the range from 15 to 30 dB are more normal than the very low loss types which are now being offered by SAW companies. Newcomers to the SAW technology must understand that SAW filters are transversal filters which operate in the time domain.

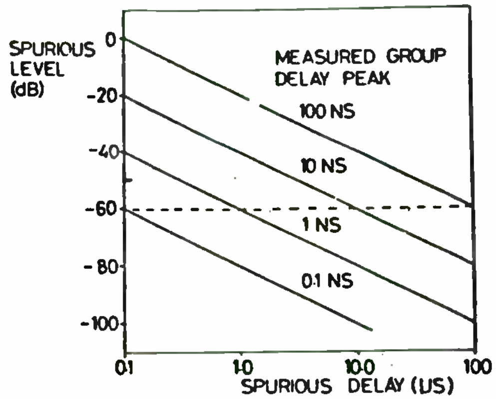


time domain response

The time domain response shows that as well as the main filter response there are other responses caused by electromagnetic radiation from the input transducer to the output transducer, the bulk wave which travels across the filter faster than the surface wave, some low level signals caused by internal reflections and finally the time echo of the main filter response which has travelled three times across the filter. There is a connection between the insertion loss, the spurious signals and the in band amplitude and group delay ripple. A very useful rule of thumb is that the triple travel echo is twice the insertion loss plus 6 dB and for this reason, the engineer usually has to tolerate a high insertion loss. It is possible to reduce the filter insertion loss by impedance matching the transducers to the source and load impedance. Usually, a simple series inductor will suffice but for broad band filters a more complex impedance transformer is necessary.

The in band amplitude ripple depends only on the amplitude of the spurious signals but the group delay variation increases with the delay to the spurious signal. The table and chart show the relationship between the spurious signal levels and the amplitude and group delay ripples.

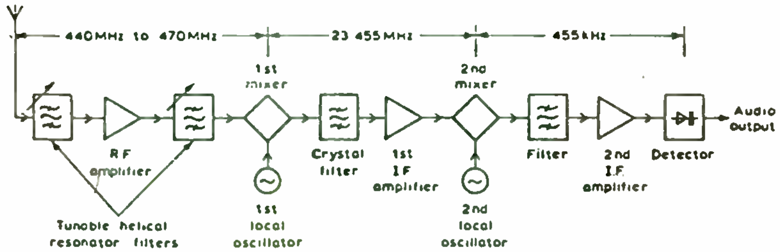
Insertion loss	-20dB	-30dB	-40dB	-50dB	-60dB	
Ampl ripple	1.7dB	.44dB	.17dB	.05dB	.02dB	p.p.



In most applications spurious suppression to 40 or 50dB should result in a satisfactory small amplitude variation. This corresponds to an insertion loss of around 20dB.

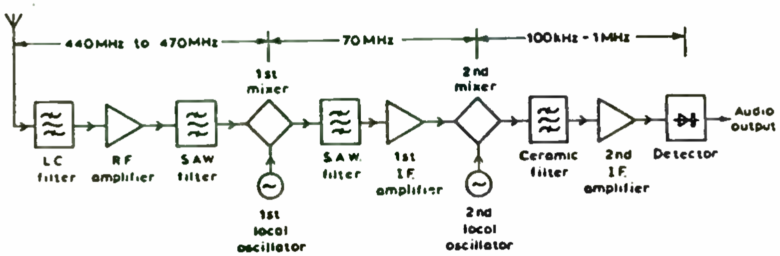
FRONT END FILTERS More modern design techniques have produced lower loss filters which are more suited to front end applications. Insertion loss has been one of the chief restrictions limiting the wider use of SAW filters and engineers can expect to see new filters emerging which can be used in portable radio telephones, paging receivers and cordless telephones.

A typical portable radio telephone operating in the 440 to 470 MHz frequency band has an architecture which looks like the following schematic diagram.



The receiver is designed to operate on a single channel and, despite its apparent simplicity, it would not be attractive to use a TRF receiver with fixed filters offering a few tens of kilohertz of bandwidth. This is because it would require a large number of filters at a large number of defferent frequencies which would be uneconomic. A double conversion superhet is used instead, with IF of 23.455 MHz and 455 kHz. Helical filters have offered the narrowest bandwidth of the conventional types of tunable filter and typically four of them are used, in two pairs, as front end filters. They would, typically, provide a 4 MHz bandwidth and 70dB image rejection for the first IF which, as they are not particularly selective, has tcbe at the high frequency of 23.455 MHz. The first IF filter would, typically, be an eigth order crystal filter composed of four monolithic duals providing all the adjacent channel rejection in a 25 kHz channel, 70 dB rejection. In this receiver, the second stage of frequency conversion is carried out mostly to reduce power consumption and ease the task of detection but it requires the crystal filter to provide 70 dB image rejection for a second IF of 455 kHz. This can be achieved with some care as crystal filters tend to suffer from spurious responses.

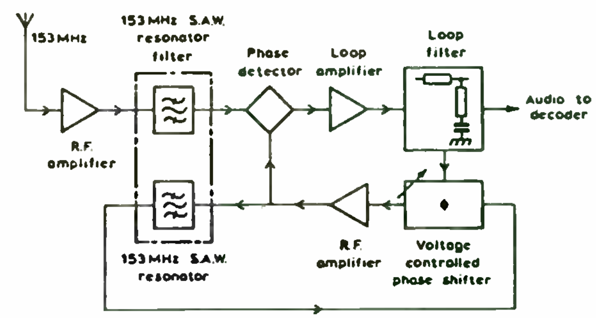
The use of SAW filters can bring about some improvement to these points and the next schematic shows the configuration of a modified receiver.



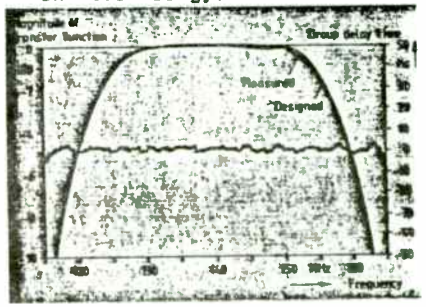
The helical filter has been replaced by a wide band low loss SAW filter offering frequency response characteristics which are not found in conventional filters. The filter would have sufficiently wide bandwidth to cover the complete 440 MHz to 470 MHz frequency band and, more importantly, a sufficiently narrow transition bandwidth to permit the use of a 70 MHz first IF. The insertion loss would be around 6 dB with 65 dB of image rejection. The filter would have 50 ohms input and output impedance and would be packaged in a TO-5 can. A dramatic reduction in volume would be achieved.

A 70 MHz first IF filter could also be fabricated in SAW. This would be a second order narrowband SAW resonator filter which, with a 25 kHz channel spacing, provides 25 dB of adjacent channel rejection and at least 60 dB image rejection for a second IF of betwenn 100 kHz and 1 MHz. The insertion loss would be around 10 dB which would be deliberately increased from minimum in order to keep the size down and hence the costs. A furthur 45 dB of adjacent channel rejection would be provided by a ceramic filter at the second IF, which could be achieved in a very compact and economical manner at 455 kHz.

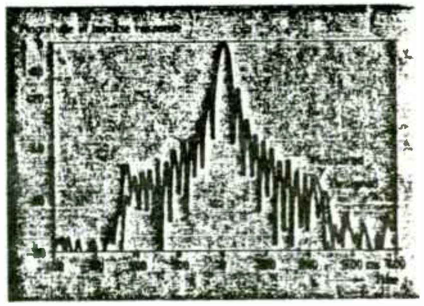
Paging receivers can also benefit from SAW filters. With pagers, there are a large number of receivers and few channels. Narrowband SAW resonators can be used in dual conversion receivers but since low cost is the most important parameter together with low power consumption an alternate receiver architecture of the direct conversion type are worthy of more consideration. The block schematic shows how SAW could be incorporated into a paging receiver.



DATA TRANSMISSION The highest bit rate planned for transmitting digital signals via radio link systems is 140 Mbit/s, which is sufficient for 1920 telephone channels. By means of a 16-stage quadrature amplitude modulation (QAM) process , the 3 dB bandwidth of the spectrum is narrowed down to 35 MHz. Half of the spectrum shaping is performed an the modulator and half in the demodulator section using filters with cosine shaped edges. The filter in the modulation section also compensates the sinx/x shaped spectrum of the digital modulation signal. The high data rate and the 16 possible QAM states make high demands on spectrum-shaping filters for optimum reception. Within the passband of the filters the magnitude of the transfer function should be undistorted and the group delay time constant.This application is ideal for the SAW technology.



Magnitude and group delay time of the transfer function of a spectrum-shaping filter for the demodulator stage of a 140 Mbit/s digital radio link system



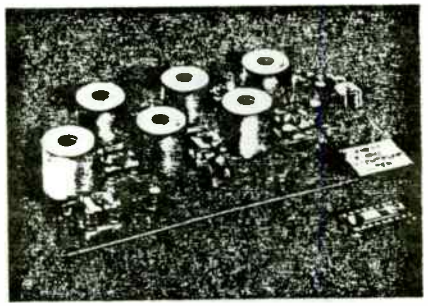
Magnitude of the impulse response of a spectrum-shaping filter for the demodulator stage of a 140 Mbit/s digital radio link system

The curves show a 35 MHz bandwidth SAW filter which is intended to be used in the demodulator section of a 140 Mbit/s digital radio link system. The transfer function deviates by only 0.1 dB from its design value. The group delay time as constant apart from a ripple of 1 ns. The magnitude of the measured and designed impulse response of the filter shows a good agreement at the specified zeros.

Yet another example of the application of SAW technology to telecommunication engineering is the narrow-band spectrum-shaping filter for digital radios with low bit rates. The LC filter requires selected, bulky and elaborately tuned components. The SAW, in contrast, are fabricated on quartz, they are small, precise and thermally stable.

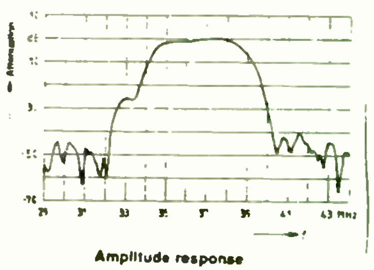
Comparison of narrowband spectrum-shaping filters for 2 Mbit/s digital radio systems:

LC filter (rear)
SAW filter on quartz (front)

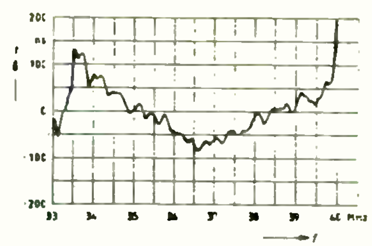


Optical fiber systems promise economical transmission of long-haul digital traffic at bit rates of 100 to 1000 Mbit/s and more. These systems use regenerators along the route to restore the magnitude, shape and timing of message pulses which are degraded in the fibre optic transmission spans. The regeneration process is controlled by a timing waveform that is extracted from the pulse train itself. The task of timing extraction at the high bit rates is accomplished by the use of a SAW filter. As well as meeting the short term jitter requirements, the SAW filter has the long term aging and thermal stability to make it suitable for use in ocean systems which have 20 to 25 year lifetimes.

TELEVISION The single biggest example of commercial success of the SAW technology is to be found in television transmission and reception. Because SAW filters can be manufactured in high volume using techniques similar to integrated circuits they are extremely repeatable and sufficiently low cost to attract the consumer electronics industry. The TV IF filter incorporates the nyquist slope and the sound shelf. Other options are parallel sound (seperate vision and sound channel) for stereo TV sets. All these applications place high demands upon the 2 T pulse step signal.



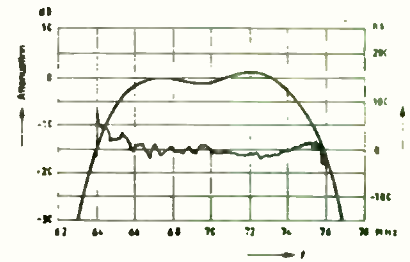
Amplitude response



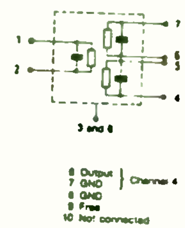
Group delay

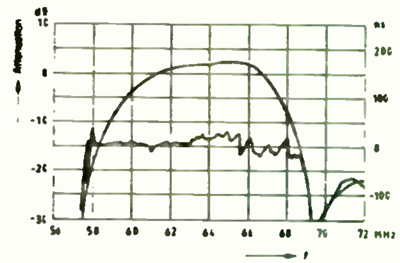
When ever a TV signal has to be modulated onto or demodulated from a RF carrier, two sidebands will be generated. SAW filters can have extremely good shape factors with flat passbands. These filters are called vestigal sidband filters and are used in the TV transmitter and the TV transposer. The full range of transmission standards have been made available by the SAW industry ; B/G for western europe, I for Great Britain, L for France, M for USA etc.

The advent of cable television has brought about the need for channel filters. The cable converter supplies the TV set with chrominance, luminance and sound information onto a single UHF channel. Usualy this is either of channels 2, 3 or 4. Some SAW filters contain two channel filters in one package having one input and two outputs. An interesting new development would be to have a channel filter with seperate sound and vision outputs which would allow the sound level to be controlled in the cable converter instead of the TV IF.



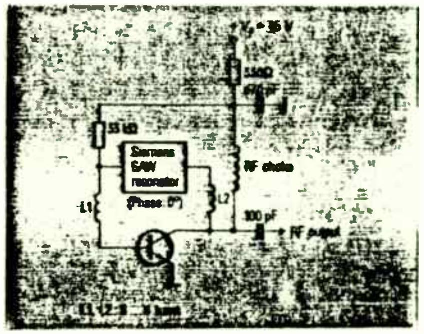
Channel 4
Attenuation
and group delay



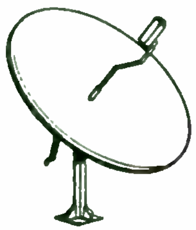


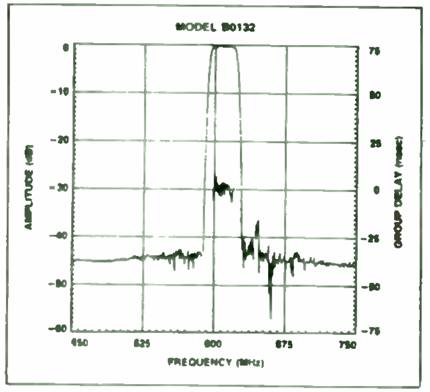
Channel 3
Attenuation
and group delay

The cable converter also uses another SAW device, the SAW resonator, as the local oscillator to hetrodyne the TV signal onto the channel. These devices are small, stable and inexpensive. They are simple to use, requiring only one transistor to maintain oscillation.

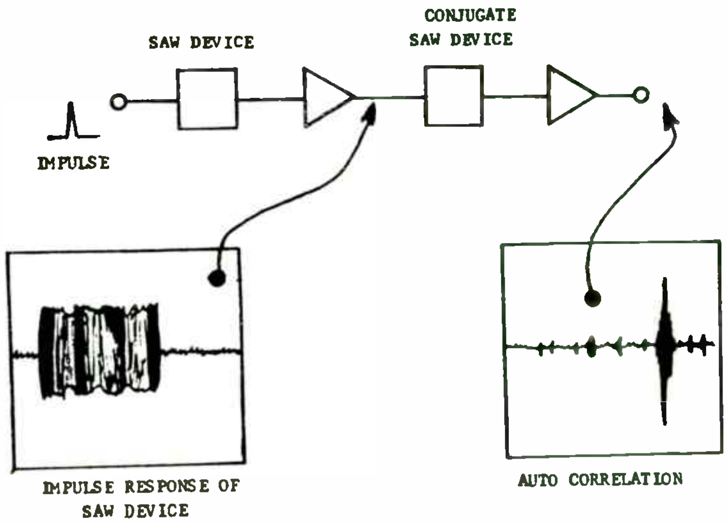


SATELLITE RECEIVERS The satellite receiver is becoming a piece of consumer electronics just like the VCR was a few years ago. The acceptance by the consumer is due, in part, to the price reduction of the satellite receiver over a short period of time. SAW IF filters have become standard components, mostly at a 70 MHz, with a bandwidth of 36 MHz. This is an extremely wide band and difficult to implent using LC components. The real advantage of SAW in this application is small size, rugged and very well shielded and requires no tuning. Some IFs are at higher frequencies, for example 134 and 610 MHz but with the same bandwidth of between 24 and 36 MHz.





SPREAD SPECTRUM COMMUNICATIONS Spread spectrum modulation is an important communication technique and is also one area of communication technology which has derived great benefit from SAW technology. At the transmitter in a direct-sequence spread-spectrum (DS-SS) communication system, a periodic code sequence (usually binary), with a digit rate which greatly exceeds that of the message data, is used to expand the transmitted signal bandwidth. A receiver can use either an active correlator or a passive matched filter (e.g. SAW), matched to the code sequence, to "de-spread" the signal to the original data bandwidth. The receiver therefore has a processing gain, given approximately by the ratio of the spread-to-despread bandwidths. The receiver is a matched filter which maximizes the signal-to-noise ratio at the bit decision or epoch event. The concept of epoch is shown below.



The generation and matched filter reception of a N chip spread spectrum sequence using SAW devices is shown. The receiver epoch is defined as the time when the received sequence exactly fills the matched filter, the conjugate SAW device. The filter is maximized as shown by the large correlation peak. In a simple data transmission system, where data is conveyed by the presence or absence of a group of PN chips, the receiver examines the matched filter output at each epoch instant for the presence or absence of a correlation peak.

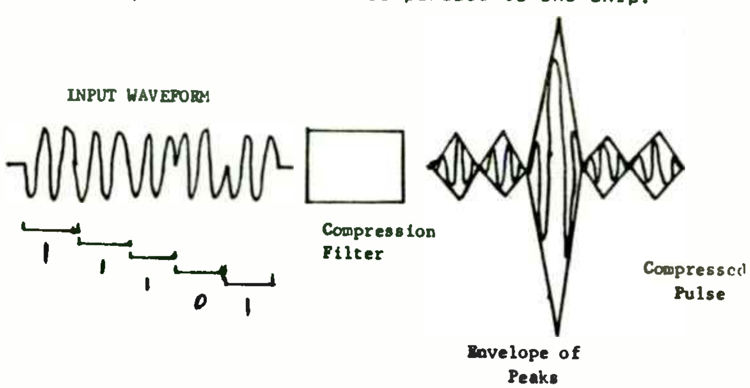
The output of a transversal filter is actually the convolution of the input waveform with the impulse response of the filter

$$w(\tau) = \int_{-\infty}^{+\infty} u(t)v(\tau - t)\,dt$$

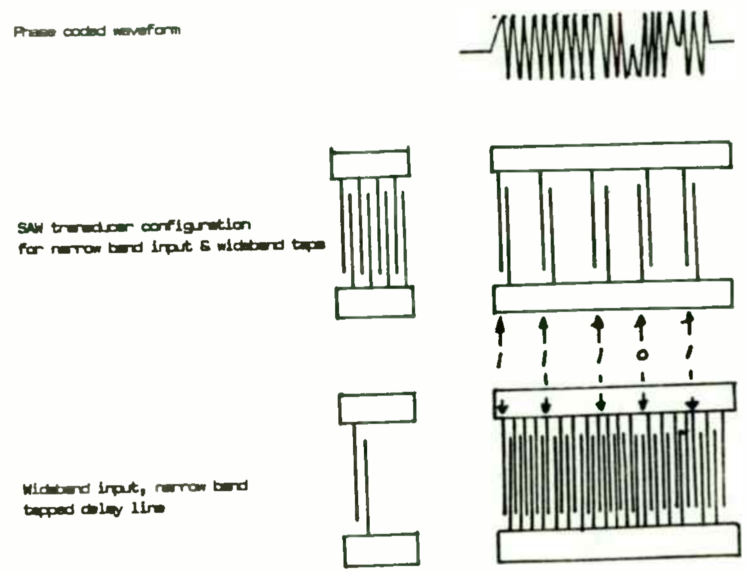
where u(t) is the input waveform and v(t) is the impulse response of the filter. Here, τ is the time at which the output response is to be measured and includes the time delay necessary in a practical SAW device. To achieve the autocorrelation function, for a matched filter, it is necessary to realize instead the correlation integral

$$g(\tau) = \int_{-\infty}^{+\infty} u(t)u(t - \tau)\,dt$$

Thus, it is necessary to construct a filter with in impulse response v(t) that is the time inverse of the signal u(t) to be correlated or matched (excepting the arbitrary time delay). The correlation of a five-chip coded sequence, below, results in a compressed pulse with a width comparable to one chip.

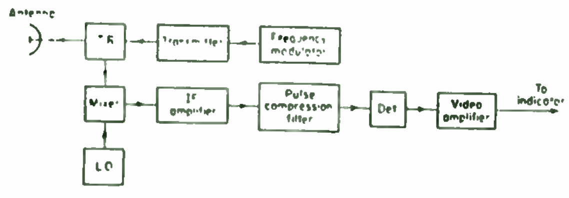


The SAW matched filter is fabricated, normaly as a phase coded delay line. The simplest case is where a 180 degree phase shift represents the difference between a 1 or a 0 (most codes considered are biphase). However, SAW has the flexibility to achieve phase coding with arbitary values of phase as well as arbitrary values of amplitude. A simple 5 chip biphase sequence generator and its implementation in SAW is shown below.



The conventional SAW device is a linear transversal filter having one electrical input port and one electrical outport. An electrical impulse applied to the input transducer produces a sinusoidal waveform similar to the coded waveform shown above.

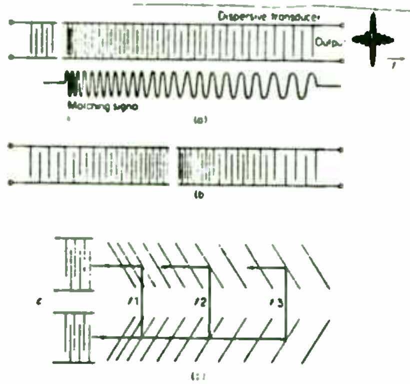
RADAR The phase coded delay line, discussed earlier has application in radar systems. The major difference is that the phase code in normally continiously swept and not changed in discreet chips. If the optimum use is to be made of the transmitting power, the solid-state output stages have to be operated with the greatest possible pulse duty factor. Pulse compression techniques allow high transmitting power and good target discrimination at the same time. Furthurmore, pulse compression reduces the effect of noise sources by the amount of the compression gain.



Block diagram of an FM pulse compression radar

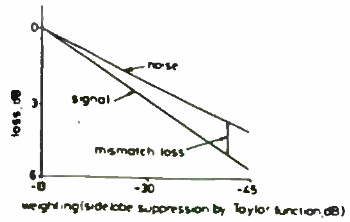
SAW dispersive filters are now commonly used in modulation of the transmitted signal and compression in the receiver. The high accuracy of the amplitude and phase in the filter pulse response is a prerequisite for small secondary sidelobes in the compressed pulse and thus jointly determines the dynamic range of the target acquisition.

SAW pulse compression filters are produced both in interdigital transducer (IDT) and reflective array compressor (RAC) format. The RAC version is shorter in length and also exhibits a number of inherent advantages; the input and output impedances are independent of the filter transfer function and secondary effects from bulk waves, diffraction and undesired reflections are negligible. Also, the RAC has the facility for a phase correction metalisation film to be incorporated.

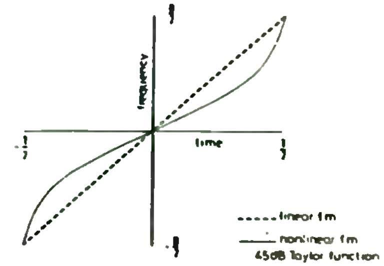


Three basic forms of SAW interdigital transducers for linear FM pulse compression (a) Dispersive delay line with dispersion designed into one transducer; (b) dispersion in both transducers, and (c) a reflective array compressor (RAC)

When sidelobe suppression is achieved by spectral weighting in a receiver which is otherwise matched to a linear chirp, the signal is reduced in amplitude. The graph below shows the loss in signal and noise when passing through a receiver providing both pulse compression for a linear-chirp input and sidelobe suppression by means of a series of Taylor functions. As the design-sidelobe suppression is taken toward the limit of -45 dB, the signal is attenuated more repidly than the noise. The difference in the two amounts of attenuation relates directly to a loss of system sensitivity and is called mismatch loss.



Signal and noise loss through filters giving sidelobe suppression with Taylor weighting functions



Frequency-time laws for chirp signals employing linear f.m. and nonlinear f.m. with 45 dB sidelobe suppression

To recover the mismatch loss, a non-linear FM waveform must be considered. The frequency-time excursions for both linear and non linear FM chirps are shown above. The non-linear law is derived using a 45 dB Taylor function. The non-linear chirp is seen to have a higher spectral content at the centre of the band by having a chirp rate in that region. The other point which can be seen is the increased sensitivity of no-linear chirp to doppler shift.

An example of a RAC device pair (expander and compresser) which have been optimised for doppler insensitivities and mismatch loss is shown on the next page. With the PC-set that has been produced, it was possible to shorten the transmit signal duration through optimisation by some 20 micro secs at a given signal/noise ratio ; with a required duration of the compressed pulse of 1.7 micro secs and a real compression gain of 14 dB, the transmit pulse duration would be between 60 and 70 micro secs if Linear FM were used. The combination of amplitude and phase weighting has produced, as a result of the reduction in mismatch loss, a signal of only 44 micro secs.

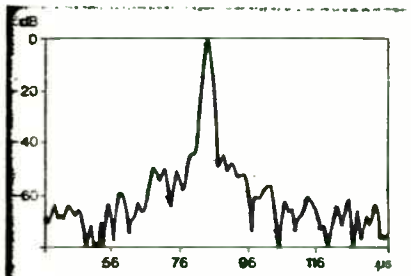


Fig 8:
Compressed pulse of the measured pulse responses

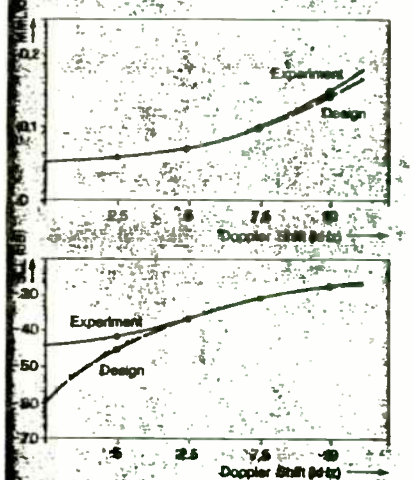
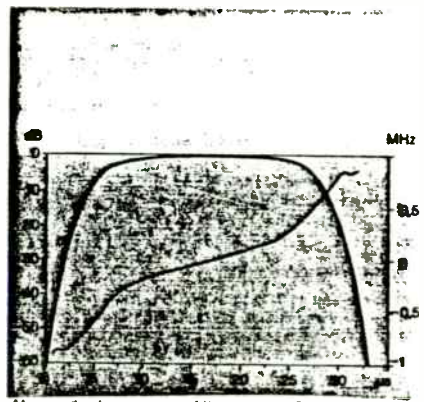
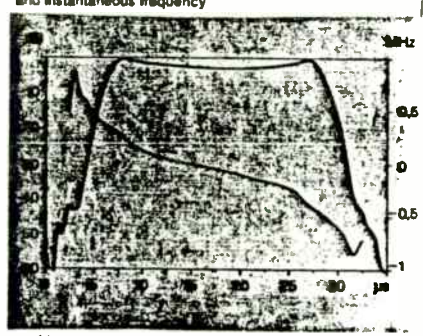


Fig. 9:
Mismatch loss and suppression of secondary sidelobes in the presence of Doppler shift



Measured pulse response of the expander filter in magnitude and instantaneous frequency



Measured pulse response of the compressor filter in magnitude and instantaneous frequency

PRIME APPLICATIONS OF SAW DEVICES

DEVICE	APPLICATIONS
Delay Line	Fusing, MTI Radar, Communications Patch Length Equaliser, Altimetry, Time Ordering
Wideband Delay Line	Recirculating digital storage
Bandpass Filter and Resonator	Color TV, Radar, Communications Satelite Repeaters, ECM, Frequency Synthesis
Oscillator	Stable Source VHF to Microwave - Communications and Radar
Tapped Delay Line	Fourier Transformation, Acoustic Image Scanning, Clutter-reference Radar, SSR, ECM Deception
Dispersive Delay Line ('Chirp')	Radar Pulse Compression, Variable Delay For Target Simulation, Fourier Transformation (Spectral Analysis). Compressive Receiver, Group Delay Equalisation
PSK Filter	Spread Spectrum Communications, Radar, Military ATC
Convolver	Synchroniser For Spread Spectrum Communications, Fourier Transformation

SAW STABILIZED OSCILLATORS

Thomas O'Shea and Jonathan Ladd
Sawtek Inc.
Post Office Box 18000
Orlando, Florida 32860

INTRODUCTION

The purpose of this paper is to present an overview of applications for SAW stabilized oscillators.

The number of applications for SAW delay line and SAW resonator controlled VHF-UHF Oscillators has been growing exponentially in the last several years. The primary reason for this growth is the simplicity of system design that results from their use. A SAW stabilized oscillator often eliminates all multiplier stages. This not only reduces complexity but also improves phase noise, short term stability and reliability. SAW resonator and SAW delay line oscillators are also more immune to mechanical shock than a bulk wave oscillator. A SAW local oscillator can be operated at higher drive levels than bulk wave oscillators which means that an intermediate amplifier is not required to drive a mixer.

The frequency versus temperature stability of a SAW controlled oscillator is often criticized as not being as good as AT cut bulk wave crystal controlled oscillators. Several means of improving the temperature stability of a SAW oscillator will be discussed in this paper.

A later section describes sub-systems utilizing SAW oscillators and comments on the emerging field of SAW sensors.

OVENIZED SAW OSCILLATOR

This section discusses a precision 740 MHz ovenized SAW oscillator that was developed by Sawtek and described in Reference 1 for application as a fixed frequency local oscillator in the transmitter of a navigation satellite. By its nature, this application requires high reliability, very good short term stability, good long term stability and low spurious and harmonic content over all environmental conditions. This application also requires a precise ability to set the carrier frequency.

Figure 1 shows a schematic of this SAW resonator controlled oscillator and buffer/amplifier. A single-port SAW resonator was chosen as the feed back element of this Pierce configuration oscillator rather than a SAW delay line or two-port SAW resonator. The single-port device has the lowest loss and highest Q of these choices and; therefore, results in the best phase noise. A second stage buffer/amplifier is lightly coupled to the oscillator stage to provide a stable +10 dBm output and immunity to load pulling. The output of 2nd stage is fed through an attenuator that serves as a minimum load on the amplifier and limits the output impedance. The signal is finally fed through a low pass filter to further limit the oscillator harmonics.

It is necessary to ovenize an oscillator of this type in order to provide optimum short term stability. However, to minimize power consumption a component oven is used to ovenize only the SAW resonator. This oven maintains the SAW resonator temperature at its turnover point which is a region of minimum deviation (Figure 2) of frequency vs. temperature. Figure 3 summarizes the achievements of this program. Most goals listed there were achieved and several firsts were established, particularly in set accuracy and long term stability.

Figure 4 is a plot of the short term stability expressed as Allen's Variance $[\sigma_y(\tau)]$ versus measurement averaging time for a second generation prototype of this oscillator. Parts in 10^{-11} are routinely achieved with these units. Figure 5 is the single side band phase noise at 10 Hz to 1000 Hz from the carrier. The SSB phase noise is -115 dBc/Hz at 1000 Hz offset and the rolloff rate is 30 dBc per decade. Figure 6 is included to demonstrate the long term frequency stability or "aging" of this oscillator. The curve is logarithmically decaying which is the characteristic of good aging. The first years aging is approximately 5 ppm. Five year aging is extrapolated to less than 15 ppm. This performance is believed to be "state-of-the-art" for a high frequency SAW oscillator of this size and power consumption. To the author's knowledge this is the first SAW oscillator scheduled to be launched in a satellite.

VOLTAGE CONTROLLED SAW OSCILLATORS

Many RF programs require that the frequency of the oscillator be variable. The frequency deviation requirement varies from a few parts per million (PPM) to over one thousand PPM. The choice between a SAW resonator or a SAW delay line as the frequency controlling feedback element primarily depends on required frequency pulling, oscillator phase noise, and short term stability. Table I gives a comparison of SAW resonator voltage controlled oscillators (VCO's) and SAW delay line VCO's.

TABLE I

COMPARISON OF RESONATOR VCO AND DELAY LINE VCO

Offset Frequency	RESONATOR VCO	vs. DELAY LINE VCO
	Phase Noise	
10 Hz	-55 dBc/Hz	-20 dBc/Hz
100 "	-85 "	-50 "
1000 "	-105 "	-85 "
10,000 "	-130 "	-105 "
	Short Term Stability	
(τ = 1 sec.)	1.10^{-9}	1.10^{-8}
	Frequency Tuning	
	10 - 100 PPM	100 - 2500 PPM

NOTE: Data presented is for oscillators with center frequency between 500 MHz and 600 MHz.

Frequency pulling of the SAW oscillator is accomplished by use of a varactor diode connected in series with the base tuning capacitor. The SAW resonator VCO schematic is shown in Figure 7 and the SAW delay line VCO is shown in Figure 8. The delay line circuit is more complex due to the necessity of impedance matching the device as well as achieving a 360° shift in the feedback loop. The resonator controlled oscillator is simpler in design because the resonator has fairly low loss and most good RF transistors have sufficient gain to overcome the unmatched resonator loss, hence, only the phase need be adjusted.

For wide band tuning a hyperabrupt junction varactor diode is chosen as the tuning diode. Hyperabrupt varactor diodes typically have a 10:1 capacitance change for a voltage variation of 12 volts. The disadvantage of using the hyperabrupt junction varactor is the frequency response curve often is logarithmic as shown in Figure 9. If this is undesirable to the system designer a varactor diode with an abrupt junciton can be used. This diode gives more linear tuning but the capacitance change is on the order of 5:1. Therefore, the frequency shift achievable with the abrupt diode is half of that achievable with the hyper abrupt diode.

SAW delay line VCO's are being used in wideband frequency synthesizers and communication systems. They are also ideal for use over very broad temperature ranges since the frequency shift due to temperature can be compensated by the control voltage input. Another advantage of the delay line VCO is it can often be used as both receiver local oscillator and transmit master oscillator in a radar synthesizer or communication system because it is broadband and can be tuned (or "slewed") very rapidly.

SAW resonator VCO's are finding their way into systems where the frequency stability and phase noise are paramount. Most systems used to-date center around frequency shift keying (FSK) or narrowband FM systems. The reduced circuit complexity makes the resonator controlled VCO very cost efficient.

TEMPERATURE COMPENSATED SAW OSCILLATORS

The frequency of a SAW oscillator varies with temperature, in a parabolic fashion, according to Equation 1 as

$$(1) \qquad f(PPM) = \left|\frac{T - To}{K}\right|^2$$

Where To is the temperature at which the frequency of the oscillator is maximum and K is a material constant that ranges from 5.4 to 5.8. Figure 2 shows the frequency vs. temperature characteristics of a SAW device.

In many applications the uncompensated SAW oscillator frequency drift is too much for the system to tolerate. In this case, a component oven is often utilized to raise the SAW device to a higher temperature and thus, limit the tempeature variation that the SAW would experience. However, component ovens are costly, relatively large and require several watts of DC power. An alternative to ovenizing the SAW device is to use a voltage controlled SAW oscillator and merely apply a tuning voltage to cancel the temperature induced frequency drift.

In some instances it is desirable to generate the tuning voltage internally in the oscillator by using temperature sensitive devices such as thermistors. Another approach, suggested by Kinsman[2], is to use the temperature sensitive base emitter junction of a bipolar transistor to generate a linear voltage versus temperature curve is shown in Figure 10, to generate a positive parabolic frequency shift that cancels the negative parabolic frequency dependence of the SAW device. This approach can be modified to compensate only one portion of the temperature dependent frequency curve as detailed by these authors in references 3. In this approach a single varactor is driven by the linear voltage from the temperature sensor (Figure 11) and the frequency drift is compensated on one side of the parabola. This approach has resulted in a frequency accuracy of ±4 PPM on a 950 MHz hybrid oscillator.

SUB-SYSTEMS

In response to requests from system users Sawtek has begun development of sub-systems in which the SAW device is the key component. We have chosen two examples for illustration: 1) A SAW resonator controlled PCM transmitter for a radiosonde, and 2) SAW sensor modules.

The radiosonde is used to provide meteorological data which is used in calculations for directing artillery fire. The radiosonde is launched on a helium ballon and is powered by water activated battery. It provides an RF pulse encoded output of atmospheric temperature and pressure and it is radar tracked to provide wind velocity information. This system substantially increases the probability of hitting a target during the first round of fire. Similar systems have been in use in the past but utilized L.C. controlled single stage transmitters. The SAW controlled unit is able to improve the temperature versus frequency stability of this transmitter which will allow it to be used in the European theater.

The block diagram of this system is shown in Figure 12. It consists of a hermetically sealed 560 MHz hybrid SAW resonator controlled oscillator which provides a +10 dBm output. The signal is then frequency tripled with a Class C amplifier. The 1680 MHz output from the tripler is then filtered by a microstrip line bandpass filter to remove sub-harmonics and is finally fed to a two stage power amplifier which provides a +24 dBm signal to the antenna. A modulator stage inverts the trigger input and pulses the power amplifier stage.

In addition to providing 1/4 watt output, this system suppresses harmonics to -40 dBc and provides an AM modulation ratio of 30 dB. The frequency versus temperature stability and set-on accuracy is better than ±200 KHz over a temperature range of -70°C to +50°C.

The SAW sensor is an emerging technology which is being driven by a need to develop increasingly more accurate accelerometers, pressure sensors, and gas detection systems. The accelerometers are used in inertial guidance systems of missiles. Pressure sensors are being developed for jet engine instrumentation where an increase in accuracy means improved efficiency. Gas sensors are used in the detection of hazardous gases.

Common to most of these sensors is the concept of the dual oscillator system which is very effective in eliminating temperature instability and can be used to improve sensitivity. Referring to Figure 13, we can see how this works. In this figure illustrating a SAW accelerometer two SAW delay line controlled oscillators are located on opposite sides of two quartz cantaleivered beams. When the system experiences an acceleration perpendicular to the major plane of these beams one of the delay lines undergoes a compressive force while the other is de-compressed. The strains induced by these forces cause a variation in the SAW acoustic velocity, which in turn, will modify the operating frequency of these oscillators.

The outputs of these oscillators are then fed into a mixer. The mixer output is passed through a low pass filter (LPF) to eliminate second order mixer products and the L.O. signal. The resultant difference frequency is directly counted to provide a digital indication of the acceleration. By this method two advantages are obtained: 1) the acceleration sensitivity is doubled by virtue of the fact that the strains are of opposite sense in each delay line, and 2) even more importantly, the frequency vs. temperature drift is eliminated to the first order. The temperature drift of both oscillators is in the same direction since both delay line units are made of the same substrate material. Therefore, the difference frequency observed at the output of the mixer is independent of temperature drift.

Figure 14 is a schematic of a SAW pressure sensor. This sensor again uses the dual oscillator concept to eliminate temperature drift. In this case, however, the SAW controlling element is a resonator. The SAW resonator (SAWR) is an alternative to the delay line where increased sensitivity is required and where oscillator simplicity is desired. As we can see in this case only one of the SAWs undergoes a pressure modulation since it is located over a thinned portion of the crystal. The other serves as a reference for temperature stability. In this case it would not be practical to fabricate a SAW on the opposite side of the membrane because of photo fabrication difficulties and because it would expose the unit to a possibly corrosive test gas. This reduction in sensitivity is compensated by the use of the SAWR.

The final example is shown in Figure 15. This is a SAW gas sensor. It consists of the now familiar dual oscillator, mixer and low pass filter. Unlike the mechanical sensors, however, the velocity shift in one of the paths is obtained by a mass loading effect associated with absorbtion or chemabsorbtion of a gas as it passes over a thin film of deposited material which is chosen to select the desired gas. Paladium for example has been used in hydrogen sensors. Sawtek has developed delay lines and resonators for these applications as well as the electronics modules for driving them.

The main thrust today in the development of this sensor is to obtain a reliable, discriminating thin films that will permit detection of the desired gas only. A regeneration feature is also being sought after which will permit multiple use of the sensor.

HYBRID SAW OSCILLATORS

More and more SAW oscillator designs are requiring hybridization to meet size and cost requirements. Hybridized SAW oscillators have the advantage of better reliability, lower EMI and greater immunity to environmental effects since most units are hermetically sealed in metal enclosures. Because of the planar construction of the inductors used in the hybrids they are potentially more immune to vibration.

Two of the applications mentioned earlier, the 560 MHz RF source for the radiosonde sub-system and the entire 950 MHz TCXO have been hybridized. Single stage SAW resonator controlled oscillators are being considered for many applications that fall under FCC part 15 Rules. In the medical area wireless patient monitoring systems are being considered. In aviation they are being considered for search and rescue appliations as emergency locator beacons aboard downed aircraft. They improve the stability and allow higher frequency operation of cordless telephones. They are effective local oscillators in RF modems in local area networks. They have been shown to improve the harmonic suppression of wireless security systems to within FCC guidelines and to improve the reliability of these systems.

Figure 16 provides a summary of the typical operating characteristics of such a single stage SAW resonator controlled hybrid oscillator. The frequency range has been extended to over 1100 MHz. Spurious rejection is usually beyond the dynamic range of the measurement system. The supply voltage can range up to 24 volts and efficiencies range from 4 - 6% for a single stage system which provided 5-10 dBm output power. The package size shown in that figure is for a 4 Pin-DIP which would be the least expensive version. Surface mountable flatpacks are also available which are smaller in size and compatible with wider operating temperature ranges.

The cost of the SAW hybrid oscillator drops below ten dollars in large quantity. The main cost drivers of this unit are frequency set tolerance and possibly the package if a surface mountable unit is required.

CONCLUSION

SAW oscillator design has matured to the point of becoming the choice of systems designers who used to rely on bulk wave crystal or LC technology. Custom designed oscillators for commercial applications no longer require extensive development effort and therefore, their cost is dropping. In addition, performance is superior to the LC oscillator. For precise frequency accuracy and stability, ovenized SAW oscillators provide an attractive alternative to bulk crystal based oscillator/multiplier systems. The ability to upgrade the performance of a system has been enhanced by SAW technology and should account for an increased demand for SAW oscillators in the future.

REFERENCES

1. T. O'Shea, V. Sullivan, R. Kindell, "Precision L-Band SAW Oscillator for Satellite Application," Proceedings 37th Annual Symposium on Frequency Control, U.S. Army ERADCOM, pp 394-404.

2. R. C. Kinsman, "Temperature Compensation of Crystals with Parabolic Temperature Coefficients," Proceedings 32nd Annual Symposium on Frequency Control, U.S. Army ERADCOM, pp 102-107.

3. J. Ladd, C. Abdallah and T. O'Shea, "Hybrid SAW Oscillator and Resonator Filter," Proceedings of the 1984 IEEE Ultrasonics Symposium.

Figure 1

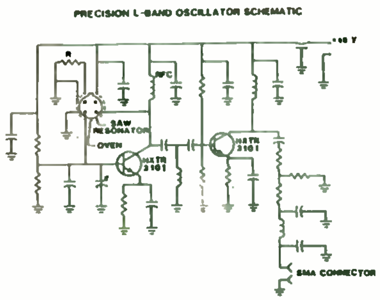
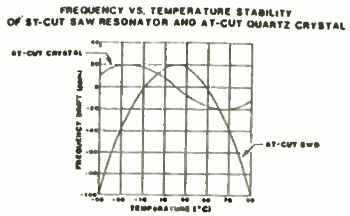


Figure 2



$$\frac{\Delta F}{F} = \left(\frac{T - T_0}{5.8}\right)^2 \text{ ppm}$$

OPERATING TEMPERATURE RANGE °C	0 to 60	-10 to 70	-30 to 85	-55 to 85
TOTAL $\frac{\Delta F}{F}$ ppm REF. FREQ. at T_0	-28	-48	-100	-146
$\frac{\Delta F}{F}$ ppm REF. FREQ. at 20°C	±14	±24	±50	±73

Figure 3

PRECISION L-BAND OSCILLATOR

SPECIFICATIONS	GOALS	ACTUAL
FREQUENCY	740.000 MHz ±3.4 ppm	-0.59 ppm
R.F. POWER (50Ω)	+10 dBm MINIMUM	+11.60 dBm at 25°C
SHORT TERM STABILITY	AVERAGE TIMES (100 samples)	
	1 sec 1 X 10^{-10}	1.473 X 10^{-10}
	10 sec 1 X 10^{-10}	1.882 X 10^{-10}
	100 sec 1 X 10^{-10}	2.896 X 10^{-10}
LONG TERM STABILITY	±5 ppm for 1st year	5.03 ppm/year Customer verified at 4.5 ppm/year
HARMONICS	<-40 dBc	2nd <-60 dBc 3rd <-55 dBc
SPURIOUS	<-60 dBc	<-60 dBc (100 Hz to 1500 MHz)
FREQUENCY VARIATION WITH ±1% VOLTAGE CHANGE	<1 ppm	<0.2 ppm
MAXIMUM POWER CONSUMPTION	3.5 W	3.47 W

Figure 4

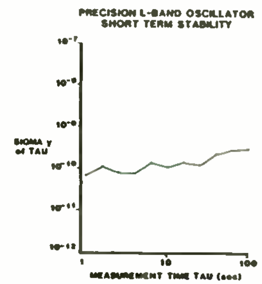


Figure 5

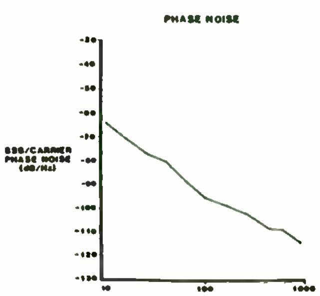
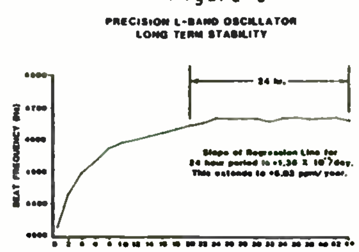


Figure 6



10 OF 12

Figure 7

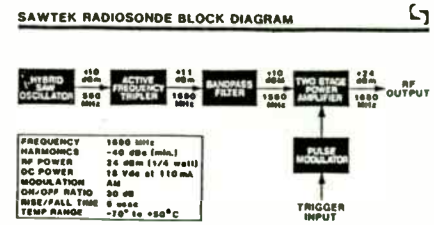


Figure 8

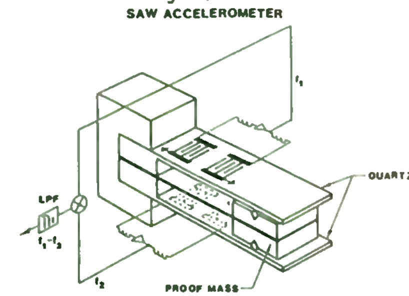


Figure 9

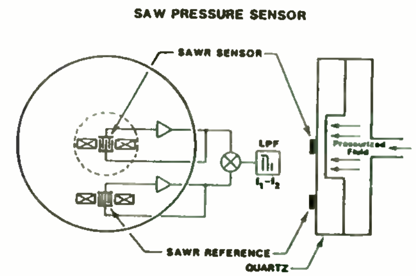


Figure 10

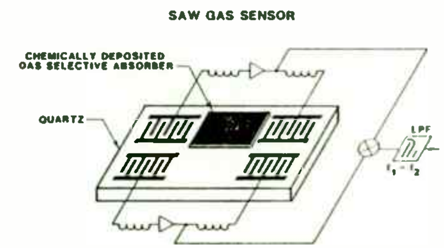
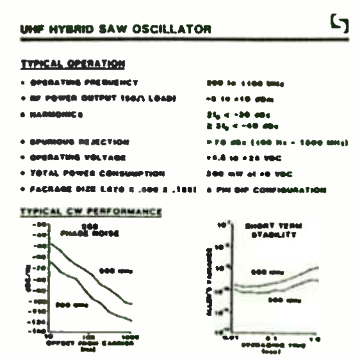


Figure 11



11 OF 12

Figure 12

SAW 1-PORT RESONATOR VOLTAGE CONTROLLED OSCILLATOR
PIERCE CONFIGURATION

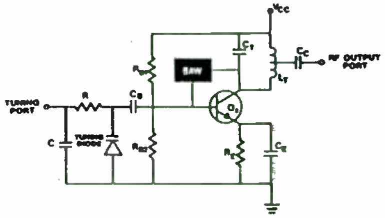


Figure 13

SAWTEK'S DELAY LINE OSCILLATOR
CIRCUIT SCHEMATIC DIAGRAM

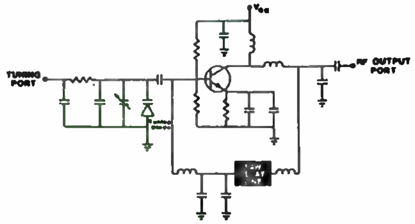


Figure 14

SAWTEK'S
600 MHz DELAY LINE VCO

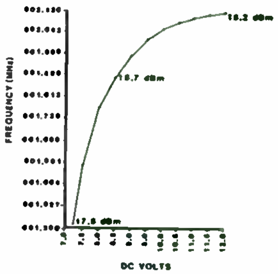


Figure 15

TEMPERATURE SENSOR

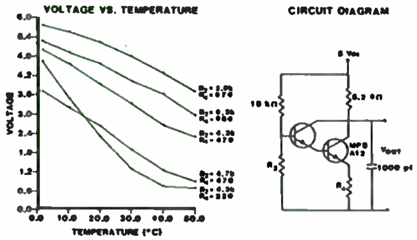
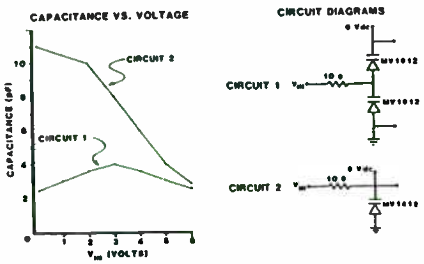


Figure 16

VARACTOR DIODE CIRCUITS



Peter Chadwick

DESIGN COMPROMISES IN SINGLE LOOP FREQUENCY SYNTHESISERS

1. INTRODUCTION

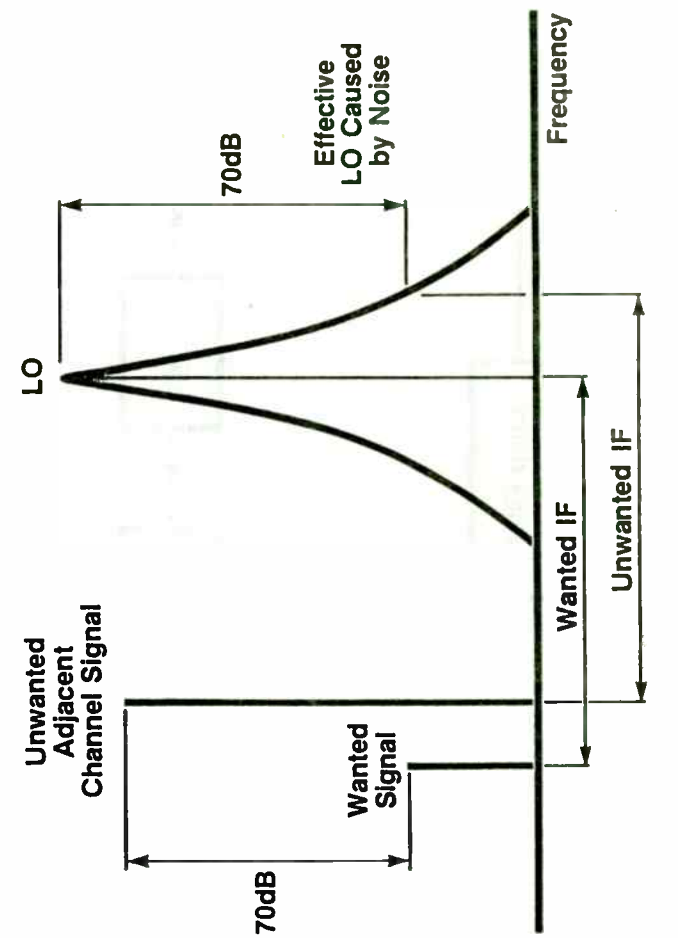
The single loop frequency synthesiser is justly popular as an approach to frequency synthesis. It has the merit of simplicity, and because of this, low cost, especially as a large amount of the circuitry is easily produced in monolithic integrated circuit form.

Certain performance parameters of the synthesiser are defined by the equipment performance. For example, a marine VHF radio frequency synthesiser has requirements for phase noise and discrete spurious outputs defined by the adjacent channel specification, and the phase noise performance may well need to be several dB better than would at first be expected. If the adjacent channel rejection is 70dB for example, then a single sideband phase noise level in the receiver bandwidth must be more than 70dB - see Fig. 1. In fact, the translated noise level should be reduced by an amount dependant upon the performance of other areas of the equipment and these specification levels are typically determined by the system architect. Frequently, however, during design of a project, some modifications in architecture become apparent, but an understanding of practical limitations is vital at an early stage if delay and consequent expense is to be avoided. For further details on the effects of phase noise on receiver performance, see Ref. 1.

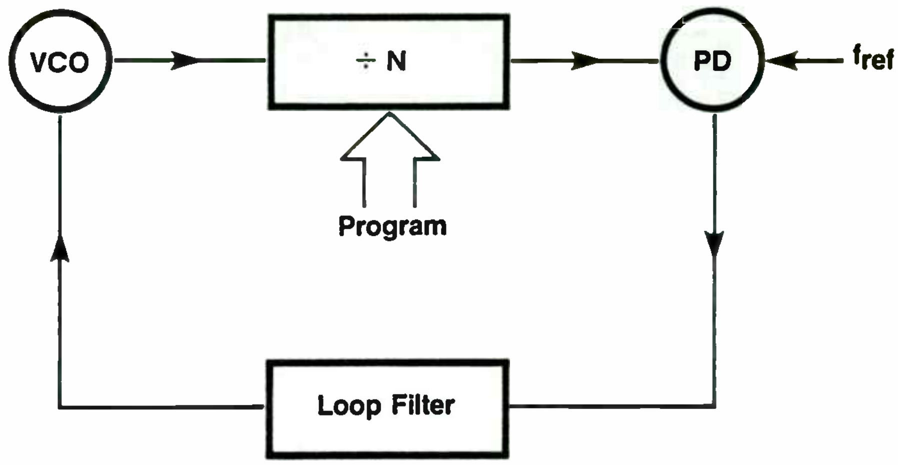
2. DIVIDERS

Single loop synthesisers using direct division as in Fig. 2 suffer from certain limitations. Fully programmable dividers are not generally available for frequencies above about 50MHz without high power consumptions, and even CMOS dividers currently available are limited in applications at low (5volt) supply voltages and extreme temperatures. Newer devices are appearing, however, and experimental 250MHz operation has been observed.

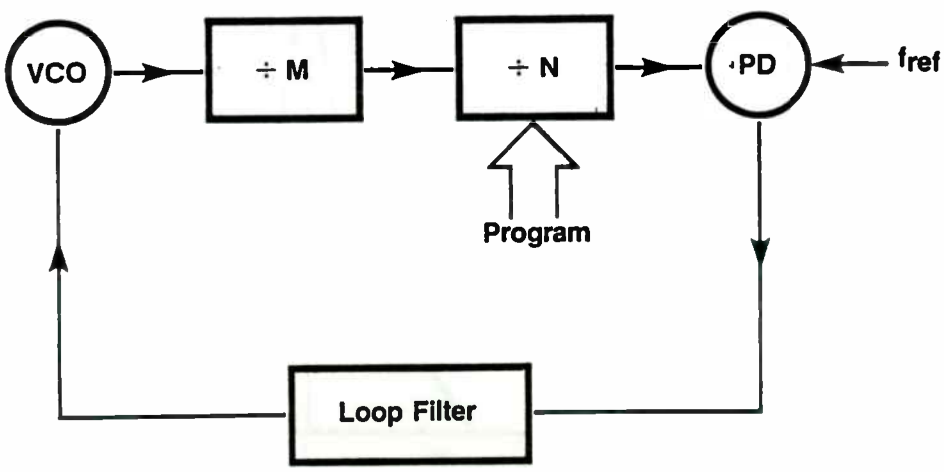
Early synthesisers used fixed pre-scalers to divide the VCO down to a suitable frequency for the programmable counter as in Fig. 3, or used mixing techniques as in Fig. 4. Indeed, a large number of CB radios use the mixing technique, but this system can suffer from spurious products unless carefully designed in choice of frequencies, input levels and particular mixers used. - Ref 2,3,4,5. In addition, the large variation in subsequent division ratio may give problems with loop dynamic performance.



Phase noise and adjacent channel rejection fig 1



Simple PLL fig 2



Use of a fixed prescaler fig 3

A major area of conflict lies in the choice of reference frequency. In synthesisers such as Fig. 3, the output frequency step size is M times the reference frequency, where M is the prescale ratio. In a system where every channel is used, the problem is then that the reference frequency has to be decreased by a factor of M, and as a result, the bandwidth of the feedback loop must decrease. The bandwidth and damping factor of the loop filter are vitally important parameters in determining such loop characteristics as lock up time as well as the phase noise characteristics. (The effects of loop bandwidth on phase noise will be discussed later). In general, the widest possible loop bandwidth is required to minimise lock up time and to confer the greatest immunity to shock and vibration. However, the loop bandwidth cannot be greater than the reference frequency and so the use of a fixed prescaler is obviously somewhat limited. The alternative is the widely used "Two Modulus" or "Pulse Swallowing" prescaler system, illustrated in Fig. 5. In this method, the prescaler is able to divide by two integers N and N+1. The two counters A and M are programmable and are clocked in parallel, the divider being set initially to the N+1 ratio. When the A counter is full, the divider is set to divide by N until the M counter is full, giving a total division ratio of MN+A. This system is limited to a minimum division ratio of N^2-N if every value of N is to be achieved (no "skipped" channels) and the M counter must always be programmed to a bigger number than the A counter. Within these limitations, however, a fully programmable divider is achieved and so f ref can now equal the channel spacing.

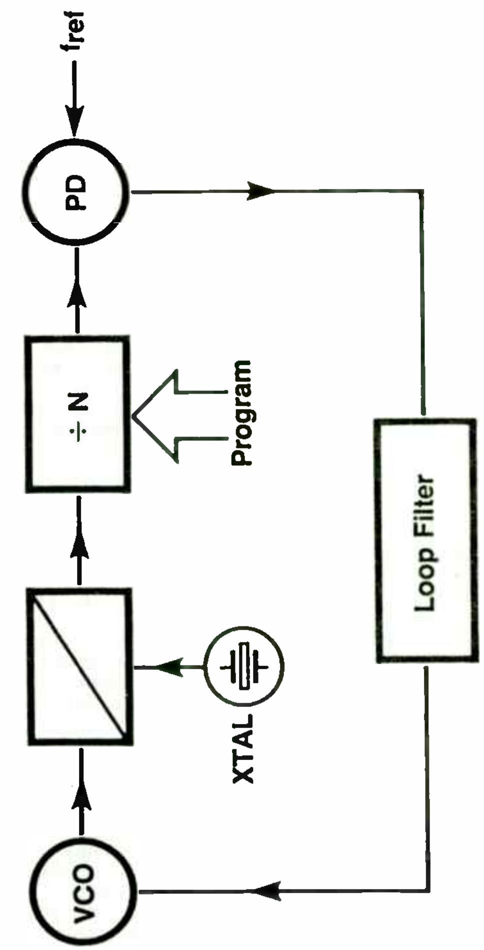
Another and more subtle limitation is in the delay times of the various components within the loop. When the circuit (Fig. 5) has counted down so that the M counter has been filled, the whole system is reset, and quite obviously, must achieve this in a time equal to N+1 cycles of the input frequency e.g. in a ÷64/65 prescaler, at 1GHz, the reset of the M and A counters must be achieved in 65 cycles or in this case, 65nS. This means that the propagation delays plus set up/release times plus reset delays must not exceed 65nS and it is this area where trouble can often be expected, especially at temperature extremes. Although a 1GHz synthesiser with a 64/65 divider only sees an input frequency of 15MHz for 1GHz input, the set up/release time and delays may well easily reach 85-90nS and the system will thus fail.

If the propagation through the divider = td
the set up time = ts
the release time = tr
the propagation delay through the A and M counters = tc

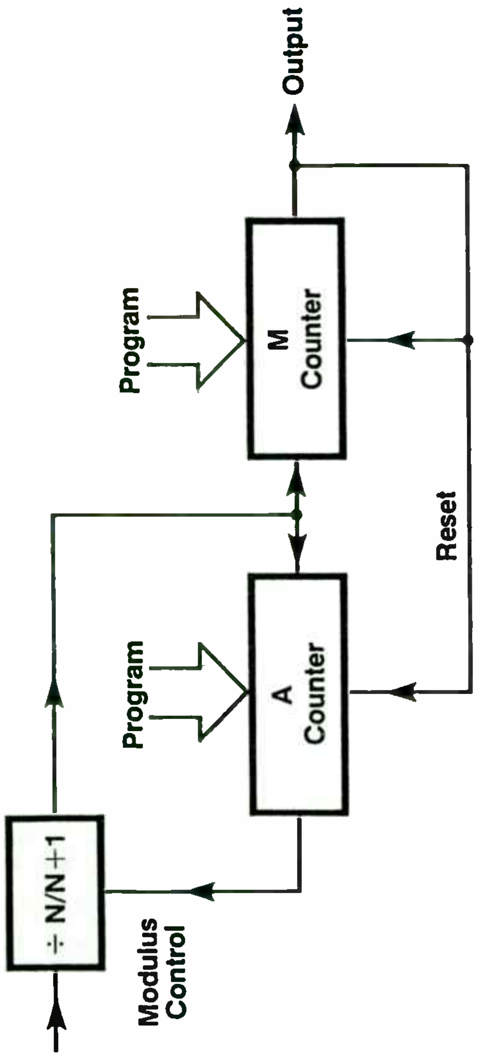
then

$$f_{max} = \frac{N}{(t_d + t_s + t_c)} \quad \text{or} \quad \frac{N}{(t_d + t_r + t_c)}$$

whichever is least.



Mixing in the loop fig 4



Two modulus divider fig 5

One of the areas in which an increase in LOOP DELAY TIME can inadvertently occur is if the A and M counters trigger from a different edge to the dual modulus prescaler. This can cause a major diminution in available loop delay, as can an attempt to physically separate the divider and control circuits. Other deleterious affects have been noted, such as radiation of the divider output to the VCO, producing high frequency sidebands, so practical synthesisers are best produced with little physical spacing between divider and control circuit.

The control circuit is a practical device in a number of technologies, although modern devices exclusively use CMOS to minimise power consumption. Prescalers are still mainly exemplified by bipolar technology, advances in which have seen major reductions in power consumptions in recent years - for example from 65mA at 5v for a divide by 10/11 operating at 250MHz in 1976 to 4mA at 5v for a 40/41 operating at 225MHz today. Some equipments still build up the A and M counters from discrete IC's and then add phase detectors, reset circuitry and so on, but such equipments are by now obsolete in design and extremely expensive to manufacture. Nevertheless, the lessons of tolerancing delays necessary in such designs should not be forgotten just because the majority of circuity is now hidden inside a block of silicon.

The choice of prescaler ratio is governed by a number of factors. Discussed so far have been minimum ratio and loop delay. However, the output frequency of the divider must be low enough for the A and M counters to function. Summarising

1. Fin $\leqslant$ N Fmax control
   where N is the divider ratio
   Fmax control is control circuit maximum operating frequency.

2. Fin $\leqslant \dfrac{N}{\text{total loop delay}}$

3. Pmin = N^2-N
   where Pmin is the minimum divide ratio.
   N is the dual modulus divider ratio.

Various values for N exist in proprietry devices. These range from 3/4 to 128/129: binary values (32/33, 64/65, 128/129) are popular for ease of pogramming from ROM's and microprocessors, while decimal and BCD are used for thumb wheel switch programming.

Programming is a straightforward exercise for binary division and the following method is recommended.

1. The A counter should contain x bits such that
$2^x = N$

2. If more bits are included in the A counter, these should be programmed to zero.

e.g. N = 64 = 6 bits
A = 10 bits
then the 4 MSB are programmed to zero.

3. The M and A counters are treated as being combined so that the MSB of the M counter is the MSB of the total and LSB of the A counter is the LSB of the total.

e.g. A synthesiser operating from 430-440MHz in 25KHZ steps uses a 64/65 divider, and the control circuit uses binary counters.
$P = {}^{F}/Fref$ and Fref = channel spacing = 25KHz
$Pmin = {}^{430}/.025 = 17200$
$Pmax = {}^{440}/.025 = 17600$

Minimum possible divide ratio is $N^2 - N = 4032$
where N is two modulus divider ratio
maximum allowable loop delay $= \frac{64}{440 \times 106} = 145nS$

Total divide ratio, P, is given by
P = NM + A
N = 64, as a 64/65 divider is used
Pmin from above is 17200
Therefore 17200 = 64M + A
And $M \geq A$
Let A = O Then $Mmin = \frac{17200}{64} = 268.75$
= 268
and $Mmax = \frac{17600}{64} = 275.0$

Thus the M counter must be programmable from 268 to 275 as required: the M counter must have at least 9 bits.

For a frequency of 433.975MHz

P = 433.97/.025 = 17359
therefore $M = \frac{17359}{64} = 271.2343$

The A counter is programmed for the remainder i.e.
0.2343 x 64 = 15

From this, the A counter is programmed to 15 and the N counter to 271. The output frequency can now be checked.

P = NM + A
= 271 x 64 + 15 = 17359
and this is the required divider ratio.

The two modulus prescaler is therefore able to offer the advantages of producing a programmable divider operating at a very high frequency, but consuming a fraction of the power of such a divider. This enables the reference frequency to equal the channel spacing, thus allowing maximisation of loop bandwidth with its concomitant faster lock up time. It is limited by total loop delay, maximum operating frequencies of dividers and counters, and in minimum count values, but is nevertheless a powerful tool for the synthesiser designer.

The limitation on the value of Pmin, the minimum divide ratio can be avoided by the use of three and four modulus dividers. The use of a four modulus counter allows a very wide frequency range to be covered with one device, but at the expense of a much higher power dissipation. Typical of such devices are the Plessey SP8901 and SP8906. Power consumptions typically range for 2 modulus dividers from 4mA at 200MHz (Plessey SP8792 and 3) through 11mA at 520MHz (Plessey SP8716/8/9) to 25mA at 1GHz (Plessey SP8703EXP) for two modulus dividers.

3. LOOP BANDWIDTH AND PHASE NOISE

As stated earlier, phase noise is a very important parameter in frequency synthesisers. Too many early synthesisers suffered from phase noise problems which manifested themselves as poor equipment performance in such areas as multiple signal selectivity and ultimate signal to noise ratio. The performance of the synthesiser may be degraded or improved by changing the loop bandwidth, depending upon the characteristics and parameters involved.

The general characteristics of a phase locked loop (PLL) are that for signals injected into the loop it acts as a low pass filter for signals inside the loop bandwidth, and as a high pass filter for signals outside the loop bandwidth. To analyse the performance, consider modulation of the VCO at very low frequencies. The output of the phase detector will be a low frequency signal of phase such as to attempt to remove the modulation imposed on the VCO. As the modulation frequency increases, the error component of the phase detector output is not passed by the loop filter, and so the modulation is not removed by the loop. Note that the modulation is phase modulation (PM) up to the filter break point, and frequency modulation (FM) thereafter. In the "in-between" range, some interesting distortion effects can occur, especially when excessive group delay exists in the loop filter.

The relationship of loop filter bandwidth to phase noise is now apparent. Phase noise from the oscillator corresponding to frequencies below the filter bandwidth will be removed by the loop, while phase noise components outside the loop bandwidth will be unaffected by the loop. Under these circumstances then, the VCO output spectrum will be cleaned up by the loop. However, for frequencies inside the loop bandwidth, other factors enter. Variations in the reference frequency cause variations in output frequency from the synthesiser, and phase noise components at the reference frequency are purely the frequency domain transforms of time domain frequency instability (ref. 6,7,8). These phase noise affects are multiplied in the loop by the divider ratio. An example (admittedly using gross instability for demonstration) is shown.

If the 430MHz synthesiser in section 1 has an instability of +1Hz in the 25KHz reference frequency, this is multiplied by P.

i.e. for operation at 433MHz
$P = 433/.025 = 17320$
Therefore if +1Hz at 25KHz gives +17.32KHz at final frequency.

Phase noise at the reference frequency is derived from two sources:-
a) the system standard oscillator
b) the reference chain divider

Oscillators for standards are available with very low phase noise characteristics, and -130 to -170dBc/Hz at 1KHz offset covers the usual range. This phase noise is modified by the reference divider and multiplied by the division ratio as explained above. Of course, phase noise at any offset is reduced by division until the phase noise floor of the divider is reached. Little has been published on the causes of phase noise in dividers, although various measurements have been made. (Ref. 9). It has been suggested that TTL and CMOS dividers are better than ECL and CMOS is better at low (10-20Hz) offsets. At a 1KHz offset, ECL levels of about -145 dBc/Hzand CMOS levels of -155 to -165 dBc/Hz appear usual. The explanations for the occurance of phase noise is intuitively regarded as being jitter in the transition point of the signal: on this basis, one would not expect CMOS to be so good as TTL insofar as the rise and fall times will be somewhat slower. Regrettably, the difficulty and cost of making meaningful measurements is an inhibiting factor: data on the phase noise performance of Gallium Arsenide dividers would be of considerable interest, especially at small frequency offsets.

From the above discussion, a phase noise floor of some -150dBc/Hz can be expected at the end of the reference frequency divider chain if a good frequency standard is used, while a low cost one may well be at about -130 dBc/Hz. In our 430MHz synthesiser, a degradation at 1KHz (if the loop is wide enough) of some 84dB will be seen, so inside the loop bandwidth, the noise performance will be limited to -130 + 84 = -46 dBc/Hz. At lower offset frequencies, the phase noise of dividers and frequency standards is worse, so the phase noise performance is now being defined by the loop, rather than the VCO. These are worst case figures, but the ultimate signal to noise ratio of an FM receiver can clearly be seen to be easily limited at UHF by multiplied phase noise. Fortunately, the noise enhancement by the loop is such that pre-emphasis of the modulation provides major improvements in signal to noise ratio.

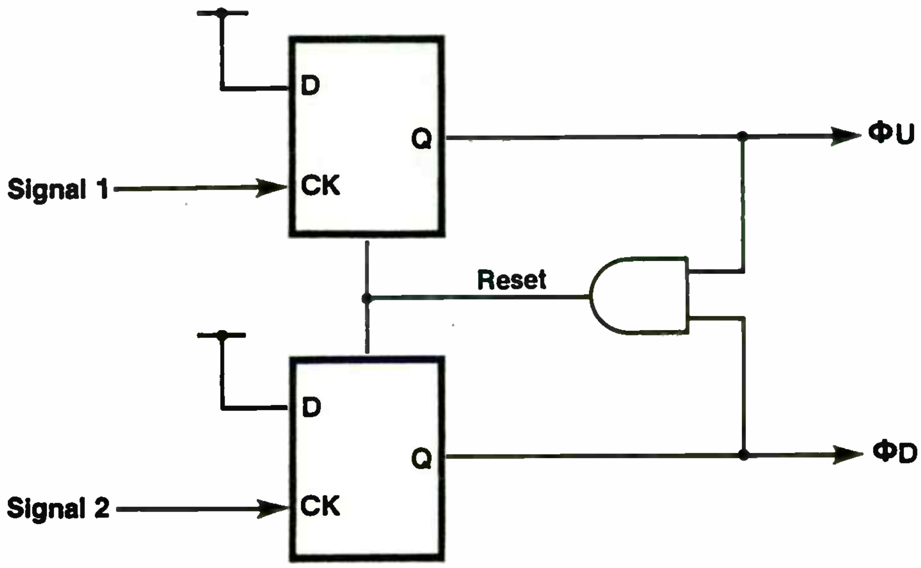
Nevertheless, it is obvious that the choice of loop bandwidth is compromised by the ultimate signal to noise level required by the system and that such factors as reference oscillator noise level and divider noise cannot be totally disregarded. Operation in the usual cellular radio bands at 800 or 900MHz makes the situation some 6 dB worse than that analysed above and the use of a psophometric audio weighting in the equipment is advisable. Sub audible tones may well need fairly high deviation if signal to noise performance is not to be severely limited on them, although modern decoders will work with a negative signal to noise ratio - Ref. 10.

In the single loop synthesiser, the phase noise in adjacent channels, which determines the adjacent channel performance, is, to a first order, unaffected by the loop and its parameters. Second order effects such as noise modulation by such loop components as high value resistors and operational amplifiers may be negated by the use of a passive low pass filter prior to the VCO. Phase noise in the oscillator will be discussed in section 4.

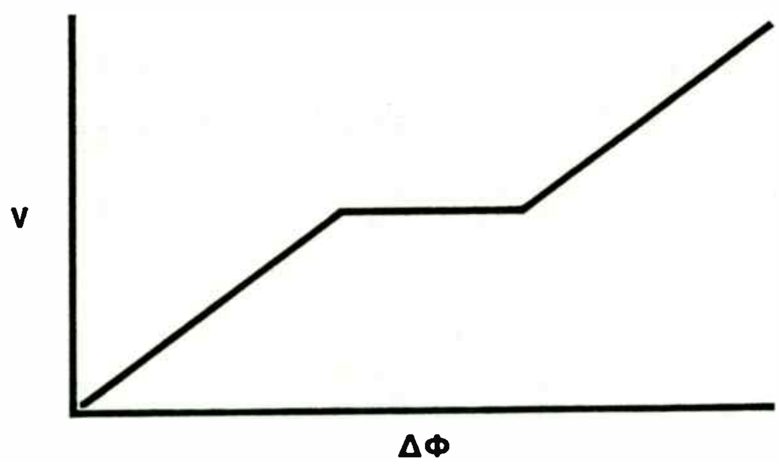
Even where the effects of multiplied phase noise may be ignored, such as where the reference divider chain noise is sufficiently low, certain other problems occur in the loop filter design. Many of these are associated with the phase detector employed, which in many areas has been a digital phase/frequency detector. Various types of detector have been used over the years, from an OR gate producing a variable mark space ratio to the well known 2 D type detector. The first of these used integration of the variable mark-space ratio to produce the required output, while the latter (Fig. 6) produces minimal width pulses on both Øu and ØD when in the zero phase error condition. Unfortunately, the zero phase error state exists for a degree of phase error dependant upon the propagation delays of the gates and a phase error/output voltage characteristic such as Fig. 7 is achieved. The performance in the central flat portion of the characteristic means that the loop gain falls to zero when the phase error reaches some small but finite value, and this leads to an increase in the low frequency phase noise of the loop. This phenomena is of course related to the reference frequency of the loop, being worse at high comparison frequencies.

Although a number of approaches have been made to minimise this problem, including the provision of a leakage path across the VCO control line (Ref. 16), the better approach is to use a linear phase detector of high gain to "fill in" the gap in the response. An additional benefit of this method is that if the digital phase detector has a "tri-state" output for the area in which the dead zone occurs and the linear phase detector operates, then the phase detector output at comparison frequency is reduced, allowing either a wider loop bandwidth for the same comparison frequency sideband rejection, or increased rejection, or to some extent, both. The analogue phase detector may easily be given a very high gain and narrow range of operation - say a 2 degree range with a gain of 600 volts/radian, but only a limited lock range. It is however, essential to ensure that saturation of this detector, and indeed of the loop filter/amplifier is minimised, as under channel change conditions, the control line and thus the filter amplifiers can be driven hard into saturation. A long recovery time here may well make a mockery of any lock up time calculations. It is this approach which has been adopted in the NJ8820 series of CMOS control circuits from Plessey with a large degree of success.

The choice of loop bandwidth is also governed by the time to change channel, and here again, compromise is often necessary. For example, a lock up time of 1mS and a loop bandwidth of 100Hz are apparently mutually incompatible. By using the two detector approach outlined above however, the loop bandwidth for the digital detector may be made much wider than the analogue detector, thus providing a form of adaptive filtering. The basic loop equation for a type 2 2nd order loop is



Dual D type phase discriminator fig 6



Transfer characteristic of phase discriminator with a charge pump fig 7

$$Wn = \sqrt{\frac{KoKv}{N\,t_1}}$$

where Wn = loop natural frequency, Kv = VCO gain in Rad/S-v, Ko = phase detector gain in volts/rad, N = division ratio and t_1 = integrator time constant, shows the dependance of Wn, the loop natural frequency on N. It should be noted the 3dB bandwidth of the loop and the natural frequency Wn, are not identical - except for a damping factor, D = 3.02.

It was stated earlier that noise caused by the phase detector and loop filter is easily filtered to avoid noise in adjacent channels. Noise inside the loop bandwidth is another matter, and the use of low noise components in loop filters (NOT a 741!) is advisable. Where possible, time constants should use large capacitors and small resistors to minimise KTBR noise. $1/f$ noise can be a problem with operational amplifiers, and where loop bandwidth is high, slew rate is important if the dynamic loop bandwidth is to bear any relationship to the small signal case.

To summarise, the choice of loop bandwidth affects close in phase noise and lock up time. Phase noise is produced by dividers, phase detectors and filters, and when multiplication ratios are high, the reference frequency phase noise can be dominant when multiplied. To minimise this effect, the loop bandwidth can be narrowed, since noise outside the loop bandwidth is determined solely by the VCO. Typical divider phase noises of -150 or -160 dBc/Hz can be expected, so low cost reference oscillators can dominate the noise performance.

4. VOLTAGE CONTROLLED OSCILLATORS

Many engineers consider VCO design to be a black art, and although some art is occasionaly involved, VCO's are amenable to analysis.

In the single loop synthesiser, the phase noise performance outside the loop bandwidth is dominated by the VCO, with the noise generation by passive components in the loop filter generally being of lesser importance.

Scherer, Leeson (ref. 12) and Robins (ref. 13) have analysed oscillator phase noise performance and Scherer (ref. 14) has demonstrated the applicability of Leeson's equations and uses the equation

$$L_{(f)} = 1/8 \left(\frac{FKT}{Ps}\right)\left(\frac{fo}{f}\right)^2 \left(\frac{1}{Q} + \frac{Po}{\frac{1}{2}CV^2 2\pi f.}\right)^2 \quad \text{eq 1.}$$

where $L_{(f)}$ is the SSB phase noise at an offset F
- F is the Noise Figure of the amplifier in the oscillator
- K is Boltzmann's Constant
- T is the Temperature
- Ps is the available signal power
- fo is operating frequency
- f is the offset at which the power is to be calculated
- Q is working Q of the tuned circuit
- C is tank capacity
- V is tank current peak voltage
- Po is rf output power

By inspection of eq 1, it may be seen that the phase noise is proportional to Q^{-2} and also to (frequency offset)-2. This means that for each octave decrease in the offset frequency, the noise power will increase by 4 times or at 6dB/octave. As the frequency offset decreases $1/f$ or flicker noise becomes important: this "break" frequency can be as high as 50MHz with GaAs devices. From eq.1, it may be determined that a low phase noise oscillator will have a large voltage swing, a high working Q and provide little output power to the load. There is of course a limit as to the level of power required, as the noise of any subsequent buffer amplifiers will degrade the oscillator.

A major compromise in the design of equipment is the choice of VCO frequency. If, for example, a 800MHz cellular radio type of receiver is considered, some fairly straightforward calculations will serve to act as a guide. Starting with the receiver parameters, we will assume that a 70dB rejection of a signal two channels (60KHz) away is required. A number of receiver sub system parameters are involved.

(a) Synthesiser phase noise
(b) IF filter performance
(c) Co channel rejection ratio
(d) Gain compression of stages before the main IF selectivity.

Of these parameters, (c) is the least obvious in its applicability. Ref. 1 showed how oscillator noise was mixed onto a wanted signal by a strong unwanted signal. The degradation of a wanted signal by this noise obviously depends upon the relative levels of signal and noise, and because the noise is on the same frequency, the Co-channel rejection. Typically, this means that a noise level within the IF passband of some 8dB less than the signal is required. Thus for the 70dB rejection, oscillator noise at -78dB is required, and 80dB would thus be the design aim.

Conversion of this level to dBc/Hz is not straightforward because of the non linear slope of the phase noise. However, for narrow bandwidths at large offsets, little error is obtained by approximating the phase noise slope to a straight line. This may be illustrated as follows:-

From equation 1, the power spectrum at an offset beyond the flicker noise knee is given by:-

$$P_o = kf^{-2} \qquad -(2)$$

where P is the noise power
K is a constant
f is the offset

For a frequency band bounded by f_{lower} and f_{upper}, the noise power is:-

$$P_t = \int_{f_L}^{f_u} kf^{-2}\,df = \Big[-kf^{-1}\Big]_{f_L}^{f_u}$$

$$= K\,(f_L^{-1} - f_u^{-1})$$

Therefore

$$K = \frac{P_t}{(f_L^{-1} - f_u^{-1})}$$

P_t has been defined as the phase noise in the band = -80dB

therefore

$$K = \frac{10^{-8}}{\left[\frac{1}{53.5 \times 10^3} - \frac{1}{67.5 \times 10^3}\right]} = 2.58 \times 10^{-3}$$

To find the phase noise in a 1Hz bandwidth at an offset F

$$P = kF^{-2}$$

so at 53.5KHz

$$P = \frac{2.58 \times 10^{-3}}{(53.5 \times 10^3)^2} = 0.901 \times 10^{-15}$$

$$= -120.5 \text{ dBc/Hz}$$

At 60KHz

$$P = -121.4 \text{ dBc/Hz}$$

and at 67.5KHz

$$P = -122.5 \text{ dBc/Hz}$$

If the 'break point' for $1/F$ noise is above 60KHz, then the spectral density is determined by noise rising at F^{-3}. Similar procedures are followed:-

$$P_o = K'F^{-3}$$

$$P_t = \int_{f_L}^{f_u} K'f^{-3}\,df = -K'\left[\frac{f^{-2}}{2}\right]_{f_L}^{f_u}$$

$$= \frac{-K'}{2}(f_u^{-2} - f_L^{-2})$$

$$= \frac{K'}{2}(f_L^{-2} - f_u^{-2})$$

Using similar figures, the performance required is:-

53.5KHz	-120 dBc/Hz
60 KHz	-121.5 dBc/Hz
67.5KHz	-123 dBc/Hz

The error by assuming a linear relationship is given by:-

IF bandwidth = 15KHz
therefore noise power is 10 $\log_{10}$ 15 x 10^3 dB greater than in a 1Hz bandwidth
which is 41.8dB
therefore if the noise power is 80dB down on the signal,
total carrier to noise power ratio is -121.8 dBc/Hz at 60KHz.

This in fact gives a requirement some 0.4 dB higher than previously calculated and in 120dB is obviously negligible.

Having decided upon the level of allowable oscillator noise, it is now possible to calculate the best methods of achieving this level. Using Scherer's figures from Ref. 13 for a 400MHz oscillator which will be doubled, using parameters of:-

Q	= 200	
C	= 23pF	
V	= 10v pk	
$\frac{FKT}{P_s}$	$= \left(\frac{6nV}{1v}\right)^2$	where 6nV is the noise voltage and 1v is the input before limiting.

The noise power P at a 30KHz offset is, from eq 1, -135 dBc/Hz.

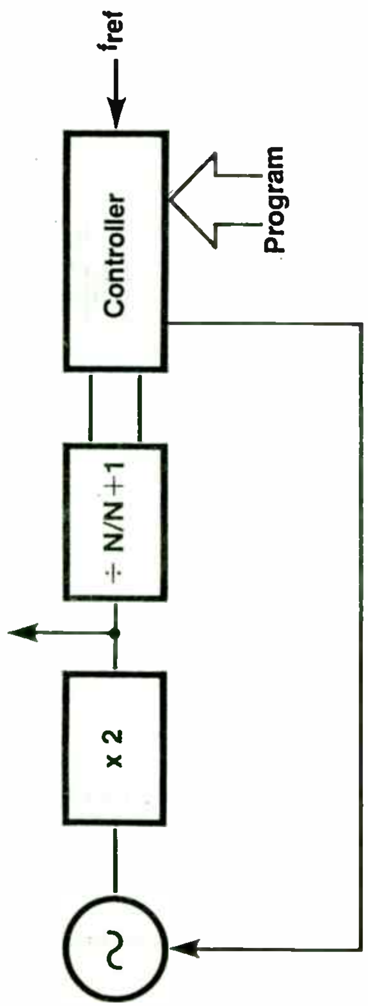
So far flicker noise has been ignored. Flicker noise is a low frequency phenomonen which causes problems by intermodulation with the carrier frequency to produce noise sidebands. The "break point" at which flicker noise becomes dominant varies but a UHF VCO of the type under consideration would probably have a break point at about 50 - 150KHz offset from the carrier. Eq 1 needs some modification to include this factor and a multiplicand of

$$\left(1 + \frac{f_e}{f}\right)$$

may be used, where f_e is the $1/f$ noise corner frequency.

The previously calculated noise will now be degraded by about 8 dB under these conditions, (assuming f_e = 150KHz) and will now be -127 dBc/Hz. This is about 5 dB inside the previously calculated requirement. Note that calculations have been made on the basis of a 30KHz offset to allow for doubling the oscillator frequency.

Considering an oscillator with a fundamental frequency of 800MHz, a number of problems appear. Ignoring for the time being the increased noise figure of the device, the available Q of components is considerably less - for example high quality chip capacitors can offer Q's of about 200, leading to working Q of about 100. Calculating noise levels for a 60KHz offset with all other parameters constant except tank capacity which is 12pF (half the 400MHz oscillator) the noise at 60KHz is -105 dBc/Hz or about 17 dB outside the requirement. Obviously, these figures are no more than a guide, but the suggestion is that the doubled 400MHz oscillator will meet requirements, while the 800MHz oscillator will not.- fig 8.



Use of a lower frequency oscillator for improved phase noise

fig 8

Flicker noise can be reduced by the inclusion of local DC negative feedback, such as an unbypassed emitter resistor, but a major requirement is to choose a suitable device. In general a low phase noise oscillator will run at high power, using a device with both low flicker noise and low high frequency noise, and with high gain and minimum damping on the tuned circuit. In fact, in many applications, the thermionic tube is attractive! Q should be as high as possible, and where VCO's are concerned, the MHz/Volt should be minimised. This is because of the effects of noise - at 10MHz/Volt, 1 microvolt of noise will produce 10Hz of FM deviation.

Where relatively wide frequency ranges are concerned, the variation in loop bandwidth may cause problems.

$$W_n = \sqrt{\frac{K_o K_v}{N t_1}}$$

where W_n = natural loop frequency
K_o = VCO constant
K_v = phase detector constant
N = divider ratio
t_1 = integrator time constant

W_n varies with N, and where desirable to maintain equal lock up times and loop bandwidth, K_v may be designed to vary with N. Several methods exist, but the use of a transmission line VCO can prove useful, as the effective inductance increases with frequency. The use of a suitable length of transmission line can provide an oscillator tuneable from 130 to 190MHz with a coarse tuning trimmer, and electrically tuneable over 6MHz at the bottom on the band to 8.75MHz at the top, thus maintaining W_n sensibly constant. The use of PIN diodes to switch capacitors is possible, although care must be taken not to degrade Q e.g. a 10pF capacitor at 150MHz has Xc = 106ohms. A PIN diode with an ON resistance of 0.5ohm will give QMAX = 212, assuming a perfect capacitor, and as considered earlier, this can have disastrous effects on phase noise performance.

An initially attractive method of realising the transmission line VCO is shown in fig. 9, where a length of line is used as a reactance inverter, changing the capacity into an inductance. The use of a Smith Chart will, however, show that the resulting inductance will have a low reactance unless the terminating capacitor is large and the line relatively long (greater than 1/8 wavelength). This leads to a low Q circuit as the resistance of the line is constant, and measurements made using a 16cms rigid coax 75ohm line with a loss of 4dB/100ft at 150MHz gave a Q of less than 100. This line was terminated with an air spaced trimmer. The same line as a shortened capacitively loaded resonator as in fig. 10 had a Q of over 250.

5. SUMMARY

The compromises in the synthesiser design are now apparent: a narrow bandwidth is required to minimise multiplied reference noise, but a wide bandwidth is needed to minimise lock up time. A high oscillator frequency may be required to avoid spurious outputs and multiplier chains, while a low frequency and multiplier chain give the best performance on system phase noise and possibly power consumption. The classical way to minimise these problems is the two loop synthesiser, but cost is a determining factor effecting the compromise finally reached. Power consumption is always a problem and unfortunately is more demanding at high frequencies while increasing channel occupancy will lead to ever tighter performance requirements in terms of phase noise and switching time.

Modern integrated circuits help the designer by providing better phase detectors and faster lower power dividers. Nevertheless, the single loop synthesiser has been shown to involve a number of compromises in its design, and in some cases, these compromises may limit the final equipment performance level. The single loop synthesiser is very useful, but is not universally applicable.

REFERENCES

1. Dynamic Range, Intermodulation and Phase Noise
   P.E. Chadwick, Radio Communication March 1984, PP223 - 228

2. The SL6440 High Performance Integrated Circuit Mixer
   P.E. Chadwick, Wescon 1981, Session 24 Record, entitled "Mixers for High Performance Radio". Published by Electronic Conventions Inc., 999 N. Sepulveda Blvd, El Segundo, CA 90245.

3. High Performance Integrated Circuit Mixers, P.E. Chadwick, rF Design, June 1980 pp20 - 23.

4. High Performance Integrated Circuit Mixers, P.E. Chadwick
   Clerk Maxwell Commemorative Conference on Radio Receivers and Associated Systems, Leeds 1981 (I.E.R.E. Conference Publication No. 50, ISBN 0 903748 45 2).

5. Frequency Synthesisers, Theory and Design. 2nd Edition.
   Vadim Manassevitsch, Wiley, 1980, ISBN 0 471 07917 0.

6. Characterisation of Frequency Stability: A Transfer Function Approach and its Application to measure via Filtering of Phase Noise.
   J. Rutman, Trans. IEEE on Instrumentation and Measurement, Vol 22 (1974) pp 40-48.

7. Phase Noise Measurement using a High Resolution Counter with On Line Data Processing. Peregrine, Ricci. Proc 30th Annual Symposium on Frequency Control, U.S. Army Electronics Command Ft. Monmouth N.J. 1976.

8. Phase Noise in Signal Sources. W.P. Robins, Peter Peregrinus Ltd., London 1982 ISBN 0 906048 76 1. pp 173 - 202.

9. Digital PLL Frequency Synthesisers, Theory and Design.
   Ulrich L. Rohde, pub Prentice Hall, 1983. ISBN 0 132 214239 2. pp 86 - 87

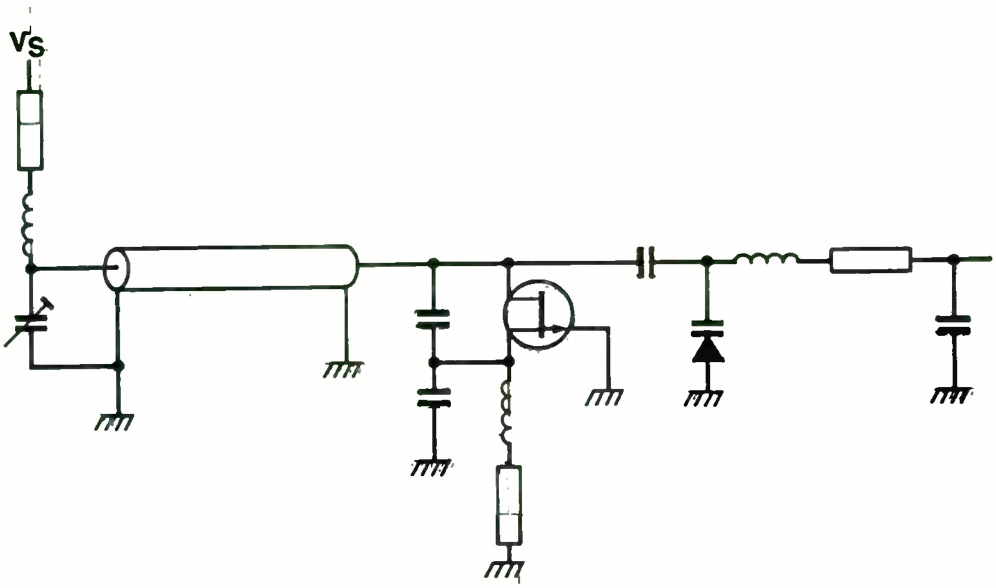
10. State of the Art Signalling Devices for Mobile Radio Systems - Selective Call, Tone Squelch and Digital Signalling.
    L.G. Litwin, Proceedings Communications 84, IEE Conference. Publication No 235, ISBN 085296292 4, 1984./

11. AN1006, A VHF Synthesiser using the SP8906 and NJ8811.
    Plessey Semiconductors Ltd. Swindon, England.

12. A Simple Model of a Feedback Oscillator noise Spectrum.
    D.B.Leeson, Proc IEEE Vol. 54, February 1966

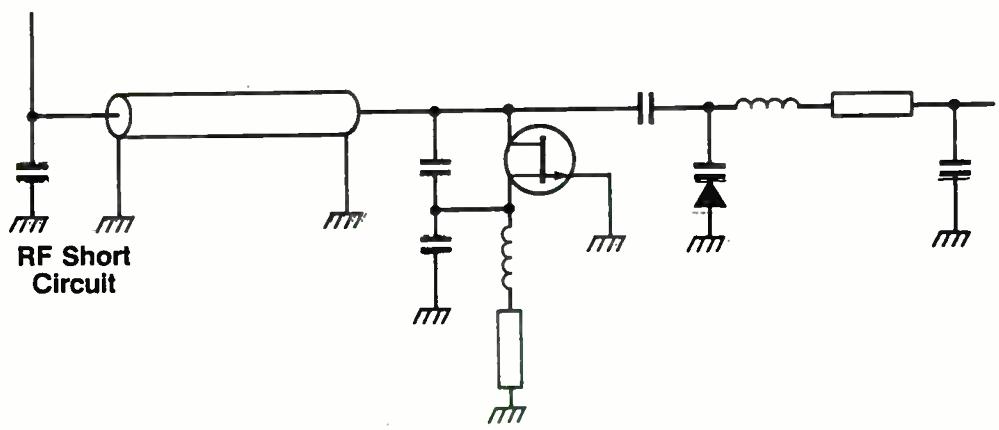
13. Phase Noise in Signal Sources, W.P. Robins pp 47 et seq.

14. Learn About Low Noise Design. Dieter Scherer
    Microwaves, April 79 pp 116 - 122 May 79 pp 72 - 77.



Transmission line VCO using the line as an impedance inverter

fig 9



Transmission line VCO using a shortened $\lambda/4$ line capacitively loaded

fig 10

PAPER TITLE: Temperature Stabilized RF Power Detector

Mr. George M. (Mike) Walley
Lead Engineer - Technical Staff
Harris Corp. Government Satellite Comm. Division
Box 93000 MS 1-1000
Melbourne, FL 32902
(305) 727-6344

ABSTRACT:

Temperature Stabilized RF Power Detector

Often in the design of high performance communications equipment a means of measuring an RF power level is needed. Consider the usual application of automatic level control of an RF signal. This application can arise from the requirement to stabilize the output power level from a modulator or from a high power amplifier stage. Input automatic gain control (AGC) is another illustration of a use for an RF power detector as a measurement device to derive a loop error signal. Most applications also have the requirement that the stabilized RF level is relatively insensitive to temperature. This article introduces a circuit topology utilizing common RF detector diodes that is inherently more temperature stable than simple envelope detectors. Included in the article is test data taken to characterize the performance of the new detector topology at different operating ambient temperatures.

DEFINING THE REQUIREMENTS:

Often in the design of RF communication equipment or systems the need arises for a simple RF power detector. A very common application of this class of simple power detector is as the control feedback detection element in an automatic gain control loop. In this application the RF detector will operate most of the time with an input at some specific power level. Since the function of the loop is to keep the power level at the detector input constant the detector only has to have well characterized behavior in a small operating range. Hence, this typical application usually has no requirement for a specific relationship between input power level and the output voltage level of the detector. The more stringent requirements are for repeatability, manufacturability, and temperature stability.

The repeatability problem has been approached in a number of ways. One method is to use high precision components and / or calibrated compensation networks. The overall problem is made easier if the power detector circuit is a simple, minimum component count implementation. If appropriate application of available precision components is made, the compensation of residual temperature and ageing effects is simplified. For example, temperature stable resistors, OP-AMPS and other components can be economically designed into a system. The detector diodes are a different matter entirely. No matter how much they cost they are inherently unstable in regards to their behavior at different temperatures. Any shift in their operating characteristics with temperature leads directly to a shift in the operating point of the leveling loop they are contained in.

It should be clear also that some facets of the repeatability solution yield positive results in terms of manufacturability also. If there is a minimal number of components in a proposed circuit this implies a simplified manufacturing process. More important, the subject of producability is tied to whether or not there are any select-and-test (SAT) elements in the design. These require a higher degree of skill and knowledge from the manufacturing personnel and slow down the process. Any adjustments necessary also lead to cost increases. The cost impact of adjustments can be reduced to a tolerable level if the adjustment procedure is simple and the adjustment itself is easy to make and remains at a stable setpoint. This goal in itself demands temperature and time stabilization of the circuit.

DEVELOPMENT OF THE CIRCUIT:

Consider the circuit shown in figure 1.0 . This is a basic envelope detector using a fast recovery diode. A limitations of this circuit topology is the need for a relatively high level signal to overcome the threshold of the diode. Also there is no form of temperature compensation. Extending this circuit to the design shown in figure 2.0a. helps with the threshold problem by providing the diode with DC biasing. A disadvantage of this approach is the reliance on the stability of the power supply voltages over time and temperature. Temperature induced changes in the threshold are not compensated for and can lead to significant shifts in an AGC operating point when using this detector. First order temperature compensation of the threshold shifts can be improved with the addition of diode D2 shown in figure 2.0b. Since the threshold of the detector diode is related to the forward DC bias current, the current flow can be compensated by D2. If the current changes, the forward voltage drop across D2 moves in the opposite direction which stabilizes the voltage

Temperature Stabilized RF Power Detector

across D1. This technique has been used with AGC loops having a stability requirement of about ±1.5 dB. To hold a tighter tolerance requires a new circuit topology taking into account both power supply variations as well as temperature shifts in the diode characteristics.

The temperature effects can be reduced significantly with a differential amplifier type of detection circuit. Shown in figure 3.0 is a circuit that will allow the use of low cost precision OP-AMPS and will still yield improved temperature stability. Temperature shifts in the DC offsets of the diodes are cancelled out and the overall output stability with temperature is a function of how closely the characteristics of the diode pair match. The operating temperature of both diodes also needs to be well matched. Adjacent location of the diodes on a printed circuit board is an easily realizable minimum requirement. The two diodes could be on the same monolithic die to provide near perfect temperature tracking. With the addition of a constant current source for providing DC bias, the effects of DC impedance changes in the diodes and power supply variation are reduced. Figure 4.0 is a schematic diagram of a complete temperature compensated RF power detector. The voltage reference is itself temperature compensated and is a relatively low cost addition to the circuit.

CIRCUIT PERFORMANCE:

This basic circuit topology has found many applications in the design of precision RF leveling loops in the past and has been well characterized over temperature. Figure 5.0 is a plot of the response of this basic circuit at 0°, 25°, and 70°C. As can be seen from the graph only a small shift in the output voltage of the detector takes place with changes in temperature. What this means for an AGC loop can be understood by looking closely at a section of the graph around a possible operating point. Figure 6.0 is an enlarged section of the graph of figure 5.0 about a hypothetical operating point of +2.0 dBm. This shows that an AGC loop using this detector could be expected to exhibit a temperature drift of only about ±0.4 dB. This is a significant improvement over the typical value of ±1.5 dB for the simple detector of figure 2.0b. For certain applications this level of performance is still not satisfactory. For these instances one more step in the quest for temperature stability can be taken. That is to place the diode pair in a temperature controlled environment, namely a component oven. The component oven is operated at a temperature higher than the maximum operational temperature the circuit is being designed to. A component oven that will keep a pair of RF diodes at a constant 75°C was developed for a past project and has found itself in several later designs. The oven is an Ovenaire part number PC2 - 82. The response of the circuit with the component oven is shown in figure 7.0. As can be seen the temperature drift to be expected in a closed loop application has dropped to about ±0.1 dB. This variation is near the limit of measurement with general purpose test equipment. The residual operating point shifts are from imperfect balance between the diode bias currents as well as offset drift in the differential amplifier.

CONCLUSIONS:

The subject of the application of temperature compensated RF power detectors was discussed. The major points to consider being the requirements for repeatability, manufacturability, and temperature stability. It was shown that for modest increases in component costs and complexity large improvements in all three areas are to be gained.

The development of a circuit topology that is inherently more temperature stable was presented as well as test data taken from a typical implementation of that topology. The further enhancement of the circuit was discussed from the stand point of obtaining near perfect control of output level variation with temperature. This was obtained at the cost of controlling the diode junction temperature directly. It is important to note that simply controlling the junction temperature of the diodes in a simple detector is not sufficient. This is because the other elements of the instability with temperature would then become dominant. Thus making the mere addition of the component oven of lesser value than if used with the new circuit topology.

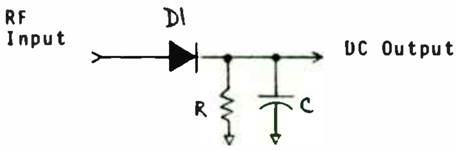


Figure 1.0 Basic Envelope Detector

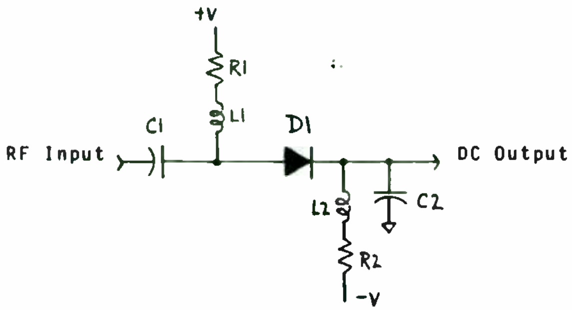


Figure 2.0a Basic Dtector with Bias

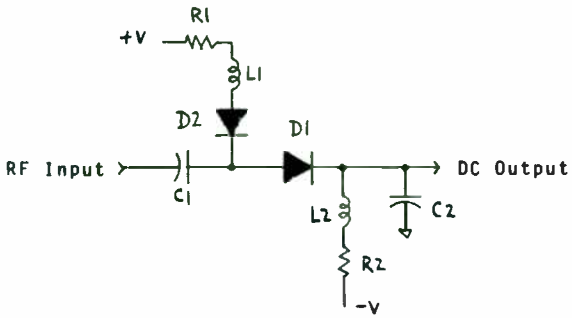


Figure 2.0b Simple Temperature Compensation

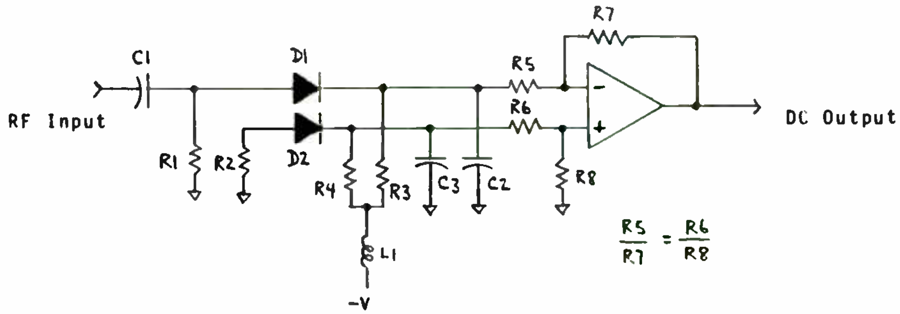


Figure 3.0 Temperature Effects Cancellation
With Differential Amplifier

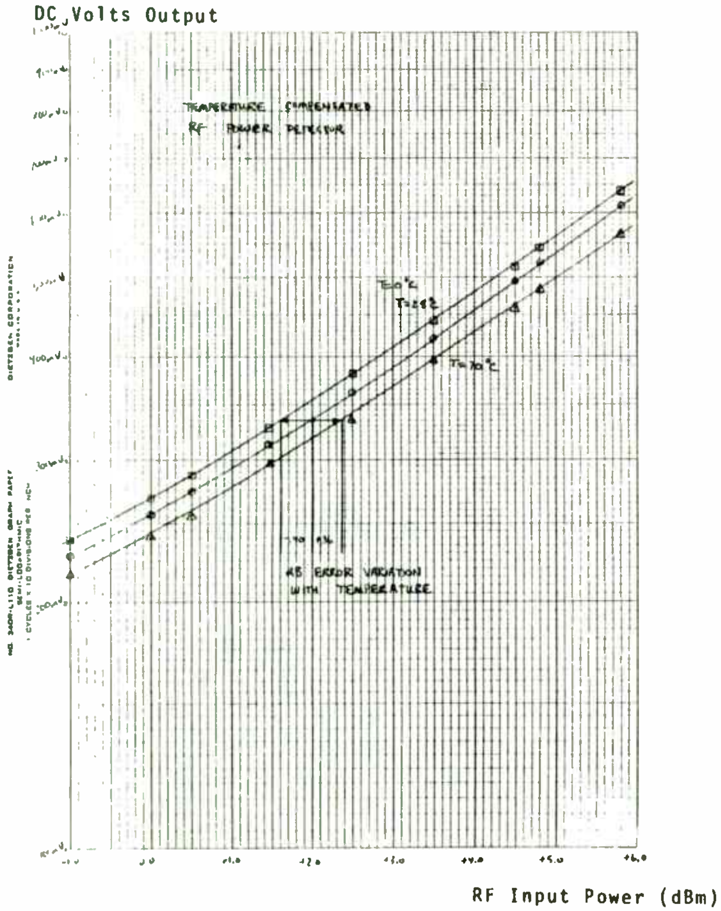


Figure 6.0 Expanded Scale of Detector About +2.0dBm

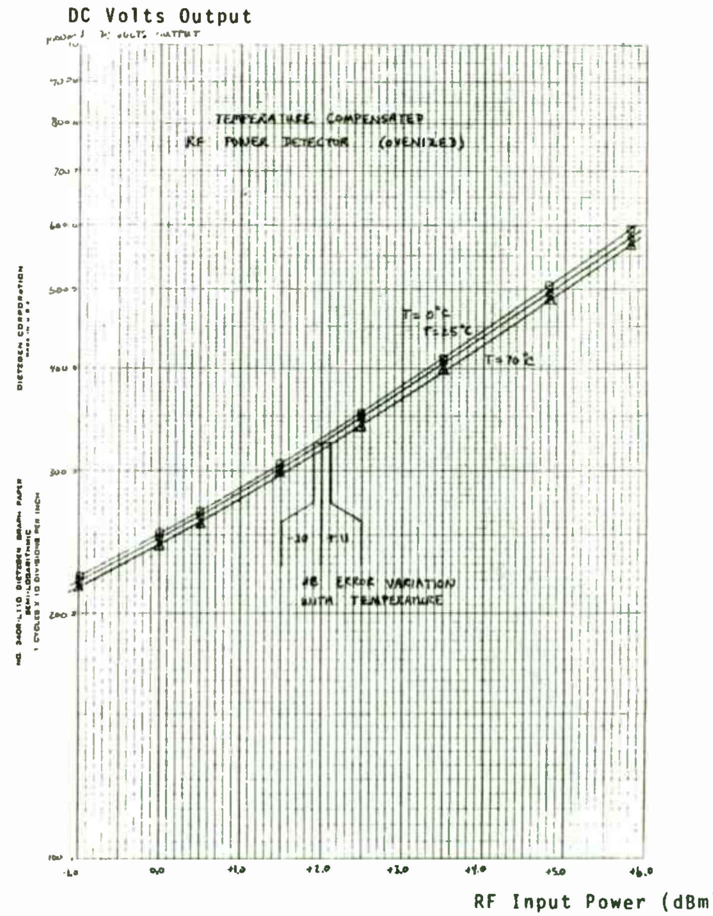


Figure 7.0 Expanded Scale Detector Response with Ovenized Detector Diodes

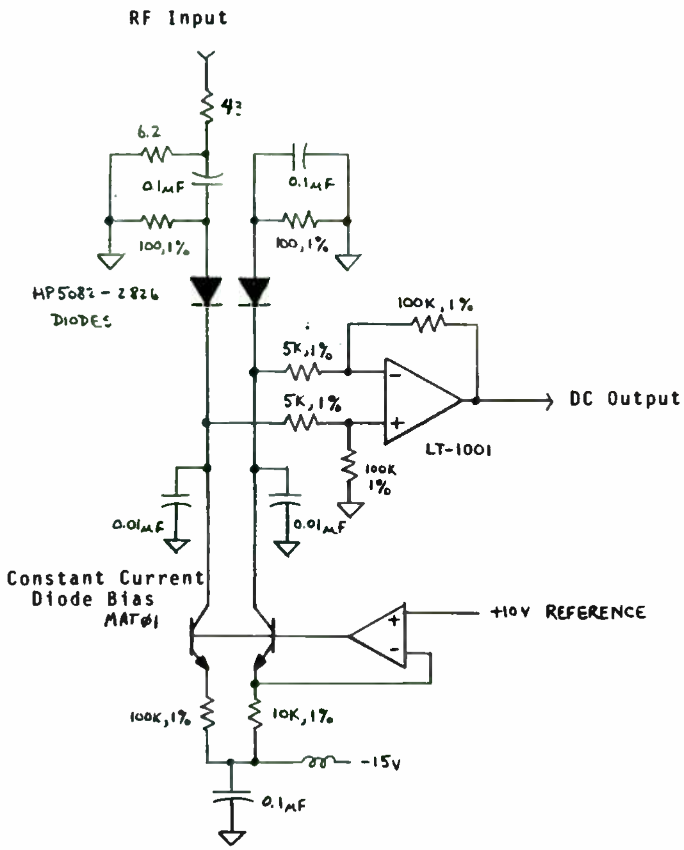


Figure 4.0 Complete Temperature Compensated RF Power Detector Circuit

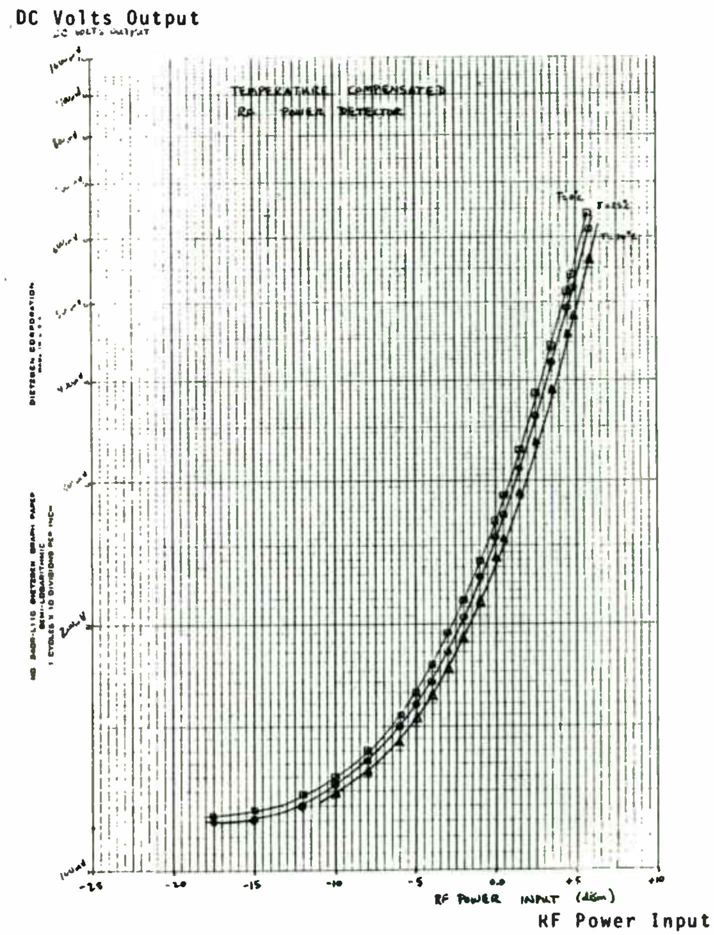


Figure 5.0 Characteristics of Temperature Compensated RF Power Detector

AN ULTRA-FAST UHF VOLTAGE CONTROLLED ATTENUATOR
WITH 35 dB OF LINEAR DYNAMIC RANGE

Daniel L. Gerlach
California Amplifier Inc.

ABSTRACT: The analysis of a fast PIN diode voltage controlled attenuator is presented in this paper. Design Applications, D.C. and R.F. features, and mathematical derivations for the driver section have been included. Complete circuit schematics and response curves are provided. A temperature compensation network is incorporated to provide system linearity over temperature.

INTRODUCTION

There are presently several independent manufacturers of wide band voltage controlled attenuators, (VCA). The product described here is just as good, or better than most of the VCA's on the market and can be built for a fraction of the cost of thin-film or hybrid VCA's.

Another benefit to the VCA described includes extreme versatility in selection of control voltage range. The VCA runs off of positive control voltages but can easily be modified to accept negative signals.

APPLICATIONS

The versatility of this design makes it suitable for many applications. In the laboratory it can aid in compression point measurements of amplifiers and transmitters. The VCA can also be used for power and gain matching of two or more amplifiers. The attenuator can be cascaded with an LNA front-end to produce a low noise voltage controlled amplifier. In the field the VCA can be used to determine receiver sensitivity by incorporating it into the transmitter. Mobile transceiver systems can be analyzed with the help of this VCA. The unit can also be used as a variable RF limiter to protect small signal amplifiers and receivers from excessive power levels.

DESIGN FEATURES

The VCA described is basically a two section softboard design utilizing an RF (attenuator) section and a driver (linearizer) section. The unit can be controlled by digital, analog, or software generated signals. The VCA described here utilizes a modified Pi configuration to achieve a zero-to-maximum or maximum-to-zero change in 2.5 microseconds with less than 0.5 dB overshoot. The 0 to 100% switching speed is 3.0 microseconds. The maximum input and output VSWR is 1.6:1 with an insertion loss of 2.5 dB. The linearizer described utilizes an operational amplifier and two Zener diodes to achieve an overall linearity spec. of ±1.5 dB. All of these specifications are held over the -30 to +60°C temperature range and 100 to 1000 MHz frequency range by incorporating a standard temperature compensation network into the overall design. The VCA utilizes only stock components that are available from most distributors and it will fit on a 1.5" x 2.5" P.C.B. Power requirements are +15 VDC at 10 mA and -15 VDC at 5 mA.

SYSTEM DESIGN

The system was designed using component design techniques. Both components were designed and tested separately, prior to integration. Standard temperature compensation network techniques were employed and added to the VCA after integration. The temperature compensation network will be

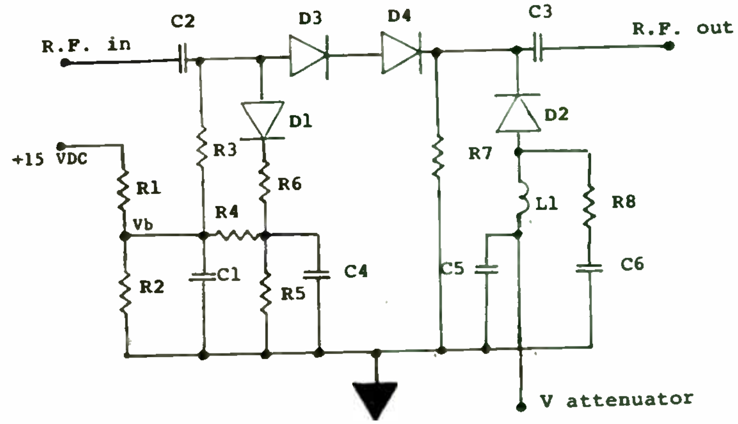
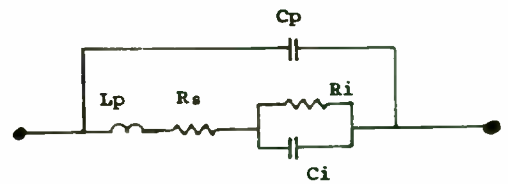


FIGURE 1
Attenuator Section



Cp = Package capacitance

Lp = Package inductance

Rs = Substrate resistance

Ri = Insulating layer resistance

Ci = Insulating layer capacitance

FIGURE 2
PIN Diode Equivalent Circuit

described at the end of this section.

ATTENUATOR DESIGN

The attenuator section was designed with the help of Touchstone, a microwave CAD software tool manufactured by Eesof Inc. of Westlake Village, CA. The actual design is a modified Pi attenuator network incorporating 4 PIN diodes instead of the usual 3 needed in most Pi networks. The fourth diode, D4, mounted in series in the R.F. line, is used to reduce the overall parasitic package capacitance in the series arm by a factor of 2 when D3 and D4 are conducting (See Figure 1). This reduction in package capacitance serves to reduce the insertion loss of the attenuator at frequencies below 500 MHz. A second modification included incorporating a 50 ohm chip resistor, R6 and R8, in each of the shunt arms. These resistors aid in maintaining good input and output VSWR during periods of maximum attenuation, i.e. shunt arms conducting. A final modification, an offset bias voltage, Vb, and corresponding voltage divider networks, R4, R5 and R3, R7 serve two purposes. First, this bias voltage allows for the condition of minimum attenuation to exist as well as providing for uniform attenuation at low control voltage settings. And secondly, the divider network allows for a wide selection of control voltage ranges. The initial resistor values were obtained from [1]. The PIN diode equivalent circuit shown in Figure 2 was used to implement the computer model. Values for the various lumped elements in the model were supplied by the diode manufacturer. Slight changes (less than 10%) were needed in the values for Cp, Lp, and Rs to obtain agreement with lab results. Various PIN diodes were evaluated and all performed well. Hewlett Packard, Alpha, Metelics or SDI PIN diodes can all be used in this design. Figure 3 shows the typical exponential response for the circuit of Figure 1.

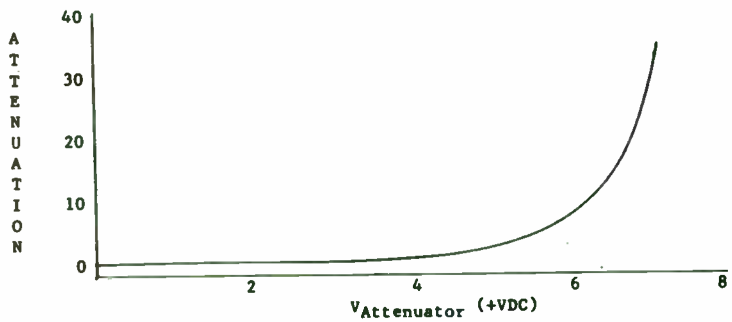


FIGURE 3
Typical Response Curve For Attenuation vs. $V_{Attenuator}$

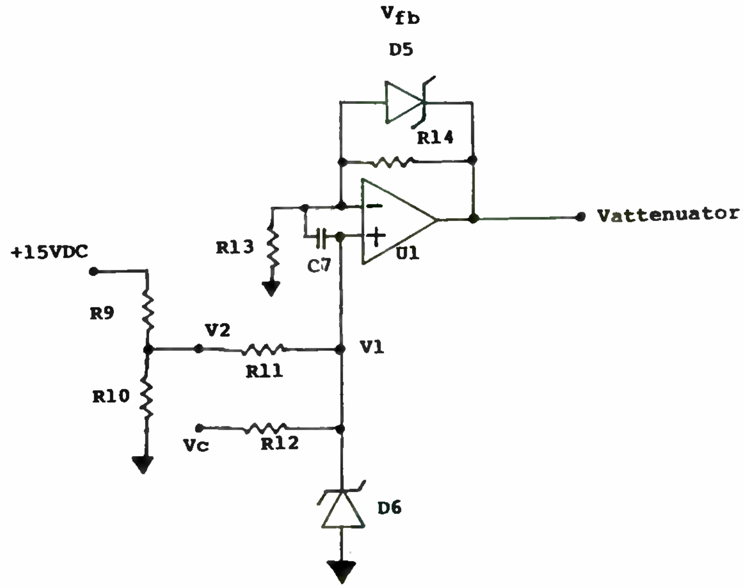


FIGURE 4
Linearizer Section

LINEARIZER DESIGN

The Linearizer section shown in Figure 4 was developed for three reasons. First, the circuit is fast, with rise and fall times of 2.5 micro-seconds, typical for 0 to -35 dB and -35 to 0 dB attenuation ranges. Compare this to 5 mili-seconds for the circuits depicted in [2] and [3]. Second, the circuit is extremely versatile. By changing R12 and the divider networks in the attenuator, any control voltage range can be easily obtained. And third, the circuit is compact, incorporating only 10 components in contrast with 19 and 14 for [2] and [3] respectively.

In order to maintain a linear response over the 35 dB attenuation range, the linearizer section must produce a logarithmic response that is the inverse of the exponential response of the attenuator section. The Zener diodes, D5 and D6, accomplish this task as by providing two breakpoints shown in the following derivation:

As Vc, (overall control voltage), is increased, the desired logarithmic response is realized by turning on D5 and D6 at the correct times. For low values of Vc, both D5 and D6 are off and appear as open circuits yielding maximum gain from the linearizer. For medium values of Vc, D5 begins conducting but D6 is still off. (V1 < the Zener voltage of D6) This yields in the ideal case a closed loop gain of unity since the feedback resistance of U1 is effectively reduced to zero. Finally, for high values of Vc, corresponding to the upper range of maximum RF attenuation, both D5 and D6 are conducting and V1 remains constant for increasing Vc. This results in a constant output voltage for increasing input voltages and completes the ideal logarithmic response curve shown in Figure 5. The mathematical derivation for the ideal case described above is now shown.

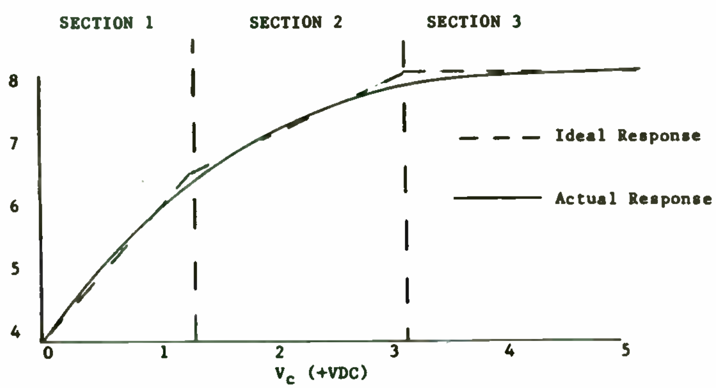


FIGURE 5
Typical Response Curve For $V_{Attenuator}$ vs. Vc

CASE 1: Both D5, D6 off

The input voltage to the op-amp can be calculated as the additive sum of V2 and Vc, as shown below:

$$(1) \quad V1 = \frac{V2\ R12}{R11+R12} + \frac{Vc\ R11}{R11+R12}$$

Note that V2 acts as an offset voltage and is used to set the starting value of Vc which corresponda to minimum RF attenuation. The value of $V_{attenuator}$ can now be found by using the non-inverting op-amp gain equation shown below.

$$(2) \quad V_{attenuator} = \frac{R13 + R14}{R13} V1$$

Equation 2 sets the initial gain of the circuit corresponding to Section 1 of Figure 5,

CASE 2: D5 on, D6 off

In this case, since R14 is ideally equal to zero, equation 2 can be rewritten as:

$$(3) \quad V_{attenuator}' = V1$$

During case 2 operation, the overall gain of the Linearizer ($V_{attenuator}/Vc$) can take on any value from zero to unity simply by choosing the correct ratio of R11 and R12. Equation 3 sets the medium value gain of the circuit corresponding to Section 2 of Figure 5.

CASE 3: Both D5, D6 on.

In this final case, no further increase in $V_{attenuator}$ can be realized since V1 is being held constant by the breakdown voltage of D6. Therefore, equations 2 and 3 become:

(4) $V_{attenuator}'' = K =$ the breakdown voltage of D6

Equation 4 corresponds to Section 3 of Figure 5.

In order to select the proper values for R11-R14 and V2, a relationship between RF attenuation and $V_{attenuator}$ must first be developed. To accomplish this, the circuit of Figure 1 must be built and data taken on attenuation level vs. $V_{attenuator}$. The data corresponding to Figure 1 is shown in the first two columns of Table 1 below. The next step is to add the third column shown in Table 1. The choice of values for Vc determines the voltage control range of the VCA (0 to 5 VDC in this case).

RF Attenuation -dB	V Attn. +VDC	Vc +VDC
2	4	0
9	5.8	1
16	6.6	2
23	7.1	3
30	7.25	4
37	7.3	5

TABLE 1
Attenuator Board Analysis

We have now developed the relation between $V_{attenuator}$ and Vc which will be used to solve for the unknown resistor values and voltage level. A final step is to incorporate eq. 1 into eq. 2 and put the resulting equation into the form shown below:

(5) $V_{attenuator} = A[B*V2 + C*Vc]$

where $A = (R13 + R14)/R13$
$B = R12/(R11 + R12)$
$C = R11/(R11 + R12)$

(6) and $B/C = 1/(B/C)$

Using eq. 5, eq. 6 and three data points from Table 1, the values of R11-R14 and V2 can be obtained.

PRACTICAL CIRCUIT CONSIDERATIONS

In reality, the Zener impedance of D5 never reaches zero, therefore Equation 3 must be rewritten as shown below:

(7) $V_{attenuator}' =$

$$\frac{R13 + Z_{D5}//R14}{R13} V1$$

where: $Z_{D5}//R14 =$

The maximum Zener impedance of D5 in parallel with R14.

Usually Z_{D5} is less than 50 ohms for most Zener diodes. The values of R13 and R14 should be chosen such that this 50 ohm is less than 2% of the value of R14. It is also apparent that equation 4 is in error and that because of the finite value of the impedance of D6, $V_{attenuator}$ never reaches a constant value. This error is shown as the slight positive slope of Section 3 of the actual curve in Figure 5. As with D5, the values of R11 and R12 should be chosen such that the Zener impedance is less than 2% of the value of R12 (or R11 + R10, whichever is smaller). The high impedance of the non-inverting input of U1 compared to the Zener impedance of D6 also helps to reduce this source of error. As Figure 5 shows, these sources of error help to smooth out the logarithmic response of the linearizer which will actually prove beneficial to the overall attenuator circuit.

TEMPERATURE COMPENSATION NETWORK

The above described circuit will provide at least 35 dB of linear dynamic attenuation at room temperature only. As the temperature increases, $V_{attenuator}$ drops, the dynamic resistance of the PIN diodes (Rs and Ri) increases, and the overall resultant attenuation level decreases. In order to compensate for this, the temperature compensation network shown in Figure 6 was used to replace R5 in the attenuator section. Other resistors could have been replaced with this network, but this would adversely affect the linearity characteristics. The actual design equations are well known and can be found in either [4] or [5]. If the unit is to be operated under severe temperature conditions the Zener diodes can be replaced with Temperature Compensated Reference diodes to aid in temperature compensation.

SYSTEM TEST AND EVALUATION

The attenuator section was laid out on .031" OAK-601 fiberglass and PTFE Laminate with 1 oz. copper clad on 2 sides. The linearizer section was also laid out on OAK-601. Chip components were used throughout the RF section except the PIN diodes, which were axial lead, glass packages. No tuning in the RF section was required. The attenuation vs. control voltage characteristics are shown in Figure 7. R10 can be adjusted to set the initial value of Vc if required. Similarly, R13 and/or R12 can be adjusted to change the linearity of the VCA. Capacitors C4, C6 and C7 can be adjusted to improve the switching speed. The circuit described operates linearily from -30°C to +60°C with no degredation to the specifications given in the Design Features Section.

The noise figure results show good agreement with the predicted insertion of 2.5 dB when measured in a 50 ohm system. The switching speed is limited primarily by the Zener diodes, D5 and D6.

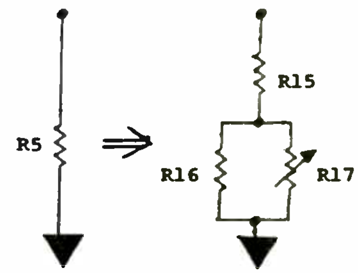


FIGURE 6

Temperature Compensation Network

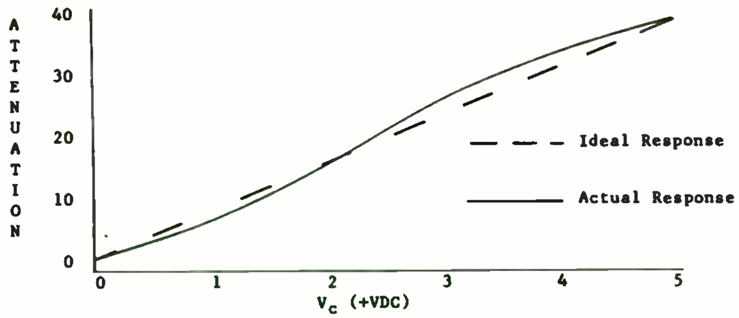


FIGURE 7
Attenuation vs. Vc

CONCLUSION

The VCA design presented here exhibits very good switching characteristics and linearity over temperature. Using the design equations and schematics provided, this VCA can be built and ready for laboratory or field use in less than one weeks time. The design allows for complete versatility in control voltage settings. By modifying the driver section, negative control voltages can also be used. By incorporating only off-the-shelf components, this design has proven to be both cost effective and efficient. Comparisons of experimental and theoretical results show this design to be accurate over a wide frequency range.

REFERENCES:

1. Hewlett Packard Application Note 936 "High Performance PIN Attenuator for Low Cost AGC Applications", October 1971.
2. National Semiconductor Corporation "Linear Applications" November 1967, pg. AN30-1.
3. Jacob Millman "Micro-Electronics Digital and Analog Circuits and Systems, 1979, McGraw-Hill, pg. 601
4. Stan Jaffe, "Temperature Compensation Using Thermistors", Microwaves & RF, April 1984, pg. 101.
5. Thermometrics Catalog No. 181-D, "Technical Applications and Data", pg. A-3

Design, Characterization and Application of a GaAs Power FET

Y. Hwang

Gould Inc. Dexcel Division

I. Introduction

Design considerations of a C-band GaAs power FET are described in section II. Fabrication process is described in section III. Section IV presents dc parameters and microwave performance of the FET. Section V covers circuit development. This includes the design and performance of a three watt IMFET by combining two FET and the development of a 21 dB gain two stage amplifier.

II. Design considerations

To design a power FET, several important factors must be considered including:

1. Proper total gate width to achieve the desired output power.
2. Obtain a sufficient power gain by reduction of gate length, source inductance and parasitic elements.
3. Maintain high drain - source breakdown voltage.

Design goal is to have a FET capable of generating two watt and good gain at 8 GHz. The device chosen to meet the goal is a single chip having 5mm total gate width, 40 parallel gates with 4 gate bonding pads and 4 drain bonding pads. Gate length is 0.5 micron and gate width is 125 micron per finger.

III. Device fabrication

Fabrication process of the GaAs FET is described in this section.

1. Mesa isolation: Epitaxial layer except for the channel under the gate is etched.
2. Ohmic contact: After mesa etching, the source and drain contacts are fabricated.
3. Gate recess: Prior to the gate metal evaporation, the conductive active layer is etched down to achieve a recessed gate structure.
4. Gate metallization: The gates are formed by a lift-off procedure.
5. Nitride protection: The device active region is covered with a thin layer of silicon nitride to protect the device from scratches.
6. Air bridge: Source fingers and source pads are connected by air bridge crossover.
7. Via hole/plated heatsink: Holes are etched through the substrate until the source pad is reached. These holes are then plated at the same time as the ground plane.

IV. Device evaluation

1. dc Characterization

Most dc characteristics can be obtained with a curve tracer. Fig. 1 shows the drain current Ids versus drain-to-source voltage Vds characteristics of this 5mm FET. Typical dc parameters of the FET are: Idss = 1100 mA, gm = 300 mmho, and Vp = 4v, where Idss, gm and Vp indicate the saturated drain current, transconductance and pinch-off voltage, respectively.

2. RF performance

The purpose of RF characterization is to obtain either an equivalent circuit for the device or a set of parameters which describes the transistor operation by an equivalent circuit. In both cases, the first step is the measurement of the scattering parameters. (s-parameters) of the device under different bias conditions and in various frequency range with a network analyzer. Fig. 2 shows measured s11 and s22 from 2 to 12 GHz of this 5mm FET. The bias is set at Vds=9v and Ids = 500mA which is typical bias used on this device. Power gain and output power at 1 db compression point are demonstrated in Fig. 3. Two watt output power and 6 dB gain have been achieved at 1 db compression.

V. Use of the power FET

The power GaAs FET devices can be used in amplifier or oscillator circuit. S-parameter and power measurements of the FET presented in section IV are integral parts of the microwave circuit development. The circuit development effort has been concentrated on establishing the microwave circuit techniques for implementing the power FET in internally matched FET and amplifier circuit.

1. Computer aided circuit design

The results of the s-parameter measurements presented in section IV have been used in conjunction with computer optimization to generate the circuit elements of the matching networks for IMFET and a two stage amplifier. Fig. 4 shows the schematic diagram of IMFET. The input circuit consists of three lumped inductors(L0, L1, and L2), a low impedance transmission line TRL1, and a shunt capacitor C1. The Output circuit includes lumped inductor(L3), tramsmission lines TRL2 and TRL3.

2. Experimental data

Fig. 5 shows the output power versus frequency of this IMFET without external tuning. The IMFET has a 3 watt power output at 1dbc with a linear gain of 8 dB from 5.9 to 6.4 GHz. A unique feature of the 2-stage amplifier is its compactness. The size of the amplifier module is 0.5 inch X 0.5 inch. A gain of 21 dB from 4.2 to 4.4 GHz with an output power of 1 watt has been achieved.

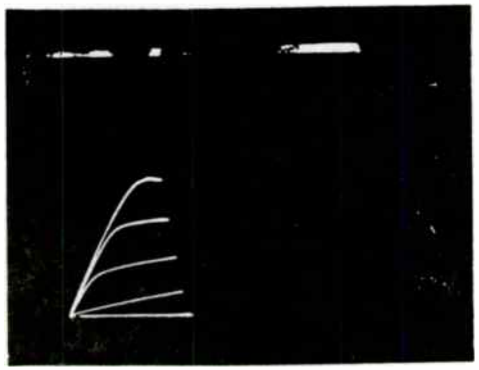


Fig. 1 Drain current Ids versus drain-source voltage Vds characteristics (total gate width: 5mm).

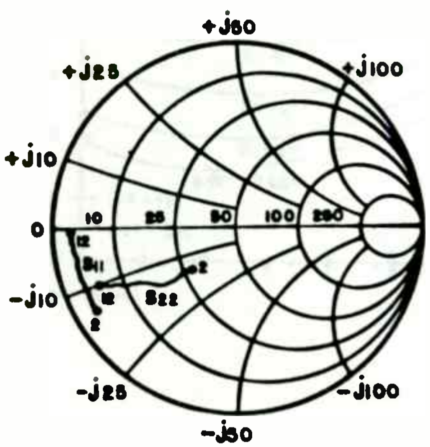


Fig. 2 Measured 2-12 GHz S_{11} and S_{22} of the 5 mm power FET.

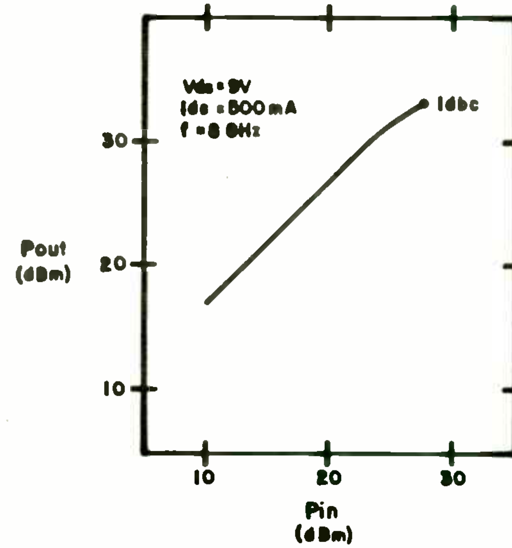


Fig. 3 Input/Output power response of the 5mm FET at 8 GHz.

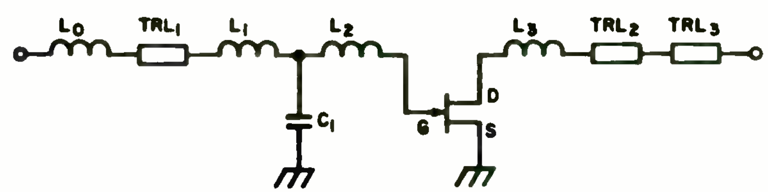


Fig. 4 Schematic diagram of the internally matched power FET.

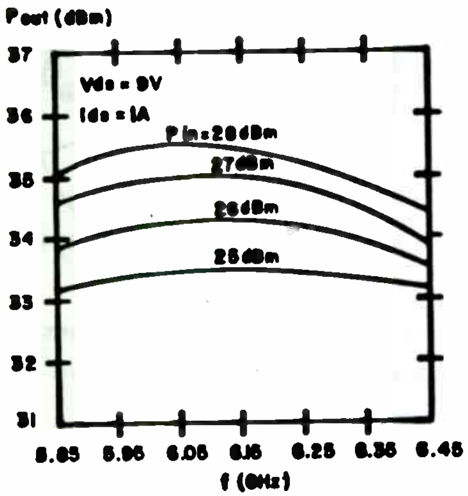


Fig. 5 Output power versus frequency of the 3 watt IMFET without external tuning.

ASPECTS OF DISCRIMINATOR DESIGN

Joseph F. Lutz

ABSTRACT

The purpose of this paper is to summarize the various discriminator types. The performance of two divergent implementations (tuned circuit and delay line) is compared through design examples. The component values are calculated and comparison of the actual hardware is made.

The paper begins by defining the various methods of demodulating an FM signal: differentiation followed by envelope detection and delay line detection. These are discussed in a conceptual framework. The pertinent equations are then developed and discussed. The actual design procedure is developed for a balanced slope demodulator followed by a design example in which L, R and C values are calculated. A calculator analysis of this circuit is presented and compared to the performance of the Lenkurt 48011 discriminator which also uses a balanced slope demodulator.

The delay line discriminator is developed and analyzed. Design equations are developed and a design example worked out. This example is compared to the Lenkurt 57040 delay line discriminator specification. The schematic of the 57040 is also examined to highlight the amount of design necessary to implement such a discriminator.

Also discussed are the techniques for measuring the performance of wideband FM discriminators used as performance monitors in long-haul microwave networks. An example is GTE Sprint.

THE FM DEMODULATION PROCESS

The purpose of a discriminator is to convert carrier frequency variation into an undistorted baseband output. There are three basic methods to accomplish this task: 1) differentiation followed by envelope detection; 2) synchronous detection; 3) FM feedback-PLL tracking. In case number one, the differentiation process translates the frequency variations into amplitude variations. The envelope detector is then used to remove the modulating signal. This is shown mathematically by equations 1 and 2. Notice that

$$v_{fm}(t) = \cos[Wct + \int m(t)] \qquad \text{eqn. (1)}$$

$$\frac{dv_{fm}(t)}{dt} = [Wc + m(t)] \sin[Wct + \int m(t)]$$

$$\frac{dv_{fm}(t)}{dt} = Wc \sin [Wct + m(t)] + m(t) (\sin Wct - \int m(t) \qquad \text{eqn. (2)}$$

after differentiation the signal consists of both AM and FM components. It is from the varying amplitude that the envelope detector recovers the modulating signal.

In synchronous detection a quadrature version of the received FM signal and the FM signal are mixed together. In essence what is happening is that the carrier wave is being synchronously beat to zero frequency with only the desired modulating signal remaining at the mixer output. Higher order terms will also be present, but they can be removed by a low pass filter. Later the equivalence of delay line detection and synchronous detection will be established. The delay line detector is easier to implement and provides much the same performance as a true synchronous detector.

The third way that an FM signal can be demodulated is to place an FM modulator in the feedback loop of an amplifier to perform the inverse of the modulation process. This can be done with a phase locked loop (PLL). For the PLL to track the FM waveform, the control voltage to the VCO must be a replica of the modulating waveform . This paper will primarily concentrate on the first two demodulation methods mentioned. Before continuing we will discuss some of the sources of noise and distortion in an FM demodulator.

NOISE AND DISTORTION IN THE FM DEMODULATION PROCESS

A close look at equation (2) will reveal why it is necessary to place a limiter ahead of an FM detector. Any carrier amplitude variations will show up as distortion in the baseband output. In wideband FM detectors the limiter can further introduce noise into the baseband by not sufficiently suppressing second order and higher harmonics, AM to PM conversion and through what are known as coupled distortions. A coupled distortion will result if there is a large IF response slope or group delay variation across the IF passband. In effect, what one requires is a flat amplitude and delay characteristic ahead of the limiter.

The most significant source of distortion in an FM demodulator is of course its linearity response. Ideally the baseband output will be directly proportional to the deviation of the carrier frequency, (See figure 1a). A good measure of the non-linearity of the response curve of figure 1a is to take the derivative. This is shown in figure 1b. This is known as the linearity or derivative response. The standard design goal is for the linearity to be within 1% over the desired band of interest. The actual measured affects of discriminator non-linearity is shown in figure 2. The linearity can be broken into two components (as shown in figure 2): Percent Slope and Percent Parabolic Linearity. Plotted on the ordinate is the signal-to-noise ratio in a 3 kHz slot at 70kHz. This corresponds to the lowest frequency in an FDM baseband multiplex group. Across from the S/N ratio on the opposite ordinate is the noise in pWpO. This is also a measure of signal-to-noise ratio or more precisely a noise-to-signal ratio. pWpO stands for picowatts psophometrically weighted relative to test tone. The psophometric weighting represents the bandpass characteristic of a 3 kHz telephone channel as defined by the CCIR. This unit is peculiar to the telephone industry. The total noise is thus the sum of the slope and the parabolic linearity contributions. A state-of-the-art discriminator can be expected to function somewhere in the neighborhood of 20 pWpO.

THE DIFFERENTIATION/ENVELOPE DETECTION METHOD

A block diagram of this system is shown in figure 3.

Let the modulated signal now be written as

$$v_{fm}(t) = \cos(w_c t + \beta \sin w_m t) \quad \underline{\text{eqn. 3}}$$

where:

$\beta = \frac{\Delta F}{w_m}$, modulation Index

ΔF = peak frequency deviation

w_m = modulating frequency

and the amplitude of $v_{fm}(t)$ has been normalized to 1. The mathematical analysis of figure 3 shows that the process of differentiation transfers the angle variation of the cosine function into an amplitude variation. The output of the differentiation is equation 4, 4 is $v_a(t)$. Notice that the signal at this point has both amplitude modulation and frequency modulation impressed upon it. The bandwidth of the signal $v_a(t)$ has been increased by $2w_m$ over the bandwidth of the original FM signal. The purpose of the envelope detector is to rectify the signal $v_a(t)$. This process is shown mathematically by equation 5. The second term of $v_a(t)$ is the original modulating signal. Not surprisingly, the amplitude of this signal is directly proportional to the peak frequency deviation. The low pass filter of figure 3 passes only the modulating signal. The final output of the discriminator will be equation 6: $v_{BB}(t)$.

The differentiation process, if shown graphically, can add a great deal of insight to what is actually physically happening. Consider splitting $v_{fm}(t)$ into 2 paths as shown in figure 4. The R leg is passed straight through while the L leg undergoes a 90° phase shift at the carrier frequency by virtue of the delay line placed in the L leg. t_o is the constant of the delay line which is later shown to be the group delay of the network. The R and the L legs are summed together by a differential amplifier with gain A_d. As the deviated carrier swings above w_c, the delay line retards the phase as is shown in the left half of figure 4. At the maximum positive frequency excursion, the value of the signal amplitude is R_{min}. Likewise, as the signal is swept to its maximum negative frequency excursion, the amplitude takes on the value R_{max}. The amplitude modulation imparted to the signal can now be clearly seen.

After a somewhat lengthy algebraic manipulation the RMS value of the envelope can be shown to be

$$V_{E\ RMS} = \frac{Ad}{2} \left(\sqrt{1+\sin(\Delta\omega p t_o)} - \sqrt{1-\sin(\Delta\omega p t_o)}\right) \qquad \text{eqn. 7}$$

A more conventional way of illustrating the same process is to use the tuned circuit slope detector of figure 5. Here differentiation is accomplished by taking advantage of the frequency response of a tuned circuit. In this case, however, note that the frequency response is different above the center frequency compared to below the center frequency and the resultant output is slightly distorted.

A more useful approach in wideband discriminators is the balanced slope discriminator, also known as the Travis or Round-Travis detector. A practical circuit and its frequency characteristic is shown in figure 6. The Travis circuit uses two tuned circuits; one tuned above the other and below the center frequency. The circuits are tuned such that the non-linear portions of their slopes cancel each other.

As an example, assume it is desired to have a 12 MHz 1% linearity bandwidth with an IF frequency of 70MHz.

A standard practice is to let

$$\delta f = 3/2\ BW \qquad \text{eqn. 8}$$

where:

δf = displacement of the tank resonant frequencies from the IF center frequency.

BW = desired bandwidth (1%)

For 12 MHz BW:

$$\delta f = 3/2(12) = 18\ \text{MHz}$$

The high-side tank will therefore be resonant at 88 MHz and the low-side tank at 52 MHz.

Choose the low tank values first by starting with a physically realizable value of L_L. Choose

$$L_L = 0.4\ \text{uH}$$

then

$$f_L = \frac{1}{2\pi\sqrt{L_L C_L}} \qquad \text{eqn. 9}$$

and

$$\left.\begin{aligned} C_L &= 23.4\ \text{pF} \\ L_L &= 0.4\ \text{uH} \end{aligned}\right\} \quad f_L = 52\ \text{MHz}$$

In order to balance the overall characteristic, the impedance magnitude of the low-side tank must equal the impedance of the high side tank at 70 MHz.

$$|Z_L(70\ \text{MHz})| = 216\ \text{ohms}$$

Now in determining the values of the high side tank we need to solve two simultaneous equations:

$$|Z_L| = |Z_H| = \frac{1}{j\omega_c C_H + \frac{1}{j\omega_c L_H}} \qquad \text{eqn. 10}$$

$$\omega_H = \frac{1}{\sqrt{L_H C_H}} \qquad \text{eqn. 11}$$

$$\omega_c = 2\pi\ (70\ \text{MHz})$$

Now doing a little algebra we have

$$L_H = |Z_L(f_c)|\ \frac{f^2_H - f^2_C}{2\ f^2_H\ f^2_C} \qquad \text{eqn. 12}$$

where

f_C = IF center frequency

f_H = resonant frequency of high side tank

$Z_L(f_c)$ = impedance of low side tank at IF center frequency

Now in our present example we have

f_C = 70 MHz $|Z_L(f_C)|$ = 216 ohms

f_H = 88 MHz

f_L = 52 MHz

From this it is found that

L_H = 0.22 uH

C_H = 15 pF

Now R1 and R2 will be set equal to one another and their value will be determined by how much buffering the low impedance driver requires. A good choice in this case would be 215 ohms.

Figure 7 is a calculator generated transfer function, also known as an "S" curve for the values of L, C, and R just calculated. This procedure, if necessary, can then be iterated until a linear characteristic of the desired bandwidth is obtained about the IF center frequency.

SYNCHRONOUS DETECTION

The output of an envelope detector was given by eqn. 5 in figure 3. Spectrally this output is shown in figure 8. As seen from figure 8, the ability to demodulate information signals with a large bandwidth can be enhanced by increasing w_C (the IF frequency), eliminating the AM-FM term centered at w_C, or both. The w_C term can be eliminated by employing a synchronous detector. In which case:

$$w_m \leq 2w_C - \frac{BW_2 + 2w_m}{2}$$

$$w_m \leq w_C - \frac{BW_2}{4}$$

where BW_2 has twice the deviation of the original FM signal and is centered at 2 W_C. The block diagram of a synchronous detector is shown in figure 9.

Figure 9 also shows mathematically how the synchronous detector eliminates the w_C term. The input signal $V_{FM}(t)$ is split into two paths, one containing a differentiator, the other a 90-degree phase-shift network. For convenience in Eq. 13, w_Ct + β sin w_mt is written as (τ). The differentiated and phase-shifted signals (Eqs. 14 and 15) are multiplied in Eq. 16, producing the output given in Eq. 17. The output contains two terms, the low-frequency information or baseband signal, and a term centered at $2w_C$. The baseband term is multiplied by sin $\pi/2$. Should the phase shift be more or less than 90 degrees, the output level will be reduced by the sine of the phase shift. The low pass filter then passes only the baseband signal to the following stages (Eq. 19).

More intuitively what the mathematics is saying is that we generate a quadrature reference for use as a local oscillator. The mixing process now produces an output at twice the IF frequency and at DC, but not at the original IF frequency. If the local oscillator were not synchronous, the baseband level would become proportional to the random phase angle of the LO.

The problem with the synchronous detector is that it also requires a differentiator for operation. This differentiation can be implemented through the use of single and double tuned circuits as mentioned previously. The bandwidth of the synchronous detector is therefore primarily a function of the bandwidth of the differentiator.

Bandwidths of 40 percent can be achieved by employing a synchronous detector implemented in delay-line form. Mixer and delay-line bandwidth now are the factors limiting the overall detector bandwidth. The design is aided also by the fact that a wideband delay line is much easier to construct than a wideband differentiator.

The delay-line detector is shown in Figure 10. The input signal is again split into two paths, one passing through a delay network and the other applied directly to the R port of the mixer. The delay-line introduces a constant delay of $t_o = T/4$ seconds over the bandwidth. This makes t_o equal to a phase shift of $-\pi/2$ at w_c. Therefore t_o is the group delay (Dg) of the network, as shown in Figure 11. Phase linearity over the IF bandwidth is essential, and can be achieved by using lumped delay lines or, preferably, an actual quarter wave line.

Thus the signals travel along a parallel path and are equal; however, the phase of the L port leg varies in proportion to the frequency of the modulating signal. The delay line acts as a frequency-to-phase converter. The output of a phase comparator should therefore yield the desired baseband output signal.

Looking again at figure 10, the output of the delay line (Eq. 21) is applied directly to the L port of a double-balanced mixer. The delay line, with its linear phase characteristics, shifts the phase of $V_L(t)$, which is directly proportional to the instantaneous frequency of $V_{FM}(t)$. The instantaneous frequency deviation is shown in Eq. 22, where Δwp is the peak deviation. The mixer R port receives $V_{FM}(t)$ directly; the output, $V_I(t)$, is shown in Eq. 24. When the signal is not modulated, the mixer I-port output is the product of two sinusoids in quadrature, and is therefore zero. As modulation is applied, the L and R ports deviate from quadrature by an amount equal to $\Delta w(t)t_o$, where $\Delta w(t)$ is the instantaneous deviation from the carrier frequency. It is the deviation from the quadrature that determines the sensitivity of the detector.

The output contains two components, one at baseband and the other at $2w_c$. However, the information signal, $\Delta w(t)$, is now embedded in a sine function because the mixer is acting as a phase detector with a transfer function that is sinusoidal. The output term is now:

$$\frac{K_m A^2}{2} \sin\ (\Delta w(t)\ t_o) \qquad \text{eqn. 27}$$

over some range of phase deviation:

$$\sin\ (\Delta w(t) t_o) = \Delta w(t) t_o \qquad \text{eqn. 28}$$

and:

$$v_o(t) = \frac{K_m A^2\ \Delta w(t) t_o}{2} \qquad \text{eqn. 29}$$

$$= \frac{K_m A^2}{2}\ \Delta w_p t_o \sin w_m t \qquad \text{eqn. 30}$$

where:

$\Delta w(t) = \Delta wp \sin W_m t$

Δwp = peak frequency deviation, rads/s

The linearity of the detector is now a function of the peak frequency group-delay product. Linearity is generally required to be less than 1 percent, which requires that:

$$t_o. \leq \frac{.24}{2\pi fp} \qquad fp = \text{peak frequency deviation} \qquad \text{eqn. 31}$$

This must also be consistent with the earlier constraint ($t_o = T/4$) imposed on the delay line for a given phase. Figure 12 shows a phase detector's deviation from linearity for a given phase. Since Δwp is usually set at the transmitter in accordance with some level of FM quieting, then, t_o, the group delay of the network, is left as the primary determining factor of detector sensitivity and linearity. Delay networks can be implemented in several forms. Bandpass filters provide large group delay at the expense of bandwidth. For systems where FM bandwidth is 40 percent, a delay line composed of a constant-k low-pass filter is good and can be made nicely in two sections at an IF of 70 MHz (t_o = 3.57 ns at 70 MHz).

PERFORMANCE MEASUREMENTS

The performance of wideband discriminators is generally measured with a white noise test set. In this type of test arrangement white gaussian noise is used to modulate a test modem of known linearity. The white noise level is adjusted to simulate the levels which would be encountered under live traffic busy hour conditions. This level is called nominal loading. A typical test set up is shown in Figure 13. By using a notch filter a small band that is transmitted contains no energy. Any energy which can be measured in that slot at the receiver is the result of distortions in the discriminator.

Actually, it is impossible to tell how much distortion is really due to the discriminator and how much distortion is due to the reference modulator. For purposes of simplification we can assume here that the reference modulator distortion products are negligible when compared to the discriminator.

The power output of the discriminator is now adjusted 10 dB above and below the nominal loading to observe the discriminator distortion mechanisms. As the deviation is backed off (white power level reduced) the intermodulation distortion products become less and less and the only noise in the system is the thermal noise. As the deviation is increased, the discriminator can exhibit all orders of non linearity, 2nd order, 3rd order, 5th order, etc.

The Lenkukrt 57040 delay line discriminator is used by GTE Sprint as a performance monitor in their long distance microwave radio network. The discriminator is used to drop the signal to baseband where it is passed along to an out-of-band noise slot monitor. Should the noise in that slot at anytime exceed a prescribed level, an alarm is activated and automatic protection measures are initiated.

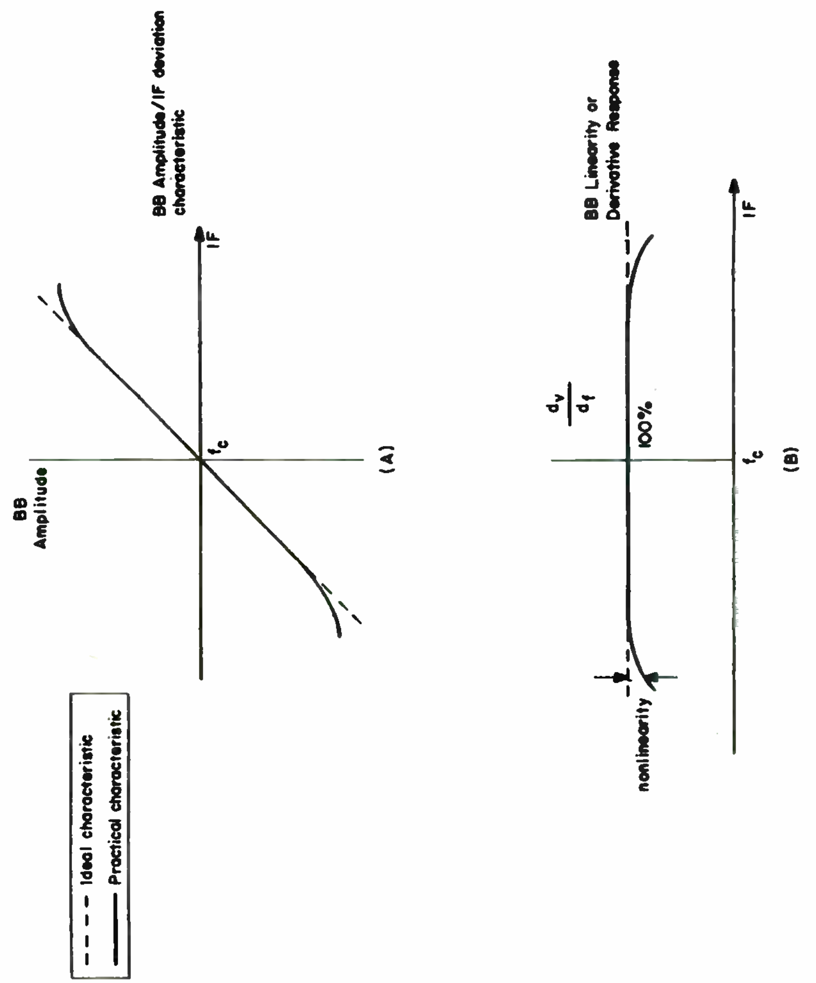


Figure 1

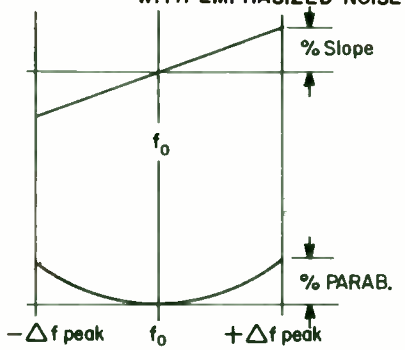
BIBLIOGRAPHY

1. Clarke-Hess, Communications Circuits Analysis and Design, Addison-Wesly 1971.

2. Lutz, J.F., "Synchronous Delay Line Detector Provides Wideband Performance", Microwaves and RF, Nov. 1982, Vol. 21, No. 12, pp. 71-79.

3. Harp, M.C., "Quantitative Data on Microwave System Degradation", Lenkurt Microwave Technical Report No. 41, Dec. 28, 1966.

NOISE DUE TO DEVIATOR/DISCRIMINATOR LINEARITY
WITH EMPHASIZED NOISE LOAD PER CCIR



RMS DEV. PER CH.	CH'N'LS	Δf peak
200 Khz	600	3.9 Mhz
200 Khz	960	4.9 Mhz
200 Khz	1200	5.5 Mhz
140 Khz	1200	3.9 Mhz
140 Khz	1800	4.8 Mhz

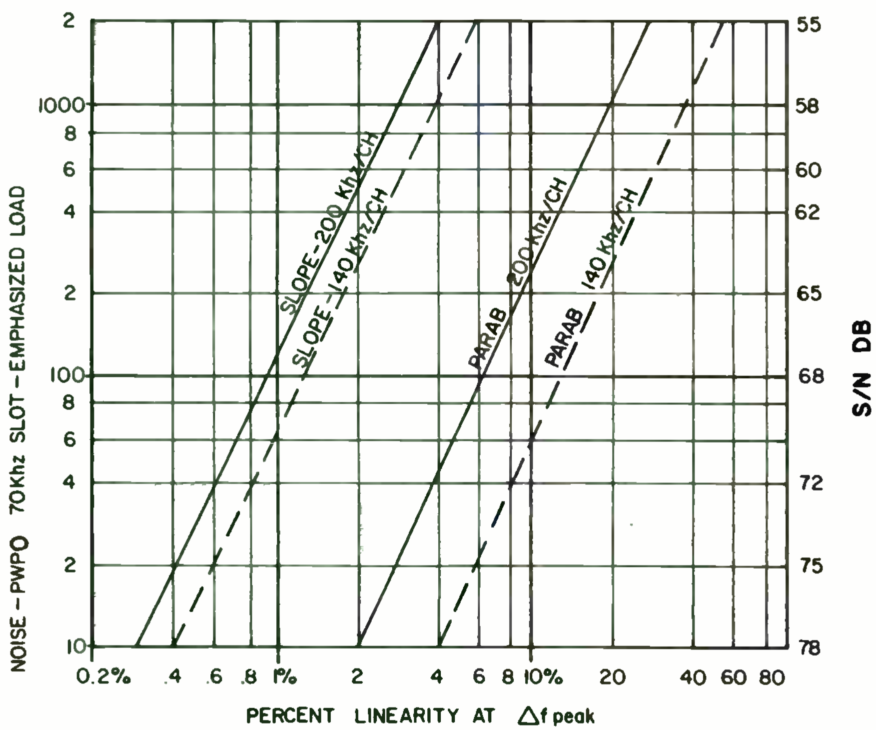
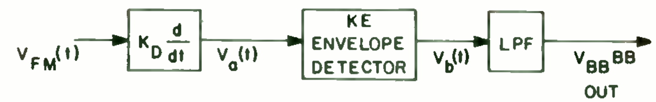


Figure 2



$$V_{FM}(t) = \cos(W_c t + B \sin Wmt)$$

AM-FM SIGNAL $$V_a(t) = \frac{dV}{dt} = K_D(W_c + B\,Wm \cos Wmt)\cos(W_c t + B \sin Wmt) \qquad \text{eqn 4}$$

$$V_b(t) = K_D K_E W_c + K_D K_E \,\Delta F \cos Wmt \qquad \text{eqn 5}$$
$$+ (K_D K_E \Delta F \cos Wmt)\cos(W_c t + B \sin Wmt)$$
+higher order AM-FM terms

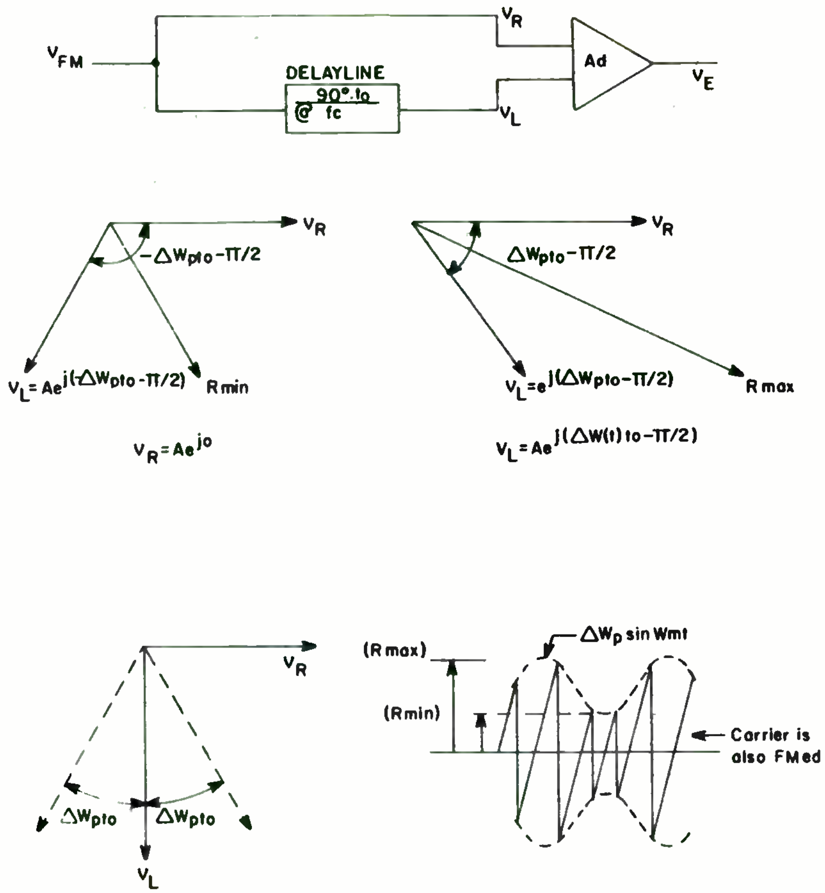
$$V_{BB}(t) = K_D K_E \Delta F \cos Wmt \qquad \text{eqn 6}$$

Where K_D = differentiation constant

K_E = envelope detector constant

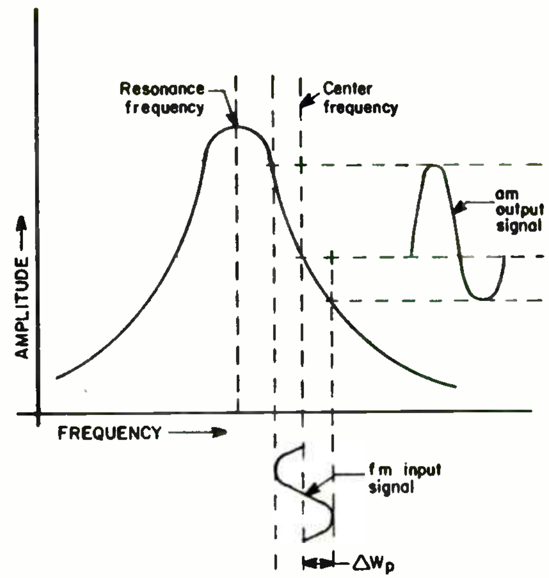
FM DETECTION BY DIFFERENTIATION AND
ENVELOPE DETECTION

Figure 3



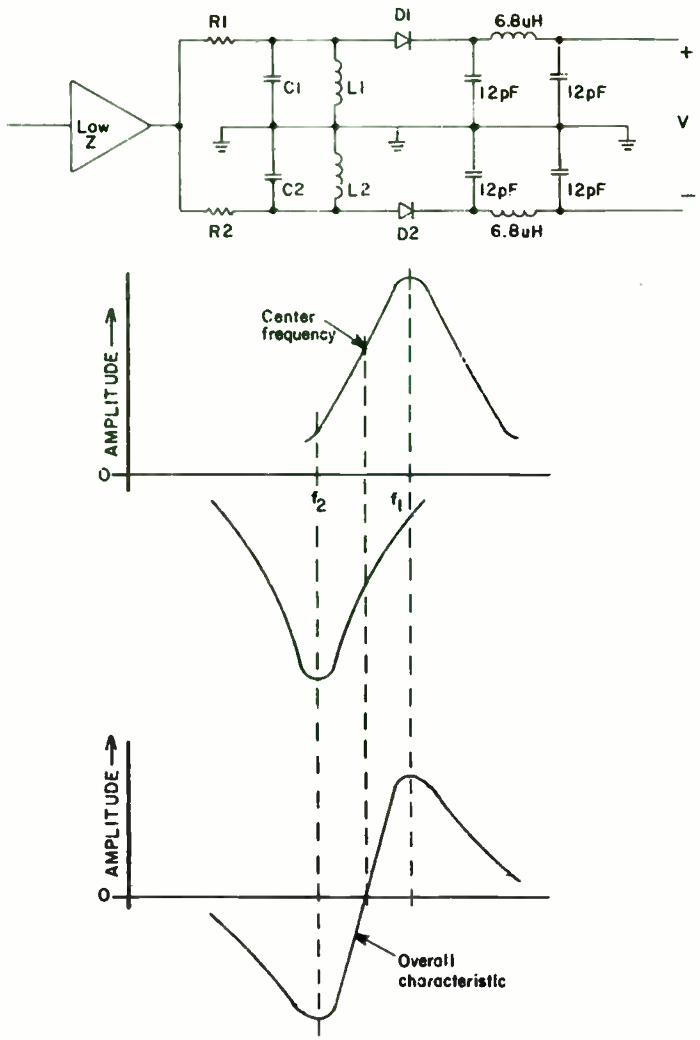
DIFFERENTIATION OF THE FM SIGNAL

Figure 4



DIFFERENTIATION BY SLOPE DETECTION

Figure 5



BALANCED SLOPE DISCRIMINATOR

Figure 6

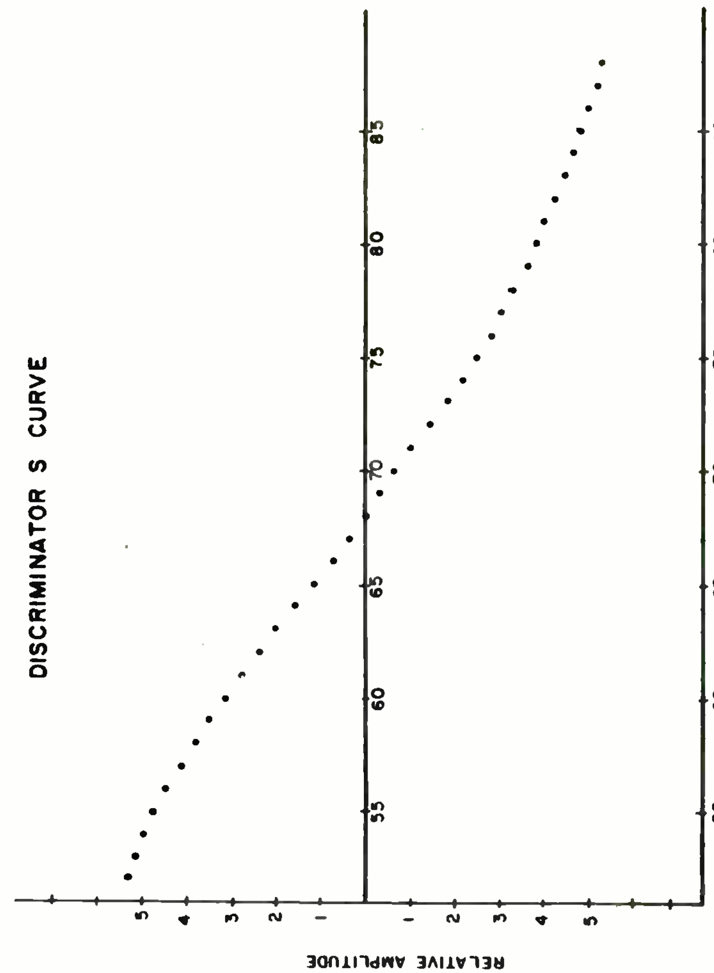
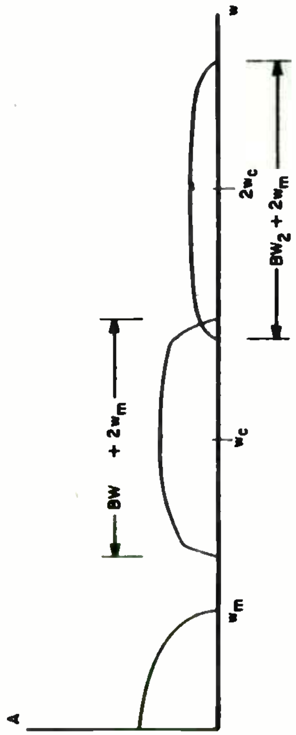


Figure 7

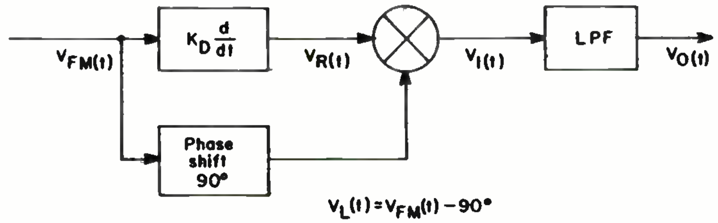


$$w_m \le w_c - \frac{BW_{FM} + 2w_m}{2}$$

$$w_m \le \frac{1}{2}\left(w_c - \frac{BW_{FM}}{2}\right)$$

ENVELOPE DETECTOR OUTPUT SPECTRUM

Figure 8



$$V_{FM}(t) = A\cos(w_c t + B\sin w_m t) = A\cos(\tau(t)) \qquad 13$$

$$V_R(t) = -AK_D(w_c + Bw_m\cos w_m t)\sin(\tau(t)) \qquad 14$$

$$V_L(t) = A\cos(\tau(t) - \pi/2 \qquad 15$$

$$V_J(t) = V_R(t) \times V_L(t) \qquad 16$$

$$V_I(t) = \frac{-A^2K_D}{2}(w_c + Bw_m\cos w_m t)\left[\sin(\pi/2) + \sin(2\tau(t) - \pi/2)\right] \qquad 17$$

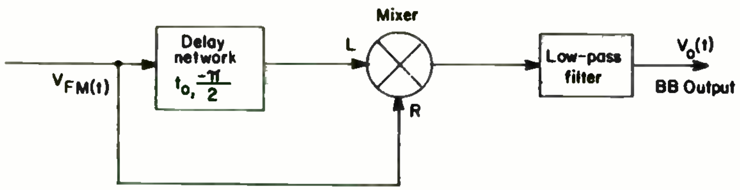
$$V_o(t) = \frac{-A^2K_Dw_c}{2} - \frac{A^2K_DBw_m}{2}\cos w_m t \qquad 18$$

$$V_o(t) = K_1 + K_2\Delta w_p\cos w_m t \qquad 19$$

$$B = \frac{\Delta w}{w_m} \qquad K_1 = \frac{-A^2K_Dw_c}{2} \qquad K_2 = \frac{-A^2K_D}{2}$$

SYNCHRONOUS DETECTION

Figure 9



$$V_{FM}(t) = A\cos(w_c t + B\sin w_m t) \tag{20}$$

$$V_L(t) = A\cos(w_c t + B\sin w_m t + \Delta w(t)\, t_o - \pi/2) \tag{21}$$

$$V_L(t) = A\sin(w_c t + B\sin w_m t + \Delta w(t)\, t_o) \tag{22}$$

$$\Delta w(t) = \Delta w_p \sin w_m t \tag{23}$$

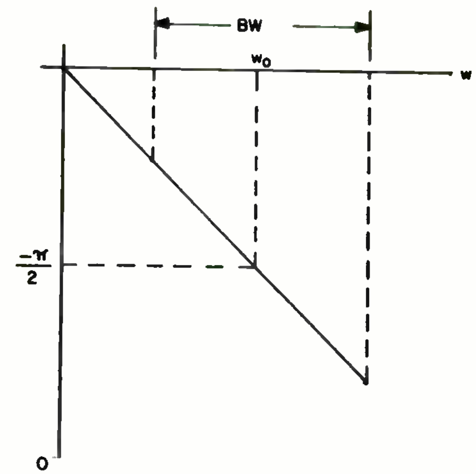
$$V_i(t) = \frac{K_m A^2}{2}\left[\sin(\Delta w(t)\, t_o) + \sin(2w_c t + 2B\sin w_m t + \Delta w(t)\, t_o)\right] \tag{24}$$

$$V_o(t) \approx \frac{K_m A^2}{2}\, \Delta w(t)\, t_o \tag{25}$$

$$V_o(t) \simeq \frac{K_m A^2}{2}\, \Delta w_p t_o \sin w_m t \tag{26}$$

DELAY LINE DETECTOR

Figure 10

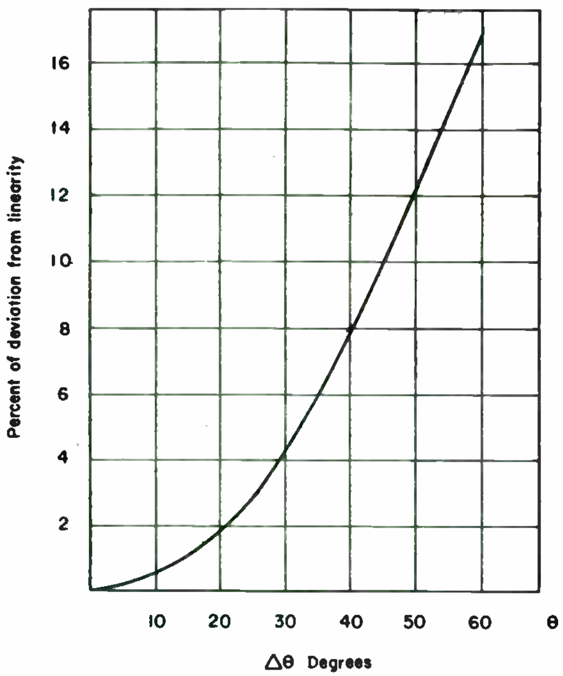


$$C_g = \frac{\Delta\theta}{\Delta w} = t_o$$

The group delay of the network is represented by t_o (in seconds) when the phase frequency characteristic is linear.

DELAY LINE CHARACTERISTIC

Figure 11



Linearity of a phase detector decreases dramatically with a change in phase.

PHASE DETECTOR LINEARITY

Figure 12

DISCRIMINATOR PERFORMANCE TESTING

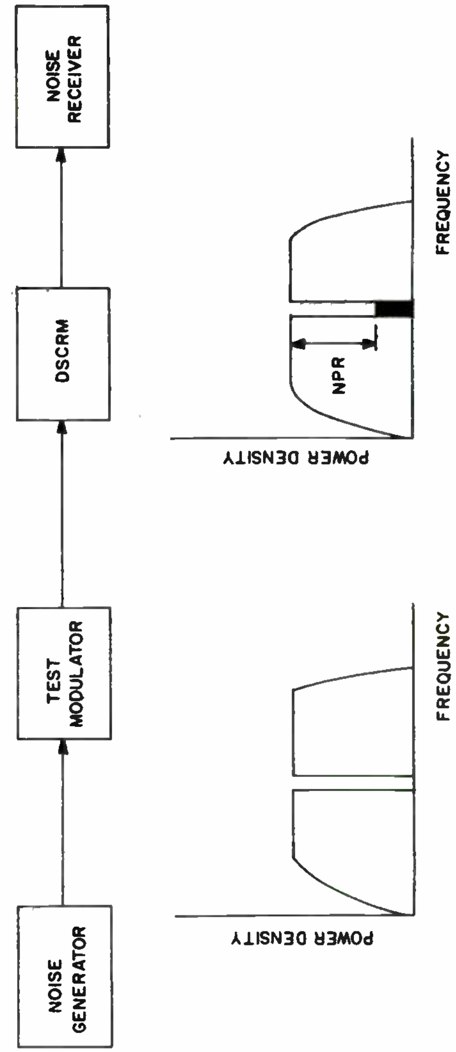


Figure 13

PAPER TITLE: Precision Glitch - Free RF Step Attenuator

Mr. George M. (Mike) Walley
Lead Engineer - Technical Staff
Harris Corp. Government Satellite Comm. Division
Box 93000 MS 1-1000
Melbourne, FL 32902
(305) 727-6344

ABSTRACT:

Precision - Glitch Free RF Step Attenuator

Glitch free step attenuators can be used to solve several difficult problems in both system design and test equipment applications. Embedding a standard step attenuator in the feedback loop of an AGC can yield a step attenuator with the accuracy of the original with no glitch during level switching. Circuit features that lead to enhanced capabilities with improved temperature compensation are described at block diagram level. Examples of applications are included by way of illustration of the power of the technique.

Precision Glitch - Free RF Step Attenuator

Application of the Technique:

The increased use of computer controlled RF communications equipment has led to the need for variable attenuators controllable in the discrete steps understood by the controlling equipment. Some of these applications do not have a requirement for no interruption in the RF signal path during level switching. These applications have the luxury of using standard step attenuators driven with RF relays or FET switches. The systems that require uninterrupted RF while switching have had to rely on continuously variable attenuators with some sort of digital to analog conversion added to yield a step response. This introduces a severe problem, which is the mapping of the digital control words to the response of the attenuator. The use of matching proms and analog linearizers has had some success but is in general an inferior solution to the problem. This is because of the failure of the linearizing schemes to adequately allow for temperature and time dependencies of the attenuation element. Once calibrated does not imply always calibrated with these schemes.

What is the possible application of this glitch free attenuator? Picture the typical satellite communications network with geographically separated ground stations. In a military environment as well as critical civilian applications it is important that communications continue at error rates above some predetermined threshold even in the presence of local rain fading conditions. Now the problem becomes one of raising the transmit power on the uplinks to a level sufficient for adequate communications performance with fading. Unfortunately all the stations are not experiencing fading at the same time. Since the satellite is primarily down link power limited and not bandwidth limited, the effect of the link overhead for intermittent fading is to limit the number of users of a single transponder based on global worst case conditions. With some means of controlling the link power as conditions change the user overhead power can be reallocated to additional users. A pool of residual power capability could be allocated on a real time or near real time basis and still meet the requirement for error rate performance in the user community with more users for a given system. It should be clear that any interruption in the signal would have a tremendous adverse effect on the usefulness of the power control system. The details of the operation of this adaptive link power control system is not the point of this paper but it does point out a prime application of a glitch free RF attenuator as an uplink power controlling element.

One possible system level use for the glitch free attenuator has been described above. What about applications of a more mundane nature? The glitch free attenuator finds application in the lab and on the assembly line when it comes time to test high performance communications equipment at the limits of it's required performance specifications. For example, When testing a high performance modem it is often necessary to characterize that modems performance at signal to noise ratios where it's required to remain locked but not to perform initial acquisition at that S/N. The typical test set consisting of step attenuators and noise source has a weakness in this type of test scenario. That weakness is in the area of transients introduced to the signal path when the test S/N is changed. These transients can cause the modem to lose lock at a signal to noise ratio they are required to operate at. These transients constitute an unfair test and degrade the reliability of the test data. With a glitch free step attenuator it is possible to sneak up on the S/N ratio so that

the modem only loses lock when it is its own fault and not due to the manner in which the test was conducted.

Description of the Circuit Topology :

The key to creating a glitch free step attenuator is controlling the transition in level with a high degree of certainty. Suppose that a conventional step attenuator is placed in the feedback loop of an AGC circuit as shown in figure 1. It can be seen that the output level is controlled as an inverse function of the attenuation setting of the step control element as the loop strives to keep the level at the input to the detector constant. While making the loop filter parameters such that the response to changes in the state of the step control element is very slow, then very little switch transient will appear at the output. The complication comes in when it is desirable for the loop to seek the new quiescent point rapidly. It becomes necessary to include a track and hold circuit to the input to the loop filter that can be commanded to "hold" just prior to changes in the state of the control element and then released after all transient behavior of the control element has abated. The improved circuit is shown in figure 2. Also shown is the addition of a differential power detector for the loop error signal generation. The circuit of figure 1 is only a programmable output level device and not a true programmable attenuator. The addition of the differential power detector allow changes in the input level to be carried through to the output independent of the action of the other leg of the loop. It is worth noting however that any fluctuations in the input power show up at the output only after having been detected and passed through the loop filter. This feedforward technique can place some restriction on the rate of change of the input power level due to the bandwidth associated with the control loop. This is not normally a severe constraint and can be partially countered by making the feedforward input of the power detector have a higher gain than the feedback input of the detector. So far this has not been required in any system where this technique has been employed. Since the feedforward action of the loop relies on the action of the loop control elements to accurately convert the detector input level change to a corresponding loop error term this places a constraint on the linearity of both the power detector and the loop control element. For most applications these constraints can be met with off the shelf components at relatively low cost.

Loop Control Element

Shown in figure 3 is the simplified schematic of a voltage variable attenuator with a linear range of approximately 30 dB. The prime control elements are a pair of Watkins-Johnson WJ-G1's. The WJ-LG1 is a linearizer circuit that is specially designed to yield a linear attenuation versus voltage when used in concert with a G1 PIN diode attenuator. The linearity of the control element is more important in this application than in the typical AGC loop due to a couple of reasons. As mentioned before, the accuracy of the feedforward of changes in the input power level is a function of the control element linearity. Also if the step attenuator is to respond to both the feedforward control and step attenuator changes in a timely manner it is desirable to have the loop bandwidth relatively wide. The actual requirements will be different for each application. It can be shown that the loop bandwidth for a first order type loop is inversely proportional to the gain constants of the detector and control element. Nonlinearities in the response of either element can

Precision Glitch - Free RF Step Attenuator

lead to undesirable changes in loop bandwidth dependent on the operating point of the loop. Hence the need for the linearizer in the control element. This circuit has no tuned elements and has a usable frequency range of 5 to 1000 MHz. The OP-AMP is used to convert the 0 to -10 volt control range of the LG1 into a 0 to +10 volt control range and provide a low impedance driver for the linearizer. If the output impedance of the driver is too high the slew rate of the control voltage is fairly dramatically limited.

Differential Power Detector

Shown in figure 4 is the simplified schematic of the differential input power detector with the track and hold switch included. If low offset OP-AMPS are used in the loop filter section and the assumption is made that the hold time is small compared to the track time there is no requirement for a hold capacitor as such. Simply disabling the input to the loop filter integrator will cause the loop filter to remain fixed at the value just prior to the onset of level switching. Once the transients on the power detector have died out the loop filter input is reenabled and the loop readjusts to the new quiescent state is an orderly manner. Careful design of the loop filter response guarantees no overshoot during the readjustment of the loop operating point and hence no glitch on the output signal level. By buying the detector diodes as lot matched pairs the D.C. offset between the two detector legs is virtually zero and the response to input power input changes is symmetrical. The diodes are D.C. grounded at the input by a relatively low impedance so that the bias current is unaffected by differing power level inputs. The use of temperature compensating voltage references for generating the bias currents for the diodes as well as the loop set point adjustment contributes to the exceptional temperature stability of this loop. The symmetrical arrangement of the control loop and the power detector make the question of age stability less of a concern. The net result is a glitch free step attenuator with the superior accuracy and repeatablity performance of readily available step attenuators with the smooth transition in output level normally associated with analog attenuators.

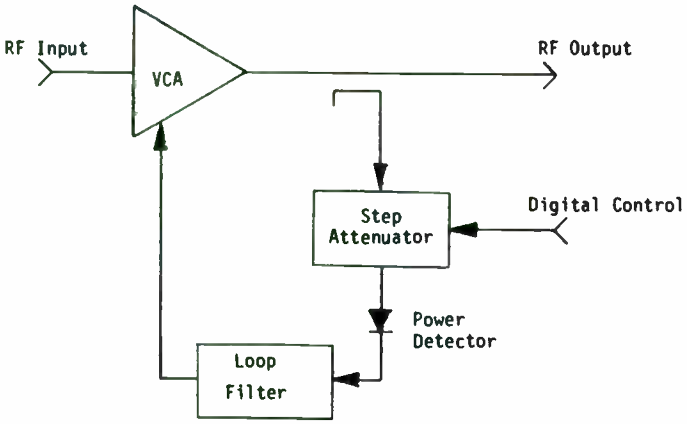


Figure 1. Basic Step Attenuator Block Diagram

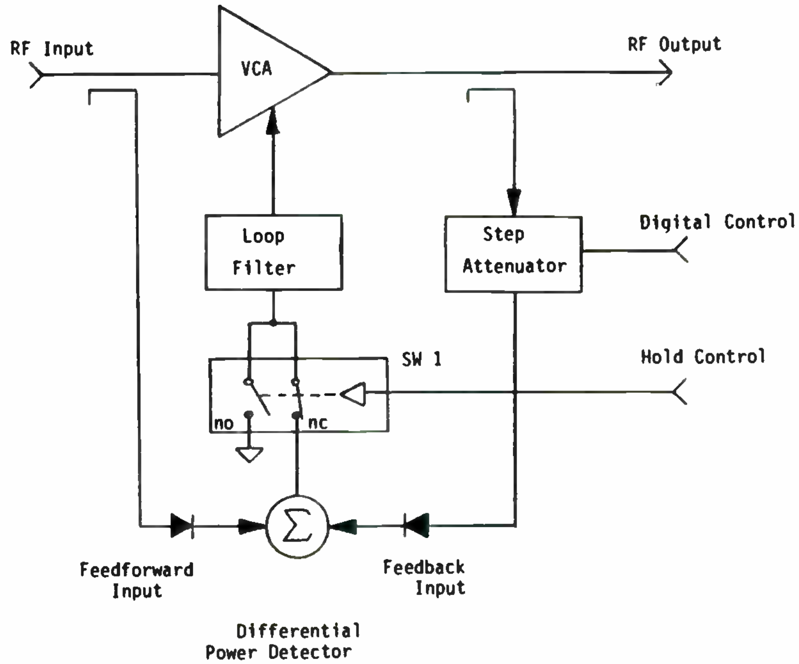


Figure 2. Improved Glitch Free Step Attenuator

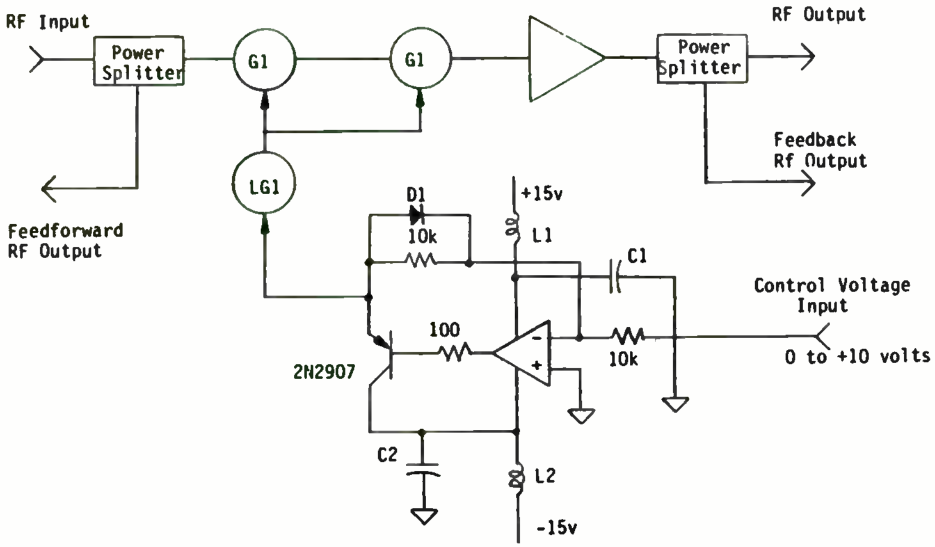


Figure 3. Simplified Schematic of Voltage Controlled Amplifier

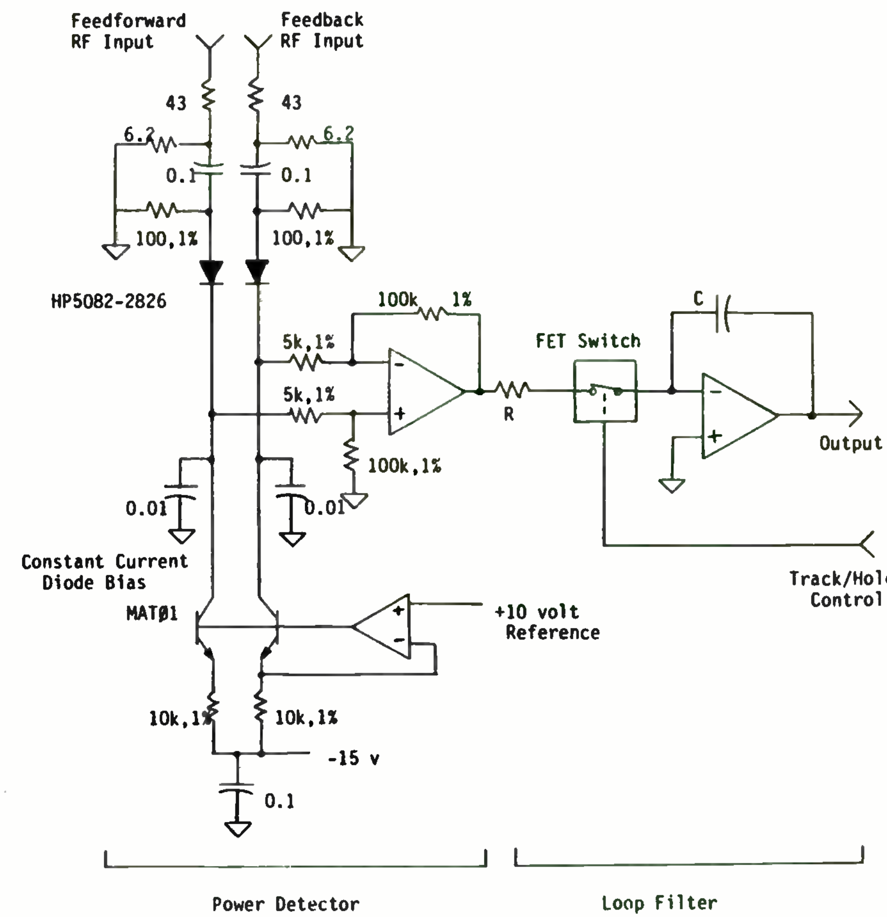


Figure 4. Simplified Schematic of Differential Power Detector

900 MHz, 220 WATT AMPLIFIER DESIGN
(BASE STATION APPLICATIONS)

K4212

By David Miller

Motorola Semiconductor
Products Sector

ABSTRACT

900MHz, 220 WATT AMPLIFIER

This high power amplifier utilizes the design advantages of the MRF898 to produce a high power, high efficiency broadband ampifier. Four Motorola 24 volt/60 watt (MRF898) transistors are paralleled using power dividing and combining techniques to produce a 220 watt, 900 MHz power amplifier.

The MRF898 is a state of the art device which incorporates double section input and output internal matching to achieve superior perfromance. Due to the internal matching of the MRF898, simple microstrip trans- mission line techniques are all that are required to externally match the device. In addition to simplifed matching, the MRF898 also exhibits good thermal performance (typically less than 1 C/watt) as well as high efficiency (typically 60% across the 800 MHz band).

The amplifier is driven by a three stage, 60 watt amplifier which utilizes an MRF898 as it's third stage. The end result is a multistage amplifier which produces a 220 Watt output for a 250 mW drive across the 850 MHz to 900 MHz band.

INTRODUCTION

The 800-960 MHz band is rapidly becoming the most active area of new product development for the UHF spectrum. Applications such as cellular telephone, paging systems, and trunck.ing/dispatch systems are all placing new demands on the design and production of RF components and circuitry. To help meet these demands, Motorola Semiconductor's RF Land Mobile Group has recently introduced several new 800 MHz RF transistors into its products portfolio.

Included in this introduction are two 24 Volt devices slated for base station applications. They are the MRF891 and the MRF898 (see figures 1 & 2). The MRF891 is a 5 watt high gain (typically 9.3dB) device packaged in a CS12 flange. The MRF898 employs a new design approach to UHF transistors by incorporating 2 sections of internal input and output matching. The use of such a matching scheme gives elevated input and output impedances, higher broadband efficiencies, and improved thermal performance over previous devices.

This paper describes the design, construction and performance of a 220 Watt, 850-900MHz, multistage amplifier which uses both the MRF891 and MRF898 as well as the MRF892. It is not the intent of this paper to demonstrate a new exotic design that will squeek a trifle more performance out of yesterday's transistors. Instead, it is intended to show the simplicity and reliablity that may be achieved using tomorrow's technology.

PROJECT OVERVIEW

The amplifier lineup consists of a 3 stage amplifier driving a high power single stage output amplifier. Though the two amplifiers have been designed to operate together, the driver amplifier also serves as a good 60 Watt stand alone amplifier. This is particularly useful in lower power applications such as cellular base stations where peak power requirements are somewhat lower.

The driver amplifier provides the majority of the gain

MRF891

NPN SILICON RF POWER TRANSISTOR

. . . DESIGNED FOR 24 VOLT UHF LARGE-SIGNAL, COMMON-EMITTER AMPLIFIER APPLICATIONS IN INDUSTRIAL AND COMMERCIAL FM EQUIPMENT OPERATING IN THE RANGE OF 800 ~ 960 MHz.

- SPECIFIED 24 VOLT, 900 MHz CHARACTERISTICS
  OUTPUT POWER = 5 WATTS
  MINIMUM GAIN = 9.0 dB
  EFFICIENCY = 50%
- GUARANTEED TO WITHSTAND 20:1 VSWR LOAD MISMATCH AT RATED OUTPUT POWER AND SUPPLY VOLTAGE
- GOLD METALLIZED, EMITTER BALLASTED FOR LONG LIFE AND RESISTANCE TO METAL MIGRATION
- SILICON NITRIDE PASSIVATED
- CS12 PACKAGE FOR TOP SURFACE MOUNTING

K4214

FIGURE 1.

FIGURE 2.

MRF898

NPN SILICON RF POWER TRANSISTOR

. . . DESIGNED FOR 24 VOLT UHF LARGE SIGNAL, COMMON-BASE AMPLIFIER APPLICATIONS IN INDUSTRIAL AND COMMERCIAL FM EQUIPMENT OPERATING IN THE RANGE OF 800 ~ 960 MHz.

- MOTOROLA ADVANCED AMPLIFIER CONCEPT PACKAGE
- SPECIFIED 24 VOLT, 900 MHz CHARACTERISTICS
  OUTPUT POWER = 60 WATTS
  MINIMUM GAIN = 6.3 dB
  EFFICIENCY = 60%
- DOUBLE INPUT/OUTPUT MATCHED FOR WIDEBAND PERFORMANCE AND SIMPLIFIED EXTERNAL MATCHING.
- GOLD METALLIZED, EMITTER BALLASTED FOR LONG LIFE AND RESISTANCE TO METAL MIGRATION.
- SILICON NITRIDE PASSIVATED

K4215

and consists of three series RF transistors (see Figure 3 for device lineup). The output amplifier is four MRF898s paralleled together to produce a high output power.

DRIVER AMPLIFIER

The driver amplifier is a 3 stage, class C, 24 Volt, 60 Watt amplifier. The amplifier specifications are shown in Figure 4 and is shown schematically in Figure 5.

Driver Amplifier Design.

Each stage of the driver amplifier is designed as an independent amplifier. Instead of designing a direct, low impedance match between stages, the input and output of each stage are transformed up to a 50 ohm intermediate impedance. Although direct matching may result in fewer components and less p.c. board real estate*, a 50 ohm intermediate impedance has several marked advantages.

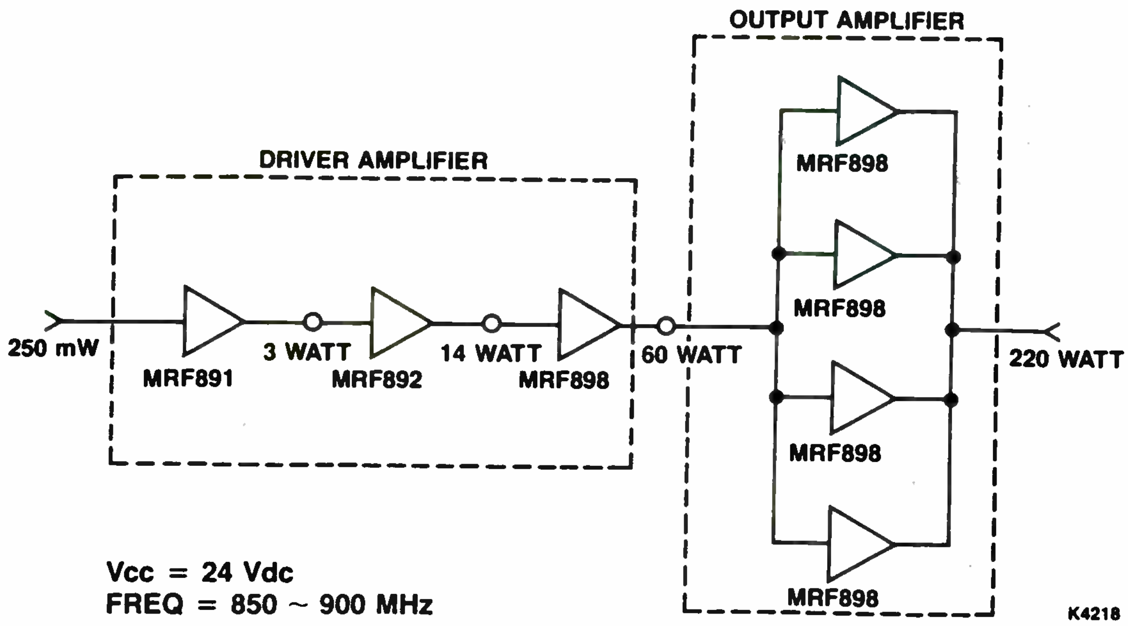
First, each stage may be independently tuned and tested. This provides the designer the necessary means of determining how well each stage is functioning. During the independent testing of each stage, collector currents should be monitored . This allows one to determine how well each stage is functioning once the stages are connected together.

---

* It should be noted that this ampifier was designed for base station applications where size is not as critical as in the case of mobile and hand-held units. In the case of these units size reduction may be the overriding design consideration.

Figure 3.

AMPLIFIER LINE UP



164

Figure 4.

TWO SEPARATE AMPLIFIERS

DRIVER AMPLIFIER (3 STAGES; MRF891, MRF892, MRF898)

Pout = 55 WATT
Gain = 23 dB (MIN)
V_{CC} = 24 Vdc
E_{ff} = 45% (MIN)
Band = 850 ~ 900 MHz

APPLICATIONS
- DRIVE OUTPUT AMPLIFIER
- CELLULAR BASE STATION (URBAN)

OUTPUT AMPLIFIER (4 PARALLEL MRF898)

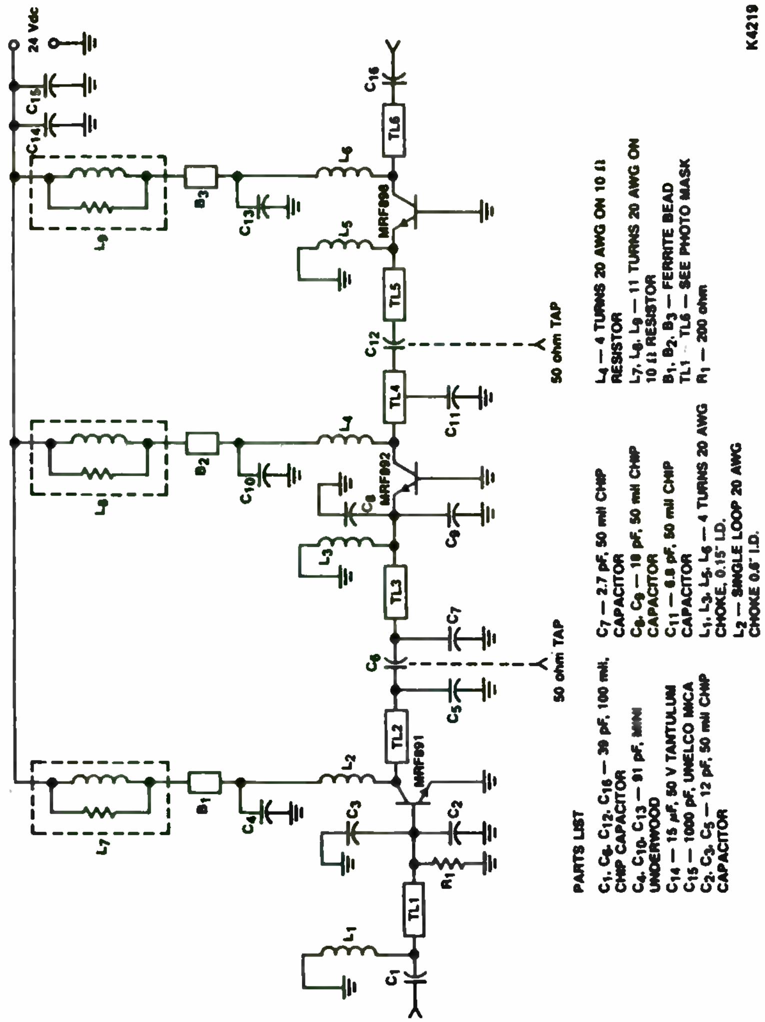
Pout = 220 WATT
Gain = 6.0 dB (MIN)
V_{CC} = 24 Vdc
E_{ff} = 55% (MIN)
Band = 850 ~ 900 MHz

APPLICATIONS
- PAGING BASE STATION
- CELLULAR BASE STATION (RURAL)

K4216

FIGURE 5.

DRIVER AMPLIFIER SCHEMATIC



For example, if the performance of the amplifier is poorer than expected, and it is noticed that the 3rd stage is drawing less collector current that it did when it was independently tested, then one may assume that there is a problem with the matching between the 2nd and 3rd stage.

The second advantage of a 50 ohm intermediate impedance is the added flexibility it provides for leveling the frequency response of the amplifier. Being able to independently tune each stage, allows the designer to adjust the frequency response of the pre-driver, or driver stage to compensate for any gain slope that may be present in the 3rd or final stage.

Low pass filter design is used throughout the circuit to achieve the desired matching. Microstrip transmission lines and chip capacitors constitute all of the matching components. Epsilam-10 (3M) p.c. board material with its high dielectric constant is used to hold the physical dimensions of the microstrip transmission lines to acceptable sizes.

It should be noted that with the elevated impedances of the MRF898 (see Figure 2 for device impedances) impedance matching can be achieved simply by using a single section of transmission line. No other matching elements are needed. This simplicity not only reduces the design effort and cost of the amplifier, but adds to its reproducibility.

DRIVER AMPLIFIER PERFORMANCE

The driver amplifier performance is shown in Figure 6.

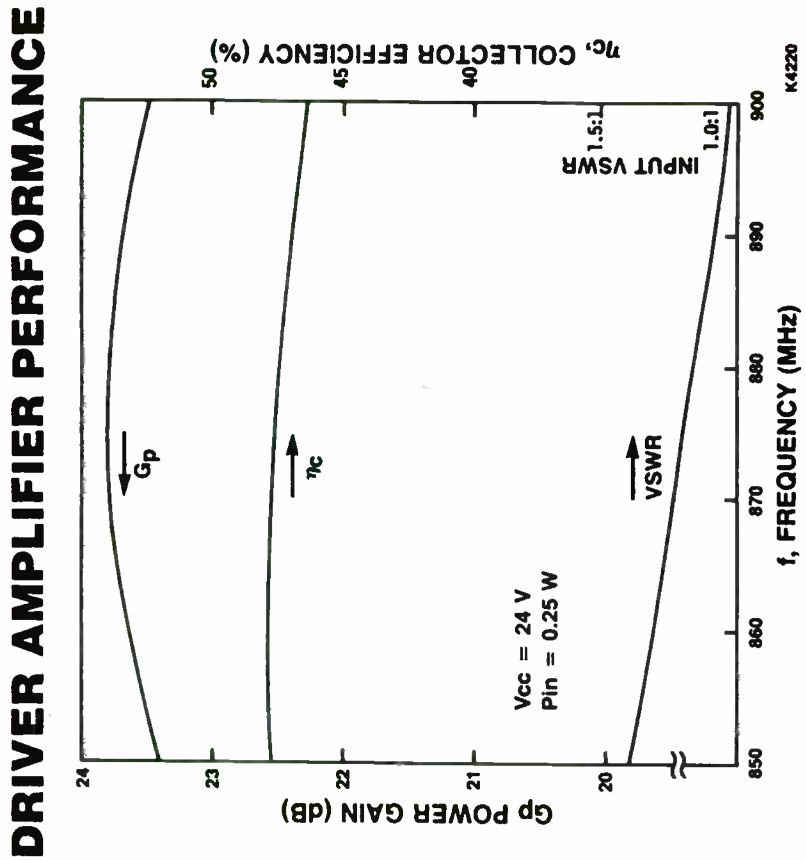


Figure 6.

The gain of the amplifier is greater than 23.4dB and is flat within + .2dB. The flat gain slope is accomplished simply by device matching. The amplifier efficiency is greater than 43% across the operating band. To realize this efficiency each individual stage must operate at greater than 60% efficiency. The figure shows the input VSWR to be less than 1.8:1.

OUTPUT AMPLIFIER

The output amplifier consists of four MRF898 transistors paralleled together using Wilkinson hybrid power splitting and combining techniques. The elevated impedances of the MRF898 allow for 100% of the circuit matching to be accomplished with microstrip transmission lines. The circuit is shown schematically in Figure 7.

OUTPUT AMPLIFIER DESIGN

The output amplifier design is separated into two general areas. The first area is the device matching and the second area is the Wilkinson hybrid matching.

A single section of transmission line is used to transform the input and output device impedance to a 12.5ohm impedance needed to match into the Wilkinson hybrid (see Figure 8). The characteristic impedance and electrical length of the matching transmission line can readily be determined using simple mathematical or Smith chart techniques. Computer analysis and/or line iterations may be used to fine tune the match. However, due to the simplicity of the match this

FIGURE 7.

OUTPUT AMPLIFIER MATCHING SCHEMATIC

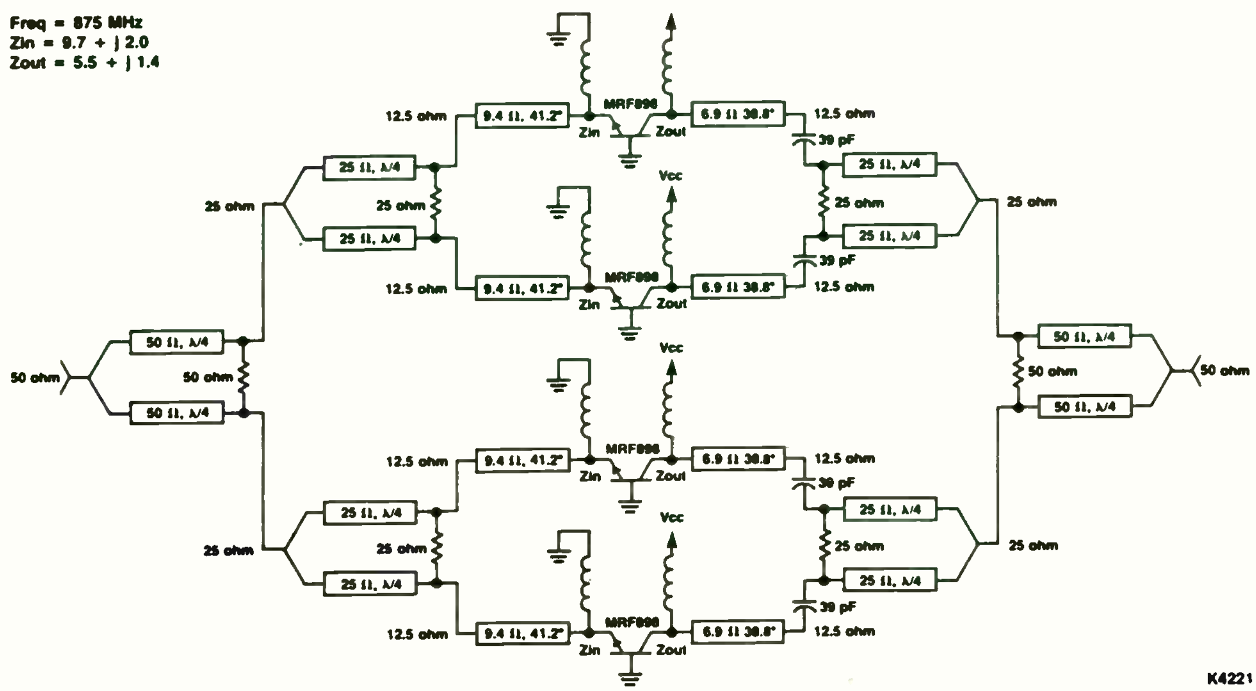
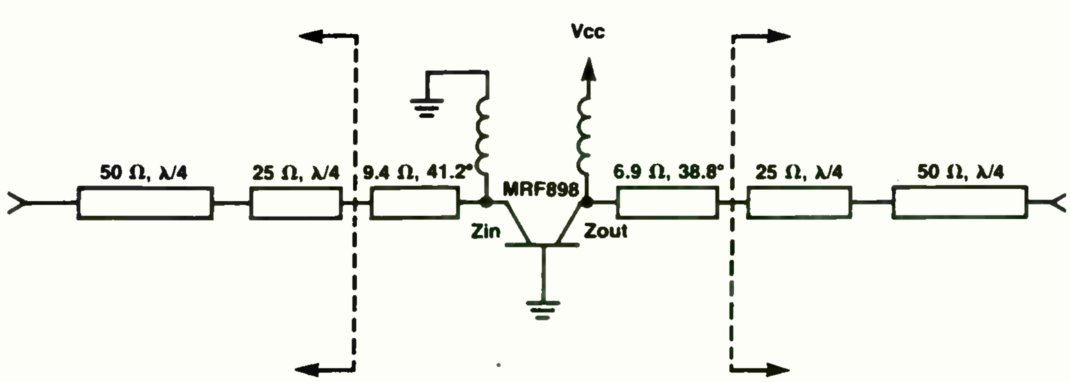


FIGURE 8.

SINGLE DEVICE TUNING



Freq = 875 MHz
Zin = 9.7 + j 2.0
Zout = 5.5 + j 1.4

K4222

will generally only provide slight improvements.

Once the device impedance has been transformed to 12.5 ohms, the remaining impedance matching is provided by a Wilkinson hybrid (see Figure 9). A multi-tier Wilkinson is used to first combine the devices together into pairs and then combine the pairs.

The first tier transforms the 12.5 ohm impedances to 50 ohms. This is accomplished with 2-25 ohm, 1/4 wavelength transmission lines, one for each device. The two 50 ohm impedances are paralelled together to yield a combined 25 ohm impedance. This 25 ohm impedance serves as the starting point for the second tier of the Wilkinson. A 25 ohm balancing resistor is shunted between the two 25 ohm transmission lines on the lower impedance 12.5 ohm end. This resistor balances any amplitude and phasing differences that may exist between the two devices.

The second tier of the Wilkinson combines the pairs of the first tier. This tier utilizes 2-50 ohm, 1/4 wavelength transmission lines to transform 25 ohms to 100 ohms. The two 100 ohm impedances are then combined in parrallel to give the final hybrid impedance of 50 ohms. Again a balancing resistor is used between the two pairs at the 25 ohm impedance end of the 50 ohm lines to smooth any amplitude and phasing differences.

The symmetry of the Wilkinson makes its design a simple and straight forward process. With both the input and output device impedances transformed to 12.5 ohms, the same Wilkinson design may be used on the input and output

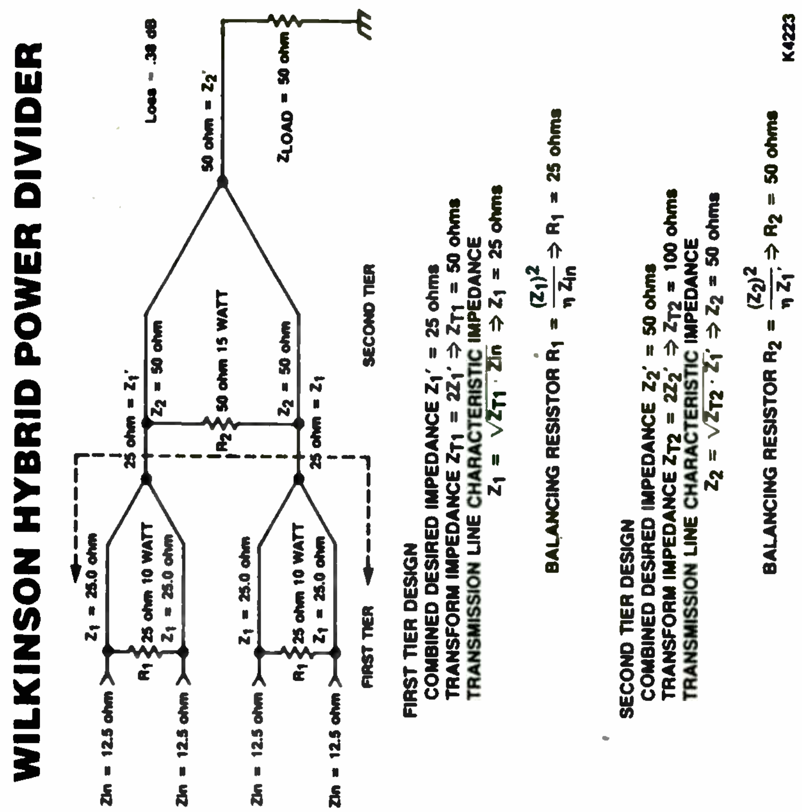


FIGURE 9.

of the amplifier.

OUTPUT AMPLIFIER CONSTRUCTION

The output amplifier is constructed on an 8" x 12.5" aluminum block attached to heat sink fins. Two types of p.c. board materials are used. The device impedance matching board is Epsilam-10. Epsilam-10 is used to hold the dimensions of the matching transmission lines to respectable sizes. Glass teflon board is used for the Wilkinson hybrid. With its lower dielectric constant the glass teflon board widens the high impedance 50 ohm transmission lines to a width capable of carrying high RF power with a minimum of loss.

OUTPUT AMPLIFIER PERFORMANCE

The performance of the output amplifier is shown in Figure 10. As shown in the figure the amplifier exhibits better than 6.4dB gain across the operating band. The efficiency is greater than 55%, with an input VSWR of less than 1.2:1.

COMPLETE AMPLIFIER LINEUP

The driver chain combined with the output amplifier results in a 4 stage high gain, high power amplifier. The performance for the complete amplifier lineup is shown in Figures 11 & 12. Figure 11 shows the amplifier to have better than 29dB gain with a +.2dB gain slope. Efficiency for the four stages is greater than 39% across the band. Figure 12 shows the linearity of the amplifier at varing

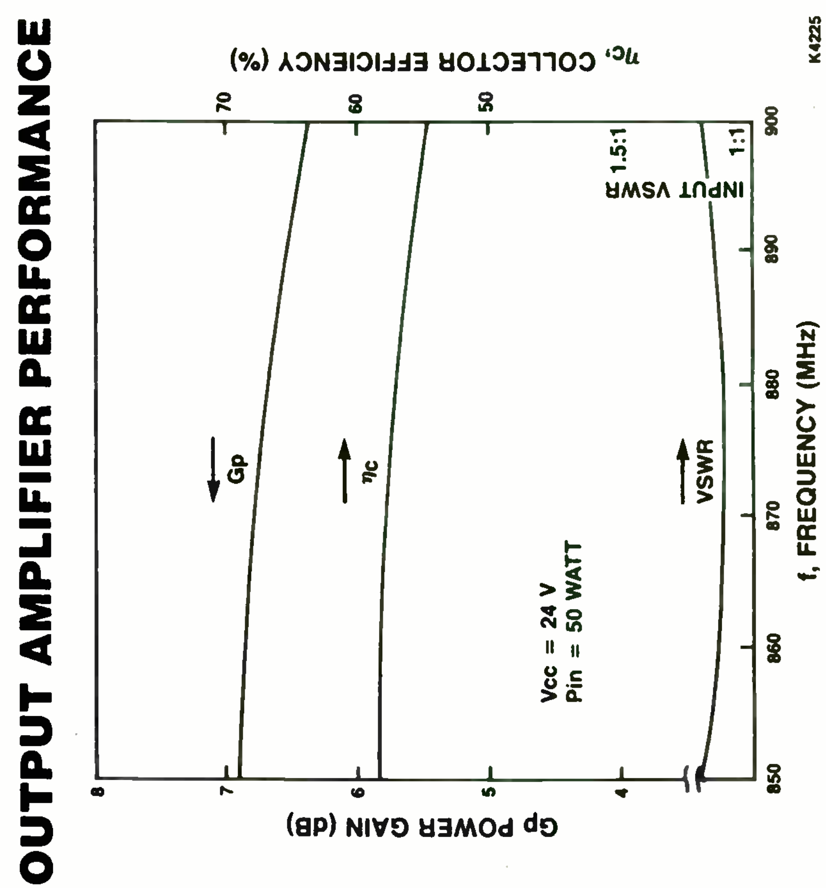


FIGURE 10.

FIGURE 11.

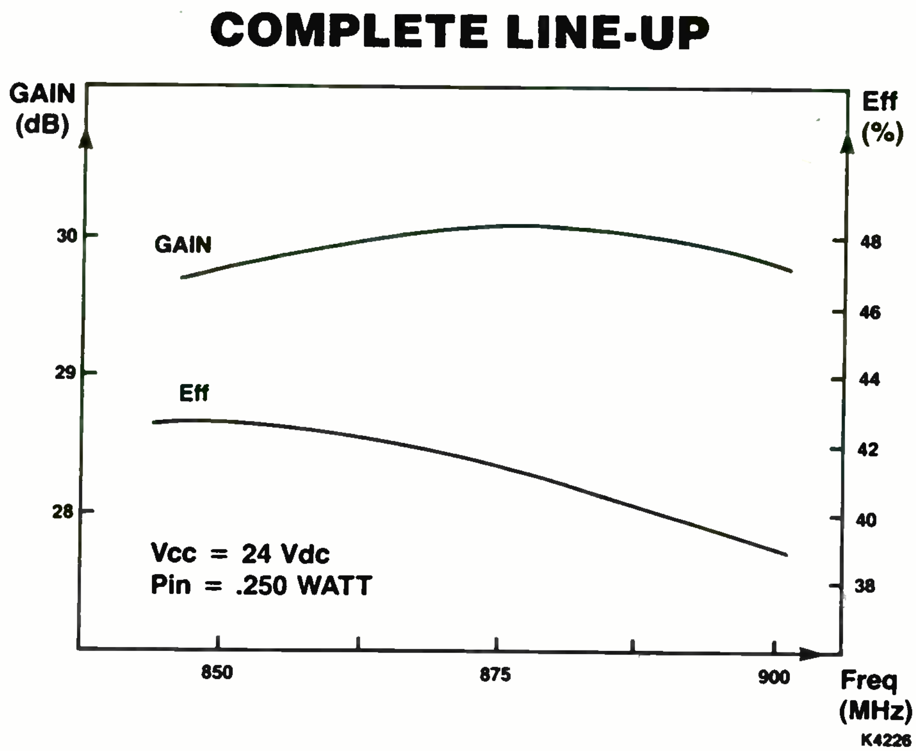
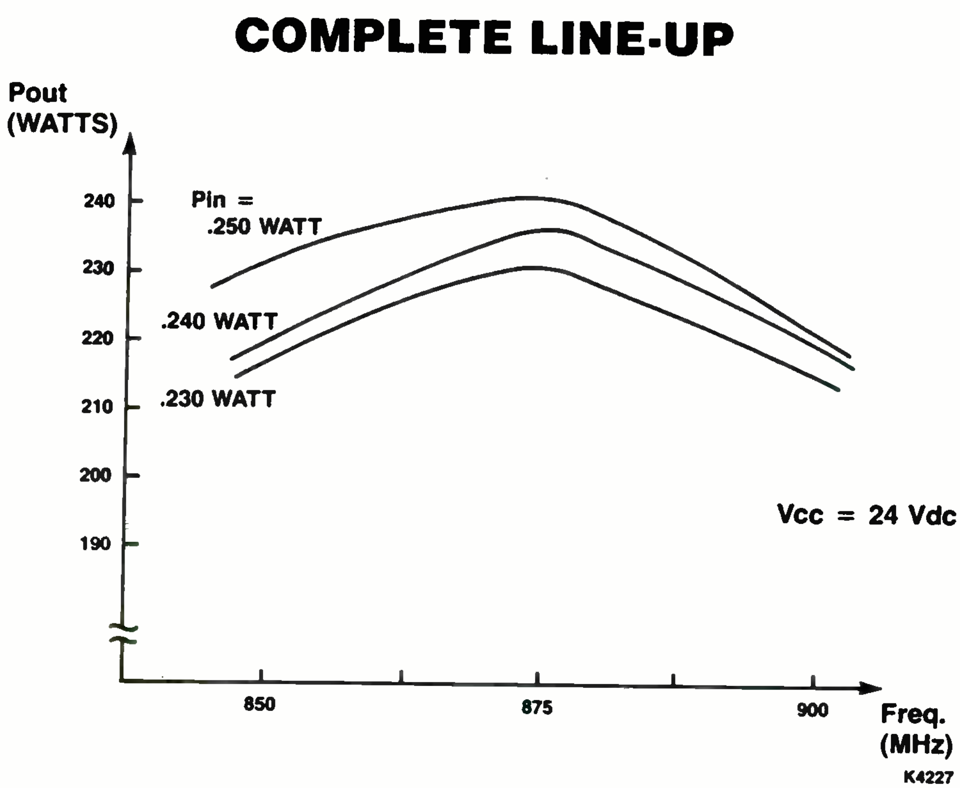


FIGURE 12.



170

drive levels.

CONCLUSION

By using tomorrow's technology today, it is possible for the designer of RF power amplifiers to design-in the simplicity and reproducibility needed to meet the demands of an ever expanding communications market. With finer line die geometries, new packaging schemes, and additional internal matching, the MRF891 and MRF898 are in fact examples of tomorrow's technology.

A HIGH STABILITY WIDEBAND FREQUENCY MODULATOR

TABLE OF CONTENTS

BY DAVID L. SOLT, P. E.

ELECTRONIC ENGINEERING CONSULTANT

43 WOODHILL DRIVE
WILLOW GROVE, PA 19090

JANUARY 1, 1985

INTRODUCTION

This paper presents the design of a high stability wideband FM modulator. This modulator was designed specifically for use in S-Band Telemetry Transmitters but, it is relevant to a wide range of communication equipment. In this application high stability is defined as less than ± 30 parts per million (ppm) frequency variation over the temperature range of -40 to +100 degrees Centigrade. Wideband is defined as frequency deviations of ± 0.03 percent from DC to 2 MHz. Different techniques for implementing the modulator are discussed, as well as the design parameters involved. A comparison of analytical results and experimental data obtained is included.

For efficient angular modulation (phase modulation or frequency modulation) there are three criteria that must be met. First, the long-term frequency stability of the carrier component portion of the modulated signal must not change in frequency during the application of modulation. Second, the frequency or phase deviation of the carrier must be directly proportional to the amplitude of the modulating signal. Third, there can be no amplitude change of the angular modulated carrier during modulation (1).

Angular modulation is generated by the variation of any reactive element or parameter on which the phase or frequency is dependent. In frequency modulation the generator which produces the carrier is generally a tuned circuit oscillator. In this case the frequency is controlled by a tuned circuit with an output frequency determined by the relationship $w^2 = 1/(L*C)$. The capacitor in this tuned circuit is often a varactor diode, a diode whose reversed biased junction capacitance depends on the amplitude of the reversed biasing voltage, creating a voltage controlled oscillator (VCO). The frequency change is proportional to the amplitude of the biasing voltage (the modulating signal). In phase modulation the modulation signal is applied after the carrier generator. A phase sensitive parameter is varied where it will not affect the frequency of the carrier.

There are two methods for generating frequency modulated carriers, Indirect FM and Direct FM. When implementing Indirect FM the carrier generator is a crystal oscillator to ensure stability. The modulating signal is applied to a variable reactance circuit, apart from the crystal oscillator, which varies the phase of the carrier. Since the derivative of phase deviation is frequency deviation, the modulation signal is integrated and then applied to the phase modulator. Indirect FM contains inherent distortion unless the deviation ratio is very small. Because of this distortion Indirect FM is only used to generate narrowband FM. To obtain the desired frequency deviation, frequency multiplication is used.

In Direct FM the transmitter output frequency is modulated directly by varying the instantaneous frequency of the carrier generator, which can be a voltage controlled oscillator (VCO). The VCO is an oscillator whose output frequency is directly proportional to the amplitude of the input signal. The principle advantage of Direct FM is the wide frequency deviations that are possible without multipliers.

The two methods of generating frequency modulation both have advantages. Indirect FM can have a crystal controlled oscillator as a carrier generator which gives it extremely high frequency stability. It has several drawbacks, including a lack of DC modulation frequency response, being narrowband, and needing multiplier stages to increase the frequency deviation to the required frequency deviation. The advantages of Direct FM are a DC modulation frequency response and wideband FM generation without frequency multipliers. Unfortunately, the frequency stability of the VCO's used in Direct FM is very poor. This paper presents a design which combines Indirect and Direct Frequency Modulation, incorporating the advantages of both, to achieve a very high stability wideband frequency modulator.

The high stability wideband frequency modulator is required to meet several stringent specifications. A flat frequency response ($\pm$1dB) from DC to 2 MHz is desired. The modulator is required to operate over the temperature range of -40 to +100 degrees Centigrade with a frequency stability of less than 30 ppm. In addition, wide frequency deviations are required (up to 0.07 percent) with excellent linearity (better than 1 percent).

Currently, a crystal controlled oscillator is the only means to attain this high stability. However, to achieve a DC modulation frequency response a voltage controllable oscillator is needed. The solution is to use a voltage controlled crystal oscillator (VCXO). This achieves both the DC response and the high stability requirements.

II. LOW FREQUENCY MODULATOR

The VCXO must have stable amplitude and frequency characteristics. The frequency is stabilized by use of a quartz crystal in a positive feedback loop of the oscillator. To achieve amplitude stability negative feedback is desirable. The oscillator chosen is a two transistor self-biasing oscillator known as the Butler oscillator. It is shown in Figure 1. The two transistors, Q1 and Q2, bias each other through R4 and receive positive feedback through the transformer coupled crystal network between their emitters. Negative feedback for amplitude stability is obtained through resistor R4.

For an oscillator to oscillate it must generate a negative real resistance that is greater than the real part of the load it is driving. The frequency of operation occurs where the oscillator's imaginary impedance sums with the load's imaginary impedance to zero. The negative resistance of an oscillator is generated by using positive feedback. This is accomplished with a low impedance path between the emitters of Q1 and Q2. By looking at the Thevenin equivalent impedance at any point in the oscillator it can be determined if a negative real impedance is generated and if so, the frequency of operation. This analysis will determine a minimum value for R4 for which the oscillator will oscillate.

For various reasons it is desirable to operate the VCXO at as high a frequency as possible. Unfortunately, quartz crystals have distortion causing anharmonic spurious responses near their

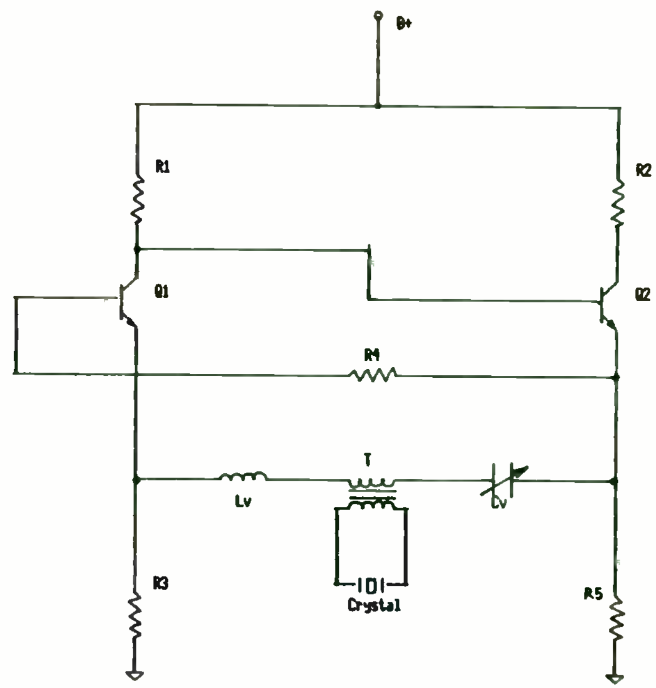


FIGURE 1. BUTLER OSCILLATOR

resonant frequency. The anharmonic spurs are due to energy trapping modes in the piezoid. These are introduced by mechanical resonances along any of the axis in flexural, torsional, shear and extensional modes of oscillation (2). Manufacturers exercise some control over where the spurs will be and their magnitude by how they shape and mount the piezoid. Oscillator designers can maximize spur free operation by keeping in mind that coupling to these undesired modes increases as the drive to the crystal increases. The spacing of these spurs from the resonant frequency is a function of several parameters. The larger the motional capacity the mounted piezoid has, the closer the spurious responses will be to resonance. Operation of the piezoid on one of its odd-order harmonic vibration modes moves these anharmonic spurious responses closer to the resonant frequency. These constraints force utilization of the highest frequency possible for fundamental mode crystals, about 25 MHz, and the smallest motional capacitance possible. Hence, the VCXO will operate at 25 MHz.

The motional capacitance of the crystal is determined by the conductive surface area deposited on the quartz blank for the electromechanical connection. It is possible to make this capacitance quite small moving the anharmonic spurs away from the resonant frequency. However, linearity constraints require the motional capacity to be as large as possible.

The design of the oscillator makes use of the high Q of the crystal to obtain frequency stability. However, this high Q causes difficulty in modulating the VCXO. Lowering the Q of the crystal will increase the oscillator's pulling range and hence, the frequency deviation of the VCXO. Two tradeoffs must be considered. Decreasing the Q, which can be done by increasing the motional capacity $(Q=R/(w*C))$, causes the anharmonic spurious responses to move closer to the resonant frequency. Decreasing the Q also causes the stability over temperature to deteriorate. A method of lowering the Q of the crystal without changing the frequencies of the anharmonic spurious responses is to transformer couple the crystal to the oscillator. The impedance of the crystal is stepped down, lowering the Q while not affecting its spurious frequencies. While this degrades the frequency stability of the oscillator, if care is exercised in the turns ratio, an acceptable amount of stability is retained.

The frequency of the oscillator is controlled by the impedance of the crystal network. This impedance can be varied by introducing a variable capacitor (Cv) in series with the transformer coupled crystal. However, this only allows pulling the VCXO frequency above the crystal resonant frequency. To pull above and below the desired frequency, the crystal would have to be cut with a series resonance below the desired frequency. A solution to this is to introduce an inductor (Lv) in series with the variable capacitor. This allows the imaginary part of the crystal network impedance to be varied both positive and negative. Consequently, the crystal can be cut to operate on its series resonant frequency. This is desired because best linearity is obtained when the crystal operates around its series resonant frequency (3). The variable capacitor is a varactor diode whose junction capacity is dependent on the reverse bias applied.

The distortion caused by the anharmonic spurs prevents modulating the VCXO above 10 KHz. Above this modulating frequency the oscillator tends to mode on one of the anharmonic spurs. A resistor can be introduced in parallel with the crystal to help reduce moding on the anharmonic spurs (4).

The frequency domain analysis of the oscillator investigates where the resistance of the oscillator circuit is negative and at what frequency does the imaginary impedance of the oscillator go to zero. The oscillator can also be analyzed using network equations (5) or transfer gain analysis. The crystal is modeled as a series motional capacitance (Cm), a series motional inductance (Lm), and a series resistance (Rs). This approximation is valid around the resonant frequency. In shunt with this series impedance is the case capacitance (Co). The transformer is modeled as a leakage-free transformer with a coupling coefficient of unity. The primary inductance is L1, the secondary inductance is L2, M is the mutual inductance, and n is the turns ratio. This is a valid assumption when working with small signals and the two transformer windings are tightly wound, one on top of the other, around a toroid with a known permeability. The small signal common-emitter transistor model is used in analyzing the oscillator. It consists of a base resistance, Rb, and a dependent current source B*Ib. Here the small signal current gain is B and Ib is the transistor base current. The varactor diode mathematical model used is

$$Cv = C_o / \{(1 + (V/r))^N\}$$

where Cv is the capacity of the reversed biased junction, C_o is the capacity of the junction at zero volts bias. V is the reverse bias applied, r is a constant that defines the slope of the capacitance versus voltage curve, and N is a constant, known as the law of the diode, that also defines the slope of the curve.

In the frequency domain analysis the crystal impedance is defined, then the transformer coupled crystal impedance is derived, and finally the other circuit components are added to obtain the impedance for the entire crystal network. This impedance is used to determine at what frequency oscillation will occur. The following equation gives the feedback loop impedance:

$$Z = (1 / j * w * Cv) + j * w * Lv$$
$$+\{n^2*(Rs+j*Xc) / \{Rs*(n^2-w^2*L1*Co)+j*(n^2*Xc+w*L1*(1-w*Co*Xc))\}\}$$

where Xc = w*Lm - (1/w*Cm). The transformer inductance is designed to resonate with the crystal case capacitance Co and other parasitic shunt capacitances in the circuit. This improves the linearity of the oscillator. By setting the imaginary part of the impedance Z equal to zero and solving for w the frequency of oscillation is obtained.

The occurrence of negative resistance can be determined by simultaneously solving the nodal current equations of the AC equivalent circuit of the oscillator. When negative resistance is generated, the circuit becomes an oscillator (6). It is possible to determine the value of R4 once the other circuit components have been chosen. Since R4 contributes negative feedback to the oscillator, the value determined is a minimum resistance for which oscillation will occur.

III. HIGH FREQUENCY MODULATOR

To obtain a modulation response above 10 KHz Indirect FM using a variable reactance phase modulator is implemented. Due to the inherent distortion of Indirect FM this is a narrowband modulator. A low capacitance varactor diode is used to vary the reactance of the phase modulator. Because of the low capacitance of the varactor diode the phase modulator will not cause any loading of the VCXO. This is an important consideration in attaining high frequency stability. A topology that utilizes a series resonant varactor and inductor was chosen. Before the carrier is applied to the tuned circuit it is divided and phase shifted. One part of the carrier is phase shifted 180 degrees and is applied to one end of the series tuned circuit. The second part of the carrier is applied to the other end of the tuned circuit. This topology increases the frequency deviation sensitivity of the modulator. The frequency domain analysis investigates the arctangent of the imaginary part of the output signal divided by the real part of the output signal. This arctangent function gives the phase of the modulated carrier. The above varactor model is used to calculate the carrier phase versus the varactor bias voltage. The phase modulator schematic is shown in Figure 2. Transistors Q1 and Q2 are used to generate the in-phase and out-of-phase signals. The modulator's output voltage (Vo) and output voltage phase (p_d) are as follows:

$$Vo = Vin * \{(1-w^2*L*C - j*w*R*C) / (1-w^2*L*C + j*w*R*C)\}$$

$$p_d = -2 * \arctan\{w*R*C / (1 - w^2*L*C)\}$$

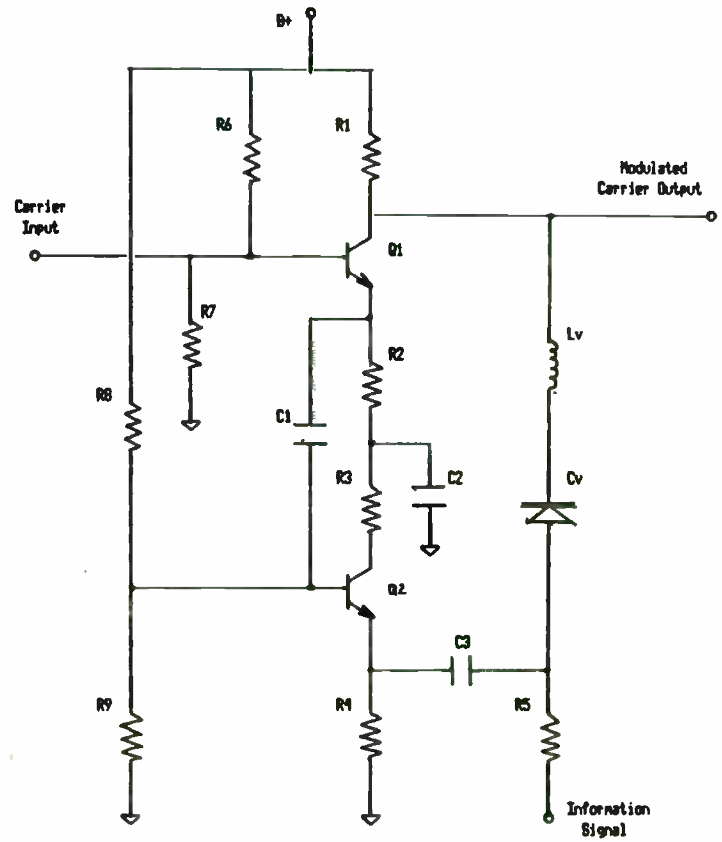


FIGURE 2. VARIABLE REACTANCE PHASE MODULATOR

By substituting the mathematical model for the varactor for the tuned circuit capacitance a plot of the varactor bias voltage versus carrier phase can be generated. It can be shown that the maximum phase deviation sensitivity to voltage change occurs when $1/(L*C)=w^2$. That is to say, when the inductor L and the varactor capacitance C are at series resonance. Because limiter stages follow the modulator, the amplitude characteristics of the modulator are not transferred to the output and hence, are not important.

One last issue must be addressed for the phase modulator and that is the circuitry required to integrate the input modulating signal so as to have frequency modulation instead of phase modulation of the carrier. This can be accomplished by introducing a low pass filter function into the modulation input line. This is implemented with the components R5 and C3 as a low pass filter with one pole. The low frequency cut-off for this filter is 1/(R5*C3). This low frequency cutoff is the lowest frequency that the modulator can be used as a frequency modulator.

IV. SIMULTANEOUS OPERATION OF BOTH MODULATORS

If the frequency deviation sensitivities of the low frequency and high frequency modulators can be matched, a high stability narrowband modulator with a modulation frequency response from DC to several Megahertz can be obtained. To implement simultaneous operation a frequency domain analysis of the two modulators operating together was performed. Figure 3 shows the topological layout. The modulation signal Vm is divided and sent to both modulators. It is low-pass filtered before being applied to either modulator. The output of the VCXO low-pass filter is

$$e_a(s) = Vm/(1 + s / w_{col})$$

where w_{col} is the 3dB cut-off point of the filter. This is a simple RC filter, series resistor with shunt capacitor. The VCXO is a device where the output frequency is proportional to the input voltage. Hence,

$$e_b(s) = Kv * Vm / (1 + s/w_{col}) = w_d / (1 + s/w_{col})$$

where Kv is the VCXO constant in frequency change per volt. This constant multiplied by the input voltage gives the frequency deviation of the VCXO (w_d). The phase of the output of the VCXO is the integral of the frequency, which in the frequency domain is the same as dividing by s. Therefore, the phase (p) becomes

$$p = w_d / \{s * (1 + s / w_{col})\}$$

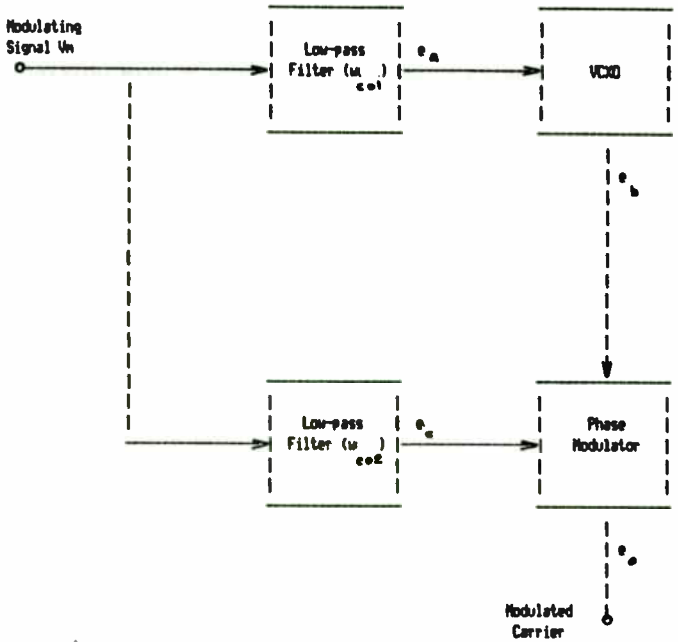


FIGURE 3. SIMULTANEOUS OPERATION OF BOTH MODULATORS

The modulation signal also goes through a low-pass filter before being applied to the phase modulator.

$$e_c(s) = Vm / (1 + s/w_{co2})$$

Here w_{co2} is the cut-off frequency of the phase modulator's low-pass filter. This filter, because of its 1/s characteristic, performs the integrating function needed to obtain frequency modulation from the phase modulator. The phase modulator output voltage's phase varies directly with the amplitude of the modulating voltage. Therefore, the phase change of the carrier due to the input $e_c(s)$ to the phase modulator is

$$Kp * Vin / (1 + s / w_{co2} = p_d / (1 + s / w_{co2})$$

where Kp is the phase modulator's constant in radians per volt. The product Kp * Vin is equal to the phase deviation (p_d) in radians.

The output signal has a phase deviation equal to the sum of the carrier phase deviation caused by the VCXO and the phase deviation introduced by the phase modulator. This phase can be represented as

$$P_{od} = w_d / \{s * (1 + s / w_{co1})\} + p_d / (1 + s / w_{co2})$$

The frequency deviation of the output signal is the derivative of its phase deviation. In the frequency domain this is accomplished by multiplying the output phase deviation by s. Hence,

$$w_{od} = w_d / (1 + s / w_{co1}) + s * p_d / (1 + s / w_{co2})$$

The plotted functions for these two terms are shown in Figure 4a. It is desirous to have as the output a constant deviation insensitive to the modulating signal's frequency. This can be accomplished by setting $w_{co1} = w_{co2}$ and $w_d = w_{co2} * p_d = w_{co1} * p_d$.

Physically, this involves setting the two low-pass filters to the same 3dB cut-off frequency and setting the VCXO and phase modulator constants so that the output frequency deviation contribution of each modulator is exactly the same. Substituting these values into the equation for the output signal's frequency deviation gives,

$$w_{od} = w_d / \ (1 + s \ / \ w_{col} + s * (w_d/w_{col}) \ / \ (1 + s \ / \ w_{col}) = w_d$$

The individual responses for the modulators and the composite response are plotted in Figure 4b. It can be seen that there is no phase distortion in crossing over from one modulator to the other. The output frequency deviation is dependent only on the amplitude of the modulating signal and the VCXO constant. This holds true as long as the two modulator's roll-off slopes are equal. This is an important result as it implies linear time delay through the modulator and hence no distortion of a wideband modulating signal, such as pulse train.

V. WIDEBAND FM GENERATION

The previous sections have discussed generation of narrow-band FM. This narrowband FM is transformed to wideband FM by frequency multiplication of the narrowband signal. When the signal is multiplied, the carrier frequency as well as the frequency deviation is multiplied. The frequency multiplication factor is determined by two parameters. They are, the final carrier frequency and the ratio of the required wideband frequency deviation to the narrowband frequency deviation obtained from the modulator.

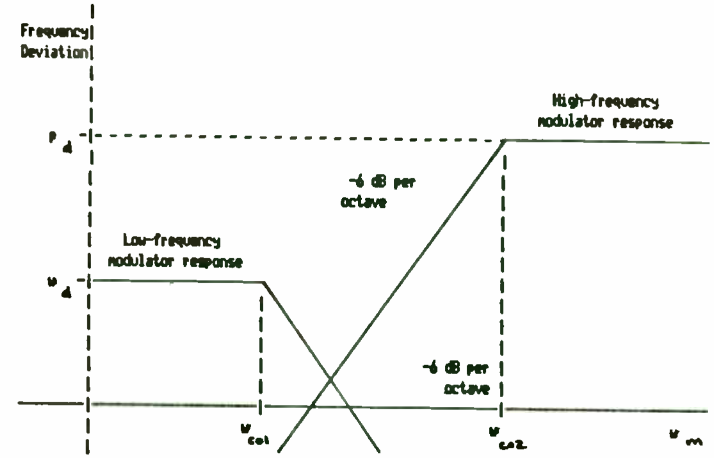


FIGURE 4A. INDIVIDUAL FREQUENCY DEVIATIONS VERSUS MODULATING FREQUENCY

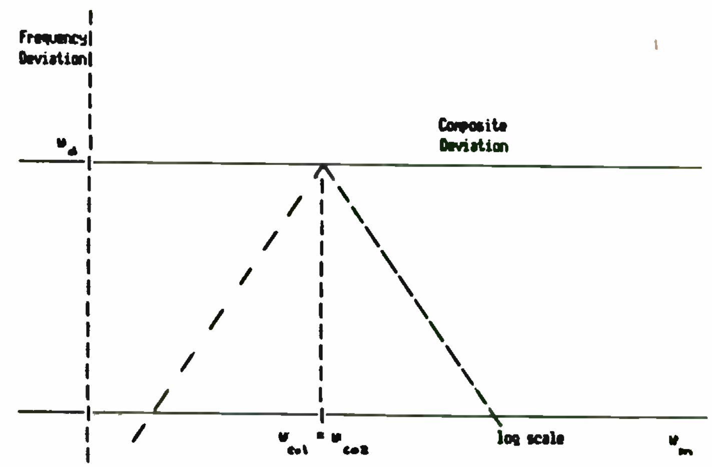


FIGURE 4B. MATCHED FREQUENCY DEVIATIONS VERSUS MODULATION FREQUENCY

The multiplication method implemented in the high stability wideband modulator is the single stage transistor multiplier. The transistor stage is a common emitter stage biased class-A with a large capacitance tied from emitter to ground. This capacitance keeps the emitter voltage constant, forcing any decrease in the base voltage to put the transistor in cutoff. Essentially, this is a class-C grounded emitter stage. The current in the collector circuit is a pulse with high harmonic content which resonates a tuned circuit on the collector. Multiplication up to a factor of four can be obtained with good results. Several stages were operated in series to provide frequency multiplication by a factor of 96. This produces the required S-Band (2.3 GHz) carrier and wideband frequency deviation.

VI. LINEARITY ANALYSIS

The linearity of the two modulators is an important parameter in producing undistorted waveforms. The linearity of the VCXO can be analyzed from the previous derivations on the oscillator circuit. Using the mathematical model for the varactor capacitance in the VCXO's impedance equation gives varactor voltage versus carrier frequency deviation. This can be programmed on a computer and simulations run to observe the effect varying individual component values has on linearity.

It can be shown that the frequency deviation (df) of the VCXO is given by the following equation:

$$df = f * \{ ((1 + Cr * (1 + V/(b+r))^N / (1 + Cr))^{1/2} - 1 \}$$

where f is the center frequency of the oscillator, Cr is the ratio of the crystal motional capacitance to the varactor capacitance at its rest bias point b, V is the incremental change in the varactor bias, N is the law of the diode, and r is the varactor diode shaping parameter.

Computer simulation shows that the larger the ratio of the motional capacitance to the varactor capacitance is, the more linear the frequency change is with voltage. This capacitance ratio can be controlled in two ways. The crystal motional capacity can be increased or the nominal varactor capacitance can be decreased. Maximum linearity occurs if the transformer inductance resonates with the case capacitance of the crystal and the varactor, Cv, and the inductor in series with it, Lv, resonates at the crystal series resonance frequency.

It can be shown that the phase deviation, p_d, of the phase modulator as a function of varactor voltage is:

$$p_d = -2 * \arctan \{ R / (-w*L + (1 + V/r)^N / (w*C)) \}$$

The linearity of the phase modulator with respect to voltage is maximized when the series resonant frequency of the varactor and inductor is at the frequency of the VCXO. By taking the first derivative of p_d it can be seen that the phase modulator has the maximum sensitivity when the varactor and inductor of the phase modulator resonate. This allows minimum change of the varactor voltage to obtain the necessary deviation. The varactor approximates a linear device only when used over a small portion of its operating curve. The modulator is designed to operate in this region (around 4 volts) giving the best linearity.

VII. TEMPERATURE STABILITY

Due to the large temperature range and the transformer coupling of the crystal, the VCXO needs a temperature compensation circuit. Rather than use a reactive component with a temperature coefficient in the VCXO frequency determining feedback loop, it was decided to inject a temperature varying voltage onto the varactor biasing voltage. This allows more flexibility to match a wide range of temperature compensation coefficients. Two approaches were considered. One uses the -2.2 mV per degree Centigrade temperature coefficient inherent in a forward biased diode junction and the second approach uses a thermistor. Both circuits worked acceptably.

VIII. COMPARISON OF ANALYTICAL AND EXPERIMENTAL DATA

The high stability wideband modulator was constructed and its operation evaluated. The main areas of interest in the comparison of experimental to analytical data are; simultaneous operation of the low-frequency and high-frequency modulators, wideband operation, linearity evaluation, and temperature stability.

The low-frequency modulator modulates exactly as predicted up to 12 KHz. Above this frequency the response was not flat enough to meet the 1 dB specification. The high-frequency modulator works well from 200 Hz to 2 MHz. However, below 5 KHz, it does not have a high enough phase constant to be used as a frequency modulator. The low-pass filters were added on the inputs of both modulators with the cutoff frequency set to approximately 7 KHz. The cutoff frequencies were matched within 0.3 dB by sweeping the applied modulating frequency across the crossover point and adjusting the high frequency modulator's cutoff frequency to match the low frequency modulator's cutoff frequency. This can be observed by receiving the modulated carrier on a calibrated receiver. Frequency response was flat across the band with rolloff beginning above 1 MHz. This is due to the high-Q in the 90 MHz multiplier resonators and is not a function of the high frequency modulator, which is relatively flat up to 2 MHz. The modulation frequency response was consistently set to ± 0.5 dB.

The 25 MHz modulator is capable of ± 0.03 percent deviation. This provided the required 750 KHz deviation required for wideband operation at S-Band (2.3 GHz). The multiplier's high-Q filters and resonators have no amplitude modulation effects on this swing because of the narrow percentage bandwidth and the limiting effects of the multiplier circuits.

The linearity of the modulator begins to deteriorate above ± 0.03 percent. This is due to the nonlinear characteristics of the varactor diodes. There are two important parameters that affect linearity of the low frequency modulator. They are the ratio of the crystal motional capacitance to the varactor diode rest bias capacitance and the law of the varactor diodes. Because of the Q desired, the varactor capacitance was fixed in the 20 pF to 60 pF range. Therefore, in order to increase the crystal motional capacitance (Cm) to varactor capacitance (Cv) ratio, the crystal motional capacity must be increased. The limiting factor on how large the motional capacity can be is the presence of anharmonious spurs. A value of 0.008 was chosen as a compromise between linearity and spurious response criteria. The law of the varactors is harder to manipulate. It can be shown that as the law approaches unity, the non-linearity disappears. The closest device found has a law of 0.826. Linearity with this device was within the one percent requirement. One other factor affects the linearity, the transformer winding ratio. This is chosen mostly by the requirements for frequency stability. For best linearity the crystal wants to be loosely coupled to the oscillator circuit, and for best stability the crystal needs to be coupled directly to the feedback loop of the oscillator. By experimentation it was found that the maximum turns ratio possible, still meeting the 30 ppm frequency stability requirement, is four.

The phase modulator contribution to non-linearity is marginal due to the high sensitivity of the modulator. Because of this high sensitivity the voltage swings on the phase modulator's varactor diode are small and the varactor non-linearity is not evident. Linearity on the order of 0.1 percent is common.

It was determined that the temperature drift of both modulator's modulating parameters are due mostly to the varactor diode's +200 ppm temperature coefficent. In addition, the low frequency modulator has the added effect of the temperature compensating voltage changing the rest bias point of the varactor diodes. While the response does vary over temperature, it is within the required ±1 dB flatness for both modulators.

The modulation set-up of this modulator is very simple. Once the initial component values are chosen and their tolerances set there are no other adjustments. The greatest advantage of this modulator over other implementations is that there is no assembly line set-up adjustments for linearity or distortion.

The VCXO is required to operate over the temperature range of -40 to +100 degrees with a frequency stability of ±30 ppm percent. The temperature instability is due to several factors, all of which could not be identified. The temperature drift is not consistant and individual temperature runs on each modulator are required. The data is then entered into a computer program which solves the resultant compensating circuit component values. Both the diode and the thermistor compensating circuits work as predicted. Component value tolerances of two percent are acceptable. A final temperature run is performed to verify temperature stability. Using this procedure, frequency stability is maintained better than the required ±30 ppm.

XI. SUMMARY

The design of a high stability wideband frequency modulator has been presented. It provides high stability by psuedo-locking the carrier generator oscillator to a crystal. The flat frequency response from DC to 1 MHz is obtained by simultaneously operating a frequency modulator and a phase modulator. Wideband modulation is obtained by the use of frequency multipliers. This design solution is not unique, but it has some significant advantages over other implementations. Other designs require substantial trained technician time to make adjustments for acceptable linearity and distortion. With this design for a high stability wideband frequency modulator there are no linearity and distortion adjustments to be made. Setting frequency stability requires an extra data gathering temperature run than some designs. However, once set, the temperature stability is good. One important consideration in circuit design is consistent and repeatable performance. Experimental results have revealed that this design performs consistently with minimal trained technician time.

REFERENCES

1. Panter, P.F.: Modulation, Noise, and Spectral Analysis; McGraw Hill Book Company, New York, pp. 381, 1965.

2. Magnavox Company: Quartz Crystal Oscillator Circuits Design Handbook; U.S. Army Electronics Laboratories, AD 460377, Ft. Monmouth, New Jersey, pp. 3, 1965.

3. Driscoll, N.M. and Healey, D.J.: Voltage Controlled Oscillators; IEEE Transactions on "Electron Devices, 8:528, 1971.

4. Arakelian, R. and Driscoll, M.M.: Linear Crystal Controlled FM Source for Mobile Radio Application; IEEE Transactions on Vehicular Technology, 2:43, 1978.

5. Frerking, M.E.: Crystal Oscillator Design and Temperature Compensation; Van Nostrand Reinhold Company, New York, pp. 5 & 177, 1978.

6. Palladino, G.J. and Richter, K.A.: End Guessing in Oscillator Design; Microwaves, October 1970.

"DESIGN CONSIDERATIONS FOR A 10 KILOWATT, SOLID-STATE, 1.5 TO 30 MHZ POWER AMPLIFIER"

BY: Lee B. Max; Vice President Power Amplifier Products at Microwave Modules and Devices (MMD)

and

Orville B. Pearce; Staff Engineer at MMD

INTRODUCTION

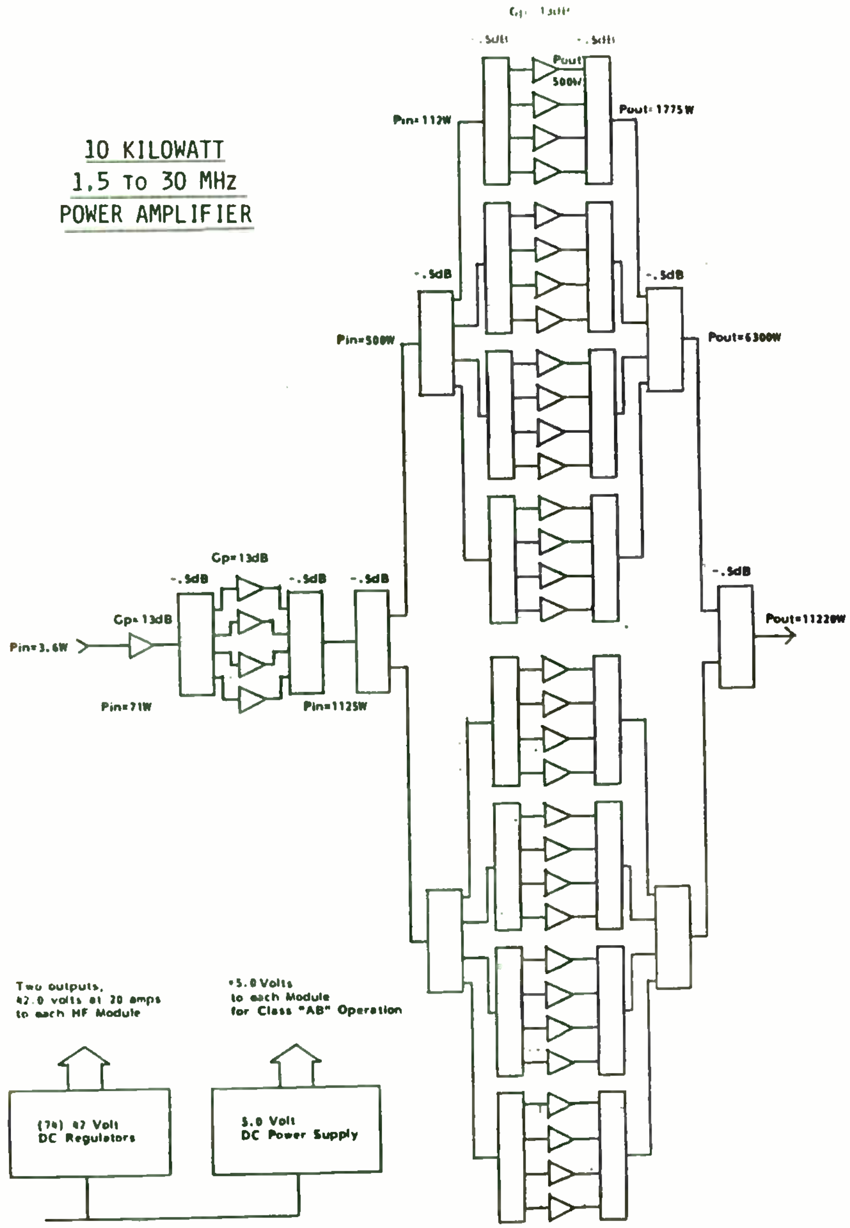
RF technology has reached the level where multi-kilowatt solid-state transmitters are not only possible, but necessary, practical, and cost effective. The paper addresses the design considerations for a solid-state, HF power amplifier to be used in a 10 kilowatt transmitter. The system block diagram, power amplifier architecture, the high power combiners, and the selection of transistor die and package must be addressed. The design considerations covered in the text are based not only on the 50 ohm RF output power requirement, but also on the thermal considerations, antenna VSWR, and reliability. The heart of the power amplifier is the 500 watt, 50 ohm basic building block module since the 10 kilowatt unit uses thirty-seven such modules.

10 KILOWATT POWER AMPLIFIER

PERFORMANCE OBJECTIVES

- FREQUENCY RANGE 1.5 TO 30 MHZ
- 10 KILOWATTS MINIMUM (INTO 2.0:1 VSWR)
- IMD_3 $<$-32 DBC
- HARMONICS: EVEN $<$-30DBC; ODD $<$-20DBC
- OPERATING TEMPERATURE -40°C TO + 71°C

10 KILOWATT
1.5 TO 30 MHz
POWER AMPLIFIER



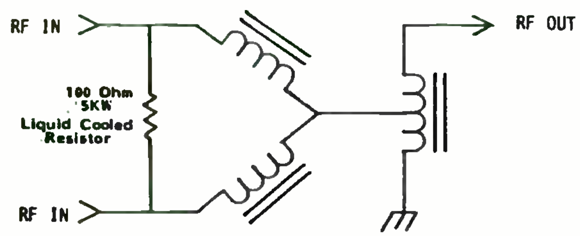
HF COMBINERS

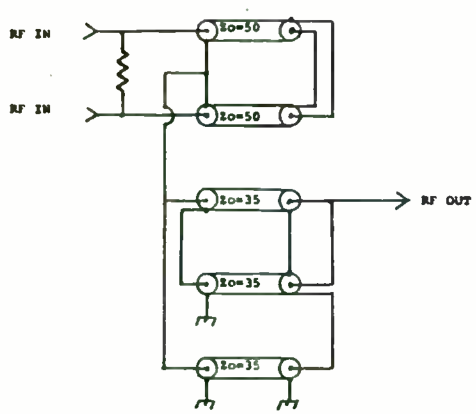
PERFORMANCE OBJECTIVES

- 10 KILOWATT 2 WAY COMBINER
- 7 KILOWATT 4 WAY COMBINER
- INSERTION LOSS .3DB MAXIMUM
- AMPLITUDE BALANCE ±.2DB
- LIQUID COOLED DIFFERENCE LOADS
- ISOLATION 20DB MINIMUM
- PHASE BALANCE ±4.0°

2 - WAY 10 KILOWATT COMBINER

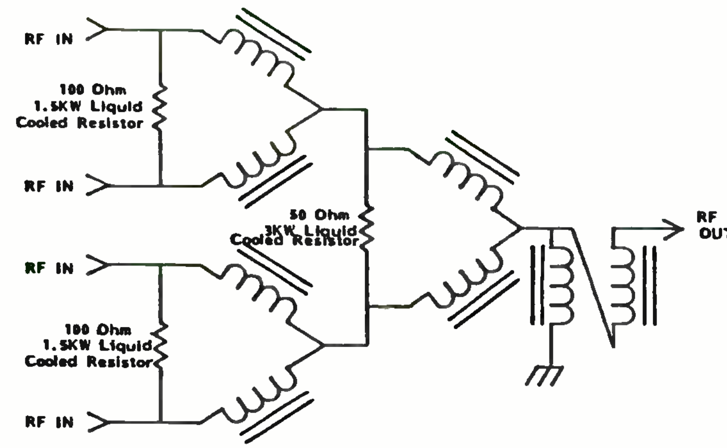
SCHEMATIC

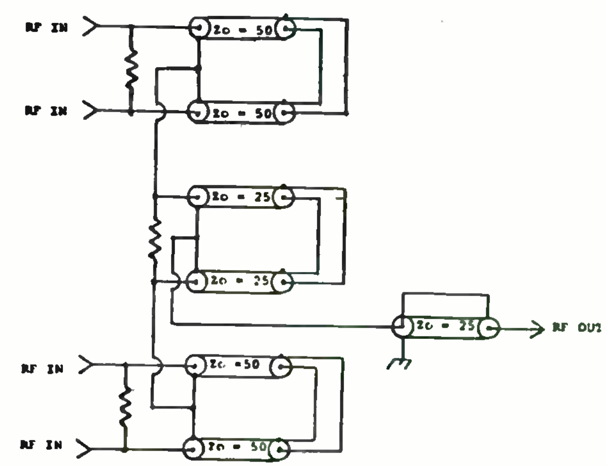




4-WAY 7KW POWER COMBINER

SCHEMATIC





TRANSISTOR DIE CHARACTERISTICS

- Large Silicon Area/Large Base Area
- High Level of Emitter Ballasting
- High Breakdown Voltages
- High F_T
- Proper EPI Thickness

TRANSISTOR DIE OBJECTIVES

- Low Thermal Resistance
- High Current Capability
- Very Rugged
- Linear Class AB Performance

TRANSISTOR PACKAGE OBJECTIVES

- Mechanically Reliable
- Low Parasitic Inductances
- Good Die Attach Capability
- Low Thermal Resistance

AREAS OF MAJOR THERMAL CONCERN

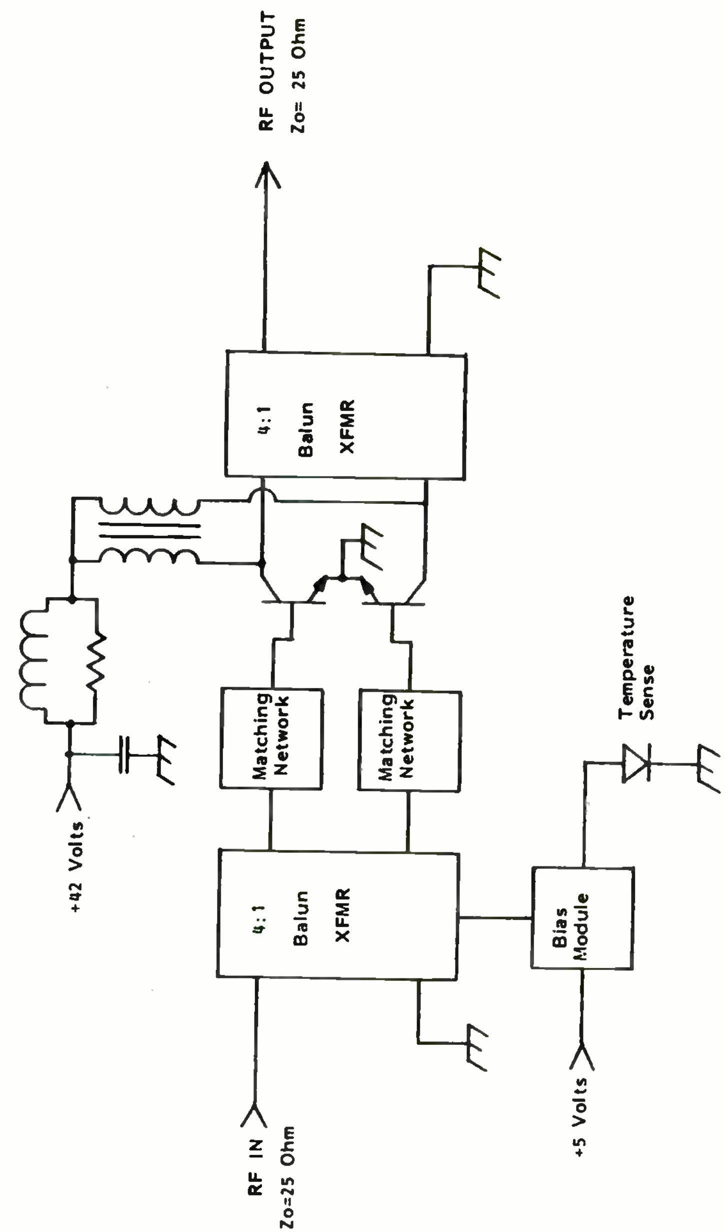
- Transistor Chip Junction
- Output Transformer Ferrite
- Power Loads for Isolation
- Output Combiner Ferrite
- Bias Point Stability

HF SUB-MODULE

PERFORMANCE OBJECTIVES

- FREQUENCY RANGE 1.5 TO 30 MHZ
- 300 WATTS MINIMUM
- GAIN 14DB MINIMUM
- IMD_3 <-35DBC MAXIMUM
- HARMONICS EVEN <-30DBC ODD <-15DBC

300W HF SUB-MODULE

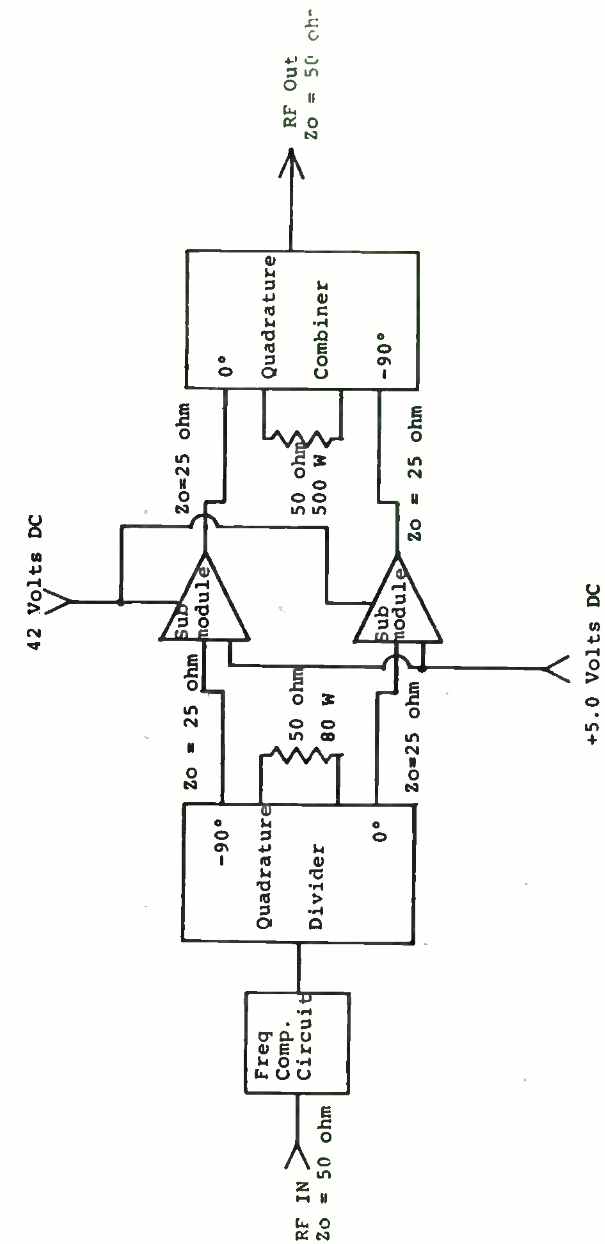


HF MODULE

PERFORMANCE OBJECTIVES

- Frequency Range 1.5 to 30 MHz
- 500 Watts Minimum (into 2.0:1 VSWR)
- Gain 13dB Minimum
- Gain Flatness ±1.5dB
- Input & Output VSWR 1.5:1 Maximum
- $IMD_3 < -35$dBc
- Harmonics: Even < -30dBc ; Odd < -20dBc

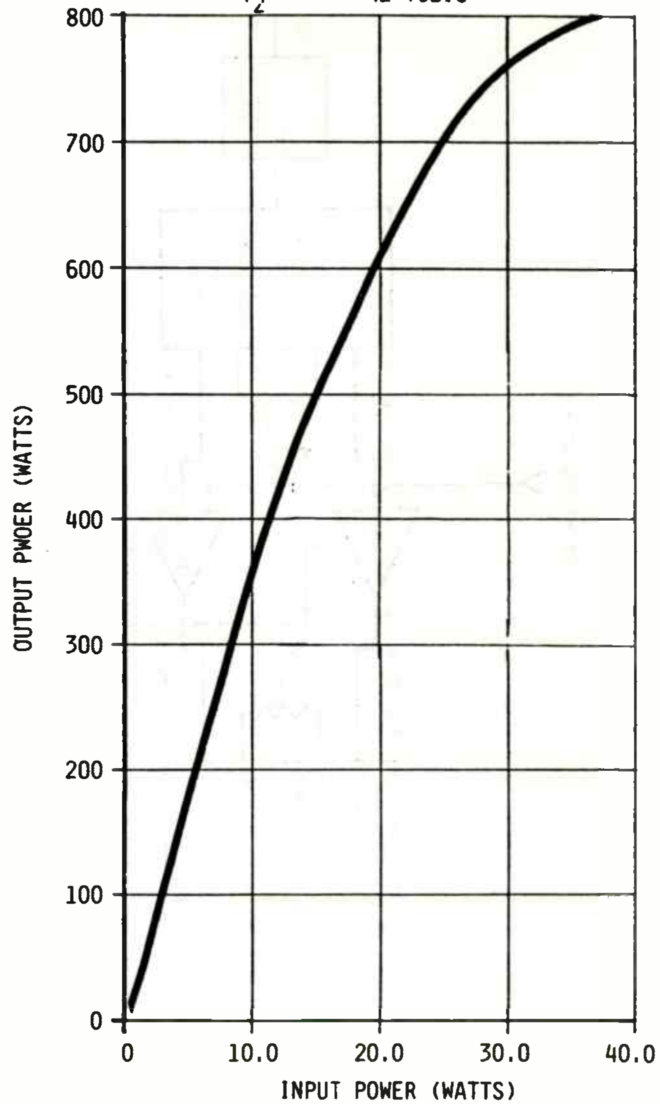
500 WATT HF MODULE

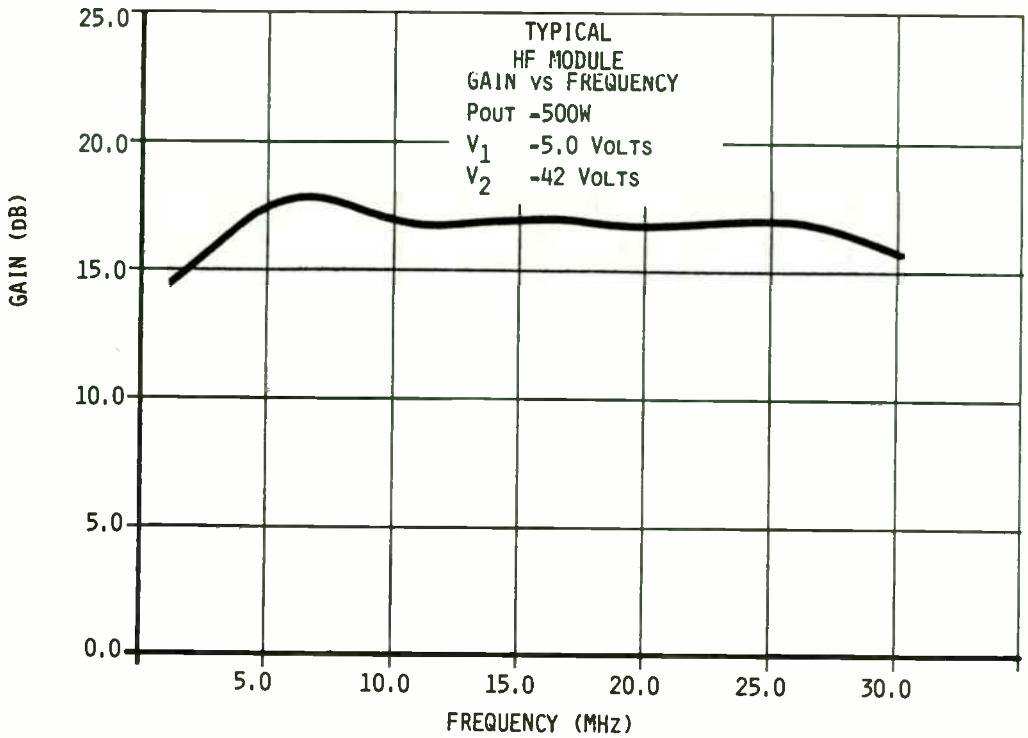


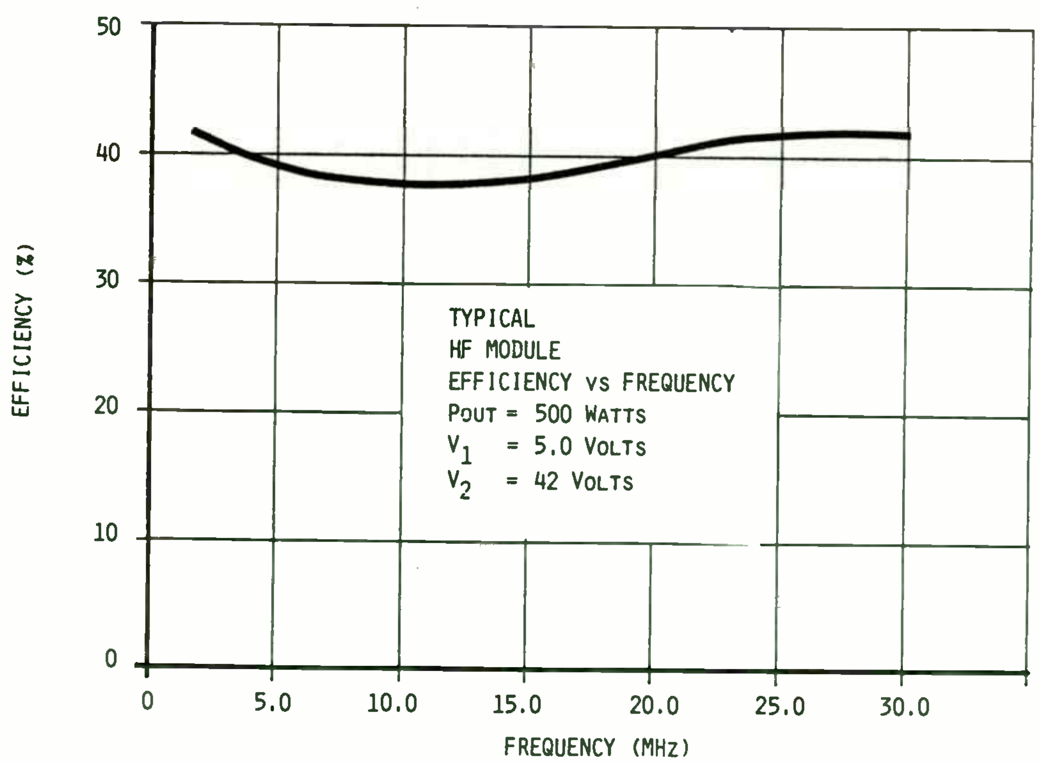
FEATURES OF MMD "QUAD MODULE"

- 1.5:1 VSWR In and Out
- Improved Power into a Mismatched Load
- Improved Back IMD
- Improved Efficiency into a Mismatched Load
- Low Even Harmonic Levels
- Reduced Odd Harmonic Levels

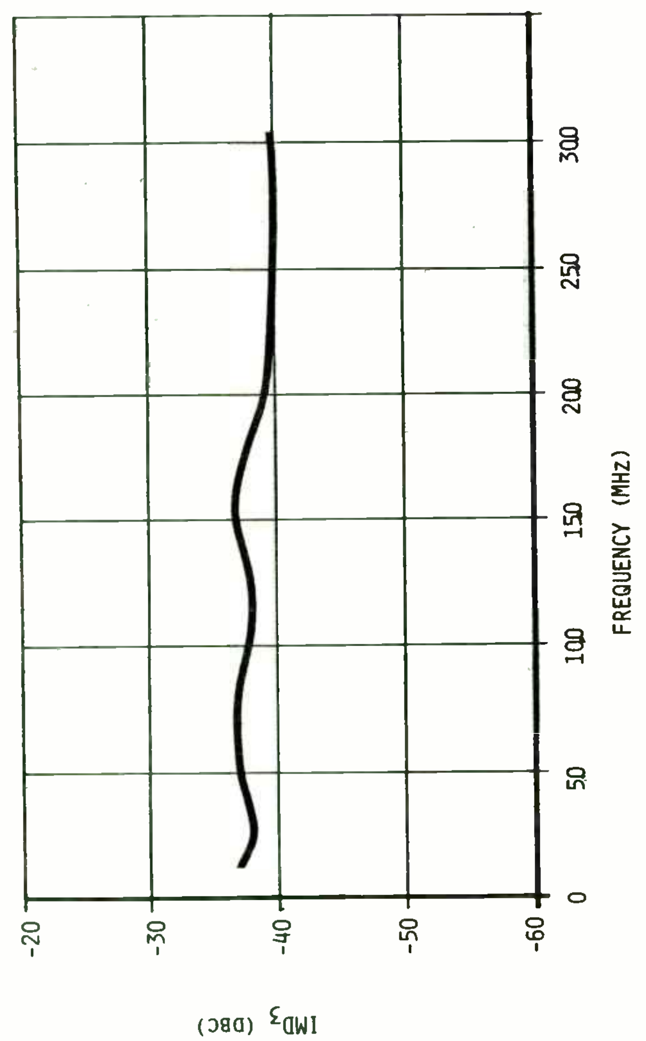
TYPICAL HF MODULE
OUTPUT POWER vs INPUT POWER
FREQ. = 30 MHz
V_1 = 5.0 Volts
V_2 = 42 Volts




TYPICAL
HF MODULE
GAIN vs FREQUENCY
POUT -500W
V1 -5.0 VOLTS
V2 -42 VOLTS
GAIN (dB)
25.0
20.0
15.0
10.0
5.0
0.0
5.0
10.0
15.0
20.0
25.0
30.0
FREQUENCY (MHz)


TYPICAL
HF MODULE
EFFICIENCY vs FREQUENCY
POUT = 500 WATTS
V1 = 5.0 VOLTS
V2 = 42 VOLTS
EFFICIENCY (%)
50
40
30
20
10
0
0
5.0
10.0
15.0
20.0
25.0
30.0
FREQUENCY (MHz)

TYPICAL
HF MODULE
IMD_3 VS FREQUENCY
POUT = 500 WATTS PEAK
V_1 = 5.0 VOLTS
V_2 = 42 VOLTS



SUMMARY

The individual technologies of high power RF transistor chips, high power wideband combiners, low thermal resistance transistor package, and wideband high power impedance matching circuitry have all reached a high level of reliability and reproduceability With these individual pieces comfortably available, multi-kilowatt solid-state transmitters are practical and cost effective. To achieve transmitters that are reliable and manufacturable, the system design engineer must intelligently specify and select the right building blocks,so when interconnected, the amplifier system will achieve the required performance and reliability levels. The design considerations outlined here should help in both the definition of a realizable system for your individual requirements and in the selection of the right modular blocks needed to smoothly construct your deliverable system.

AN ADAPTIVE RANGE GATED RADAR - SYSTEM AND CIRCUIT ELEMENTS

Dr I J Dilworth
Department of Electrical Engineering Science
University of Essex
Colchester
Essex
UK

SUMMARY

A universally adaptive range-gated radar system has been developed. In our application the system is used in conjunction with 34 and 9 GHz pulsed radars to obtain detailed information on the spatial and temporal reflectivities of particulate media in the troposphere eg Rainfall and upper atmospheric ice crystals. However, the accurate quantative measurements produced by such a system have many potential applications.

The main area presented in this paper (and oral presentation) will be the circuitry developed for the (fast) range-gating system and an outline of the control system which allows the circuitry to be adaptably controlled via a small (and cheap) microprocessor.

The RF circuitry is designed for use at an intermediate frequency and therefore may be used in principal, on any pulsed superheterodyne based system.

Introduction

The purpose of the experiment is to investigate electromagnetic scattering levels at various angles from hydrometeors in the troposphere, such as rain, the melting band and upper atmospheric ice crystals.

Because of congestion in the lower microwave communications bands future terrestrial and satellite systems will be forced to use ever higher frequencies. Apart from the increasing absorbtion (and hence attenuation) with frequency of EM radiation - from gases and liquid water, there is also the problem of interference from hydrometeor induced scattering and radio refractive index beam bending. Such interference is potentially most troublesome between co-frequency terrestrial and satellite systems as illustrated in figure (1).

The CCIR have adopted a model for the calculation of such bistatic scattering interference based upon the effective reflectivity of the common scattering volume and the effective distance through the rain. This model uses the bistatic cross section of the raindrops within the common-volume between the two radio systems. The model is based upon work by CRANE in 1974 in which the rain scatter and hence bistatic cross-section was calculated using the Rayleigh approximation for spheres. The formulation of the modelling within the common volume thus uses isotropic scattering from individual scatterers. It would thus seem reasonable to investigate the validity of this rather simple formulation for use in the frequency bands currently under consideration (20-40 GHz) for both ice, water and conglomerate particle scatters. The assumptions in need of verification were:

(i) the single-particle bistatic scattering model and
(ii) the common volume model.

Our theoretical work so far has concentrated on the former but the latter is also currently under consideration

However some measured data suggest that scattering above about 10 GHz is not Rayleigh-like. The main aim of this experiment is thus to measure various regimes of scattering and to quantitively model certain regimes in order to establish whether, at 34 GHz, a Rayleigh scattering model is accurate enough to be used for planning and coordination purposes. For example (single raindrop) theoretical calculations at 30 GHz indicate only small differences between Rayleigh and exact calculations using Mie and Fredholm Integral formulations as shown in figure (2). The data in this figure assumes an incident vertically polarised signal and shows the radiated power versus azimuthal angle. Note the null for the Rayleigh Model.

Experimental Systems

For a slant satellite link there are two possible interfering situations; one from terrestrial links and the other from satellites in near orbits. In the first case, as well as possible rain scatter interference, tropospheric radio refractive index irregularities (RRI) can cause significant problems. Scattering from the melting band and upper atmospheric ice crystals will not be important in the terrestrial case except for 'over the horizon' or 'overshoot' propagation paths, where the terrestrial link crosses and hence illuminates the fresnel zone region of the slant satellite path, in this case usually above any rainfall. Clearly, RRI effects may influence the frequency of this occurence.

For the case of adjacent satellite interferers, rain, ice, and the melting band may all contribute to scattering. Hence, for both the case of (1) scattering from terrestrial links into slant satellite paths and (2) scattering into adjacent satellite-earth paths it is necessary to investigate all three regions of the troposphere.

From the systems viewpoint, measurements approximating to likely interference angles, both in azumith and elevation, and path lengths are desirable. Such angles and paths are not necessarily those one would choose when making measurements to evaluate and improve a theoretical model of scattering which can then be more generally applied.

System Constraints

The system constraints leading to the final experimental set-up will now be outlined. The conclusion we reach is that the system and experimentalists set-up are not coterminous.

(i) Elevation angle

The Radar elevation angle to be = geosynchronous satellite elevations from the UK and ideally to be the same as currently available satellite beacons.

(ii) Bistatic scattering volume

To be well defined so that, as far as possible, there are no ambiguities as to the region and hence hydrometeors we are observing - implies narrow beam antennas.

(iii) Bistatic angle

Interference into satellite earth stations from terrestrial systems can be from any azimuthal direction. The elevation angle is, however, constrained. The resultant possible scattering region is approximately a hemisphere situated below the height of the highest scatterers. For rainfall, theoretical calculations indicate that significant scattering can occur for all bistatic angles within the proposed measurement zone. (We must take care with polarisation, however.)

(iv) Path Length

The primary constraint on this bistatic angle, however, is the path length from the scattering region to the receiver(s). The further away the receiver(s) then the smaller the bistatic angle from any particular scattering region becomes. However, if attenuating hydrometeors are present in this path the minimisation of its length is important in order not to mask the scattered signal amplitudes we are trying to measure. Reference to tables 1 and 2 shows for given rain rates how much downpath attenuation results, assuming various scattering heights, and the corresponding scattered signal amplitude.

The attenuations have been calculated using the results of NORBURY and WHITE who used a short path (2x224m) and close spaced rain gauges (45m). The attenuations are therefore based on average rainfall rate over the whole path.

These results are comparable with HOGG using a 30 GHz, 3 km link in Bedford, England, where the climate conditions should be similar to our Essex site. From these results it would appear that a specific attenuation of 5 dB/km (rain rate 20 mm/hr) will not be exceeded for more than 0.01% of the time.

The scattered signal amplitudes have been based on the plots of GUNN and EAST and isotropic re-radiation has been assumed.

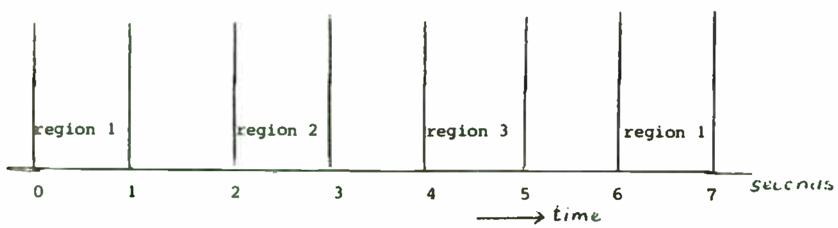
Tropospheric changes

The melting band height will change between 0.5 - 3 km on a seasonal basis. It will also change over a smaller range in the short term.

Nodding of Antenna

To accommodate this we 'nod' the receive antenna in elevation. In this way we can observe all scattering regions from one antenna and receiver system. Fixed dishes of narrow beamwidth do not offer the same versatility or data acquisition possibilities. In order to produce detectable scattered signals at the receiver we need to look at each scattering volume for ~1sec.

In practice we look sequentially at three regions initially, one in, one above and one below the melting band. To do this the antenna must move swiftly to the next region.



Determination of Melting Layer

'3-D' time profiles of the range gated radar backscattered signals eg figure 4 show a peak of reflectivity at constant height from the melting band. Such data will be simultaneously available from the X-band (already working) and the Q band radars. We are able to use the bistatic receiver in a 'search' mode, where the dish is continually sweeping in elevation, and hence obtain similar time profiles of bistatic scattering. We revert to the sampling of 3 regions upon command from the computer system or by the operator.

Terrestrial and Q band radar

By splitting the Q band radar output and transmitting a component along a terrestrial path to the bistatic receiving site and using a range gated radar receiver we can obtain data relating to the rainfall along the path. Also by receiving the forward scattered radar pulses we get direct measurements of the path averaged attenuation. Secondly, we get information on the spatial rainfall rate along the path from the range gated radar receiver. This will furnish us with rainfall path length data which will assist in analysing the bistatic scatter signal amplitudes we simultaneously receive.

The Range Gated Radar System

A block diagram of the whole system is shown in figure (8). In order to extract sufficient quantitative data for exact modelling of the interference phenomena spatial as well as temporal data are required. The spatial data is most easily processed if it is quantised in range (and amplitude). In order to record detailed reflectivity profiles from the 9 GHz terrestrial system we employ 60 gates of 80 metres length. This produces sufficient accuracy for the 5 km path. See figure (5).

Turning now to the 34 GHz Bistatic system a unique feature of this is the ability to be able to spatially separate the scatter from the mainlobe common volume and the sidelobe to mainlobe common volume as shown in figure (6).

This is made possible by triggering the Bistatic receiver gating circuit (30 gates) by the terrestrially received 34 GHz pulse. Under certain circumstances it is possible to get more scattering from the sidelobe - mainlobe volume than the mainlobe common volume. With this system we are able to identify such situations. Since the bistatic receive dish 'nods' in elevation the intersection of the 30 gate 'gating window' needs to be moved in order to be in the correct position to coincide with the common volume. This is achieved, under software control from a shaft encoder input, by ROM based data calculated using trigonometry. This is illustrated in figure (7). In order to correct for inverse square law loss the bistatic receiver employs gain-sweeping as illustrated in Figure (9) and Table 3.

Detailed block diagrams of the transmit system and bistatic receive system are shown in figures (10) and (11) respectively. The diagrams are largely self explanatory except for the microprocessor controlled range-gating system which will now be described.

Hardware Design for Range Gating

In order to resolve spatial distances down to, say, 15 metres in the Troposhpere switching speeds of 50ns (~20-30 MHz) are required. Because the common volume in this experiment is small ie the volume in which scatter's are illuminated by the transmit beam, the scattered received power is also small. It is therefore necessary to integrate the received signal, for a period, to achieve a useable S/N level. We employ a pulsed Magnetron, for the 34 GHz transmitter, with a pulse repitition frequency of 1 KHz. Calculations show that good sensitivity is achieved by using < 1 second integration time. The video processing system needed for range-gating the radar is shown in figure (12). The driver hardware on the left of the diagram shapes the incoming terrestrial pulse and from this pulse provides variable switching and delay for the gating.

The main problem with this system is providing the current source driver with sufficient speed to obtain the required resolution. Two systems were tried, shown in figure (13). The final circuit realisation is shown in figure (14).

Circuit Description

Video signals at levels between 0.17v - 1.8v are fed to the scaling voltage amplifier (2N2369 - 2N2369) to obtain a level suitable for driving the tail current source transistor. This transistor then supplies current, proportional to video input, to either integrator or bypass transistor. When the gate is disabled, all current flows in bypass transistor - this is held on by logic high level on GE.... control line.

The enable pulses rapidly turn off the bypass transistors and direct current through integrator transistor to charge the integrator capacitor (.01µ). Since this current is dependent upon the video input voltage, the capacitor voltage will be an integral of this, taken over the integration time.

This integration time is chosen to be 1 sec giving approximately 1000 samples of the video waveform, of 50ns width.

Gate voltages are read on a common line via 4066 transmission gates, one to each gate, - these are selected via the gate read multiplexer by the microprocessor control software. Resetting is achieved when the gate is accessed by applying +12v to the common read line via paralleled 4066 gates. However, to avoid a large shunt capacitance the gate circuitry is arranged into banks of 12 gates. Each bank is selected individually by decoding the gate which is being addressed.

High impedance on the common read line is achieved by feeding this into a 3140 fet input op-amp. Adjustment of integrator gains is obtained by variable resistors in tail transistor emitter circuits. Performance of the gates are shown in figure (15).

Timing and Software Controlled Switching Systems

The drive system for the gates is shown in figure (16). This system supplies 50ns gate enable pulses (GE...) to the high speed gate system (which can be of arbitrary length). These pulses are obtained from a 20 MHz oscillator which clocks 74LS161 counters to obtain binary gate 'addresses'.

Synchronisation is derived from the terrestrially received pulse.

Both the scan delay, via a 74192, and the gate widths, can be driven under software control and hence may be changed dynamically. The sequence of operation is shown in the lower diagram in figure (16).

Terrestrial Processing

The terrestrial signal pulse is used to derive sampling pulses for reading its own level. The sampling pulses are obtained by supplying the input pulse to a comparitor with a threshold of ~0.2v (o/p from log amp) and this gives a satisfactory sample pulse down to -80 dBm signal level.

Because the pulse width as received, is too wide and may vary with multipath effects, the initial transition of the pulse is used to trigger a simple monostable using a differentiator and Schmitt trigger gate (22_pF - 1.5k) thus the pulse width is determined by this time constant.

An integrator gate was employed to determine the terrestrial signal level to improve recording accuracy and to be compatable with the bistatic range gates. The integration time is therefore 1 sec and sampling pulse width 0.1µs.

The threshold control on the comparitor is set up to give no triggering for a noise level coming from the log amp - adequate sensitivity is then obtained.

CONCLUSIONS

A versatile signal processing system for a pulsed radar system has been developed. The system may be reconfigured under software control. Performance is such that resolution of the order of 15 metres are readily detectable although in this particular application long integration times appear necessary due to relatively low signal levels. The main advantage of this pulsed system over C W Bistatic systems is, although the system sensitivity is theoretically less, ability to spatially resolve received (scattered) signal amplitudes and hence detect sidelobe couplings will prove useful in improving theoretical models of scattering and hence communications systems design.

Acknowledgements

The author gratefully acknowledges the UK Science and Engineering Research Council, and the Rutherford and Appleton Laboratories for funding this work. Manifold thanks are also due to Dave Johnson who developed much of the hardware for this project.

Fingringhoe Site

TABLE 1

	receive power			attenuation		
Height h	1mm/hr	5mm/hr	20mm/hr	1mm/hr	5mm/hr	20mm/hr
1km	-57.9dBm	-47.9dBm	-38.8dBm	2.7dB	5.4dB	27dB
2km	-55.2dBm	-45.2dBm	-36.2dBm	3.3dB	6.5dB	32.5dB
3km	-56.8dBm	-46.8dBm	-37.8dBm	4.5dB	9.0dB	45dB

Receive power levels include 3.7dB feeder and duplexer loss at transmitter and 1dB feeder loss at receiver.

TABLE 2

Expected power levels at backscatter receiver and attenuation for rain

	receive power			attenuation		
height h	1mm/hr	5mm/hr	20mm/hr	1mm/hr	5mm/hr	20mm/hr
1km	-56.0dBm	-46.0dBm	-37.0dBm	2.0dB	3.0dB	20dB
2km	-62.0dBm	-52.0dBm	-43.0dBm	4.0dB	6.0dB	40dB
3km	-65.5dBm	-55.5dBm	-46.5dBm	6.0dB	9.0dB	60dB
4km	-68.0dBm	-58.0dBm	-49.0dBm	8.0dB	12.0dB	80dB
5km	-69.9dBm	-59.9dBm	-50.9dBm	10.0dB	15.0dB	100dB

Receiver power levels include 3.7dB feeder and duplexer loss at transmitter.

TABLE 3 Gain Sweep Calculations

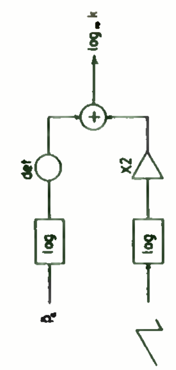
Gain Sweeping

Need to compensate for the inverse square law for backscatter ie. remove r dependency.

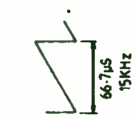
For constant rainfall over entire path length, put in constant k

$\bar{P}r = \frac{k}{r^2}$

for log IF $\log_{10} \bar{P}r = \log_{10} k - 2 \log_{10} r$ = det O/P



Gain sweep to 10km



$\frac{20 \cdot 10^3 \text{ secs}}{300 \cdot 10^6}$

$= \frac{200}{3}\mu S.$

Required compensation

Find log amp function from I/P v O/P plot

-30dBm $0{\cdot}85 = u + x \log_{10} 10^{-6}$

-60dBm $0{\cdot}37 = u + x \log_{10} 10^{-9}$

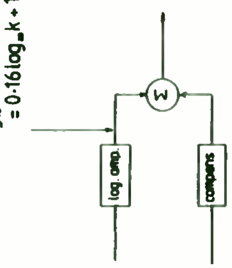
$0{\cdot}85 = u - 6x$

$0{\cdot}37 = u - 9x$ $x = 0{\cdot}16$

$0{\cdot}85 = u - 6.\,0{\cdot}16$ $u = 1{\cdot}81$

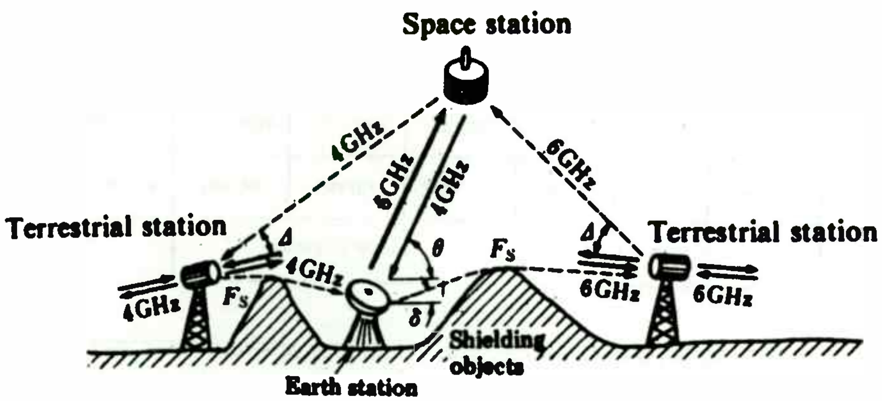
det O/P $= 1{\cdot}8 + 0{\cdot}16 \log_{10} \bar{P}_R$

$= 0{\cdot}16 \log_{10} k + 1{\cdot}8 - 0{\cdot}32 \log_{10} r$



System will give range ramp (incorrect data) if reflections to max range fall below dynamic range of log amp

FIGURE 1 Scattering Geometrics



Solid line : Wanted signal
Broken line : Interfering signal

Interference between satellite and terrestrial systems sharing the same frequency bands (4/6 GHz bands)

FIGURE 2 Scattering Amplitudes Versus Angle for Rayleigh, Mie and FIM

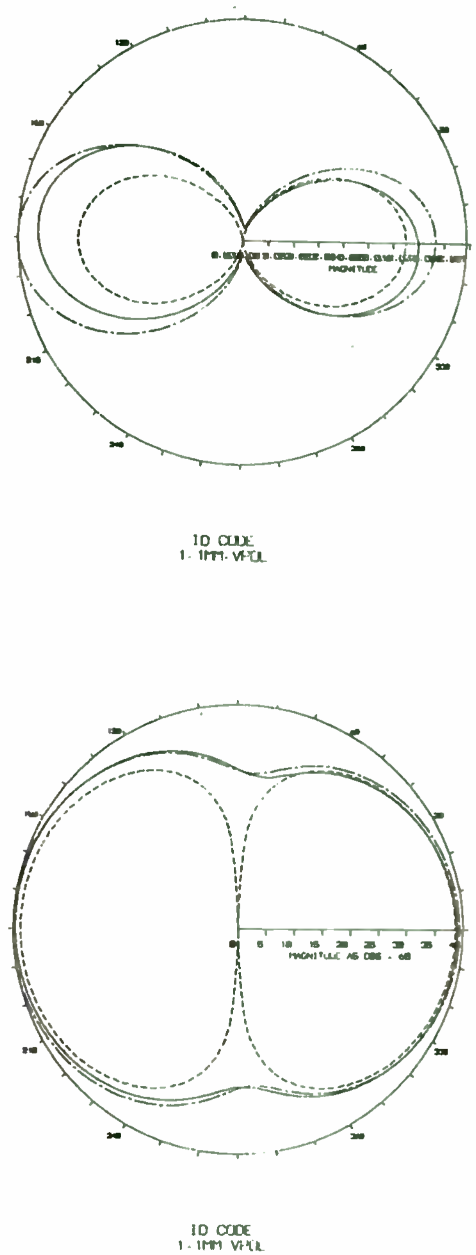


FIGURE 3 Experiment Path and Geometry

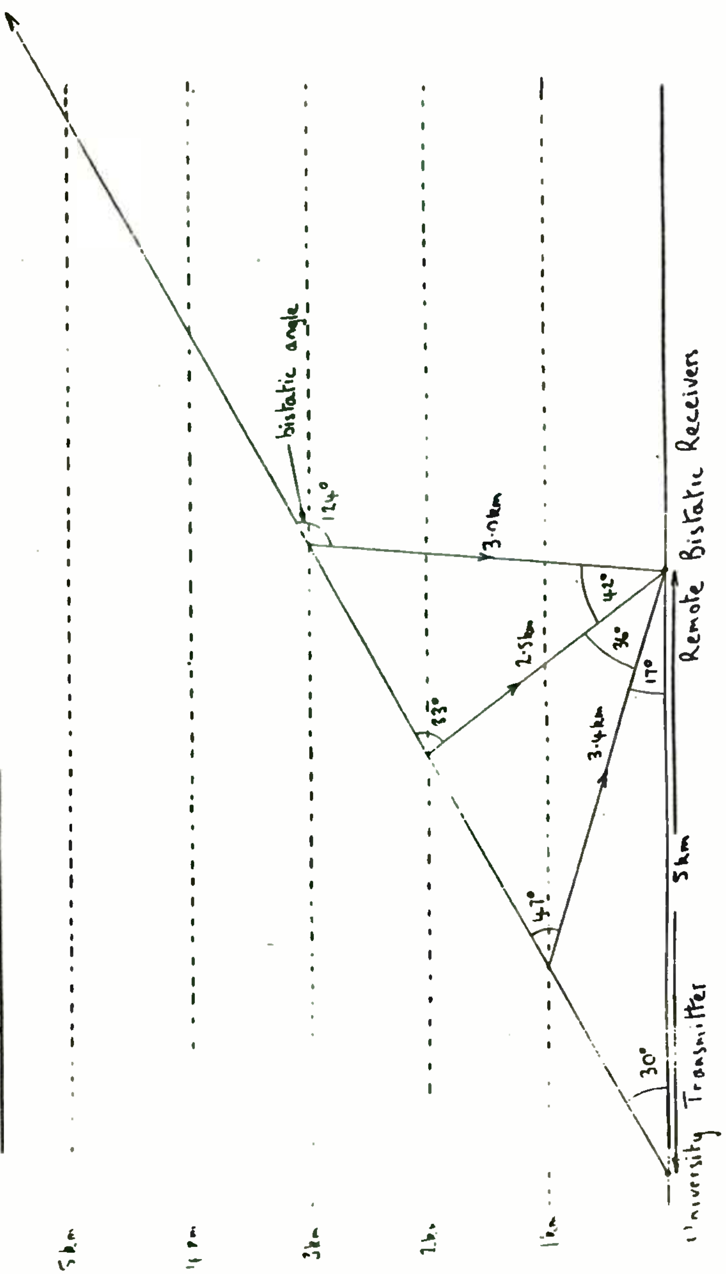


FIGURE 4 '3D' Isometric Graph of Radar Reflectivity

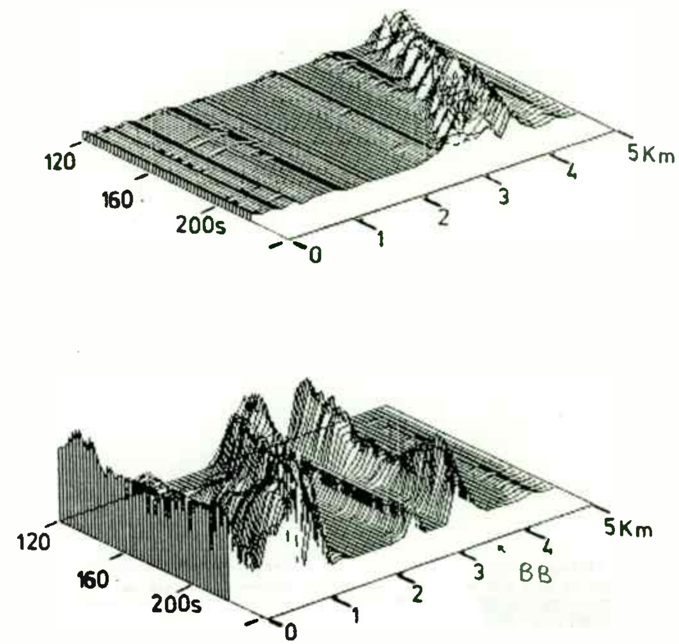


FIGURE 5 Path Profile

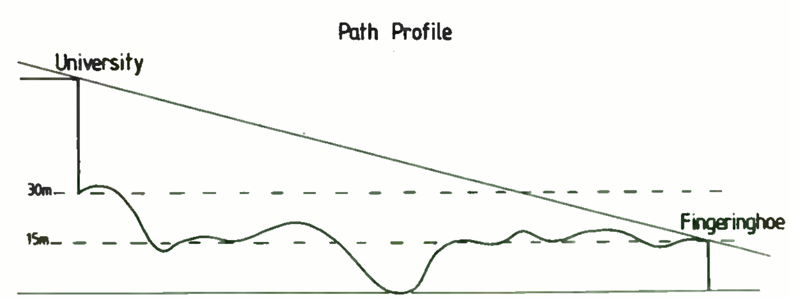


FIGURE 5

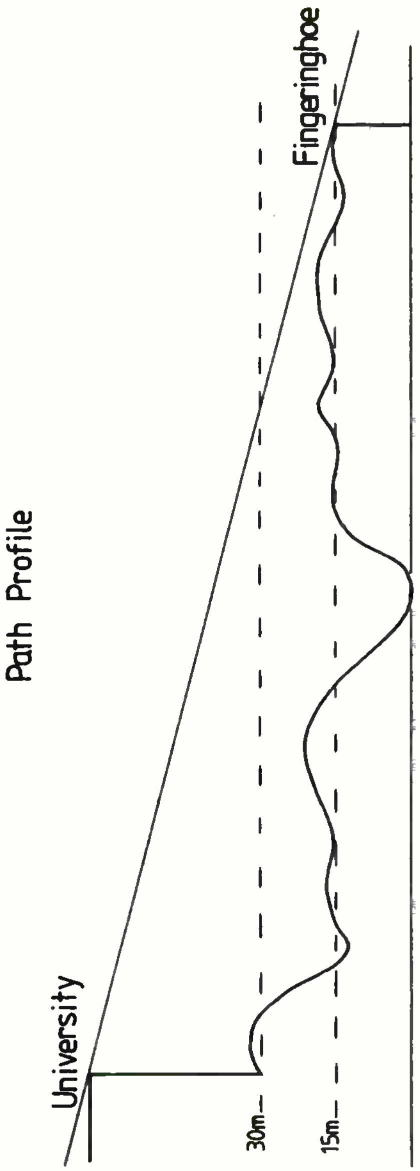


FIGURE 7 Variable Gate Window Timed by Terrestrially Received 34 GHz Radar Pulse

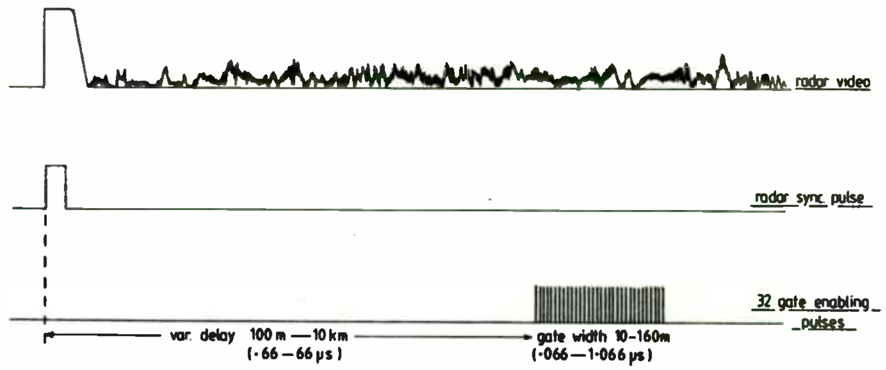


FIGURE 8 Block Diagram of Experiment

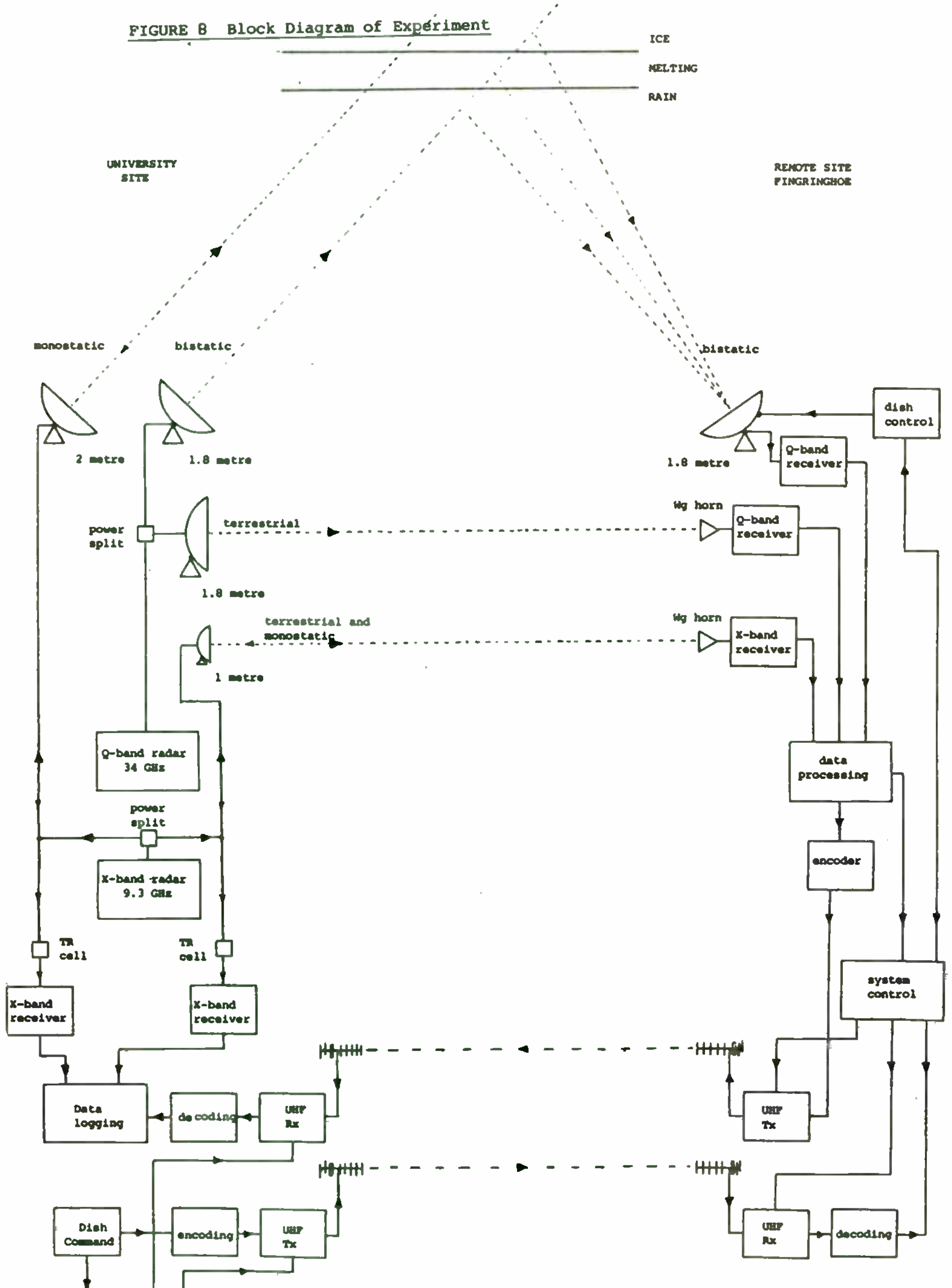


Fig. Bistatic, monostatic and terrestrial experimental system

FIGURE 9 Gain Sweeping System

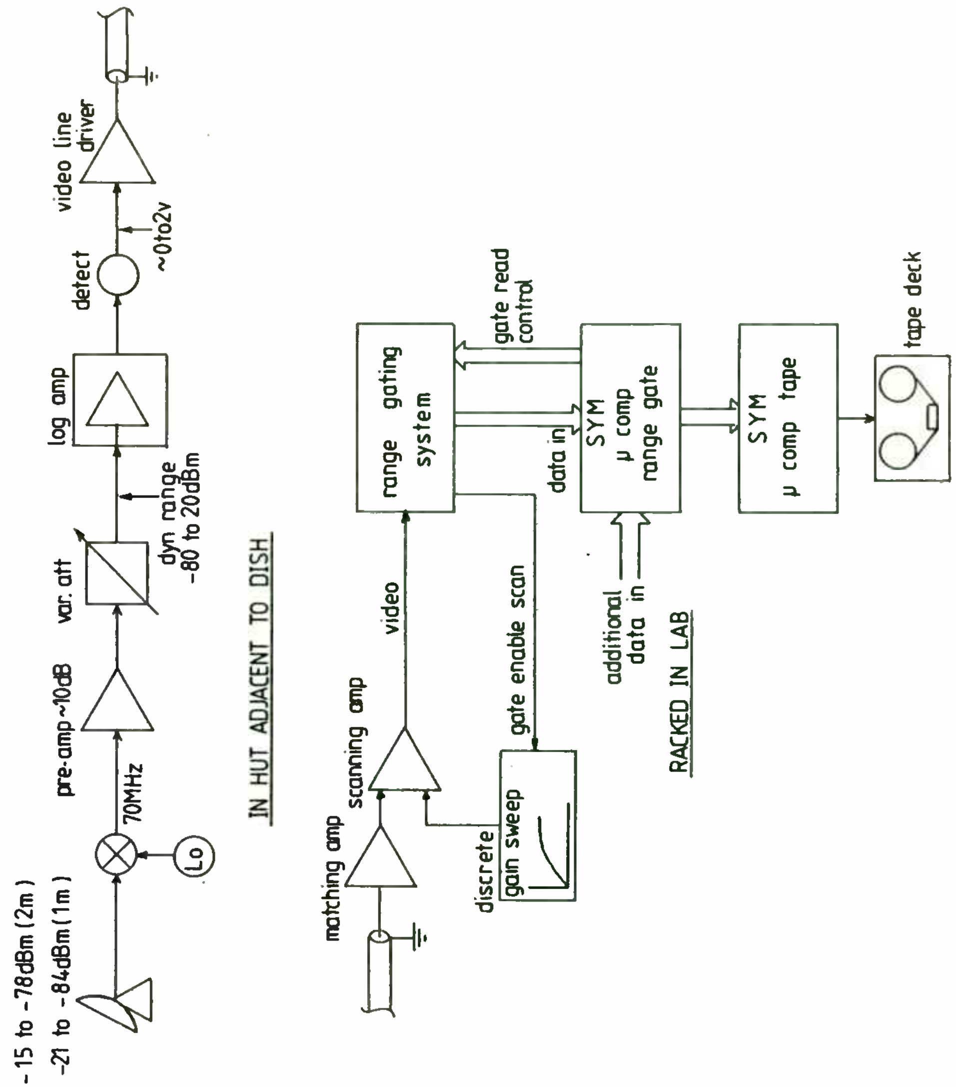


FIGURE 10 Transmit System

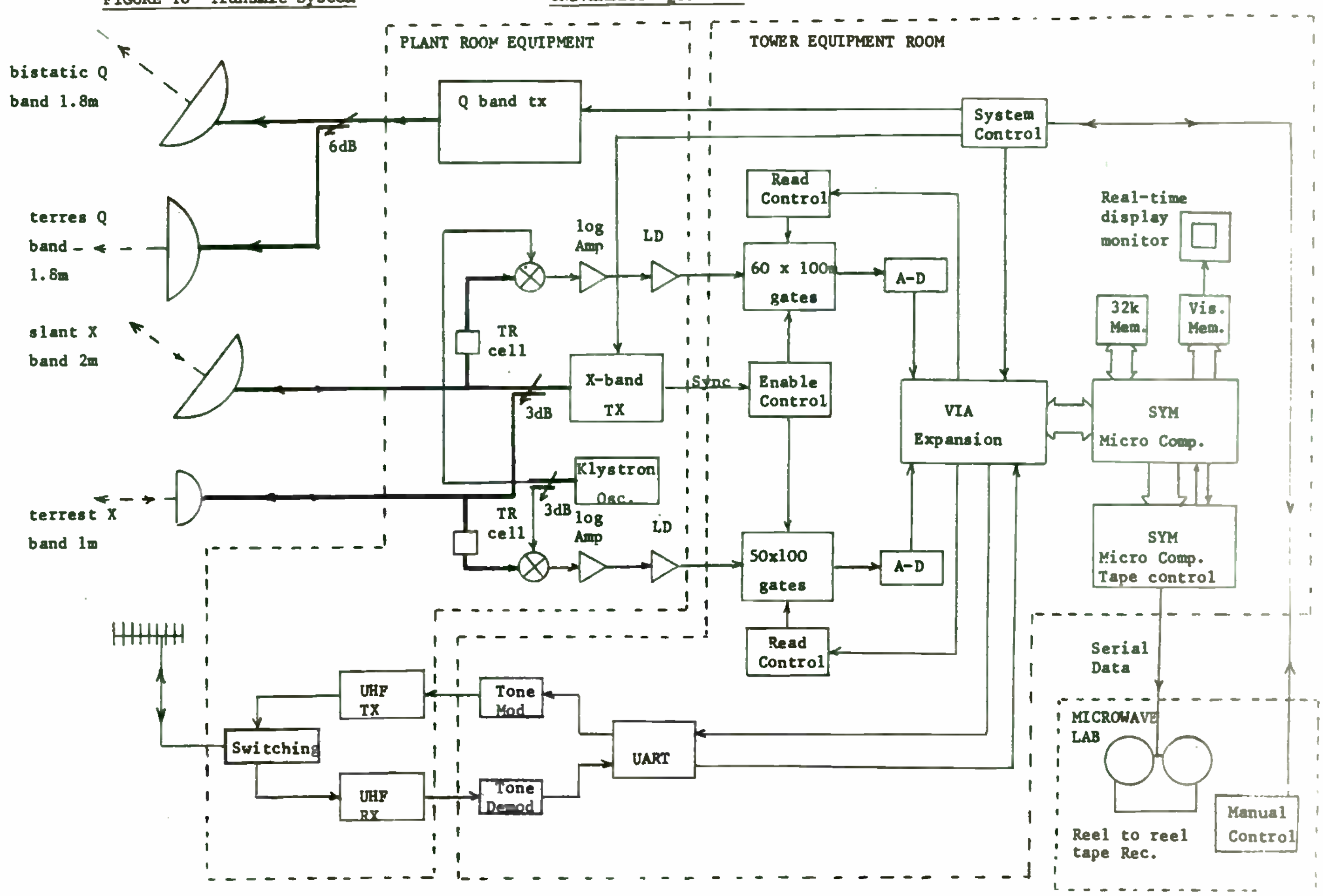
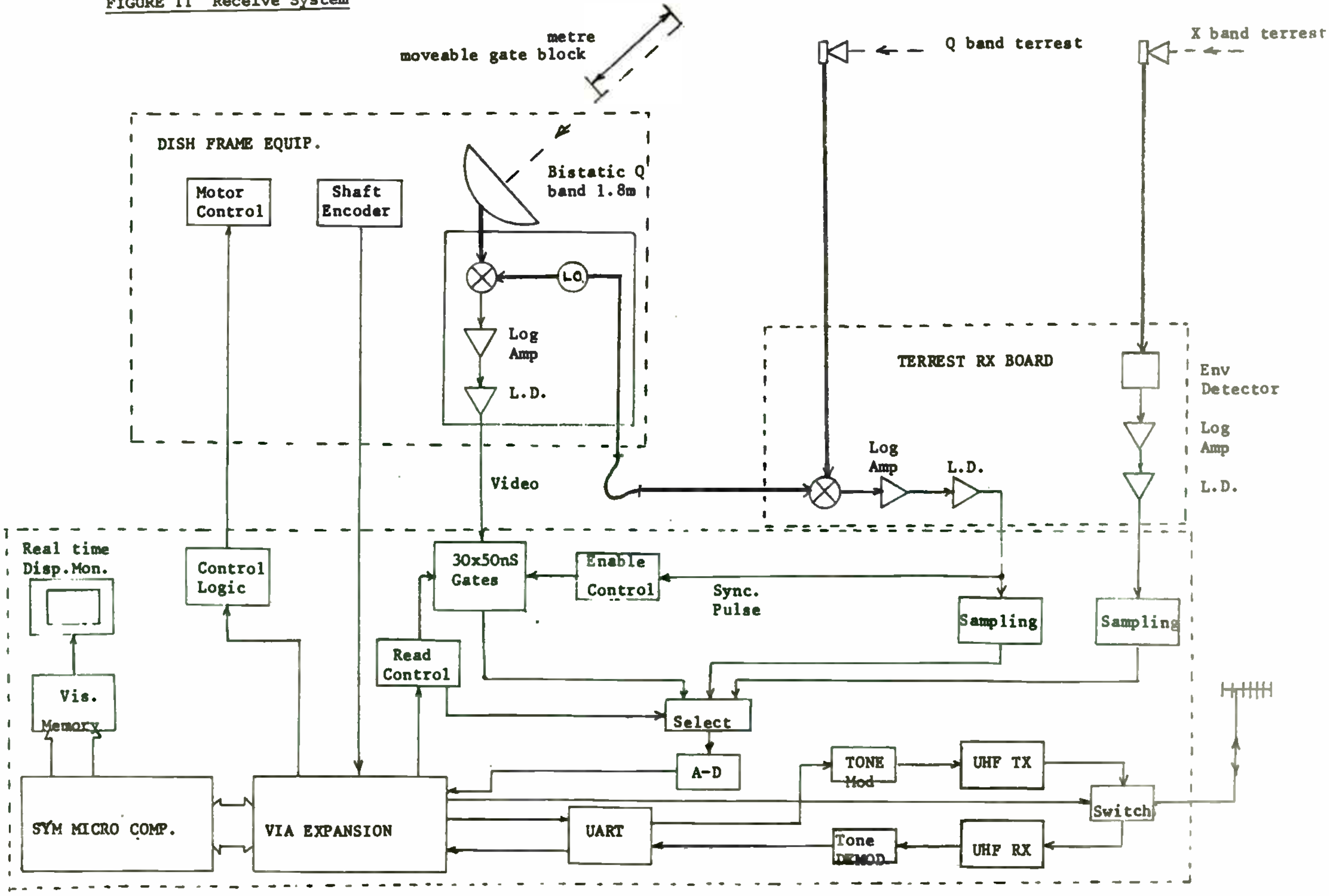


FIGURE 11 Receive System



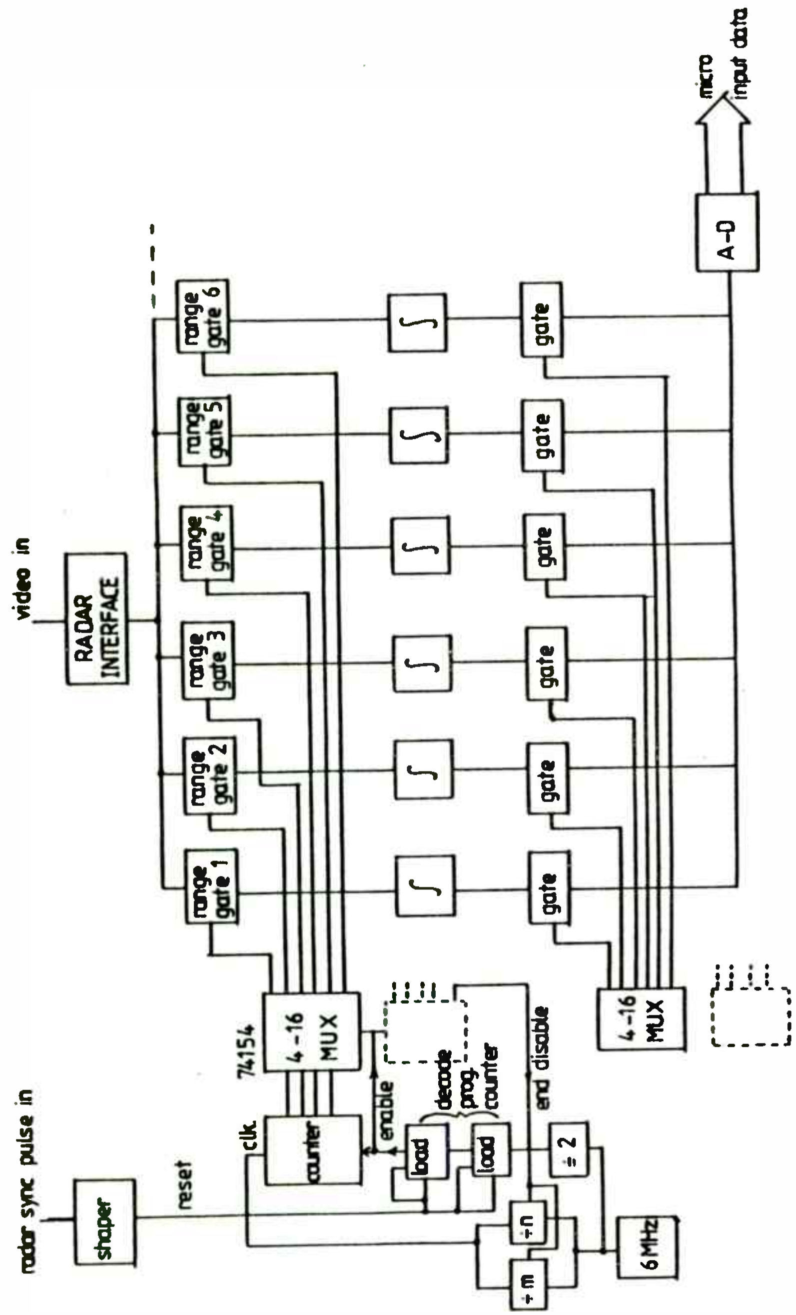
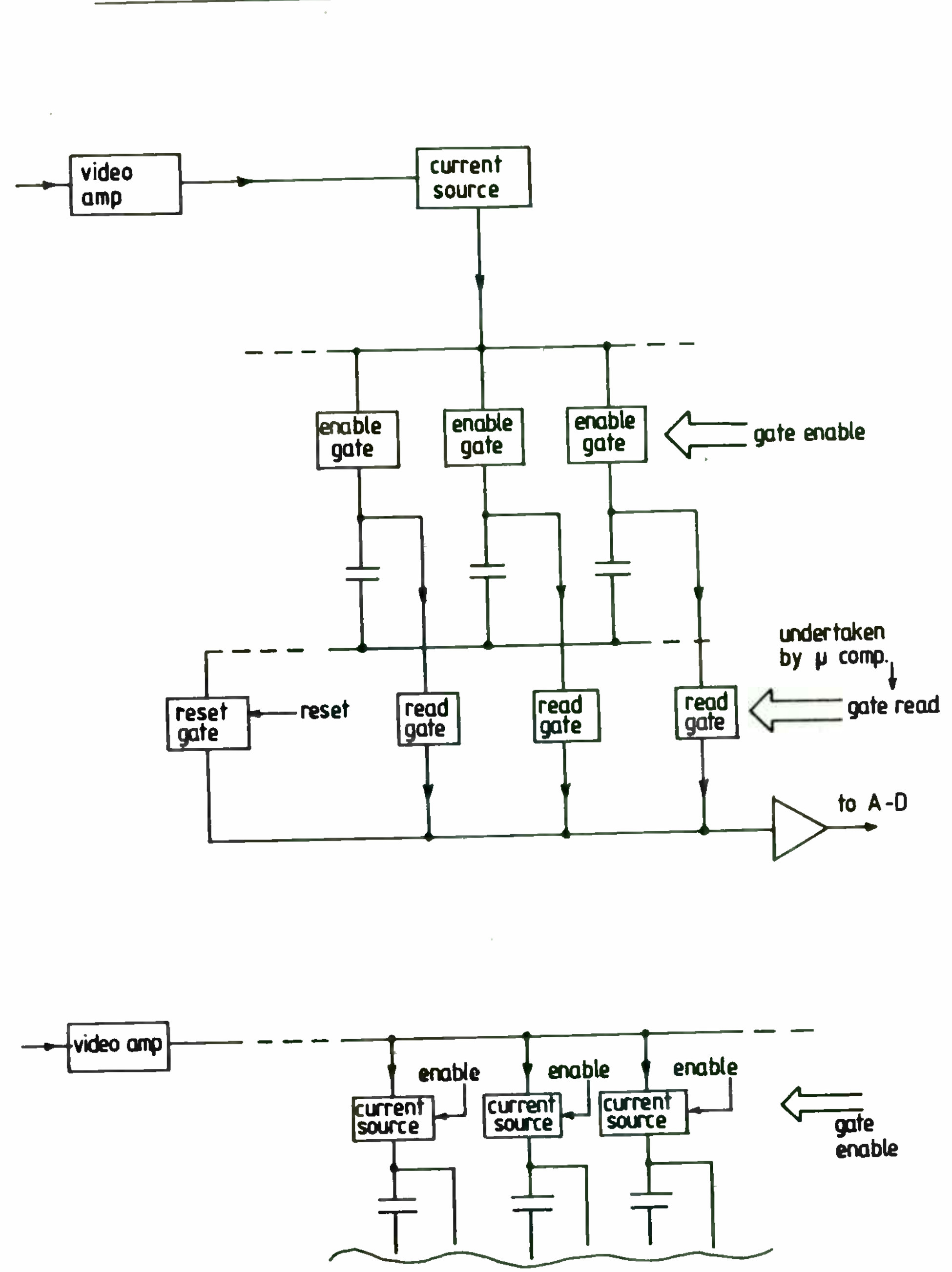


FIGURE 12 Video Signal Processing

FIGURE 13 Gate Drive Systems



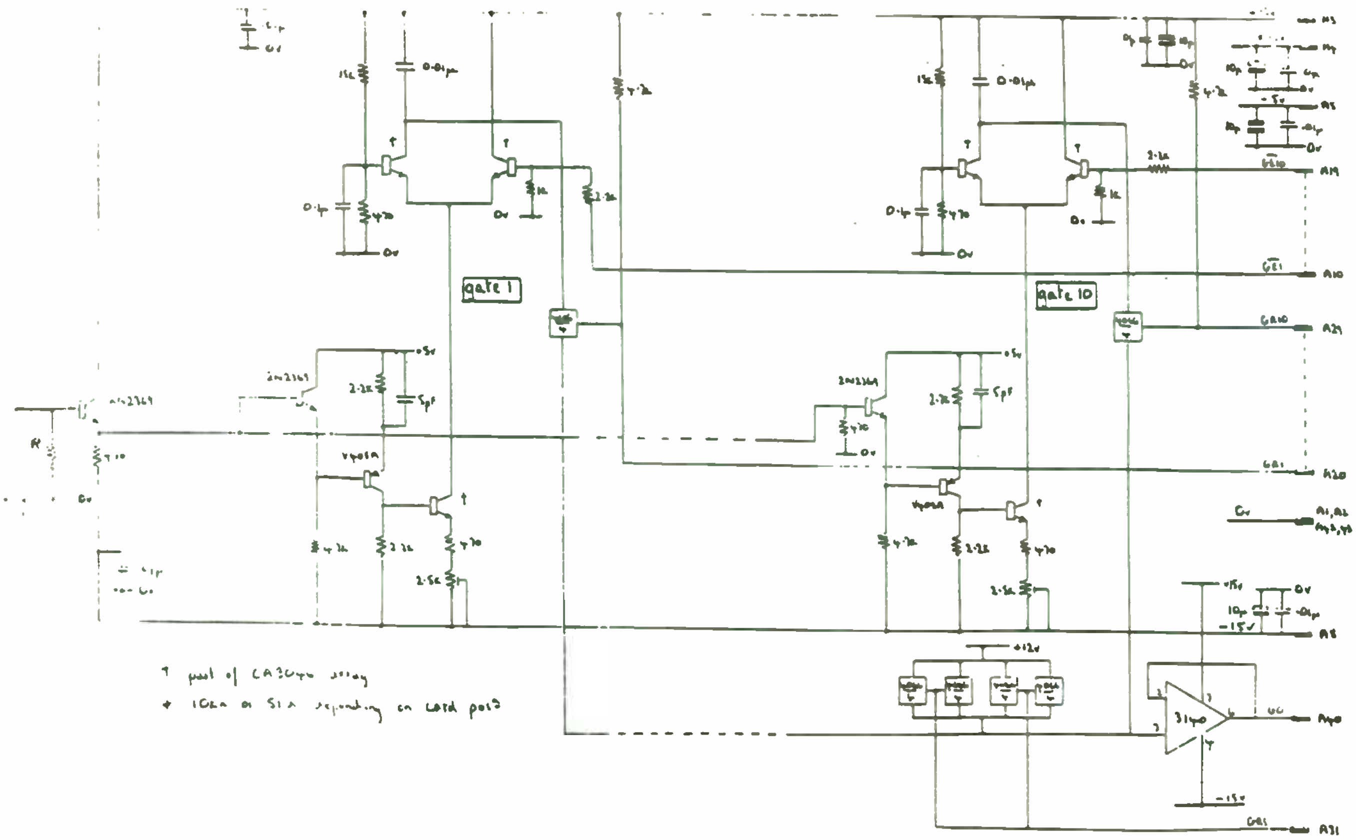
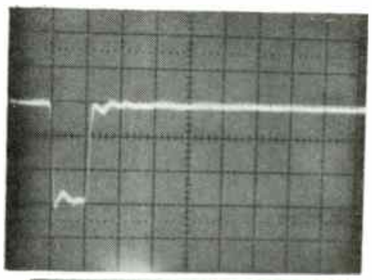
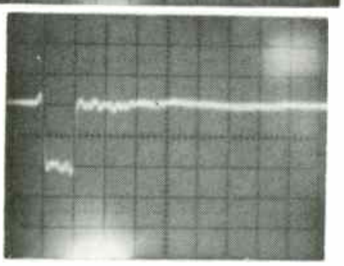


FIGURE 14 High Speed Gate System

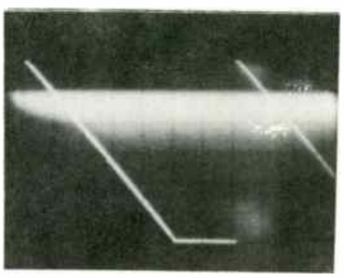
HIGH SPEED GATE PERFORMANCE



Pulse Drive to $\overline{GE}$
50ns / div
1v/div



Gate Current
100Ω collector R
50ns/div
0·2v/div

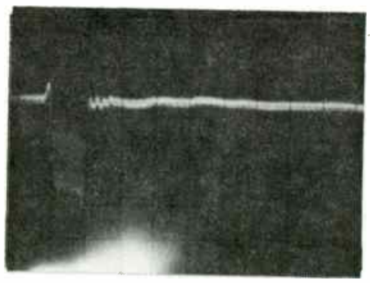


Capacitor Charging
Ramp (0 01μF)
50ns sampling pulse
1sec. integ. time
0·2sec/div
2v/div

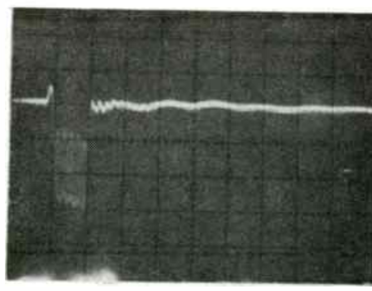
HIGH SPEED GATE
VIDEO FREQUENCY RESPONSE

100Ω Coll. load
I/P 1·5v rms
50ns/div
2v/div

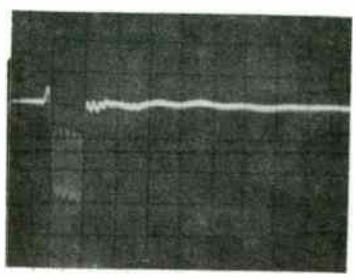
Eye diag.



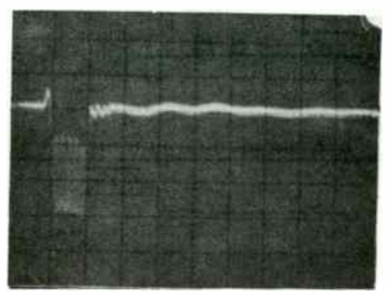
10 MHz



1MHz



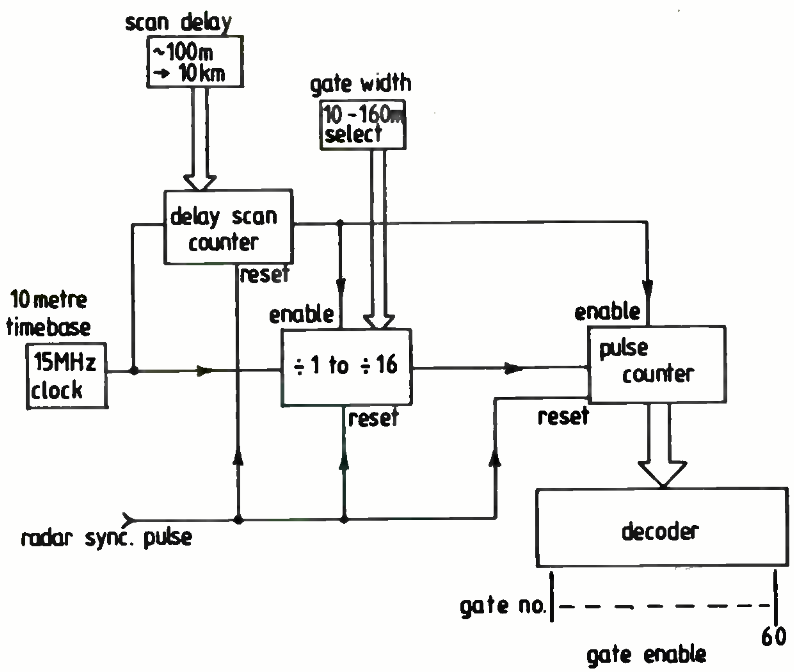
10 KHz

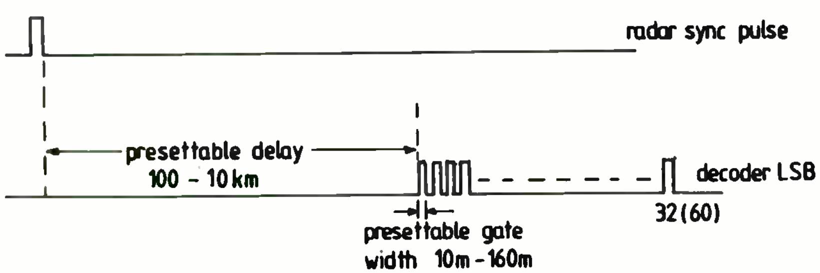


100 Hz

FIGURE 15 Timing and Switching System

FIGURE 16 Gate Drive System





The Logical Completion of the Cascode Idea as it Relates to Radio Frequency Amplifiers

William F. Griffith

THE EXTENDED CASCODE AMPLIFIER

This work was first applied to mobile radio audio output at Bendix Radio Division in 1962 as a means to utilize the earliest engineering samples of Silicon transistors , as described in Ref. 5 . In 1981 it appeared that the excessively high resistivity of the base material exhibited by those early bipolars might not be a disadvantage in a context such as radiation hardening of power amplifiers .

About the years 1979 and 1980 the Extended cascode was experimentally applied at R.F. since it appeared that the configuration would yield higher non-overloaded output than the conventional circuits using a given bipolar device , as well as higher wide- band gain and better distortion characteristics near overload . These augmented gain and Max. output power were both desirable in interface amplifiers for fast digital data links using coaxial cable . The gain advantage was instrumental to realization of interfaces having immunity to digital pulse interference due to proximity of other interface cables . Short links could be thus protected by use of outer coaxial conductor having high magnetic permeability as well as reasonably low conductance in virtue of the enhanced skin effect . (A given thickness of high permeability braid contains many more skin depths than an equal thickness of copper at I.F. and R.F. parts of the spectrum .)

Radio frequency application was by means of transmission line coupling transformer; based on several earlier succesful uses of the 1:1 ratio transformer of this type to configure wide - band push-pull class B and class A amplifier modules in both the R.F. and audio ranges of frequency . (Ref. 12). The 1:1 ratio transformer of this type was widely used in the infancy of the "R.F.I." and "E.M.C." disciplines as a general purpose "fix" for common-mode coupling problems and it was subjected to careful analysis (Ref. 13 , for example) which disclosed the greater bandwidth of the "common - mode choke " connection compared with the isolating transformer connection . The device was imbedded in a push - pull amplifier module using two identical type bipolar transistors to produce an AC substitute for the complimentary pair ; I.e., the inputs were paralleled in anti-phase and so were the outputs . This tightly - coupled

class B package is probably worthy of a detailed report in its own right since it obviates the use of complimentary devices ~~xkxkk~~ and so may offer cost and performance improvements at V.H.F. or microwave frequencies . It would constitute the preferred R.F. output circuit for an extended cascode amplifier . It is shown in a modified form in Fig. 5 of Ref. 5 , in which only the outputs are in parallel . As a three terminal device , the two bipolars coupled by a "common mode choke " can be utilized in common emitter , common collector or common base modes of operation with the devices in parallel or series D.C. connection . Only the series D.C. connection ,however , permits realization with a miniature transmission line transformer at audio and R.F. For higher outputs , parallel D.C. connection involves a 1:1 output transformer of higher current rating and larger size .

Detail design data for an automotive radio output amplifier are given to define the component sizes somewhat , although the transistors used are not representative of current types . See Fig. 1 . The output transistor of this 2 Watt amplifier was a 2N1227 - 2 with a 0.3 Ohm CUPRON positive temperature coefficient emitter resisitor , which was typical of Bendix practise of the time . The audio output transformer , designed to operate at supply D.C. Voltages between 10. V. and 16. V. , had nominal primary impedance of 26 Ohm when loaded with a standard 8 Ohm speaker . The driver transformer was strikingly smaller than the typical one for the function , being a small input unit having impedance ratio 50,000 / 30 and DC winding resistances of 2600 Ohm and 6 Ohm, respectively . Its turn ratio was 41 / 1. With an output transistor of 28 dB. minimum gain , this driver had to deliver 4.2 mW. The operating point of the Hyperm driver core was kept near zero magnetisation by approximate cancellation of the primary DC Ampere-turns by bias current in the secondary . Sensitivity at 1 K.Hz. was 22 m.V. for 1.0 W. into 8 Ohms with 3 dB. frequencies at 33 Hz. and about 4. K. Hz. A feature worth noting was the maximum Watts output above clipping level , which ranged from 4.5 W. at 1 K.Hz. to 3.7 W. at 5 k.Hz.

The prototype audio amplifier incorporated local positive Voltage shunt feedback inherently and over-all negative Voltage feedback . (Ref's. 5 & 11 .)

THE EXTENDED CASCODE AMPLIFIER Page 3

A 160 Mbit/s. interface amplifier was realized recently by connecting a Fairchild generio bipolar chip in the cascode configuration with the "extended cascode " feature in the form of a miniature toroidal ferrite-cored transmission line transformer externally connected . Turn ratios of 4/1 and two higher values were used . In the initial approach , lumped - element 75 Ohm filters were assembled and tested to be connected to the amplifier output , which were designed in accordance with Ref's. 7 & 8 . The desired low - pass characteristics were realized , but subsequent examination of the performance on a 0 - 600 M Hz. network analyzer revealed that the amplifier itself , with a single shunt capacitor across the output , actually realized the same pass-band and band edge responses with an improved stop band performance .

The good bandwidth , extending to the transition frequency of the bipolar chip transistors , is attributable to the impedance of the output emitter junction increasing with frequency at a rate that approximately compensates the decrease in frequency response of the input common emitter devices as detailed in Ref's. 6 & 9 . In connection with the local positive feedback inherent in the extended cascode , a general conclusion regarding the gain-sensitivity product of positive feedback schemes in integrated-hybrid active filters appears to be relevant ; namely , that they exhibit much lower gain-sensitivity products than do negative feedback schemes . Ref. 11 is concerned with linearity improvement attained by enclosing a positive f.b. loop within a negative f.b. loop .

A final remark about this configuration appears to be as relevant today in the automotive electronics context as it was in 1962 . Then , as now , there existed electronic modules behind the dashboard which required auxilliary DC-to-DC converters to generate 35. Volt supply from the nominal 12. Volt battery output . By way of contrast , the extended cascode can be designed to operate at Voltages lower than this by an order of magnitude . It is not integratable as it stands , but it poses a challenge to chip designers to realize an integratable circuit to replace the lumped element. Then they would have the needed "gain cell" capable of operating at Voltages well below the five Volt level , a declared prime objective for V.L.S.I.

EXTENDED CASCODE AMPLIFIER BIBLIOGRAPHY

1. Transient Radiation Effects on the Electronics of Aircraft Systems , by J.T. Lambarski , in I.E.E.E. Soc. Circuits & Systems 1978 Symposium Record , pp. 119-123.

2 Controlling Secondary Breakdown in Bipolar Power Transistors , by R.J. Widlar ,in 1981 I.E.E.E. International Solid-State Circuits Conference Digest ,pp.44,45 & 258 .

3. (Tutorial Re. Transmission Line Transformers) A good review is in the article "Broadband 60W. R.F. Linear Amplifier", by O. Pitzalis,Jr., Rob't. E. Horn and R.J. Baranello, in I.E.E.E. Journal of Solid State Circuits , vol.SC-6,No. 3 ,June 1971, pp. 93-103. Biblio.

4. Analysis of the Transistor Cascode Configuration , by J.R. James in Electronic Engineering , Jan. 1960 ,pp. 44-48. Biblio . Also pp. 62 for French language abstract and p. 66 for German language abstract .

5. United States of America Patent No. 4,435,686 , Wm. F. Griffith , issue date 6 Mar., 1984 .

6. Principles of Transistor Circuits ,Ed. by R.F. Shea , Wiley, 1953, pp.207 -.

7. Realizable Matched Filters for Minimal Intersymbol Interference , by L.F. Lind S.E. Nader , in I.E.E.E. 1978 International Symposium on Circuits & Systems Proceedings, pp. 441-446 .

8. Optimum Data Transmission Filters , by L.F. Lind & Said E. Nader, in I.E.E.E. Trans. Circuits & Syst., vol.CAS-26 , No. 1 ,Jan. 1979 ,pp. 36 - 45 .

9. A Single Chip 200 Mb/s. Fiber Optic Receiver Circuit , by Dennis L. Rogers , in 1981 Digest of Technical Papers of I.E.E.E. Int'l. Solid-state Circuits Conference, pp. 172 - 173.

10. A 50 to 400 Mbit/sec. Single Chip Fiber Optic Receiver Circuit , by Dennis L. Rogers and Albert X. Widmer and Joseph Mosley ,

11. United States of America Patent No. 2,652,458 , J.M. Miller ,assignor to The Bendix Corp.

12. U.S. Patent No. 3,100,878 , Wm. F. Griffith , 1963 .

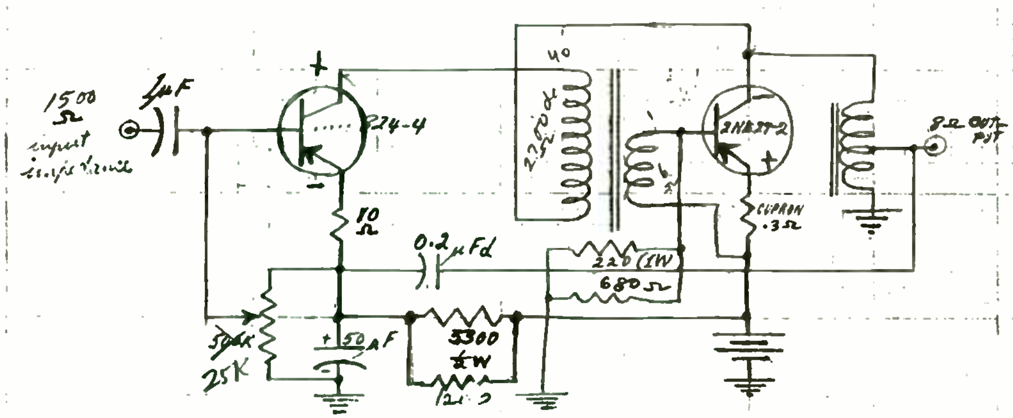
13. The Common Mode Choke , by Tom Herring , in E.M.C. Symposium Record ,I.E.E.E. ~~197~~ 1968 pp. -

W.G./wg 85/01/23

447-44
W. F. Griffith
12 Feb 1960

PARTS SKETCH

2.5 Watt High Gain Audio System (Cascode)



$$Z_i = \frac{23}{120} \times (10K + Z_i)$$

c, 2 μF 10

$$K_1\beta = \frac{\cancel{6005}}{15\cancel{000}} \times \frac{2.82}{\cancel{x}} = \frac{28.2}{15} \doteq 2$$

$$1 + K_1\beta \doteq 3 \rightarrow$$

$$\frac{1}{R} = \frac{1}{680} + \frac{1}{220} = \frac{1+3.1}{680}$$

$$R = 166 \text{ ohm}$$

Sens at 1 Kc/s 22 mV in for 1.0 W in 8.0 ohm

Resp. at ¼ Watt: −4dB at 100 C/S, −7.5dB at 60 C/S ; −3dB at 5Kc/s, −10dB at 10Kc/s

F c/s	MAX UNCLIPPED WATTS	MAX WATTS
1000	2.0 (28mV)	4.5 (90mV)
400	2.0 (32mV)	3.5 (54mV)
200	2.0 (31mV)	2.5 (42mV)
100	1.0 (35mV)	1.5 (72mV)
5000	2.4 (60mV)	3.7 (145mV)
10,000	1.9	

United States Patent [19] [11] 4,435,686
Griffith [45] Mar. 6, 1984

[54] EXTENDED CASCODE AMPLIFIER

[76] Inventor: William F. Griffith, 920 Jefferson St., Apt. 3, Rochester, Ind. 46975

[21] Appl. No.: 267,341

[22] Filed: May 26, 1981

[51] Int. Cl.[3] H03F 1/00
[52] U.S. Cl. 330/311; 330/165
[58] Field of Search 330/112, 165, 166, 167, 330/171, 291, 311

[56] References Cited

FOREIGN PATENT DOCUMENTS

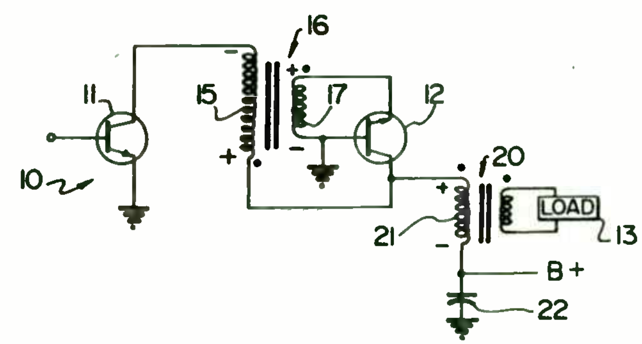
970614 9/1958 Fed. Rep. of Germany 330/311

Primary Examiner—James B. Mullins
Assistant Examiner—Gene Wan
Attorney, Agent, or Firm—Wenderoth, Lind & Ponack

[57] ABSTRACT

Disclosed is a cascode amplifier circuit with improved current gain, linearity, overload performance and bandwidth which uses a step-down (of voltage) transformer between a driver stage and an output stage. One end of the primary of the transformer is connected to a driver transistor's collector and the other end is connected to the collector of a grounded-base output transistor. The transformer secondary is connected across the base and emitter of the output transistor. The transformer windings are phased such that a large driver current is forced through a high impedance, i.e. the primary winding impedance which is a multiple of the common base input impedance of the output transistor, resulting in power gain. The amplifier is capable of high maximum overloaded power output free of "break-up" of the waveform compared with conventional cascode operation, as well as greater power gain. By using transistors with high resistivity base material the amplifier operation is resistant to degradation by ionizing radiation and protected against secondary breakdown.

2 Claims, 5 Drawing Figures



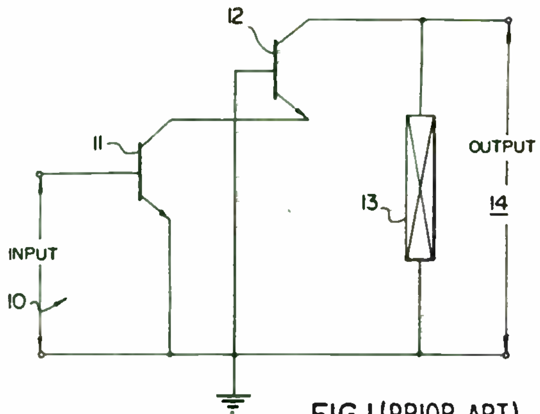


FIG.1 (PRIOR ART)

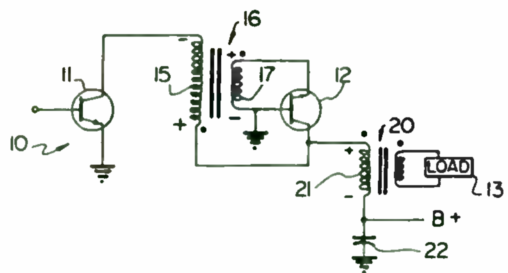


FIG. 4

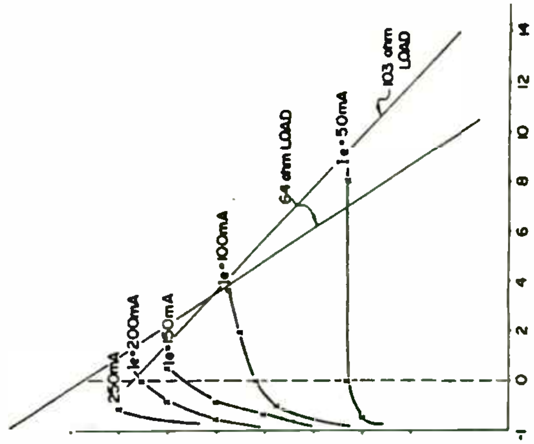


FIG. 3

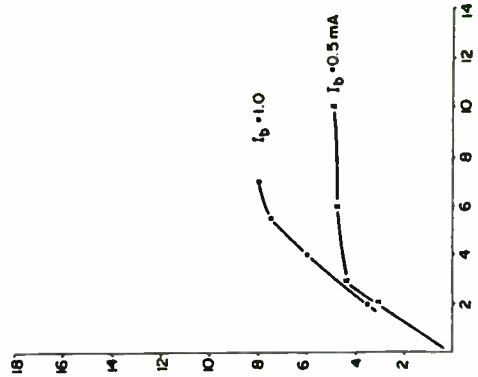


FIG. 2

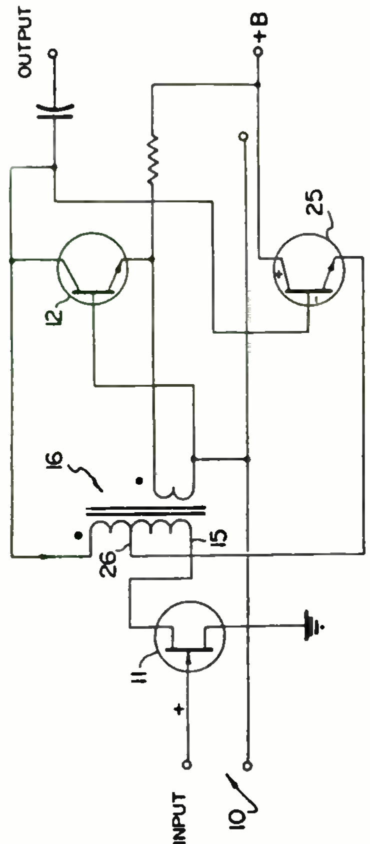


FIG. 5

United States Patent [19] [11] 4,435,686
Griffith [45] Mar. 6, 1984

[54] EXTENDED CASCODE AMPLIFIER

[76] Inventor: William F. Griffith, 920 Jefferson St., Apt. 3, Rochester, Ind. 46975

[21] Appl. No.: 267,341

[22] Filed: May 26, 1981

[51] Int. Cl.³ H03F 1/00
[52] U.S. Cl. 330/311; 330/165
[58] Field of Search 330/112, 165, 166, 167, 330/171, 291, 311

[56] **References Cited**

FOREIGN PATENT DOCUMENTS

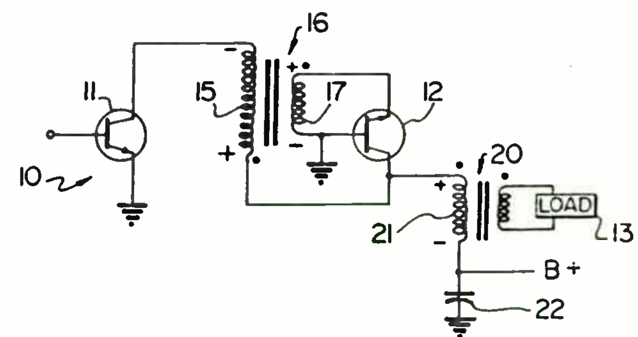
970614 9/1958 Fed. Rep. of Germany 330/311

Primary Examiner—James B. Mullins
Assistant Examiner—Gene Wan
Attorney, Agent, or Firm—Wenderoth, Lind & Ponack

[57] **ABSTRACT**

Disclosed is a cascode amplifier circuit with improved current gain, linearity, overload performance and bandwidth which uses a step-down (of voltage) transformer between a driver stage and an output stage. One end of the primary of the transformer is connected to a driver transistor's collector and the other end is connected to the collector of a grounded-base output transistor. The transformer secondary is connected across the base and emitter of the output transistor. The transformer windings are phased such that a large driver current is forced through a high impedance, i.e. the primary winding impedance which is a multiple of the common base input impedance of the output transistor, resulting in power gain. The amplifier is capable of high maximum overloaded power output free of "break-up" of the waveform compared with conventional cascode operation, as well as greater power gain. By using transistors with high resistivity base material the amplifier operation is resistant to degradation by ionizing radiation and protected against secondary breakdown.

2 Claims, 5 Drawing Figures



BASIC RECEIVER DESIGN

Jim Eagleson
IDX, Inc.

January 1985

This paper is going to cover the basics of receiver design.

It includes most parameters that should be looked at to evaluate whether a particular design will work for the intended purpose, in the likely environment, and for the desired cost.

Included in the Appendix is a simple BASIC program for use with a small computer (in this case, a TIMEX 1000). It should be readily adaptable to any machine supporting BASIC.

DESIGN GOAL

The very first thing that must be discussed is not really technical at all. The designer must sit down with management and/or the customer and determine the design **goal.**

While this should go without saying, the success or failure of a given design may not relate to the skill of the design engineer nearly as much as it does to how well the design goal was defined before starting.

This is especially true of cost goals since most engineers will tend to over design if they are not sure how much more the unit will be required to do beyond the original goals.

The following items should be considered:

1) **How much should the product cost?**

2) **What must the unit do?**

3) **What must the unit NOT do?**
(Could be as important as 2)

4) **What does it fit inside or mate with?**

5) **What should it look like?**
(Military, Commercial, Consumer)

6) **What kind of environment will it see?**
(Temperature, Humidity, Vibrations, Chemicals)

7) **What kind of person will be using it?**
(Consumer, Industrial, Technical)

8) **What level of servicing will it see?**
(In-place, Local, Regional, Factory Service)

9) **When is pre-production prototyping due?**

10) **Is this a preliminary or permanent design?**

SPECIFICS REQUIRED

In the "what must the unit do" category, the engineer may need to make somwhat of a nuissance of himself before he can get everyone who must use the end product to clearly define their specific needs.

All the same, he must know:

Voltages available and power drain allowances

Mechanical constraints - most RF circuits are layout sensitive. They are not always conveniently changed in mid-stream...especially at VHF and above!

Output Requirements - including **levels, impedances, video response, distortion allowance, etc.**

Interconnection Requirements - including **cables, connectors, teminals, etc.**

Control Requirements - including **external, internal, manual, automatic, trimmer** and control **signal levels** if from other sources.

Environmental Specifications - Temperature, Humidity, Moisture, Dust, Chemicals, Vibration, Shock.

Electrical/RF Environment - IMD, CRSS, Desense, Adjacent Channel, Co-channel Interference potentials. Power source conditioning, EMI filtering, Surge Protection, etc.

Cosmetic Constraints - While this could be placed under Mechanical Constraints, often even the look **inside** a box can be important to its effectiveness (in this case in saleability). Certainly, cosmetics of the outside of the box can impact the placement of controls and connectors. Don't be fooled! Even in Military markets the look can make a difference. The look may be different, but the requirement that it "look right" is the same!

EXAMPLE OF PRE-DEFINITION

Unlike some of the cheap, built-by-the-millions Taiwanese or Japanese portable "Ghetto-Blaster" FM radios, the receiver in an FM Broadcast translator is much more demanding in its environmental and performance preevaluation.

Areas of concern here are:

1) The receiver must be able to handle the presence of a **1 Watt** or a **10 Watt** transmitter located in the same box. This demands good **isolation** and **overload capability.**

2) The receiver is likely to see **strong adjacent channel signals.** This requires good **selectivity.**

3) The receiver will often be **located near other strong broadcast transmitters.** Again, requires good **isolation** and **overload capability.**

4) The entire translator will typically be designed for **minimum power drain.** Significant numbers will find service on **solar** or **propane** power.

5) For the reasons in **4,** capability for running the unit from **13.8** or **28 Volts** should be a goal.

6) Since the site will often be **remote, reliability** and **ease of field servicing** is important. **Protection against lightning damage** should be provided (both antenna strike and power surge damage).

7) **FCC Type-Approval is required.** This has definite impact on **Frequency Stability, Selectivity** (don't want to retransmit other than the desired station), and **Gain Stability.**

8) For ease of maintenance and field service, **Full Metering of Signal Level, PA Status,** and **VSWR** should be provided.

Obviously, the translator's receiver is going to be some 50-100 times the cost of the mass-produced "Ghetto-Blaster"!

APPLYING DESIGN GOALS TO BASIC DESIGN

Every receiver has certain characteristics which must be looked at in relationship to its application.

These characteristics are:

1) **Input Frequency (FI)**
2) **Intermediate Frequency(ies) (IF, IF1, IF2..)**
3) **Image Frequency (IMG)**
4) **Frequency Tolerance (TOL)**
5) **Second Order Intermodulation (IM2)**
6) **Third Order Intermodulation (IM3)**
7) **Crossmodulation (CRSS)**
8) **Compression Point (CP, CP1, CP3)** (Usually referenced to 1dB or 3dB limiting)
9) **Desense, Blocking** (Can relate to CP, IF limiting in stages prior to filtering, or AGC depression by strong signals)
10) **Noise Figure (NF)**
11) **Noise Bandwidth (dB)**
12) **Selectivity (B / Xdb)**
13) **Half-IF and One-Third IF (IF/2, IF/3)**

Related to the above characteristics:

1) **Minimum Discernable Signal (MDS)**
2) **Tangential Signal Sensitivity (TSS)**
3) **Sensitivity (Xuv/Ydb SNR)** (Usually established at the minimum useful SNR for a given mode.)
4) **Dynamic Range (Xdb)**
5) **Intermodulation Dynamic Range (IMDR, IMDR2, IMDR3)** (Sometimes called Spurious Free Dynamic Range.)
6) **External Spurious Responses**
7) **Internal Spurious Responses**

Also worth considering:

1) **AGC Threshold**
2) **AGC Control Range**
3) **AGC Attack/Release/Hold Times**
4) **IF Limiting Threshold** (Typically 1dB or 3dB point)
5) **Log/Linear Tracking** (dB/V variation in a Log IF)
6) **Detector's Video Bandwidth**

TWO DESIGNS - TWO APPROACHES

A look at two different designs...one complex and one simple, will provide insight into how the characteritics relate to the design.

The first design is for an ACSB transceiver for VHF Land Mobile use.

WHAT DOES THE DESIRED OUTPUT LOOK LIKE?

The output from an ACSB receiver is communications quality audio. Thus we are looking for a fairly flat, low distortion response from 300-2700 Hertz in the audio channel with detector output including frequencies up to 3200 Hertz.

For the sake of simplicity I am going to assume that we are looking at a video response, then, of 3200 Hz or so. (Actually, with ACSB the noise bandwidth is narrower than this due to De-emphasis)

Given the above:

1) **Video Response:**

a) **Most Audio devices would be O-K.**

b) **Need at least 5 Watts AF Output for mobile.**

c) **Need high level speaker** (which can impact the mechanical design due to size).

2) **Mode Impacts:**

a) **IF Bandpass must be 3.2 KHz.** ACSB has a Pilot Tone plus audio.

b) **IF must be Linear.** Some form of AGC control must be provided to keep the RF and the IF stages in their optimum operating region. This must also keep signals at the detector within a 30-35 dB dynamic range so that the ACSB processor can keep within its control range.

c) **The RF AGC must have Fast Attack** to limit overload prior to processing. It should also have a **moderately fast release time** for following of moderate fades but long enough not to cause instabilities associated with **feedback AGC systems.**

d) **Requires high Adjacent Channel Selectivity.** Since the whole point of ACSB is the use of 5 KHz seperated channels, selectivity is a prime design goal. If our -6dB response is going to be 3.2 KHz, it is probably a good idea to use a filter with a 1.8:1 Shape Factor (Battenuate/Bpass...6/60dB). This will give us 1.8 X 3.2 or 5.76 KHz bandwidth at - 60 dB. This gives us 60 dB of Adjacent Channel protection on all signals further than ± 2.88 KHz from the center of the desired channel.

e) **Requires tight frequency** tolerance. Current ACSB designs can pull in signals from ± 800 Hz or so. Thus our transmit to receive drift must stay under 800 Hz if pull-in is to be reliable. This means that each must stay within ± 400 Hz in case one drifts down at the same time the other drifts up in frequency.

Furthermore, Adjacent Channel Selectivity will be affected by significant drift beyond a few hundred Hertz.

This suggests the use of **Crystal Ovens** and/or **Temperature Controlled Crystal** Oscillators(TCXO).

f) **AGC detection and Pilot Tone processing suggest high level detection requiring moderate RF/IF gain...100-120dB.** Furthermore the design of the Phase locking of the Pilot tone is best done at a single frequency at the second Local Oscillator/Mixer stage.

For best performance of the Front End the gain in this section should be restricted to less than 20 dB. Thus the RF is 20 and the IF is 100 dB.

FRONT END DESIGN

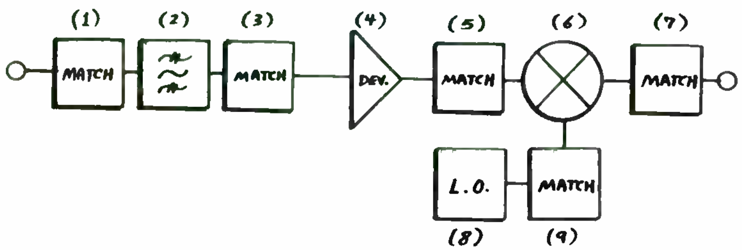
Since we now know what we expect to come out of the receiver and the approximate gain distribution, bandwidth required (IF), and the intended use (Land Mobile), we can define the **Front End** of the receiver.

We'll fist pick a design goal for **Sensitivity.** Typical Land Mobile FM equipment yields 0.25uV/20dB Quieting. We'll select this as our design goal for the ACSB radio.

Due to the processing used in ACSB, 0.25uV/20dB "Quieting" occurs at an input SNR of about 7dB in an ACSB system. Thus we would have a **Sensitivity of 0.25uV/7dB SNR** or, more coventionally, **0.35uV/10dB SNR.**

Also related to the application (Land Mobile) is the RF Bandwidth required. Typically a given radio will need to cover several channels within 2 MHz of each other. Thus, we will require a minimum **RF Bandwidth** of 2-2.5 MHz...also a good design goal for other reasons.

FRONT END BLOCK DIAGRAM



As illustrated, the **Front End** (or **First Converter**) consists of several elements:

1) **Antenna Matching**
2) **RF Band Pass Filtering**
3) **BPF to Device Matching**
4) **Device Gain**
5) **Device to Output Matching**
6) **Mixer (or Converter)**
7) **Mixer to IF Matching**
8) **Local Oscillator (Multipliers, Filters)**
9) **L.O. to Mixer Matching**

Each of these elements will have some level of impact on the performance of the **Front End.** Usually, however, there will be some level of integration of these elements into multiple function circuits. This can simply be a resonant L-C tuned circuit tapped to match the antenna and the amplifying device, or it could be a more complex, multi-pole Band Pass Filter or a Helical Resonator.

Since the design application is ACSB, ie, Land Mobile, and since this mode is most likely to be required in markets already too crowded for wideband modes, it seems likely that **Helical Resonator** or **Multiple Section Band Pass Filtering** will be used. Naturally, integration of antenna match and device match into this filter would be a design goal. Further, expanding the passband to 2.5 MHz to make alignment simpler and ensuring that the unit can be tuned to **any** 2 MHz segment of the 150-165 MHz spectrum is also a goal. Each of these factors influences the choice of filter style.

DEVICE SELECTION

Now comes the heart of Front End design. In order to see what kind of device we can use for the active gain stage, we must evaluate several things.

The first item is required **Noise Figure.**

NOISE FIGURE

We have established that the **Sensitivity** required is **0.35uV/10dB SNR.**

This translates to a level 10dB below 0.35uV (-116dBm) or **-126 dBm.**

Put another way, the **Equivalent Noise Input** in the **specified 3.2 KHz bandwidth is -126 dBm.** The formula for this is:

$$Ne = -174\text{dBm/Hz} + (10 \text{ Log } Bn) + NF$$

Where Ne is Noise(equivalent) (dBm)
Bn is Noise Bandwidth (Hz)
NF is Noise Figure (dB)

To solve for the required **Noise Figure,** however, requires re-arrangement of the formula to:

$$NF = Ne - (-174\text{dB/Hz} + (10 \text{ Log } Bn))$$

Thus for the ACSB radio:

$$NF = (-126) - (-174 + 35) = (-126) - (-139)$$
$$= 139 - 126 = \mathbf{13\ dB\ Noise\ Figure}$$

But, in reality, we need to know the **Front End Noise Figure.** What we have just calculated is the total **System Noise Figure.**

IF NOISE FIGURE

First, however, we will need to determine what the IF/Detector **Noise Figure** is going to be.

Typically, the IF Noise Figure will generally fall in the 6-13 dB region. It can be worse. It can be better.

Most Integrated Circuit IF devices have noise figures in the 7-9 dB area and most IF Filters will show 3-10 dB loss, depending on kind and frequency. Thus we should see 10-19 dB overall Noise Figure in the IF with IC Amps but somewhat better results...6-16 dB Noise Figure, using common transistors or FETS.

Since for **ACSB** we are using a highly selective filter (3.2 KHz), we will probably use a **Crystal Filter** having a typical loss of 6 dB. Since we are keeping Front End gain at a minimum, an IF Noise Figure in the stages after the filter should be kept in the 4-5 dB region. This is easy to achieve with FETS or MOSFETS.

After Second Coversion we can use an IC gain block such as the MC1350AP since there is enough gain in front of it by this point in the system to overcome its 8 dB Noise Figure.

Thus our **ACSB** radio's total IF Noise Figure will be in the 10-11 dB region.

CALCULATING NOISE FIGURE

The formula for calculating Noise Figure is:

$$F(system) = F1 + ((F2 - 1) / G1)$$

Where F1 is First Stage Noise **Factor** (ratio, not dB)
F2 is Second Stage Noise **Factor** (ratio)
G1 is First Stage Gain (ratio)
and F(system) is System Noise **Factor** (not dB)

But, of course, we want **F1**, not **F(system)**. Thus we re-arrange the formula to:

$$F1 = F(system) - ((F2 - 1) / G1)$$

Now all that we need to do is to convert our previous NF(system) and G1 from dB into ratios. We can use a chart for this or use the following relationships:

$$F = 10^{(NF/10)}$$

$$G = 10^{(G/10)}$$

Where the F and G in the exponent are in dB and the F and G after conversion are in power ratios.

Actually, in the example given these conversions are quite easy to do in ones head since:

13 dB = 10 dB + 3 dB or **10 X** and **2 X** power.

thus 13 dB is 2 X 10 times or **20** times the power **ratio.**

Similarly, **20** dB is 10 dB + 10 dB or 10 X 10 or **100 times the power ratio.**

By keeping 1dB, 2dB, 3dB, and 10dB ratios in ones head, almost any power (or voltage/current) ratio can be estimated quite easily.

Plugging in our data in the formula for **F1** we get:

$$F1 = 20 - ((10 - 1) / 100) = 20 - 0.09 = 19.91$$

or very close to 13 dB Noise Figure.

Even if the IF had a 16 dB Noise Figure we would get:

$$F1 = 20 - ((40 - 1) / 100) = 20 - 0.39 = 19.61$$

which comes out to 12.93 dB Noise Figure, still very close to our 13 dB requirement!

Obviously, IF Noise Figure is not the predominant factor when moderately high gain is available in the Front End. In fact, as a rule-of-thumb one can assume that if First Stage Gain (G1) exceeds the Second Stage Noise Figure by around **10** dB, the First Stage Noise Figure will establish the System Noise Figure within a few tenths of a dB. This is insignificant at VHF frequencies and probably not significant for mobile operation even at UHF where other factors such as external noise, flutter, etc. have greater impact.

NOISE VERSUS BANDWIDTH

Just in case it was not apparent from the formula previously used, **Equivalent Noise Input (Ne)** is directly proportional to the **Bandwidth used.** If one doubles the bandwidth, noise output doubles (up 3 dB).

By example, in comparing an **FM** receiver with an ACSB receiver, the difference in bandwidth is 18 KHz to 3.2 KHz. Plugging into our Ne formula:

Ne = -174 + (10 Log 18000) + NF = -118.5 dBm for FM
Ne = -174 + (10 Log 3200) + Nf = -126.0 dBm for ACSB

Put another way, the difference in Ne relates directly to Bandwidth ratios. Thus:

SNR improvement = 10 Log B1 / B2
= 10 Log 18000/3200
= 10 Log 5.63
= **7.5 dB SNR Improvement**

Since the **SNR Improvement Threshold for FM** and **ACSB** are both in the **10dB** CNR region, theoretically **ACSB** should have a 7.5 dB advantage on **FM, given the same Noise figure.** In fact, field tests show a frequent extension of operating range using this mode though other factors modify the theoretical 2:1 range extension that should occur to only 10-20%. This is largely due to the fact that ACSB has only 1/4 the average power of **FM.**

CNR, by the way, is **Carrier-to-Noise-Ratio** or the SNR coming in to the detector. The actual audio SNR after detection and/or processing will be higher for these two modes, however.

INPUT LOSSES - BAND PASS FILTER

Since our Noise Figure requirement is hardly difficult to achieve with modern devices even at UHF frequencies, we have a broad selection of transistors and FETS to choose from.

Before setting our Noise Figure spec to 10-11 dB for this device, however, we must come to terms with the Antenna Matching and Band Pass Filter requirements.

We have suggested that the input bandpass filter should be 2 MHz wide (perhaps 2.5 MHz to simplify tuning). It should have good out-of-band rejection since Land Mobile frequencies are very congested.

This suggests use of either a **Helical Resonator,** which at these frequencies would be moderately large, physically, or a **Multi-Section Band Pass Filter.** Helical Resonators will have between 0.5-1.5 dB losses, typically. Multi-Section L-C filters will have between 1.0-6.0 dB losses depending on bandwidth, component "Q", and required rejection.

Since we only need a 13 dB Noise Figure, the filter loss is not particularly important. Like all good engineers, however, it is probably a good idea to give ourselves a 3 dB "fudge factor" so that we'll aim for 10 dB Noise Figure to ensure our 13 dB goal.

It is likely that a reasonable filter can be made with less than a 4 dB loss. As this loss adds directly to the device noise figure, to give us the desired 10 dB Noise Figure we will then need a device having a 6 dB Noise Figure (6 + 4 = 10 dB).

Since devices with 6 dB Noise Figures are readily available at VHF frequencies, we should now define other factors affecting device choice.

DEVICE LIMITATIONS

All amplifiers have limits. They generate internal noise due to thermal effects. They have power limitations set by their voltage and current capability.

If an amplifier is bias for a certain power drain, it can only deliver some percentage of that power to a load. Just how much power it can deliver depends on device efficiency and matching efficiency between the device impedance and the load impedance.

Obviously, Front End devices must be biased in the "Class A" mode of operation. This is required if amplification is to occur down to the noise level of the device.

Unfortunately, "Class A" is not the most efficient bias point we could use. We only see 5-30% efficiencies, typically.

The problem, of course, is that the power handling capability at the high end of the scale can only be achieved by increasing the quiescent (or idle) current. Increased device current, naturally, increases device temperature which also increases the amount of noise generated internal to the device and thus increases its Noise Figure.

So there is always some level of tradeoff between Noise Figure and high level signal performance. In portable equipment, we may also have a further limitation in power drain.

BLOCKING, DESENSE, COMPRESSION POINT

Related to power handling is what an amplifier will do in the presence of a weak desired signal and an undesired strong signal.

Obviously, if the strong signal is strong enough it will overload the amplifier effectively stealing power from the weaker signal. We call this **Gain Compression.**

When such a signal reduces the amplifier gain by 1 dB we call that the **1 dB Gain Compression Point (CP1).** (Some reference the Compression Point to a gain reduction of 3 dB (CP3)).

The Compression Point may be referenced to input or output levels. For communications circuits we are mainly concerned with input **Compression Point** though most data sheets will give **output Compression Point** in their specs. It is a simple matter to subtract out the device gain (not stage gain which could be lower due to matching losses) in order to derive the input CP.

EXAMPLE:

J310 FET at 10 mA has an output CP1 = **10 dBm**
" " " " has a gain = **14 dB**

= **-4 dBm**

SECOND ORDER INTERMODULATION

When **two** (or more) very strong signals are present in an amplifier, at some signal level they will modulated each other producing sum and difference products. These products will occur at some level below the level of the two tones producing them. The **exact** level of these products is proportional to the level of the original signals (though device characteristics may modify the curve).

When the **fundamental** of one signal mixes with the **fundamental** of another signal we call this **Second Order Intermodulation (IM2).** In fact, in one sense a **mixer** could be considered to be **intentionally producing Second Order Intemodulation** (FI - LO = IF). In this special case, however, we use a **very large** Local Oscillator "signal" to mix with our weak to moderate **RF Input** signal.

Thus, **IM2** can occur when two **in-band** signals mix to create a signal falling on the input to the **IF Amplifier.**

Alternately, two signals well removed from the desired **RF Input** signal can mix together to produce an **IM2** product falling at the **RF Input** frequency. In most cases, however, Front End Selectivity does not allow such widely dispersed signals to enter the amplifier. A special case would be in very broadband receivers using no preselection.

EXAMPLES:

A broadband receiver covering 500-2000 MHz.

Desired Channel:	**1000 MHz**
Signal # 1:	**1700 MHz**
Signal # 2:	**700 MHz**

1700+700 = 2400 MHz.............no problem.
1700-700 = 1000 MHz........on desired chan.

Our ACSB receiver has an input filter which is 2.5 MHz wide at - 6 dB.

Desired Channel:	**155.000 MHz**
Signal # 1:	**160.700 MHz**
Signal # 2:	**150.000 MHz**

160.700 + 150.000 = 310.700 MHz...no problem.
160.700 + 150.000 = 10.700 MHz....potential problem.

Two things are in our favor in the ACSB receiver. First, both Signal #1 and Signal #2 fall outside of our RF Band Pass Filter. They will most likely see 15-20 dB attenuation due to the BPF selectivity.

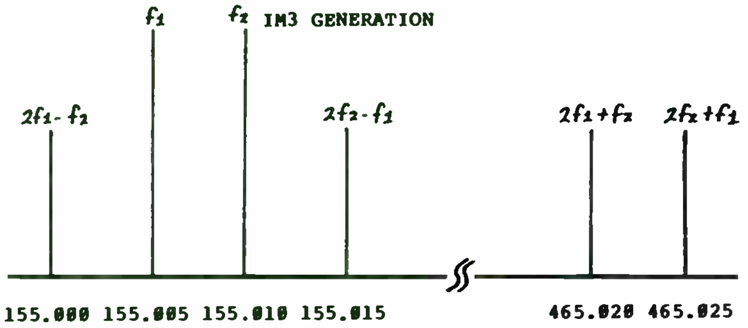
Furthermore, a properly designed Mixer will not pass signals coming into it at the IF frequency very well. Isolation will be in the 15-30 dB region in typical VHF mixers. (This may not be true of some Bipolar or single-diode mixers designs, however.)

At any rate, IM2 is not normally a significant factor except in certain broadband designs.

THIRD ORDER INTERMODULATION

If two very strong signals are present in our RF Amplifier they will also cause **Third Order Intermodulation Products (IM3).** Again these products will be at a level below the level of the distortion causing signals by an amount proportional to the levels of the IM generating tones. The specific ratio is **3:1.** In other words, as the IM generating tones go up by 1 dB, the IM3 products will increase by 3 dB.

However, in the IM3 case these products fall just above the higher tone and just below the frequency of the lower tone by an amount exactly equal to the frequency difference between the two original signals. Thus two signals which fall +5 KHz and +10 KHz above our desired ACSB channel (at 155.005 and 155.010) would generate IM3 products falling at 0 (5 KHz below the lower tone) and at +15 KHz (5 KHz above the higher signal).

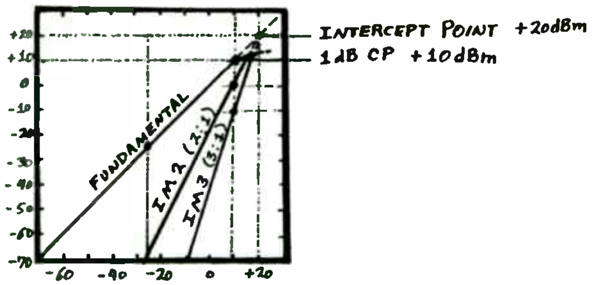


Signals will fall at 2F1 - F2, 2F2 - F1, 2F1 + F2, and 2F2 + F1. The additive signals obviously fall way above the usual passband of most receivers and, like IM2, can generally be ignored.

INTERCEPT POINT

The concept of Intercept Point is useful in predicting IM2 and IM3 levels. If we graph Fundamental response along with IM2 and IM3 products it becomes very apparent that the 1:1, 1:2, and 1:3 "curves" will ultimately meet at some point above the 1 dB Compression Point. In reality, of course, the output levels will compress beyond the CP1 so that this theoretical meeting of curves can never occur.

Fundamental, IM2, and IM3



In the "real world" one may find variations in these theoretical curves and in the IM2 and IM3 Intercepts relative to the Fundamental. For general evaluation, however, assuming a common Intercept Point about 10-15 dB above the 1 dB Compression Point is appropriate.

(It would be a good idea to point out, here, that CP1 may occur due to other factors if the measurements are made in a total receiver. Obviously, leakage around filters can allow strong enough signals to depress the AGC. Similarly, stages of IF amplification prior to the IF filter and also the mixer stage can go into limiting...another word for "gain compression." (This latter problem is common in Automobile FM receivers.)

So we can take the published data and find the IM2 and IM3 information, if given, or we can extrapolate it from the CP1 or CP3 data (usually given). Merely add 10-15 dB to CP1 or 7-12 dB to CP3. It is best to work with the lower level (10 or 7 dB) for conservative estimates.

Again, the same caution as before...we are interested in input, not output levels. The input CP1 or CP3 levels will be lower by the amount of the device gain.

EXAMPLE:

CP1 = +10 dBm (J 310 FET output)
IP = +20 dBm (CP1 + 10 dB output)

but...

IP(input) = IP - G = 20 - 14 dB = +6 dBm

IM2, IM3 FORMULAS

At the Intercept Point (theoretically) IM2 will be equal to the levels of the two tones producing the IM2 products. At a 1:2 "curve", if the level of the generating tones falls 20 dB (to -14 dBm in our J 310 case), the IM2 levels will fall 40 dB (to -34 dBm), or twice as far.

Thus:

IM2 (dBm) = IP2 - 2 (IP - IM generating tones)
= +6 - 2 (+6 - (-14)) = +6 2 (20)
= +6 -40 = -34 dBm

and:

IM3 (dBm) = IP3 - 3 (IP - IM generating tones)
= +6 - 3 (+6 - (-14)) = +6 - 3 (20)
= +6 - 60 = -54 dBm

The opposite is also true, of course. If the tones increase by 20 dB, the IM2 will increase by 40 dB and the IM3 will increase by 60 dB.

Also clear is that signals at the high levels used in the illustration could easily be seen at a broadcast site where output power can be in the tens of kilowatts (+60 to +80 dBm). Units designed for this kind of use had better have good Intercept Point performance or be placed in a very well shielded box with a well designed Notch Filter on the frequency of any transmitters sharing the site. It would take 74 to 94 dB of isolation to lower the transmitter levels to - 14 dBm!

INTERMODULATION DYNAMIC RANGE

Essentially Dynamic Range is the ratio between some pre-determined reference (say 12 dB SINAD) and the level of signal required to degrade the desired signal by some amount (in the SINAD case, by 6 dB).

For design purposes it is much easier to talk about **IMDR , or Intermodulation Dynamic Range.**

The most common definition for **IMDR** is the ratio between the Equivalent Noise Input level (in this case usually called the **MDS** or **Minimum Discernable Signal** level) and the level of two tones required to generate 2nd Order or 3rd Order IM products at the same (MDS) level.

Obviously, any IM2 or IM3 products falling below the MDS level will not be heard. Any above this level will have a direct affect on the SNR of the desired signal.

MDS, has the same formula as Ne:

$$MDS = -174 \text{ dBm/Hz} + (10 \text{ Log Bn}) + NF$$

We can estimate our Intercept Point from the CP1 point of the device being used:

J 310 = + 10 dBm CP1
= + 20 dBm IP

Since IM2 products fall off at a 2:1 rate, the 2nd Order IMDR will occur half way between IP2 and MDS. That is, if the IM2-creating tones are at a level half-way between IP2 and MDS, the resulting IM2 distortion products will be at a level just equal to the MDS level. IMDR is the dB ratio required above MDS to generate IM2 tones just at

Thus:

$$\begin{aligned} IMDR2 &= (IP2 - MDS) / 2 \\ &= (+6 - (-126)) / 2 \\ &= 132 / 2 = 66 \text{ dB} \end{aligned}$$

However, we previously pointed out that one set of IM2 products falls so far outside our Input BPF that they can be ignored and the other IM2 generating signals would be attenuated some 15-20 dB by the BPF. Thus we can safely add another 15 dB or so to the IMDR2 figure making it 81 dB.

Thus it would take two tones 81 dB above - 126 dBm to create an audible IM2 product. This is a level of -126 + 81 or **-45 dBM (1260 uV)...a healthy set of signals** not likely to occur too often at just the right frequencies for our IM2 situation (particularly when one throws in Mixer isolation).

IMDR3, on the other hand, could be much more of a problem.

Even though we have a 3:1 drop off in IM3 products versus input signals, we do not have the selectivity working for us like we do in the IM2 case. In Land Mobile urban environments there will be many instances when signals will be spaced the correct distance apart in frequency to cause this problem.

These signals could be one channel up plus two channels up, two channels up plus four channels up, four channels up plus eight channels up (or down) from the desired channel and so forth.

Like IMDR2 we can relate IMDR3 to the Intercept Point.

Since IM3 drops off at a 3:1 rate, two tones falling 2/3 of the way between Intercept Point and Minimum Discernable Signal level will just cause IM3 products equal to the MDS level.

Thus:

$$\begin{aligned} IMDR3 &= 2 (IP3 - MDS) / 3 \\ &= 2 (+6 - (-126) / 3 = 2 (132) / 3 \\ &= 264 / 3 = 88 \text{ dB} \end{aligned}$$

One last factor should be put into the formula if we are going to get an accurate picture of actual IMDR3. Our Input Band Pass Filter has about 4 dB of insertion loss. This affects both IMDR2 and IMDR3 by 4 dB.

Thus our IMDR2 will become 85 dB and IMDR3 becomes 92 dB. This means that two signals must exceed **-41 dBm(2000uV)** to create IM2 and would need to exceed **-34 dBm(4500 uV)** to create IM3 products which would be audible above the receiver's own noise level.

HALF-IF, ONE THIRD-IF

Half-IF and One Third-IF are related to IM2 performance. As a general rule, 2nd and or 3rd harmonics generated in an amplifier by a single tone will have a level about 6 dB below what the IM2 levels would be for two tones at that same level.

Thus a signal at half the frequency of our desired input frequency would double in the RF Amplifier creating its second harmonic on the input channel. Of course, this does not happen in normal tuned amplifiers, but could happen in very broadband amplifiers.

A special case that can happen is when a signal is removed from the Local Oscillator frequency of the receiver by one-half or one-third of the IF frequency. In other words, given a 10.7 MHz IF, any signal 5.35 or 3.5667 MHz above or below the L.O. frequency will be converted to one-half or one-third of the IF frequency by the mixer. If this signal is strong enough and IF selectivity prior to the IF Amplifier(s) is not narrow enough to reject this signal (often the case in older, double-conversion designs which used only 455 KHz Ceramic Filters), the IF stage (and, in some cases, the mixer) will create the second or third harmonic of the signal. This, of course, will fall right in the middle of the IF and will interfere with desired signal.

This is not likely to occur with modern designs using selectivity immediately after the mixer (except with some active mixers, perhaps).

IMAGE FREQUENCY

Any mixing scheme will have two RF Inputs which will mix with the L.O. to create the IF frequency. One will be L.O. + IF above the L.O. and one will be L.O. - IF below the L.O.

EXAMPLE:

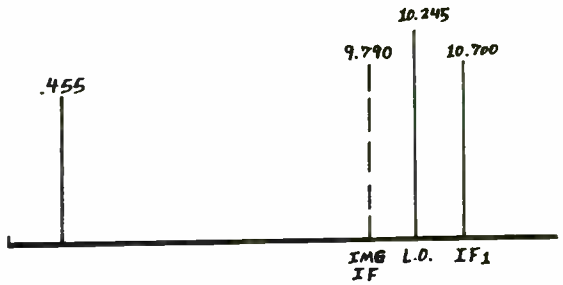
Our ACSB radio receives, say, 155.700 using an L.O. of 145.000 to obtain an IF of 10.700 MHz. The Image Frequency would be 145.000 - 10.700 or 134.300 MHz...not much of a problem with our BPF unless its ultimate rejection fails us at this frequency for some reason.

It is probably wise to check for the Image Frequency since it is so far removed from our filter passband that skirt selectivity (ultimate rejection) could see variations in simpler filters).

IF IMAGE

Once again, the double conversion radio could have a case where a signal only, say, 910 KHz away from the desired signal (910 KHz = 455 x 2) would be converted to 9.790 MHz. If our second IF is 455 KHz, the desired 10.700 MHz signal requires a 2nd L.O. of 10.700 - .455 or 10.245 MHz. But, 10.245 MHz - 9.790 MHz also give us .455 MHz (455 KHz). Obviously, the selectivity of the 10.7 MHz stage must be narrower than 2 x 910 KHz or 1820 KHz (assuming the IF is peaked at 10.7 and is symmetrical at plus and minus 910 KHz).

IF IMAGE



As before, however, use of a selective filter in this stage totally prevents the problem. Monolythic crystal filters are inexpensive enough to make this problem obsolete in modern designs.

In a given case when this occurs, it can be eliminated merely by shifting the 2nd L.O. to the other side of the input signal. In other words, if low side injection is being used (10.245 for a 10.7 MHz input), high side injection at 11.155 MHz would move the IF Image to 11.610 MHz. Providing that no interfering signal converts to this frequency, this move would cure the problem.

A word of caution, however. This cures the problem for FM or AM since the sideband relationships (upper versus lower) are unimportant in double-sideband modes. With USB or LSB one would also need to move the BFO oscillator to the opposite side of the IF passband to restore proper carrier to sideband relationships. The same is true of UHF to VHF conversions of TV signals. Since the audio sub-carrier is above the visual carrier, use of alternate side L.O. injection to prevent IF Image problems would not work!

CROSSMODULATION

Crossmodulation is a term that seems to be defined differently by different people.

The original concept seems to describe the condition where one strong signal imparts its modulation to a weaker signal on the desired channel.

This, of course, is related to the Gain Compression characteristic of the amplifier and is best applied to AM systems.

Some have also used the term with reference to out-of-band signals mixing together to produce an in-band IM product. This, of course, is really IM2 or IM3.

Of course, in FM systems the IM2 or IM3 product of two unequal signals where one is much stronger than the other can also create a similar condition. FM's capture effect will suppress the weaker of the two IM sources so that it appears that only one signal is causing the problem.

This would be intermittent, however, since both signals would not be present simultaneously all the time. Then again if one was a broadcast station (TV or FM) which would be present most of the time it could appear that the problem is CRSS.

SENSITIVITY

The term **Sensitivity** has various definitions depending on mode and/or signal-to-noise requirements. With FM systems it is usually specified as a specific level of microvolt input for a given background noise **quieting**.

Example: **0.25uV / 20 dB Quieting** (Communications)
1.35uV / 30 dB Quieting (FM Broadcasting)

These represent the lowest SNR levels considered "useful" to the average user for the intended purpose.

With **AM, Pulse, SSB** or other **Amplitude Modulated** modes the typical measurements are:

TSS = Tangential Signal Sensitivity

MDS = Minimum Discernable Signal

SNR = Microvolts input for a given SNR output

TSS is the level at which the envelope of the signal plus noise just doubles the level (peak-to-peak) of the noise without the signal present. This is a 6dB peak-to-peak SNR or 8dB peak-to-peak signal + noise to RMS noise level.

The TSS level represents the threshold of operation for some systems. Mainly, however, it represents a convenient reference point when using an oscilloscope!

MDS is the level where the signal is just detectable in the noise. Theoretically it represents 0dB SNR, but this discernment by ear or with a scope varies so much from person to person that it is really mainly a reference used for theoretical calculations.

SNR specifications are an attempt to present a level at which a signal just becomes useful. This can be 1dB SNR for Morse Code, 6dB for SSB, 12dB SINAD on NBFM, or 20dB on Broadcast FM.

Some references are the level at which the signal becomes "fully useful". This could be 20dB SNR on AM or SSB, 20 dB quieting on NBFM, or 30dB quieting on Broadcast FM.

As you can see, there is some level of ambiguity associated with Sensitivity specs. The key word is "standard reference"...the use of methods and reference SNR's which have "universal" acceptance.

SECOND RECEIVER DESIGN EXAMPLE

A quick look at a second receiver will show the kind of variations one finds due to application.

We wish to receive signals at 2000 MHz at minimum cost with commercial application in an amplitude modulated data transmission system with interface to a microprocessor controller.

Signal levels are in the -70dBm to -30dBm region.

FCC regulations on the associated transmitter allow it a ± 8 MHz frequency drift and we are to receive a signal at the 2nd Harmonic of the transmitter.

There are no FCC radiation requirements for the receiver since it is operating totally above 1000 MHz.

DESIGN IMPACTS OF SPECIFICATIONS

Again, starting from the receiver output:

1) **Video response = 500 KHz square wave, TTL** level

a) **Gain must be provided** to bring the detected signal level up to TTL level.

b) **Requires no power devices (TTL).**

c) **AM signal requires 2X data bandwidth.** A 500 KHz square wave has required harmonics at least to 10X or 5 MHz. Thus, the IF bandwidth must be at least twice this, or 10 MHz.

d) **Square wave response suggests a Gaussian IF response curve.**

e) **The allowed transmitter drift suggests a wider than required IF** selectivity...perhaps 20 MHz rather than 10 MHz.d

2) **Mode and frequency suggest:**

a) **The transmitter should be restricted to minimal drift, say** ± 3 MHz so that the required receiver IF will need only ± 6 MHz extra bandwidth to make up for transmitter drift. This will help keep sensitivity from being impacted adversely by too wide a bandwidth.

b) **The receiver should have minimal drift and should be designed to track the transmitter drift as much as possible.**

c) **The cheapest possible Front End should be designed since the 2 GHz portion is likely to be the most expensive part of the receiver.** This suggests use of a sub-harmonic mixer where the oscillator injection is at half the desired L.O. frequency.

d) **Sufficient IF gain to allow diode detection of the weakest desired signal is required.** For our purpose an SNR of 15 dB is adequate. Thus an MDS 15 dB below 60 dBm is required, or -75 dBm. For safety, design for MDS = -80 dBm (-74 dBm, TSS).

e) **Signal can be subject to very rapid fading up to a 10 mS rate.** Thus an A-C coupled Logarithmic Amplfier is suggested. This will ensure relatively constant P-P voltage output for a given modulation percentage regardless of received signal level. Some A-C coupled limiting after detection may also help present a more constant signal to the TTL converter at the output.

f) **An effect High Pass Filter and/or 1000 MHz Notch Filter should be used to prevent mixer overload by the transmitter.**

GAIN, NOISE FIGURE, AND FRONT END

Now we can figure out the Front End requirements.

Gain required relates to the minimum detector level required.

For a diode detector this will probably be in the 100 mV region at an impedance of around 500 ohms.

This represents a power level of:

$$P\ (mW) = ((\ 0.1\ )^2\ /\ 500\ )\ *\ 1000 = 0.02\ mW$$

$$P(dBm) = 10\ \text{Log}\ 0.02 = \mathbf{-16.9\ dBm}$$

Since our lowest level requiring detection is to be **-75 dBm**, the gain required is:

$$\mathbf{P(dBm) - P(MDS) = (-16.9) - 75 = 58\ dB}$$

We've suggested a harmonic mixer for economy and since typical L.O. requirements for this kind of mixer is in the -16 dBm region, the input CP1 will be about -22 dBm... 6 dB lower. Thus the IM3 level will be about -12 dBm.

Furthermore, this mixer is going to be lossy... -8 dB or so. Thus our IF gain will need to be 8 dB higher than calculated, or about **66 dB** (58 + 8 dB).

Since 2 GHz is not a band where it is likely that too many strong signals will be found in most applications of this equipment (certainly not as high as its own transmitter!), IM3 can probably be disregarded.

IMDR3 will be:

$$\mathbf{2 * (\ (-12dBm) - (-80dBm)\ )\ /\ 3 = 45\ dB}$$

This works out to an actual level 45 dB above -75 dBm, or -30 dBm per tone... a pretty healthy signal at 2 GHz!

For similar reasons, IM2 does not seem problematical. On the other hand, IM2 does provide us a look at 2nd Harmonic generation within the Front End and since this can be a problem for us we should take a close look at it.

We've previously stated that IM2 products will be about 6 dB higher than harmonics levels created in a given amplifier. Since we wish to have any harmonic of the transmitter fall below our MDS level of -75 dBm, this suggests that the signal from the transmitter must be attenuated so that its level falls below:

MDS + ((IP2 - MDS) / 2)
= ((-75 + ((-12)-(75))/2
= -75 + 31.5 = **-43.5 dBm**

Thus, the total isolation from TX output to RX input must be:

P(tx) - (-43.5 dBm)

In our design case the maximum transmitter power is 4 Watts (+36 dBm). Thus our isolation must be **36dBm - (-43.5dBm)** or **79.5 dB.**

IF FREQUENCY

The last problem is the choice of an IF frequency.

This is limited by two factors: **Required Bandwidth** and **Mixer Bandwidth.**

Since we need an IF bandwidth 10-20 MHz wide to accept the modulation and allow for oscillator drift, we will need to use a frequency at least twice as high as the bandwidth. Otherwise it can be difficult to filter out the IF component of the detected signal and gain changes over the passband can become serious when the percentage of bandwidth to center frequency becomes too high.

Thus we need at least a **40 MHz IF** and since **45 MHz** is a standard TV IF frequency, this seems like a good choice.

An IF at **60** or **100 MHz** might be better, of course, but our "cheap" mixer-only front end might not have the kind of isolation necessary to keep out strong TV or FM stations that can be on these frequencies in industrial (ie, urban) areas. Furthermore, it is best to keep the percent of bandwidth at the harmonic mixer as low as possible for best performance.

CONCLUSIONS

It is hoped that these sample designs and discussion of the parameters of a receiver which impact performance in a given situation has aided the reader.

One can copy a previous design, over design, under design, or he can take into account all of the things the unit is supposed to do (or not do) so that each parameter is balanced against cost, complexity, and need.

This can make the difference between an "acceptable" design and a good one.

BIBLIOGRAPHY

1. H. Paul Shuch, "Quiet! Preamp at Work", Ham Radio Magazine, November 1984 pp 14-20 (with computer program in basic)

2. W. Richardson, "IMD and Intercept Points of Cascaded Stages", Ham Radio Magazine, November 1984 pp 28-34 (with basic computer program)

3. J. Reisert, "VHF/UHF WORLD... high dynamic range receivers", Ham Radio Magazine, November 1984 pp 97-105 (with references)

4. J. Eagleson, "Linear Translators", Ham Radio Magazine, September 1983 (cover article)

5. J. Duncan, "BIRD CHART", Tele-Tech Corp. 1982, P.O. Box 1827, Bozeman, MT 59715

6. VARI-L Company, Application Notes, Double Balanced Mixers, 11101 East 51st Ave, Denver, CO 80239 (ph: (303)371-1560)

7. AVANTEK, "Designing With Modular Amplifiers" 1976, ATP-1035/8-77, 3175 Bower Ave., Santa Clara, CA (408)249-0700

8. Hewlett-Packard, Ap Note 57-1, "Fundamentals of RF and Microwave Noise Figure Measurement", 5952-8255, July, 1983

```
0 REM  B AND B ELECTRONICS
       RXD3N  OCT 1983
  5 LET A$="
"
  10 PRINT A$,,TAB 8,"RECEIVER D
ESIGN",,,A$,,"RECEIVE FREQUENCY=
";TAB 29;"MHZ",,"I.F. FREQUENCY
  =";TAB 29;"MHZ";AT 20,0;A$
  15 INPUT R
  20 PRINT AT 6,20;R
  25 INPUT I
  30 PRINT AT 8,20;I
  75 LET L=R-I
  76 LET LH=R+I
  60 PRINT ,,"LOCAL OSC.(L)     =
 ";TAB 20;L;" OR",TAB 20,LH;" MH
Z"
  85 LET L=R-I
  90 PRINT ,,"IF BANDWIDTH(B)   =
 ";TAB 29;"KHZ"
 100 INPUT B
 110 PRINT AT 13,20;B
 120 PRINT ,,"NOISE FIGURE(N)   =
 ";TAB 30;"DB"
 130 INPUT N
 140 PRINT AT 15,20;N
 150 PRINT ,,"COMPRESSION PT(C)=
 ";TAB 29;"DBM"
 160 INPUT C
 170 PRINT AT 17,20;C
 180 PRINT ,,"( CONT )"
 185 SLOW
 190 IF INKEY$="Z" THEN COPY
 195 IF INKEY$="C" THEN GOTO 205
 200 IF INKEY$<>"C" THEN GOTO 19
0
 205 FAST
 207 CLS
 210 PRINT ,,A$,,"IMAGE=";L-I;"
MHZ(LOW SIDE INJ)",TAB 6,LH+I;"
MHZ(HIGH SIDE INJ)"
 240 LET E=10*LN (1E3*B)/LN 10+N
-174
 250 PRINT ,,"EQUIV NOISE INPUT=
 ",,INT (E*10)/10;" DBM"
 360 PRINT ,,"DYNAMIC RANGE= ",I
NT ((C-E)*10)/10," DB"
 270 PRINT ,,"IM DYN. RANGE= ",I
NT (10*(C+12-E)*2/3)/10," DB"
 280 PRINT ,,"HALF IF= ",L+I/2;"
 AND ",,L-I/2;" MHZ"
 290 PRINT ,,"ONE-THIRD IF= ",L+
I/3;" AND ",,L-I/3," MHZ",,,A$
 295 STOP
 300 SAVE "RXDSN"
 305 FAST
 310 RUN
```


```
0 REM  B AND B ELECTRONICS
       RXD3N  OCT 1983
  5 LET A$="
"
  10 PRINT A$,,TAB 8,"RECEIVER D
ESIGN",,,A$,,"RECEIVE FREQUENCY=
";TAB 29;"MHZ",,"I.F. FREQUENCY
  =";TAB 29;"MHZ";AT 20,0;A$
  15 INPUT R
  20 PRINT AT 6,20;R
  25 INPUT I
  30 PRINT AT 8,20;I
  75 LET L=R-I
  76 LET LH=R+I
  80 PRINT ,,"LOCAL OSC.(L)     =
 ";TAB 20;L;" OR",TAB 20,LH;" MH
Z"
  85 LET L=R-I
  90 PRINT ,,"IF BANDWIDTH(B)   =
 ";TAB 29;"KHZ"
 100 INPUT B
 110 PRINT AT 13,20;B
 120 PRINT ,,"NOISE FIGURE(N)   =
 ";TAB 30;"DB"
 130 INPUT N
 140 PRINT AT 15,20;N
 150 PRINT ,,"COMPRESSION PT(C)=
 ";TAB 29;"DBM"
 160 INPUT C
 170 PRINT AT 17,20;C
 180 PRINT ,,"( CONT )"
 185 SLOW
 190 IF INKEY$="Z" THEN COPY
 195 IF INKEY$="C" THEN GOTO 205
 200 IF INKEY$<>"C" THEN GOTO 19
0
 205 FAST
 207 CLS
 210 PRINT ,,A$,,"IMAGE=";L-I;"
MHZ(LOW SIDE INJ)",TAB 6,LH+I;"
MHZ(HIGH SIDE INJ)"
 240 LET E=10*LN (1E3*B)/LN 10+N
-174
 250 PRINT ,,"EQUIV NOISE INPUT=
 ",,INT (E*10)/10;" DBM"
 260 PRINT ,,"DYNAMIC RANGE= ",I
NT ((C-E)*10)/10," DB"
 270 PRINT ,,"IM DYN. RANGE= ",I
NT (10*(C+12-E)*2/3)/10," DB"
 280 PRINT ,,"HALF IF= ",L+I/2;"
 AND ",,L-I/2;" MHZ"
 290 PRINT ,,"ONE-THIRD IF= ",L+
I/3;" AND ",,L-I/3," MHZ",,,A$
 295 STOP
 300 SAVE "RXDSN"
 305 FAST
 310 RUN
```

A SIMPLIFIED ANALYSIS OF THE SUPERREGENERATIVE RECEIVER

By

Robert C. Skar

Chamberlain Manufacturing Corporation
December 12, 1984

Although the super-regenerative receiver (SRR) still has many useful applications in consumer electronics products, radio control and amateur radio, little has been written about it in recent years, so many people in this audience may not be familiar with the concept. Also, most of the papers which have been written either assume outdated concepts such as the vacuum tube or concentrate on a specific solution to some of the many problems associated with this type of circuit. My primary purpose in this paper is not to reveal an elegant solution to all these problems (I don't have one) but to impart a basic understanding of how the circuit works so that the ingenious r-f designer will be able to apply those principles to present-day technology. However, for those who like formulas, a few of the more important equations are included, where applicable, and I also have included a bibliography of all the sources that I have been able to uncover in my research on the subject. Of these, by far the most important is a 170 page book written just after the World War II:

SUPER-REGENERATIVE RECEIVERS
by J. R. Whitehead
Cambridge University Press, 1950

This book is <u>the</u> definitive reference book on the subject as it was understood at the time and appears to arrive at mathematically rigorous conclusions consistent with its basic premises, but, unfortunately, many of the assumptions it makes in arriving at its conclusions are based on the characteristics of vacuum tubes and do not readily apply to devices like bipolar transistors. Therefore, it is more important to understand the derivation and meaning of the equations than it is to try to use them for quantitative calculations.

The first step in our understanding of this concept is to define it. In essence, the SRR consists of an r-f oscillator tuned to the carrier frequency of the desired signal. This oscillator is turned off, or "quenched" periodically either by an external modulator or by an internal circuit modification which will cause it to squeg or self-quench at a rate much lower than the carrier frequency but higher than the highest modulation frequency of the received signal. The quench frequency must be low enough so that any residual signal in the oscillator tank circuit will be below the noise level before the next on-period starts. Operation of the circuit depends on the basic fact that the oscillator will restart more quickly in the presence of an externally introduced signal than it would in the presence of low level random noise. This means that the following period of oscillation will either last longer (in the externally quenched case) or that the quench frequency will increase (in the self-quenched case).

At this point, one is tempted to draw and analyze a typical r-f oscillator consisting of an RLC tank circuit driven by some active device but to do so would destroy the generality I am trying to convey at this time. So, in order to present the concept in a way which I also believe will be more intuitively obvious, I have chosen the analogy of a pendulum modified in such a way that it will act as an externally quenched oscillator. This system is illustrated in Figure 1. The pendulum consists of a mass M attached to a light rod of length ℓ suspended from a fixed point by bearing B. We will assume that there is a small amount of friction associated with the system which results in a system Q defined by:

$$Q = 2\pi \frac{\text{maximum energy stored per cycle}}{\text{energy dissipated per cycle}}$$

If the pendulum moves an angle θ from vertical, the energy stored is $E = Mg\ell \sin^2\theta \approx Mg\ell\,\theta^2$ if θ is not too large and is expressed in radians. The period of the pendulum or time to complete one oscillation is

$$T = 2\pi\sqrt{\frac{\ell}{g}} \text{ seconds.}$$

Let us assume that the friction in the system is such that Q is approximately 100. It can be shown that if the pendulum is caused to start swinging at some initial amplitude and then allowed to coast to a stop, the amplitude will decrease to 1/e (36.79 percent) of its initial value after Q/π cycles and to 10^{-8} times its initial value after 5.86Q cycles. (We will assume that this is below the noise level; if the initial displacement of the mass were one inch, the final displacement would be about 10^{-8} inches and the stored energy 10^{-16} times its initial value.)

The pendulum is kept in motion by an electromechanical system consisting, first, of a motion sensor whose electrical output consists of a series of short positive pulses coinciding with the point where M is nearest to the motion sensor and whose amplitude is proportional to the displacement of M from its rest position.

The output of the motion sensor is connected to an output pulse integrator and to a gain controlled amplifier. The output pulse integrator will provide a d-c output proportional to the total charge contained in the pulses coming from the motion sensor during what we will call the quench period. The output of the gain controlled amplifier is used to drive a solenoidal electromagnet (S_1) positioned near a soft iron slug fastened to the pendulum shaft. The gain controlled amplifier can be gated on by an external signal generator whose waveshape we will assume here to be a square wave. Let us assume that the mass is initially at rest but that the gain controlled amplifier has just been turned on and has sufficient gain so that the system will slightly reinforce any motion of M. The motion of the pendulum will then build up exponentially as shown in Figure 2. It should be noted that the original amplitude was the noise level which, in this case, we will define as being less than 10^{-8} inch displacement.

The gain controlled amplifier has been designed to saturate before the mechanical system becomes non-linear. However, since we are illustrating what we will later call the linear externally quenched SRR, we will gate off the amplifier while it is still well below the limit of its linear gain region and the pendulum will never approach the limits of its swing. If we were illustrating the so-called logarithmic mode, the oscillation envelope would approach a fixed upper limit asymptotically. In either case, after the amplifier is turned off, the motion of the pendulum will damp out exponentially again. The amplifier will not be turned on again for a period of at least 5.86 Q times the pendulum period so that the motion of the pendulum is again less than 10^{-8} inch.

Now we will introduce another component into the system: an external pulse generator driving a second magnet, S_2. This is analogous to the introduction of a small external r-f signal in the case of the normal SRR. If the output of the external pulse generator is at the natural resonant frequency of the pendulum, and if the energy imparted to the system each cycle is greater than the energy lost, the amplitude will increase with time until the energy input per cycle is equal to the energy lost in friction, so that when the gain controlled amplifier is turned on again, the previous cycle will repeat, but will start at a higher level as shown in the dashed line of Figure 2. The difference in area between the two envelopes represents the output of the system.

We are now in a position to draw some conclusions that are applicable to an r-f system. First, let us consider the parameters of the pendulum itself, using physical intuition as our guide. It is immediately obvious that the mass of the pendulum should be as small as possible for maximum sensitivity since the total energy input required to accelerate the pendulum to a given displacement is directly proportional to its mass. If time is of no consequence, we also will obtain maximum sensitivity if the friction in the system is very low, which implies that a high-Q circuit is desirable; however, it should be very clear that high Q should <u>not</u> be obtained by increasing the pendulum mass. Since the pulsations from the external generator must remain in phase, it is also clear that high Q will result in a narrower band width. Other conclusions can be drawn concerning the rate of change and the amount of feedback (in this case, the characteristics of the gain controlled amplifier). The condition of the system shown in Figure 2 is where the amplifier gain is adjusted so that the energy added to the system by S_1 is exactly twice the energy dissipated by friction so that in electrical terms (which we will discuss later) the absolute value of the net negative resistance in the circuit during oscillation is the same as the positive resistance during the quench period. This results in the symetrical buildup and decay of the oscillation shown in Figure 2.

I am sure that we could do much more with the pendulum analogy, such as simulating a self-quenched oscillator, installing a damping system to shorten the recovery time and using multiple steps in the on-cycle to speed up the oscillation period, but time limitations dictate that we leave these variations to the r-f model. I wish to state again, however, that I believe the mechanical analogy makes clear some fundamental truths which sometimes are obscurred by the mathematics in Whitehead's book.

Again, because of time, I will now start with the most fundamental elements of the SRR and built it into a working unit, digressing as I go to mention some less important variations and ending up with some theoretical formulas from Whitehead and my comments on them.

Figure 3 shows a standard parallel resonant circuit (or tank circuit as it is commonly called) with all the losses absorbed into the parallel conductance G. We use G rather than R because it is easier mathematically to go from +G thru 0 to -G. during circuit operation than to go from +R thru ∞ to -R. Whitehead is very rigorous and develops complex equations concerning this circuit but for the time being I will content myself with referring to the resonant angular frequency, ω_o

$$\omega_o \approx \sqrt{\frac{1}{LC}} \quad \text{and} \quad Q = \frac{\omega C}{|G|}$$

It is important to notice the absolute value sign around G because a large negative G, indicating strong feedback, will accelerate the buildup of oscillations in the tank circuit the same way that a large positive G will cause rapid damping.

We could add a current generator at the resonant frequency to the circuit of Figure 3, but it can be shown that this is at least qualitively the same as adding a negative G in parallel with the existing positive G. To finish our model we must also add some means of introducing the r-f input signal to the circuit, but before discussing how this might actually be accomplished, we will just assume that the signal exists; for example, we can assume that L is a ferrite rod antenna in an r-f field.

The SRR mode which is easiest to analyze mathematically is called the slope-controlled externally quenched linear SRR. This mode is illustrated in our Figure 4 which is taken from Whitehead's book. In this example, the control signal is the grid voltage of a triode tube which is originally biased beyond cutoff: G is positive. Next, the oscillator is gradually brought into conduction and G goes from positive to negative at a rate designated G'(t). This is called the regenerative period and any external signal existing in the circuit will be amplified, particularly right around the time labeled t_1 when G becomes 0. This is called the sampling period. During a relatively long period of time called the buildup period the feedback is increased to a maximum value and then decreased so that at time t_2 the oscillation will have reached a peak value. The net value of G has now returned to 0. Then, for a still longer interval called the damping period, the active element is turned off and G returns to its initial positive value. The active device must remain off until any residual oscillation in the tank circuit is below the noise level. The oscillation pulse shape is Gaussian as shown in Figure 4.

The significant feature of this mode of operation is that the active device is turned on a relatively short time so that the circuit never saturates. It can be shown that under these conditions the peak r-f output voltage is a linear function of the signal input voltage so that the SRR can be used to receive an A.M. modulated signal with low distortion. However, there are at least two major disadvantages: extra circuit cost and complexity because of the required external quench circuit and, in some cases, an r-f detector circuit must be used to detect the r-f pulses, because if the active element is running Class A linear, the anode or collector current will not change as a function of r-f level so self-detection is not possible. Another mode, called the sinusoidally quenched SRR, can be analyzed by almost the same equations. It is most commonly used because the quench voltage is relatively easy to generate. I have found, however, that it is not feasible to use this method with some high frequency transistors because the base to emitter reverse breakdown voltage is too low. Whitehead also discusses, at almost equal length, the linear rectangular quench mode, but since this mode is not used as often and our time is limited, I will not discuss it in detail. It is similar to the slope controlled mode except that the active device is brought instantaneously from cutoff to the desired "on" state. The resulting pulse is not Gaussian but is like the shape shown in the pendulum analogy, Figure 2. Aside from any other considerations this mode is seldom used because, at least at high frequencies, it is difficult to generate a pulse with the sharp transition times required.

By using a longer on-time in the quench sequence, the SRR can be operated in what is called the logarithmic mode because an analysis will show that the output voltage from such a receiver is proportional to the logarithm of the input signal

$$\frac{V_{2out}}{V_{1out}} = \log_e \left[\frac{V_{2in}}{V_{1in}}\right]$$

In cases where audio fidelity is not important this has several important advantages; better immunity to interference from other signals, better apparent gain for small signals and an apparent a.g.c. action.

A receiver operating fully in the logarithmic mode will reach saturation even when the oscillator starts from noise. However, if there is a signal present, the oscillator will start sooner, so the r-f output from a receiver in this mode will appear as shown in Figure 5. Detection methods are the same as for the other modes. It is also possible to adjust the receiver so that it operates in the linear mode for small signals and in the logarithmic mode for large signals. By far the most common mode at the present time is the self-quenched oscillator in which the bias circuit is deliberately designed to operate somewhat like a blocking oscillator, that is, to bias itself to cutoff at the quench frequency. Since RC time constants are involved, it will operate in the slope-controlled mode and since the oscillator must saturate before it can turn itself off it will operate in the logarthmic mode. However, instead of increasing the width of the quench pulses the system operates by decreasing the distance between pulses and hence, increasing the pulse repetition rate or PRR. Since the active device will consume more current when it is on than when it is biased off, this system is self-detecting. An illustration of the quench pulses from this class of operation is shown in Figure 6.

There are many other variations of the SRR principle that have been discussed in the literature, but I will have time to mention only a few. One class of these variations is a method of speeding up the quench cycle by using a more complex quench waveshape, but it requires so much more circuitry and requires such careful adjustment that it is probably not worth the effort. Figure 7 shows a typical quench cycle. Starting with the active device biased below cutoff, at time t_o the bias voltage is rapidly made less negative to a point where the loop gain of the circuit is slightly less than one. The voltage is then gradually made more positive so that it passes thru the point where oscillation starts to a point where the loop gain is slightly greater than one. The time from t_o to t_1 is long enough for the oscillation amplitude to increase to fully encompass the sensitive region. At time t_2 the grid is made more positive so that the circuit has maximum gain, and so that at t_3 the r-f envelope has reached its maximum value. At this point, the grid voltage is again dropped below cutoff. This system has all the advantages of optimum slope control but shortens the total time the oscillator is on.

A more modern improvement shunts a good r-f diode across a portion of the tank circuit. At low voltages the diode acts like a lossless capacitor, but the diode will limit the amplitude of the peak r-f voltage developed in the circuit. This helps in two ways: it reduces r-f radiation and also reduces the required quench time. Its main disadvantage is the cost of the diode which, for optimum performance, should have a low forward voltage, low capacitance and low loss at the signal frequency.

I will now present, without derivation, some of the more important equations from Whitehead's book, and then give you some of my own comments and conclusions. It will be up to each of you to decide whom to believe. In any case, where a tuned circuit is involved, the circuit of Figure 3 will apply. If this circuit is excited with a sine wave current generator tuned to the resonant frequency of the circuit, the complete solution for the voltage across the circuit is

$$V = \frac{A}{G}\,\frac{\omega_o}{\omega_d}\,e^{-Gt/2C}\sin(\omega_d t) + \frac{A}{G}\sin(\omega_o t) \qquad (1)$$

where A is the amplitude of the exciting signal, ω_o is the natural resonant frequency, ω_d is the frequency of the damped oscillator and α is the damping factor.

$$\alpha = G/2C$$

$$\omega_d = \sqrt{\frac{1}{LC} - \left(\frac{G}{2C}\right)^2} = \sqrt{\omega_o^2 - \alpha^2}$$

The first term of Equation (1) is the transient term, which will be very small for most signals. For the case of a vacuum tube used as the active element G can be defined as

$$G = G_o - Kq_m \qquad (2)$$

where G_o represents the losses associated with the circuit with the tube turned off and $-Kq_m$ is the negative conductance associated with positive feedback. The circuit will oscillate whenever G is negative. In the case of other active elements, q_m can be replaced, at least in principle, by the appropriate gain parameter. Equation (2) may also be expressed in the form $G(t)=G_o[1-F(t)]$.

He then goes on at great length to solve the differential equations representing the operation of the circuit during the quench cycle, but the results are much too complicated to present here. One of the conclusions he draws, however, is that the conditions for the slope-controlled mode of operation are that

$$t_1 > \frac{120}{2\pi}\,T_o \qquad (3)$$

where t_1 is as shown in Figure 4, T_o is the period of one oscillation, and

$$Q = \frac{\omega C}{G} \qquad (4)$$

$$Q_o = \frac{\omega_o C}{G_o} \qquad (4a)$$

Equation (3) is important because it defines the maximum rate of change of conductance, $G'(t)$, for the circuit to be in the slope-controlled mode; the required active period must include more than 20 r-f cycles.

He then goes on to derive expressions for the gain of the receiver operating in the slope controlled mode and divides it into two parts: The first part

$$Ns = -\frac{1}{2C}\int_{t_1}^{t_2} G(x)\,dx = \frac{a^-}{2C} \text{ nepers} \qquad (5)$$

he defines as the super-regenerative gain which is expressed in nepers:

1 neper = 8.7db

a^- is the area shown in Figure 4 where G is negative.

The other portion of the gain, called the slope gain, or regenerative gain is

$$N_o = \frac{1}{2}\log_e\left[\frac{\pi\,G_o^2}{C/G'(t)/}\right] \text{ nepers}$$

or, (6)

$$4.35\log_e\left[\frac{\pi\,G_o^2}{C/G'(t)/}\right] \text{ db}$$

It is appropriate that I stop at this point and make a few comments about Equation (6), which I consider highly misleading. I agree that the terms C and $G'(t)$ should be in the denominator, but I disagree with the conclusions the casual reader may draw from the appearance of the term G_o^2 in the numerator. He might conclude that you will increase the gain (and therefore presumably improve the receiver) by deliberately increasing the losses in the circuit. After some thought, I have been able to resolve the conflict in two steps.

First, the author defines gain as the ratio of the r-f voltage in the tank circuit at the end of the period under consideration compared to the voltage at the beginning of the period. Equation (6) refers only to the increase of voltage between t_o and t_1 in Figure 4. It should first be recalled, from Equation (1), that the steady state voltage at time t_o is inversely proportional to G_o. Hence, one of the factors of G_o is immediately cancelled if we define gain in terms of A rather than A/G (A is the amplitude of the signal current impressed on the tan: circuit). Secondly, the period t_o to t_1 is defined as the time taken to go from $G=G_o$ to $G=0$ at the rate $G'(t)$. Hence, $G_o/G'(t)$ will define the time T_1. So, the term in the brackets of Equation (6) can be rewritten, dimensionally, as $(KT_1)/C$, where K is a constant and gain is defined in terms of output voltage over input current. What the equation now says, dimensionally, is that gain is inversely proportional to the circuit capacitance and directly proportional to the time the external signal can influence the circuit before it is swamped by oscillation, both of which are reasonable conclusions. An equation found later in the book says that bandwidth is

$$bs = \frac{1}{\pi}\sqrt{\frac{G'(t_1)}{C}} \qquad \text{at } -8.7\text{db} \qquad (7)$$

Equation (6) says that if you want to maintain a constant gain, and increase G_o, you must increase $G'(t)$, but this will increase the bandwidth and hence increase the noise bandwidth of the receiver. Hence, the ultimate effect of increasing G_o is to decrease the sensitivity (if not the gain) by increasing the noise in the receiver. It should also be noted that if we accept the proposition that we use normal r-f design rules about using low loss tuned circuits and hence do not make G_o a variable in Equation (6) and if we also accept the idea that C should be as small as possible, and therefore not variable, that the only variable left is $G'(t)$ and we get back to the usual rule that there is a tradeoff between gain and bandwidth.

Whitehead himself goes on later in the chapter to show that the total voltage gain in the receiver is

$$u_t = \frac{b_e}{b_{es}} \exp\left(\frac{a^-}{2C}\right) \qquad (8)$$

where

$$\frac{b_e}{b_{es}} = G_0 \sqrt{\frac{\pi}{C|G'(t_1)|}} \qquad (9)$$

Remembering that his definition of voltage gain is V_2/V_o and that $V_o=A/G_o$, where A is the amplitude of the signal current, then

$$V_2 = u V_o = \frac{A}{G_o}\frac{b_e}{b_{es}} \exp\left(\frac{a^-}{2C}\right)$$

$$= \frac{A}{G_o} G_o\left[\sqrt{\frac{\pi}{C|G'(t_1)|}}\right] \exp\left(\frac{a^-}{2C}\right)$$

$$= A\left[\sqrt{\frac{\pi}{C|G'(t_1)|}}\right] \exp\left(\frac{a^-}{2C}\right)$$

It has been shown, therefore, that even using his own equations, G_o does not appear as a factor in the final gain equation if gain is based on the ratio of the final output voltage to the original exciting current.

It will be noted that this gain expression depends only on C and $G'(t)$ as design parameters, and that both of them should be made as small as possible. The area factor, $\exp(a^-/2C)$, relates only to how long the oscillation is allowed to proceed before quenching and has more to do with the required dynamic range than with receiver sensitivity. For example, if the receiver is required to have a linear dynamic range of 20db, then the factor, $\exp(a^-/2C)$ must be 20db smaller than it could be made in a receiver which would saturate on small signals.

I will say very little about the step-controlled receiver except to say that the major gain factor, the super-regenerative gain, is the same as for the slope-controlled receiver, $\exp(a^-/2C)$.

The bandwidth is given by

$$b_s = \frac{G_o G_1}{G_o+G_1}\frac{1}{\pi C} \qquad \text{at } -6\text{db} \qquad (10)$$

where G_1 is the negative resistance during the oscillation period. The so-called natural bandwidth of the tuned circuit (passive state) is

$$b = \frac{G_o}{2\pi C} \qquad \text{at } -3\text{db}$$

He also shows that the step-gain factor is

$$\mu_o = \frac{G_o+G_1}{G_1} \qquad (11)$$

so that

$$\mu_o = \frac{G_o+G_1}{G_1} = \frac{1}{2}\frac{b}{b_s} \qquad (12)$$

This compares to the factor

$$2\sqrt{\pi\left(\frac{b}{b_s}\right)}$$

already given for the slope-controlled case. Therefore, the additional gain attributable to using the slope-controlled mode is approximately $4\sqrt{\pi}$, or approximately seven times better. This is an entirely reasonable conclusion based on the fact that in the slope-controlled mode the original small signal can be gently amplified for a longer time before being jolted into full oscillation.

We will next consider the logarithmic mode, which, since it includes the self-quenched type of SRR, is currently the most popular mode of operation. The output waveform of the externally quenched logarithmic mode SRR is shown in Figure 5. In this mode the amplitude of the r-f signal builds up to the saturation level, V_e, along an exponential curve from an initial voltage:

$$V_e = V_1 e^{at_1} = V_2 e^{at_2} \tag{13}$$

where 1/a is the time constant of the buildup.

It can be shown that the difference in area between these two envelopes

$$\Delta A \approx \frac{V_e}{a} \log_e \left(\frac{V_2}{V_1} \right) \tag{14}$$

so that the audio output, which is proportional to ΔA, is logarithmically related to the signal voltage. This is ideal for systems such as garage door openers, which use pulse modulation, since it provides a sort of low cost AGC action, but is unsuitable for ordinary AM reception due to the severe distortion that would result. For signals near the noise level, Equation (14) becomes

$$\Delta A \approx \frac{V_e}{a} \log_e \left[\frac{\sqrt{\overline{V_n^2} + V_s^2}}{\sqrt{\overline{V_n^2}}} \right] \tag{15}$$

or, if the signal is much larger than the noise

$$\Delta A \approx \frac{V_e}{a} \log_e \left[\frac{V_s}{\sqrt{\overline{V_n^2}}} \right] \tag{16}$$

where $\sqrt{\overline{V_n^2}}$ is the rms noise voltage.

The apparent frequency response of a receiver is a rather complex function which depends on the signal to noise ratio, but, in general, is narrower at the nose and wider at the skirts than the same circuit operating in the linear mode. A comparative plot showing the frequency response for various values of signal to noise rates is shown in Figure 8.

Another useful equation is one for the maximum quench frequency which can be used while still staying in the logarithmic mode. This is

$$fq = \frac{\pi f_o}{20 \log_e (V_e / \sqrt{\overline{V_n^2}})} \tag{17}$$

where

f_o is the r-f carrier frequency,

V_e is the limiting r-f voltage, and,

$\sqrt{\overline{V_n^2}}$ is the rms noise voltage at the input of the receiver.

This equation should only be used as guide.

An equation illustrating the relationship between the demodulated output voltage V_m and some of the design parameters is:

$$V_m \propto \frac{2V_e f_q C}{G_1} \left[\log_e (1 + M\cos W_m t) \right] \tag{18}$$

where M is the degree of modulation, and G_1 is the negative conductance during oscillation.

This seemingly contradictory equation (C is in the numerator) appears because a, the rate of buildup of the oscillation, appears in the denominator in the equation from which Equation (18) is derived and $a = G_1/2C$. Instead of considering C a variable, it would be best to concentrate on making G_1 as small as possible. It should also be noted from Equation (17) that the maximum quench frequency will be lowered if C is made larger (since Q is directly propoportional to C) so the factor C will cancel out in Equation (18).

The most popular SRR system at this time (because of its simplicity and low cost) is the self-quenching receiver in which the bias system of the single active device is designed to bias the device to cutoff periodically when it reaches saturation. This state has already been illustrated in Figure 6. Since the self-biasing system will almost certainly involve time constants, the self-quenched receiver will almost certainly operate in the slope-controlled state. It also can be seen that it operates in the saturated or logarithmic mode, but, instead of prolonging its stay in the saturated state, as shown in Figure 5, it changes the quench frequency as shown in Figure 6. The net current thru the system increases in the presence of an input signal because the duty cycle increases just like in the externally quenched case. For signals much greater than noise, the output voltage

$$\Delta V \propto \frac{f_q^2}{a} A \log_e \frac{V_s}{\sqrt{\overline{V_n^2}}} \tag{19}$$

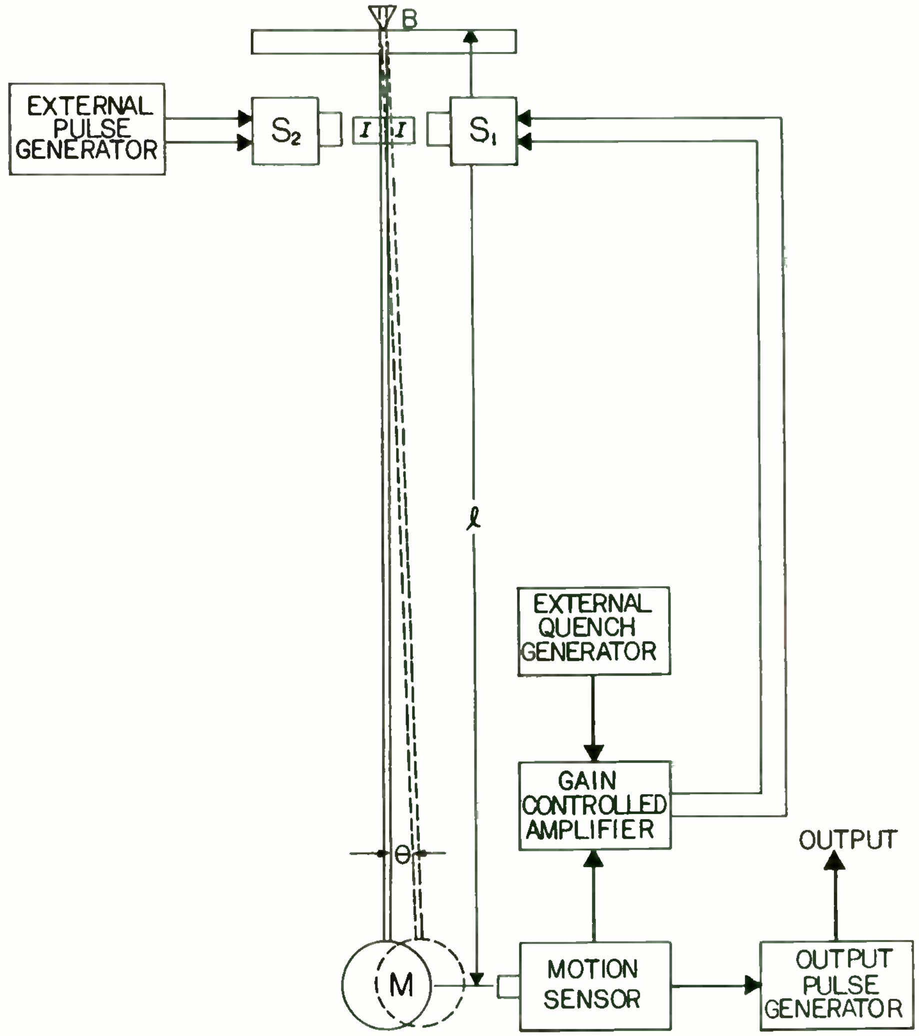
where A is the area under the voltage-time envelope, f_q is the quench frequency, and 1/a is the time constant of buildup.

Having covered the equations and conclusions in Whitehead's book as thoroughly as time will permit, I will end this paper with some thoughts of my own which I am not in a position to prove mathematically, but which you might want to consider in light of what you have learned from the rest of the paper.

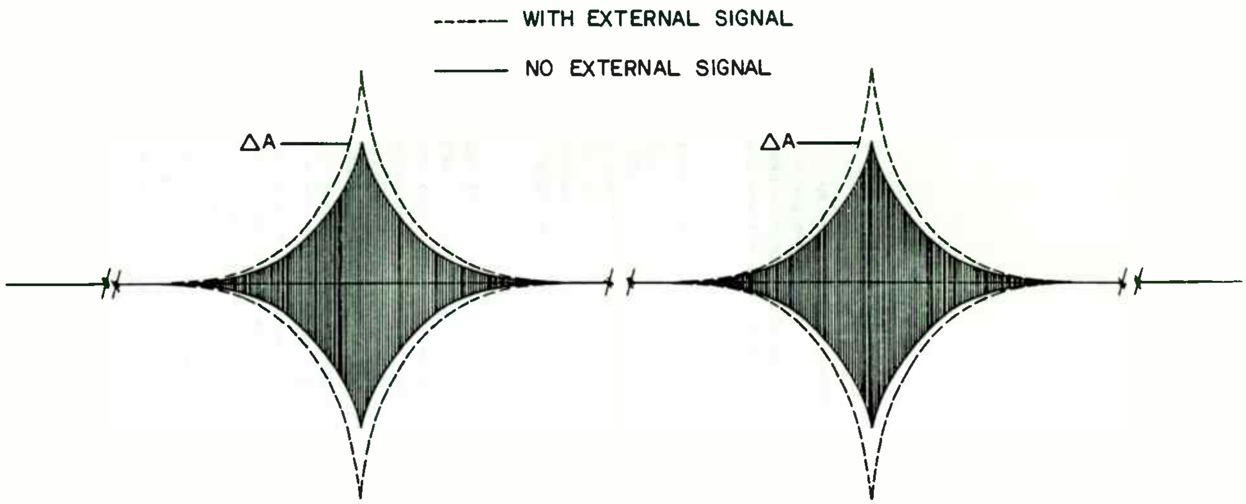
My principal disagreement with Whitehead is that he continually talks about gain instead of sensitivity, altho all of his gain equations ultimately break down if the input signal is smaller than the noise level at the point where the signal is introduced. So my opinion is that the job of the designer is first of all to treat the system as a well designed r-f amplifier where the signal to noise ratio at the input of the amplifier is optimised as well as possible. Although at one point Whitehead confronts this problem and effectively ends up saying that it doesn't matter, I think that if you are operating at a frequency low enough that significant gain can be obtained from the active device being used, and also if the device at that frequency has a high input impedance, that the best circuit design plan would be to put the tuned circuit at the input of the amplifier and to use a low positive feedback factor to maintain high gain in the active device. Whitehead's comment on this is that high gain requires high currents which leads to high shot noise, but I believe that the best compromise should be found so that the noise in the system compared to the input signal at that point is as low as possible.

The other matter which Whitehead says very little about but which has a major effect on the gain of the system is to operate in such a way that the change of d-c current in the active device when going from a passive to oscillatory state is as great as possible. This is especially true if the following audio amplifier is a bipolar transistor which is current driven. Another major problem for which Whitehead offers no solution except the use of an r-f amplifier is that of r-f radiation. This is another opportunity for a clever designer to show his ingenuity.

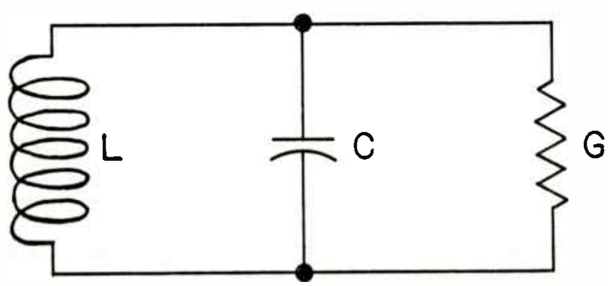
I am sure that I have not covered all the questions which might arise on this topic, but I will have achieved my goal if, after studying the paper, you feel that you understand what parameters and methods are important in the design in the same way that you understand how to design the more common types of r-f circuits.



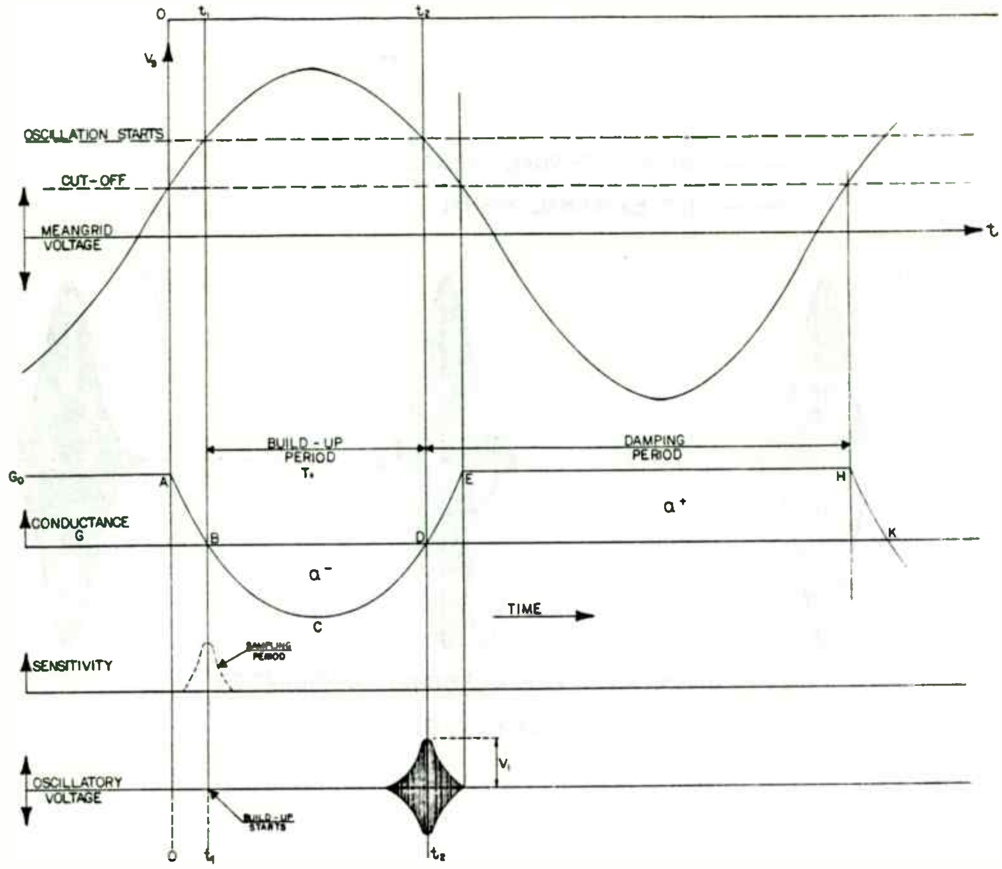
PENDULUM ANALOGY
FIGURE 1.



QUENCH CYCLES: RECTANGULAR EXTERNAL QUENCH
FIGURE 2.

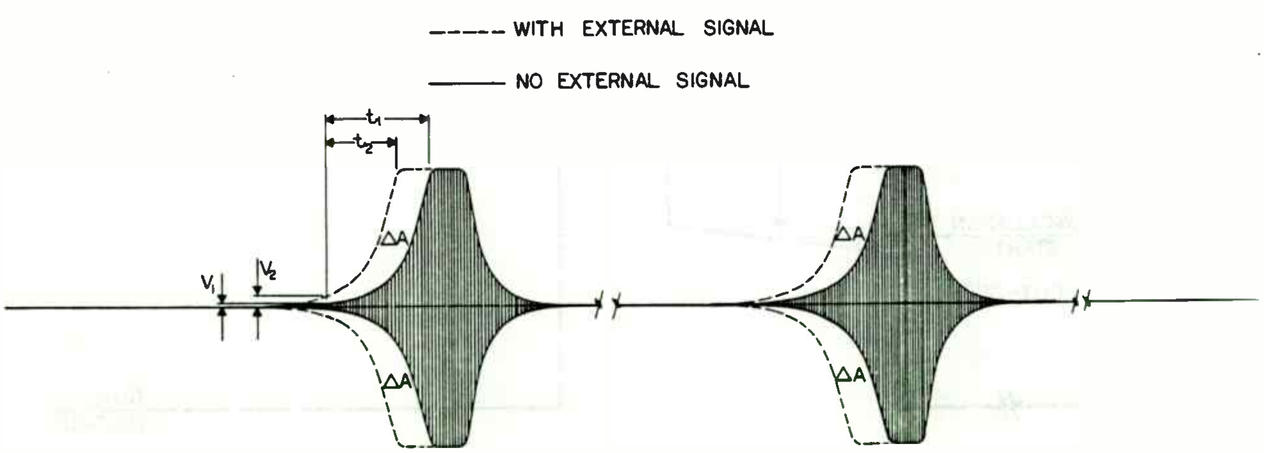


PARALLEL RESONANT CIRCUIT
FIGURE 3.

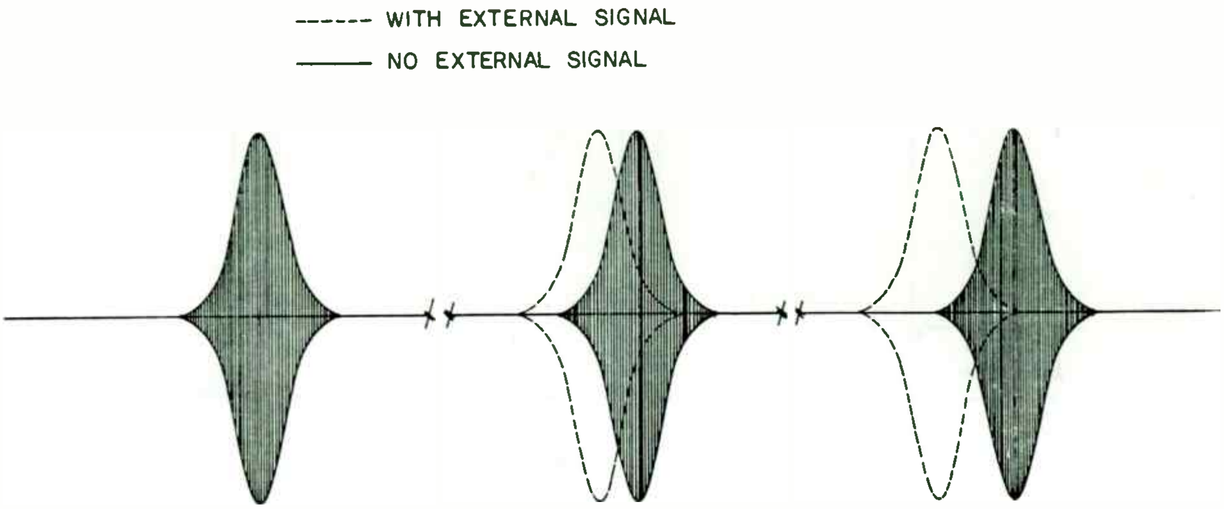


FROM: WHITEHEAD SUPER-REGENERATIVE RECEIVERS

QUENCH AND CONDUCTANCE CYCLE : SLOPE-CONTROLLED STATE.
FIGURE 4

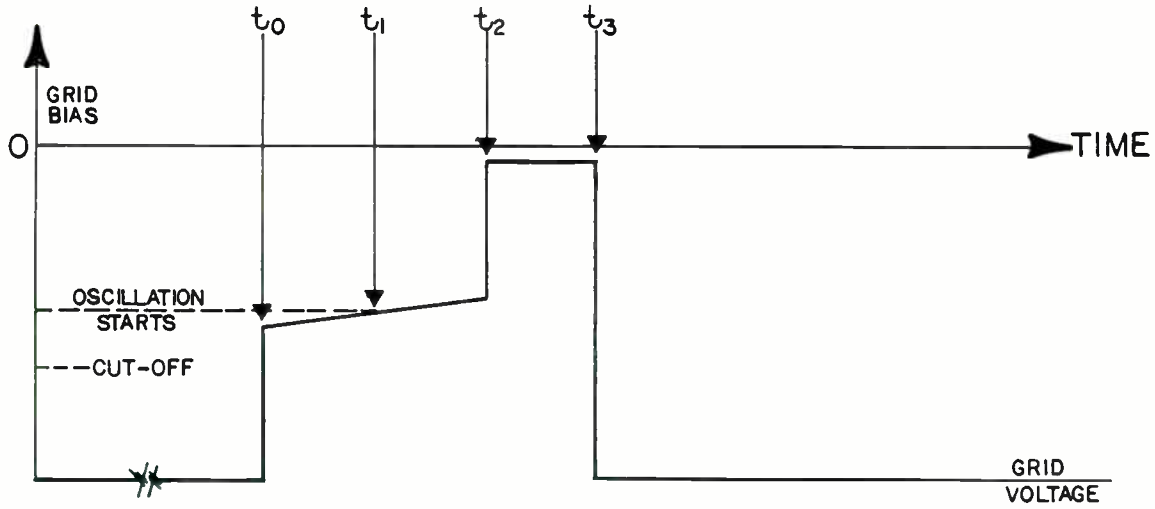


QUENCH CYCLES: LOGARITHMIC MODE, EXTERNALLY QUENCHED
FIGURE 5.



QUENCH CYCLES: SELF QUENCHED OSCILLATOR

FIGURE 6.



COMPLEX QUENCH CYCLE

FIGURE 7.

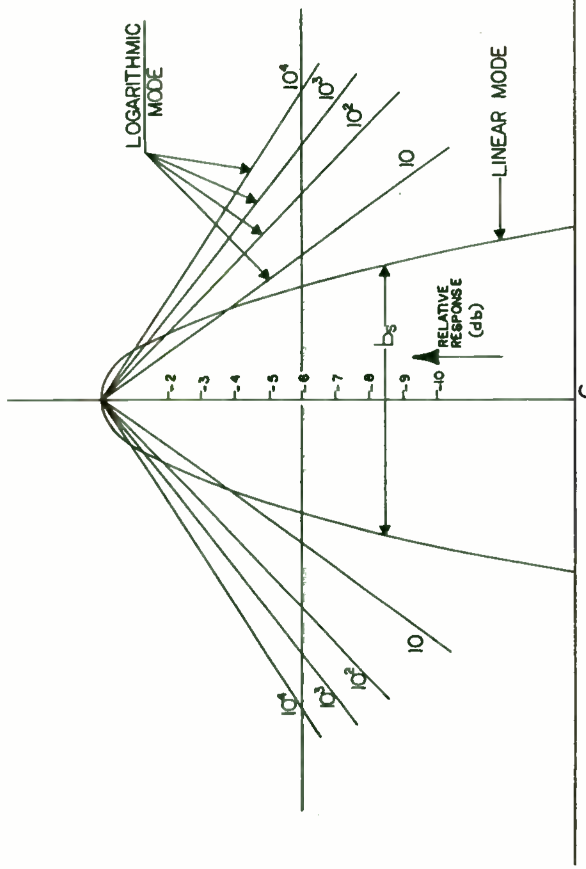


FIGURE 8.

FREQUENCY RESPONSE COMPARISON LOGARITHMIC VS LINEAR MODE

FROM: WHITEHEAD SUPER-REGENERATIVE RECEIVERS

BIBLIOGRAPHY

1. Harold A. Wheeler, A Simplified Theory and Design Formulae for Superregenerative Receivers. Wheeler Monograph 3, Wheeler Labs, Inc. Great Neck, N.Y., June 1948
2. W.E. Bradley, Superregenerative Detection Theory, Electronics, Vol. 21, Pages 96-89 September 1948
3. A. Hazeltine, D. Richman, and B.D. Loughlin, Superregenerator Design. Electronics Vol. 21, Pages 99-102, September 1948
4. Edwin H. Armstrong, Some Recent Developments in Regenerative Circuits. Proceedings of the I.R.E. August 1922, Pages 244-260
5. F.W. Frink, The Basic Principles of Superregenerative Reception. Proceedings of the I.R.E., January 1938, Pages 76-106
6. G.G. MacFarlane and J.R. Whitehead, The Theory of the Superregenerative Receiver Operated in the Linear Mode. Journal of the I.E.E. (London), Part III, May 1948, Pages 143-157
7. H. Ataka, On Superregeneration of an Ultra-Short-Wave Receiver. Proceedings of the I.R.E. August 1935, Pages 841-884
8. F.R.W. Strafford, The Superregenerative Detector: An Analytical and Experimental Investigation. Journal of the I.E.E. (London Part III, No. 21, January 1946, Pages 23-28.
9. E.P. Tilton, A Non-radiating Superregenerative Receiver for Two Meters, QST, February 1946, Pages 53-56 and 108
10. A. Easton, Superregenerative Detector Selectivity. Electronics Volume 19.3 March 1946, Page 154 and following
11. H. Stockman Superregenerative Circuit Applications Electronics, Volume 21.2 February 1948, Page 81 and following
12. H.A. Robinson, Regenerative Detectors, QST, Volume 17.2, February 1933, Page 264 and following
13. H.P. Kalmus, Some Notes on Superregeneration With Particular Emphasis on Its Possibilities For Frequency Modulation. Proceedings of the I.R.E. October 1944, Pages 591 to 600
14. L.A. Hazeltine, Discussion on the Shielded Neutrodyne Receiver. Proceedings of the I.R.E. June 1926, Pages 395-412
15. Hazeltine FreModyne FM Circuit. Tele-Tech, December 1947, Pages 41,85 and 86
16. C.L. Ring, Diode Improves Performance in Superregenerative Circuit. Electronic Design, March 19, 1969, Page 230
17. A. Iwakami, Improved Superregenerative Receiver. Ham Radio, December 1970, Pages 40-49
18. C.L. Ring, Optimizing the Superregenerative Detector, Ham Radio, July 1972, Pages 32-35
19. H. Olson, Enter the Strange But Useful Blacking Oscillator, Ham Radio, April 1969, Pages 45-49
20. W.S. Percival, Superregenerative Receiver, U.S. Patent No. 2171148, Nov. 20, 1935 - August 29, 1939 (Zero Conductance for a Fraction of the Quench Period)
21. R.C. Emerson, Superregenerative Receiver, U.S. Patent No. 2398214, Feb. 14, 1944 - April 9, 1946 (Zero Conductance for a Fraction of the Quench Period)

A General Computer Algorithm
for Mixer Spurious Analysis

Submitted by:
Alan M. Victor
Motorola Communications Div.
Plantation, Florida 33322
9/22/84

Introduction

This paper (article) discusses an approach to the calculation of mixer spurious. The analysis and the computer algorithim which follow allow the prediction of various mixer spurious including general harmonic analysis, cross-over spurious loacation, and self-quieting analysis. The insight gained by this analysis can save you future headaches and costly re-design. The approach outlined is applicable to receiver designs, synthesizers, and other systems which use a multitude of mixers, oscillator sources, and intermediate frequencies, (I.F.'s.).

In the design of receivers or frequency synthesizers a number of mixer stages, fixed sources, or sweeping sources are present. In receiver design the selection of the (I.F.) is critical in minimizing unwanted responses. In frequency synthesizers a similiar problem exists when a multiple loop approach is used. In this case the intermediate frequency is analogous to the synthesizer loop bandwidth. Just as in a receiver, if a spurious response is present in the I.F. no additional selectivity will reduce the level of spurious. In a synthesizer one could be forced to narrow the loop bandwidth or side-step around the interfering condition to reduce the level of spurious, but this is at the expense of increased lock time or V.C.O. phase noise degredation. Instead the motivation is to select the right choice of oscillator frequency, I.F., and mixer type (1).

Equations Developed

As an introduction, we shall define some of the most prevelent types of mixer spurious as they occur in a communications receiver. The ideal mixer would behave as perfect product multiplier and if the input signals (local oscillator and R.F. source) are sinusoidal then the only I.F. components are the sum and difference frequency. Simple low-pass or high pass filtering removes the wanted I.F. frequency. Unfortunatly the sources are rarely perfect and the mixer is not a perfect product multiplier. Instead the mixer output consists of harmonic multiples of the local oscillator source, the R.F. source, and any other fixed or sweeping source which might couple into the mixer inputs. The harmonic multiples at the mixer output follow an integer relation among the various mixer input signals and the I.F. output can be expressed as

$$f_{I.F.} = N \cdot f_{R.F.} + M \cdot f_{L.O.} + K \cdot f_X \qquad (1)$$

In equation (1) N, M, and K are integer values which are plus or minus and N, M, or K can be zero. The term f_X is any

Alan Victor: Computer Algorithim for Spurious Analysis

third frequency present and might represent a fixed source or a sweeping one. Equation (1) can be extended to any number of sources and I.F.'s. Proper choice of the source coefficents N, M, and K permit both high and low side oscillator injection (local oscillator source is either above or below the R.F. signal) into the mixer. Therefore, up conversion or down conversion can be analyzed. Let's consider the case in Equation (1) where f_X is absent. If N is +1 and M=-1 we have low side injection and the desired mixing term is

$$F_{I.F.} = F_{R.F} - F_{L.O.} \qquad (2)$$

For N=-1 and M=+1 we have high side injection but F I.F. remains the same. In one case Equation (1) gives the desired result; the other result is the image frequency and vica-versa. A particularly troublesome spurious is given by M=N=2 and is referred to as the 1/2 I.F. spurious. This spurious lies one half I.F above or bel the desired R.F. frequency (depending on the whether high or low side injection is used) and is difficult to filter if a wideband receiver is contemplated.

Troublesome spurious occur for M=N +/- 1 or N=M +/- 1. When N and M differ by only an integer value close in spurious can occur. These undesired signals are again close to the desired R.F. receive frequency and are difficult to filter. Proper selection of the intermediate frequency as well as the mixer type helps to reduce the level of these spurious.

Equation (1) can be handled graphically (2,3); but quickly becomes a nightmare as the number of sweeping sources or I.F.'s increase. Instead a computer program is presented which handles the above equation and also allows all three sources to be swept. Two I.F.'s are checked in the computer routine, but any number of I.F.'s as well as any number of swept sources can be accomodated. Figure (1) illustrates the system under consideration. To keep the number of possible spurious to a reasonable level the I.F. bandwith or the search range is made adjustable. Additional reduction in the number of M,N, and K spurious response data points is possible if selectivity or mixer supression is factored into the program (4). Furthermore, the starting and ending values for the coefficents of each source are independent and also adjustable in range.

Several problems are addressed in the program which allow all possible spurious to be detected without causing an undue increase in computation time. Crossover spurious for example can be particularly troublesome especially if swept sources are being used. These spurious can move through the I.F. and therefore, have zero offset from the specified I.F. center frequency. Instead of trying to calculate where these crossovers precisely occur, we need only determine that a crossover spurious does exist. Reference to Figure (2) illustrates the technique used. The program checks the sign of the calculated I.F. frequency, and notes if the signal

Alan Victor: Computer Algorithim for Spurious Analysis

produced by the spurious lies within the I.F., above it, or below it. This calculation is done for the extreme end frequencies of each swept source. Clearly if the spurious produces an I.F. signal within the specified search range then the spurious condition is met. If the signal lies below the I.F. range then a negative flag is set, otherwise a positive flag is set. The state of the flag is checked on each calculation as the swept sources are moved to their next extreme frequency point. If any of these set flags change sign from one calculation to the next; the I.F. signal produced by the spurious must crossover the I.F. center frequency and a spurious condition exists. Note that equation (1) is really quite general. The algorithim does not care what the signal terms are, and it is up to the individual to determine what shall be identified as the R.F. signal, the L.O., and finally the I.F.

As mentioned previously the program will handle a multitude of signal sources and I.F. frequencies. The present routine will handle three sources and two I.F. frequencies. All of the sources can be swept or they can be fixed. The program steps the sources in a Grey-Code sequence, given in Figure (3). Thus, each source is essentially swept one at a time and the program looks for a valid spurious condition.

Given three oscillators, we find that 16 possible mixing conditions exist, refer to Appendix A. These conditions take into account the sweeping of the three oscillators as well as the type of mixing, i.e. sum mixing or difference mixing. In addition we need to consider both high and low side injection. Since we can take the absolute value of the result for high or low side injection and arrive at the same spurious response, only 8 combinations are neccesary. In addition, 4 combinations are needed to account for the kind of mixing, i.e. sum or difference.

Computer Routine Developed

To save computation time the 4 combinations of mixing are presented as a menu and the user is asked to select the appropiate case type one thru four, or case five which selects all combinations. The main subroutine in the program is SEARCH which increments the coefficents of the three oscillators: f1, f2, and f3 using the alpha variables N, M, and K. The mixing "type" is selected by the variables X, Y, and Z, and is controlled by selecting the desired case routine.

The spurious analysis does not actually sweep the three oscillators, but instead makes a spurious computation at both band edges for f1, f2, and f3. The four subroutines IFRANGE, FLAGTSET, FLAGTEST, and SPUR, compute if a spurious response falls exactly in our specified I.F. bandwidth. Not only is the I.F. bandwidth checked but so is the mirror image of the I.F. since this represents a valid response. Since the I.F. band and the mirror image I.F. band are utilized in a

Alan Victor: Computer Algorithim for Spurious Analysis

calculation we are able to check for spurious which could move through the I.F. for a small change in any one of the signal source frequencies. Crossover spurious are checked in this manner. The subroutine IFRANGE checks for in band spurious. The program requests the desired I.F. bandwidth or search range. The subroutines FLAGSET and FLAGTEST monitor the movement of the I.F. spurious as the sign of the spurious changes from one side of the I.F. to the other; i.e. the spurious passes thru either the image I.F. or the I.F. The last major subroutine is SPUR which outputs the final result. This routine recognizes wheather a single conversion analysis is being preformed or if a general harmonic analysis is requested. In the latter all three sources can be present and they may be swept or fixed in frequency.

In single conversion analysis only one I.F. is allowed. A single conversion analysis is possible using a swept R.F. source (f1), a swept L.O. source using f2, and a fixed I.F., f I.F.(2). The analysis using the I.F. and the image I.F. detect spurious at the R.F. image frequency, the half I.F. frequency and at many M, N frequencies where M and N differ by only unity. The significance of these spurious is the fact that they occur close to the desired R.F. receive frequency. f1. Portions of the subroutine SPUR calculate the location of these spurious and the R.F. range over which they move. The designer is then presented with a good picture as too how much R.F. filtering is required.

The computer routine discussed is written in BASIC and two versions are now complete. One is written for the series 200 HP desk top personal computer and the other on the APPLE II. Other versions are running in PASCAL and on a mainframe are clearly faster than the personal computer versions. Never-the-less a 3 source, 30 th order spur search takes less than 60 seconds (and even faster if their are few spurious found) on a personal computer.

As an add in understanding the spurious problem, let us work some examples which are shown in Figure (4) and should help illustrate the material covered.

Consider a F.M. receiver covering 88-108 mhz. The first I.F. is 10.7 mhz and the second I.F. is 455 khz. A microprocessor is used to obtain a clock display function as well as controlling the tuning of the receiver. The uP oscillator is a 3.579545 mhz (f3) crystal. Using the program we find that the general spur analysis points out a self-quieting condition (the receiver will essentially be quieted by it's own internal oscillators) at 3*f3 which is 38.6 khz below our 10.7 I.F. Clearly this is a potential problem especially with the first I.F. over 100 khz wide. Other spurious are noted including some which also effect the second I.F.

Consider a communications receiver for the 25-88 mhz band. Low side injection is used with up-conversion and the first I.F. is 90 mhz. A single conversion spurious analysis points up a number of harmonic related spurious, indicating the need for sub-octave bandpass filters. A low order N,M spur exists

Alan Victor: Computer Algorithim for Spurious Analysis

(2,1) but is at least 25 mhz away from the desired receive frequency. Also a (1,2) spur exists and is more trouble-some as it is within 1 mhz of the desired receive frequency.If high side injection is used then this spurious is no longer present.

Finally a VHF receiver is contemplated. Several I.F frequencies were chosen and low and high side injection tried. The resultants shown in Figure 4c,d indicate the trend. An optimum I.F. is about 1/7 the desired receive frequency when the (M,N) spurious differ by unity. If a wideband receiver is contemplated then the image frequency and half I.F. frequency could be more of a problem than the M,N, spurious. Higher I.F.'s allow these spurious to be moved out, while the M,N spurious move in closer to the desired R.F. frequency. High side injection raises the order of the (M+N) spurious and makes the filtering task easier.

These analysis can be re-evaluated as other sources are involved. Trade-off's will be required but the computer algorithim and approach outlined for handling these conditions should make your job a bit easier.

Conclusions

An approach to the analysis of spurious signals present in a system with multiple mixers, and sources was outlined. This analysis along with a computer algorithim aid the designer in the proper selection of signal frequencies in multiple conversion receivers or similiar devices. The algorithim is general in nature and can be expanded to handle a multitude of sources , mixers, and intermediate frequencies.

Appendix A

Equation (1) is expanded to consider all possible mixing situations and also to account for the sweeping of sources one, two, and three. Using variables N,M, and K as the multiple coefficents for the three sources and variables X,Y, Z to account for the mixing type (sum mixing or difference mixing) we have the following:

$$f\ I.F.(1) \text{ or } f\ I.F.(2) = N \cdot X \cdot f1 + M \cdot Y \cdot f2 + K \cdot Z \cdot f3 \quad \text{(A1)}$$

Since f I.F. = ABS (-f1 -f2 -f3) = (f1 + f2 + f3) only half of the 16 possible sweeping cases need to be analyzed. So we have the following equations: (A2)

$$\begin{aligned} f\ I.F.(1) \text{ or } f\ I.F.(2) &= N \cdot X \cdot f1\ min + M \cdot Y \cdot f2\ min + K \cdot Z \cdot f3\ min \\ &= N \cdot X \cdot f1\ min + M \cdot Y \cdot f2\ min + K \cdot Z \cdot f3\ max \\ &= N \cdot X \cdot f1\ min + M \cdot Y \cdot f2\ max + K \cdot Z \cdot f3\ max \\ &\ \ \vdots \end{aligned}$$

and the remaining 5 terms follow the Grey-Code sequence generated by Figure (3).

References

(1) Henderson, Bert, Mixer Design Considerations Improve Performance, Microwave News Inc., October 1981, vol 11, no 10, pp 103-118

(2) Shores, M.N., Chart Pinpoints Receiver Interference Problems, EDN, January 15, 1969, pp 43-46

(3) Capraro, G.T. and Perini J., A Graphical Approach to the Intermodulation and Spurious Response Problem, I.E.E.E Electromagnetic Compatability, Febuary 1974, pp. 38-43, vol 16, no 1

(4) Meixner, Raymond R., Basic Computer Algorithim Spots Spurious Responses, Microwaves, March 1977, pp 42-46

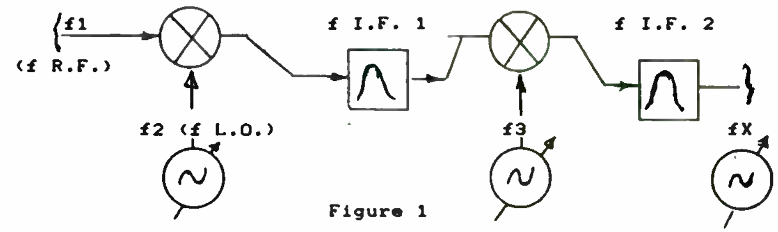


Figure 1

Three source, dual I.F. (dual conversion) system

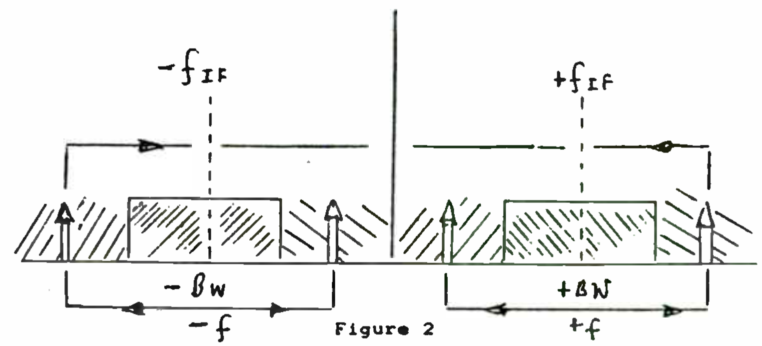


Figure 2

Movement of f I.F. which might constitute a crossover spurious is detected by monitoring the sign of f I.F. and setting and clearing flags. Arrows show possible direction of movement for generating a spurious frequency detect.

IF SPURIOUS DEVIATION

Define the difference frequency $F_{IF} = F_D$ as:

$$F_D = F_{IF} = |m F_{RX} - n F_{OSC}| \qquad (1)$$

The difference between the desired receive freq F_{RX} and the spurious receive frequency is ΔF

$$\Delta F = F_{RX} - F_{SPUR} \qquad (2)$$

Using low side injection with $\Delta F = 0$

$$F_D = F_{IF} = F_{SPUR} - F_{OSC} \qquad (3)$$

With ΔF not zero from (2)

$$F_{RX} = \Delta F + F_{SPUR} \qquad (4)$$

$$F_D = | m (F_{SPUR} + \Delta F) - n F_{OSC} | \qquad (5)$$

Replacing the absolute value signs and with m = n

$$F_D = \pm [m (F_{SPUR} + \Delta F) - m F_{OSC}] \qquad (6)$$

$$F_D = \pm [m (F_{SPUR} - F_{OSC}) + m \Delta F] \qquad (7)$$

$$F_{IF} = F_D = \pm [m (F_{IF}) + m \Delta F] \qquad (8)$$

$$F_{IF} - [\pm m (F_{IF})] = m \Delta F \qquad (9)$$

$$\boxed{-F_{IF} (m \pm 1) / m = \Delta F \quad \text{LOW SIDE}} \qquad (10)$$

$$\boxed{F_{IF} (m \pm 1) / m = \Delta F \quad \text{HIGH SIDE}}$$

M, N + 1 OR N, M + 1 SPURIOUS

Define the difference frequency $F_D = F_{IF}$ (low side injection) as:

$$F_D = F_{IF} = | m F_{RX} - n F_{OSC} | \qquad (1)$$

The difference between the desired receive frequency F_{RX} and the spurious receive frequency F_{SPUR} is ΔF.

$$\Delta F = F_{RX} - F_{SPUR} \qquad (2)$$

Now replacing n = m + 1 in (1) and using (2)

$$F_D = F_{IF} = \pm [m (\Delta F + F_{SPUR}) - n F_{OSC}] \qquad (3a)$$

$$F_{IF} = \pm [m (F_{SPUR} + \Delta F) - m F_{OSC} - F_{OSC}] \qquad (3b)$$

$$= \pm [m F_{SPUR} + m \Delta F - m F_{OSC} - F_{OSC}] \qquad (3c)$$

Substitute for F_{OSC}, $F_{OSC} = -F_{IF} + F_{SPUR}$ (4)

$$F_{IF} = \pm [m F/_{SPUR} + m \Delta F + m F_{IF} - m F/_{SPUR} + F_{IF} - F_{SPUR}] \qquad (5)$$

$$F_{IF} = \pm [m \Delta F + (1 + m) F_{IF} - F_{SPUR}] \qquad (6)$$

$$\frac{-F_{IF} (\pm 1 + m + 1) + F_{SPUR}}{m} = m/\Delta F \qquad (7)$$

$$\Delta F = \frac{F_{SPUR} - (m + 1 \pm 1) F_{IF}}{m} \qquad (8)$$

$$\boxed{\Delta F = (F_{SPUR} - m F_{IF}) / m, (F_{SPUR} - (m + 2) F_{IF}) / m \quad \text{LOW SIDE}} \qquad (9)$$

A similar condition occurs for high side injection.
Now $F_D = | n F_{OSC} - m F_{RX} |$

$$F_D = F_{IF} = \pm [- m (\Delta F + F_{SPUR}) + n F_{OSC}] \qquad (10)$$

and m = n + 1

Solving as in (3) through (6), (10) yields

$$\Delta F = (m - 1 \pm 1) F_{IF} - F_{SPUR} / m$$

$$\boxed{\text{Thus, } \Delta F = m F_{IF} - F_{SPUR} / m, (m - 2) F_{IF} - F_{SPUR} / m \quad \text{HIGH SIDE}} \qquad (11)$$

Equations (9) and (11) yield two different spurious responses associated with each odd value of m. Letting ΔF go to zero produces the crossover spurious since the spurious crosses over the operating frequency.

If we let the spurious frequency or operating frequency be X (times) the IF frequency, then the order of the spurious for low side injection is 2X - 3 and 2X + 1. For high side injection, we obtain an order of 2X - 1 and 2X + 3. Thus, high side injection permits a higher IF frequency for a given RF operating frequency.

	sources: f 5	f 4	f 3	f 2	f 1
	0	0	0	0	0
	0	0	0	0	1
	0	0	0	1	1
	0	0	0	1	0
	0	0	1	1	0
	0	0	1	1	1
	0	0	1	0	1
3 sources:------>	0	0	1	0	0
combinations	1	1	1	0	0
	1	1	1	0	1
	1	1	1	1	1
	1	1	1	1	0
	1	1	0	1	0
	1	1	0	1	1
	0	1	0	0	1
4 sources:------>	0	1	0	0	0
combinations	1	1	0	0	0
	1	1	0	0	1
	1	1	0	1	1
	1	1	0	1	0
	1	1	1	1	0
	1	1	1	1	1
	1	1	1	0	1
	1	1	1	0	0
	1	0	1	0	0
	1	1	1	0	1
	1	1	1	1	1
	1	1	1	1	0
	1	1	0	1	0
	1	1	1	1	1
	1	1	1	0	1
5 sources:------>	1	0	0	0	0
combinations					

Figure (3)

Every zero accompanied by a one indicates that the frequency term (f 1 through f 5) is stepped from a minimum frequency to the maximum. If a 1 is accompanied by a zero then the frequency term is stepped back from f maximum to f minimum. The program routine includes 3 sources but any number is handled by this technique.

```
+++++++++++++++++++++++++++++
+        SPURSEARCH         +
+                           +
+ GENERAL SPURIOUS ANALYSIS +
+        VERSION  1.9       +
+    A.M. VICTOR   8/17/84  +
+++++++++++++++++++++++++++++

  COMPUTING CASE 1 (+,-,-)

N*F1               M*F2               K*F3
 0                  0                  3

DELTA IF1                         DELTA IF2

-38.6 KHZ                         -10300 KHZ

COMPUTED SPURIOUS FOR CASE 1 (+,-,-)

N*F1               M*F2               K*F3
 1                  4                  8
DELTA IF1                         DELTA IF2

 16.4 KHZ                         -10200 KHZ

COMPUTED SPURIOUS FOR CASE 1 (+,-,-)

COMPUTED SPURIOUS FOR CASE 1 (+,-,-)

N*F1               M*F2               K*F3
 1                  5                  8
DELTA IF1                         DELTA IF2

 10300 KHZ                          16.4 KHZ
```

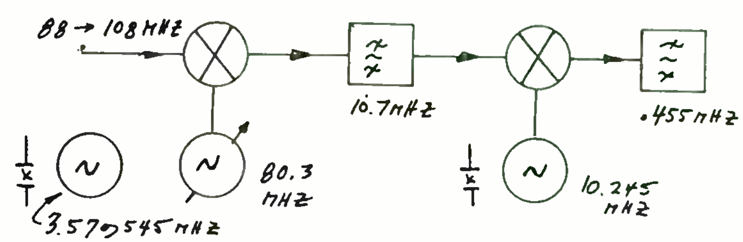


Figure 4 a

-25-88 MHz RECEIVER -

LOW SIDE INJECTION

HIGH SIDE INJECTION

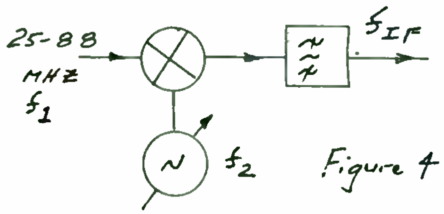


Figure 4 b

- VHF RECEIVER -

17.9 IF
LOW SIDE INJECTION

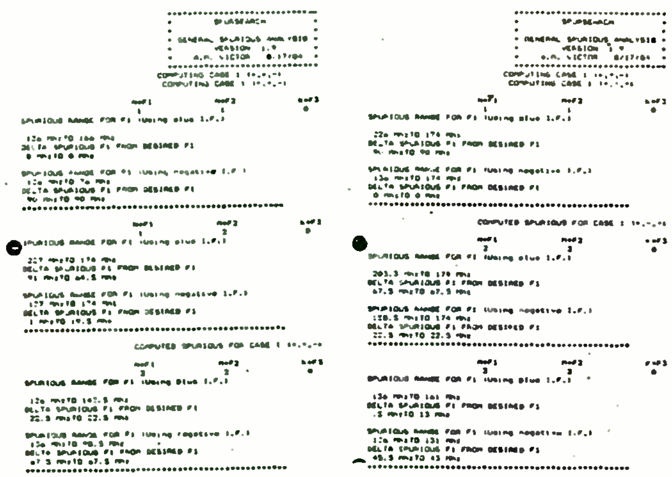
17.9 IF
HIGH SIDE INJECTION

21.4 IF
LOW SIDE INJECTION

21.4 IF
HIGH SIDE INJECTION

Figure 4 c

- VHF RECEIVER -



45.0 IF
LOW SIDE INJECTION

45.0 IF
HIGH SIDE INJECTION

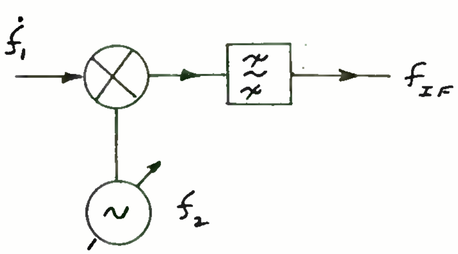


Figure 4 d

STRONG SIGNAL OVERLOAD OF BROADCAST RECEIVERS - THE PROBLEM DEFINED

by Don Jones
Mississippi Authority for
Educational Television
January 8, 1985

STRONG SIGNAL OVERLOAD OF BROADCAST RECEIVERS - THE PROBLEM DEFINED

By Don Jones, Mississippi Authority for Educational Teleivison

Television viewers living near broadcast transmitters operating in the television and FM services frequently experience interference to one or more television channels from strong signal overload. This condition may appear in the antenna preamplifier or in the receiver tuner. This paper provides information that defines the problem observed in the field, past attempts to resolve it, and recommendations of techniques to be applied in the future. It is intended for persons engaged in the design of receiving systems and particularly the rf front end portion of television receivers.

A Brief History of the Problem

Ever since the first broadcast stations began operation, problems have been observed with respect to front end overload. Stations operating in the AM, FM, and television services are not the only operators affected by this difficulty. Interference has been observed due to strong signals completely out of the desired band of operation. In the 1950s and 60s, when two-way radio communication with taxicab and police services became popular, an increase was noted in the number of complaints received concerning interference of these services to television reception. Today we have many sources to contend with in the rf spectrum, from video games to cable television rf leakage from trunk-line cables. A great number of complaints regarding television interference have been traced to users of the citizens band on 27 MHz and the frequent use of high power linear amplifiers by CB operators. In short, the potential for interference to television reception is greater today than in the past, and may be caused by a very large number of different sources operating singly or sometimes as an effect of multiple interfering signals. Today we will look at one portion of the problem and how it relates to receiver front end design.

Receiver overload from television transmitters is common in areas very near the offending transmitter site. It is important to note that the offending transmitter may be (and usually is) operating well within the spurious emission standards prescribed by the Federal Communications Commission. The observed problem has been shown to be due to simple fundamental signal overload of the receiver front end causing the transmitted picture and sound to appear on several, or all, channels on the television receiver nearby. The FCC recognizes this problem and has called the effect "blanketing." Broadcasters are required to relieve any cases of blanketing that occur near transmitter sites through the use of solutions appropriate to the specific case. Some of these solutions have been effective, and many have not. While it is true that reported cases of blanketing are usually eventually solved, it sometimes takes longer than desired to locate a workable remedy and apply it to the receiver affected. Further, a change in receivers by the resident living near the transmitter often has meant starting over on the blanketing problem and devising a solution to cure the "new problem" that comes with a change in receivers.

Blanketing most often occurs in the VHF portion of the spectrum where channels 2-13 operate. Lower channels are affected more severely than the higher ones, and channels 2-6 are affected more frequently than any others. UHF blanketing does sometimes occur, but it is relatively infrequent and usually easily solved by the installation of a directional antenna or simple trap on the receiver. UHF television transmitters have been observed to cause interference to other UHF services in the business radio portion of the spectrum and in a few specific cases to aircraft navigation equipment, but these are cases with some specific anamoly such as proximity to the transmitter site.

Interference to television reception is also caused by transmitters in the FM broadcast band. The FM band runs from 88-108 MHz and the lower portion of the allocated band is reserved for noncommercial broadcasters. Blanketing effects have been observed near stations operating on any frequency in the FM band, both commercial and noncommercial. Interference from FM stations is usually limited to one or two channels on the television receiver unless strong signal overload drives the receiver rf amp into saturation. When this occurs, interference is observed on all VHF channels on the receiver. Television channel 6 is most often affected by FM interference and is most severely affected by the noncommercial broadcaster. Interference can be classified into two types:

1. Fundamental overload effects
2. Adjacent channel effects

Each type produces similar results on the screen of the receiver, and both can be relieved by a single solution applied to the receiver. The symptoms are very similar, but the causes are very different. The rf front end designer must know and understand the difference in these kinds of interference in order to apply proper design techniques to relieve these symptoms.

The Problem Defined

One of the significant difficulties in dealing with broadcast interference problems is the lack of accurate information about the problem as it occurs in the typical residential installation. The FCC and the National Association of Broadcasters have conducted some rather elaborate and extensive laboratory studies of interference threshold in various receivers and the effect of nearby interfering rf carriers. Most of these studies are quite dated and do not represent the performance of modern receivers in the residential environment. The study techniques which are used and the information gathered are valid and accurate, but it does not correlate with what is observed in the field. Because of this lack of useful data, the Mississippi Authority for Educational Television conducted an extensive series of field studies to determine the extent of the problem as it occurs in a viewer's home, and to gather and document the information needed by rf engineers to design front end circuitry that would be relatively immune to interference from TV and FM broadcast operations. These studies were begun in late 1983 and are continuing at the present. The information provided by these studies will define the observed residential receiving system and the rf environment around it. Armed with accurate field data, designers should be able to improve front end circuitry only as necessary, and not to counter some laboratory defined threat that simply does not exist in the natural environment.

Viewers affected by interference will usually suffer a degradation of the television picture that appears as a fine herringbone pattern on the screen. It is rare for interference to occur to the television sound unless the interference mechanism is fundamental overload. When fundamental overload occurs, the picture is completely blocked and the sound will be severely distorted and garbled. Under this condition, the viewer will usually hear the offending station in the television audio. It is worthy to note that these effects occur from television transmitters also and not just FM operations. For viewers living near enough the transmitter site to experience fundamental overload, the offending television picture and sound frequently appear on the screen and in the sound along with the desired channel.

Received rf levels observed in the residential installation vary widely. The most significant factor in the level of interfering signal received by a viewer is simply his location with respect to the interfering transmitter. Levels in excess of +50 dB have been recorded in homes within one mile of the transmitter site. (Measured on a resonant dipole antenna, 30 feet high, and oriented for maximum signal strength of the offending transmitter. 0 dB = 1,000 microvolts in 75 ohms.) With such levels of rf present, problems occur with receiving systems picking up rf from sources not normally a factor. For example, 300 ohm downlead is commonly used in rural installations where man-made sources of noises are minimal. It is cheap and, when properly installed, offers superior transmission characteristics over coaxial cable because it is a balanced system. Common mode noise or rf does not easily penetrate the line. However, it is rarely installed properly and consequently the feedline itself will pick up significant amounts of rf because it becomes unbalanced. Coaxial line is only marginally better because it is an unbalanced system subject to ground loop effects. Further, the common crimp type connectors used by installers of coaxial line are often insufficiently crimped and allow a high resistance ground connection to be made to the shield on the coaxial cable. This results in significant rf ingress into the transmission line which cannot be reduced by proper orientation of the receiving antenna.

Another difficulty encountered in homes is the use of consumer installed equipment in the receiving chain such as antenna preamplifiers and multiple set couplers. Our experience in the field has been that most homes with interference problems have caused it themselves by improperly installing connectors, using the wrong preamp or splitter, and sometimes failing to use splitters at all; wires are just twisted together and insulated with "scotch" tape. All of these examples are typical of residential installations experiencing interference. It may be difficult to see, for the moment, how the rf designer can account for these difficulties in front end design; but there are techniques that are effective in offsetting some of these receiving system defects, and they have been successfully applied in field studies. Although the rf designer cannot correct defects caused by improper antenna system installation, he can correct many of the results. The common denominator in interference cases is simply controlling the strength of the undesired interfering signal. It is unimportant how this signal gets into the receiving system; it is only necessary to control it once ingress has occurred.

Control of fundamental signal overload effects is obtained by reducing the signal strength of the undesired signal prior to the first rf amplifier stage. The first rf amp may be in the television tuner itself or it may be the antenna preamplifier, if one is used in the installation. Notch filters, or traps as they are more commonly called, have been used with a virtual 100% success rate in controlling overload. They have relieved cases where the undesired signal is in excess of +60 dB. It is important to note that conventional trap technology is not usually successful in relieving such overload because of the limited notch depth obtained with simple broadband traps or various hi-pass or lo-pass filters. A four pole notch filter, cut to the frequency to be controlled with a notch depth in excess of -70 dB and excellent insertion loss characteristics at other frequencies up through the UHF television band, has been successful. The difference is that these traps are not tunable in the field. They are cut during manufacture so it is necessary to know the frequency of the signal to be controlled when the trap is ordered. Since the traps are only about 200 kHz wide, minimal disruption of the spectrum occurs on either side of the notch; and the viewer does not experience perceptible degradation of adjacent channel signals that are usually desired.

One other type of fundamental overload has been observed which requires the use of more than one trap. In a receiving system where an antenna preamp is used, the preamp nearly always suffers from overload. Placing a trap ahead of the preamp will prevent overload of the preamp itself and will attenuate the undesired signal by approximately -65 dB. If the interfering signal is very strong (for example, +50 dB), the resulting residual signal through the trap will be about -15 dB. Most preamps will give about +17 dB of gain at VHF. The signal sent down to the television tuner will be about +2 dB. A level of +2 dB will not overload a normally functioning receiver and will not cause adjacent channel interference. Problems arise when a secondary amplifier stage is located inside the house for distribution. Some distribution amplifiers have gains in excess of +40 dB. Use of such an amplifier brings the interfering signal back up to +42 dB and sends it on to the receiver where overload occurs. The solution to this is simply to use an additional trap at the distribution amp input to retain control of the undesired signal. Our studies indicate that this is a special case to be dealt with, because it has only been observed once in over 3,000 field tests. It is mentioned here because it is the only case observed that requires application of multiple traps in the system to correct the overload. All other cases to date have been corrected with the use of one single trap. (These traps are manufactured by PICO, Inc. and cost approximately $8.00 each in quantities of 100 up.)

Selectivity effects caused by strong signal overload can be almost mystical in the symptoms that they cause. A selectivity effect is any undesirable interference to reception that is caused by lack of selectivity in the receiving equipment itself. Television signals occupy about 6 MHz of spectrum space. Tuners in most receivers are quite broad in response. In the early days of television, when stations were few and far between, adjacent channel performance was not so important; there simply wasn't much to cause interference. The need for improvement became apparent with the introduction of cable television systems using all 12 VHF channels. Adjacent channel performance became a necessity for those viewers connected to cable in order to adequately separate all those stations on the cable. Manufacturers did improve the selectivity somewhat, at least to the point that receivers would work adequately on cable systems. Many people thought that such performance was all that was needed in television tuners. But what about selectivity at high signal strengths? Many tuners in use today are varactor tuned designs and offer good selectivity at signal levels typical of those provided by a cable connection, about +6 dB. These tuners exhibit a characteristic at high signal strengths that has the effect of reducing selectivity. The rf signal itself has a tendency to "swing" (for lack of a better term) the varactor and broaden the response. This effect has been observed to occur well below the third order intercept point and may very likely account for the apparent reduction in selectivity when a strong signal is received off-the-air. We have not progressed far enough in our studies to accurately define the mechanism at work here, only far enough to report the phenomenon and call it to the attention of the rf community for further study.

Selectivity effects manifest themselves as interference to the visual part of the television signal at relatively low signal strengths. For most receivers tested, interference appears as diagonal herringbone patterns in the picture. Interestingly enough, the sound portion of the television signal is rarely affected. We do not understand this effect well because the visual carrier is located 4.5 MHz further down in the spectrum than the aural carrier, and it would be expected that interference would occur to the carrier closest to the undesired signal. It is thought that the effect is observed in the picture because the visual portion of the television signal is essentially amplitude modulated, while the sound is frequency modulated. The capture effect of discriminators and the limiting effect of the audio i-f in the receiver may simply prevent the effects of the nearby carrier from appearing at the output of the audio chain. In any case, we have not seen cases where interference occurs only to the sound. Levels of +20 dB of undesired signal are generally required to produce this kind of effect. In some cases where the desired television signal is very weak, around -20 dB, interference from selectivity effects has been observed to occur with as low as +5 dB of undesired signal. This happens because the automatic gain control circuits in most television receivers are quite nonlinear in response. For example, a given desired-to-undesired ratio of rf levels may be observed to produce interference when the desired signal is at +10 dB; but, when the desired signal is reduced to +3 dB and the same ratio maintained, the observed effects are not the same. If the desired signal is increased in strength, a much larger ratio can be tolerated without visible interference appearing on the screen. For this reason, the use of desired-to-undesired ratios in establishing spectrum allocations for TV and FM services does not work. The FCC continues to predict possible interference on the basis of desired-to-undesired ratios even though the field results observed have never correlated with the predictions.

A part of our continuing studies is to determine a better method for predicting possible interference from adjacent channels without using ratios as the sole basis of calculation. Many other effects occur in the field that are not taken into account by the Commission in their predictions, including the effects of local man-made noise, utility line noise, and the effects of co-channel interference from another distant television station. One of the surprises in our findings was that these other effects quite often mask any selectivity effects that may otherwise be observed. In other words, the desired signal already has enough degradation from other sources that selectivity effects cannot be seen on the screen of the receiver. It is worth remembering that the real test of interference is whether or not the viewer can perceive any degradation in the desired signal as it appears on the screen and comes out of the speaker of his television receiver. One can certainly observe undesired products after the first mixer using a spectrum analyzer, but most viewers don't watch spectrum analyzers. In that sense, the performance desired in the tuner is simply that level which provides an interference-free picture and sound on the television receiver. This is the practical and proper measure of receiver performance in its intended use environment and not as may be measured in the sterile rf spectrum of a laboratory test.

Solutions to the Problem

Past efforts at solving the TV-FM overload problem have largely been successful due to the persistence of those attempting to eliminate it. Broadband traps are almost universally applied as the remedy and, unfortunately, frequently are not reliable as installed. For most devices, such parameters as lead dress, tuning, proximity to metal objects, and temperature all affect the trap adversely and require some attention on the part of the viewer. Because these difficulties are well known, the FCC and others have been reluctant to allow any factor for traps in calculating and predicting potential interference when spectrum is allocated. A secondary problem when using traps to relieve overload is that someone has to install them on the affected equipment. Historically, this has fallen at the feet of the offending transmitter operator; and it has usually been the station technical staff that has actually done the field work of installing and adjusting traps. Once installed, maintenance of the trap becomes the responsibility of the viewer.

Current successful approaches to control of overload and interference also make use of trap technology, but in a different manner. The trap devices in use are state-of-the-art, first manufactured about 1983 for interference control purposes. These devices are temperature compensated, enclosed in metal cases, fix tuned, inexpensive, and offer low insertion loss. They were originally developed as cable television premium channel traps, to prevent subscribers not wanting pay channels from receiving them. Since the technology to produce such devices in quantity was already in place, we made arrangements with the manufacturer to produce some of these cut to specific frequencies in the FM band to be used for interference control. The results have been consistent. Modern traps provide a level of interference control virtually not obtainable by any other means. For the first time ever, this performance can be had for less than ten dollars. However, traps are not the long-term solution to the problem at hand. They do represent a very well established interim solution to the selectivity dilemma, but more work is needed by front end designers to produce tuners that are not susceptible to selectivity effects. Tuners with adequate performance are the essential long-term solution to this part of the problem. Strong signal overload, on the other hand, will continue to be with us for quite some time until someone makes an rf amp with a +60 dB intercept point at a price cheap enough to be used in consumer receivers. Since this isn't very likely to happen in the near future, the use of traps to control fundamental overload will continue to be the only effective means of dealing with very strong signals.

Recommendations for RF Designers

Rf designers working on receiver front ends will find that the following techniques are helpful for improving the ability of a tuner to reject undesired signals:

1. Make the third order intercept point as high as practical. A third order point of +25 dB will achieve substantial improvement especially where the interfering signal is at VHF.

2. Use a split band design for the rf amplifier. Television channels 2-6 are located in the spectrum from 54 to 88 MHz. Channels 7-13 are located between 174 and 216 MHz. Most television tuners use a broadband rf stage that covers the entire band from 54-216 MHz. Some of the newer cable ready receivers go up to 300 MHz. Since most interference and overload problems are from transmitters in the FM broadcast band, a split band design inherently leaves this portion of the spectrum out and gains substantial improvement over the broadband type. Observations of broadband versus split band designs in preamplifiers have shown at least an 18 dB improvement in ability to reject undesired strong signals when the split band design is used. In split band designs, the critical parameter is good selectivity at the low end of the FM band where noncommercial stations operate. Since the television channel 6 sound carrier is located at 87.75 MHz and the noncommercial broadcaster may operate on 88.1 MHz, the slope of the response curve in the split band design is critical at the lower end of the FM band. Field tests to date have established that as much as -10 dB of attenuation is well tolerated by receivers of the channel 6 sound carrier without any perceptible degradation to the television audio. In fact, studies reveal that as much as 18 dB of attenuation can be introduced without perceptible effects; but the 10 dB limit is recommended to protect future use of the aural carrier by muiti-channel (stereo) sound television broadcasters.

3. Provide adequate i-f filtering behind the mixer. Most receivers today are using a saw device at the input of the i-f string to establish the response of the receiver i-f. This works very well and, in particular, cleans up most of the genuine selectivity effects that are not due to simple overload. The i-f saw device will not help in cases of fundamental overload; but, since these only occur to residents living very near the transmitter site, a relatively small number of viewers are affected.

4. Provide coaxial inputs on the receiver. Even though traps are available in 75 ohm versions and are easily installed in 300 ohm systems using transformers, the 75 ohm system is the preferred method because it encourages the use of coaxial cable rather than twin lead with its attendant problems. The use of coaxial cable inputs has become almost mandatory because of widespread penetration of cable into American households. Designers can expediate the conversion to coaxial systems by including coaxial inputs on the receiver.

5. Provide overload and selectivity specifications as a part of factory servicing information. Our studies have revealed that many television service technicians are not familiar with overload and selectivity effects and do not know how to obtain information on dealing with them. Manufacturers can include, for example, the third order intercept point of the receiver and a typical selectivity curve which will assist the technician in determining whether he has a genuine receiver defect (malfunction) or simply has a case of overload beyond the capability of the receiver. Similar specs would be extremely helpful from manufacturers of antenna preamplifiers and would assist the antenna installer in choosing a unit with sufficient dynamic range to tolerate the signal levels present at the site of installation.

6. Manufacturers should include in the operating manual for their product a short description of overload effects and how they appear on the screen of the receiver. This will help the user of the receiver to determine whether he has a case of interference or not. Frequently, a complaint will be registered about interference that will turn out to be a receiver malfunction. When the unit is repaired, the interference disappears. Advice from the manufacturer of the receiver to the user is simply one additional step that can be taken to help educate the public about the interference phenomenon and how it can be resolved.

7. Share what you learn with other rf designers that you know. The interference problem is likely to continue to worsen, because we continue to place more and more interfering signals in the spectrum. Earlier I mentioned video games. These have been a major problem because many people connect them directly to the receiver in parallel with the antenna lead; and, when the game is powered up, the receiving antenna now becomes a transmitting antenna. We have on file cases where one such arrangement caused interference for nearly a quarter mile in all directions. As devices such as this proliferate, we must provide front end designs that will reject out-of-band signals and handle strong signals nearby without causing perceptible degradation of the viewer's desired channel. The recommendations in this paper, if followed, will do much to provide the television industry with long-term solutions to a chronic problem.

RF OPERATION OF 450 VOLT VERTICAL POWER MOS TRANSISTORS IN AN ULTRA-LIGHTWEIGHT HF TRANSMITTER

Robert W. Vreeland
Senior Development Engineer
University of California
Research and Development Laboratory
4th and Parnassus Avenues, Room U-10
San Francisco, California 94143

Modern solid state transceivers are ideal for battery powered applications. However, the extremely heavy power supplies required, make them unsuitable for lightweight portable AC powered use.

We have taken a lesson from the AC-DC vacuum tube radios of the World War II era and have utilized high voltage MOS FETS in a two pound RF amplifier for the 3.5, 7.0 and 14.0 MHz amateur bands. The transformerless power supply weighs an additional pound and runs on either 220 volts or 117 volts AC. The amplifier is driven by a one pound dry battery powered heterodyne VFO. By using tuned push-pull operation we have reduced all harmonics sufficiently so that no low pass output filters are needed. We have used this transmitter to work Japan from San Francisco. The complete station can be carried in an attache case or in a camera gadget bag.

Traditionally, AC-DC receivers have utilized a half wave rectifier in a transformerless unregulated power supply. The supply voltage is usually in the 150 to 170 volt range. By utilizing a bridge rectifier on 220 volts or a voltage doubler from 117 volts it is possible to develop sufficient voltage to run a simple linear voltage regulator with an output of 200 volts.

This is an ideal power supply for an RF amplifier utilizing a pair of 450 volt MOS FETs. The allowable drain to source voltage swing is zero to 400 volts leaving a 50 volt margin of safety.

Some of the older vertical power MOS FETs have a rise time as short as five nanoseconds making them suitable for use in the 14 MHz amateur band. Unfortunately, the present design trend is toward higher power, lower frequency MOS FETs.

There are obvious safety problems resulting from transformerless design. The amplifier must have RF link coupling in and out. All circuits must be insulated from the metal panel and the panel must be grounded. Since the amplifier cannot be safely keyed, we must key the VFO. The VFO must be dry battery powered because a safe isolated power supply would add unnecessary weight.

These safety requirements were easily satisfied by building the RF power amplifier into a 2 x 4-1/2 x 6-1/2 inch plastic box with a grounded aluminum panel. The box has a plastic cover which closes to form a neat easily carried two pound package. Separate 2 x 3-3/8 x 5-1/2 inch plastic boxes contain the power supply and the VFO which weigh a pound each. The VFO is powered by a pair of nine volt MN1604 alkaline batteries. This three package configuration makes the transmitter easy to carry in a variety of containers.

The transmitter is designed to operate in the 7 and 14 MHz amateur bands using fourteen foot long inductively loaded dipoles. For 3.5 MHz operation, we use a tuned loop antenna formed from a fourteen foot length of RG-8/U coaxial cable.

The foregoing looks like a fairly straightforward design problem. Let us now look at the real world of power MOS FETs. They are marvelous devices when properly used but they do present a number of unusual problems. Unfortunately their true identity has been shrouded in a certain amount of mythology.

Myth Number One:

"Due to the absence of second breakdown and thermal runaway, MOS FETs are virtually indestructable." This may be true if one adheres strictly to the manufacturer's ratings. However, they will blow in microseconds if the peak current rating or the gate to source voltage rating is exceeded. To illustrate, let us look at the simple series regulator shown in Fig. 1. This is a good regulator with one percent regulation from no load to a full load of several hundred milliamperes. However, the pass transistor will blow when the switch is moved from "R" to "S". The charging current of a bypass capacitor as small as 0.02 microfarads exceeds the MOS FETs peak current rating. Since the source is momentarily grounded through the capacitor, the gate to source voltage rating is also exceeded resulting in a punctured gate. Obviously some form of very high speed current limiting is required. This is most easily done by the use of series limiting resistors as shown in Fig. 2. Excessive limiting will of course defeat the regulation for which the power supply was designed. It is usually possible, however, to divide the load so that the high current portion which does not require much regulation does not interfere with the low current well regulated loads. One must remember of course that the total peak charging current for all of the bypass capacitors must not exceed the MOS FETs peak rating. The reverse diode (DO) is built into the MOS FET and need not concern us for normal drain to source voltage excursions.

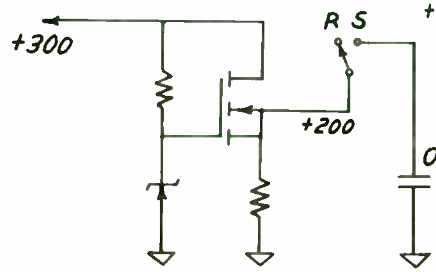


Fig. 1. A bypass capacitor as small as 0.02 Mfd will blow the pass transistor.

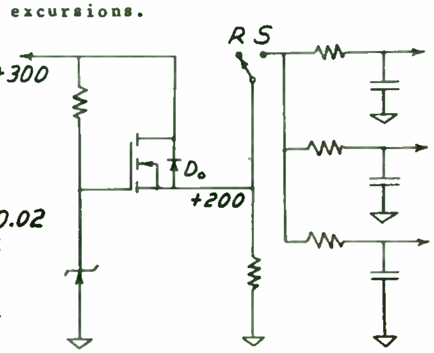


Fig. 2. By dividing the load, it is possible to stay within the peak current rating without excessive degradation of regulation.

Myth Number Two:

"Due to their insulated gate construction, MOS FETs require virtually no driving power." While this is true for DC operation, it certainly does not apply to RF amplifiers. Let us look at Fig. 3. Here the MOS FET is represented by a resistive channel. Current from source to drain through this channel is controlled by an electric field set up by the gate capacitor plate. As the driving generator frequency is increased, the capacitive reactance of the gate structure goes down resulting in an increase in driving current into the lossy structure of the MOS FET. Our amplifier is noticeably harder to drive at 14 MHz than on the lower frequency bands. The drive requirements are moderate, however, when compared to a bipolar transistor amplifier.

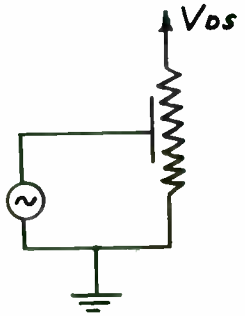


Fig. 3. The drive requirements for for MOS FETs increase with frequency.

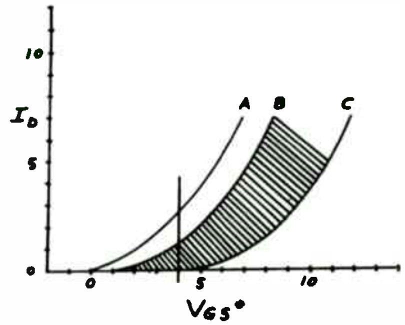


Fig. 4. Semiconductor manuals usually show a normalized transfer curve (Curve A).

Myth Number Three:

"Due to their self compensating thermal characteristics, MOS FETs will automatically adjust to share the load equally when connected in parallel."

This is true only when the MOS FETs are selected in matched pairs. Actually the gate threshold voltage may vary from one and a half to five volts from transistor to transistor. For this reason manufacturers print normalized rather than actual transfer characteristics curves. In Fig. 4, curve A is the normalized transfer curve. The actual curve may fall anywhere in the shaded region between curve B and curve C. Let us connect a curve B MOS FET in parallel with a curve C MOS FET. If we bias the parallel combination at plus four volts, the curve B MOS FET will carry a current of about one and a third amperes whereas the curve C MOS FET will not even turn on. Actually the manufacturer is probably over conservative in specifying such wide range of gate threshold voltage. Inspection of table I will show that the majority of IVN6000KNT's fall within the two to three volt region.

TABLE I

MEASURED IVN6000KNT
GATE THRESHOLD VOLTAGES (VOLTS)

4.538	3.897	2.981
4.478	3.815	2.964
4.219	3.804	2.898
	3.243	2.801
	3.199	2.748
	3.129	2.699
	3.079	2.677
		2.666
		2.578
		2.406
		2.403
		2.376
		2.069

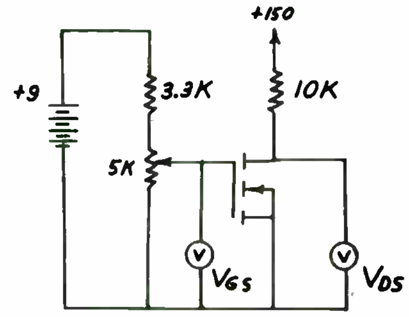


Fig. 5. The circuit used to measure gate to source threshold voltage.

In our power supply we use a parallel pair of MOS FETs as pass transistors. The parallel pair was selected using the circuit shown in Fig. 5. The gate to source voltage (VGS) is gradually increased until the drain to source voltage (VDS) drops from 150 volts to 100 volts. The VGS meter is then read. This is then the gate threshold voltage; VGS(th).

We did not select matched MOS FETs for our RF amplifier because individual bias voltage controls were provided. The amplifier was designed so that 10,000 ohm load resistors could be plugged into the output transformer jacks. The bias controls could then be used to determine the gate threshold voltages of the transistor pair using the previously described technique.

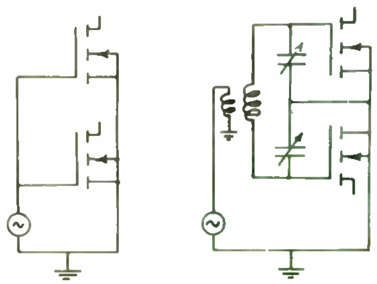
Drain to source voltage, drain current, rise time and input capacitance are perhaps the most important factors to be considered when selecting MOS FETs for RF use. The drain to source voltage rating must be high enough to allow the drain to swing up to double the power supply voltage with fifty volts or so to spare. The drain current rating should be sufficient for the desired power level. Also, don't overlook the peak drain current rating as exceeding this value will instantly destroy the MOS FET. The manufacturers provide safe operating area curves which are helpful for selecting the DC drain current operating point.

Rise time is perhaps the most critical factor affecting high frequency operation. It must be less than ten nanoseconds for fourteen MHz operation.

Input capacitance is also extremely important. It can range anywhere from about 200 to 300 picofarads for a good RF MOS FET such as the IVN6000KNT to more than ten times that value for the low frequency switching MOS FETs.

How to drive an input capacitance even as low as a couple of hundred picofarads can be a real problem if two or more MOS FETs are connected in parallel. In parallel operation, the input capacitances of course add as shown in Fig. 6A. If, however, we use a tuned push-pull circuit as shown in Fig. 6B the input capacitances appear in series across the tuning capacitor. This effectively cuts the input capacitance in half. Furthermore, the series combination becomes a part of the input tuning capacitor thereby serving a useful purpose. Push-pull operation effectively reduces the input capacitance by a factor of four over parallel operation.

The gate to drain feedback capacitance (Cgd) is important because it introduces positive feedback which can lead to oscillation and transistor destruction. In a tuned push-pull circuit conventional crossed capacitor neutralization can be used as in the case of vacuum tubes. Neutralization must be done with the full operating supply voltage applied because the gate to drain capacitance is a function of drain to source voltage. Neutralization is done with an oscilloscope connected across a dummy load on the amplifier output. The gate to source voltages are reduced to that no drain current flows. A small amount of RF excitation is then applied and the neutralizing capacitors are adjusted for a null on the oscilloscope. This adjustment is only approximate and must be touched up for stable operation at full output.



A B

Fig. 6. Tuned push-pull operation (B) effectively reduces the input capacitance by a factor of four over parallel operation (A).

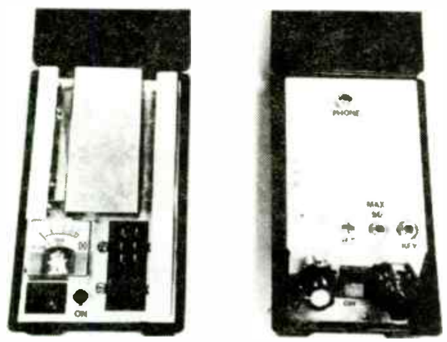


Fig. 7. The power supply (left) and the VFO (right) weigh one pound each. The turns counting dial is connected via a bellows to the slug which tunes the VFO. A frequency counter plugs into the "Max 50" jack. The "phone" jack is for the sidetone output.

Our power supply is packaged in a small covered plastic box (Fig. 7 left). Note the cooling fins and the miniature meter which is shunted for one ampere full scale. The regulated output and the transmitter safety ground are on the four pin connector. Separate grounded three wire power cords are used for 117 volt and 220 to 240 volt operation. Connecting the appropriate power cord to the six pin connetor makes the necessary circuit changes.

A simplified circuit of the power supply appears in Fig. 8. Diodes D1 and D2 form a voltage doubler with C1 and C2 for 117 volt operation. For 220 volt operation, the power cord completes the bridge circuit via jumper J1. The triangular symbols represent a floating common (not ground). The safety ground is not shown. A pair of Intersil IVN6000KNT MOS FETs serve as pass transistors and the zener is a 1N5388. The regulation is one percent from no load to a full load of 500 miliamperes.

Selecting diodes that would handle the inrush current into the 160 microfarad capacitors (C1 and C2) turned out to be a bit of a problem. We finally chose RCA SK3051's. It was necessary to limit the AC inrush current with a series five ohm, five watt resistor in order to prevent the on-off switch from burning out. Changing the power cord selects a 1-1/2 ampere fuse (F1) for 117 volts or a 3/4 ampere fuse (F2) for 220 volt operation. Output current limiting is provided in the transmitter package.

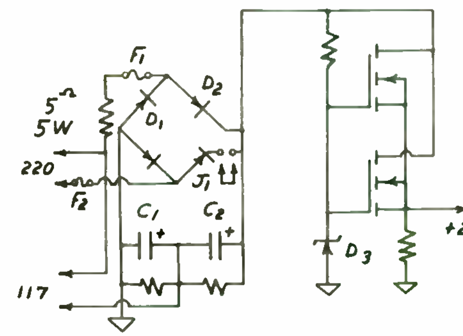


Fig. 8. A simplified circuit of the power supply. Regulation is one percent from no load to a full load of 500 milliamperes.

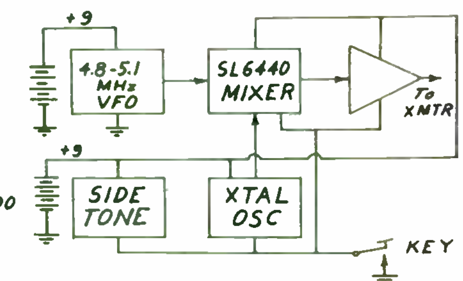


Fig. 9. A block diagram of the heterodyne VFO. Chirps are prevented by leaving the variable oscillator on during "key up" and while receiving.

For stability, we chose a heterodyne VFO (Fig. 7 right). It is powered by two nine volt alkaline transistor radio batteries (Fig. 9). The signal from the variable oscillator is mixed with one from a crystal oscillator to produce an output in the desired amateur band. The mixer is followed by a tuned output amplifier. Also included is an audio side tone oscillator.

The ground returns for all VFO circuits except the variable oscillator are keyed. An R-C circuit was inserted in the positive nine volt lead to slow the rise of the keyed signal in order to reduce key clicks. By not keying the variable oscillator we have avoided any chirp problems. Since none of the harmonics from this oscillator fall within the bands used it is left on while receiving. It is, however, very important to keep the harmonics low while transmitting as they will mix with the crystal oscillator to produce spurious outputs.

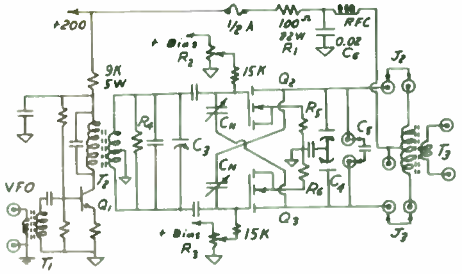


Fig. 10. A simplified circuit of the neutralized push-pull transmitter. For safety it is link coupled in and out. The triangular symbols represent a "hot" floating common (not ground). Note the separate bias controls (R2 and R3).

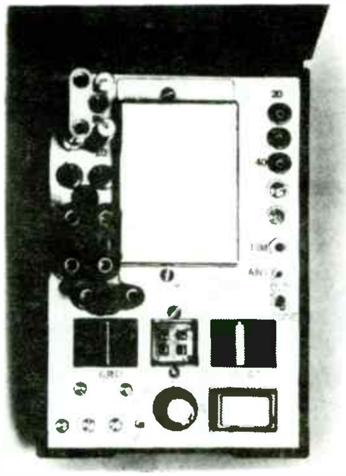


Fig. 11. The two pound transmitter set up for 3.5 MHz operation. Double banana plugs are used to select the tank coils. Note the thumbwheel tuning controls and the rocker type send-receive switch. The power MOS FETs are mounted in sockets under the aluminum cover.

A simplified circuit of the final amplifier is shown in Fig. 10. Since the power supply is transformerless, toroidal isolation transformers T1 and T3 are used for the amplifier input and the output. The triangular symbols represent a hot floating common (not ground).

The driver transistor (Q1) is an RCA SK3044. It is coupled via T2 to a push-pull pair of IVN6000KNT's (Q2 and Q3). Both the primary and the secondary of T2 are tuned. A separate transformer (T2) is used for each amateur band. The transformers are wound on toroidal cores and are selected by a band switch (not shown).

The positive bias on Q1 is increased slightly for fourteen megahertz operation in order to compensate for the increased drive requirements on that band.

The input of the push-pull pair (Q2 and Q3) is tuned by a 365pf plastic dialectric variable capacitor (C3). It is shunted by a fixed silvered mica capacitor. A two gang 100pf air variable capacitor (C4) tunes the amplifier tank transformer (T3). Both C3 and C4 are thumbwheel tuned. In Fig. 11 the thumbwheels are labeled "grid" and "plate".

By tuning both the input and the output of the final amplifier, we have reduced all spurious signals to more than 40dB below the carrier level. The effect of tuned push-pull operation on even harmonics can be seen in Fig. 12. The second harmonic is 56dB down whereas the third harmonic is only 45dB down. For 3.5 MHz operation the tank circuit Q is slightly better. There, the second harmonic is 58dB down and the third harmonic is 49dB down.

Separate toroidal output transformers (T3) are used for fourteen and seven MHz operation. They are selected by jumpers J2 and J3. The jumpers are on double banana plugs. For simplicity only one pair of jumpers and one output transformer are shown in Fig. 10. On 3.5 MHz a polystyrene fixed capacitor (C5) is plugged into another pair of banana jacks to bring the 7 MHz output transformer down to that frequency.

The jack at the center tap of the primary of T3 is of special interest. It is used to substitute a pair of 10,000 ohm, ten watt fixed resistors for T3. A dual trace oscilloscope is connected to monitor the DC levels on these resistors. The individual bias controls (R2 and R3) can then be used to find the gate threshold levels for Q2 and Q3. Having determined these levels, one can decrease the positive bias the same amount for both Q2 and Q3 in order to set the desired class C operating point. For safety, an insolation transformer should be used for the above adjustments.

A pair of 15pf air variable neutralizing capacitors (CN) is used. The neutralizing procedure has already been covered.

Additional stabilization is provided by a swamping resistor (R4). A separate resistor is used for each amateur band.

The negative feedback source resistors (R5 and R6) are each 3.3 ohms, one watt.

As previously mentioned, the 0.02 microfarad bypass capacitor (C6) is sufficient to blow a pass transistor in the power supply. This is prevented by the 100 ohm twenty-two watt resistor (R1). The resistor is a pair of eleven watt units mounted on a double banana plug. This arrangement permits free flow of air to cool the resistors. Careful insulation of the resistor leads is required for safety. The one-half ampere fuse protects the power supply in the event of failure of Q2 or Q3. It should be noted that the driver stage (Q1) derives its power via a separate series resistor so that its power supply regulation is not adversely affected by the heavy current through R1.

The transmitter puts out about twenty to thirty watts as shown in Table II. The final, amplifier efficiency is typically about thirty-five percent. While this might not be acceptable for high power equipment it is not a problem for our transmitter which requires less power than a one hundred watt light bulb.

Since the transmitter is not mismatch protected, tuning is done with a combination dummy load and antenna impedance bridge. This also prevents unnecesary radiation while tuning (1).

We are indebted to Frank Rittiman of Intersil for his assistance in getting us started.

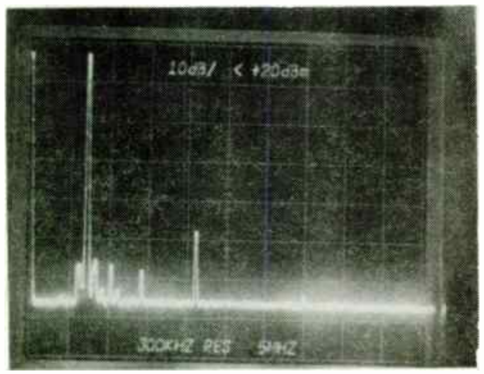


Fig. 12. The spectrum of the 7MHz signal. Tuned push-pull operation has reduced the second harmonic to a level below the third harmonic. The vertical scale is 10dB per division and the horizontal is 5MHz per division. Note the second harmonic of the variable oscillator in the VFO which is 55dB down at 10MHz.

TABLE II

PUSH-PULL [illegible] RF AMPLIFIER

Frequency (MHz)	Power Output (Watts)	Drain To Source Voltage (Volts)	Drain Current (Milliamperes)	Power Input (Watts)	Efficiency (Percent)
14.10	22.5	182	332	60.4	37.2
7.10	19.5	179	315	56.4	34.6
7.05	22.0	180	345	62.1	35.4
3.75	32.5	176	370	65.1	49.9
3.70	32.0	175	380	66.5	48.1
3.65	27.0	172	415	71.4	37.8
3.60	22.0	173	420	72.7	30.2
3.55	22.0	179	360	64.4	36.1

REFERENCES

(1) Volpe, J. Confidential; For Antenna Tuner Users Only; CQ August 1983: 56-59

LOW NOISE UHF TRANSMITTER DESIGN

by

Dennis A. Sweeney / Judd O. Sheets

E-Systems ECI Division
P.O. Box 12248
St. Petersburg, Florida 33733

ABSTRACT

The increasing reliance on frequency agile transmitters and receivers as a means to defeat jamming has created new interference problems for platforms with multiple R/T units. The problems can be solved by a variety of means but the most optimum way is by improved transmitter and receiver design. This paper will focus on the design of a low noise UHF transmitter that will provide 10 watts of RF output while maintaining a noise floor of less than -140 dBm/Hz and spurious levels greater than 120 dB below the carrier.

PROBLEM

The threat of jamming to military communications is real and has led to the development of many different anti-jamming (AJ) systems. Many of these systems (such as HAVE QUICK, SINCGARS, JAGUAR (V), and others) employ frequency hopping as the technique to defeat the jamming threat. Frequency hopping is a very effective means of AJ, but does create problems that didn't previously exist, or were easily solvable with the old single channel radios when installed in facilities or platforms that have a need for multiple communication sets.

The problem can best be illustrated by a simple example. For the installation shown in Figure 1A, a pair of R/T units are colocated, operating on separate antennas. The numbers shown are typical of present day UHF R/T units. There are two primary forms of interference that must be eliminated – broadband noise and

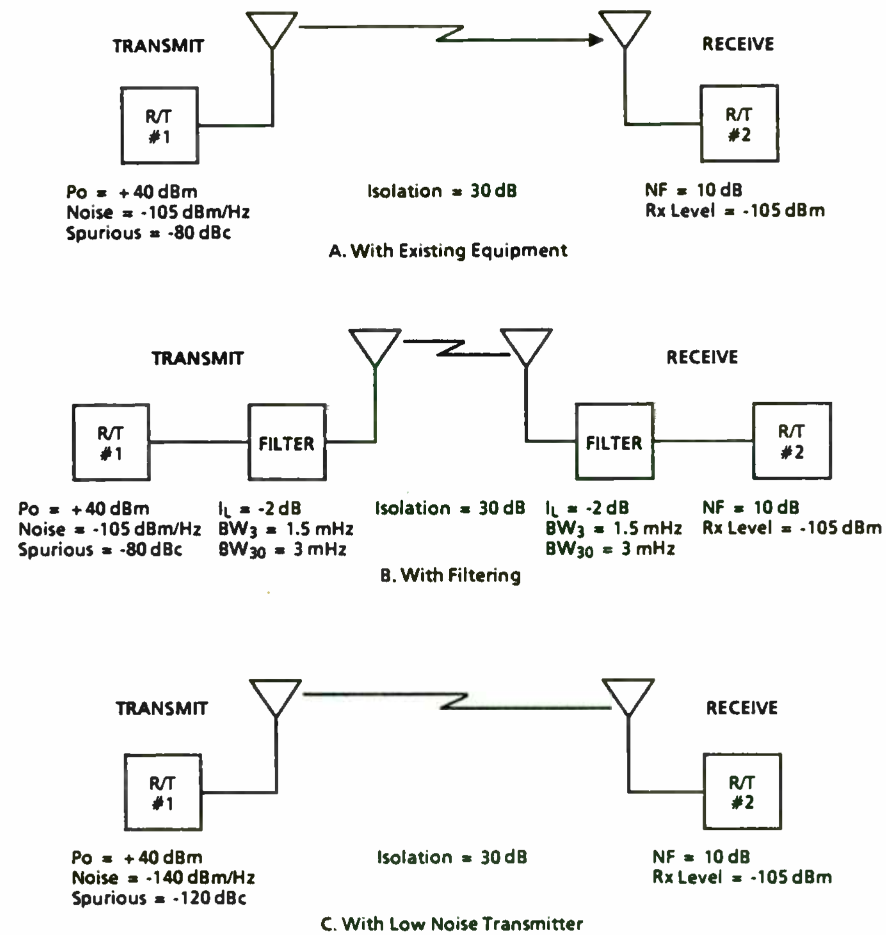


Figure 1. Colocated Communications

spurious outputs from the transmitter that are coupled into the receiver front end and mask out desired signals. There are other forms of interference that also exist such as transmitter and receiver intermodulation, receiver spurious response, and receiver desensitization, but these won't be addressed here since they are basically associated with receiver performance.

The numbers presented in Figure 1A show that the transmitter's broadband noise, as seen at the receiver's input, is -135 dBm/Hz, which is 29 dB higher than the receiver's noise floor. Since the broadband noise is present across the entire frequency band, this interference is present at every frequency to which the receiver is tuned. In addition to the broadband noise, there are spurious outputs from the transmitter that also interfere with the receiver. As shown in Figure 1A, the transmitter's spurious level at the receiver is -70 dBm compared to a receiver sensitivity of -105 dBm. Thus, at every point at which a transmitter spurious occurs, that receiver channel will experience interference.

SOLUTIONS

The problem of interference between colocated transmitters and receivers has been shown to be real. The two best solutions to the interference problem are to filter the output of the transmitter or to improve the performance of the transmitter itself. Each of these solutions will be discussed and some of the advantages of each will be pointed out.

The first approach to be considered is to filter the output of the transmitter to reduce the level of broadband noise and spurious. As can be seen from the signal flow of Figure 1B, the transmitter broadband noise level, as seen by the receiver at its tuned frequency, is reduced to -169 dBm/Hz, which is below the receiver noise level of -164 dBm/Hz. The transmitter spurious, as seen by the receiver, has been reduced to -104 dBm, which is equivalent to the minimum sensitivity of the receiver. This performance is acceptable and is achieved through the use of a low loss, high Q fixed tuned filter following the transmitter to reduce the level of broadband noise and spurious output.

When existing single channel R/Ts are replaced by frequency hopping R/Ts, the problem becomes more difficult. One approach is to maintain the same performance on the R/T and to try to clean up its output with a frequency agile filter. However, to achieve the same characteristics as the fixed tuned filter requires a fairly large complicated frequency agile filter. Because the filter sections in a tunable filter are more lossy, they must be intermixed with amplifier sections to maintain the necessary low insertion loss. The result is a filter that meets the requirements but is large and power consumptive. For some applications, this is a good approach since it works on both the transmit and receive signal paths and improves performance for all types of interference.

However, when space, weight, and power are at a premium, the best solution is to improve the performance of the transmitter to the point where a separate filter is no longer required. As can be seen in Figure 1C, if the transmitter's broadband noise can be reduced to -140 dBm/Hz, then the noise seen by the receiver is -170 dBm/Hz, which is well below the receiver's own internal noise floor. Additionally, if the spurious output of the transmitter is kept to -120 dBc, then the receiver will only see -110 dBm of spurious signal, which is also below the receiver's sensitivity.

The remainder of this paper will describe the approach taken to achieve the above mentioned performance.

APPROACH

The transmitter developed derives its performance from a low noise synthesizer voltage controlled oscillator (VCO), a frequency agile bandpass filter, and a low noise figure power amplifier (PA) stage. Each of the stages must have characteristics chosen to not degrade the overall signal-to-noise ratio (SNR) of the transmitter output.

Originally, it was proposed to use a recently developed high output power VCO driving a low noise power amplifier. The oscillator itself exhibited an SNR in excess of 185 dBc/Hz, slightly better than the required performance of 180 dBc/Hz (all SNRs taken 7 MHz from center frequency). While this approach seemed feasible from a block diagram standpoint, the possibility of spurious signals from the phase locked loop (PLL) divider modulation and switching power supply ripple, led to the introduction of a bandpass filter module between the synthesizer and power amplifier.

The low noise transmitter consists of four functional blocks, divided into five separate modules – power supply, control interface/synthesizer PLL, bandpass filter, and power amplifier (PA). Each of the modules is contained in an individual aluminum chassis, which are stacked together and interconnected to form a complete transmitter. A simple block diagram and physical outline is shown in Figure 2.

The electrical specifications imposed on each of these modules was chosen by beginning at the signal source, the VCO, and determining its output SNR, then the rejection necessary in the bandpass filter, and finally the required gain and noise figure of the PA.

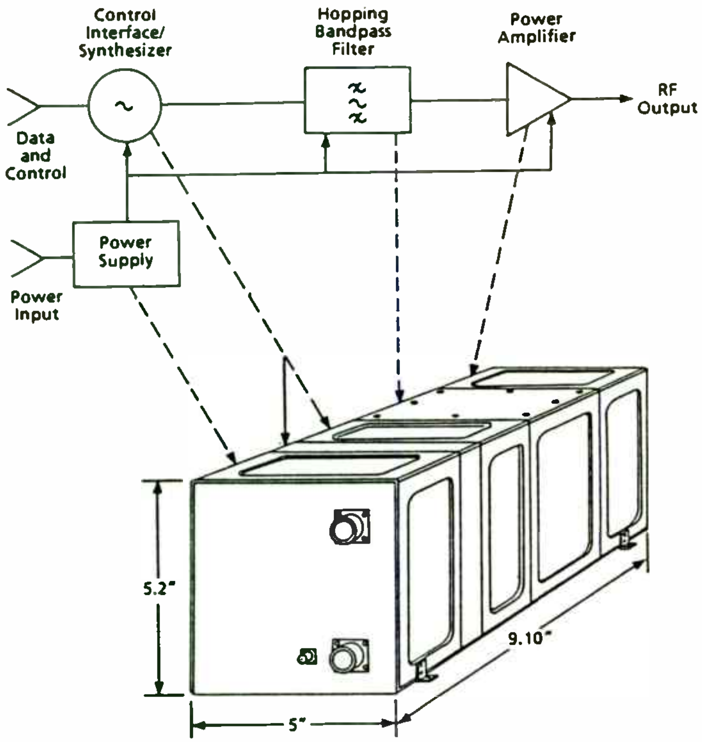


Figure 2. Low Noise Transmitter Block Diagram

The VCO chosen for the transmitter was selected for its low noise and frequency hop characteristics. This VCO has an output signal-to-noise of 165 dBc/Hz, while the required transmitter output signal-to-noise is 180 dBc/Hz. Therefore a minimum of 15 dB filter rejection is required to reach the output noise requirement. A prototype of the filter provided in excess of 25 dB of rejection at a 7 MHz offset with an insertion loss of less than 4 dB. The performance of this filter is used in the overall analysis.

The VCO output level is 0 dBm, requiring 40 dB of amplification to reach the necessary output power of 10 watts. Gain must be included as well to overcome bandpass filter loss, PA power control insertion loss, and output harmonic filter loss. Several stages are required to achieve this gain and an appropriate point must be found to insert the bandpass filter. The location selected must be at a level low enough to allow reasonable filter PIN bias voltages yet high enough so that the filter output noise level will be above thermal noise and its rejection not "wasted".

A signal flow diagram showing the noise and signal levels, block interconnection, and block performance is presented in Figure 3. It can be best understood by following the signal level backward through the chain, dropping down in signal level by the SNR of the VCO and following the noise level forward. As shown on the output end of the diagram, the resultant SNR is in excess of the 180 dBc/Hz requirement which, at 10 watts output level (+40 dBm), provides a noise floor of less than -140 dBm/Hz.

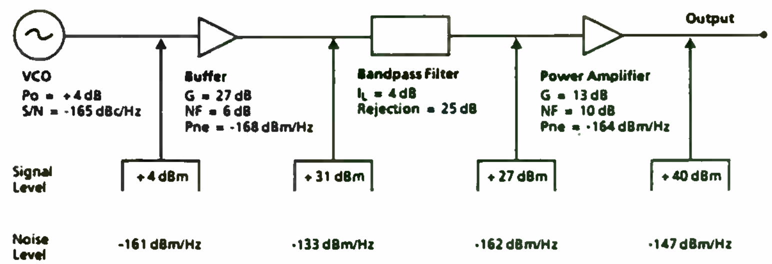


Figure 3. Signal and Noise Level Analysis

MODULE DESCRIPTIONS

Power Supply

The power supply is contained in the interface connector end of the transmitter. It converts the 28 Vdc primary power to +28, +12, +8, and -12 for general use in the transmitter and +300 and -2 for PIN diode bias in the bandpass filter module. The two sections of the power supply each employ a flyback converter operating at over 200 kHz to ease filtering and reduce component size. Each of the outputs is always within 5% of nominal voltage and noise plus spurious levels are at least -60 dBV (+300 and -2 Vdc outputs -45 dBV). Of particular concern is the +28 Vdc output noise as any ripple on this line can directly modulate the PA collector and increase the transmitters broadband noise. To meet transmitter output noise and spurious requirements, this output was constrained to a maximum of -90 dBV ripple.

Synthesizer

The Control Interface and PLL modules make up the synthesizer. The control interface module receives the balanced pair control lines, performs serial to parallel conversion of the frequency data, and converts the modulating data stream to a filtered analog voltage. The analog data stream is used to FM the PLL reference to provide FSK modulation of its output. The PLL uses a 100 kHz reference and achieves a 25 kHz channel spacing by SSB modulating the programmable divider input. Critical to noise performance, the VCO uses a transmission line resonator to maximize resonator shielding and microphonic rejection. Coupling this to a low noise gain device yields an oscillator of good noise performance and high repeatability. A low noise amplifier buffers the oscillator output up to a level of 1.25 watts while maintaining the oscillator 165 dBc/Hz SNR.

Hopping Filter

A frequency agile bandpass filter is used to improve the output SNR of the PLL and remove any incidental spurious output. The filter selected to satisfy the design requirements is a two pole combline structure tuned via PIN diode switched capacitors. A maximum insertion loss of 4 dB and a minimum selectivity of ± 7 MHz of -25 dB were the primary requirements and in large part determined the minimum size of the module. The filter module's cavity volume was designed to yield an unloaded cavity Q of approximately 1500 which, in conjunction with a tuning network Q of greater than 500, provides ample design margin on the insertion loss and selectivity requirements. The loaded Q of the filter was designed to be 108 at the high end of the band and varies from this value by less than 10% over the entire band. The general trend is that of decreasing Q with decreasing frequency. A 7-bit, binary weighted electronic tuning network provides a tuning increment of 700 kHz between adjacent tuning codes. A 3 dB bandwidth of approximately 2.5 MHz, ensures that the filter is used over a relatively small portion of its nose bandwidth. All required control and PIN diode driver circuitry is included in the module.

Power Amplifier

Gain is necessary after the hopping filter to boost the filtered output to the required output power. However, the gain, even at this high signal level, must exhibit a good noise figure to maintain the SNR of the drive signal. Filter output level must also be chosen high enough so that a reasonable PA noise figure does not result in signal-to-noise degradation at this point. At the 10 watts power level, a gain of approximately 15 dB coupled with a noise figure of 10 dB gives good margin on the output noise level requirement of 180 dBc/Hz. This level of performance is obtainable with UHF power FET devices. Exceptional care must be taken to protect this noise level from drain and bias supply noise contamination.

Transmitter spurious output level is also dependent on this module. A -120 dBc output requirement and estimated -25 dBc PA harmonic levels necessitate a output low pass filter with 100 dB of second harmonic rejection to allow 5 dB of margin. Allowance must be made for the consequences of this unusual filter in terms of increased passband loss, large size, and cost.

The PA module also contains circuitry to level power output over a range of drive levels and a system for controlling power output rise and fall times.

TEST RESULTS

Broadband Noise

To confirm the predicted performance of the transmitter chain, prototypes of the modules were used to assemble a nonhopping transmitter. A transmission line VCO, 31 dB gain buffer, hopping bandpass filter, and PA were cascaded and the resultant intermediate and output SNRs measured. An E-Systems manufactured cavity bandpass filter was used in these measurements to reject the carrier power and prevent spectrum analyzer front end overload and extend dynamic range. A bandpass filter approach is preferred in some cases over notching the carrier out with a band-reject filter to prevent analyzer overload from harmonic energy.

At the input to the hopping bandpass filter, signal level was +31 dBm and noise 7 MHz away was -135 dBm/Hz, an SNR of 166 dBc/Hz. Output signal level (including 2 dB additional loss from a circulator to prevent filter interaction) was +26 dBm and output noise -163 dBm, an SNR of 189 dBc/Hz. The difference between the two SNRs (23 dB) is the improvement due to the hopping filter. This figure is slightly less than the 25 dB minimum filter rejection due to the present noise measurement limit.

The 15 dB gain PA, minus low-pass filter losses and automatic level control circuitry losses, yields a PA net gain of 13 dB. Power out of the PA module was +41 dBm, 1 dB over predicted because of less than expected filter loss. Noise at this point measured -146 dBm/Hz, giving an SNR of 187 dBc/Hz, well within requirements.

Spurious

To ensure that no spurious output levels would exceed -120 dBc, measurements were taken of harmonic levels of order 1 - 30. These were taken with a fundamental frequency at the lower band edge to simulate worst case low pass filter rejection. By the 10th harmonic (2.25 GHz), levels were down to -78 dBc and by the 20th (4.5 GHz) harmonic levels were below -100 dBc. The filter selected maintains 100 dB of attenuation to 1 GHz, but is allowed to have decreased rejection above this frequency to minimize size and passband insertion loss. Above 7 GHz no filter rejection is necessary as comparatively low PA switching speed prevents generation of significant energy and the PA output network reduces this level even further. The combination of the PA harmonic rolloff and output filter rejection guarantee an output spurious level of less than -120 dBc at all frequencies as shown in Figure 4.

SUMMARY

The introduction of frequency hopping receivers and transmitters into platforms with multiple communications links has compounded an already severe mutual interference problem. Adding a frequency hopping filter on the output of the transmitter can solve the problem, but its size, weight, and power consumption prohibit it from being used in many applications. The use of a low noise transmitter can also solve the primary interference problem of broadband and spurious outputs. This paper has described the design of a low noise transmitter and has presented test results verifying its performance. A prototype of this transmitter is being built and a production version will be available in the near future.

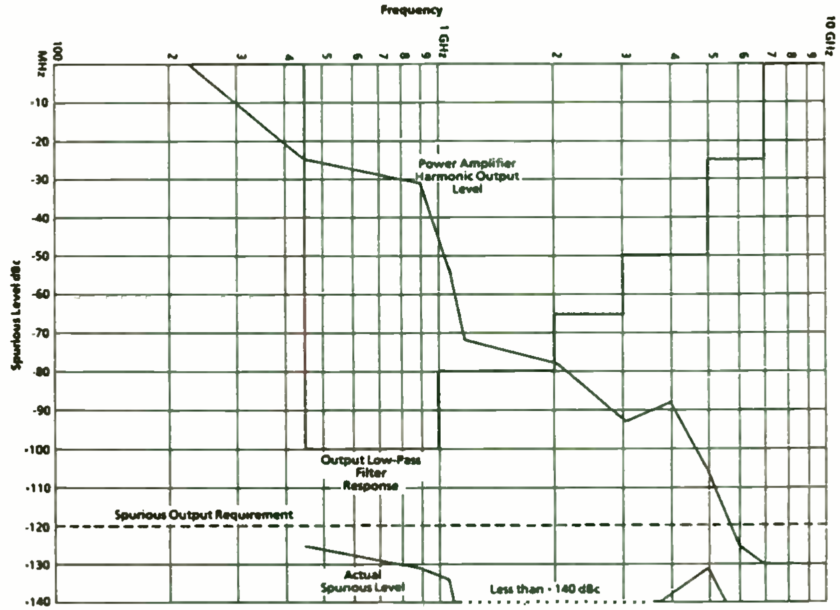


Figure 4. Spurious Output Levels

2 GHz AMPLIFIER PROVIDES TELEMETRY CAPABILITY

Joseph F. Lutz

ABSTRACT

This paper describes a 2 GHz amplifier which is used in a solar powered repeater. The amplifier also has a feature for the transmission of alarm telemetry through the use of low level AM. The alarm telemetry allows the user to monitor the repeater's status and also various site parameters. The PIN diode circuit which functions as an amplitude modulator also performs temperature compensation of the amplifiers gain.

INTRODUCTION

The amplifier to be described was designed to be used in a 2 GHz solar powered repeater. Solar powered repeaters are used in remote locations such as mountain tops, deserts and also in rural areas. Being located in such geographic areas means that commercial power is not available. A solar powered repeater, therefore, offers a considerable advantage over a conventional terminal, which would require the installation of generators and frequent maintenance trips. The solar powered repeater also offers a significant advantage over a passive reflector in that the repeater is 1/10 the cost of a passive installation. A repeater, as distinct from a radio terminal, uses no up converters, down converters, oscillators or modulators. The principal components are bandpass filters, isolators, circulators, and amplifiers. The amplifier is, in fact, the only active component in the repeater. Because of this, the repeater is capable of withstanding large temperature variations thus eliminating the need for shelters and in some cases, active temperature control, i.e., heating or air conditioning. Thus, the repeater is then a straight-through device which amplifies the RF signal and then retransmits it. No provision is allowed for accessing the information which is being transmitted.

The amplifiers are linear and the repeater is, therefore, capable of retransmitting any modulation format. A unique feature of the repeater is a patented technique using a combination of isolators, circulators, and filters which allows a single amplifier to simultaneously amplify east and west path signals. A block diagram of the repeater is shown in figure 1.

Based on system requirements, the amplifier has to meet the following requirements:

Frequency Band	1.7-2.3 GHz
Noise Figure	4 dB
Power Output (1dB compression)	+23.5 dBm
Gain	62 dB min.
Gain Flatness	±1.5 dB max.
VSWR input/output	2.1:1 max.
AM/PM conversion	40 dB max.
Group Delay	1ns max.
Battery Voltage	10Vdc-18Vdc
Current Consumption	300mA @ 12.6Vdc
Temperature Range	-40°C to +60°C
Alarm Telemetry	Low level AM

THE OVERALL AMPLIFIER

The amplifier is a 7-stage bipolar transistor amplifier operating in the 1.7 to 2.3 GHz frequency band to provide a gain of 62 dB and output power of a quarter watt. The first four stages are low noise transistors which provide 40 dB of gain. The fifth and sixth stages are medium power transistors which provide 20 dB of gain. The seventh stage consists of a linear high power transistor to provide 8 dB of gain. The bias circuit is a voltage feedback and constant base current source design to provide DC stability. The amplifier draws nominally 300 mA of current and has a maximum noise figure of 4 dB.

The modulator is placed between the third and fourth stages to provide a wideband attenuation of 6 dB at room temperature, giving the overall amplifier a nominal gain of 62 dB at room temperature. To compensate for the transistors, the modulator provides 4dB of attenuation at 60°C and 8 dB attenuation at -40°C which results in a constant output level over the temperature range.

TELEMETRY

Giving the amplifier a telemetry capability allows users to monitor the status of a variety of conditions at the repeater site without an operator being present.

This is a very important feature of the amplifier. As previously mentioned, repeaters tend to be in remote locations such as mountain tops: some sites being accessible only by helicopter. A trip to a site can, therefore, be expensive and difficult and is to be avoided except when absolutely necessary. Alarm telemetry, therefore, becomes essential.

The telemetry system is a one-way system communicating information from the remote RF repeater by impressing low-level amplitude modulation on the microwave signal as it passes through the repeater. The status information is retrieved at the IF section of the next terminal in the system. This can be done either with envelope detection or by monitoring the AGC voltage. The status information may thus be sent to the nearest system supervision or maintenance center by any of the conventional supervisory systems. Status information at the repeater is encoded in a repetitive sequential format: the rate is 10 bits/second. This data rate is satisfactory for telemetry and low enough so as not to interfere with normal radio link performance of an analog FM or PSK, QPRS, or other digital modulation format. Conditions to be monitored can include:

1) Battery voltage
2) Battery temperature
3) RF output
4) Battery charging
5) Low waveguide pressure
6) Security

A security alarm would indicate such things as a site gate that was open or also an open repeater cabinet. Alarms can be programmed as necessary by the user.

ENCODING FORMAT

Alarm information is transmitted by sequential frames. The basic frame length is 16 bits. Of these, 7 bits are assigned to on-off status or alarm conditions (X1 thru X7). Three bits are for telemetry of battery voltages and temperature. Each voltage and temperature reading is transmitted as an 8-bit word. Two voltage words, one for each battery, are transmitted in sequence over X9. The X8 bit is used to identify the two voltages. Thus it takes 16 frames to update the voltage readings and 8 frames for the temperature. There are 6 "F"-bits for frame synchronization of the receiver. The basic frame format, therefore, is

F1F0X1F1X2X3F0X4X5X6F1F0X7X8X9X10.

The encoded information is then used to control the attenuation of a PIN diode attenuator in series with the RF path imparting a low level AM to the signal.

The telemetry transmitter is shown in figure 2.

The alarm inputs enter on the left. Alarms are allowed to have only two states, either on or off. The alarms are buffered and then sent to a 16 channel multiplexer. Two battery inputs are shown as each site usually has two batteries for redundancy. The A/D converter converts the battery voltage into an 8 bit word for serial transmission. The battery temperature is also converted into an 8 bit word for transmission.

As the 8 bit counter counts up, F1 through X10 is sequentially selected. The multiplexer output assumes the selected value and causes the modulator to assume an attenuated or unattenuated state thus imparting AM to the microwave signal. LED indicators are also present to give visual alarm status to repair personnel that have been dispatched to the site.

TELEMETRY DECODER

The telemetry decoder is shown in figure 3. Here the signal is shown coming from the IF strip for envelope detection. The signal could just as well have come from the AGC detector of the radio--in which case an envelope detector would not be necessary. The decoder is located at a terminal site. Once sync is achieved, the alarm data is read out of the 16 bit shift register and latched. Decoded alarm data is presented to the monitoring equipment as a PNP open collector output. A non-alarm condition is signified by an open circuit while an alarm places the output into a grounded condition.

The temperature and battery words are sequentially read out of the shift register, decoded and displayed. A D/A converter also provides an analog output of both the temperature and battery voltage. This allows for continued alarm transmission via analog service channel, or redigitized in the case of a digital overhead channel such as is typically used in PSK radios.

PIN DIODE MODULATOR

As previously stated, amplitude modulation is impressed upon the microwave signal by varying the attenuation of a PIN diode attenuator which is in series with the signal path.

The attenuator is shown schematically in Figure 4. A balanced configuration of two AL67003-7 PIN diodes is connected in series between two branch-line couplers. The couplers function in the following way. First, input power is split evenly to pass into the diodes. Power passing through the diodes is recombined by the second coupler. This power reinforces at port 3 (output port) and cancels at port 4 (isolated port). When the diodes are fully turned on, nearly all of the input power goes to the output. At zero bias and points in between, power is reflected from each diode back into the first coupler. Because of the quadrature phase relationships in the coupler, the reflected power cancels at port 1 (input port) and reinforces at port 2 (isolated port). This ensures a good match at the input for all attenuation settings; subject to good balance in the couplers and good match between the diodes. The two isolated ports (ports 2 and 4) are terminated with 50-ohm chip resistors. RF grounds for the resistors are provided by open-circuited stubs a quarter-wavelength long at 2 GHz, the center frequency of the amplifier.

TEMPERATURE COMPENSATION - DC CIRCUIT

TEMPERATURE COMPENSATION CIRCUIT PROVIDES THREE DEGREES OF FREEDOM

Originally it was attempted to temperature compensate the amplifier by using a single thermister to vary the current through a PIN diode attenuator similar to that used for telemetry. After much effort, it was finally concluded that the thermister could not track the amplifier gain variation over the entire temperature range (100°C). What was required was a temperature compensation circuit which assumed the characteristics of figure 5. Furthermore, it was required that the individual temperature breakpoints, T_1, T_2 and T_3 to be set independently. A very simple and quite effective circuit to accomplish this is shown in figure 6. Below -30°C a minimum of current (0.05 mA) is required in the PIN diode to produce an attenuation of approximately 8 dB. At temperatures below -30°C, V_T is below the reference voltages on inverting terminals of U1A and U1B. Thus the output voltages of U1A and U1B, are driven toward the negative rail (OV) This leaves only R10 to provide a constant current of 0.05 mA to the diodes. By using gain-lowering type feedback resistors, R4 and R5, the actual transitions were much smoother than shown in figure 5. At temperatures above -30°C U1B is turned on and its output voltage begins to move toward the positive rail and drives an additional current of 0.11 mA through the PIN diode producing an overall attenuation of 6 dB. As the temperature moves above 30°C U1A conducts supplying an additional 1.9 mA of current and reducing the attenuation to 4 dB.

The circuit of figure 5 can be modeled as shown in figure 7. Here, V_A, R_A, V_B, and R_B are the Thevenin Equivalents looking back from the op-amp inverting terminals. V_{T1} and V_{T3} are the voltages at which op-amp B and A, respectively begin supplying current to the PIN diodes. We have assumed here that the op-amp circuits do not load the voltage divider string. As a practical matter, a voltage vs. temperature curve should be generated so that V_T will be well known. The precise voltage which corresponds to desired trip temperature can then be substituted for V_{T1} and V_{T3}. V_B and V_A should then be set equal to V_{T1} and V_{T3}, respectively. Working backwards then R1, R2, and R3 can be calculated. The R4 and R5 can be chosen based upon how rapidly one desires to make the transition from one attenuation state to the next; remember V_T is continuously varying with temperature. Resistors R8 and R9 can be calculated with load line analysis but is more easily and accurately done empirically. R10 is made very large to simulate a constant current source. R10 is also best adjusted empirically after R8 and R9 have been determined.

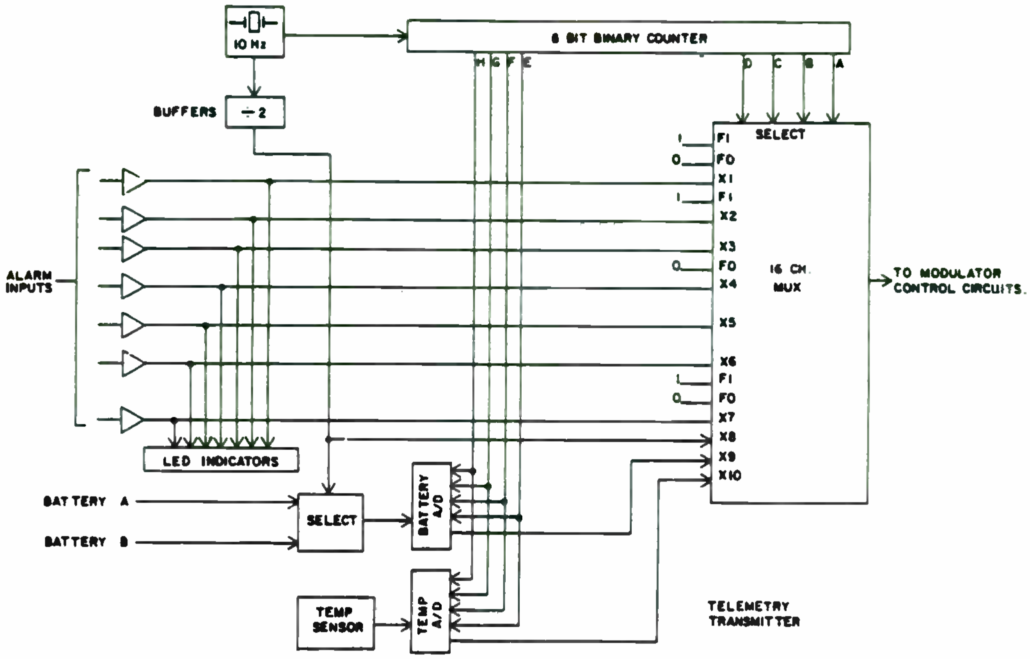


Figure 2

SOLAR POWERED REPEATER

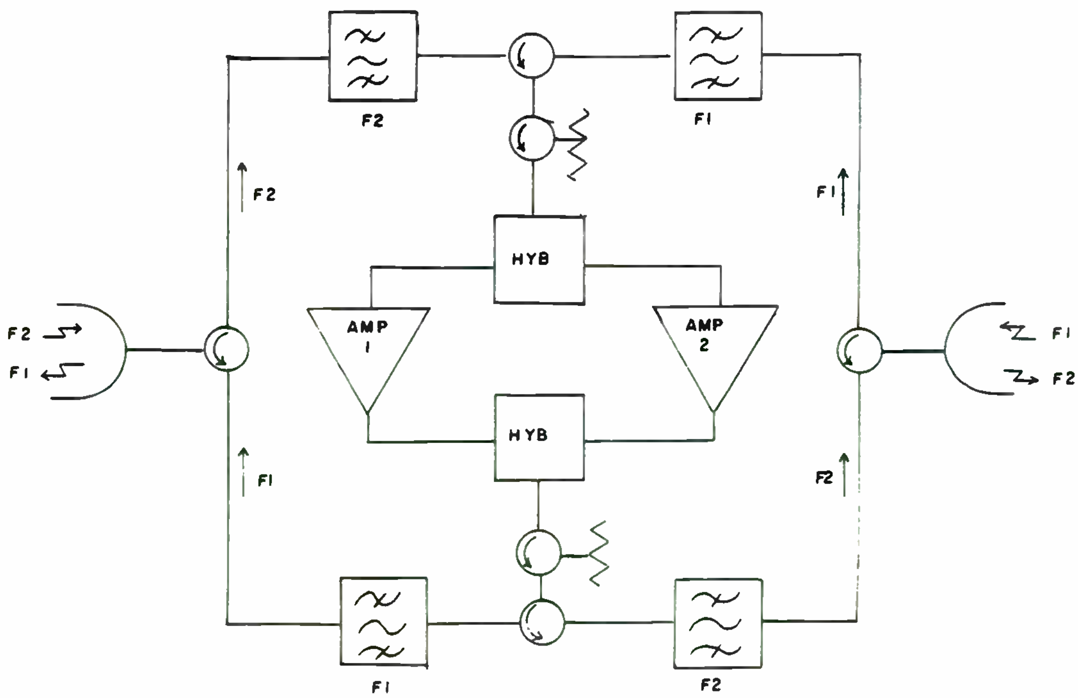
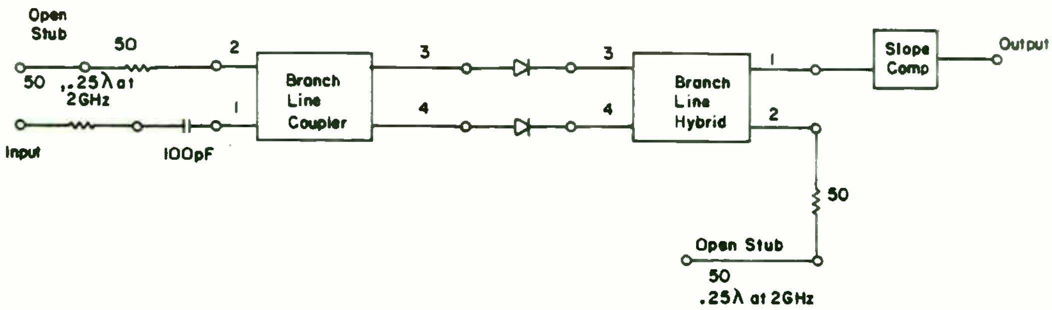
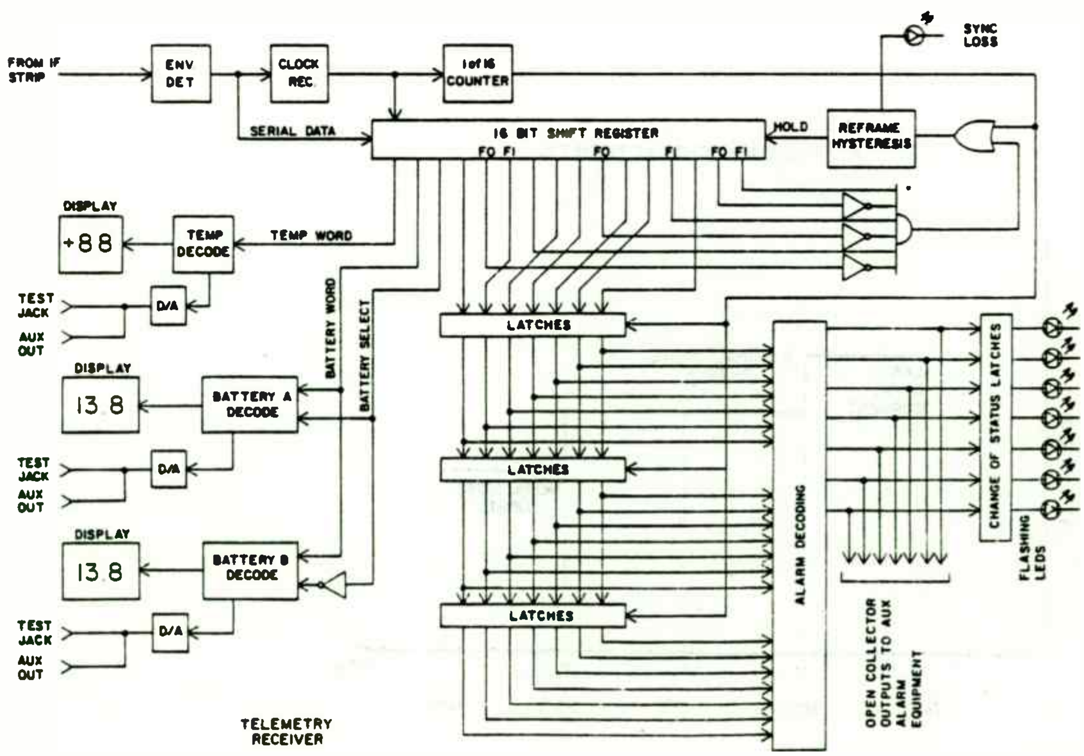


FIGURE 1



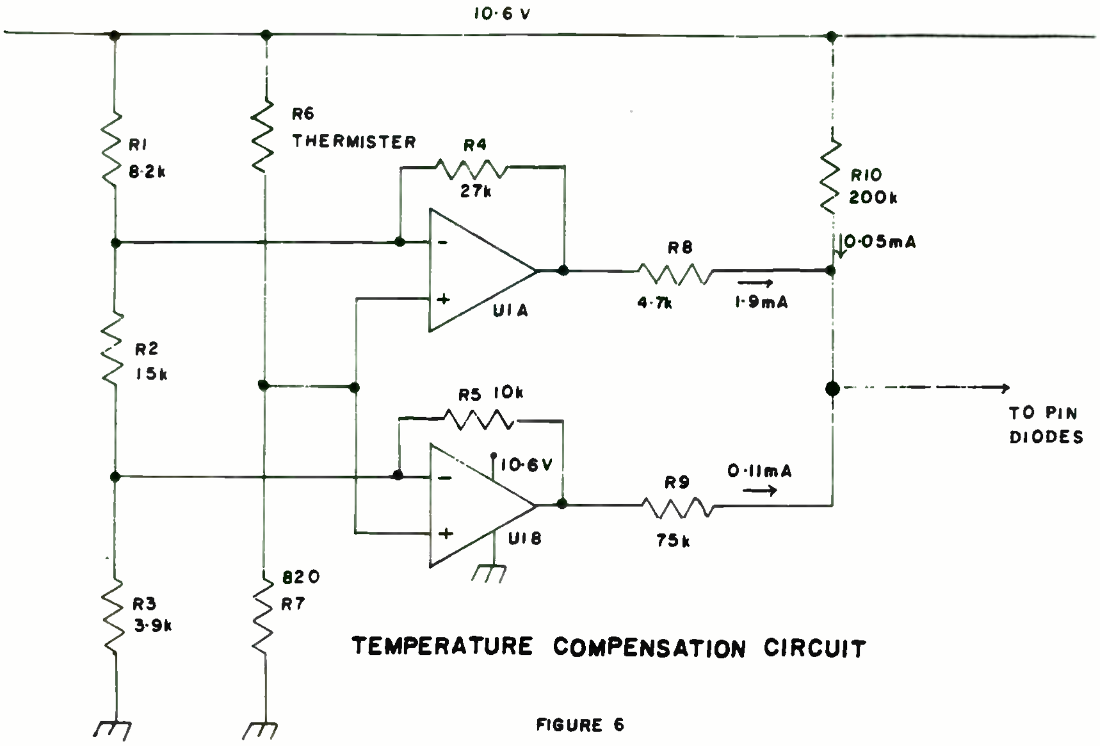
PIN DIODE ATTENUATOR

Figure 4



TELEMETRY RECEIVER

Figure 3



TEMPERATURE COMPENSATION CIRCUIT

FIGURE 6

ATTENUATION CHARACTERISTIC

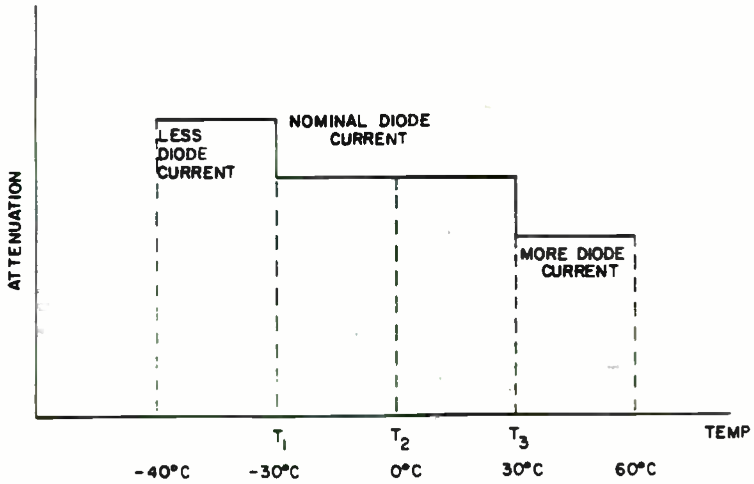


Figure 5

TEMP COMP EQUIVALENT CIRCUIT

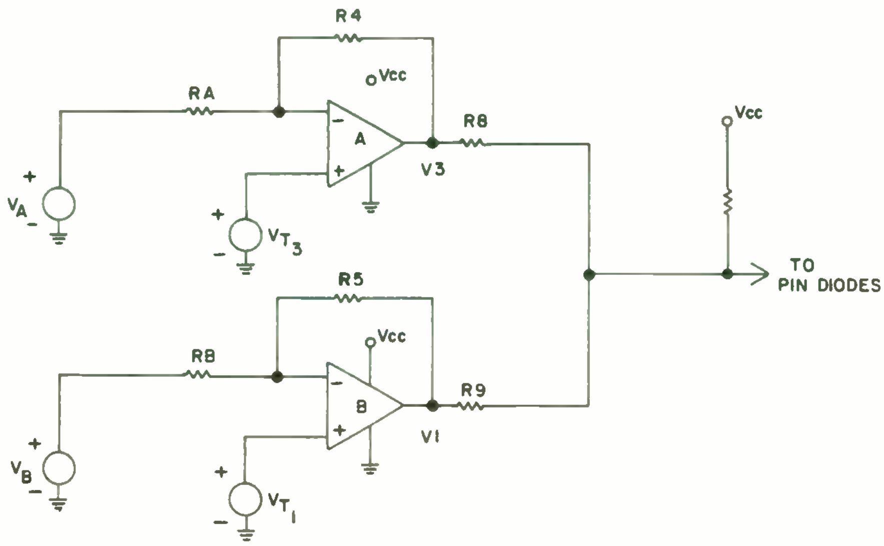


Figure 7

AN INTEGRATED KA-BAND POWER AMPLIFIER

Albert N. Pergande and Johnathan Ladd*

Martin Marietta Orlando Aerospace, Orlando, Florida, USA

Abstract

An integrated Ka-band pulse power amplifier consisting of a Gunn diode injection locked oscillator (ILO) with input and output ferrite isolators is described. The isolators and circulator used for injection locking of the Gunn diode oscillator are of a compact design, constructed on a single quartz substrate, and are biased by a common magnet. The pulsed Gunn diode ILO operates with pulsewidths of 40 nanoseconds to 12 microseconds, and provides a minimum of 8 dB gain over 500-MHz.

Introduction

Modern millimeter-wave radar guidance systems require miniature radar hardware capable of being produced in large quantities at extremely low costs. Much work on miniaturization and cost reduction of millimeter-wave receivers has been reported. However, transmitter miniaturization has received limited attention to date.

The amplifier described in this paper is a hybrid design employing a combination of resonant cavity and substrate designs. The goal was to develop a miniature Ka-band pulsed-power amplifier capable of operating in either a short- or long-pulse mode. The building blocks for this amplifier are a thin-film hybrid circulators with a common bias magnet and thin-film terminations, and a variable pulse-width Gunn diode oscillator. With these components, the amplifier takes up 1.8 cubic inches of space. The amplifier has demonstrated a maximum of 8 dB gain over a 500-MHz bandwidth at Ka-band when operated with a pulsewidth of 40 nanoseconds to 12 microseconds.

The amplifier module uses a combination of a suspended stripline and E-plane split waveguide. Short runs of waveguide are used as connections between components, and to interface with the oustide world (i.e., test equipment or antennas). This E-plane split block construction has several near-term advantages for hybrid systems. Each component can be developed to operate in its optimum transmission medium. For example, mixers and switches can be developed in finline, oscillators in waveguide cavities for low noise and high power, and ferrite devices can be created in stripline. Individual components can be tested separately prior to incorporation into the assembly. Finally, the system can have as much or as little integration as needed. As a result, the probability of having a working radar at any point in development is increased.

*Mr. Ladd is now with Sawtek, Inc., of Apopka, Florida.

The Ka-band amplifier takes advantage of the split block construction concept by using a compact suspended stripline circulator manifold and a medium power cavity oscillator. The loss of the suspended stripline at Ka-band is 0.3 dB per inch. Waveguide has a loss of between 0.03 and 0.06 dB per inch, while microstrip has a loss of 0.38 dB per inch[1]. These results show that the loss of the suspended stripline is slightly better than that of the microstrip, but much worse than that of the waveguide. The reduced size and shorter interconnect length of suspended stripline partially compensates for the increased loss.

A probe transition from the WR-28 waveguide to the suspended stripline was developed. A quarter-wavelength probe was inserted into the E-plane of the waveguide at a fixed distance from a waveguide short circuit. The return loss and insertion loss for a pair of transitions and 0.75 inch of line is shown in Figure 1. The line impedance is 70 ohms[2].

Suspended Stripline Circulator

The design of the stripline junction circulator is based on the work of Wu and Rosenbaum[3]. By using their approach, the electrical performance can be predicted from such physical parameters as ferrite puck diameter, coupling angle, magnetization factor ($4\pi Ms$), and bias field. Certain parameters, such as $4\pi Ms$ and bias field, are limited by available materials, while puck diameter and coupling angle are varied to achieve the desired frequency response.

A cross-section view of the circulator is shown in Figure 2. Two ferrite pucks are attached to either side of the quartz substrate. Input lines and ground plane for the suspended stripline and the metal resonator for the ferrite junction are printed on quartz. The coupling angle and resonator size are designed so that the output line width is almost identical to that of a 70-ohm line. A tapered transformer matches the impedance of the ferrite junction to that of the transition to waveguide. Figure 3 shows a disassembled view of a circulator substrate in a test housing. The transitions and one of the two ferrites stand out in the diagram.

The stripline circulator uses ferrites that are 70 mil in diameter by 30 mil thick. The $4\pi Ms$ is 5250 Gauss (Trans-Tech TT2-111), and the coupling angle is 84 degrees. The measured response of the circulator is shown in Figures 4 and 5. Insertion loss is between 0.6 and 1 dB, with a 15 dB isolation over a 6-GHz bandwidth.

Three circulator junctions and two thin-film terminations were integrated on a common substrate to form the circulator manifold. The junctions were spaced 0.2 inch apart. The junction proximity permits the use of a single magnet for bias. As a result, the size of the manifold can be reduced. Two of the circulator junctions serve as isolators in conjunction with the thin-film terminations, while the third junction acts as the duplexer for the amplifier. Three transitions to the waveguide allow interconnections to the reflection amplifier and to the outside world.

The thin-film termination consists of lossy transmission lines 2 wavelengths long. The titanium-tungsten metal layer provides a loss of 5 dB per λ, resulting in a total return loss of 20 dB for each load over the 26- to 40-GHz frequency range.

When the circulator was assembled in the module, there was a loss of 1.5 dB through two junctions. The isolation was better than 20 dB across the band (Figures 6 and 7). Return loss was 20 dB (Figure 8). Alignment of the six ferrites was aided by a simple assembly jig. The jig held the ferrites fixed with respect to the substrate. Wax was used for ferrite attachment in this prototype. However, cyano-acrylate adhesive is expected to be used in the future. The circulator was assembled and disassembled several times, showing that these actions are easily duplicated.

Gunn Diode Injection Locked Oscillator (ILO) Design

The Gunn diode (Microwave Associates Part 49837-138) is mounted on the diode pedestal, as shown in Figure 9. The pedestal is then press-fit into the coaxial cavity. The cavity couples to a standard height waveguide through a resonant iris. A dc pulse is applied via the center conductor, which also serves as a dumbbell radio frequency (RF) choke. The length of the inner conductor controls the center frequency of the pulsed RF signal, while the quality factor (Q) of the cavity can be adjusted by changing its diameter. In order to achieve a wide-locking bandwidth, a relatively low-Q cavity must be used. The locking bandwidth of the ILO is proportional to the free-running frequency chirp observed, with no input signal presented to the ILO. The cavity Q must be lowered in order to increase the locking bandwidth, which increases the frequency chirp[4].

In order to achieve the same amplifier characteristics for both long pulse and short pulse modes, the voltage pulse applied to the diode via the inner conductor is varied from approximately 7 volts for long pulse (up to 12 microseconds) to approximately 10 volts for short pulse (40 nanoseconds). No changes in tuning, coupling, operating temperature, or any other parameters are required for long pulse/short pulse operation.

The modulator used to drive the oscillator consisted of an integrated circuit clock driver (DS 0026) connected to the gate of a VMOS power FET (IRF 131) used in a source follower configuration. This combination exhibited extremely fast switching times (less than 5 nanoseconds) and good power handling capability.

Integrated Power Amplifier Performance

The power amplifier assembly was measured on the balanced phase bridge shown in Figure 10. This setup is useful for making intrapulse phase and amplitude measurements of an ILO. Typical waveforms of the ILO are shown in Figure 11. The top trace is detected RF, and the bottom trace is the phase of the oscillator during the pulse relative to the locking source. The phase characteristic is important in pulse compression systems, as well as in systems using power-combining techniques. Most of the change in phase relative to the locking source occurs during the first portion of the RF pulse. In the long pulse mode, there is a power drop related to the bias on the diode and the thermal environment of the oscillator. Normal power variation is approximately 1 dB for a 10 microsecond pulse at room temperature.

The overall performance of the power amplifier is described in Table 1. No change in the tuning of the ILO or of the stripline circulators was required to achieve this range of pulsewidth, duty cycle, and bandwidth. For an input power level of +20 dBm, peak output powers of +28 dBm for pulsewidths of 12 microseconds, and +31 dBm for pulsewidths of 40 nanoseconds were achieved. The locking bandwidth for an input signal of +20 dBm was 600 MHz for the range of duty cycle and pulsewidths stated in Table 1. The locking bandwidths for both long pulse and short pulse operation is 500 MHz.

Conclusion

The integrated power amplifier is a compact, relatively low-cost circuit, with performance comparable to similar circuits using only waveguide. The E-plane split block techniques discussed yield an optimum trade-off between cost, size, ease of testing, and performance. Further integration with other substrate-based components such as mixers, switches, couplers, and attenuators will result in a flexible system to meet millimeter-wave transmitter/receiver requirements.

We would like to acknowledge the invaluable assistance of Gene Allard, Tony Lazarski, and Jim Stonebraker in the assembling and testing of this device.

REFERENCES

1 Kajfez, D. and Tew, M., "Pocket Calculator Program for Analysis of Lossy Microstrip", Microwave Journal, pp 39-48, December 1980

2 Yamashita, E. and Atsuki, K., "Stripline with Rectangular Outer Conductor and Three Dielectric Layers", IEEE Transactions on Microwave Theory, Vol MTT-18 No. 5, pp 238-244, May 1970

3 Wu, Y.S. and Rosenbaum, F.J., "Wideband Operation of Microstrip Circulators", IEEE Transactions on Microwave Theory, Vol. MTT-22 No. 10, pp 849-856, October 1974

4 Kurokowa, K., "Injection Locking of Microwave Solid State Oscillators", Proceedings of IEEE, Vol 61, No. 10, pp 1386-1410, October 1973.

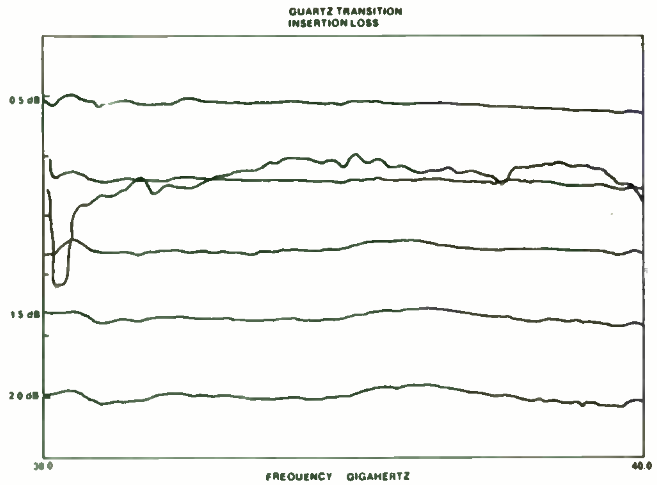


Figure 1a. Insertion Loss for Two Transitions and 0.75 inch of Stripline

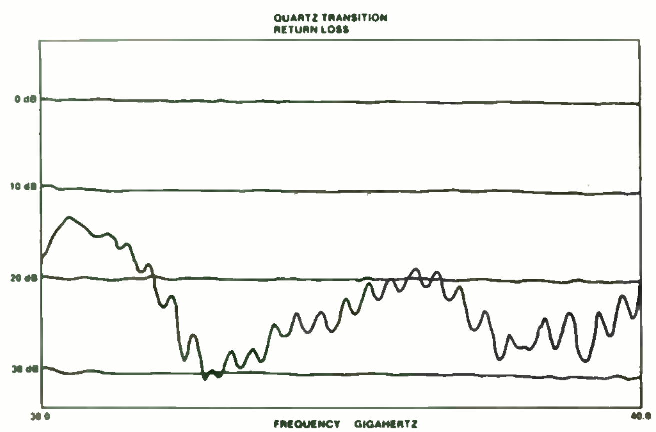


Figure 1b. Return Loss for Two Transitions and 0.75 inch of Stripline

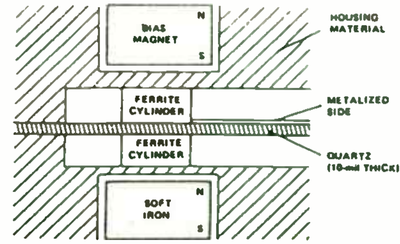


Figure 2. Circulator Junction Cross Section

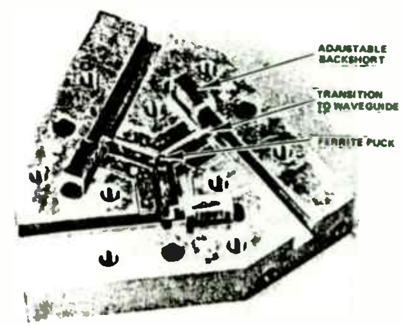


Figure 3. Single Stripline Circulator Substrate and Test Housing

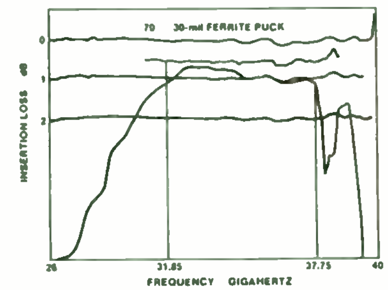


Figure 4. Insertion Loss versus Frequency of Single Stripline Circulator (Worst Case - Any Two Ports)

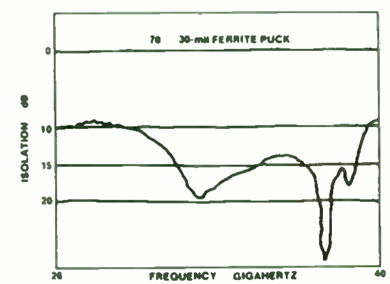


Figure 5. Isolation versus Frequency of Single Stripline Circulator (Worst Case - Any Two Ports)

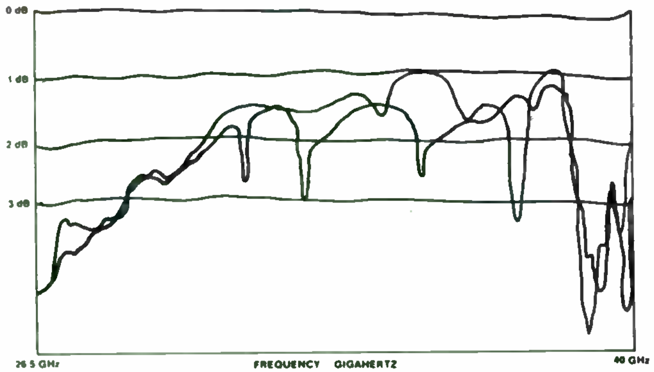


Figure 6. Insertion Loss from Input to Amplifier and from Amplifier to Output

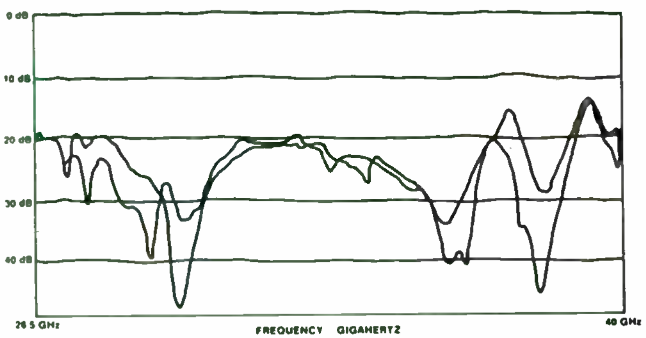


Figure 7. Isolation from Input to Amplifier and from Amplifier to Output

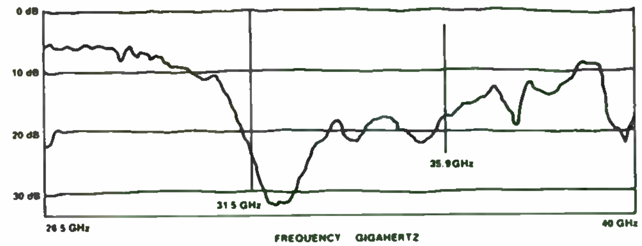


Figure 8. Return Loss at Amplifier Port

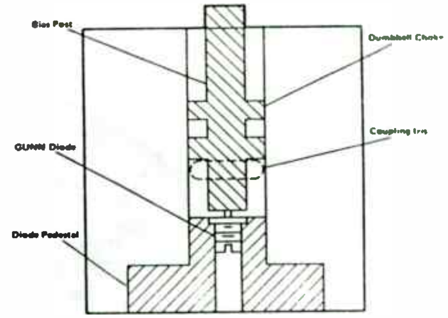


Figure 9. Gunn Diode Coaxial Cavity

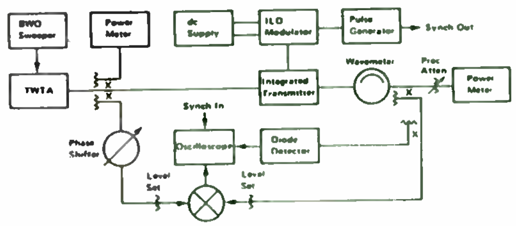


Figure 10. Balanced Bridge Test Set-up

TABLE 1

Ka Band Transmitter Performance

	12 microseconds Pulsewidth	40 nanoseconds Pulsewidth
Input Power	+20 dBm	+20 dBm
Output Power	+28 dBm	+31 dBm
Duty Cycle	33 percent maximum	5 percent maximum
Intrapulse Amplitude Variation	1 dB	1 dB
Amplitude Variation over Bandwidth	± 0.25 dB	± 0.25 dB
Intrapulse Phase Variation (from Injected Signal)	40 degrees	40 degrees
Injection Locked Bandwidth	600 megahertz	600 megahertz

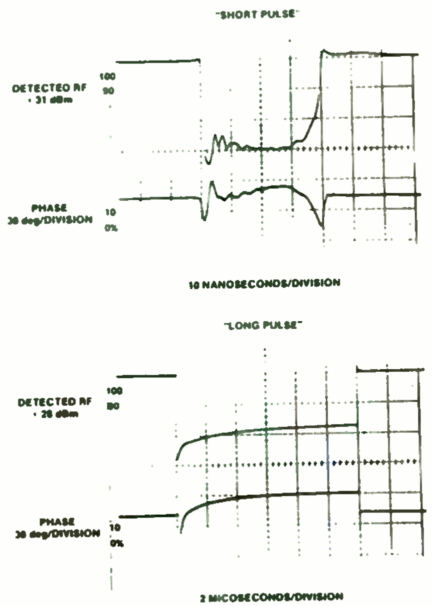


Figure 11. RF Waveforms

CLASS-D POWER-AMPLIFIER LOAD IMPEDANCE FOR MAXIMUM EFFICIENCY

by

Frederick H. Raab, Ph.D., Consultant
Green Mountain Radio Research Company
50 Vermont Avenue, Fort Ethan Allen
Winooski, Vermont 05404

Presented at

RF Technology Expo 85
Disneyland Hotel
Anaheim, California
January 23 - 25, 1985

ABSTRACT

The principal causes of inefficiency in a class-D power amplifier (PA) are (1) power dissipation due to output current flowing through the saturation voltage or on resistance of the transistors, and (2) power dissipation due to charging the transistor output capacitance. The desired power output can often be achieved by several different combinations of supply voltage and load resistance. A design trade-off exists because use of a higher supply voltage and load resistance reduces the saturation loss but increases the capacitor-charging loss. This paper derives formulas for power dissipation and the supply voltage and load resistance that produce the maximum efficiency for a given set of circuit parameters. Formulas are derived for both BJTs and FETs with either linear or voltage-variable output capacitance.

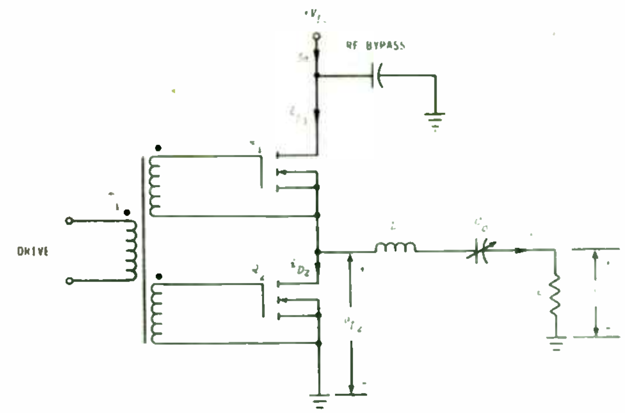
1. INTRODUCTION

The class-D RF power amplifier (PA) [1 - 3] employs a pair of transistors as switches to generate a squarewave voltage (Figure 1). The series-tuned output circuit allows only fundamental-frequency current to flow into the load. A class-D PA with ideal transistors (zero saturation voltage or resistance, instantaneous switching, and no shunt capacitance) achieves a dc-to-RF conversion efficiency of 100 percent. It is also capable of producing twenty-seven percent more power output than a class-B PA with transistors of the same ratings.

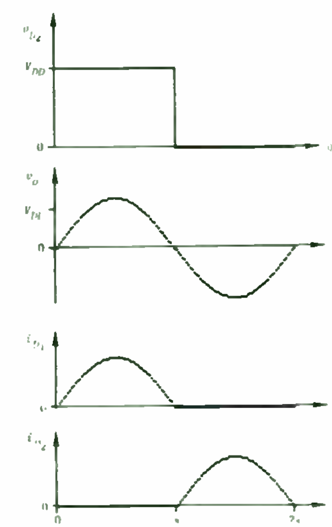
Recent developments in RF-power transistors (especially MOS FETs) have made practical high-frequency class-D amplifiers with power outputs in the 10- to 300-W range [4, 5]. However, practical class-D PAs do not achieve the ideal 100 percent efficiency, primarily because of saturation voltage or on resistance and shunt capacitance. (There are, of course, many additional sources of power loss, including the inductor in the tuned output circuit.) The saturation voltage of a BJT or on resistance of an FET is always effectively in series with the load resistance, thus reducing the fraction of the input power delivered to the load. The collector or drain output capacitance must be charged once per RF cycle, resulting in additional dc input power.

Many class-D power amplifiers use a matching network to convert the actual load impedance into a resistive load impedance suitable for the available transistors. The output power is a function of both the supply voltage and the effective load resistance. Consequently, the load resistance (or

GMRR TP80-5



(a) Circuit



(b) Waveforms

Figure 1. Complementary voltage-switching class-D PA.

supply voltage) may be regarded as a design parameter when a specified power output is desired. Increasing the load resistance decreases the power dissipated in the BJT saturation voltage or FET on resistance. However, the associated increase in supply voltage that is required to maintain the specified power output produces an increase in the power dissipated in charging the shunt capacitance.

This paper begins by presenting design equations for class-D PAs using both BJTs and FETs. The power dissipations associated with either fixed (linear) or voltage-variable output capacitance are then derived. The load resistances and supply voltages for maximum efficiency are then determined for both BJT and FET PAs with either fixed or voltage-variable output capacitances.

2. DESIGN EQUATIONS

Equivalent circuits for class-D power amplifiers using BJTs and FETs are shown in Figure 2. Saturation of a BJT produces a nearly constant collector-emitter voltage V_{sat}, while saturation of an FET produces a nearly constant drain-source resistance R_{on}. In both cases, the power P_{ds} dissipated due to device saturation is the result of the output current flowing through the saturated transistors.

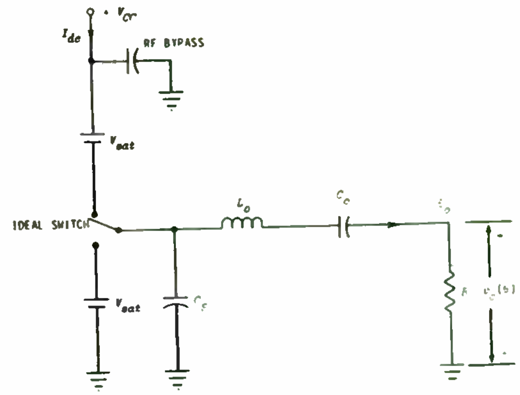
The output capacitance from both transistors (as well an any other stray capacitance) is assembled into a single equivalent shunt capacitor C_S. This capacitor is charged when Q_1 switches on, and discharged when Q_2 switches on. If the transistors are capable of charging and discharging C_S quickly, the switching and output waveforms are not affected by its presence (or value). However, additional dc input current is drawn to supply the required charge.

Since the shunt capacitance does not affect the output power, the supply voltage and load resistance can be related to the desired power output without regard to the shunt capacitance. The dissipation due to charging the shunt capacitance can then be parameterized as a function of either the supply voltage or load resistance, as is more convenient.

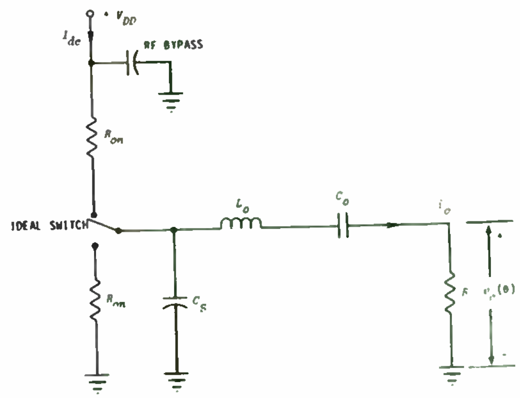
The amplitude of the fundamental-frequency component of the output is $2/\pi$ times the peak-to-peak voltage of the squarewave produced by the switching action of Q_1 and Q_2. The power output is therefore [3]

$$P_o = \frac{2}{\pi^2}\frac{V_{eff}^2}{R} , \quad (1)$$

where V_{eff} is the effective supply voltage. The effective voltage required for a given power output into a given load is determined by rearranging (1).



(a) With BJTs



(b) With FETs

Figure 2. Equivalent circuits.

Bipolar-Junction Transistors

The peak-to-peak voltage of the squarewave is reduced by saturation voltage V_{sat} at both the high and low ends. The effective supply voltage for a class-D PA with BJTs is therefore

$$V_{eff} = V_{CC} - 2\,V_{sat} \quad . \tag{2}$$

The true supply voltage is determined from the effective supply voltage and the saturation voltage by rearranging (2).

The dc input power is divided into dissipated and output power according to the ratio of the saturation and effective-supply voltages, thus the power dissipated due to saturation is

$$P_{dS} = \frac{2\,V_{sat}}{V_{eff}}\,P_o \quad . \tag{3}$$

Field-Effect Transistors

The equivalent series connection of load R and FET on resistance R_{on} acts as a voltage divider that reduces supply voltage V_{DD} to the effective supply voltage

$$V_{eff} = \frac{R}{R + R_{on}}\,V_{DD} \quad . \tag{4}$$

Given a specified power output and load resistance, V_{eff} is determined from (1) and V_{DD} is then determined from (3).

However, the determination of R given specified P_o and V_{DD} is somewhat more complicated. Insertion of (4) into (1) produces

$$(R^2 + 2\,R\,R_{on} + R_{on}^2)\,P_o = \frac{2R}{\pi^2}\,V_{DD}^2 \quad . \tag{5}$$

Rearrangement yields a quadratic equation in R, whose solution is

$$R = \frac{1}{2}\,[-b + (b^2 - 4R_{on}^2)^{1/2}] \quad , \tag{6}$$

GMRR TP80-5

where

$$b = 2\,R_{on} - \frac{2\,V_{DD}^2}{\pi^2\,P_o} \quad . \tag{7}$$

The dc input power is divided into dissipated and output power according to the ratio of the resistances, thus the power dissipation due to saturation is

$$P_{dS} = \frac{R_{on}}{R}\,P_o \quad . \tag{8}$$

3. LINEAR OUTPUT CAPACITANCE

The energy W_i required to charge a capacitor from 0 to v Volts is

$$W_i = \int v\,i(t)\,dt = v \int i(t)\,dt = v\,Q = 2\,W_s \quad . \tag{9}$$

Note that W_i is twice the stored energy W_s, regardless of the charging waveform or the dependence of capacitance upon voltage. Charging a capacitor through a resistance (e.g., the transistor) therefore requires twice the energy ($C_S^2 v^2/2$)that is stored in the capacitor.

In this section, C_S is assumed to be independent of voltage. The power expended in charging C_S to voltage V_{DD} once per RF cycle is therefore

$$P_{dC} = C_S\,V_{DD}^2\,f = \frac{V_{DD}^2}{2\pi X} \quad , \tag{10}$$

where $X = 1/2\pi f C_S$ is the reactance of the capacitor at the frequency f of operation.

The energy required to charge the capacitance is taken directly from the power supply, and therefore does not decrease the power output. The dc power input is therefore

$$P_i = P_o + P_{dS} + P_{dC} \quad . \tag{11}$$

Bipolar-Junction Transistors

In a class-D PA using BJTs, the voltage across C_S is increased by V_{eff}

each time Q_1 turns on. The associated power dissipation is [by analogy to (10)]

$$P_{dC} = \frac{V_{eff}^2}{2\pi X} \quad . \qquad (12)$$

The dc power input is found by insertion of (12) and (3) into (11). The derivative of the dc power input with respect to the effective supply voltage is therefore

$$\frac{\partial P_i}{\partial V_{eff}} = -\frac{2\, V_{sat}\, P_o}{V_{eff}^2} + \frac{V_{eff}}{\pi X} \quad . \qquad (13)$$

Setting this derivative equal to zero yields

$$V_{eff_0} = (2\pi X V_{sat} P_o)^{1/3} \qquad (14)$$

for minimum power dissipation, hence maximum efficiency. Load resistance R can now be found from (1), and the true supply voltage is obtained by adding $2V_{sat}$ to V_{eff}.

Curves of efficiency and power dissipation for a "typical" 25-W PA are shown in Figure 3. The transistors have V_{sat} = 1 V and together produce a shunt capacitance whose reactance is 100 Ω at the frequency of operation (e.g., C_S = 106 pF, f = 15 MHz). The benefits of the proper choice of R and V_{CC} are apparent.

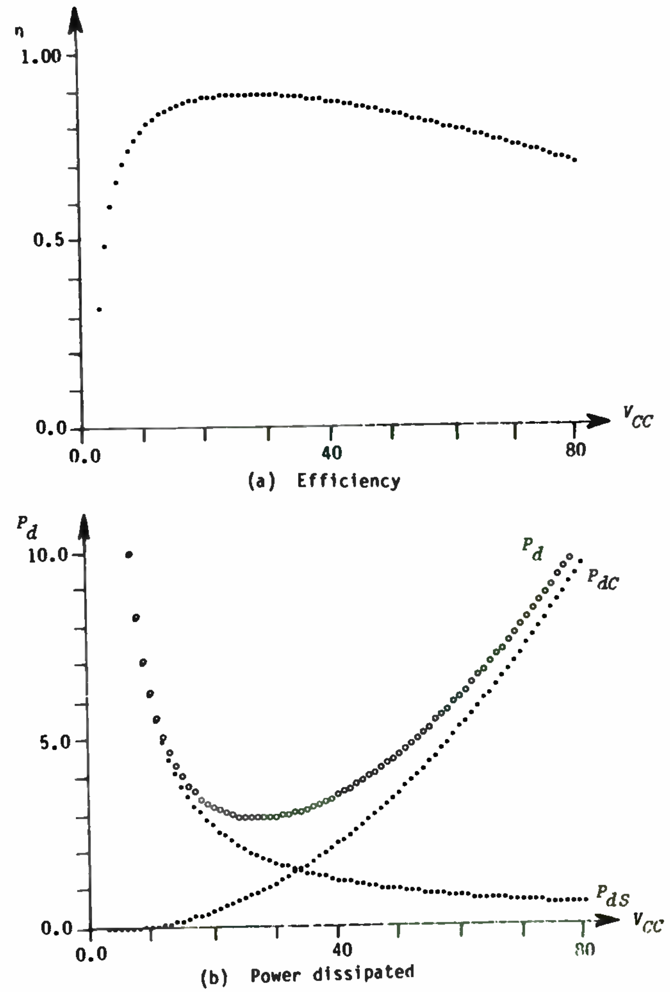
Field-Effect Transistors

The voltage across the shunt capacitance in an FET PA is increased by V_{DD} each time Q_1 turns on, hence the associated power dissipation is given by (10). However, P_{dC} must be converted to a function of R for compatibility with P_{dS} as given by (8). To produce the desired output, the supply voltage must be

$$V_{DD} = \frac{\pi}{2}\left(1 + \frac{R_{on}}{R}\right) V_{om} \quad . \qquad (15)$$

Since $P_o = V_{om}^2/2R$, substitution of (15) into (10) yields

GMRR TP80-5



Parameters: P_o = 25 W, V_{sat} = 1 V, X_0 = 100 Ω

Figure 3. Efficiency and dissipation of BJT PA with linear capacitance.

$$P_{dC} = \frac{\pi}{8X}\left(1 + \frac{R_{on}}{R}\right)^2 V_{om}^2 = \frac{\pi}{4}\left(1 + \frac{R_{on}}{R}\right)^2 \frac{R}{X} P_o \quad . \qquad (16)$$

Setting $\partial P_i / \partial V_{om} = 0$ then produces

$$R_0^2 = R_{on}^2 + \frac{4}{\pi} X R_{on} \quad . \qquad (17)$$

The other circuit parameters can now be calculated from the equations given previously.

Curves of efficiency and power dissipation for a "typical" 25-W PA are shown in Figure 4. The FET on resistance is 1 Ω and the total shunt capacitance is again assumed to produce a reactance of 100 Ω at the operating frequency.

4. VOLTAGE-VARIABLE OUTPUT CAPACITANCE

Most of the output capacitance of RF-power transistors (whether bipolar or field-effect) is abrupt-junction capacitance. The variation of this capacitance with voltage is given [6, 7] by

$$C(v) = C_0 \, (1 + v/\phi)^{-1/2}, \; v > 0 \qquad (18)$$

where C_0 is the zero-voltage capacitance and ϕ is the "barrier potential."

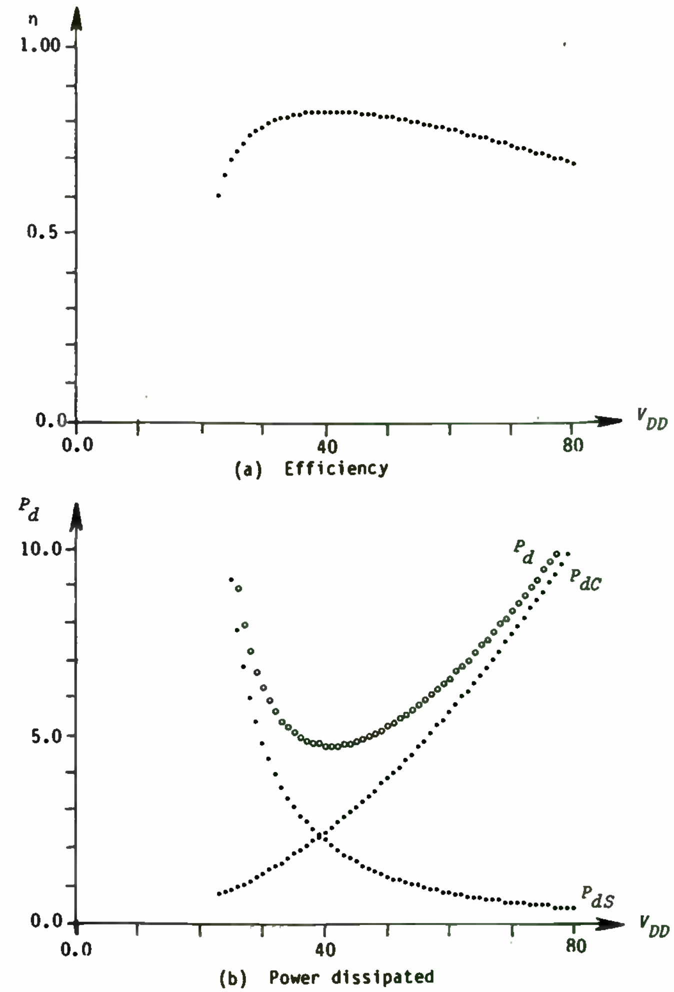
The value of ϕ is easily determined from two small-signal impedance measurements with different collector-emitter or drain-source bias voltages. For most RF-power transistors, $\phi \approx 1$ V. Figure 5a depicts a typical variation of capacitance with voltage.

Charge Stored

The ac capacitance given by (18) is defined by $C(v) = dQ(v) \,/\, dv$, where $Q(v)$ represents the total stored charge. The variation of total charge with voltage is therefore

$$Q(v) = \int_0^v C(u)\, du = 2\, C_0 \, \phi \, [(1 + v/\phi)^{1/2} - 1] \quad . \qquad (19)$$

Figure 5b shows the variation of total charge with voltage for both linear and abrupt-junction capacitors. It is apparent that if the zero-voltage capacitance of the abrupt-junction capacitor is equal to the fixed capacitance, considerably less charge is stored in the abrupt-junction capacitance



Parameters: P_o = 25 W, R_{on} = 1 Ω, X_0 = 100 Ω.

Figure 4. Efficiency and dissipation of FET PA with linear output capacitance.

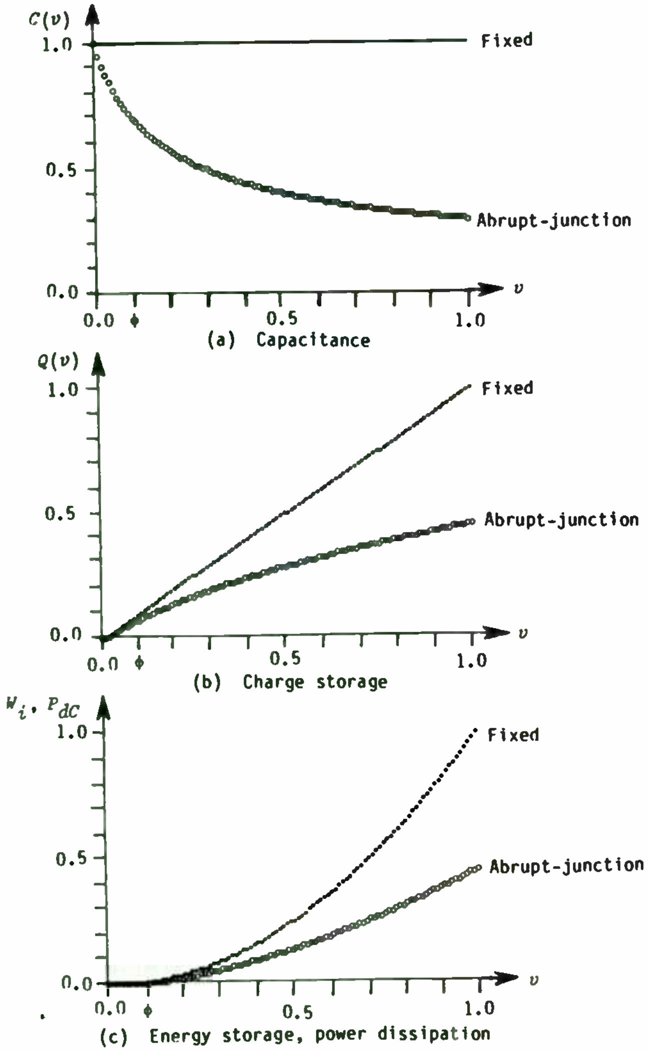


Figure 5. Normalized characteristics of capacitors.

as the voltage increases.

The energy required to charge C_S to voltage v is obtained by insertion of (19) into (9). The power dissipated in charging C_S f times per second is by analogy to (10)

$$P_{dC} = W_i \, f = 2 \, f \, C_0 \, v \, [(1 + v/\phi)^{1/2} - 1] \tag{20}$$

$$= \frac{v \, [(1 + v/\phi)^{1/2} - 1]}{\pi X_0} \, , \tag{21}$$

where X_0 is the impedence of C_0 at frequency f.

The relative power dissipated in linear and abrupt-junction capacitors is shown in Figure 5c. It is apparent that significantly more power is dissipated in a linear capacitor than in an abrupt-junction capacitor of equal capacity at zero voltage.

Bipolar-Junction Transistors

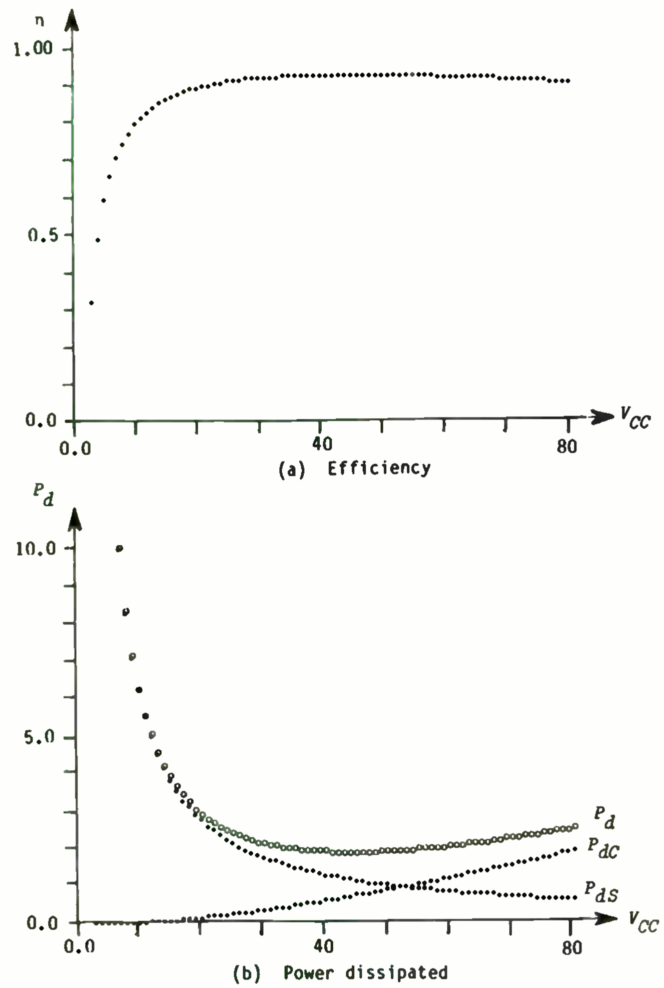
The saturation loss in a complementary class-D PA using BJTs is given by (3), and the effective supply voltage is given by (2). Since the output capacitance must be charged from V_{sat} to $V_{CC} - V_{sat}$, the associated power dissipation is [from (20)]

$$P_{dC} = \frac{V_{eff} \, [(1 + V_{eff}/\phi)^{1/2} - 1]}{\pi X_0} \, . \tag{22}$$

The relationships among the supply voltage, efficiency, and power dissipation of a typical 25-W BJT power amplifier are shown in Figure 6. A maximum efficiency of 92.8-percent is obtained with $V_{CC} = 44.3$ V and $R = 14.5\ \Omega$.

The total power dissipated is the sum of the powers dissipated due to saturation and output capacitance. The maximum-efficiency point can be found by setting the derivative of P_d with respect to V_{eff} equal to zero, thus

$$0 = -\frac{2 V_{sat} \, P_o}{V_{eff}^2} + \frac{1}{\pi X_0} \left\{ [(1 + \frac{V_{eff}}{\phi})^{1/2} - 1] + \frac{V_{eff}}{2\phi} \, (1 + \frac{V_{eff}}{\phi})^{-1/2} \right\} \, . \tag{23}$$



Parameters: P_o = 25 W, V_{sat} = 1 V, X_0 = 100 Ω, ϕ = 1.0 V

Figure 6. Efficiency and dissipation of BJT PA with voltage-variable capacitance.

Analytical solution of (23) is difficult or impossible. However, in most applications $V_{eff} >> \phi$, hence $1 + V_{eff}/\phi \cong V_{eff}/\phi$ and $(V_{eff}/\phi)^{1/2} - 1 \cong (V_{eff}/\phi)^{1/2}$. Use of these approximations in (23) yields

$$0 = -\frac{2V_{sat} P_o}{V_{eff}^2} + \frac{3V_{eff}^{1/2}}{2\pi X_0 \phi^{1/2}} \quad , \qquad (24)$$

and rearrangement produces

$$V_{eff} \cong \left(\frac{4\pi}{3} X_0 \phi^{1/2} V_{sat} P_o\right)^{2/5} \quad . \qquad (25)$$

The value of V_{eff} obtained from (25) is converted to supply voltage V_{CC} by (2) and to load resistance R by (1). The accuracy of (25) increases as V_{CC} increases. For the example of Figure 6, (25) predicts a 92.8-percent maximum efficiency at V_{CC} = 42.5 V and R = 13.3 Ω, which are fairly close to the values (V_{CC} = 44.3 V, R = 14.5 Ω) obtained by numerical evaluation of P_d via (3) and (22).

Field-effect Transistors

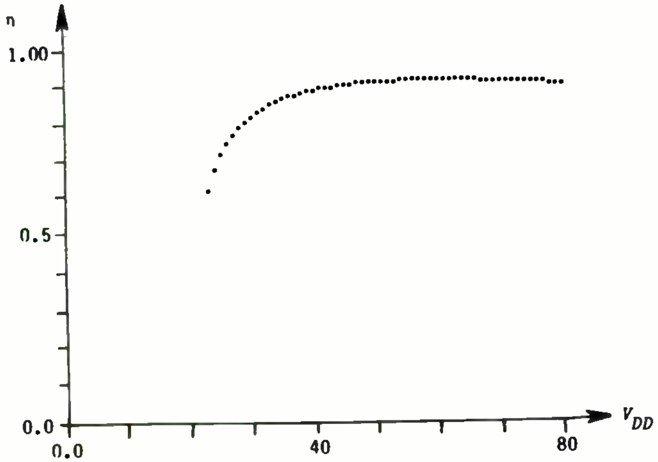
The saturation loss in a class-D PA using FETs is given by (8). Since the output capacitance is charged from 0 to V_{DD}, (20) gives

$$P_{dC} = \frac{V_{DD}\left[(1 + V_{DD}/\phi)^{1/2} - 1\right]}{\pi X_0} \quad . \qquad (26)$$

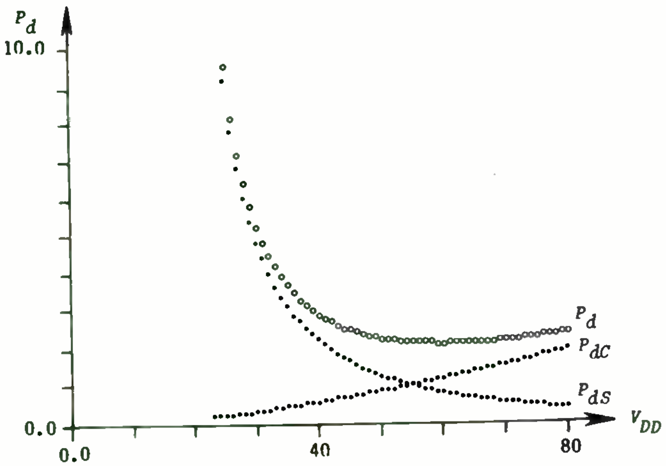
for the related power dissipation.

Curves showing the variation of efficiency and dissipation with voltage for a typical 25-W PA are shown in Figure 7. A maximum efficiency of 91.8 percent is obtained with V_{DD} = 59.6 V and R = 26.8 Ω.

To determine the maximum-efficiency operating point, it is necessary to convert (8) into a function of V_{DD}. Use of the exact relationship given by the quadratic formula (6) leads to unmanageable equations. It is therefore necessary to rearrange (1) and to assume $V_{DD} \cong V_{eff}$ to obtain



(a) Efficiency



(b) Power dissipated

Parameters: P_o = 25 W, R_{on} = 1 Ω, X_0 = 100 Ω, ϕ = 1 V

Figure 7. Efficiency and dissipation of FET PA with voltage-variable capacitance.

GMRR TP80-5

$$\frac{1}{R} \cong \frac{\pi^2 P_o}{2V_{DD}^2} \quad , \tag{27}$$

hence

$$P_{dS} \cong \frac{\pi^2 R_{on} P_o^{\,2}}{2V_{DD}^2} \quad . \tag{28}$$

The minimum power dissipation therefore occurs when

$$0 \cong - \frac{\pi^2 R_{on} P_o^{\,2}}{V_{DD}^3} + \frac{3V_{DD}^{1/2}}{2\pi\, X_0\, \phi^{1/2}} \quad , \tag{29}$$

hence

$$V_{DD} \cong \left(\frac{2\pi^3}{3} X_0\, \phi^{1/2} R_{on} P_o^{\,2}\right)^{2/7} \quad . \tag{30}$$

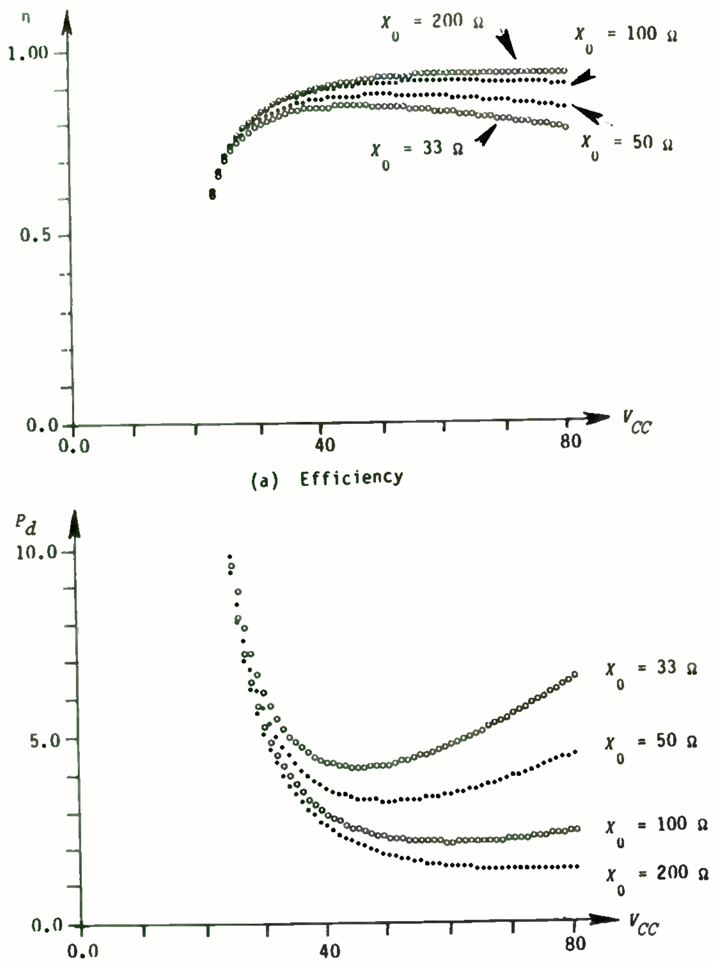
Once V_{DD} has been calculated from (30), R can be calculated from (6). For the example of Figure 7, (30) yields a maximum efficiency of 91.8; V_{DD} = 55.7 V and R = 23.1 Ω, which are in fair agreement with the values obtained by numerical evaluation of (26) and (28).

Effect of Frequency

The power dissipated in charging the shunt capacitance varies directly with the frequency of operation. In contrast, the saturation losses remain relatively constant. Consequently, the maximum-efficiency operating point shifts to lower supply voltages and lower load resistances as frequency increases. Examples of efficiency and total power dissipation are shown in Figure 8.

5. COMMENTS AND CONCLUSIONS

The efficiency of class-D power amplifiers is reduced from the ideal of 100 percent primarily by power dissipation due to saturation and charging of the output capacitance of the transistors. When load resistance and supply voltage can be varied to achieve a desired output power, a design trade-off



Parameters: P_o = 25 W, ϕ = 1 V

Figure 8. Effect of frequency.

is possible between the two principal power dissipations. Formulas have been derived for the load resistance and supply voltage that produce maximum efficiency in class-D PAs with both BJTs and FETs and either linear or voltage-variable output capacitance.

While the output capacitance of real class-D PAs is primarily abrupt-junction capacitance, some linear capacitance and other types of voltage-variable capacitance are also present. Consequently, the maximum-efficiency point for a real PA is generally between the maximum-efficiency conditions predicted for fixed and abrupt-junction capacitances.

In theory, it is possible to use the load network to charge and to discharge the shunt capacitance [8], thus eliminating all losses associated with the shunt capacitance, as in the ideal class-E amplifier [3]. However, it is in practice rather difficult to achieve the precise three-state switching and load impedance necessary for ideal "class-DE" operation. Nonetheless, a small improvement in efficiency can often be obtained by reducing drive slightly and tuning the output network for a slightly inductive net reactance.

6. REFERENCES

1. P.J. Baxandall, "Transistor sinewave oscillators: Some general considerations and new developments," *Proc. IEE*, vol. 106, part B, suppl. 16, pp. 748 - 758, 1959.

2. F. H. Raab, "High efficiency amplifications techniques," *IEEE Circuits and Systems Newsletter*, vol. 7, no. 10, pp. 3 - 11, December 1975.

3. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, New York: Wiley, 1980.

4. F. H. Raab, "Get broadband dual-mode operation with this FET power amplifier," *EDN*, vol. 23, no. 19, pp. 117 - 124, October 20, 1978.

5. H. O. Granberg, "Power MOS FETs versus bipolar transistors," *R. F. Design*, vol. 4, no. 6, pp. 11 - 15, November/December 1981.

6. J. Costa-Freire, and P. Teixeira, "Modeling of epidrain effects in VMOS power transistors for CAD," *IEE Design Automation Conference Abstract (EDA)*, Brighton, UK, 1981.

7. S. M. Sze, *Physics of Semiconductor Devices*, New York: Wiley and Sons, 1969.

8. S. A. Zhukov and V. B. Kozyrev, "Double-ended switching generator without commutating (switching) loss," *Poluprovodnikovye Pribory v Tekhnike Elektrosvyazi*, vol.15, Moscow: Svyaz, 1975.

CAD METHODOLOGY FOR MICROWAVE OSCILLATORS

Arturo Velazquez*, Jose Luis Medina*, Arturo Serrano**
(*CICESE Research Center, Ensenada, Baja California, Mexico
**University of Baja California in Ensenada)

ABSTRACT

This paper describes a computer program implemented to obtain in a simple but effective way, microwave oscillators using bipolar or GaAs FET devices. The program uses a small signal approach to give a first order approximation of the selected device's capabilities to generate negative resistance for oscillation. Using this computer program several oscillator prototypes have been constructed and have given adequate characteristics for local oscillator applications in down conversion stages for satellite communications earth station systems.

INTRODUCTION

In our computer program implemented to design microwave oscillators, the small signal S parameters are used to know the oscillation capabilities for a given device in different circuit configurations. The main point in the stability analysis that is made with this purpose, consists in the evaluation of the instability areas both at the input and output ports of the active device in order to determine if it is convenient to add some kind of feedback element and try of maximize those areas with the variation of the value for that feedback element.

When the unstable area is maximized, a wider reflection coefficient area for oscillation is assured; this implies in turn a larger number of combinations of impedance values for resonance for both the output and input ports. (Oscillation conditions and the associated circuit are shown in fig. 1). Once the maximum unstable area is defined, the feedback element value is set and then the procedure for negative resistance generation is initiated by means of varying the reflection coefficient of the resonator port. With this reflection coefficient an output negative resistance is obtained to deliver maximum output power.

In the following section of this paper we describe the main functions of the subroutine programs of the methodology.

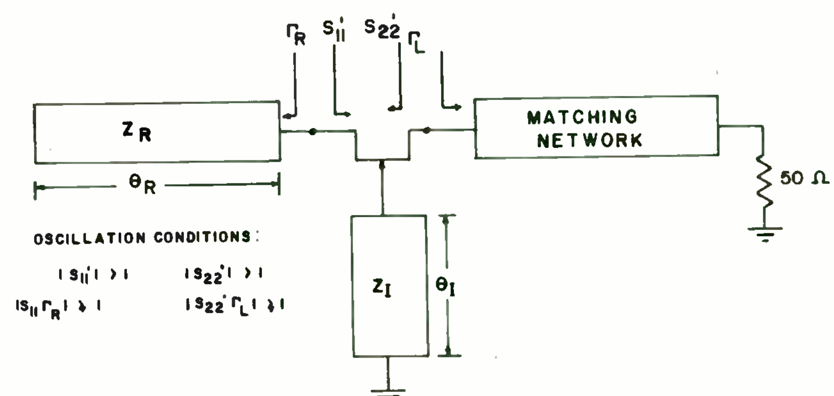


Fig 1) Configuration of transistor oscillator

MICROWAVE OSCILLATOR PROGRAM

This program uses a small signal approach to provide a first order approximation of the negative resistance generation capability of a given device to produce oscillations. The program is written in FORTRAN IV and has an interactive structure to assist the user in the design process. Fig. 2 shows a block diagram of the main program and associated subroutines which main functions are:

a) Stability analysis of active device;
b) Maximize unstable areas;
c) Change circuit configuration;
d) Adhere feedback element;
e) Input and output matching network design;
f) Microstrip circuit dimensions definition;

The program's input data are: S parameters of the active device, frequency of operation and required bandwidth, limits of the feedback element values. The corresponding output data are: Impedance and electrical length of the input and output matching circuit elements, resonator and feedback element dimensions.

SUBROUTINE "ESTAB"

Subroutine "ESTAB" is described in more detail because it is the most important part of the present design methodology. This subroutine is called each time stability analysis of the active device is required. Stability analysis is realized from a given circuit configuration that can be changed in order to obtain information about the instability characteristics of the device. Rollet's factor K and stability circles equations are used in our analysis.

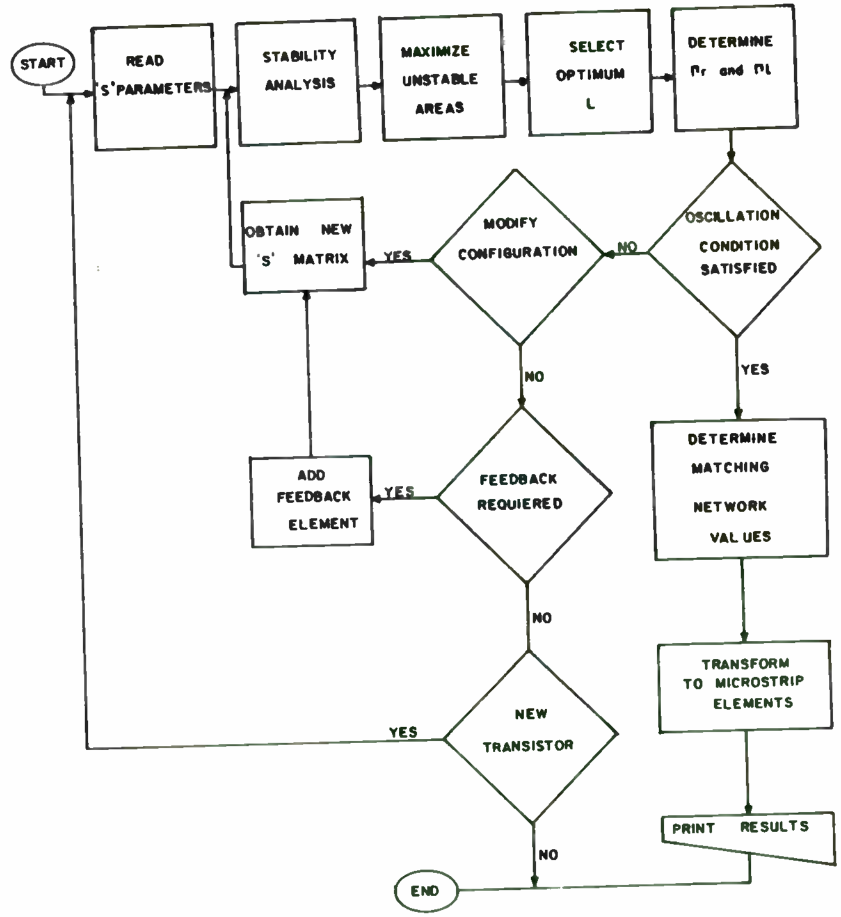


Figure 2).- Flow diagram of the CAD methodology .

For each circuit configurations is possible to include a feedback element and analyze the stability characteristics of the device varying the feedback element value and storing that one that gives the maximum unstable area.

Unstable areas are calculated taking into account the different cases that could arise in the stability circles and on the Smith Chart. Figs. 3A) and 3B) show an example of unstable area calculation for two typical cases.

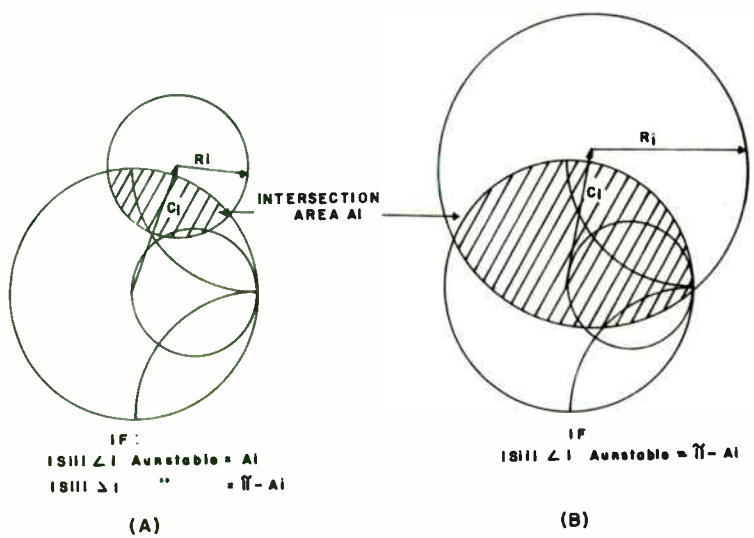


Fig 3) Two typical cases of unstability areas

The search for maximum unstable area has the objective of having a larger number of reflection coefficient for oscillation.

To maximize and unstable area, we use a routine called MINMAX that provides maximum and minimum values for an unidimensional array. In this case, three arrays are formed: Output unstable area, input unstable area and K factor. All these three for each feedback element value. Up to this point, the program user might choose among the three array elements the one that is more important for him and from here define the feedback element.

SUPPORT SUBROUTINES

a) Subroutine CONFIG. This subroutine provides circuit configuration change. From the initial common emitter configuration other configurations can be obtained from the unstable area point of view. The associated S parameters for each configuration are also provided in this part of the program.

b) Subroutine RETRO. This subroutine includes a series inductive or shunt capacitive feedback element in the circuit. This process is realized with appropiate matrix operations of the elements involved.

c) Subroutine ACOPL. Matching network elements are calculated in this subroutine using Przedpelski's [1] equations. The output impedance seen by the active device is calculated using Maeda's formula [2]:

$$Im(Z_L) = -Im(Z_{out}) \qquad \text{-----------------} \qquad 1)$$

$$Re(Z_L) = 1/3 \left| Re(Z_{out}) \right| \qquad \text{-----------------} \qquad 2)$$

where Z_{out} is the resultant impedance when a resonator reflection coefficient is set.

d) Subroutine MICROS. In this subroutine, the physical dimensions of the circuit elements defined in the subroutine ACOPL are obtained. Input data for this subroutine are dielectric constant of the substrate, dielectric and metal thickness and frequency of operation. Hammerstaed's [3] formulas are used to evaluate the physical dimensions of the circuit elements.

e) Subroutine AREA. This subroutine has the specific function to calculate the intersection area between the Smith Chart and the stability circle defining in this way the unstable area concept described in figs 3a and 3b.

RESULTS AND CONCLUSIONS

The program described in this paper has been very helpful in the design and manufacturing of several oscillator prototypes. Fig. 4 shows the picture of one of the designed prototypes using a bipolar transistor. Fig. 5 shows the associated frequency response with 10dbm output power for 3.95 GHz central frequency. We can conclude that the program provides useful results for the manufacturing of microwave oscillator circuits.

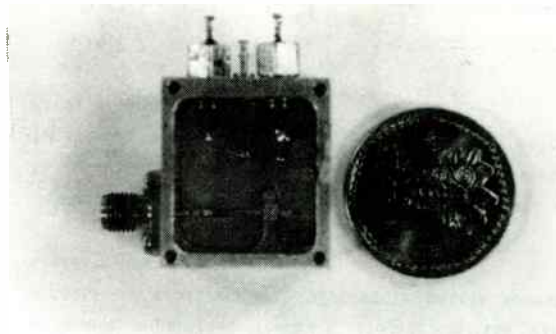


Fig 4) Photograph of an oscillator prototype

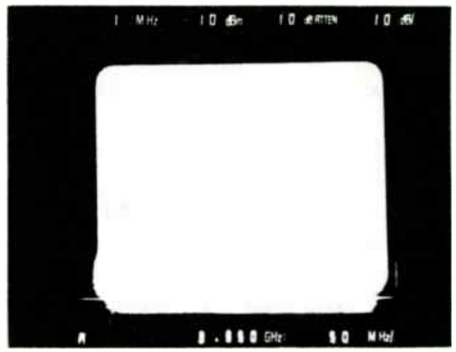


Fig 5) Measurement of output power and frequency response for the oscillator of fig. 4.

REFERENCES

1. A. Przedpelski; "Simple Transmission Line Matching Circuits"; RF Oesign; Summer 1980. pp 27-32.

2. M. Maeda, K. Kimura and H. Kodera; "Oesign and Perfomance of X-band Oscillators with GaAs Schottky-gate Field-effect Transistors"; IEEE Trans. on MTT. Vol MTT-23, No. 8. pp 661-667, Aug. 1975.

3. E. Hammestaed; "Equations for Microstrip Circuit Oesign"; Proc. of the V European Microwave Conference. Hamburg, Germany. pp 268-272. 1975.

RF Design on a Small Computer

James Eagleson
IDX, Inc.

January 1985

This paper is not designed for the experienced computer-aided-design whiz. I'm not one. Many RF engineers fall into the same "not a computer whiz" category.

What this paper attempts to do is to allow those who currently use calculators for simple design tasks to use their home computer for similar tasks when the added flexability of the computer is desireable.

The paper is also not going to deal with the more sophisticated programs and computers such as Apple, IBM, etc., but is keyed more towards the Atari, Commodore, TRS, Timex end of the spectrum. Of course, since BASIC is basic, much of what is covered here can be applied to the more sophisticated machines.

COMPUTER VERSUS CALCULATOR

A Calculator can be an indispensable tool for the working engineer.

A Computer, on the other hand, is not generally considered "indispensable" by many engineers.

First, those who need a computer in their work are generally provided a computer by their company. It is also rare in most "high tech" companies not to have at least an Apple or IBM level computer available for which fairly sophisticated CAD programs are available (though access to the computer may be limited to times when accounting or inventory control isn't using the machine!)

Second, if a $5000 computer is not justifiable, programmable scientific calculators are readily available in the $50-250 price range complete with printer and memory expansion. These are often adequate for most needs of an R. F. Engineer.

Then again, a fairly sophisticated computer can be purchased for as little as $100-200 today...$600-800 with disc drives, modem, and a printer, making the use of these "home" computers very attractive. Used systems, particularly such notables as Osborne and other "fall out" systems can often be picked up complete with sophisicated software for $600-800.

Thus one must consider that a small computer may not be all that "dispensable", after all!

CALCULATOR USES

Calculators have the following strengths:

1) **Provide quick checks on New Ideas**

2) **Provide quick conversions**

3) **Provide a resource for simple iterration**

4) **Provide straight forward programming**

5) **Compact**

6) **Portable**

7) **Large selection of published program material**

COMPUTER USES

A given Computer, on the other hand, may not be as portable, compact, or have as large a selection of specialized program material.

A computer is generally much faster, **more flexable, and can do things most calculators** cannot do, on the other hand.

Computers have the following strengths:

1) **Fast execution time.** The slowest computer is several times faster in execution than most hand calculators.

2) **More Memory.** Even the smallest current home computer has **16K** of memory available. While some calculators can expand to this kind of power (and **RPN** machines are very

memory efficient, to boot), most computers can store a group of related programs in one **Integrated Program** which can be **Menu** driven (that is, a given program can be selected from a list of all programs in the set. Results can be transferred from one program to another, as well.)

3) It is frequently more convenient to do multi-level iterrations on a computer than with the calculator. For example, data can be retrieved from a data base without physically attending the system (no magnetic cards, etc.)

4) Most computers have much more Graphics and Formatting capability than the most sophisticated of the hand calculators. Some can provide high resolution graphs, charts, circuit diagrams, response curves, even pictures beyond any of the calculators.

5) Computers can be used for Word Processing and Data Base Management. Most of the $150, 64K machines (C-64, Atari 800XL) have programs available in these areas that rival programs available for the more expensive PC's! Even a Timex 1000, with appropriate hardware and software, can provide reasonable performance in this area! This is certainly not available with calculators.

6) Any computer (even the Timex) can be used for communications with other systems. This can be to another, bigger computer via RS-232 ports, to data bases via telephone Modems, to A-to-D conversion interfaces for automatic testing, data gathering, or control, and other input/output configurations.

THE PROBLEM WITH COMPUTERS

The problem with computers, of course, is that you've got to program the things!

While so-called "reverse-Polish" is hard enough to grasp so that some hand calculators have continued to provide "arithmetic" programming (generally conceded to be slower and to require many more steps), BASIC and all other computer languages tend to be as bad as, if not worse than reverse-Polish!

Fortunately, many users of the newer Hewlett-Packard calculators will find the transition somewhat easier since HP's adoption of BASIC-like structure in many of their newer machines. Users of arithmetic machines will find BASIC to be quite similar in use of parenthesis, etc.

Unfortunately, there are differences, such as BASIC's insistence on Radians for Trig Functions, which remain to be dealt with as well as machine-sensitive commands relative to print functions, labels, stacks, input of data, etc.

PROGRAM DEVELOPMENT

As with a calculator, programming a computer requires some thought as to exactly what data you wish to input and what kind of answer(s) you wish to have come out.

A simple program for determining Mixing Spurious Outputs from a hetrodyne transmitter's mixer demonstrates several programming techniques.

FORMULAS

The essential formula for Mixing Products is:

F1 + F2 and **/F1 - F2/**

Where **F1 = Frequency #1** and F2 = **Frequency #2**

For convenience, we have taken the **Absolute Value of F1 - F2** so that if F2 is the larger number it will come out as "positive frequency", not negative. (This is not necessary with F1 + F2, of course, since you'll never go negative adding two positive numbers together!)

In **BASIC**, however, Absolute Value is generally taken as follows:

ABS(F1 - F2)

(Note that the F1 - F2 is bracketed so that the ABS function works on the result of F1 - F2 rather than on F1, alone. Also note that for some functions a given computer will use a different abbreviation than another computer...example, the Commodore 64 uses LOG whereas the Timex uses LN even though both stand for Natural Log.)

The above formula gives us the fundamental responses at the desired output frequency and the image frequency but these are not usually the ones we are concerning about. It is the higher order mixing products which concern us.

Thus a little modification of the formula:

nF1 + mF2 and /nF1 - mF2/
or
in BASIC, ABS(nF1 - mF2)

The n and m represent the harmonic of F1 and F2 for a given calculation. They will vary from 1 to the highest value you wish to set.

We could just set an arbitrary value to n and m, of course, but it would provide more flexability if we can set a **limit** on the number of harmonics we wish to use.

Since it is likely that we won't be interested in harmonic numbers for F2 higher than those for F1, however, we can probably set the same limit for both. We will call this limit H (for Harmonic number).

VARIABLE NAMES

At this point it would be wise to say a few words about the **name** of **variables** in computer programming.

This is an area that varies from computer to computer even when using BASIC.

Some computers would let you use **ANTI-ESTABLISHMENTARIANISM** if you wished to do so. Others require the use of only a few characters and/or numbers as names of variables. Anything else may show a SYNTAX error or be automatically shortened (truncated) by the computer.

It is probably a good practice to use one to three letters with appropriate numbers to name variables.

In that this is not unlike calculator programming, it should not be hard to get used to.

SETTING LIMITS

We may wish to look at all possible mixing products generated by our two frequencies in the sample program, but it is most likely that we are only interested in a limited range of possible output frequencies as determined by the Band Pass Filter being used on the output of the mixer.

Therefore we will want to set and upper and lower frequency limit beyond which we don't care if there is a mixing product since the BPF will eliminate it.

We'll call these limits **FL** for **FREQUENCY LOWER** limit and **FU** for **FREQUENCY UPPER** limit.

INPUTTING DATA

Here is where the computer has an additonal advantage over the calculator...it can tell you what input it requires **in plain English** (that is, it is **self-prompting**).

A typical computer input programming sequence would be as follows:

```
      1 REM   ENTER DATA

HD    5 PRINT TAB 11; "SPURIOUS RESPONSES"

     10 PRINT "<ENTER> FREQUENCY #1= ";
F1   20 INPUT F1
     30 PRINT F1;" MHZ"

     40 PRINT "<ENTER> FREQUENCY #2= ";
F2   50 INPUT F2
     60 PRINT F2;" MHZ"

SPC  70 PRINT

     80 PRINT "<ENTER> HARMONIC LIMIT= ";
H    90 INPUT H
    100 PRINT H

    110 PRINT "<ENTER> UPPER FREQUENCY LIMIT= ";
FU  120 INPUT FU
    130 PRINT FU;" MHZ"

    140 PRINT "<ENTER> LOWER FREQUENCY LIMIT= ";
FL  150 INPUT FL
    160 PRINT FL;" MHZ"
```

This will ask you line by line for Frequency #1, **Frequency #2, Harmonic, Upper Frequency Limit, and** Lower Frequency Limit. On most machines this will start in the upper left hand corner of the screen (sometimes refered to as "home") and go down one line for each "PRINT" statement (line 70, for example, exists only to produce a blank line between the F1 and F2 data and the Harmonic, FU, and FL data).

Each line will prompt an input, say, "<ENTER>Frequency #1= " then print the value next to it immediately upon your entering the data. It will then print the next prompt, "<ENTER>Frequency #2= ".

SPURIOUS RESPONSES

<ENTER> FREQUENCY #1= 155.000 MHZ
<ENTER> FREQUENCY #2= 10.700 MHZ

<ENTER> HARMONIC LIMIT= 5
<ENTER> UPPER FREQUENCY LIMIT= 175.000 MHZ
<ENTER> LOWER FREQUENCY LIMIT= 135.000 MHZ

This is fine but doesn't fully utilize the computer's potential for graphically pleasing display of results.

Formatting the data is one thing easily done on a computer but not as easily done on many calculators...even the printing variety.

A more pleasing way to display the above would simply require use of **TAB** funtions in the computer. Instead of just using the ";" to print the data immediately following the <ENTER> statement, we can use **TAB 37** for the data and **TAB 47** for the tag (MHZ).

The exact placement would depend on your computer's line length and the number of decimal points typically printed out (on input data we have some control over this, of course.) At any rate, using this techniques would make our data look like:

```
<ENTER> FREQUENCY #1          =  155.000    MHZ
<ENTER> FREQUENCY #2          =  10.700     MHZ

<ENTER> HARMONIC LIMIT        =    5
<ENTER> UPPER FREQUENCY LIMIT=  175.000    MHZ
<ENTER> LOWER FREQUENCY LIMIT=  135.000    MHZ
```

. I've also moved the "=" to line up at **TAB 34.**

One could use a subroutine to align the decimal places as well, but I think the blocking of data as shown provides sufficient neatness for most purposes.

In some cases it is nice to be able to see all required inputs before inputting data. On my rather slow **TIMEX**, printing all of the above except the data inputs as one massive **PRINT** statement is much faster than via the single line prompt method.

With the **TIMEX** this is relatively simple to do. The machine uses **PRINT AT** statements with an **X, Y Coordinate** system so that I can place the beginning of any line or word at a given number of lines down and/or a given number of columns in from the left edge. By judicious use of the **TAB** function, the basic input chart would print as follows:

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
      SPURIOUS RESPONSES
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

FREQ #1           =           MHZ
FREQ #2           =           MHZ

MAX HARM          =

FREQ MAX          =           MHZ
FREQ MIN          =           MHZ

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

The program for this is:

```
10 LET G$="XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX"
20 PRINT G$; TAB 6; "SPURIOUS RESPONSES";
   TAB 0;G$; TAB 1; TAB 0; "FREQ 1"; TAB
   18; "=" ; TAB 29; "MHZ"; TAB 0; "FREQ
   2"; TAB 18; "="; TAB 29: "MHZ"; TAB 1;
   TAB 0; "MAX HARM"; TAB 18; "="; TAB 1;
   TAB 0; "FREQ MAX"; TAB 18; "="; TAB
   29; "MHZ"; TAB 0; "FREQ MIN"; TAB 18;
   "="; TAB 29; "MHZ"; TAB 1; TAB 0; G$
```

Of course I would use one of the TIMEX graphics characters rather than the "X" for the border and the 32 character format is established by the screen and the printer limitations on that machine. The "=" and "MHZ" are set in position by the **TAB** as are the line feeds (TAB 0; TAB 0 would move the print location two lines down if the printing location was already past the TAB 0 on the current line).

This format allows input of up to 7 digits in the frequency which should allow for any typical set of frequencies up to 1000 MHz and up to three decimal places to 9999.999 MHz.

To get the entered data to print at the right location, the **TIMEX "PRINT AT X,Y"** statements are used after the **INPUT** statements:

```
 30 INPUT F1
 40 PRINT AT 4,20;F1
 50 INPUT F2
 60 PRINT AT 5,20;F2
 70 INPUT H
 80 PRINT AT 6,24;H
 90 INPUT FU
100 PRINT AT 9,20;FU
etc.
```

What this does is to place the frequency, F1 at location 4,20. This would be 5 spaces down from the top (remember, computers count from 0, not 1) and 20 spaces over from the left. This is one space beyond our equals sign which is at 18 spaces over from the left. Providing we do not input a number longer than 8 digits (9999.999 including the decimal), we will not overwrite our **"MHZ"** previously put at 29 spaces from the left.

THE PROGRAM

While all of the above is part of the program, **THE PROGRAM**, that is, the heart of the calculations we want to do, is the next module to set up.

We know that we want to look at F1 and F2 with respect to their additive and subtractive products over a limited set of harmonics and (usually) a limited set of possible frequencies. We have defined our variables: F1, F2, H, M, N, FU, and FL.

Now we must determine how to put these into a calculation that will compute what we want to know.

LOOPS WITHIN LOOPS

It is obvious at once to a programmer that the thing to use to make the required calculations is a set of **Nested Loops**.

If you want to repeat something, you go back to the beginning of that sequence by telling the computer to go back to the beginning and doing it over.

Unfortunately, there is no way to get out of the loop once started unless some limitation is place on it **inside** the loop.

In our case, we want to look at all values of F1 up to its H harmonic in relationship to all values of F2 up to its H harmonic...H being the maximum harmonic of interest.

In other words, if H is 3 (the third harmonic), we want to look at:

```
1F1+1F2     ABS(1F1-1F2)
2F1+1F2     ABS(2F1-1F2)
3F1=1F2     ABS(3F1-1F2)

1F1+2F2     ABS(1F1-2F2)
2F1+2F2     ABS(2F1-2F2)
3F1+2F2     ABS(3F1-2F2)

1F1+3F2     ABS(1F1-3F2)
2F1+3F2     ABS(2F1-3F2)
3F1+3F2     ABS(3F1-3F2)
```

These are the 18 possible mixing products for F1 and F2 up to their third harmonic.

We would use a very basic **FOR-NEXT** loop to do this job. The loop would be:

```
180 LET N=1
190 LET M=1
200 FOR I=1 to H
210 LET A=N*F1+M*F2
220 LET B=ABS(N*F1-M*F2)
230 PRINT A,B
240 NEXT I
```

I, by the way, is frequently used in FOR/NEXT loops since is conveniently stands for Increment.

But, of course, this does not really do it for us, we want to have N and M change as in the chart above.

As shown in the chart, we want to have N change with each calculation, then step M once, then have N change through each harmonic again, etc.

This is done with a **nested loop.**

We want N to be limited to H as its highest harmonic number so we use:

```
200 FOR N=1 TO H
210 LET X=...etc.
230 NEXT N
```

This will step the value of N from 1 to H (which we have set to 3) then go to the next step.

Thus, this loop will give us the F1, 2F1, 3F1 steps for N*F1.

Each time we go through one of these N=1 to 3 cycles, we then would like to step M by 1 and repeat this until M becomes 3 (or H). To do this we place a second FOR-NEXT loop around the N=1 TO H loop.

Thus N will step from 1 to H (3), go to the next step in the program, which is to increment M by 1, then go through N=1 to H again, increment M by 1 again, until both N and M are equal to H.

This set of loops looks like:

```
200 FOR M=1 TO H
210 FOR N=1 TO H
220 LET A=N*F1+M*F2
230 LET B=ABS(N*F1-M*F2)
240 PRINT A,B
240 NEXT N
250 PRINT
260 NEXT M
```

This prints out all possible mixing products exactly a shown earlier in the chart. The 260 PRINT statement provides the space between each value for M.

Next we want to limit what we print out on the screen to only those values that fall within our frequency limits.

We merely state that if A is greater than the Lower Frequency Limit and less than the Upper Frequency Limit, PRINT A. The same statement can be made concerning B.

In BASIC this would be:

```
IF A >= FL AND A <= FU THEN PRINT A
IF B >= FL AND B <= FU THEN PRINT B
```

So we can add these statements into our loops:

```
200 FOR M=1 TO H
210 FOR N=1 TO H
220 LET A=N*F1+M*F2
230 LET B=ABS(N*F1-M*F2)
240 IF A >= FL AND A <= FU THEN PRINT A
250 IF B >= FL AND B <= FU THEN PRINT B
260 NEXT N
270 PRINT
280 NEXT M
```

MORE OUTPUT

Actually, it would be nice if we could print out a bit more information than just what spurious frequencies will be generated.

Ideally, the printout should give us more information than we get from the above statements. We have already calculated just exactly which harmonics of which input signals create which spurious outputs. We should have the computer print this information out for us, too.

Here, then, is a set of PRINT statements that will accomplish this goal:

```
240 IF A >= FL AND IF A <= FU THEN PRINT TAB 0;
G$; N;" X F1 + ",M;" X F2 ="; TAB 18; A; TAB 29;
"MHZ"

250 IF B >= FL AND IF B <= FU THEN PRINT TAB 0;
G$; N;" X F1 - ",M;" X F2 ="; TAB 18; B; TAB 29;
"MHZ"
```

This tells the computer to check the results of the calculation to see if it falls in our FL-FU range, print the the Graphics String (G$), the harmonic number (N) " X F1 + " (M) " X F2= " then the result (A) "MHZ". A similar thing happens for (B):

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX      G$
N X F1 + M X F2 =  RES.ULTA  MHZ      A
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX      G$
N X F1 - M X F2 =  RES.ULTB  MHZ      B
```

It might be good to eliminate the PRINT step between NEXT N and NEXT M unless it is desired that there will be a space (or one could use a graphics line) between printouts of M=1, M=2, etc.

ENDING THE PROGRAM

At the end of all calculations and comparisons the FOR/NEXT loop will, of course, go to the next step in the program which will be either an ending statement of some kind or a request for more data.

To end the program the following line would be adequate:

```
260 PRINT G$
```

This, of course, would print a single Graphics String to close off the bottom of the last result of the program.

Some may wish to change the frequencies, say, check the impact of a 21 MHz IF versus a 10.7 MHz IF. Or one might wish to look at the spurs over a wider or narrower frequency limit. Rather than re-entering all of the data, the program could end in a way which would allow just one or two items to be changed.

One way to do this would be:

```
        260 PRINT G$

        270 PRINT "CHANGE VARIABLES? (Y/N)
CHG?    280 INPUT C$
        290 IF C$="Y" THEN GOTO 350

STOP    300 STOP (END)

CLS     350 CLS (CLEAR SCREEN)

INSTR   360 PRINT "ENTER NEW VALUE OR 0"

        370 PRINT "FREQ 1= ";F1
F1      380 INPUT FC1
        390 IF FC1 <> 0 THEN LET F1 = FC1

        400 PRINT "FREQ 2= ";F2
F2      410 INPUT FC2
        430 IF FC2 <> 0 THEN LET F2 = FC2

        440 PRINT "HIGHEST HARM= ";H
H       450 INPUT HC
        460 IF HC <> 0 THEN LET H = HC

        470 PRINT "UPPER FREQ= ";FU
FU      480 INPUT FCU
        490 IF FCU <> 0 THEN LET FU = FCU

        500 PRINT "LOWER FREQ= ";FL
FL      510 INPUT FCL
        520 IF FCL <> 0 THEN LET FL = FCL

CLS     530 CLS (CLEAR SCREEN)

RESTART 540 GOTO 20
```

But we have a problem. If we just go back to line 20 we will need to input each piece of data again. To avoid this, we will change the INPUT lines (lines 30, 50, 70, 90, 110) to IF C$ <> "Y" THEN... This will remove the problem. Since C$ is not given prior to line 30, however, we must add a line to do this. Thus we will have:

```
  5 LET C$=""

 30 IF C$ <> "Y" THEN INPUT F1

 50 IF C$ <> "Y" THEN INPUT F2

 70 IF C$ <> "Y" THEN INPUT H

 90 IF C$ <> "Y" THEN INPUT FU

100 IF C$ <> "Y" THEN INPUT FL
```

HP 25 PROGRAM

Just to show the difference between computer implementation of this program and the same program on an HP 25 calculator, let's briefly look at the HP 25 version:

SET-UP:

1) STO 0	= 0	
2) STO 1	= 0	
3) STO 2	= F1	Frequency 1
4) STO 3	= F2	Frequency 2
5) STO 4	= FL	Lower Limit
6) STO 5	= FU	Upper Limit
7) STO 6	= 0	
8) STO 7	= H	Harmonic Limit

PROGRAM:

1) RCL 0	16) R V	31) 0
2) RCL 2	17) R/S	32) STO X 0
3) X	18) NOP	33) GTO 01
4) RCL 1	19) 1	34) NOP
5) RCL 3	20) STO+0	35) RCL 2
6) X	21) RCL 0	36) RCL 3
7) -	22) RCL 7	37) +
8) ABS	23) X>=Y	38) -
9) RCL 4	24) GTO 01	39) R/S
10) X>=Y	25) 1	40) CLX
11) GTO 19	26) STO+1	41) GTO 19
12) CLX	27) RCL 1	42) SIN -1
13) RCL 5	28) RCL 7	
14) X < Y	29) X<Y	
15) GTO 19	30) GTO 42	Change line 07 to + for other spurs.

MODIFICATION FOR OTHER PURPOSES

It will often become apparent that one program can be used for a similar, but different purpose by minor modifications.

Our Spurious Mixing Products program can be applied to Receiver Intermodulation products in a very similar program.

Here, however, we are looking for possible input frequencies which could interfere with a given receiver channel, say in a Repeater System or FM or TV translator receiver.

Thus we will want to step through all possible channels to find those that could mix together and fall within or adjacent to our receiver's passband.

I won't go into full details, but FOR/NEXT loops can be used as follows:

```
FI = Input Frequency (Desired Channel)
CH = Channel Spacing (for band being used)
FU = Upper Frequency Limit (RX Front End)
FL = Lower Frequency Limit (RX Front End)
CU = FI + CH (Upper Adjacent Channel)
CL = FI - CH (Lower Adjacent Channel)
 H = Maximum Harmonic (each frequency)

           10 INPUT FI
           20 INPUT CH
DATA       30 INPUT FU
           40 INPUT FL
           45 INPUT H
           46 LET CL=FI-CH
           47 LET CU=FI+CH

           50 FOR A=FL TO FU STEP CH
LOOPS      60 FOR B=FL TO FU STEP CH
           70 FOR C=1 TO H
           80 FOR D=1 TO H

CALC       90 LET E=ABS(C*A-D*B)

COMP      100 IF E >= CL AND <= CU THEN
              PRINT G$;C;" X ";A;" - ";PRINT;
              D;" X ";B;" = CHANNEL"

          110 NEXT D
LOOP      120 NEXT C
CLOSE     130 NEXT B
AND       140 PRINT G$
END       180 NEXT A
          190 PRINT G$;G$
          200 STOP (END)
```

This program would print out something like the following for FI=101.5, CH=.2, FU=102.5, FL=100.5, and H=3:

```
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
2 X 100.5 - 3 X 100.9 = CHANNEL
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1 X 100.5 - 2 X 101.1 = CHANNEL
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
1 X 100.7 - 2 X 101.1 = CHANNEL
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
etc.
```

COMBINING PROGRAMS

One problem with small computers is that they frequently lack a Disc Drive. That means that one are likely to be loading things from a cassette. Cassette data storage is notoriously slow.

One answer to this is to buy a Disc Drive, of course, but another solution is to buy one of the specialized programs developed to speed up loading and saving with cassettes. There are several such programs for the TIMEX, some even on EPROM, and at least one for the COMMODORE 64.

Since these quickloading programs can load 16K of data from cassette in 30-60 seconds, a combined large program made up of several related smaller programs may be an effective way to save time. This is especially true if one program needs to be run prior to running a second program.

The main consideration with interactive programs is to make sure that the correct program is run first and that the variables of one do not conflict with the variables of the other.

Since our two Spurious programs are somewhat related, we'll use them as an example of one method of combining programs:

```
10 PRINT G$;TAB 15;TAB 13;"MENU:";
   TAB 3;"1) SPURIOUS MIXING OUTPUTS";
   TAB 3;"2) RECEIVER INTERMODULATION";
   TAB 1;TAB 0;G$
20 INPUT S$
30 IF S$="1" THEN GOTO 500
40 IF S$="2" THEN GOTO 1000
50 IF S$ <> "1" AND S$ <> "2" THEN GOTO 20
```

This program MENU will be printed by line 10 and will go to 500 if "1" is selected or to 1000 if "2" is selected. The start of the Spurious Mixing Outputs program would, of course, need to be numbered between 500-999. The Receiver Intermodulation program between 1000-xxxx.

Furthermore, at the end of each should be some means of either continuing with the same program or going back to the MENU. This is most easily accomplished by replacing the "STOP" line in the end of the earlier described SPURIOUS program with:

```
300 GOTO 0
```

Thus, if the answer to "AGAIN (Y/N)" is not "Y", the computer will execute the next step, GOTO 0, which prints out the MENU. It will do this regardless of the actual letter put in S$, of course, since only "Y" will restart the program currently in use.

If you wish to be able to stop everything, add:

```
295 IF C$="S" THEN STOP
```

Then change line 270 to:

```
270 PRINT "CHANGE VARIABLE (Y/N/S=STOP)
```

This way a "Y" restarts the program, "N" takes you to the MENU, and "S" stops the computer.

SUB-ROUTINES

The last thing we will look at is the use of **subroutines.**

The simplest way to modify data... convert from dB to ratio, dBm to milliwatts or uV, and so forth, is to provide a subroutine which can be called upon in your program whenever necessary. This is particularly important if the conversion is made several times during a program.

Subroutines can be of great assistance in developing new programs by allowing use of old subroutines where appropriate. Those with UTILITY programs which allow combining programs by entering one, renumbering, then entering another and merging the two could save a lot of time by having a file of subroutines on tape (or disc) which you would then be able to call upon as required.

This **modular** approach to programming isn't always the easiest technique to dis-assemble once the program is done, however. Trying to follow the logic of a **BASIC** program using sub-routines extensively is not my idea of great fun! For simpler programs I generally don't use sub-routines unless re-use of the routine warrants this.

SOME RF / ELECTRONICS SUB-ROUTINES

CAPACITIVE REACTANCE:

```
10 LET F=X
20 LET C=Y
30 GOSUB 1000
40 (Next steps using XC etc.)
```

(X, Y and 10, 20, 30 are in main program)
(X=Frequency in Hz, Y=Capacitance in F)

```
1000 LET XC=1/(2*PI*F*C)
1005 RETURN
```

```
10 LET F=X
20 LET XC=Y
30 GOSUB 1010
40 (Next steps using C, etc.)
```

(X=Frequency in Hz,Y=Capac. Reac. in Ohms)

```
1010 LET C=1/(2*PI*F*XC*1E-6)
1015 RETURN
```

INDUCTIVE REACTANCE:

```
1020 LET XL=2*PI*F*L
1025 RETURN
```

(F in HZ, L in H)

```
1030 LET L=XL/(2*PI*F)
```

(F in HZ, XL in Ohms)

```
1035 RETURN
```

DECIBELS TO VOLTAGE:

```
1040 LET V=10**(DV/20)
1045 RETURN
```

VOLTAGE TO DECIBELS:

```
1050 LET DV=20*LN V/LN 10
1055 RETURN
(NOTE: Timex uses LN for Natural Log.  Many
       computers use LOG as in BASIC this
       denotes Natural Log.)
```

DECIBELS TO POWER:

```
1060 LET DP=10*LN V/LN 10
1065 RETURN
```

POWER TO DECIBELS:

```
1070 LET P=10**(DP/10)
1075 RETURN
```

DBM TO UV, V, MV

```
1080 LET UV=223607*10**(DBM/20)    (50 OHMS)
1085 RETURN

1090 LET UV=273861*10**(DBM/20)    (75 OHMS)
1095 RETURN

1100 LET V=.707107*10**(DBM/20)   (500 OHMS)
1105 RETURN

1110 LET V=1000*10**(DBMV/20)      (75 OHMS)
1115 RETURN
```

UV, V, MV TO DBM

```
1120 LET DBM=20*LN (UV/223607)/LN 10  (50)
1125 RETURN

1130 LET DBM=20*LN (UV/273861)/LN 10  (75)
1135 RETURN

1140 LET DBM=20*LN (V/.707107)/LN 10  (500)
1145 RETURN

1150 LET DBMV=20*LN (MV/1000)/LN 10   (75)
```

TURNS RATIO

```
1160 LET T=SQR (Z1/ZO)    (or Z1/Z2)
1165 RETURN
```

CONCLUSION

Many of us have been confirmed calculator button pushers in the past. We will continue to be so in the future.

However, there are some things that can be done better on a computer and computers are so reasonably priced today that not owning one is difficult for an engineer to justify (unless he's got a $5000 one at work, of course!)

I hope that this paper has given some insight for the beginning computer user into how programs can be put together for best utility. Many will go on to far exceed the level presented here.

A book that I recommend for the Electronics tech or engineer is:

BASIC COMPUTER PROGRAMS IN SCIENCE AND ENGINEERING

Jules H. Gilder

Hayden Book Company
Rochelle Park, N.J.

ISBN 0-8104-0761-2

Aside from the usual technical publications such as **RF Design**, another good source of RF related programs on the simple to moderate level of complexity is **HAM RADIO** magazine.

CIAO and DESIGN:

Circuit Optimization and Synthesis Programs for Personal Computers

by

STEPHEN E. SUSSMAN-FORT

Associate Professor
Department of Electrical Engineering
State University of New York
Stony Brook, New York 11794
(516) 246-5141 (516) 246-6757

and

President
SPEFCO Software
18 Bennett Lane
Stony Brook, New York 11790
(516) 751-2376

Table of Contents

ABSTRACT

1.1 ABSTRACT

The CIrcuit Analysis and Optimization program CIAO (pronounced 'chow') and the Matching Network Synthesis program DESIGN are powerful tools for RF and microwave linear circuit design that have been developed for personal computers. Versions of CIAO and DESIGN are available for the IBM-PC [1] and most other 16-bit MS-DOS [2] machines. A RAM of at least 256K is required to run CIAO, while 128K will suffice for DESIGN. Both programs will run under DOS version 1.0 or above; in addition, full support of the 8087 numeric data processor chip is optionally available. A slightly abbreviated version of CIAO as well as the complete version of DESIGN is also available for most 8-bit, 64K RAM, CP/M-80 [3] machines.

CIAO provides the ability to analyze and optimize broad classes of passive and active networks. Both the magnitudes and phases of the scattering parameters of the networks - which may consist of arbitrary interconnections of R, L, C, controlled-source, gyrator, transmission-line, two-port, and three-port elements - can be optimized over a frequency band. DESIGN performs the truly automated synthesis of broadband, gain-sloped, lumped-element or distributed-parameter matching networks. DESIGN uses tabulated load data in its synthesis process; no equivalent circuit load-modeling is required. Both CIAO and DESIGN have been shown to work efficiently and accurately in a wide variety of cases.

INTRODUCTION

2.1 CIAO's Analysis Capabilities

CIAO performs the analysis and optimization of linear circuits in the frequency domain. Specifically, CIAO can calculate the complex scattering parameters, over a specified frequency range, of a one-port or two-port network which is described by a data file. The circuit elements recognized by CIAO (for both analysis and optimization) include:

- Resistors
- Capacitors
- Inductors
- Voltage-Controlled Current Sources (VCC) with optional Delay
- Gyrators
- Two-port Transmission Lines with optional Loss
- Open-Circuit Transmission-Line Stubs with optional Loss
- Short-Circuit Transmission-Line Stubs with optional Loss
- One-ports described by tables of scattering (S), admittance (Y), or impedance (Z) parameters over frequency
- Two-ports described by tables of S, Y, or Z parameters over frequency
- Three-ports described by tables of S parameters over frequency.

Each port-reference-impedance of the one-port or two-port may be characterized by either single real value or by a one-port S, Y, or Z table description. All two-ports may have four independent terminals; the three-ports are common-terminal types. Appendix A shows how other circuit elements, such as ideal transformers, controlled-sources other than VCC, and mutual inductance can be accurately modeled by simple interconnections of a few of the library elements.

2.2 CIAO's Optimization Capabilities

Starting with an initial set of element values for a circuit design, CIAO can obtain, by means of an iterative optimization method, a new set of element values which provides an improved frequency response for the circuit. The various modes and features of the optimization process will now be described.

2.2.1 Magnitude Optimization

In the magnitude optimization mode, CIAO attempts to match, over a specified frequency band, any number of the four scattering-parameter MAGNITUDES of the two-port (or the single reflection-coefficient magnitude of the one-port) to target values or ranges specified in the data file. The program performs this task by optimally perturbing designated element values in the circuit. The target values or ranges may vary with frequency; in addition, weights may be specified to to achieve a better match with some S-parameter magnitudes than with others at certain frequencies. The program provides substantial flexibility in expressing just what the goals of the optimization should be.

2.2.2 Table Optimization

In the table optimization mode, CIAO attempts to match, over a given frequency range, both the MAGNITUDE and PHASE of the network's scattering parameters to values given in the data file's special optimizing table. Separate weights may be applied to the magnitude and phase of each S-parameter. By choosing the weighting factors appropriately, it is possible to optimize either the magnitude or phase separately, or to simultaneously optimize the magnitude and phase of any or all of the S-parameters to achieve the best match over the specified frequency band. In addition, a simple data-file command statement enables the conversion of the optimizing table target values to ranges of permissible values for the S-parameter magnitudes.

2.2.3 Delay Optimization

In delay equalizer applications, it is necessary only to match the delay through a given network to a specified shape over a frequency range, rather than to a set of absolute values. (The magnitude response is generally sloped or flat, depending on whether gain equalization is also required.) Since delay is the negative derivative with respect to frequency of the phase of the network's S_{21}, matching a network response to a given delay shape is equivalent to matching that response to the corresponding phase shape plus an arbitrary constant phase plus an arbitrary linear phase. CIAO optionally allows this type of phase optimization for S_{21} if desired.

It should be noted that COMPACT [4], the well-known microwave-circuits computer program, attempts to calculate delay directly. There are two problems with this approach. First, COMPACT uses numerical differentiation of the phase to approximate the delay of S_{21}, and numerical differentiation is a notoriously noisy and inaccurate process. The second problem is even more serious. By using numerical differentiation to calculate delay, one cannot take account of the contribution to the delay of any circuit element which is described by a table of values at distinct frequencies. This is easy to see, since an element described by a table of values has for its phase description a piecewise-linear curve, and the derivative of this piecewise-linear curve is discontinuous at just those frequency points where we would like to evaluate the derivative. In fact, COMPACT ignores the contribution to delay of any circuit element described by a table of values. We have confirmed that this causes substantial error in network-response calculations for all but very narrow-band applications or for those trivial cases where the phase of the table-described element remains constant over the given frequency range. Furthermore, in some cases, COMPACT takes the absolute value of the computed phase before performing the numerical differentiation, resulting in a sign error for the delay [5]. Our approach, using phase calculations, avoids all of the aforementioned difficulties and provides consistently accurate results over a wide range of circuit conditions.

2.2.4 The Optimization Method

To perform the optimization, CIAO first constructs an error function consisting of the sum of the squares of the errors between the CALCULATED magnitudes (and phases) and the DESIRED magnitudes (and phases) of the network's scattering parameters over a given set of frequencies. Analytically, the most general form of the error function used by CIAO is given by

$$\text{Error Function} = \sum_{f} \sum_{i,j} \Big[Wm_{ij}(f) * [Sm_{ij}(f)(\text{calc.}) - Sm_{ij}(f)(\text{des.})]^2 + Wa_{ij}(f) * [Sa_{ij}(f)(\text{calc.}) - Sa_{ij}(f)(\text{des.})]^2 \Big]$$

where:

f denotes frequency;

Sm_{ij} and Sa_{ij} denote the magnitudes and phase angles of the S-parameters;

and

Wm_{ij} and Wa_{ij} denote the magnitude and phase angle weighting factors.

For a two-port, i=1,2 and j=1,2 in the above equation, while for a one-port only the i=1, j=1 case is considered. Special methods (that we shall not discuss here) are used to handle the cases where the desired S-parameter values are represented by ranges of permissible values.

A modified version of the Fletcher-Reeves [6] optimization algorithm is then used to minimize the value of the error function by perturbing designated element values. The algorithm requires a knowledge of the partial derivatives of the error function with respect to the optimizable element values, and it is here that CIAO provides a unique ability. We calculate the EXACT partial derivatives of ALL the optimizable elements - regardless of their number - using just TWO circuit analyses. This is made possible by using adjoint network [7] techniques. The result is an extremely fast optimization process which converges under a wide variety of circuit conditions.

It should be noted that COMPACT, which employs a gradient-type optimizer similar to that of Fletcher and Reeves, uses the method of finite-differences to calculate derivatives. This is an approximate method that requires two circuit analyses per EACH optimizable circuit element, thus yielding much slower execution times than would otherwise be possible using adjoint techniques. Furthermore, the finite-difference approximation can sometimes be insufficiently accurate to obtain convergence to an acceptable solution. With CIAO, however, the optimization almost always proceeds rapidly and predictably towards satisfactory convergence; in addition, the program user may interrupt and then continue or stop the optimization process at any time.

Random-grid-search methods have sometimes proved effective in circuit optimization problems, particularly when the initial design is very far from the optimum. Typically, the error function is evaluated several times for a range of element values - i.e. over a "grid" in the vector space of optimizable element values - and from this a "better" set of circuit parameters is deduced. When the grid is large enough, convergence to local minima can be avoided, but this often requires a relatively large number of error-function evaluations. A future version of CIAO will additionally include a random-grid search method for optimization. However, CIAO's specially-modified gradient optimizer already does indeed provide the capability to achieve convergence for some rather difficult circuit problems.

2.3 The Methods and Capabilities of DESIGN

We now discuss the capabilities of the program DESIGN, which can automatically synthesize broadband, gain-sloped matching networks, and which requires only source and load impedance values as essential data. This program alone can design matching networks to meet basic system requirements. In conjunction with CIAO, DESIGN can synthesize precision matching networks that are optimized in the face of non-ideal effects such as transmission-line loss and the non-unilateral-device assumption.

The synthesis method used in DESIGN is based upon the work of Carlin and Komiak [8]. It has been shown that matching networks designed by this technique are "simpler in structure and superior in frequency response to equal-ripple designs" [9] based upon the classical approaches of Fano [10] and Youla [11] as developed by Chen [12] and Mellor [13]. To the best of our knowledge, DESIGN is the only commercially available program that employs the superior Carlin-Komiak technique for matching network synthesis.

2.3.1 DESIGN: What It Does

DESIGN synthesizes lossless matching networks to provide a specified S_{21} magnitude response (Sm_{21}) across a frequency band between a real source impedance and a complex load impedance. (The complex-source to complex-load case can be easily handled in conjunction with CIAO.) The program user interactively inputs the desired degree of the network, the source and load data, and the desired values for Sm_{21} across the passband. The frequency points specified for the complex load define the passband of the network.

Since the relationship

$$(Sm_{11})^2 = 1 - (Sm_{21})^2$$

holds at any frequency for lossless two-ports, specifying "1" for Sm_{21} across the passband will generally produce a matching network with minimum reflection coefficient values at the source and load ports. (Recall that the magnitudes of the reflection coefficients at the ports of a lossless reciprocal two-port are always equal.) Specifying a varying Sm_{21} (<=1) yields a gain-sloped matching network, with the accompanying tradeoff of higher reflection coefficients at the ports. Thus DESIGN affords substantial flexibility in its synthesis process.

The design procedure is based on a lumped, shunt-capacitor, series-inductor, lowpass topology, with an optional shunt-inductor at the load end of the network providing a bandpass response, if so desired. The program indicates which topology is most appropriate for your particular load data, and the program user has the option to follow this suggestion or not. If the bandpass structure is selected, the shunt inductor at the load must be assigned a fixed value (a suggested value is supplied by DESIGN), and then this inductor is automatically "absorbed" by the original load to create a new complex load. The synthesis then proceeds on a lowpass basis, and ultimately provides the values for series-inductors and shunt-capacitors. The program also produces an equivalent distributed-parameter matching network consisting of open-circuit and short-circuit transmission-line stubs. No data file is required; all input to the program is requested interactively.

2.3.2 DESIGN: The Method Briefly Explained

2.3.2.1 Step 1.

The Carlin-Komiak method, as implemented in DESIGN, begins with the construction of a piecewise-linear function of frequency, R(f). This function is an initial approximation to the real part of the matching-network impedance looking back into the network at the complex-load end, with the network terminated at the other end in its real source impedance. To construct R(f), a lowpass topology is assumed with R(0) being the source impedance, and $R(f_n)$ being zero at some frequency f_n above the highest frequency specified for the complex load. The frequency chosen for f_n will depend upon the degree selected for the network. Experience must dictate the degree to choose, with degrees of 3, 4, or 5 being common for many applications. A number of break frequencies f_i must then be chosen in the passband between zero and f_n, along with the corresponding resistance values $R(f_i)$, to fully define the initial guess R(f). The number of break frequencies is dependent upon the bandwidth of the network and the desired degree. It most cases it is sufficient to select fewer than 7 break frequencies which are more or less uniformly distributed across the passband. The $R(f_i)$ may initially be assumed to vary in a linear fashion between R(0) and zero at f_n. It should be noted that DESIGN makes all the aforementioned decisions for the program user (except for the degree specification); generally, the sophisticated user is allowed to override certain default values generated by the program.

2.3.2.2 Step 2.

The program then calculates the reactance function X(f), from R(f), by the method of the Hilbert Transform. This calculation is extremely efficient by virtue of the piecewise-linearity of R(f). At this point, we note that the (non-optimal) matching network is completely defined by R(f) and X(f), and could be synthesized by standard network-theoretic techniques. The next step is to calculate Sm_{21} of the matching network from a knowledge of R(f), X(f), and the source and load impedances across the frequency passband. Then an error function is constructed as the squared-error between the calculated values of Sm_{21}, and those values of Sm_{21} that we wish the network to have in the passband.

2.3.2.3 Step 3.

Finally, an iterative optimization procedure is employed to find the break-frequency resistance values $R(f_i)$ that minimize the error function. In other words, we seek an optimal R(f) and corresponding X(f) which will yield the desired S_{21} magnitude response for the network. The optimization is numerically very well-conditioned because the error function, as it can be shown, is only quadratically dependent upon the unknowns $R(f_i)$. Consequently, in almost all cases, the optimization converges satisfactorily.

DESIGN then performs the network synthesis steps necessary to yield the final element values for both the lumped and distributed versions of the desired matching network. This involves curve-fitting a rational polynomial to the optimized function R(f), finding the left-hand plane roots of the corresponding function of the complex-frequency variable s, constructing the matching-network's transducer-loss function H(s) = $1/Sm_{21}(s)$, and then expanding this function in an appropriate manner to find the actual element values. Each step of the synthesis process is summarized in DESIGN's output.

In those exceptional cases where DESIGN either initially fails to converge, or obtains physically unrealizable element values from the curve-fit, various strategies may be employed which almost always persuade DESIGN to eventually succeed in the synthesis. These include: (1) overriding certain default values established by the program; (2) selecting a bandpass network structure instead of a lowpass structure or vice versa; or (3) increasing the degree of the network.

2.3.2.4 Step 4.

CIAO can also be used with DESIGN to synthesize matching networks between a complex source and a complex load, or to compensate matching networks for a variety of non-ideal effects. In addition, CIAO can improve the response accuracy of any network synthesized by DESIGN; this may be desirable, occasionally, to compensate for the approximations inherent in the Carlin-Komiak method.

2.4 Entering Data for CIAO and DESIGN

The program CIAO always requires a disk data file, while DESIGN requests all data to be typed in at the keyboard at run-time. All data is entered for DESIGN in response to queries from the screen. Data entry is free-format. When more than one piece of data is entered on the same line, any number of spaces may separate the data fields. The program traps errors on data input, and allows revision of all data before proceeding.

Input data for CIAO is taken from disk files. Certain options are specified by the user in response to screen queries. Both CIAO and DESIGN do not distinguish between upper or lower case characters, hence all alphabetic input (from either the keyboard or a disk file) may be either upper or lower case, or even mixed upper and lower case. Numeric input data for both programs may be entered in integer, floating-point, or scientific ("E") notation (e.g. 1.23e-7).

For CIAO, a data file stored on disk needs to be constructed to describe the circuit to analyzed or optimized. Any ASCII editor may be used for this purpose, and almost any legal operating system file name may be used for the data file. (The only exception is that the file may never have the extension .BAS, as this would denote that the file has line numbers, as discussed presently.) It is also possible to use a BASIC editor to create disk data files for CIAO. One need only create a file, with BASIC's usual line numbers, according to the specification of the circuit, and then save that file in ASCII mode, as in

```
SAVE "B:CIRCUIT.BAS",A .
```

For CIAO, data files are constructed line-by-line in free format, starting in any column, with any number of blanks delimiting the data fields. Any number of blank lines may be interspersed throughout the the data file. Comment lines are permitted; they are denoted by a single apostrophe (') being the first character on the line. Comments may also be appended to the end of any line of data as long as the comment is separated by at least one space from the last valid data field on the line, and the first character of the comment is neither an integer (0-9), a plus sign (+), a minus sign (-), nor a period (.). In addition, CIAO will detect and identify many types of errors in the data file, and will instruct the user to return to the operating system to edit the file.

CIAO and DESIGN PROGRAM SESSIONS

3.1 A CIAO Sample Session

3.1.1 The Circuit

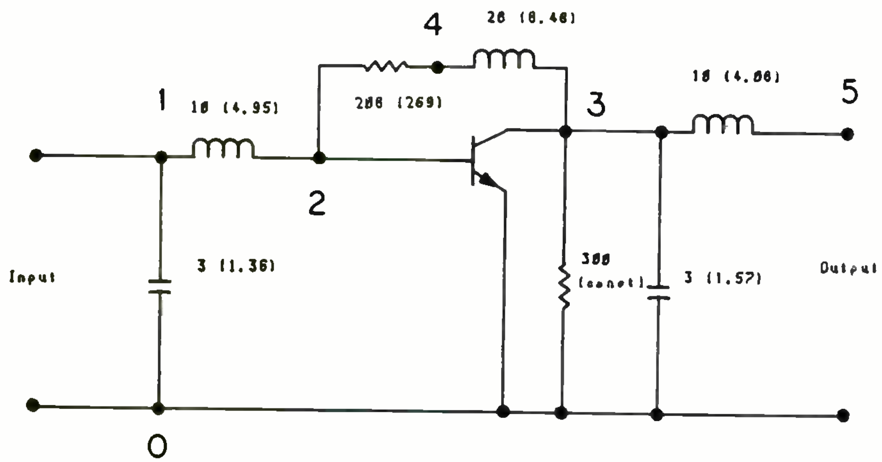
The circuit shown in Fig. 1 will be used to demonstrate some of the basic features of CIAO. The circuit has been discussed in [14] and [15]. In [14], a optimization of the circuit was performed using SUPER-COMPACT; we show here for this example that our results with CIAO duplicate, if not exceed, the performance of this well-known program.

The circuit is a single-stage microwave transistor amplifier whose circuit response is to be improved by perturbing the nominal element values shown in Fig. 1. Specifically, new element values are to be found to achieve the smallest values for S_{11} and S_{22} while achieving a flat gain for S_{21} of 3.162 (10 dB) across the frequency band 10 MHz - 250 MHz. The CIAO data file to achieve this is shown immediately following Fig. 1.

Fig. 1: CIAO Optimization Example

Initial values as shown; (optimized values in parentheses).

Units: pF, nH, and ohms.



CIAO Data File for Circuit of Fig. 1

```
'CIRCUIT DESCRIPTION SECTION  (* means optimize)

cap* 1 0 3e-12
ind* 1 2 10e-9
res* 2 4 200
ind* 4 3 20e-9
res  3 0 300
cap* 3 0 3e-12
ind* 3 5 10e-9
two 2 0  3 0  table1

port1 1 0  50
port2 5 0  50

perform twoport optimization magnitudes
END


'THE FREQUENCIES SECTION  (freq (Hz) \ |Sij|-weight/|Sij|-goal)

10e6 100e6          \  1/0  0  1/3.162  1/0
250e6 1500e6 250e6  \  1/0  0  1/3.162  1/0
END


'The TABLES SECTION

table1 sparam
50
.95   -2  .003 84.3  7.35 174.6  1.01  -1
.92  -11  .007 79.0  7.15 168.0   .99  -4
.87  -28  .015 69.2  6.83 154.5   .96 -10
.78  -54  .026 54.0  6.28 135.0   .90 -18
.69  -78  .033 41.4  5.67 123.0   .84 -25
.63  -98  .037 33.0  5.04 113.0   .79 -30
.60 -114  .038 29.3  4.42  99.9   .77 -33
.60 -127  .039 28.0  3.88  87.0   .76 -35
END
```

3.1.2 The Data File

The CIAO data file for this example is typical of the data file format for any circuit. The file contains three sections: circuit description, frequencies, and (optionally) tables.

The circuit description is given in the familiar nodal interconnect format; nodes are numbered sequentially from zero, with node #0 being the ground node. Circuit elements are described by three-letter codes; in the present case, "cap", "ind", and "res" denote capacitor, inductor, and resistor, respectively. An asterisk appended to a code indicates that the element value is optimizable. The element code "two" defines a two-port circuit element which is specified by a table of S, Y, or Z parameters in the data file. In the present example, "two" is described by a table of scattering parameters.

The circuit description section also contains the port specifications and the analysis options. In the given example, port 1 is between nodes 1 and 0 with a 50 ohm reference impedance. The port 2 description is obvious from the file. The "perform" statement says here that the circuit is to be optimized to S-parameter magnitudes given in the frequencies section. Only the first three letters of each word in the "perform" statement are significant, hence, in the data file, this statement could have been written as "per two opt mag". CIAO will detect errors arising from invalid element types, missing or incomplete "port" or "perform" statements, and the like.

Each line of the frequency section contains frequency data, S-parameter weighting factors, and S-parameter goals. In the present example, S_{11} and S_{22} are given the weight/goal specification 1/0, which means that for these parameters the error function weighting factors are 1 and the desired values are 0. S_{12} is given the weighting 0, which means that its value does not contribute to the error function. S_{21} is given the weighting 1/3.162, which means that the error function weighting is 1 and the desired absolute magnitude is 3.162 (10 dB). Note that weights are defined for every frequency band, i.e. every line of frequency data. Frequencies themselves are specified in ascending order or by a first-frequency, last-frequency, frequency-increment format. The frequencies are separated from the weight data by a backslash (\) as shown, or alternatively, by a semicolon (;). At least one space must precede and follow the backslash or semicolon. Errors in this format or incomplete weight specifications are flagged and identified by CIAO at run-time.

The tables section gives the data for tables referenced in the circuit description section. There is only one table in the present case, but in general, up to 10 tables may be referenced. The order of appearance of the tables in the tables section is immaterial. A given table may be referenced any number of times by different circuit elements. Only the first letter of the table type is significant. In the present case, where table1 contains scattering parameters, the first entry (which must begin on a new line after the table statement) is the reference impedance, and subsequent entries correspond to the magnitude and phase (in degrees) for S_{11}, S_{12}, S_{21}, and S_{22} at the successive frequencies given in the previous data section. In the table, any number of entries may be placed on a given line; we have grouped the entries here by frequency for easy readability. If a circuit element references a table that does not exist in the data file, CIAO will detect this error at run-time.

3.1.3 The Program Output: Initial Analysis, Optimization Results, and Execution Speed

As mentioned earlier, CIAO is executed by typing in, at the keyboard, CIAO plus a carriage return. The user then enters responses concerning the name of the data file and the output device (either the screen or the printer). At this point, CIAO asks whether a summary of the input data should be output, and afterwards, it performs the initial analysis of the circuit. If optimization of the circuit is indicated in the "perform" statement, the number of iterations desired is requested, and then the optimization begins. The user may interrupt the optimization at any time by holding down any key. The program has internal criteria for stopping the optimization; the user is queried from the screen as to whether the optimization should continue or stop.

CIAO's output for the example of Fig. 1 is shown on the next two pages. Note that the initial and final analyses print, at each frequency, the network's scattering parameters (magnitude and phase in degrees), and stability factor K. The initial analysis, which agrees with that of [14, p. 224], shows input and output reflection coefficients that are as high as 0.63 and 0.87 respectively, and a S_{21} which varies from 6.0 to 11 dB. Ten optimizing iterations are then performed, with the results that S_{11} has a maximum value of 0.18, S_{22} has a maximum value of 0.11, and S_{21} varies from 9.7 to 10.1 dB across the frequency band. These results compare quite favorably with those of SUPER-COMPACT from [14], where, after optimization, the maximum values for S_{11} and S_{22} are 0.19 and 0.17, respectively, and S_{21} varies from 9.8 to 10.1 dB.

CIAO Program Output for Circuit of Fig. 1

Initial Analysis

Freq. (Hz)	S11 Mag.	S11 Ang.	S12 Mag.	S12 Ang.	S21 Mag.	S21 Ang.	S22 Mag.	S22 Ang.	S21 dB	K Fact
1.000E+007	0.08	17	0.184	-0.2	2.684	175.4	0.04	138	8.6	1.2
1.000E+008	0.08	11	0.182	-15.5	2.665	163.0	0.08	103	8.5	1.3
2.500E+008	0.10	5	0.178	-41.0	2.711	140.1	0.15	78	8.7	1.2
5.000E+008	0.22	-6	0.168	-83.7	2.954	100.4	0.19	63	9.4	1.1
7.500E+008	0.44	-34	0.146	-130.5	3.200	58.7	0.23	83	10.1	1.1
1.000E+009	0.63	-71	0.119	-178.1	3.260	11.7	0.37	83	10.3	1.1
1.250E+009	0.49	-109	0.104	125.0	3.528	-51.7	0.36	94	11.0	1.1
1.500E+009	0.63	-67	0.047	42.3	1.999	-144.1	0.87	89	6.0	1.1

Iteration number: 0 - using 1 function evaluation(s).
Error function: 0.582

Magnitude errors [11,12,21,22]: 1.61E-001 0.00E+000 2.77E-001 1.43E-001

Variable	Value	Gradient
1	3.00000E-012	3.04153E-001
2	1.00000E-008	1.00701E+000
3	2.00000E+002	-6.6843E-001
4	2.80000E-008	4.56586E-001
5	3.00000E-012	9.37410E-001
6	1.00000E-008	6.42941E-001

Iteration number: 1 - using 4 function evaluation(s).
Error function: 0.172

Magnitude errors [11,12,21,22]: 6.94E-002 0.00E+000 5.24E-002 5.00E-002

Variable	Value	Gradient
1	2.72409E-012	-1.7291E-002
2	7.26567E-009	-5.6132E-002
3	2.57125E+002	-1.0394E-001
4	2.42247E-008	3.20011E-001
5	2.22836E-012	-8.3608E-003
6	8.15510E-009	2.77998E-002

Iteration number: 2 - using 3 function evaluation(s).
Error function: 0.155

Magnitude errors [11,12,21,22]: 7.06E-002 0.00E+000 3.83E-002 4.61E-002

Variable	Value	Gradient
1	2.73287E-012	2.80956E-002
2	7.33876E-009	-3.0728E-002
3	2.73082E+002	1.49098E-001
4	2.28093E-008	2.22470E-001
5	2.19309E-012	4.41039E-002
6	7.91213E-009	5.74532E-002

CIAO Program Output for Circuit of Fig. 1 (cont'd)

Iteration number: 9 - using 2 function evaluation(s).
Error function: 2.64E-002

Magnitude errors [11,12,21,22]: 1.74E-002 0.00E+000 2.67E-003 6.34E-003

Variable	Value	Gradient
1	1.42188E-012	4.53870E-003
2	5.03307E-009	-2.8783E-003
3	2.71317E+002	5.06559E-002
4	8.79645E-009	8.55861E-003
5	1.59959E-012	-4.1153E-004
6	4.04664E-009	-3.1482E-003

Iteration number: 10 - using 2 function evaluation(s).
Error function: 2.59E-002

Magnitude errors [11,12,21,22]: 1.71E-002 0.00E+000 3.12E-003 5.72E-003

Variable	Value	Gradient
1	1.36017E-012	2.01228E-004
2	4.94833E-009	6.72160E-004
3	2.68690E+002	-3.7395E-004
4	8.48246E-009	2.38870E-003
5	1.57179E-012	-6.1062E-003
6	4.08098E-009	-4.3751E-004

Final Analysis

Freq. (Hz)	S11 Mag.	S11 Ang.	S12 Mag.	S12 Ang.	S21 Mag.	S21 Ang.	S22 Mag.	S22 Ang.	S21 dB	K Fact
1.000E+007	0.18	5	0.162	0.5	3.107	175.7	0.06	21	9.8	1.2
1.000E+008	0.17	-8	0.161	-7.4	3.072	166.9	0.07	26	9.7	1.2
2.500E+008	0.16	-32	0.158	-20.5	3.060	150.4	0.08	31	9.7	1.3
5.000E+008	0.13	-69	0.151	-41.4	3.122	123.6	0.11	28	9.9	1.3
7.500E+008	0.09	-126	0.139	-63.1	3.175	100.0	0.08	17	10.0	1.3
1.000E+009	0.07	157	0.128	-84.1	3.189	75.1	0.03	4	10.1	1.4
1.250E+009	0.08	76	0.120	-105.7	3.195	44.2	0.03	-111	10.1	1.5
1.500E+009	0.13	37	0.112	-129.6	3.168	9.6	0.11	-153	10.0	1.5

It should be noted that SUPER-COMPACT used both random-grid and gradient optimization, with a total of at least 37 error-function evaluations (21 function evaluations for the grid search plus allowing 2 function evaluations for each of 8 gradient iterations), while CIAO was able to achieve its results with gradient optimization alone, using only 23 function evaluations over its 10 iterations.

CIAO's execution-time (including printout) for the 10 iterations was approximately 4 minutes on a Corona 16-bit, MS-DOS microcomputer equipped with an 8087 co-processor chip. Each of the 23 function evaluations corresponds to an analysis of the network, as well as derivative and error-function calculations at 8 frequency points. A simple calculation yields CIAO's "optimization speed" in this example of about 1 frequency point per 1.3 seconds. (CIAO's speed in analyzing the network - i.e. without derivative or error-function calculations - is about 1 frequency point per 0.65 seconds.) Without the 8087 co-processor chip, we found that the run-times increased by a factor of about 4.5 . The run-times on a Kaypro II 8-bit, CP/M-80 microcomputer were found to be just about the same as the times on the Corona without the 8087 chip.

3.2 A DESIGN Sample Session

3.2.1 Statement of the Synthesis Problem

We present here a realistic matching network synthesis problem that can routinely be handled by the DESIGN program. The specific task is stated as follows:

> The input reflection coefficient, S_{11}, of a Plessey GAT6 GaAs FET chip (common-source configuration) is shown in the table below. A network is to be designed to match this S_{11}, as a complex load, to a 50 ohm source across the given frequency band. The matching network must provide the indicated gain-sloped S_{21} magnitude response between the real source and the complex load.
>
> Both a lumped and a distributed version of the network is required.

FREQ. (GHz)	COMPLEX LOAD S11 (mag)	S11 (phase-deg)	GAIN SLOPE (Desired S21 of matching network)
8.0	0.775	-107	0.732
9.0	0.750	-118	0.809
10.0	0.730	-128	0.875
11.0	0.710	-136	0.943
12.0	0.695	-145	1.000

PROBLEM: Match the complex load to 50 ohms with the prescribed gain slope.

NOTES: 1. The complex load (S11 of the GAT6 chip) is defined with respect to a 50 ohm reference.

2. The desired S21 of the matching network is defined with respect to the prescribed 50 ohm source and the indicated S11 complex load.

Note that prescribing the S_{21} magnitude response of the matching network also determines the magnitudes of the input and output reflection coefficients of the network, as discussed in Section 1.4.1 .

3.2.2 Data Input

The data for DESIGN is entered interactively at the keyboard. The transcript of the data entry for the specific problem at hand is given the next two pages. The data required by DESIGN is mostly self-explanatory, except for a few items that we now discuss.

<u>The Degree</u>.The user must select a degree for the matching network. This may require some trial-and-error to achieve the most efficient design.

<u>The Match at DC</u>. The program asks whether the "DC input resistance ... be kept fixed at the DC source resistance value". If the frequency band for the match starts at DC, then the appropriate reply would be "Y". Otherwise, if the frequency band is of the bandpass variety, "N" is the better answer, since there is usually no need to enforce impedance matching outside the specified frequency band. However, on rare occasion, it may be necessary to require matching at DC - even if the desired response is bandpass - in order to achieve a satisfactory design.

<u>Bandpass or Lowpass Response</u>. The program indicates that either a lowpass or

Interactive Input for DESIGN

(User Response Underlined)

```
Enter S for screen output or P for printer output: s

Enter source resistance: 50

Source resistance read as:   50.00  -- OK?  (Y/N): y


Enter number of frequency points (max. 25) for load:  5

Number of frequencies read as:  5  -- OK?  (Y/N): y


How is load described?
(R)eflection coefficient, (I)mpedance, or (A)dmittance - (R/I/A): r

Load description read as: R  -- OK?  (Y/N): y


What is reference resistance for reflection coefficient data? 50

Reference resistance read as:   50.00  -- OK?  (Y/N): y


ENTER>>     Freq._(Hz)   Mag.   Phase_(deg.)    S21      <<AT FREQ. #  1 :
8e9 .775 -107 .732

Parameters read as:
8.000E+009    7.750E-001    -1.07E+002    0.7320  -- OK?  (Y/N): y


ENTER>>     Freq._(Hz)   Mag.   Phase_(deg.)    S21      <<AT FREQ. #  2 :
9e9 .75 -118 .809

Parameters read as:
9.000E+009    7.500E-001    -1.18E+002    0.8090  -- OK?  (Y/N): y

ENTER>>     Freq._(Hz)   Mag.   Phase_(deg.)    S21      <<AT FREQ. #  3 :
10e9 .73 -128 .875

Parameters read as:
1.000E+010    7.300E-001    -1.28E+002    0.8750  -- OK?  (Y/N): y

ENTER>>     Freq._(Hz)   Mag.   Phase_(deg.)    S21      <<AT FREQ. #  4 :
11e9 .71 -136 .943

Parameters read as:
1.100E+010    7.100E-001    -1.36E+002    0.9430  -- OK?  (Y/N): y
```

Interactive Input for DESIGN (cont.)

```
ENTER>>     Freq._(Hz)   Mag.   Phase_(deg.)    S21      <<AT FREQ. #  5 :
12e9 .695 -145 1

Parameters read as:
1.200E+010    6.950E-001    -1.45E+002    1.0000  -- OK?  (Y/N): y


Enter desired degree (2-12) of matching network: 4

Degree read as:  4  -- OK?  (Y/N): y


Shall the DC input resistance to the matching network
be kept fixed at the DC source resistance value?   (Y/N): n

Keeping DC input resistance fixed read as: N  -- OK?  (Y/N): y


A LOWPASS, shunt-C series-L network normally results from the synthesis.
A BANDPASS structure can be realized by cascading a single shunt inductor
across the complex-load end of the matching network.

For your load data, either a LOWPASS or BANDPASS structure is possible.

Do you want a (B)andpass or (L)owpass matching network structure?  (B/L): b


What value inductance (in Henries) should be placed across the complex load?

(You might try  2.193E-010 H, this value changes the sign of
the midband load susceptance.)

                      Enter value or X to switch to LP : 5e-10

Inductance value read as: 5.000E-010  -- OK?  (Y/N): y


Shall the default number of break frequencies be used?  (Y/N): y

Using default number of break frequencies  -- OK?  (Y/N): y


The "CURVE-FIT FACTOR" affects the relative curve-fitting accuracy between
the passband and stopband.  The default value of 1.0 is usually a good
compromise;  a smaller value weights the passband more, a larger value
emphasizes the stopband more.

    Shall the default value be used for the curve-fit factor?  (Y/N): y


Using default value for curve-fit factor  -- OK?  (Y/N): y
```

bandpass matching-network structure is possible. Then, the program indicates which choice will most likely lead to a successful synthesis by DESIGN. In the present case, DESIGN says that either a lowpass or bandpass topology would be suitable. The program offers this suggestion based upon the location of the load data on the Smith Chart, and upon the capabilities of the Carlin-Komiak synthesis algorithm. We choose the bandpass option here. The program then suggests a possible value for the added inductor; we select a somewhat larger value, as shown in the data input. Note that in the present example we initially specify a matching network degree of 4, but that selecting a bandpass topology actually increases the degree of the network by one to 5.

Number of Break Frequencies. The program asks whether the "default number of break frequencies be used". In most broadband cases, the default value will suffice. However, for narrow-band matching, it may occasionally be necessary to override the default to a larger value to achieve a satisfactory design.

Curve-Fit Factor. Finally, the program asks whether the default value for the "curve-fit factor" should be used. This parameter, which may be adjusted between 0.2 and 1.5, varies the curve-fit weighting used in approximating the optimized piecewise-linear real-part function R(f). The default value of 1.0 distributes the curve-fit points in a proportional way between the passband and stopband of the matching network; in most instances, this is adequate. In cases of convergence to non-physically realizable element values, changing the default value can often be beneficial.

3.2.3 The Resistance-Excursion Optimization and Curve-Fitting

After the data input in completed, a summary of the (normalized) load and frequency data is output, and then the two-part design process begins. The first part involves finding the optimum piecewise-linear real-part function R(f), and the second part involves fitting a rational polynomial to R(f) and then decomposing this polynomial in the proper way to find the actual lumped and distributed matching networks. The DESIGN program output for the current example is shown on the next three pages.

DESIGN Program Output

```
Degree of network to be designed:  4
Source resistance (ohms): 5.00E+001

Note: Units of frequency in program output are rad/sec.
      Frequencies are normalized so that the last load freq. = 1 rad/sec.
      Frequency scale factor used (rad/sec): 7.54E+010
      Immittances are normalized so that source resistance becomes 1 ohm.

Normalized Load Frequencies,       Re(Yload),  Im(Yload);  Desired |S21| are:
 0.667                             3.48E-001   -7.0E-001          0.732
 0.750                             5.10E-001   -2.3E-001          0.809
 0.833                             7.37E-001   2.23E-001          0.875
 0.917                             1.03E+000   5.97E-001          0.943
 1.000                             1.50E+000   9.39E-001          1.000

Note: load impedance values include effect of inductor of value  5.00E-010 H
shunted across original load.


RESISTANCE EXCURSION OPTIMIZATION


Iteration number:   0       - using   1 function evaluation(s).
RMS Percentage ERROR of  Transducer Power Gain  across passband:  26.768

Variable      Norm. Break-Freq.     RESISTANCE VALUE      Gradient
 1            0.0000                 1.0000               -2.90E-001
 2            0.3725                 0.7500               -3.02E-001
 3            0.7450                 0.5000               -3.33E-001
 4            1.1174                 0.2500               -1.48E-001


Iteration number:   1       - using   5 function evaluation(s).
RMS Percentage ERROR of  Transducer Power Gain  across passband:  13.630

Variable      Norm. Break-Freq.     RESISTANCE VALUE      Gradient
 1            0.0000                 1.1393               -2.36E-002
 2            0.3725                 1.1377               -2.92E-002
 3            0.7450                 1.1056               -3.48E-002
 4            1.1174                 0.9528               5.390E-003


Iteration number:   2       - using   3 function evaluation(s).
RMS Percentage ERROR of  Transducer Power Gain  across passband:   9.572

Variable      Norm. Break-Freq.     RESISTANCE VALUE      Gradient
 1            0.0000                 1.3647               1.146E-002
 2            0.3725                 1.4910               6.601E-003
 3            0.7450                 1.6707               -1.32E-002
 4            1.1174                 1.4957               -7.39E-003
```

DESIGN Program Output (cont'd)

**Convergence Achieved, final summary follows.

Iteration number: 12 - using 2 function evaluation(s).
RMS Percentage ERROR of Transducer Power Gain across passband: 1.815

Variable	Norm. Break-Freq.	RESISTANCE VALUE	Gradient
1	0.0000	0.9012	-1.72E-004
2	0.3725	0.8205	-1.66E-004
3	0.7450	1.9954	-1.72E-004
4	1.1174	2.0637	-3.44E-005

CONSTRUCTION OF THE IMPEDANCE Z(s)

The even part of the Z(s) of the matching network looking back from the complex load (with the network terminated in the real source impedance) is given by Ev[Z(s)] = a0 / (1 + b2*s^2 + b4*s^4 + b6*s^6 + ...)

By curve-fitting, the coefficients of Ev[Z(s)] are obtained:

a0 = 6.564E-001
b2 = 2.713E+000
b4 = 3.947E+000
b6 = 2.508E+000
b8 = 6.007E-001

The left-half s-plane roots of the denominator of Ev[Z(s)] are:

Root 1:	Real Part = -3.79E-001	Imaginary Part = -7.65E-001
Root 2:	Real Part = -2.48E-001	Imaginary Part = 1.307E+000
Root 3:	Real Part = -3.79E-001	Imaginary Part = 7.654E-001
Root 4:	Real Part = -2.48E-001	Imaginary Part = -1.31E+000

The transducer function H(s) of the matching network, with port 1 at the real source and port 2 at the complex load, is constructed from the LHP roots of Ev[Z(s)]. The result is H(s) = A0 + A1*s + A2*s^2 + A3*s^3 + ... where:

A0 = 1.290E+000
A1 = 1.703E+000
A2 = 2.874E+000
A3 = 1.254E+000
A4 = 1.000E+000

DESIGN Program Output (cont'd)

FINAL RESULTS: MATCHING NETWORK ELEMENT VALUES FROM REAL-SOURCE SIDE TO COMPLEX-LOAD SIDE.

(All values in Henries or Farads, unnormalized.)

Shunt Capacitor = 1.389E-013

Series Inductor = 8.356E-010

Shunt Capacitor = 4.154E-013

Series Inductor = 4.975E-010

The last element is the Shunt Inductor added at the start of the synthesis, of value 5.000E-010 Henries.

A corresponding DISTRIBUTED EQUIVALENT CIRCUIT for the matching network is:

OST, Zo = 25.0 ohms, length = 12.06 degrees at 9.798E+009 Hz

TRL, Zo = 120.0 ohms, length = 31.67 degrees at 1.200E+010 Hz

OST, Zo = 25.0 ohms, length = 32.59 degrees at 9.798E+009 Hz

TRL, Zo = 120.0 ohms, length = 18.21 degrees at 1.200E+010 Hz

SST, Zo = 120.0 ohms, length = 14.39 degrees at 9.798E+009 Hz

Optimization. At each iteration in the resistance-excursion optimization, a description of the normalized function R(f) is given, and more importantly, the percentage error in the transducer gain (i.e. the square of the magnitude of S_{21}) is given for the matching network that would result if R(f) could be realized exactly. In the present example, the optimization was able to reduce the error to 1.8%. The program user is able to interrupt the optimization by holding down any key; the program has internal features to stop the optimization as well. The optimization is numerically well-conditioned, and therefore almost always succeeds in reducing the error to a small value.

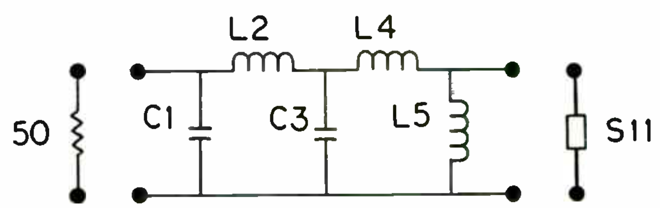
Curve-Fitting. After the optimal, piecewise-linear, resistance function R(f) is found, the (normalized) function Ev[Z(s)] is obtained as a rational-polynomial, curve-fitting approximation to R(f). A successful curve-fit causes the left-hand plane roots of Ev[Z(s)] to appear on the real axis, or, to appear as complex conjugates, with no roots on the imaginary axis. Then the function is decomposed and denormalized to yield the actual element values of the desired matching network (both lumped and distributed versions), as shown in the program output.

3.2.4 The Results of the Example

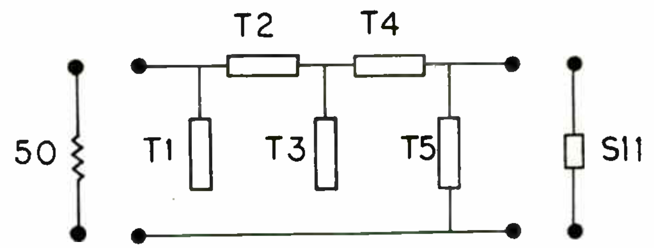
Fig. 2 shows the circuits which result from the design requirements of the present example. The results of a CIAO analysis of these networks are shown in Fig. 3, which illustrates that DESIGN has provided a reasonable approximation to the desired S_{21} magnitude response for each matching network operating between the real source and the complex load. For many cases these results would be adequate. For precision applications, a better match may be required, so we used CIAO to vary the elements of both the lumped and distributed versions of the matching network to obtain improvement. The results were excellent, as indicated in Fig. 3. After the CIAO optimization, the error between the desired and calculated S_{21} values for the matching networks was reduced, essentially, to zero. The CIAO data files to achieve this optimization, as well as the optimized element values for the networks, are shown following Fig. 3.

It should be noted that DESIGN alone can often produce matching networks that meet requirements with great accuracy. In other cases, DESIGN gets close enough to the requirements that CIAO can then be used to reduce any remaining error to negligible values. Convergence with DESIGN occurs in the large majority of cases, although certain problematic data may require resetting some of the DESIGN's default data values.

Fig. 2: Lumped and Distributed Matching Networks from DESIGN

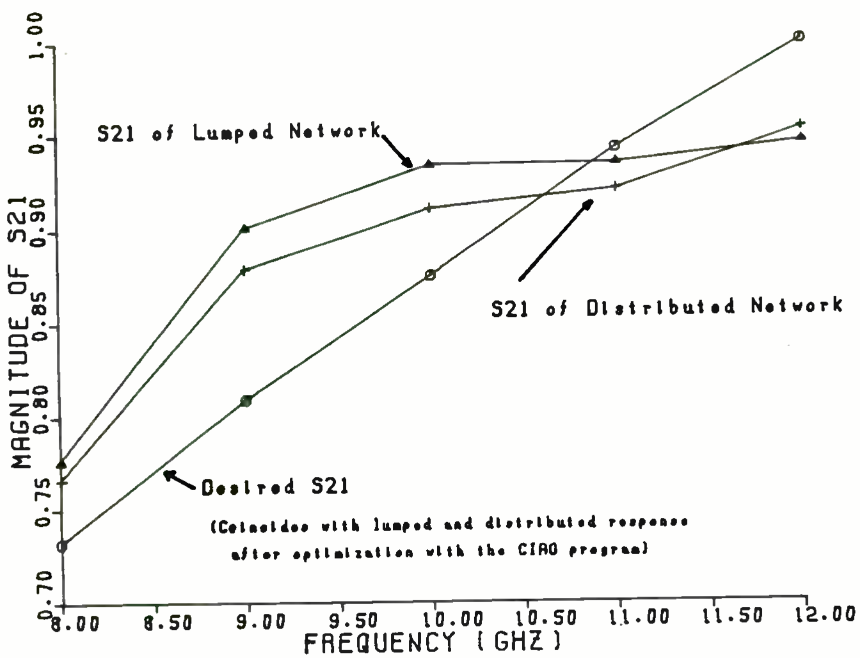


C1: 0.139 pF
L2: 0.836 nH
C3: 0.415 pF
L4: 0.498 nH
L5: 0.500 nH



T1:	Open-Ct Stub,	Zo= 25 ohms,	length=12.1 deg	at	9.80 GHz
T2:	Trans Line,	Zo=120 ohms,	length=31.7 deg	at	12.0 GHz
T3:	Open-Ct Stub,	Zo= 25 ohms,	length=32.6 deg	at	9.80 GHz
T4:	Trans Line,	Zo=120 ohms,	length=18.2 deg	at	12.0 GHz
T5:	Short-Ct Stub,	Zo=120 ohms,	length=14.4 deg	at	9.80 GHz

Fig. 3: LUMPED AND DISTRIBUTED MATCHING NETWORK SYNTHESIS FROM THE DESIGN PROGRAM



CIAO Data File for Optimization of LUMPED-ELEMENT Matching Network

```
'4th order bandpass + load inductor

cap#  1 0  .1389e-12
ind#  1 2  .8356e-9
cap#  2 0  .4154e-12
ind#  2 3  .4975e-9
ind#  3 0  .5e-9
perform twoport optimization table1
port1  1 0  50
port2  3 0  table5
end

8e9 12e9 1e9   \   0  0  1/0  0
end

table5 Sparameters
50  (char. impedance)
.775 -107
.75  -118
.73  -128
.71  -136
.695 -145

table1 Sparam
'(no char impedance needed for opt. table)
0 0   0 0   0.732 0   0 0
0 0   0 0   0.809 0   0 0
0 0   0 0   0.875 0   0 0
0 0   0 0   0.943 0   0 0
0 0   0 0   1.0   0   0 0

end
```

OPTIMIZED ELEMENT VALUES

Variable	Value
- 1	1.23980E-013
2	6.70662E-010
3	5.17870E-013
4	5.60760E-010
- 5	7.19564E-010

CIAO Data File for Optimization of DISTRIBUTED-ELEMENT Matching Network

```
'4th order bandpass plus load inductor

'(_* means optimize length but keep Zo constant)
'(*_ would mean optimize Zo but keep length constant)
'(** would mean optimize both Zo and length)

'(Note:  default length type is electrical degrees)
                                                     OPTIMIZED LINE LENGTHS
'                Zo  length  @freq.
'                --  ------  -----               Variable       Value
ost_* 1 0        25  12.06    9.798e9  ------->     1           1.13861E+001
trl_* 1 0 2 0   120  31.67   12.000e9               2           2.24814E+001
ost_* 2 0        25  32.59    9.798e9               3           3.64162E+001
trl_* 2 0 3 0   120  18.21   12.000e9               4           1.86120E+001
sst_* 3 0       120  14.39    9.798e9  ------->     5           2.09042E+001

perform twoport optimization table1
port1  1 0  50
port2  3 0  table5
end
8e9 12e9 1e9   \   0  0  1/0  0
end

table5 Sparameters
50 (char. impedance)
.775 -107
.75  -118
.73  -128
.71  -136
.695 -145

table1 Sparam
'(no char impedance needed for opt. table)
0 0   0 0   0.732 0   0 0
0 0   0 0   0.809 0   0 0
0 0   0 0   0.875 0   0 0
0 0   0 0   0.943 0   0 0
0 0   0 0   1.0   0   0 0

end
```

3.2.5 DESIGN's Execution Speed

DESIGN's execution speed is quite rapid. On a Corona 16-bit MS-DOS microcomputer equipped with an 8087 coprocessor, the output for the example discussed above was virtually instantaneous. Without the coprocessor, or on a Kaypro II 8-bit CP/M-80 machine, each iteration of the resistance optimization took a few seconds, with a somewhat longer delay for the construction of the matrix needed in the curve-fitting process.

CONCLUSION

4.1 Conclusion

Only some of the capabilities of CIAO and DESIGN have been described herein. The User's Guide for CIAO and DESIGN contains many more involved examples illustrating the power of these programs. CIAO and DESIGN have been expertly developed and thoroughly tested, and represent, we believe, the best performance to price ratio of any software of its type in the industry.

4.2 PRICING and AVAILABILITY of the CIAO and DESIGN PROGRAMS

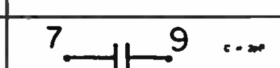
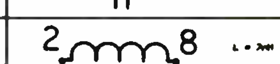
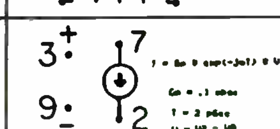
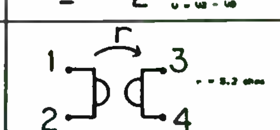
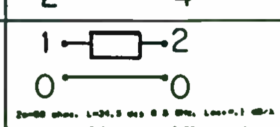
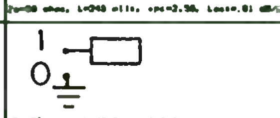
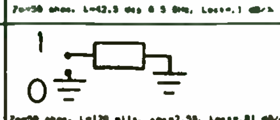
The CIAO and DESIGN programs for personal computers are currently available for immediate purchase, at a price of $1900.00 each.

Versions of CIAO are available for IBM-PCs and compatibles with at least 256K RAM, with or without the 8087 co-processor chip. (A version can be obtained that will run, with overlays, in as little as 128K of RAM.) DESIGN is available for IBM-PCs and compatibles with a minimum of 128K of RAM, again with or without 8087 support. Full-featured versions of CIAO and DESIGN are also available for CP/M-80, 8-bit, 64K RAM microcomputers.

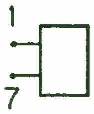
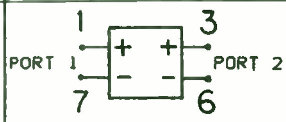
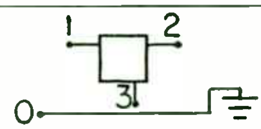
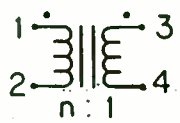
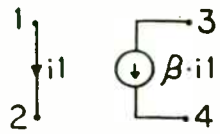
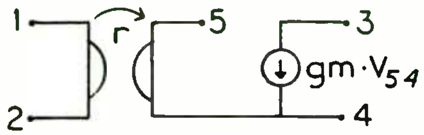
The source codes for both CIAO and DESIGN are available at an additional charge. Please contact the author for further details.

Appendix A

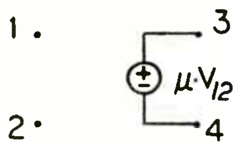
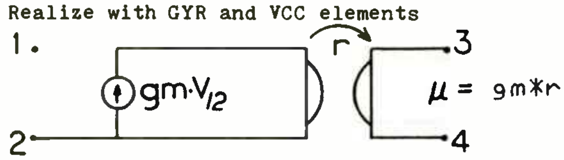
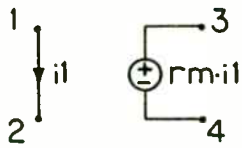
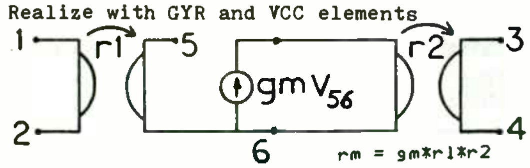
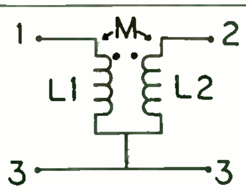
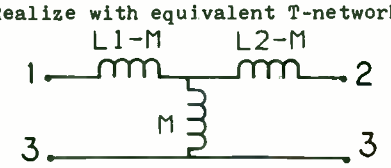
ELEMENTS of a CIAO DATA FILE (Chart)

Circuit Element Summary				
Element Type	**Code**	**Schematic Example**	**Data Example**	**Comments**
RESISTOR	RES		RES 1 2 150	RES* for optimization
CAPACITOR	CAP		CAP 7 9 2e-12	CAP* for optimization
INDUCTOR	IND		IND 2 8 7e-9	IND* for optimization
VOLTAGE-CONTROLLED CURRENT SOURCE	VCC		VCC 3 9 7 2 .1 2e-12	τ defaults to zero if not specified. VCC*, VCC_*, and VCC** optimizes gm, τ, and both gm and τ, resp.
GYRATOR	GYR		GYR 1 2 3 4 5.2	GYR* for optimization
TWO-PORT TRANSMISSION LINE	TRL		TRL 1 0 2 0 50 34.5deg 5e9 .1 or TRL 1 0 2 0 50 245mils 2.55 .01 Loss defaults to zero if not specified	
OPEN-CIRCUIT STUB LINE	OST		OST 1 0 50 42.5deg 5e9 .1 or OST 1 0 50 120mils 2.55 .01	
SHORT-CIRCUIT STUB LINE	SST		SST ... format same as for OST For all transmission lines: "*" optimizes Zo, "_*" optimizes length, "**" optimizes both Zo & length: e.g. SST*, OST_*, TRL** .	

Circuit Element Summary

Element Type	Code	Schematic Example	Data Example	Comments
ONE-PORT ELEMENT	ONE		ONE 1 7 Table5	References either an S, Y, or Z table
TWO-PORT ELEMENT	TWO		TWO 1 7 3 6 Table3	A true 4-terminal netwk References either an S, Y, or Z table
THREE-PORT ELEMENT	THR		THR 1 2 3 Table2	Common terminal 3-port References an S-param. table
IDEAL TRANSFORMER	-		Realize as equivalent cascaded gyrators  r1/r2 = n	
CURRENT-CONTROLLED CURRENT SOURCE	-		Realize with GYR and VCC elements  gm*r = β	

Circuit Element Summary

Element Type	Code	Schematic Example	Data Example	Comments
VOLTAGE-CONTROLLED VOLTAGE SOURCE	-		Realize with GYR and VCC elements  μ = gm*r	
CURRENT-CONTROLLED VOLTAGE SOURCE	-		Realize with GYR and VCC elements  rm = gm*r1*r2	
MUTUAL INDUCTANCE	-		Realize with equivalent T-network 	

Appendix B

NOTES and REFERENCES

1. IBM is a registered trademark of International Business Machines, Inc.

2. MS-DOS is a registered trademark of the Microsoft Corp.

3. CP/M is a registered trademark of Digital Research, Inc.

4. COMPACT and SUPER-COMPACT are registered trademarks of COMPACT Software, Inc.

5. Documentation of these phenomena are available upon written request to SPEFCO Software, 18 Bennett Lane, Stony Brook, NY 11790.

6. R. Fletcher and C. M. Reeves, "Function Minimization by Conjugate Gradients," Computer Journal, vol. 7, pp. 149-154, 1964.

7. S. W. Director and R. A. Rohrer, "The Generalized Adjoint Network and Network Sensitivities," IEEE Trans. Circuit Theory, vol. CT-16, pp. 318-323, 1969.

8. H. J. Carlin and J. J. Komiak, "A New Method of Broad-Band Equalization Applied to Microwave Amplifiers," IEEE Trans. Microwave Theory ant Techniques, vol. MTT-27, no. 2, pp. 93-99, Feb. 1979.

9. H. J. Carlin and P. Amstutz, "On Optimum Broad-Band Matching," IEEE Trans. Circuits and Systems, vol. CAS-28, no. 5, pp. 401-405, May 1981.

10. R. M. Fano, "Theoretical Limitations on the Broadband Matching of Arbitrary Impedances," J. Franklin Institute, vol. 249, pp. 52-83, Jan. 1950; vol. 249, pp.129-155, Feb. 1950.

11. D. C. Youla, "A New Theory of Broadband Matching," IEEE Trans. Circuit Theory, vol. CT-11, pp. 30-50, Mar. 1964.

12. W. K. Chen, "Synthesis of Optimum Butterworth and Chebyshev Broadband Impedance Matching Networks," IEEE Trans. Circuits and Systems, vol. CAS-24, pp.152-169, Apr. 1977.

13. D. J. Mellor and J. G. Linvill, "Synthesis of Interstage Networks of Prescribed Gain Versus Frequency Slopes," IEEE Trans. Microwave Theory and Techniques. vol. MTT-23, no. 12, pp. 1013-1020, Dec. 1975.

14. G. Gonzalez, Microwave Transistor Amplifiers, Prentice-Hall, Englewood Cliffs, N.J., 1984 (chap. 4 and App. A).

15. L. Besser, "Microwave Circuit Design," Electronic Engineering, Oct. 1980.

MATCHING NETWORK DESIGN USING

HP-41, HP-71 AND HP-75 COMPUTERS

Irving L. Weiner

Ortho Medical Services

Westwood, MA 02090

January 1985

MATCHING NETWORK DESIGN USING
HP-41, HP-71 AND HP-75 COMPUTERS

The Smith Chart is most probably the oldest surviving 'slide rule' still in common use. This graphical aid has served a countless number of RF engineers in the design of impedance matching networks for antennas, interstage couplers, amplifiers and transmission lines. This paper deals primarily with the use of hand held and portable (now lap held!) calculator/computers of the Hewlett-Packard 40 and 70 Series. The computational power of these machines with their specialized plug-in modules (MATH, STATISTICS, CURVE FIT, AC ANALYSIS, FORTH, FINANCE, PLOTTER, SURVEYING, TIMER, I/O and even VISICALC to name just a few pale the capabilities of the desk top computer of a decade ago; and those of us who started out with vacuum tube design fondly remember our retired sliderules as venerable antiques.

The aim of this paper is the introduce the basic features of the Smith Chart applicable to the design of impedance matching networks for newcomers to the RF domain.

SMITH CHART

The Smith Chart is basically a plot of the reflection coefficient ρ ; where $\rho = \frac{Zn-1}{Zn+1}$.

Zn is the normalized impedance, Z/Zo where Z is the load network impedance and Zo is the characteristic impedance represented by the scaled value at the center of the Smith Chart (1,0). The ability to normalize the working impedances permit the Smith Chart to scale any range of actual network impedances encountered. The range of values for ρ (utilizing passive networks only) lie between 0 and 1.

If we now write $Zn = Rn+jXn$
and $\rho = U+jV$

we have: $U+jV = \frac{Rn+jXn-1}{Rn+jXn+1}$

After some additional algebra we get: (elimination of Xn)

$$\left[U-\frac{Rn}{Rn+1}\right]^2 + V^2 = \frac{1}{(Rn+1)^2}$$

On rectangular coordinates of U and V this is the equation of a circle whose center, for any value of Rn is located at U = Rn/(Rn+1), V=0 and whose radius is 1/(Rn+1). If we repeat the above, eliminating Rn instead of Xn, the following equation

$$(U-1)^2 + (V-1/Xn)^2 = (1/Xn)^2$$

is the locus of the circle defining constant values for Xn .
In Figs. 1 and 2 we can see the construction for the family of curves that make up the Smith Chart. Note that the outer boundary for the Smith Chart can be contained within a square whose side is

unity. This will become important later when we discuss the plotter setup.

In the U, V rectangular coordinate system the center of the Smith Chart is (0,0) and the "compass" points are (North first) (0,1), (1,0), (0,-1) and (-1,0). If we now solve for Rn = f (U,V), Xn = f (U,V), U = f (Rn,Xn) and V = f (Rn,Xn), the resulting equations give us a transform between a given Zn = Rn + j Xn and ρ = U + j V. This permits us to map (in actuality, to plot) Rn, Xn in U, V; operate on U, V and return the new values of Rn, Xn.

a) $Rn = (1-U^2 - V^2) / [V^2 + (U-1)^2]$

b) $Xn = 2V / [V^2 + (U-1)^2]$

c) $U = (Rn^2 + Xn^2 - 1) / [Xn^2 + (Rn + 1)^2]$

d) $V = 2Xn / [Xn^2 + (Rn + 1)^2]$

The absolute magnitude of the reflection coefficient may be used to calculate the transmission efficiency of power; i.e., % Power Reflected = $|\rho|$ x 100. The VSWR of "voltage standing wave ratio" is an indicator of this power loss since VSWR = $(1+|\rho|)/(1-|\rho|)$. The locus of the magnitude of the reflection coefficient is a circle whose center is at the origin of the Smith Chart (Zn,0), the larger the VSWR or diameter of such circle the greater the power loss. It is the goal of impedance matching to 'operate' on the existing network such that the resulting network has an optimized minimum VSWR or for special needs a prescribed VSWR over the frequency range. The 'operations' referred to involve the addition of reactive and resistive elements placed in series or shunt within the network. The reactive elements employed can be capacitors, inductors, sections of transmission line or combinations of such. Again, the goal is to manipulate the original network locus to an optimized or prescribed locus. This is the function of the computer calculator program.

THE PROGRAM

The Program contained in this paper permits the designer to graphically display the impedance or admittance data versus frequency for a given network. The data display can be to a CRT monitor, printer or both. The graphic display is to a plotter; the programs included permit output to a HP7470, 7475, Radio Shack CGP - 115, or to a standard analog plotter using the HP-82166, 82165 IL Converter and a peripheral pair of D-A converters. The actual plotting can be done directly on pre-printed Smith Charts or on sheets of plain paper used in conjunction with Smith Chart overlays. This is a very convenient approach when many trials are to be made and the line work starts to get messy. The primary effort will utilize the 7470 due to two of its outstanding features; the ability to digitize and change scale under program control. The ability to change scale under program control permits the user to expand or 'zoom' into a region of interest for greater

accuracy. The user can prepare plastic overlays for this region using the defining equations of the Smith Chart; this technique in conjunction with plain paper or Xeroxes of the overlays can yield very precise graphic analysis. The second feature, digitizing, permits the extraction of data points from a Smith Chart generated by a network impedance analyzer or from the technical literature. The use of a two dimensional Cubic Spline interpolation program (HP41 user Library Program #1435C) can add additional data points to the data set.

Upon executing the program the user is reminded to reset the scale points P1, P2; this is done only if Xerox chart copies of a new reduction are used or if the user has chosen to alter the chart scale factor. The next two prompts requests the number of data points and the need to create new files. If the number of data points is increased it will be necessary to create new files; but not if you are modifying the values of the present data set. The two following prompts requests the device to indicate data output; D1 is HP Video Interface (82169A or 82163) connected to a Video monitor, P1 is the HP82162 thermal printer. Selecting the AutoPlot Mode produces a plot and data printout automatically after initializing and each component trial. The alternate choice permits manual selection of plotting and printing. Next, the user is prompted to input each point in the data set as follows: Frequency, Real, Reactive.

After entry of the data set and assuming manual mode the user is prompted to make a menu selection:

Next (1) - select next component for trial.

Init (2) - initialize to original data set (or start again)

Print (3) - Print original data or results

Plot (4) - Plot original data or results

If Next is chosen you will be prompted 'Delete Previous Element'; this permits the user to negate or accept a previous component trial.

The subsequent menu driven prompts are self-explanatory; after entering the 'next' component information the computer will be "Working" to get the results. The components presently available for use are series and shunt Resistors, Capacitors, Inductors. The use of a series transmission line Xfmr is available from the main menu and a scaling transformer is available from the utilities file. In addition, the utilities files contain routines for digitizing, printing and/or plotting the admittance data during the session. This is strictly a convenience feature for conceptualizing the selection of the "next" element and does not at this time place the U-V plane into an 'admittance' mode. The routine permitting plots of selected VSWR Circle is also present in the utilities file and is invaluable when plotting on plain paper. The algorithm for handling series elements (except for the Line Xfmr) is simple:

$$Znew = Rold + Xold + Rnext \text{ or } \pm j\, Xnext$$

Computation is made in the R-Z plane and converted to the U-V plane for plotting by the Init-Result routine.

The algorithm for handling shunt elements is similar:

$$Rold + j\, Xold \rightarrow Gold \pm j\, Sold = Yold$$

$$Rnext + j\, Xnext \rightarrow Gnext \pm j\, Snext = Ynext$$

$$Ynext = Yold + Ynext$$

The R,X data is transformed from the impedance plane to the admittance plane and summed by parts in the R,X plane, As before, the Init-Result routine effects conversion to the U-V plane.

The use of a series line transformer consisting of a section of transmission line effectively rotates each point in the data set about the origin of the U-V plane by an angular displacement that is proportional to frequency. The U-V data set stored in file 2 is operated on directly by the classical algebraic equations for rotation-translation and read back into file 2. The U-V data is simultaneously converted into the R, X plane for printout.

Additional network elements to be incorporated include shunt transmission lines or stubs. The ability for the program to handle double stubs with variable lengths and spacing would be an invaluable design aid. An additional utility not incorporated due to the pressure of time would permit a coarse plot on the Video monitor. This preview technique would minimize the number of Plots accumulating on the worksheet from repeated trials.

Example:

Fig. 3 illustrates a simple application of this program. It is desired to optimize the VSWR of the ORIGINAL load such that the resultant network exhibits a reasonable and constant value for VSWR for the frequency range 50-100 MHz. The first trial element added, a series inductive reactance of +.32 moves the network impedance so that it is almost symetrical about the real axis. The addition of a series resistance of 0.2 moves the network impedance downward so that it is almost symetrical about the origin. A VSWR circle of 2 drawn on the plot by the program circumscribes the major portion of the load plot and the matching network is accepted.

Appendix

a. There is an additional convenience utility contained in the program that is not part of the operating system. This utility enabling the reading of the 5 working data files is located at lines 9000-9010 and is useful for debugging when program changes are made. The function of each file is:

FILE #1	ORIGINAL data	R , X
FILE #2	Smith Map	Unew, Vnew
FILE #3	Working/Result	Rnew, Xnew
FILE #4	Previous Result	Rold, Xold
FILE #5	Previous Smith Map	Uold, Vold

b. NOTE: The Math Pack for the HP71 and HP75 contain powerful, complex function and complex variable instructions that can simplify many parts of this program. Their usage was not included in this presentation in order to maintain universality with the typical Basic instruction sets in common use.

References

P. Smith Electronic Applications of the Smith Chart, McGraw-Hill 1969, Krieger Publishing 1983

R. Thomas A Practical Introduction To Impedance Matching, Artech House 1976

T. Cuthbert Circuit Design Using Personal Computers, Wiley 1983

R. Chipman Transmission Lines, McGraw-Hill-Schaum

L. Gerig Computer Interface For Smith Chart Calculations, RF Design, January-February 1982

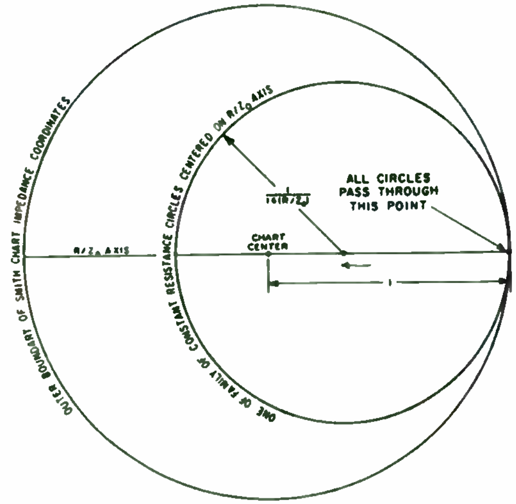


Fig 1 **Construction of normalized resistance circles for SMITH CHART of unit radius.**

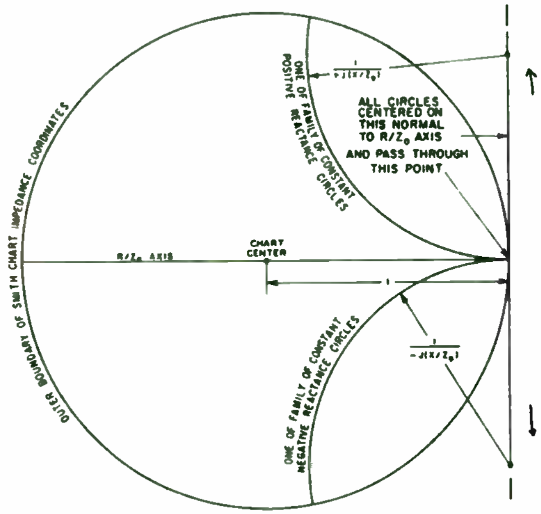
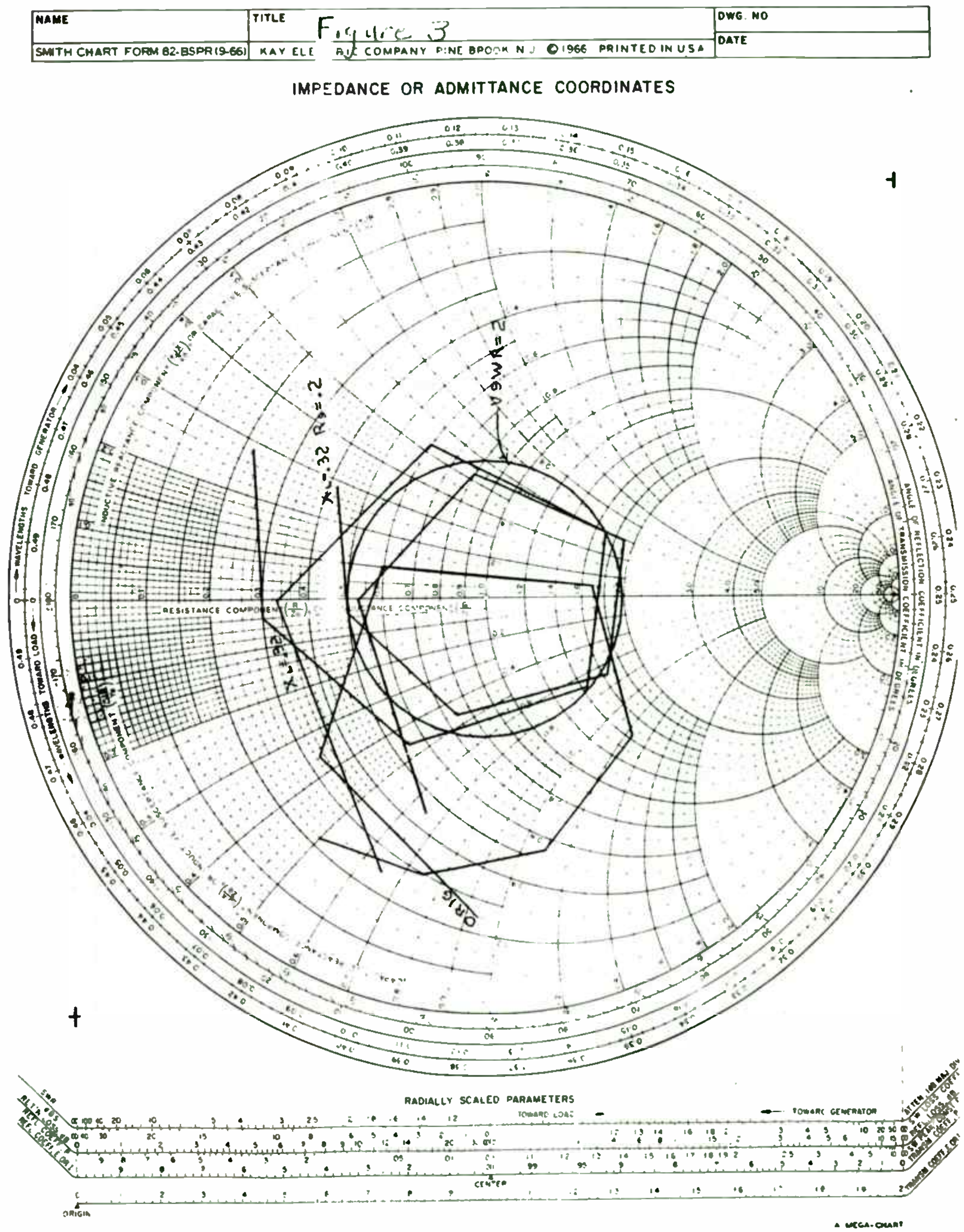
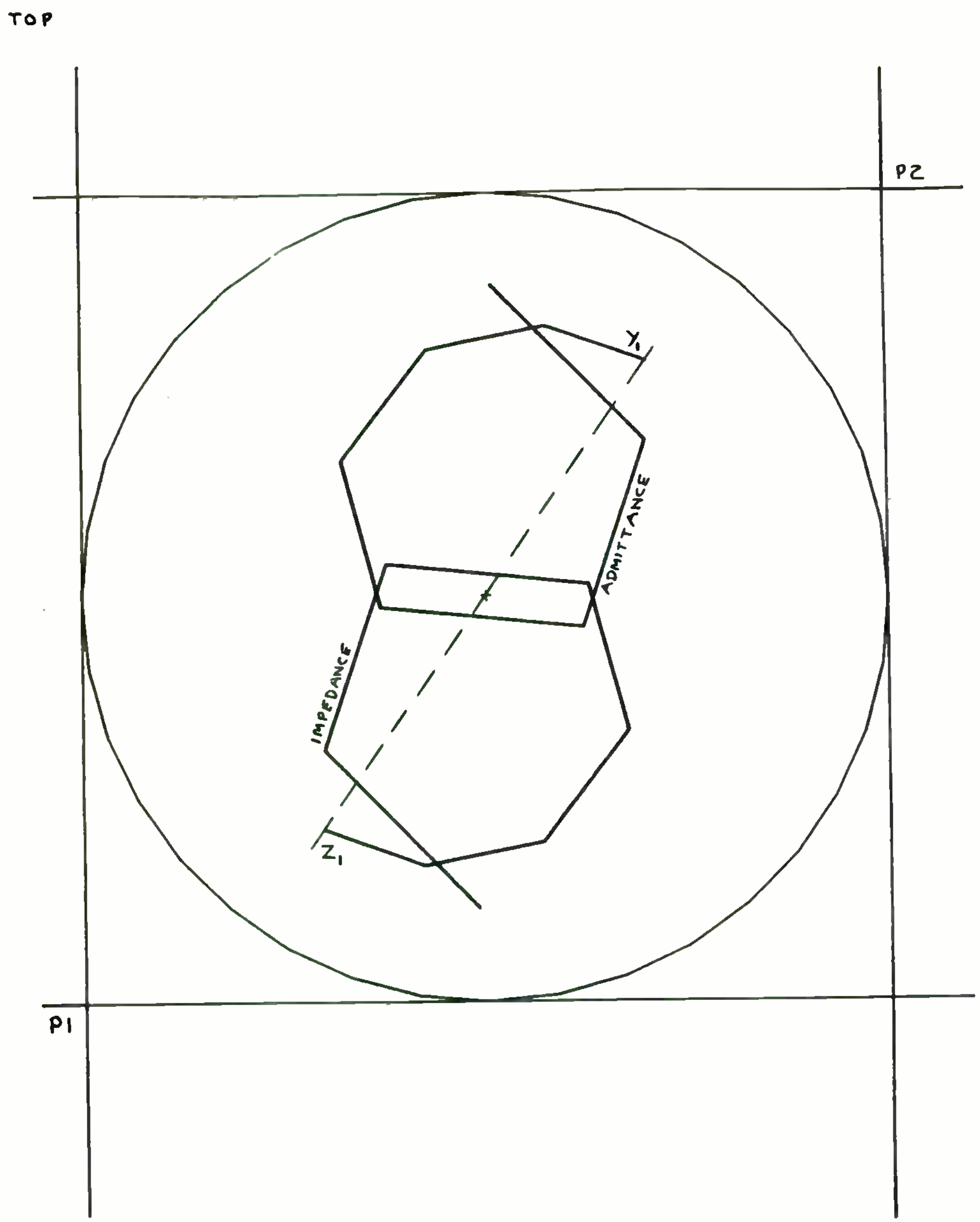


Fig 2 **Construction of normalized reactance circles for SMITH CHART of unit radius.**





Smith Chart Defining Equations

1) $\rho = \dfrac{Z_n - 1}{Z_n + 1} = U + jV$ complex reflection coefficient

2) $Z_n = R_n + jX_n$ network impedance normalized

3) $U + jV = \dfrac{R_n + jX_n - 1}{R_n + jX_n + 1}$

4) eliminating X_n $\left[U - \dfrac{R_n}{R_n + 1}\right] + V^2 = \dfrac{1}{(R_n + 1)^2}$

circle: center at $\dfrac{R_n}{R_n + 1}$, $V = 0$ radius $= \dfrac{1}{R_n + 1}$

5) eliminating R_n $(U - 1)^2 + \left(V - \dfrac{1}{X_n}\right)^2 = \left(\dfrac{1}{X_n}\right)^2$

circle: center at $\dfrac{1}{X_n}$, $U = 1$ radius $= \dfrac{1}{X_n}$

Note symmetry of positive and negative reactance circles.

Transform Equations

$(R, X) \Rightarrow (U, V)$

$R_n = (1 - U^2 - V^2)/[V^2 + (U - 1)^2]$

$X_n = 2V/[V^2 + (U - 1)^2]$

$(U, V) \Rightarrow (R, X)$

$U = (R_n^2 + X_n^2 - 1)/[X_n^2 + (R_n + 1)^2]$

$V = 2X_n/[X_n^2 + (R_n + 1)^2]$

$VSWR = \dfrac{1 + |\rho|}{1 - |\rho|}$ circle: center at $U = 0$, $V = 0$ radius $=$ VSWR

ORIGINAL

Freq	Real	Imag
40.	0.2500	-0.9400
50.	0.3300	-0.3600
60.	0.6000	0.1000
70.	1.7000	0.1000
80.	1.4500	-1.2700
90.	0.5500	-1.1000
100.	0.3000	-0.7500
110.	0.2200	-0.5000

Series X= .32

Freq	Real	Imag
40.	0.4500	-0.3711
50.	0.5300	0.3511
60.	0.8000	0.9533
70.	1.9000	1.0956
80.	1.6500	-0.1322
90.	0.7500	0.1800
100.	0.5000	0.6722
110.	0.4200	1.0644

Series R= .2

Freq	Real	Imag
40.	0.6500	-0.3711
50.	0.7300	0.3511
60.	1.0000	0.9533
70.	2.1000	1.0956
80.	1.8500	-0.1322
90.	0.9500	0.1800
100.	0.7000	0.6722
110.	0.6200	1.0644

Admittance

Freq	Real	Imag
40.	1.1602	0.6625
50.	1.1125	-0.5351
60.	0.5239	-0.4994
70.	0.3743	-0.1953
80.	0.5378	0.0384
90.	1.0161	-0.1925
100.	0.7432	-0.7137
110.	0.4086	-0.7015

HP71 Digiplot

```
1 ! SMITH CHART #7 HP71 DIGIPLOT
10 DISP 'Set Scale Pts. P1,P2 !' @ BEEP 500,.5 @ BEEP 500,.5
15 PRINTER IS ':G1' @ PRINT 'IP8598,360,1300,7650';'SC-1,1,-1,1';'VS7' ! xeroxcopy
20 ! PRINTER IS ':G1' @ PRINT 'IP7631,355,1045,6987';'SC-1,1,-1,1' ! phil's ovrly
22 ! PRINTER IS ':G1' @ PRINT 'IP8508,382,1279,7650';'SC-1,1,-1,1' !printedchart
25 DIM U2(15),V2(15)
30 SHORT A0,A1,A2,A3,B,B2,B3,D,D0,F,F0,G1,G2,N,N1,R,R0,R2,U1,V1,U,V,W,W1,X,X0,X1,X2,Y0
35 B=1 @ D$='ORIGINAL' @ INPUT '# Data Pts.=?';N1          B=1 read original data file
40 DISP 'Create New Files?(Y/N)?' @ A$=UPRC$(KEYWAIT$) @ IF A$='Y' THEN GOSUB 9020
45 INPUT 'Printer is: D1(disp) P1(prnt)?';P$ @ P$=':'&P$
50 GOSUB 9120
55 DISP 'AutoPlotMode:Yes(1)/No(0)' @ B2=VAL(KEYWAIT$)          B2=1 Autoplotmode
60 FOR N=1 TO N1
65 DISP 'Data Pt.#';N;'=';  @ INPUT F,R,X @ PRINT #1,N-1;F,R,X ! FREQ,REAL,IMAG
70 NEXT N
75 RESTORE #1 @ GOTO 1500
500 INPUT 'Delete Previous Element(Y/N)?';K$ @ IF UPRC$(K$)='Y' THEN GOSUB 9100
502 DISP 'Series(0)orShunt(1) Impedance?' @ S0=VAL(KEYWAIT$)
510 DISP 'Line(1)orReact(2)orResis(3)?' @ S1=VAL(KEYWAIT$)
515 PURGE 'file4' @ COPY 'file3' TO 'file4'  - put old result→4
520 IF S0=0 THEN ON S1 GOTO 3000,2000,2500
530 IF S0 THEN ON S1 GOTO 5000,4500,4000
1000 DISP 'Next(1),Init(2),Prnt(3),Plot(4)' @ A=VAL(KEYWAIT$)
1002 DISP 'Working!'
1005 IF B3 AND A=1 THEN GOSUB 9110 @ B3=0
1010 ON A GOTO 500,1015,1100,1300
1015 RESTORE #1 @ B=1 @ D$='ORIGINAL' @ GOTO 1500
1100 RESTORE #3 @ PRINTER IS P$ ! Print Data
1105 PRINT TAB(8);D$; @ IF B=1 OR B3 THEN PRINT ELSE PRINT '=';D0
1110 PRINT 'Freq';TAB(8);'Real';TAB(18);'Imag'
1120 FOR D=1 TO N1 @ READ #3;F,R,X          - get result←3
1130 FIX 0 @ PRINT F;TAB(7); @ FIX 4 @ PRINT R;TAB(17);X
1140 NEXT D @ STD @ IF B2 THEN RETURN ELSE 1000
1300 RESTORE #2 @ PRINTER IS ':G1' @ PRINT 'PU';'SP1' ! Plot Data HP
1310 FOR D=1 TO N1 @ READ #2;F,U,V  - get Smith←2
1410 PRINT 'PA',V,U @ IF D=1 THEN INPUT 'Start Pt.O.K.?';K$   - is plot start proper?
1420 IF D=1 AND UPRC$(K$)='Y' THEN PRINT 'PD'
1430 NEXT D @ DISP 'Plot Done' @ BEEP @ PRINT 'SP0';'PU'
1435 IF B2 THEN RETURN ELSE 1000
1500 RESTORE #2 @ RESTORE #3 ! Init @ Reslt          Initialize Results
1510 FOR D=1 TO N1 @ READ #B;F,R,X  -  get ORIG (B=1) or Results (B=3)
1520 U2(D)=(R^2+X^2-1)/(X^2+(R+1)^2) @ V2(D)=2*X/(X^2+(R+1)^2) @ PRINT #2,D-1;F,U2(D),V2(D)  (R,X)→(U,V)
1530 IF B=1 THEN PRINT #3,D-1;F,R,X  - put ORIG into result
1540 NEXT D
1550 IF B2 THEN GOSUB 1100 @ GOSUB 1300    Auto Output
1555 GOTO 1000
2000 ! Series Reactance
2010 RESTORE #3 @ INPUT 'Ref.Freq=?';F0 @ INPUT 'Xcap(0) or Xind(1),Xvalue=?';A,X0 @ D0=X0
2020 FOR D=1 TO N1 @ READ #3;F,R,X          read work← 3
2030 IF A THEN X1=X0*F/F0 ELSE X1=X0*F0/F
2040 PRINT #3,D-1;F,R,X+X1                  put result→ 3
2050 NEXT D @ B=3 @ D$='Series X' @ GOTO 1500
2500 ! Series Resistor
2510 RESTORE #3 @ INPUT 'Series Res.=?';R2 @ D0=R2
2515 FOR D=1 TO N1 @ READ #3;F,R,X @ PRINT #3,D-1;F,R+R2,X  getwork,putresult 3→3
2520 NEXT D @ B=3 @ D$='Series R' @ GOTO 1500
3000 ! Series Line Xfer
3005 RESTORE #2 @ RESTORE #3 @ INPUT 'Ref.Freq=?';F0
3007 PURGE 'file5' @ COPY 'file2' TO 'file5'  -  put Smith →5 (save2)
3010 INPUT 'Wavelengths=?';W @ D0=W @ W=W*720 ! angle rotated
3020 FOR D=1 TO N1 @ READ #2;F,U,V @ W1=W*F/F0 ! angle at F   get Smith← 2
3030 U1=U*COS(W1)+V*SIN(W1) @ V1=-U*SIN(W1)+V*COS(W1) @ PRINT #2,D-1;F,U1,V1 ! rot.u,v  put rotated Smith→2
3040 R2=(1-V1^2-U1^2)/(V1^2+(U1-1)^2) @ X2=2*V1/(V1^2+(U1-1)^2) @ PRINT #3,D-1;F,R2,X2  put rotated Result→3
```

HP71 Digiplot

```
4000 ! Shunt Resistor
4010 RESTORE #3 @ INPUT 'Shunt Res.=?';R1 @ G1=1/R1 @ D0=R1
4020 FOR D=1 TO N1 @ READ #3;F,R,X @ Y0=1/(R^2+X^2) @ G2=G1+R*Y0 @ S=X*Y0     set work ← 3
4030 R0=1/(G2^2+S^2) @ R2=G2*R0 @ X2=S*R0 @ PRINT #3,D-1;F,R2,X2             put result
4040 NEXT D @ D=3 @ D$='Shunt R' @ GOTO 1500
4500 ! Shunt Reactance
4510 RESTORE #3 @ INPUT 'Ref.Freq=?';F0 @ INPUT 'Xcap(0) or Xind(1),Xvalue=?';A,X0 @ D0=X0
4520 FOR D=1 TO N1 @ READ #3;F,R,X
4530 IF A THEN S0=-1/(X0*F/F0) ELSE S0=-1/(X0*F0/F)
4540 Y0=1/(R^2+X^2) @ G1=R*Y0 @ S1=S0-X*Y0
4550 R0=1/(G1^2+S1^2) @ R2=G1*R0 @ X2=-S1*R0 @ PRINT #3,D-1;F,R2,X2
4560 NEXT D @ D=3 @ D$='Shunt X' @ GOTO 1500
5000 DISP 'no shunt lines yet' @ WAIT 2 @ GOTO 1000
6000 DISP 'Plot Utilities & Other Aids' @ WAIT 1
6001 DISP '(1)Admittance (2)VSWR (3)Digitize (4)Xfer?'
6002 ON VAL(KEYWAIT$) GOTO 6050,6100,6150,6200
6050 DISP ' ADMITTANCE Calcs!'
6051 PURGE FILE4 @ COPY FILE3 TO FILE4 @ GOSUB 9120 @ RESTORE #3
6052 FOR D=1 TO N1 @ READ #3;F,R,X @ Y0=1/(R^2+X^2) @ G1=R*Y0 @ S1=-X*Y0 @ PRINT #3,D-1;F,G1,S1
6054 NEXT D @ D=3 @ D3=1 @ D$='Admittance' @ GOTO 1500
6100 INPUT 'VSWR Circle=?';V0 @ V0=(V0-1)/(V0+1) @ PRINTER IS ':GI' @ PRINT 'SP2'
6102 FOR N=0 TO 360 STEP 10 @ U=V0*COS(N) @ V=V0*SIN(N) @ PRINT 'PA',U,V;'PD' @ NEXT N
6104 PRINT 'SP0';'PU' @ GOTO 1000
6150 ! READ PLOTTER COORDS:X,Y
6160 PRINTER IS ':GI' @ DISP 'USER(0) PLOT(1) UNITS?' @ P=VAL(KEYWAIT$)
6165 IF P THEN A$='OA' @ FIX 0 ELSE A$='OC' @ FIX 4
6170 DISP 'ENTER NEXT PT! (or No)' @ BEEP @ A4$=KEYWAIT$
6175 IF A4$='#38' THEN 6180 ELSE 1000
6180 PRINT A$ @ ENTER ':GI' ;P1,P2
6185 DISP 'X=';P1;'Y=';P2 @ WAIT 2
6190 GOTO 6170
6200 ! Xfer Scaling
6205 PURGE 'FILE4' @ COPY 'FILE3' TO 'FILE4' @ GOSUB 9120 @ RESTORE #3
6210 INPUT 'Xfer Scale=?';A3 @ D0=A3
6215 FOR D=1 TO N1 @ READ #3;F,R,X @ PRINT #3,D-1;F,A3*R,A3*X @ NEXT D
6220 D3=1 @ D=3 @ D$='Xfer' @ GOTO 1500
9000 ! Read Data Files
9001 INPUT 'File#=?';A0 @ DISP 'File#=';A0 @ ASSIGN #A0 TO 'FILE'&STR$(A0) @ RESTORE #A0
9010 FOR D=1 TO N1 @ READ #A0;F,R,X @ DISP F;R;X @ NEXT D @ GOTO 9000
9020 FOR N=1 TO 5 @ PURGE 'FILE'&STR$(N) @ NEXT N
9030 FOR N=1 TO 5 @ CREATE DATA 'FILE'&STR$(N),N1,24 @ NEXT N @ RETURN
9100 IF D$='LineXfer' THEN PURGE 'file2' @ COPY 'file5' TO 'file2'
9110 PURGE 'file3' @ COPY 'file4' TO 'file3'
9120 ASSIGN #1 TO 'file1' @ ASSIGN #2 TO 'file2' @ ASSIGN #3 TO 'file3'
9130 ASSIGN #4 TO 'file4' @ ASSIGN #5 TO 'file5' @ RETURN ! get prev state
```

HP75

```
0 ! SMITH CHART 06 HP DIGIPLOT
1 DISP 'Set Scale Pts. P1,P2 !' @ BEEP 3000,.5 @ BEEP 3000,.5
2 PRINTER IS ':GI' @ PRINT 'IP8549,360,1268,7603';'SC-1,1,-1,1';'VS5'
4 DIM U(15),V(15)
5 SHORT F,R,R2,X,X1,X2,U1,V1,N,N1,U,V
8 INPUT 'Printer is: DI(disp) PI(prnt)?'; P$ @ P$=':'&P$
9 ASSIGN # 4 TO 'file4' @ ASSIGN # 5 TO 'file5'
10 ASSIGN # 1 TO 'file1' @ ASSIGN # 2 TO 'file2' @ ASSIGN # 3 TO 'file3'
11 DISP 'AutoPlotMode:Yes(1)/No(0)' @ D2=VAL(SKEY$)
15 D=1 @ D$='ORIGINAL' @ INPUT '# Data Pts=?';N
20 FOR N1=1 TO N
25 DISP 'Data Pt.#';N1;'='; @ INPUT F,R,X @ PRINT # D ; F,R,X ! FREQ,REAL,IMAG
30 NEXT N1
40 RESTORE # 1 @ GOTO 1500
500 INPUT 'Delete Previous Element(Y/N)?'; K$ @ IF UPRC$(K$)='Y' THEN GOSUB 9990
502 DISP 'Series(0)orShunt(1) Impedance?' @ S0=VAL(SKEY$)
510 DISP 'Line(1)orReact(2)orResis(3)?' @ S1=VAL(SKEY$)
515 PURGE 'file4' @ COPY 'file3' TO 'file4'
520 IF S0=0 THEN ON S1 GOTO 3000,2000,2500
530 IF S0 THEN ON S1 GOTO 5000,4500,4000
1000 DISP 'Next(1),Init(2),Prnt(3),Plot(4)' @ A=VAL(SKEY$)
1010 ON A GOTO 500,1015,1100,1300
1015 RESTORE # 1 @ D=1 @ D$='ORIGINAL' @ GOTO 1500
1100 RESTORE # 3 @ PRINTER IS P$ ! Print Data
1105 PRINT TAB(8);D$; @ IF D=1 THEN PRINT ELSE PRINT '=';D0
1110 PRINT 'Freq';TAB(8);'Real';TAB(18);'Imag'
1120 FOR D=1 TO N @ READ # 3 ; F,R,X
1130 PRINT F;TAB(7);ROUND(R,4);TAB(17);ROUND(X,4)
1140 NEXT D @ IF D2 THEN RETURN ELSE 1000
1300 RESTORE # 2 @ PRINTER IS ':GI' @ PRINT 'PU';'SP1' ! Plot Data HP
1310 FOR D=1 TO N @ READ # 2 ; F,U,V
1410 PRINT 'PA',V,U @ IF D=1 THEN INPUT 'Start Pt.O.K.?';K$
1420 IF D=1 AND UPRC$(K$)='Y' THEN PRINT 'PD'
1430 NEXT D @ DISP 'Plot Done' @ BEEP @ PRINT 'SP0';'PU'
1435 IF D2 THEN RETURN ELSE 1000
1500 RESTORE # 2 @ RESTORE # 3 ! Init@Reslt
1510 FOR D=1 TO N @ READ # D ; F,R,X
1520 U(D)=(R^2+X^2-1)/(X^2+(R+1)^2) @ V(D)=2*X/(X^2+(R+1)^2)
1525 U=U(D) @ V=V(D) @ PRINT # 2 ; F,U,V
1530 IF D=1 THEN PRINT # 3 ; F,R,X
1540 NEXT D
1550 IF D2 THEN GOSUB 1100 @ GOSUB 1300
1555 GOTO 1000
2000 ! Series Reactance
2010 RESTORE # 3 @ INPUT 'Ref.Freq=?'; F0 @ INPUT 'Xcap(0) or Xind(1),Xvalue=?'; A,X0 @ D0=X0
2020 FOR D=1 TO N @ READ # 3 ; F,R,X
2030 IF A THEN X1=X0*F/F0 ELSE X1=X0*F0/F
2040 PRINT # 3,D ; F,R,X+X1
2050 NEXT D @ D=3 @ D$='Series X' @ GOTO 1500
2500 ! Series Resistor
2510 RESTORE # 3 @ INPUT 'Series Res.=?'; R2 @ D0=R2
2515 FOR D=1 TO N @ READ # 3 ; F,R,X @ PRINT # 3,D ; F,R+R2,X
2520 NEXT D @ D=3 @ D$='Series R' @ GOTO 1500
3000 ! Series Line Xfer
3005 RESTORE # 2 @ RESTORE # 3 @ INPUT 'Ref.Freq=?';F0
3007 PURGE 'file5' @ COPY 'file2' TO 'file5'
3010 INPUT 'Wavelengths=?'; W @ D0=W @ W=W*720 ! angle rotated
3020 FOR D=1 TO N @ READ # 2 ; F,U,V @ W1=W*F/F0 ! angle at F
3030 U1=U*COS(W1)+V*SIN(W1) @ V1=-U*SIN(W1)+V*COS(W1) @ PRINT # 2,D ; F,U1,V1 ! rot.u.v
3040 R2=(1-V1^2-U1^2)/(V1^2+(U1-1)^2) @ X2=2*V1/(V1^2+(U1-1)^2) @ PRINT # 3 ; F,R2,X2
3050 NEXT D @ D=3 @ D$='LineXfer' @ GOTO 1500
4000 ! Shunt Resistor
4010 RESTORE # 3 @ INPUT 'Shunt Res.=?'; R1 @ G1=1/R1 @ D0=R1
4020 FOR D=1 TO N @ READ # 3 ; F,R,X @ Y0=1/(R^2+X^2) @ G2=G1+R*Y0 @ S=X*Y0
```

```
4030 R0=1/(G2^2+S^2) @ R2=G2*R0 @ X2=S*R0 @ PRINT # 3,D ; F,R2,X2
4040 NEXT D @ D=3 @ D$='Shunt R' @ GOTO 1500
4500 ! Shunt Reactance
4510 RESTORE # 3 @ INPUT 'Ref.Freq=?'; F0 @ INPUT 'Xcap(0) or Xind(1),Xvalue=?'; A,X0 @ D0=X0
4520 FOR D=1 TO N @ READ # 3 ; F,R,X
4530 IF A THEN S0=-1/(X0*F/F0) ELSE S0=-1/(X0*F0/F)
4540 Y0=1/(R^2+X^2) @ G1=R*Y0 @ S1=S0-X*Y0
4550 R0=1/(G1^2+S1^2) @ R2=G1*R0 @ X2=-S1*R0 @ PRINT # 3,D ; F,R2,X2
4560 NEXT D @ D=3 @ D$='Shunt X' @ GOTO 1500
5000 DISP 'no shunt lines yet' @ STOP
8888 ! SMITH CHART HP PLOT SCALE
9990 INPUT 'Copy file#? to file#?'; A1,A2 @ INPUT 'Xfmr Scale=?';A3
9991 RESTORE # A1 @ RESTORE # A2
9992 FOR D=1 TO N @ READ # A1 ; F,R,X @ PRINT # A2 ; F,A3*R,A3*X @ NEXT D @ STOP
9998 IF D$='LineXfmr' THEN PURGE 'file2' @ COPY 'file5' TO 'file2'
9999 PURGE 'file3' @ COPY 'file4' TO 'file3' @ RETURN ! recall previous status
```

CALCULATOR- AND COMPUTER-AIDED DESIGN TOOLS FOR THE RF ENGINEER

Steven L. March

Compact Software, Inc, 1314 Sam Bass Circle, Round Rock, Texas 78664

INTRODUCTION

For years, the Virginia Slims cigarette advertisement extolled, "You've come a long way, baby!". The same can be said about software for the RF and microwave design engineer. "You've come a long way, engineer!" This paper will examine just how far the RF engineer has progressed by examining the software and sources for programming information that he/she has available for circuit and/or system design today. It will also examine the question of "make or buy" as it applies to software for the RF designer.

THE NEED FOR DESIGN AIDS (SOME HISTORY)

In the early years of RF and microwave design, one of the features of microwave components and RF circuits had been the amount of engineering labor that went into each design. The engineer designed the component or circuit, sketched out a preliminary drawing for mechanical fabrication, made a second drawing of the circuit for electrical design and/or layout, and waited for the parts to be delivered to him for assembly. He probably even assembled the first piece himself. When the circuit did not perform as expected, optimization was required. The empirical adjustment or tuning of the characteristics was performed by inserting screws at judicious locations or by the introduction of dielectric inserts at the proper points. In some waveguide circuits, "hand optimization" meant reaching into the tool box, retrieving a ball-pien hammer, and denting the waveguide in the correct place so as to improve its performance! C-clamps were also very popular for a similar purpose. Microwave and RF design was "black magic" and each successful engineer was a Harry Houdini.

Enter the computer. Enter computer-aided design. Much has been written about the value of computer-aided design as an engineering tool. The first Special Issue of an IEEE publication devoted to computer-aided design was the November 1967 Proceedings of the IEEE. Since then, there have been ten additional Special Issues of the Proceedings of the IEEE or one of the IEEE Transactions devoted to computer-aided design, analysis, modeling, or computational methods. Even its most severe critics now concede that computer-aided design has brought about significant improvements in circuit performance

and producibility. Perhaps the most important contribution made by the existence and utilization of CAD is the vast reduction in the amount of time required to design and implement new circuits, when compared to the amount of time the same procedures would have required without using CAD techniques.

CAD - MAKE OR BUY

Do I need CAD? What CAD do I need? Do I buy a program already available or do I write my own? Mainframe, super-mini-computer, minicomputer, desktop computer, personal computer, or handheld-calculator - which do I need? These can be very complex issues. However, if the only thing that your company currently manufactures and anticipates producing in the future is 225 MHz to 400 MHz two-way lumped-element power dividers, the same as it has been making for the past decade, then you do not need CAD, and probably never will. On the other hand, if you are stretching the state-of-the-art in dc to 40 GHz distributed FET amplifiers, you definitely do need CAD.

Buying "canned" software or developing your own depends on several factors. Purchasing a software product from a company such as Compact Software requires a capital outlay. In-house development means that the expense is spread out over a period of time. Purchased software is available today (maybe tomorrow, in some cases) and has been tested, verified, and debugged. An internally developed program may require research, the development of many lines of code, verification, documentation, and probably extensive debugging. Can your company wait that long, do you have the people to put on the project, are they qualified, what could they be designing if they were not developing software, what is the effort going to cost me in terms of fully-loaded labor dollars? These questions all have to be answered before an intelligent make or buy decision can be rendered.

LET'S DEVELOP IT HERE

If you have decided to develop your own software for a specific application, you will find an abundance of printed and published material to aid you. There are books, magazines articles, reports from Government agencies, Government contractors, and universities, theses, informative information from companys that make and sell computers and/or calculators, and software directories available to make the job easier. Many of these sources contain well developed and tested software that can be easily copied and incorporated in any new product.

Books

There are at least two dozen excellent books on the market or in many libraries which deal with programming techniques and languages. Additionally, there are at least another two dozen books concerned with numerical and mathematical techniques (matrix manipulations, integration, differentiation, curve-fitting, solution of equations, etc.) when performed on a computer. There are at least thirteen books dedicated to optimization and an equal number devoted to the topic of graphics. The computer-aided design of electronic circuits and components is the theme of another dozen books, while there are twice as many written about the general topic of computer-aided design and analysis. Some of the better ones (in the opinion of this author) are listed in the bibliography. In fact, the soon to be published Antenna Design Using Personal Computers (D.M. Pozar, Artech House) will include diskettes containing several programs.

Reports

Another excellent source for written and tested software on a variety of topics is reports and technical memoranda generated by Government agencies and those written by commercial organizations and universities, especially when produced under Governmental contract. There are documents available that deal with surface acoustic waves, optimization, test equipment calibration, noise and and semiconductor device analysis, antenna and electronic component design, the characterization of various transmission media, network synthesis, and general circuit analysis and design. Several representative reports, which contain program listings, have been included in the bibliography.

Articles

Magazine articles represent an excellent source of abundant and free software, especially short programs and subroutines for the calculation, design, or analysis of individual electronic devices, components, and techniques. Over the past decade, over forty-eight articles, containing computer or calculator program listings, have been published in Microwaves magazine. In addition, articles which contain software listings for calculators or larger computing equipment include Electronic Design (28), Electronic Design News (27), Electronics (16), Microwave Journal (13), RF Design (12), and Microwave System News (4). There have also been at least one-half dozen programs or subroutines published in the Proceedings of the IEEE, various IEEE Transactions, and records from IEEE-sponsored conferences. This author has located at least 154 magazine articles which contain software of probable interest to the RF or microwave engineer.

For many years, the Microwave Journal published The Microwave Engineers' Technical and Buyers Guide. The editions published in 1974, 1975, and 1976 contained 36 subroutines, written in FORTRAN or BASIC, and of interest to the RF designer. Fifteen of the programs appeared in 1974, seventeen in 1975, and four in 1976.

The vast majority of the articles dealt with the design of specific devices and components such as helical antennas, PIN diode switches, power dividers, phase-locked loops, inductors, and transformers. Over a dozen articles dealt with the characterization and analysis of microstrip, another six with RF directional couplers, eighteen others with transistor amplifier design, analysis, and performance, and a dozen more with the characterization and/or design of filters. There were over a half-dozen articles that contained various transforms (FFT or fast Hartley transform, for example) for time-domain or waveform analysis. An equal number contained listings for the design of matching networks. There were also several which dealt with EMI and RFI characterization. Finally, there have been a number of magazine articles which contained subroutines for mathematical operations such as interpolation, convolution, matrix inversion, and Laplace transformations.

MTT Transactions

Beginning in August 1969, a new section was added to the IEEE Transactions on Microwave Theory and Techniques. The new section, Computer Program Descriptions, was introduced for the dissemination of software relative to the field of microwaves and available to the general public from the author of the program or the ASIS/NAPS service (ASIS/NAPS, c/o Microfiche Publications, P. O. Box 3513, Grand Central Station, New York, NY 10017). The last Computer

Program Description appeared in the March 1980 MTT Transactions. In that time frame, thirty-eight issues of the IEEE Transactions on Microwave Theory and Techniques contained fifty-eight listings of computer-program descriptions.

Directories and Catalogs

Computer manufacturers brag about the abundance of internally-generated or third-party developed software that is available to run on their hardware. Hewlett-Packard Company has developed and sells applications paks, solutions handbooks, and code applicable to their entire range of products - desktop computers, hand-held calculators, and minicomputers. There are applications paks for electronics, mathematics, statistics, and circuit analysis. Also, solutions handbooks have been published for antenna engineering, high level mathematics, statistics, and electrical engineering. The Hewlett-Packard Software Solutions Catalog [1] contains complete descriptions for over 275 software products developed by HP. The Autumn/Winter 1984/1985 edition of the Hewlett-Packard Computer Users' Catalog [2] lists over 1500 entries for programs which have been developed to run on HP computers. And it's FREE! Hewlett-Packard also publishes a Third Party Technical Software Solutions Catalog [3] for its Series 200 desktop computers. In addition, Digital Equipment Corporation's 'PDP-11 Software Sourcebook' [4] lists more than 1200 applications packages from over 250 sources. Probably only a small number of these programs will actually be of interest to the RF engineer.

The fields of computer programming and computer software have grown so rapidly and so large in the past decade that there is now an 'International Directory of Software' [5], being published. It lists over 5100 software packages for mainframes, minicomputers, and microcomputers. On the other side of the Atlantic Ocean, Butterworth Publishers in Surrey, England have put together the 'Computer-Aided Design Directory/Buyer's Guide' [6].

Therefore, if the decision is made to develop, in-house, some CAD software for RF and/or microwave circuit or component design, analysis, synthesis, or optimization, there are more than ample sources of previously-written and tested subroutines available. Before 'reinventing the wheel', several of these sources should be consulted. The savings in time and effort could be considerable.

THE DECISION TO BUY

If you have a need for RF CAD tools, if it is more general than just a need to design one or two specific components and if you can not wait the months (or years, depending on the complexity of the program) that an in-house development effort would take, then you probably have decided that you should purchase an already developed computer-aided design program from a software vendor.

There already exist at least sixty-one (61) programs, libraries of subroutines, and 'design kits' that can be purchased from thirty-four different companies. Most of these companies have only one commercially-available program or design aid. There are five firms that offer two CAD products and another five companies that have at least three different programs available for purchase. These are listed in Table I.

TABLE I

COMPANIES WITH THREE OR MORE SOFTWARE PRODUCTS
COMPACT SOFTWARE, INC. (5)
BV Engineering (5)
Communications Consulting Corporation (8)
Jensen Transformers (4)
Optimization System Associates (3)
COMPANIES WITH TWO SOFTWARE PRODUCTS
AB Associates
EEsof, Inc.
I R & D Electronics
Made-It Associates
Spefco Software

In the past six months, there have been several articles [7-12] that have dealt with commercially-available computer-aided design and development software for microwave and RF circuits. The most recent and comprehensive of these was the one by Barry Manz [7], which appeared in Microwaves and RF magazine. Manz presented an annotated listing of forty (40) different CAD software packages available from nineteen companies. In a companion article, March [8] presented some additional descriptive information on software available for circuit layout and mask generation. The compilation prepared by Trubitt and Podell [10] was concerned primarily with six software packages for RF circuit analysis that runs on either personal computers or desktop computers. The article compares the performance, features, and shortcomings of the six programs. The addendum briefly lists five other analysis programs.

What is Available

The sixty-one commercially-available software products can be placed into seven general classifications. These are: ac small-signal, linear analysis; programs for the design or analysis of a particular type of circuit element (except filters); programs for filter design and analysis; mathematical subroutine packages; graphics and layout; ac/dc linear/nonlinear circuit analysis; all else. Table II lists the number of programs that are available in each of these categories.

TABLE II

AVAILABLE RF CAD SOFTWARE BY CATEGORY

AC SMALL-SIGNAL LINEAR ANALYSIS	20
AC/DC LINEAR/NONLINEAR ANALYSIS	8
DEVICE (EXCEPT FILTERS) DESIGN	6
GRAPHICS AND CIRCUIT LAYOUT	6
FILTER DESIGN AND ANALYSIS	5
MATHEMATICAL SUBROUTINES	5
ALL ELSE (SYNTHESIS, OPTIMIZATION, DC-ONLY ANALYSIS, SCHEMATIC GENERATION, INTERMODULATION ANALYSIS)	11

Another way of cataloging the available software is by the type of equipment upon which it is designed to be operated. Somewhat arbitrarily, programmable computational equipment can be divided into six categories. These are: mainframe computers (such as the IBM 370 and the CDC Cyber); super-minicomputers (such as the PRIME and the DEC VAX units); minicomputers (such as the H-P 9000 series and the Apollo Domain nodes); desktop computers (such as the H-P 200 series and the Tektronix 405x series); personal computers (such as the IBM PC, the Apple IIe, and the Tandy TRS-80); handheld calculators (such as the H-P 41CV). Table III shows how many of the commercially-available RF CAD software packages are available to operate on each category of computing equipment.

It should be noted that many CAD software suppliers are willing to send a prospective purchaser or licensee a demonstration diskette, cassette, or tape for a nominal fee which can usually be credited toward the purchase price of the actual product.

TABLE III

AMOUNT OF SOFTWARE AVAILABLE BY COMPUTER TYPE

MAINFRAMES	11
SUPER-MINICOMPUTERS	13
MINICOMPUTERS	15
DESKTOP COMPUTERS	22
PERSONAL COMPUTERS	27
HANDHELD CALCULATORS	2

NOTE: The number of items exceeds the 61, because some of the commercial software is available for more than one category of equipment

It would require more pages than were used by Barry Manz to detail all of the commercially-available software packages of interest to the RF engineer. Instead, only a sampling of the available programs will be sescribed below. Additional programs will be discussed during the actual conference presentation.

AC Small-Signal Linear Circuit Analysis -- SUPER-COMPACT

SUPER-COMPACT is still the premier product for the analysis, synthesis, anad optimization of small-signal, linear networks. The program is intended for lumped-element or distributed circuits, from dc to well into the microwave range. The program supports user-interactive graphics on a variety of raster-scan terminals having a true graphics capability. The program performs analysis using either ABCD chain matrices or the indefinite admittance matrix, depending on the complexity of the ciruit and the number of accessible ports. A transmission media design section (TRL), an extensive transistor databank, a line editor, and a Lange coupler design subprogram (LANG) are internal to Super-Compact. The total capabilities are shown in Figure 1.

Circuit optimization can be performed using either a randomized pattern search procedure or gradient methods. These methods can be combined by the user for maximum efficiency and the highest probability of truly locating the global minimum of the performance error function. Performance objectives include group delay, phase, gain, reflection coefficient, etc.

Additionally, a device modeling capability allows input and output impedances to be semi-automatically calculated for any active two-port and provides reasonably accurate values which can be used directly for synthesizing input and output matching networks. Both network synthesis and tolerance analysis capabilities are contained in the program.

Output data may be plotted on polar, rectangular, or Smith charts for any of the available 2-, 3-, or 4-port parameters. Performance limits and frequency coverage can be expanded or compressed in order to more closely examine performance details. In addition, constant gain, noise, and stability circles can be calculated.

With the TIME DOMAIN option, the step, ramp, or impulse response of almost any circuit can be calculated and plotted. The time-domain response is calculated using a Fast Fourier Transform algorithm. Super-Compact contains about 93,000 lines of FORTRAN code. The program is available for large mainframe computers, super-mini-computers, and minicomputers. Super-Compact PC is the personal computer version which incorporates a flexible report writer, a full-screen editor, and a tune-mode capability, but does not have the transistor databank or the synthesis and device modeling capabilities.

Programs such as Touchstone and SUGAR contain several features not incorporated to date in Super-Compact. These include the capability to calculate internal node voltages, to plot constant gain or constant noise circles, to plot performance as the optimizer is improving the result, and to optimize differential measurements, such as directivity.

AC/DC Linear/Nonlinear Analysis -- IG-SPICE

SPICE is a general purpose circuit simulation program. It can simulate and analyze circuits composed of active and passive elements. SPICE will calculate a circuit's dc operating point, will calculate its dc transfer function, will determine its component sensitivities at dc, and will perform a noise level determination. In addition, SPICE has the ability to calculate a network's ac transfer function and to perform a distortion analysis.

Using the time-domain, Fast Fourier Transform, analysis capability in SPICE, the transient response of a network can be evaluated as a function of time. With the sensitivity analysis, the sensitivity of any frequency-domain network performance calculation with respect to any network component changes can be evaluated. With the statistical analysis, worst case performance and tolerance analysis can be evaluated for the circuit. This analysis option is based on a Monte Carlo technique for component value variation.

In its latest implementation, IG-SPICE features a fully inter-active mode of operation using CRT terminals such as the DEC VT125 and the Tektronix 40xx series, instead of the batch mode in which the original SPICE was implemented. IG-SPICE allows inputs in free-format and supports both tabular and graphical displays of the results. There is also a zoom capability for the output plots. IG-SPICE also allows the inclusion of user-defined FORTRAN subroutines. This allows new models to be added quickly and conveniently.

In addition, IG-SPICE contains a very extensive element library. When IG-SPICE is purchased from AB Associates, the user also receives a 250 page applications handbook and a 12 hour video cassette training seminar.

In addition to IG-SPICE which is available for mainframe and super minicomputers, M-SPICE is available from Mentor Graphics for the Apollo Domain minicomputer, HP-SPICE can be purchased for the HP 9000 series of minicomputers, and SPICE-D is the desktop computer version which can be purchased from Cromemco.

Filter Design and Analysis -- S/FILSYN

Most computer programs for the design of complex RF, microwave or low-frequency filters are saddled with drawbacks. Not only do they require specialized knowledge, they are usually based upon approximations. Designers who use such programs, assumung that their results will lead to circuits having realistic filter dimensions and component values, might very well be disappointed. And while more practical programs than S/FILSYN exist, most are obscure and not readily available to the engineering community.

S/FILSYN was originally conceived by George Szentirmai in 1963. It was started in 1967 and the first 5000 lines of code required five years of Dr. Szentirmai's free time and weekends. In 1977, the program contained 99 FORTRAN subroutines and more than 10,000 lines of code. Currently, S/FILSYN contains over 25,000 lines.

S/FILSYN is a general purpose filter design program that offers exact synthesis of commensurate transmission line networks as well as lumped L-C circuits. The filter types include lowpass, linear-phase lowpass, highpass, bandstop, and both conventional and parametric bandpass. The passband response can be Butterworth, Chebyshev, flat or sloped, with various stopband specifications. Both single and double terminated networks are offered; the former type is particularly suited for multiplexer design. Users can specify their own topology or the program can provide a suitable one. Predistorted design is available for lossy filter structures.

In addition, the program synthesizes, designs, and analyzes recursive (infinite impulse response) and nonrecursive (finite impulse response) digital filters, active RC filters, switched-capacitor filters, and crystal filters. All filters can be up to fiftieth degree. The structure of S/FILSYN is shown in Figure 2. Group delay equalization is offered in two ways: all-pass equalizer sections may be incorporated in the filter or they can be cascaded to the network using the Delay Equalizer subprogram. The program can provide up to 20 sections of all-pass equalization. Element values for the equalizer sections are computed along with the values of the rms delay deviation from constant.

S/FILSYN is conversational with features for both the novice and the experienced filter designer. The interactive nature of the program permits experimentation and movement between program segments. All filters can be analyzed in either the frequency-domain or the time-domain. Results can be listed in tabular form or plotted graphically. In addition, S/FILSYN offers an interface to Super-Compact.

Network Synthesis -- MICRO-AMPSYN

Micro-Ampsyn is a program available for the HP 200 series of desktop computers which is used for synthesizing RF matching networks, including input, output, and interstage circuits. The topology can include up to ten reactive elements. The program allows for the adjustment of the frequency response to include parasitic elements (reactances associated with either the generator or load impedances) and for impedance transformation in order to obtain the desired termination values.

Time-Domain Analysis -- MAMA

MAMA (Measurement And Microwave Analysis) is an HP series 200-based desktop computer program developed by Harold Stinehelfer, Sr which is used to convert frequency-domain data as measured on an HP8409C automatic network analyzer to a time-domain output. The conversion is performed with a Fast Fourier Transform (FFT). The program can also store and retrieve the raw data as measured by the ANA.

The impulse and square-wave response can be analyzed with MAMA's harmonic series analysis driver, and measurements can be subtracted from each other for de-embedding purposes. MAMA behaves very much like a computerized time-domain reflectometer (TDR), and as such can help to pinpoint and characterize individual circuit discontinuities.

There are a variety of programs within the MAMA family, including a directory program, a master catalog program, and a MAMA to Micro-Compact file conversion program.

Layout and Mask Production -- AUTOART

AUTOART is an interactive, two-dimensional drafting program for microwave and RF planar circuits. Autoart translates microwave circuit descriptions, given by physical dimensions, to mask layouts used in the fabrication of hybrid and monolithic microwave and RF integrated circuits.

In the AUTO mode, a Super-Compact circuit file is converted semi-automatically (depending on the type of circuit elements encountered in the circuit file) to a layout. The layout is displayed on the user's terminal. By using NODE mode commands, the designer can edit the design using nodal microwave terminology. Components can be added, deleted, and altered, and with a dual-mode display capability, the displayed layout is redrawn as the changes are

mode. Final changes are performed using the GEOMetric mode. The manipulations are performed directly on the layout using the keyboard or a graphics cursor. This mode allows for the inclusion of registration targets and lettering, if desired. A flowchart for Autoart is shown in Figure 3.
Autoart provides the capability of interfacing directly with precision flatbed plotting tables manufactured by Wild Heerbrugg, with automatic coordinatographs manufactured by Aristo Graphics Corporation, with photoplotters produced by Gerber Scientific Instruments, Inc., with an Applicon 860 drafting machine, and through an IGES post-processor with pattern generators and mechanical drafting equipment manufactured by CalComp and ComputerVision.

CONCLUSION

In case anyone reading this article is not completely convinced of the necessity for computer-aided design software for microwave and RF engineering, consider the 10MHz to 2.0GHz feedback amplifier amplifier design that has been used for the past few years in the 'Microwave Circuit Design' short course taught by Les Besser, Bob Wenzel and Steven March. The circuit is shown schematically in Figure 4. Figure 5 gives the initial response as calculated with Super-Compact and using educated guesses for the element values. After optimization for 10dB flat gain, the final response is shown in Figure 6. In another example, Dr. John Bandler of Optimization System Associates has applied gradient optimization to a twelve channel multiplexer to achieve the performance shown in Figure 7. The filters are assumed lossy and dispersive; waveguide junctions are assumed to be non-ideal throughout.

Computer-aided design tools, either purchased programs or in-house developments, are not supposed to replace the design engineer, only augment his capabilities. Remember, no matter how sophisticated a computer program is, the user must at least have a reasonable idea of what is going on in the design process to be able to utilize any computer program to its fullest.

REFERENCES

[1] --, Hewlett-Packard Software Solutions Catalog, Hewlett-Packard Company, Computer Supplies, P.O. Box 60008, Sunnyvale, CA 94088, telephone (800) 538-8787 or (408) 738-4133, part number 5953-5880, $10.95

[2] --, Hewlett-Packard Computer User's Catalog, Autumn/Winter 1984/1985 edition, Hewlett-Packard Company, Computer Supplies, P.O. Box 60008, Sunnyvale, Ca. 94088, telephone (800) 538-8787 or (408) 738-4133, part number 5953-2450 (D), $FREE$

[3] --, Third Party Technical Software Solutions Catalog, Hewlett-Packard Company, Computer Supplies, P.O. Box 60008, Sunnyvale, CA 94088, telephone (800) 538-8787 or (408) 738-4133, part number 5957-4315, $25.00

[4] --, PDP-11 Software Sourcebook, Second Edition, Vol. I: Applications Software and Vol. II: Systems Software, Digital Equipment Corporation, 146 Main Street, Maynard, MA 01754, telephone (800) 343-8321, part number EB25775DP, $17.00

[5] --, International Directory of Software, Computing Publications, Inc., Princeton Forrestal Center, 101 College Road East, Princeton, NJ 08540

[6] --, Computer-Aided Design Directory/Buyer's Guide, Butterworth Publishers, Ltd., Surrey, United Kingdom, 1982

[7] B. Manz, The Software Selector: A Guide to RF and Microwave CAD, Microwaves and RF, Vol. 23, Dec. 1984, pp. 70-95

[8] S.L. March, Microwave Circuit Layout: A Dynamic Plot Emerges, Microwaves and RF, Vol. 23, Dec. 1984, pp. 59-62, 158-161

[9] L. Besser, High-Frequency CAD Comes out of the Lab and onto the Shelf, Microwaves and RF, Vol. 23, Dec. 1984, pp. 65-69

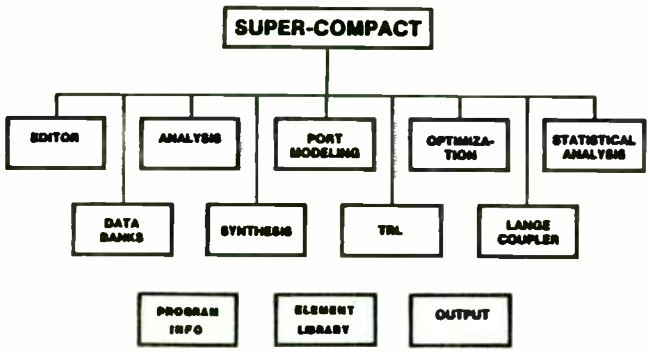
[10] D. Trubitt and A. Podell, Desktop CAD Offers Alternatives to the Microwave Design Engineer, Microwave System Design Handbook, July 1984, pp. 73-91

[11] K. Akima, Software Streamlines RF Engineering, RF Design, Vol. 8, Jan. 1985, pp. 22-23

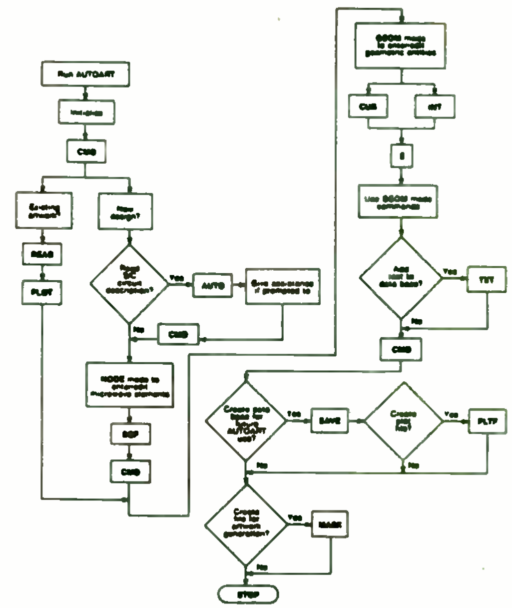
[12] L. Besser, Microwave Circuit Design Enhanced Through Use of CAD/CAM, Microwave System Design Handbook, July 1984, pp. 55-69

[13] J. Browne, Rounding up the Latest Computer Design Programs, Microwaves and RF, Vol. 23, Aug. 1984, pp. 161-162

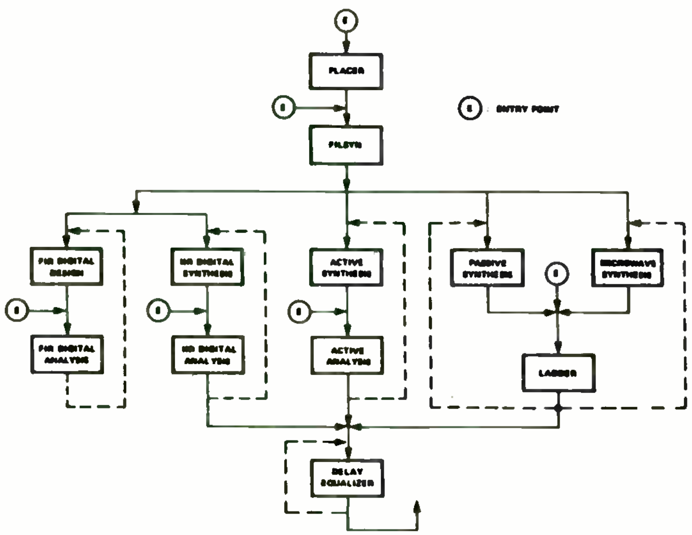
[14] R. Mittra, 'Thou Shalt not Use the Computer Blindly' - A Golden Rule for the Computer-Oriented Microwave Practitioner, Proc. 3rd European Microwave Conf., Sept. 1973, paper B8.1



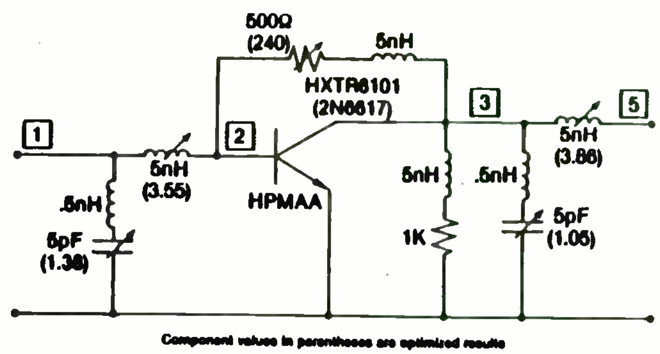
ELEMENTS OF SUPER-COMPACT



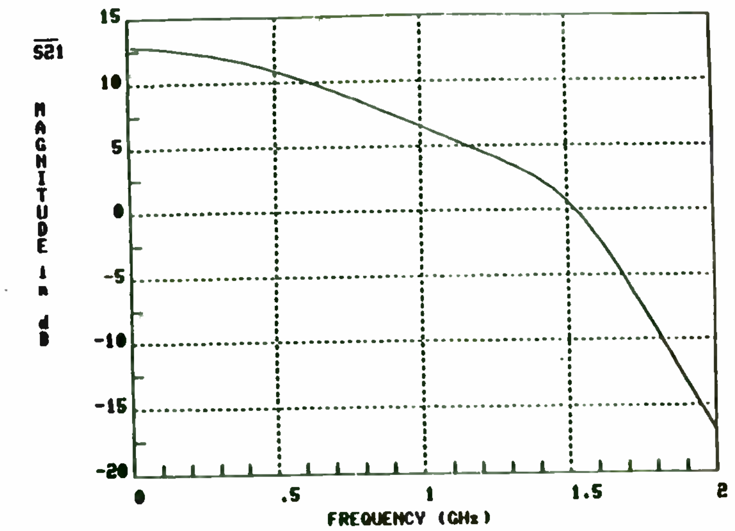
AUTOART FLOWCHART



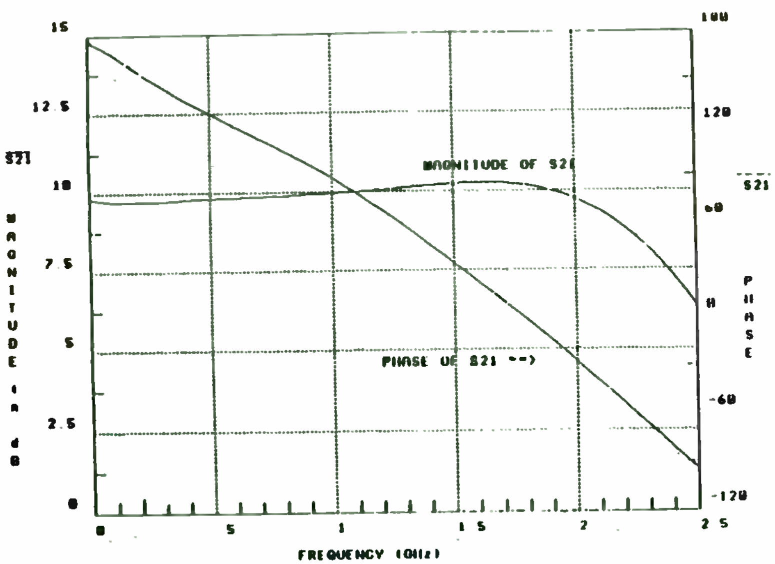
S/FILSYN STRUCTURE



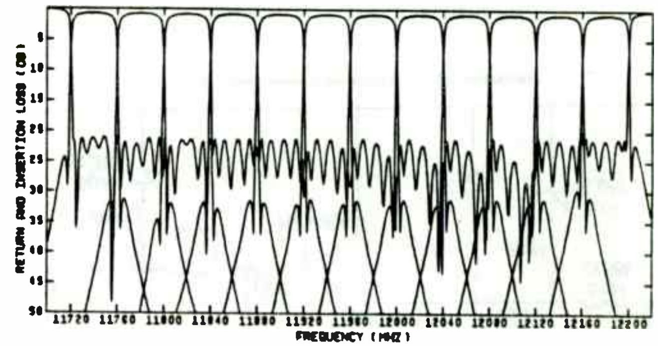
FEEDBACK AMPLIFIER SCHEMATIC



INITIAL AMPLIFIER RESPONSE



OPTIMIZED AMPLIFIER RESPONSE



Response of a 12-channel multiplexer

MAINFRAME PERFORMANCE FOR RF/MICROWAVE CAE ON PERSONAL COMPUTERS

By William H. Childs

EEsof, Inc.
31194 La Baya Drive
Westlake Village, CA 91362

Reprint of a talk at the RF Technology Expo
January 24, 1985
Disneyland Hotel
Anaheim, CA

ABSTRACT

A new RF/microwave CAE program, Touchstone, has delivered remarkable performance on microprocessor-based personal computers. The factors that influenced program design and architecture are reviewed. Benchmarks for interactive response and optimization are reported. These results, obtained on an IBM AT, indicate performance directly comparable to a VAX-based microwave CAE program on a single-user VAX 750. As a result, the cost of a professional RF/microwave design station is lowered by nearly an order of magnitude.

Introduction

Microwave CAE has become the accepted way to perform modern RF/microwave design. Factors that have generated this acceptance are the non-tuneability of thin-film and monolithic integrated circuits and the growing availability of microwave CAE programs on mainframe computers.

Microwave CAD for mask layout has considerably less acceptance, but there is a growing awareness of its utility and desireablility. Commercial and private mainframe computer programs provided a useful but costly and non-interactive design environment. Frequently, organizations would invest considerable sums of money (typically $300,000 for a VAX 780 plus software), only to find excessive response times as more and more users weighted down the time-share system. Though it isn't widely understood, a truly professional RF/microwave design chair in this environment costs close to $100,000.

Workstations and the Personal Computer as a Work Station

The microprocessor-based computer that gains widespread acceptance (the Apple II, Z80 CPM computers, and the IBM PC) approaches the status of a commodity. Manufacturing costs become so low that the hardware costs become incidental to the task at hand. It is fairly easy, today, to purchase an Intel 8088/8087-based computer with 0.5 Megabyte of Random-Access Memory (RAM) and a graphics display for less than $2500.

The acceptance of microprocessor-based computers had been growing since the mid-1970's. These early machines had limited applicability to engineering use, since the address space was limited to 64 Kbytes and floating-point calculations had to be performed in software. The advent of the de facto standard created by the IBM PC based on the Intel 8088 microprocessor changed this situation since the 8088 can address 1 Mbyte of RAM and the Intel 8087 math coprocessor can work in conjunction with the 8088/8086 microprocessor to perform floating point calculations in hardware.

In addition, IBM's presence stabilized the marketplace and encouraged software entrepreneurs to introduce new and frequently innovative computer programs.

The properties of generic engineering workstations are summarized in Table 1. As recently as four years ago, having such hardware dedicated to a single user was only becoming conceivable.

Table 1. Properties of an Engineering Workstation

- Dedicated to a Single User
- 0.5 Mbyte RAM or More
- Memory Mapped Screens
- Bit-mapped Graphics
- Dedicated Input Devices
  --Keyboard
  --Mouse
  --Pad
- Bus Architecture
- Provision for Floating Point Calculations
- Capable of Being Networked

Today, given the commodity-like nature of some computers, it is not only accepted but cost-effective. Some specific examples of engineering workstations are given in Table 2.

Table 2. Some Specific Engineering Workstations

IBM PC and Compatibles

IBM PC-AT

HP 200 Series

Sun

Apollo

VAX Stations

Daisy Logician

A New RF/Microwave CAE Program

Touchstone was designed to meet four general design objectives. (See Table 3.) The first was to offer a full and complete set of scientific features. The second was to achieve a new level of interactiveness and responsiveness. And the third was to design the program to be easily integrated into an RF/microwave CAD program (mask layout and drawing preparation).

The fourth goal was to integrate the program with the lab environment so that it could interact in real time with measurements begin taken in the lab. It is believed that in the next 2-3 years, the RF/microwave industry will be integrating the CAE, CAD, and CAT functions into a single-user workstation. The impetus for this will come primarily from reduced design times and intensified competition.

Table 3. General Design Objectives for a Modern RF/Microwave CAE Program

- Complete Set of Scientific Features
- Highly Interactive
- Integrated with RF/microwave CAD
- Integrated with RF/microwave CAT

Scientific Features

The needed technical features of a modern RF/microwave CAE program are summarized in Table 4. The early RF/microwave CAE programs of the middle to late 70's were box programs. The advent of complex feedback and distributed amplifier designs necessitated true nodal programs--so that nodal analysis is listed first on the feature list.

A set of circuit elements that is rich in physical models allows the modern RF/microwave designer to achieve success with his first mask.

Optimizers are taken for granted today and their value is greatly enhanced by general goal and weight specification as a function of frequency. Likewise, modern design requires noise figure to be calculated with loss present and for arbitrary topologies.

A modern RF/microwave CAE program should allow the engineer to perform calculations on defined networks; the phase shifter is the classical example where "ON" and "OFF" conditions can be described in the net list and then a derived network defined that displays the true quantity of interest--the net phase shift.

Finally, a modern CAE program should provide the capability for yield predictions given expected statistical variations in production components.

Table 4. Scientific Features of a Modern RF/Microwave CAE Program

- Nodal Analysis
- Complete and Accurate Model Set
- Optimizer with General Goal Specification
- Versatile Measurements and Output
- Noise Figure Calculations
- User-defined Variables
- Derived Networks
- Yield Predictions

User Features

The required user features of a modern RF/microwave program are summarized in Table 5. It is emphasized that this desired environment can best be obtained on a workstation.

For example, the rapid presentation of graphics or screens is a greatly facilitated with the bit-mapped screens of a workstation, as opposed to the traditional graphics terminal attached to a mainframe over a limited rate serial link on a mainframe computer.

Likewise, user inputs on a workstation can easily be processed to interrupt optimizers and the like; such interrupts are difficult to process on a time-share system.

It is a widely held belief that interactiveness in an engineering program contributes to the engineer's insight into the design, and that a lack of interactiveness contributes to an over-dependence on optimizers.

Speed and the resulting responsiveness to the user have a first-order effect on the cost of the hardware needed to operate the program. Very early in the development of Touchstone, speed tests were performed on the IBM PC; if the outcome of those tests had not been satisfactory, a more powerful engine would have been selected.

Table 5. User-Friendly Features of a Modern RF/Microwave CAE Program

- Speed and Responsiveness
- Interactiveness
- Instrument-like Environment
- Logical and Consistent Syntax
- Lab-like Graphics
- Menus
- Help Messages
- Reliability

Implementation Specifics

Touchstone was initially targeting on computers using the Intel 8088/8086 microprocessor along with the Intel 8087 math coprocessor. The target operating system was the highly popular Microsoft MS-DOS system.

The secondary target machine was the Hewlett-Packard 200 series computers based on the Motorola 68000 microprocessor. The program is written in Pascal with key portions in assembly language. One of the reasons that Pascal was selected was to facilitate porting of the program onto the HP 200 series computer, under the HP Pascal operating system.

The selection of compilers is a crucial one in realizing the ultimate speed of the computer. The realization of the program under MS-DOS is sufficiently robust that it easily ports onto all the IBM compatibles and only minimum development is required to port onto any MS-DOS machine such as the Texas Instruments Professional Computer.

Speed and How to Find It

RF/microwave CAE programs are rich with floating point calculations. Careful attention to the concepts in Table 6 allow surprising speed to be obtained in a CAE program. Having compute-intensive portions of code with no jumps keeps the instruction queue of the 8088/8086 microprocessor filled.

It is vital that all floating point operations be handed over to the 8087 coprocessor for greatest speed. It is important that the machine language generated by the compiler be optimum in compute-intensive portions of the programs. If it isn´t, steps must be taken to correct the situation.

It is necessary that the program be all RAM resident. Touchstone has achieved such outstanding speed because it possesses memory-intensive algorithms. This set of algorithms is called a circuit compiler--it is best described as a method of translating large data states into a technique for performing nodal analysis.

Finally, speed is obtained by the implementation of program architecture in both general and specific ways. For example, Touchstone separates the tasks of file processing, circuit simulation, and display of output. Everything possible is done to prevent unnecessary performance of these functions.

For example, if the circuit file has been processed and only a circuit data value is changed, then the file will not be processed. Likewise, if a new display is requested but there is no need to simulate the circuit, then it is not reprocessed. The stored results are merely displayed in a different manner.

A more specific example of the interaction of program speed and architecture comes in the calculation of physical models. It is a common practice in RF/microwave CAE programs to perform the whole physical model calculation at each frequency. In Touchstone, however, the physical model is divided into static and frequency-dependent parts. The static part is calculated only once and the frequency-dependent portions are calculated through the balance of the frequency sweep.

As a result, in a typical problem, the speed difference between using ideal transmission lines and microstrip transmission lines is slight.

Table 6. How to Achieve Speed on Intel 8088/8086/8087-based Computers

- In-line Code
- Tailored for 8087 Coprocessor
- Key Portions in Assembly Language
- Close Understanding of Machine Language Generated by Compiler
- Close Understanding of Linker
- All RAM Resident
- Circuit Compiler--Memory-Intensive Algorithms
- Program Architecture - General
- Program Architecture - Specific

Benchmarks

A schematic for a simple 6-18 GHz amplifier is shown in Figure 1A and the Touchstone circuit file is shown in Figure 1B. The performance results for this example are summarized in Table 7. The results are separated into the components of file processing, simulation, and output display because with Touchstone, the engineer can separate these effects.

In particular, the Tune mode in Touchstone eliminated file processing and shortens graphics output to about 0.2 seconds.

The effect to the designer of tuning this circuit is almost like a real-time response and fully comparable to his lab tuning experience.

This amplifier is put into a balanced configuration in Figure 2 and the performance results are summarized in Table 8. This problem was simulated using a major VAX-based program on a VAX 780 (accelerator plus 2 Megabytes RAM). It was reported that output did not begin until 8 seconds after the command; output presentation was limited by the aerial port to an additional 4 seconds.

Optimization times for a compute intensive filter diplexer are shown in Table 9. The comparison is for Touchstone on an IBM-AT and a VAX based mainframe program on a VAX 750. The notion of trials in both programs is the same. The Touchstone IBM-AT result of 1352 seconds compares very favorably with the VAX 750 based program result of 1073 seconds.

Figure 1A.

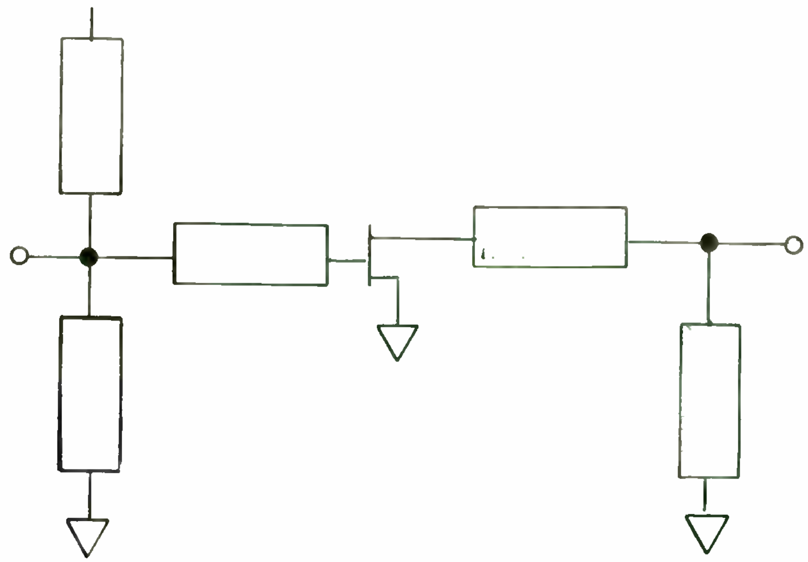


Figure 1B.

```
VAR
  F1=18
  Zmax=100
CKT
  TLOC    1  0  Z^Zmax E\64  F^F1
  TLSC    1  0  Z=70   E=141 F^F1
  TLIN    1  2  Z^Zmax E=23  F^F1
  S2PA    2  3  0      NEC70000
  TLIN    3  4  Z^Zmax E=16  F^F1
  TLSC    4  0  Z#40 51 100  E=70  F^F1
  DEF2P   1  4  AMP
OUT
  AMP DB[S21]  GR1
  AMP DB[S11]  GR2
  AMP DB[S22]  GR2
  AMP DB[S12]  GR2
  AMP ANG[S21] GR3
  AMP S11
  AMP S22
FREQ
  SWEEP  6  18   2
GRID
  GR1     5  9   .5
  GR2   -30  0    5
  GR3   -180 180 30
OPT
  RANGE  6  18
  AMP DB[S21] < 8
  AMP DB[S21] > 7.5
```

Table 7. Touchstone Performance of the Example in Figure 1.

Example Run on IBM AT

File Processing	< 2 seconds
Circuit Simulation	< 0.5 seconds
Graphics Output	< 0.5 seconds

Figure 2A.

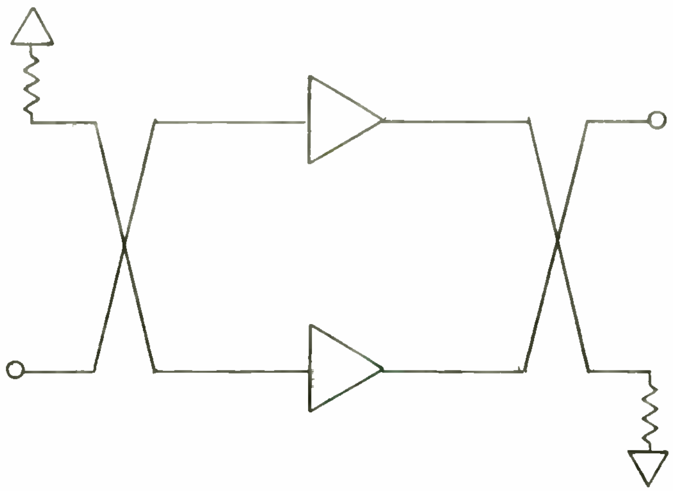


Figure 2B.

```
!BAMP6-18:  Balanced amplifier, 6 to 18 Ghz.
VAR
  F1=18
  Zmax=100
CKT
  TLOC   1  0  Z^Zmax E\64  F^F1
  TLSC   1  0  Z=70   E=141 F^F1
  TLIN   1  2  Z^Zmax E=23  F^F1
  S2PA   2  3  0      NEC70000
  TLIN   3  4  Z^Zmax E=16  F^F1
  TLSC   4  0  Z#40 51 100  E=70  F^F1
  DEF2P  1  4  AMP
  CLIN   1  2  3  4   ZE=144.3 ZO=17.3 E=90 F=12
  DEF4P  1  2  3  4   CPLR
  CPLR   1  2  3  4
  CPLR   5  6  7  8
  AMP    2  5
  AMP    4  7
  MATCH  3
  MATCH  6
  DEF2P  1  8  BALAMP
  BALAMP 1 2
  S2PB 2 3 0 FILT12
  DEF2P 1 3 FAMP
 OUT
  AMP     DB[S21] GR1
  BALAMP  DB[S21] GR1
  AMP     DB[S11] GR1
  BALAMP  DB[S11] GR1
  FAMP    DB[S21] GR2
  FAMP    DB[S11] GR2
  AMP     DB[S21] GR2
 FREQ
  SWEEP   6  18  2
 GRID
  RANGE   2  22  1
  GR1    -20 10  2
  GR2    -20 10  2
```

Table 8. Touchstone Performance of Figure 2

Example Run on IBM AT

File Processing	< 4 seconds
Circuit Simulation	< 1.0 seconds
Graphics Output	< 0.5 seconds

Table 9. Comparison of Touchstone and VAX-based Program for a Filter Diplexer

	Touchstone 1.2	VAX-based Program
Test Machine	IBM AT	VAX 750
RAM	512 Kbyte	2.0 Mbyte
Time for 250 Trials	1352 seconds	1073 seconds
System Cost	$13,500	$150,000

Conclusions

The engineering workstation environment is so attractive that it is believed that over a very short period, almost all engineering work will be done on a workstation instead of the traditional mainframe with attached terminals. (This transition is already very advanced among digital designers.) A modern RF/microwave CAE program should be tailored to take advantage of the attractive features of this workstation environment. By careful attention to program architecture and by the application of innovative nodal analysis algorithms, it is possible to exceed industry standard performance on VAX class machines in interactive design and be fully competitive with CPU-intensive performance such as optimization.

These results are obtained by a new RF/microwave program, Touchstone, on an IBM AT, PC, and compatibles as well as HP 200 series workstations.

Not only do these results decrease the cost of a professional RF/microwave design chair by nearly an order of magnitude, they show that personal computers can be used as a truly professional RF/microwave workstation.

COMPUTER-AIDED DESIGN OF PHASE-LOCKED LOOPS

Ulrich Rohde

SUMMARY

The design of phase-locked loops, while following certain mathematical guidelines, has largely been done empirically. In many cases, important data like reference suppression, locked-up time, or phase noise have been determined experimentally rather than predicted mathematically. This paper shows how modern computers can be used to reduce time and increase accuracy in calculating loop performance.

#1 Design Disk

Let us assume a phase-locked loop synthesizer operating from 110 to 210 MHz has to be designed. A reference frequency of 10 KHz is given and the tuning diode with a minimum capacitance of 6 pF and a maximum value of 6.0 pF is chosen. For reasons of performance, we will choose a type 2 third-order loop as developed and described by Andy Przedpelski and reprinted in my book, DIGITAL PLL FREQUENCY SYNTHESIZERS, THEORY AND DESIGN, published by Prentice-Hall, ISBN #0-13-21-4239-2.

The phase noise calculations use Leeson's model and the following equation with the following abbreviations:

$$\mathcal{L}(f_m) = 10 \log_{10}\left(SQR\left(\left(\tfrac{1}{2}\left[1+\frac{1}{\omega_m^2}\left(\frac{\omega_o}{2Q_{load}}\right)^2\right]\frac{FkT}{P_{sav}}\left(1+\frac{f_c}{f_m}\right)^2 + \frac{(SQR(4kT_oRdF)K_o\sqrt{2})^4}{f_m}\right)\right)\right)$$

ω_n = 2π x offset

ω_o = 2π center frequency

Q_{load} = loaded figure of merit Q of the tuned circuit

F = noise figure

kT = 4.2E-21 at 300°K (room temperature)

f_c = flicker frequency of semiconductor

f_m = frequency offset

R = equivalent noise resistor of tuning diode

dF = integration bandwidth

K_o = oscillator voltage gain

Please note that the equation determining the mean square value has to use the power 4 for the second term.

Using the mathematical model on pages 400-404 describing oscillator amplitude output, the energy available from the oscillator depending upon the normalized Fourier coefficients can be determined.

The lock-up time of a phase-locked loop can be defined in many ways. In the digital loop I prefer to define it by separating frequency lock or pull-in and phase lock. Both numbers have to be added. To determine the pull-in time, a new statistical approach has been used as described on pages 53-55 of my book defining a new gain constant

$$K' = V_B/\omega_o$$

with V_B = supply voltage and ω_o = frequency offset. The phase-locked time has to be determined from the Laplace transform of the transfer function shown on pages 32-36 of my book.

Actual Design

A set of programs was written around the equations mentioned above and is being marketed under "PLL Design Kit." This program is used in the actual design using the above-mentioned equations and/or assumptions. Table 1 shows the input data and information of lock-up time and reference suppression.

```
INPUT DATA:

REFERENCE FREQUENCY IN Hz =1000
NATURAL LOOP FREQUENCY IN Hz =50
PHASE DETECTOR GAIN IN V/rad =1
VCO GAIN CONSTANT IN Hz/V =1.00E+07
DIVIDER RATIO =160000
VCO FREQUENCY IN Hz =          1.6E+8
PHASE MARGIN IN deg =45

THE LOCK-UP TIME CONTANT IS:      1.73E-02 sec
REFERENCE SUPPRESSION IS:            44.4 dB
```

Table 1

Based on the frequency range and the tuning diode, a wideband VCO is required. The computer program interactively determines the values shown in Table 2.

VCO DESIGN

CALCULATION OF TUNING RANGE:
Fmin= 110 Mhz Fmax= 210 Mhz
CENTER RANGE IS 160 Mhz TUNING RATIO = 1.909
Cmin (at Vmax) OF TUNING DIODE= 6 pF Cmax (at Vmin) OF TUNING DIODE= 60 pF
FET CHOSEN:
CISS = 2 pF ; TRANISTOR IS OPERATED AT Id= 12 mA/12V;Gm= 17 mS
CUT-OFF FREQUENCY OF FET = 1.4 Ghz ; OUTPUT POWER IS 28.8 mW OR 15 dbm
BOARD STRAY CAPACITANCE = 1.2 Pf
Cmin OF DIODE COMBINATION= 5.44 pF; Cmax OF DIODE COMBINATION= 29.6 pF
COUPLING CAPACITOR Cs= 58.5 pF ; REQUIRED INDUCTANCE IS .0628 uH
FEEDBACK CAPACTIOR OF 1 pF CHOSEN
PARALLEL TRIMMING CAPACITANCE CT = 2.5 pF

Table 2

Depending upon the frequency range, the PLL Design Kit has four different recommended circuits. It covers a narrowband and a wideband VCO for lumped elements and a halfwave and quarterwave oscillator for UHF. The circuit configuration shown in Figure 1 has been chosen:

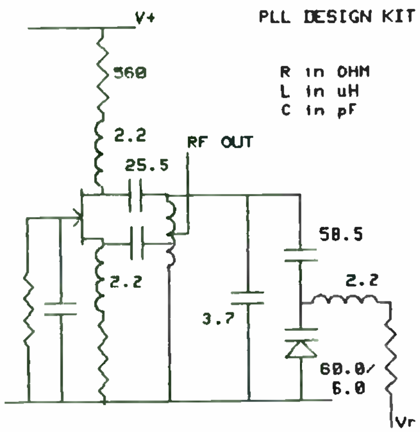


Figure 1

Based on the description of the circuit components, the computer then calculates the SSB phase noise as shown in Tables 3 and 4:

SSB PHASE NOISE CALCULATION

LO-POWER = 0 dBM, LO NF= 10 dB
The rms noise voltage per sqr(1 Hz) bandwidth =1.30E-08 V
Rn= 10000 Ohm
F= 10 dB
VCO GAIN (Hz/V)=1.00E+07
FREQUENCY OFFSET IN Hz= 10000
CENTER FREQUENCY = 160 Mhz
LOADED Q = 120
FLICKER FREQUENCY IS 150 Hz

The ssb phase noise in 10000 Hz offset is -100.76 dBc/HZ

Table 3

SSB NOISE TABLE

FREQUENCY (Hz)	PHASE NOISE (dBc/Hz)
1.00E+00	-25.50
3.16E+00	-40.44
1.00E+01	-55.25
3.16E+01	-69.70
1.00E+02	-83.31
3.16E+02	-95.60
1.00E+03	-106.68
3.16E+03	-117.09
1.00E+04	-127.22
3.16E+04	-137.26
1.00E+05	-147.19
3.16E+05	-156.41
1.00E+06	-162.17
3.16E+06	-163.58
1.00E+07	-163.75

Table 4

As tables in many cases are not very revealing, the phase noise is then plotted. This is shown in Figure 2:

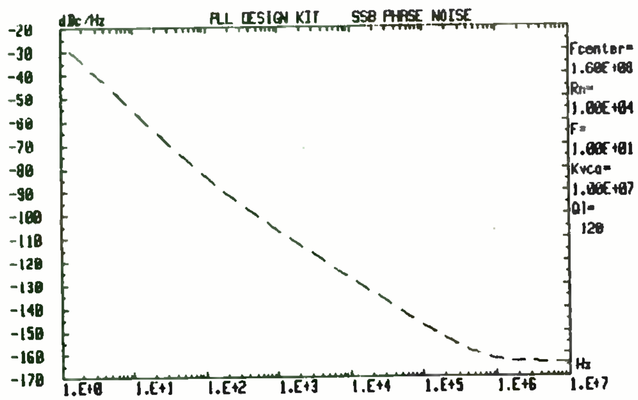


Figure 2

The free-running oscillator phase noise close-in has to be poor and can be improved by embedding it in a type 2 third-order loop, as shown in Figure 3:

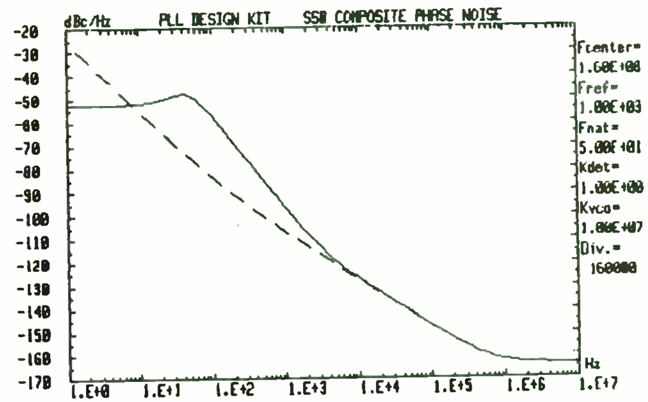


Figure 3

It is apparent that the close-in noise now has become better and that the phase noise of both graphs meets at approximately 4 KHz. The type 2 third-order loop, if properly designed, gives excellent lock-up time and reference suppression. This can be seen from the plot shown in Figure 4:

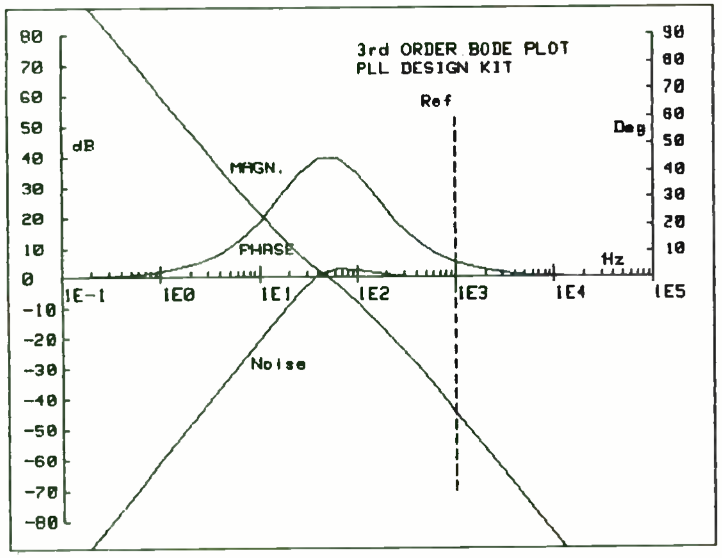


Figure 4

The phase-lock time can be obtained from Table 5. It is good practice to define the lock-up time as the point where the phase error is less than 1°. Based on the 50 Hz loop frequency, this value is 32 mS. For the total lock-up time, we have to add the frequency lock-up time of 17.3 mS and an approximate 50 mS lock-up time, for both frequency and phase lock is required:

LOCK-IN FUNCTION :

TIME/s	PHASE DET.DEV./deg
0	3.60E+02
.0016	2.58E+02
.0032	1.16E+02
.0048	2.94E+00
.0064	-7.36E+01
.008	-1.12E+02
.0096	-1.20E+02
.0112	-1.07E+02
.0128	-8.40E+01
.0144	-5.95E+01
.016	-3.79E+01
.0176	-2.12E+01
.0192	-9.55E+00
.0208	-2.37E+00
.0224	1.45E+00
.024	2.99E+00
.0256	3.17E+00
.0272	2.67E+00
.0288	1.93E+00
.0304	1.22E+00
.032	6.45E-01
.0336	2.45E-01
.0352	3.35E-03
.0368	-1.17E-01
.0384	-1.57E-01
.04	-1.49E-01
.0416	-1.19E-01
.0432	-8.34E-02
.0448	-5.13E-02
.0464	-2.65E-02

Table 5

Finally, Figure 5 shows the active integrator for the type 2 third-order loop.

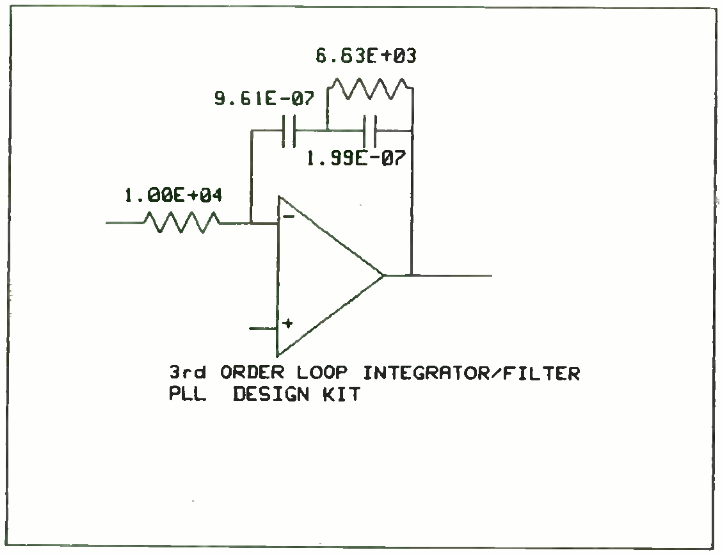


Figure 5

SUMMARY

Having a sufficiently accurate mathematical model allows the development of highly interactive computer programs taking some of the non-linearities of the phase-locked loop into consideration. The above-described analysis program is being used by many domestic and international companies with great success and reproducibility.

REFERENCES

[1] A. Przedpelski, "Analyze, Don't Estimate Phase-Locked Loop Performance of Type 2 Third Order Systems," ELECTRONIC DESIGN, May 10, 1978.

[2] A. Przedpelski, "Optimized Phase Locked Loop to Meet Your Needs or Determine Why You Can't," ELECTRONIC DESIGN, September 1978.

[3] A. Przedpelski, "Suppress Phase Locked Loop Sidebands without Introducing Instability," ELECTRONIC DESIGN, September 1978.

ABSTRACT

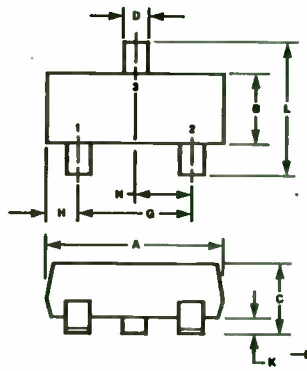
This paper describes the mounting techniques and thermal considerations for various plastic molded packages used in RF low power applications. Infrared scan data, taken under RF operating conditions, is presented to empirically demonstrate the tradeoffs of mounting techniques versus thermal performance. A thermal resistance model is derived to predict the thermal performance based on the mounting technique design and packaged device thermal specifications.

INTRODUCTION

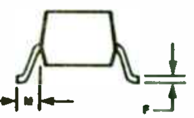
This paper compares four plastic molded packages frequently used in RF applications:

1. SOT23 (see Figure 1)
2. SOT143 (see Figure 2)
3. Macro-X (see Figure 3)
4. PowerMacro (see Figure 4).

The PowerMacro package is analyzed for thermal resistance comparing various mounting techniques. The experimental RF scan results are compared to a basic thermal resistance model and conclusions are made regarding the validity of the model.



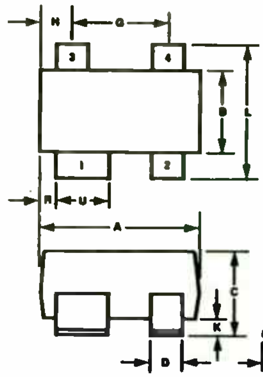
DIM	MILLIMETERS MIN	MILLIMETERS MAX	INCHES MIN	INCHES MAX
A	2.80	3.04	0.1102	0.1197
B	1.20	1.40	0.0472	0.0551
C	0.85	1.20	0.033	0.0472
D	0.37	0.45	0.0150	0.0177
F	0.085	0.130	0.0034	0.0051
G	1.78	2.04	0.0701	0.0807
H	0.51	0.60	0.0200	0.0236
K	0.10	0.25	0.0040	0.0098
L	2.10	2.50	0.0830	0.0984
M	0.45	0.60	0.0180	0.0236
N	0.89	1.02	0.0350	0.0401



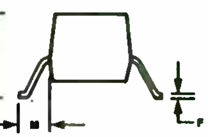
STYLE 6:
PIN 1. BASE
2. EMITTER
3. COLLECTOR

STYLE 7:
PIN 1. EMITTER
2. BASE
3. COLLECTOR

FIGURE 1 — CASE OUTLINE DRAWING FOR SOT23 PACKAGE



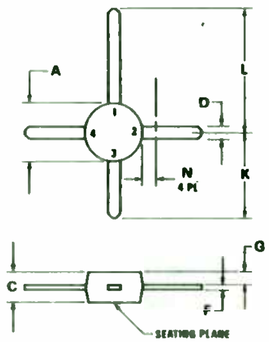
DIM	MILLIMETERS MIN	MILLIMETERS MAX	INCHES MIN	INCHES MAX
A	2.79	3.04	0.110	0.120
B	1.20	1.39	0.047	0.055
C	0.76	1.01	0.030	0.040
D	0.36	0.46	0.015	0.018
F	0.08	0.15	0.003	0.006
G	1.78	2.03	0.070	0.080
H	0.51	0.60	0.020	0.024
K	0.03	0.07	0.001	0.003
L	2.11	2.48	0.083	0.098
M	0.45	0.60	0.018	0.024
R	0.10	0.17	0.004	0.007
U	0.76	0.89	0.030	0.035



STYLE 1:
PIN 1. COLLECTOR
2. EMITTER
3. EMITTER
4. BASE

STYLE 2:
PIN 1. SOURCE
2. DRAIN
3. GATE 1
4. GATE 2

FIGURE 2 — CASE OUTLINE DRAWING FOR THE SOT143 PACKAGE



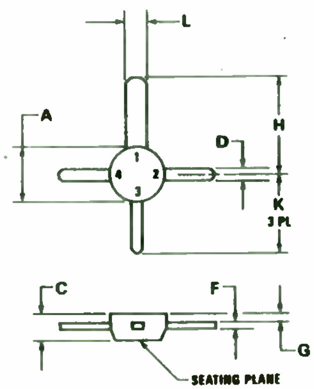
DIM	MILLIMETERS MIN	MILLIMETERS MAX	INCHES MIN	INCHES MAX
A	4.44	5.21	0.175	0.205
C	1.90	2.54	0.075	0.100
D	0.84	0.99	0.033	0.039
F	0.20	0.30	0.008	0.012
G	0.76	1.14	0.030	0.045
K	7.24	8.13	0.285	0.320
L	10.54	11.43	0.415	0.450
N	-	1.65	-	0.065

STYLE 1:
PIN 1. DRAIN
2. SOURCE
3. GATE 1
4. GATE 2

STYLE 2:
PIN 1. COLLECTOR
2. EMITTER
3. BASE
4. EMITTER

NOTE:
DIMENSION D NOT APPLICABLE IN ZONE N

FIGURE 3 — CASE OUTLINE DRAWING OF THE MACRO-X PACKAGE



DIM	MILLIMETERS MIN	MILLIMETERS MAX	INCHES MIN	INCHES MAX
A	4.44	5.21	0.175	0.205
C	1.90	2.54	0.075	0.100
D	0.84	0.99	0.033	0.039
F	0.20	0.30	0.008	0.012
G	0.63	1.01	0.025	0.040
H	8.84	9.72	0.348	0.383
K	7.24	8.13	0.285	0.320
L	2.46	2.64	0.097	0.104

STYLE 1:
PIN 1. DRAIN
2. SOURCE
3. GATE 1
4. GATE 2

STYLE 2:
PIN 1. COLLECTOR
2. EMITTER
3. BASE
4. EMITTER

FIGURE 4 — CASE OUTLINE DRAWING OF THE POWERMACRO PACKAGE

Macro-X and PowerMacro are trademarks of Motorola.

MOUNTING CONSIDERATIONS FOR RF LOW POWER PLASTIC PACKAGES
HARRY J. SWANSON MOTOROLA, INC. 1/85

COMPARISON OF PACKAGES

Many RF low power plastic packages are available to the designer for RF applications. Table 1 compares the properties of four of these packages:

TABLE 1 — PACKAGE COMPARISON

Package	Collector Lead Thickness (mils)	Collector Lead Width (mils)	Leadframe Material Composition	Max Power Dissipation at T_J (max) = 150°C	Thermal Resistance θ_{JA}* (°C/W)	Thermal Resistance θ_{JC}** (°C/W)
SOT-23	3.4 to 5.1	1.5 to 17.7	Alloy 42	200 mW	625	—
SOT-143	3.0 to 8.0	30 to 35	Alloy 42	350 mW	357	—
Macro-X	8.0 to 12.	33 to 39	Copper	2.0	—	40
PowerMacro	8.0 to 12.	97 to 104	Copper	3.0	—	25

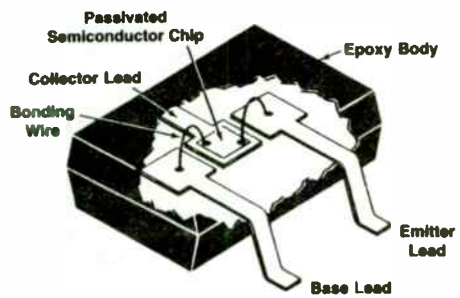
* T_A = 25°C WITH DEVICE MOUNTED IN FREE AIR.
**T_C IS SPECIFIED ON THE DEVICE DATA SHEET. THE PACKAGE IS MOUNTED TD PROVIDE CASE TEMPERATURE LESS THAN OR EQUAL TO THE GIVEN T_C AT THE GIVEN POWER DISSIPATION, P_D.

DESCRIPTION OF TYPICAL RF LOW POWER PACKAGES

The SOT23 surface mount package is similar to the SOT 143 except for the number of common leads and the width of the collector lead. Figure 5 is a cut away view showing the internal construction of the SOT23 package. The package consists of a molded Alloy 42 leadframe which is nominally 4 mil thick. The SOT23 has 3 leads all the same width of nominally 16 mils while the SOT143 has an additional common lead placed diagonally from the other one. The SOT143 collector lead is twice the width of the SOT23 collector lead. The transistor chip is silicon-gold eutectic die bonded to the coll[illegible] lead and is wirebonded for various pinouts to the other leads. Completion of the assemb.y process is accomplished by molding the leadframe and tin plating the leads.

FIGURE 5

SOT-23 Internal Construction



The PowerMacro package is similar to the Macro-X package except for the wider collector lead. Figure 6 is a cut away view showing the component parts of an epoxy molded copper leadframe which has an 100 mil wide collector lead. A Macro-X has a 35 mil wide collector lead. The transistor chip is silicon-gold eutectic die bonded to the collector lead and is wirebonded in a manner similar to the Stripline Opposed Emitter ceramic package. Completion of the assembly process is accomplished by molding the copper leadframe and tin plating the four leads.

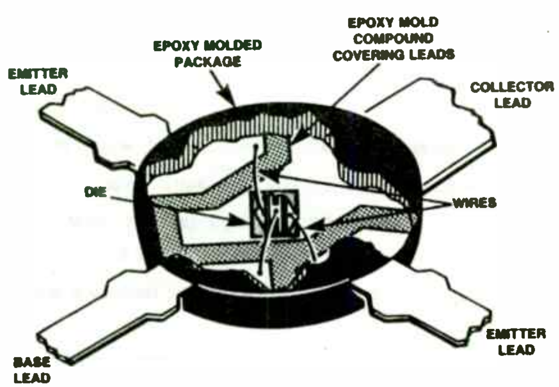


FIGURE 6 — CUT AWAY VIEW OF THE POWERMACRO PACKAGE

GENERAL MOUNTING CONSIDERATIONS

The SOT23 and SOT143 packages are often mounted to both sides of a circuit board with other surface mount components. Usually no heat sink is used and the majority of the heat transfer is accomplished by convection and radiation from the surfaces of the package and the circuit board. It is typical for a vendor to spec the Θ_{JA} in free air at $T_A = 25\ ^oC$.

Electrical and mechanical mounting of the SOT23 and SOT143 surface mount packages are justified by the automated, low cost manufacturing, and generally confines itself to lower power applications. The lack of power dissipation capability may prove to be a limitation when attempting to substitute for a Macro-X package.

The electrical performance of a RF device is difficult to maintain without practicing sound principles of circuit construction and matching techniques. The frequency and bandwidth requirements usually determine the network topology used to ensure good input and output VSWR. Lumped components are often used at lower frequencies, whereas, at higher frequencies, distributed components are effective. It is important to provide good RF grounding on the circuit board close to common leads of the part.

The thermal performance of an RF plastic transistor is dependent upon the mounting technique. The collector lead of the plastic package should offer a low thermal resistance path from the transistor chip. This lead should be utilized effectively to provide the best thermal interface. Since this lead is the output lead of the device, it is necessary to consider RF matching and biasing. The collector lead is usually mounted to a circuit board material such as G-10, glass Teflon, alumina, or beryllium oxide (BeO). The circuit board material is chosen to provide both a low thermal interface resistance and a good output match to the collector.

The thermal parameters of a device are specified on its data sheet. For example, an RF plastic transistor is guaranteed to have a certain thermal performance defined by the total device dissipation, P_D at a certain case temperature, T_C which is measured on the collector lead immediately adjacent to the body of the package. The parameters are defined

Teflon is a registered trademark of DuPont Corp.

for a condition of a maximum junction temperature, T_J of 150 °C. Thus, the task of the designer is to provide a low thermal resistance path for the heat to flow from the collector lead. A practical solution is characterized by a certain temperature rise, ΔT for a given ambient condition with a known amount of heat input or power dissipation, P_D. Table 2 lists the thermal parameters and their electrical analogs used to derived the thermal resistance model.

TABLE 2. THERMAL PARAMETER AND THEIR ELECTRICAL ANALOGS

SYMBOL	THERMAL PARAMETER	UNITS*	ELECTRICAL ANALOG SYMBOL	ELECTRICAL ANALOG PARAMETER
ΔT	TEMPERATURE DIFFERENCE	°C	V	VOLTAGE
H	HEAT FLOW	WATTS	I	CURRENT
θ	THERMAL RESISTANCE	°C/WATTS	R	RESISTANCE
γ	HEAT CAPACITY	WATT-SEC/°C	C	CAPACITY
K	THERMAL CONDUCTIVITY	CAL/SEC-CM °C	σ	CONDUCTIVITY
Q	QUANTITY OF HEAT	CAL	q	CHARGE
t	TIME	SEC	t	TIME
θγ	THERMAL TIME CONSTANT	SEC	RC	TIME CONSTANT

*NOTE THE ONE MAJOR DIFFERENCE IN THERMAL AND ELECTRICAL UNITS — Q IS IN UNITS OF ENERGY, WHEREAS q IS SIMPLY CHARGE. HENCE, H IS IN UNITS OF POWER AND MAY BE EQUATED TO AN ELECTRICAL POWER DISSIPATION.

The thermal resistance model shown in figure 7 leads to the equation [1]

$$T_J = P_D(\theta_{JC} + \theta_{CS} + \theta_{SA}) + T_A$$

Where: P_D = Power dissipated in the transistor in watts

θ_{JC} = Published thermal resistance - junction to case;

θ_{CS} = Interface thermal resistance - case to heat sink

θ_{SA} = Heat sink thermal resistance - heat sink to ambient.

MOUNTING CONSIDERATIONS FOR RF LOW POWER PLASTIC PACKAGES
HARRY J. SWANSON MOTOROLA,INC. 1/85

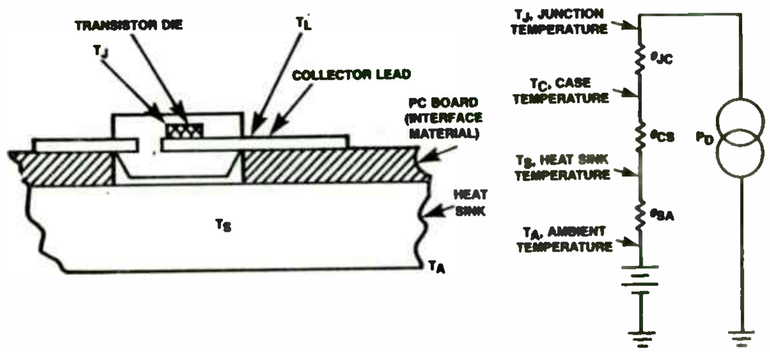


FIGURE 7 — THERMAL RESISTANCE MODEL FOR THE RF LOW POWER TRANSISTOR

The thermal resistance of the transistor (junction to case thermal resistance), θ_{JC} is not constant; it is a function of biasing and temperature as given on the data sheet.

The interface thermal resistance, θ_{CS} is affected by mounting technique and interface material used.

The power dissipation requirements of an RF low power transistor may dictate special mounting considerations to reduce θ_{CS}. The thermal resistance, θ to heat flow between parallel surfaces in a bar of conducting material (see Figure 8) is:

$$\theta = h/KWL = h/KA$$

where: h = Length of thermal path

A = Cross sectional area of thermal path, A = WL

K = Thermal conductivity

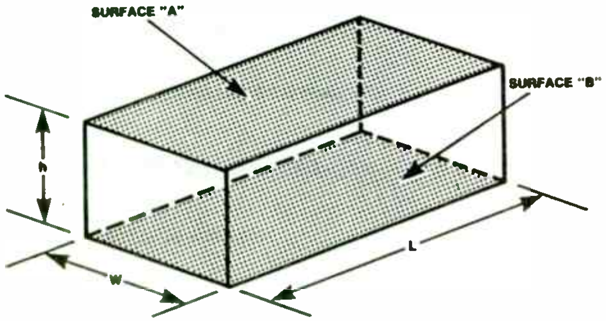


FIGURE 8 — BAR OF CONDUCTING MATERIAL

Thus, the interface thermal resistance is reduced by:

1. Decreasing the length of the thermal path
2. Increasing the cross sectional area of the thermal path
3. Using an interface material of high thermal conductivity (Table A2 lists the thermal properties of various heat sinking materials).

The experimental IR scan results of the MRF553 PowerMacro transistor in the next section compares the interface thermal resistance, θ_{CS} using two circuit board materials.

IR SCAN RESULTS

RF applications of a plastic transistor utilize microstrip techniques to obtain a good case to heatsink, θ_{CS} and to obtain good RF input and output matching.

The experimental results of an IR scan of the MRF553 (1.5W, VHF, 12.5V PowerMacro transistor) represented in Figure 9 shows the comparison of two circuit board materials using three mounting techniques:

1. G-10 circuit board mounted to aluminum heat sink (see Figure A1, circuit used in IR Scan I)
2. Alumina/Copper Socket to aluminum heat sink (see Figures A2 and A3, circuit used in IR Scan II)
3. G-10 circuit board with no heat sink (DC only scan - see Figure A4 for details of circuit used in IR Scan III).

Table A1 in the Appendix lists the typical thermal data from the IR scans for various operating conditions of P_{out}, P_D, T_J (die junction temperature), T_C (collector lead temperature), and T_S (heat sink temperature) and T_A (ambient temperature).

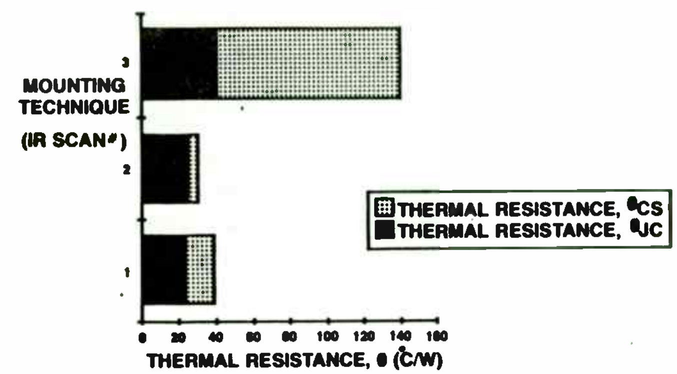


FIGURE 9 — THERMAL RESISTANCE, θ_{JC}, θ_{CS} versus MOUNTING TECHNIQUE

Figure A1 shows the circuit used to provide the thermal data of the MRF553 device mounted to a 62 mil thick G-10 circuit board with 1oz. copper on both sides. The device is soldered to the circuit board which is mounted to a 3 inch x 5 inch x 3/4 inch aluminum heat sink.

Figure A2 shows the circuit used to provide the thermal data for the MRF553 device mounted with an alumina interface. The device is soldered into a specially constructed socket (see Figure A3) which is mounted to a 3 inch x 5 inch x 3/4 inch aluminum heat sink. The socket is copper and uses 28 mils thick alumina substrates (195mils x 250mils) with 62 mils thick copper tabs (125 mils x 250 mils) on the input and output. These components are soldered together using high temperature solder.

A comparison of data in Table A1 shows the relative performance of the two mounting techniques. IR Scan II of the alumina/copper mounting technique clearly shows its superior thermal performance. Comparing the data at $P_D \sim$ 1.9 watts for IR Scan I (G-10 circuit board mounting) and IR Scan II (alumina/copper mounting) demonstrate the better thermal interface using alumina/copper. θ_{JS} for the alumina/copper mounting is 30.7 $^oC/W$ while θ_{JS} for G-10 circuit board is 39 $^oC/W$.

As expected, the θ_{JL} is approximately the same for the two mounting techniques. The difference in θ_{JS} is dependent on the mounting technique used. The resulting θ_{CS} is calculated from the IR scan data by:

$$\theta_{CS} = \theta_{JS} - \theta_{JL}$$

Thus, for IR Scan I (G-10 mounting):

$$\theta_{CS} = (39-23.2)\ C/W = 15.8\ ^oC/W$$

Whereas, for IR Scan II (alumina/copper mounting):

$$\theta_{CS} = (30.7-24.4)\ C/W = 6.3\ ^oC/W$$

Therefore, the IR scan results show a marked improvement in thermal performance when using the alumina/copper.

COMPARISON OF IR SCAN RESULTS WITH THERMAL RESISTANCE MODEL

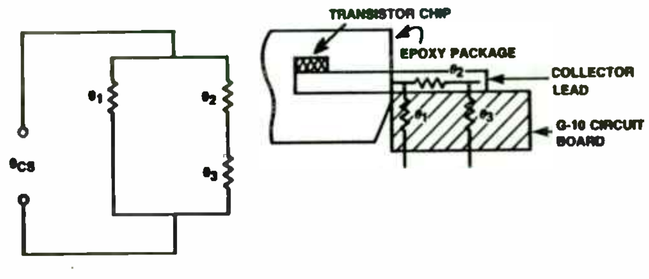
The thermal resistance model represented in Figure 7 and the basic equation, $\theta = h/KWL$, which relates the thermal resistance to the heat flow between parallel surfaces will be used to verify the IR scan results of IR Scan I and II in the following examples. The following assumptions are made to simplify the approach:

1. The calculations deal only with uni-directional heat flow
2. Heat spreading is assumed to be at a 45 degree angle
3. Parallel heat flow paths are considered.

In IR Scan I, the thermal resistance network is predicted by considering the heat flow paths from the collector lead (immediately adjacent to the package body):

1. θ_1, directly through the circuit board (spreading at a 45 degree angle) to the heat sink
2. θ_2 and θ_3, by spreading along the collector lead and then through the circuit board and to the heat sink.

The following illustration in Figure 10 shows the thermal paths and resulting model in detail:



$$\theta_{CS} = \theta_1 \parallel (\theta_2 + \theta_3)$$

FIGURE 10 — G-10 CIRCUIT BOARD MOUNTING THERMAL MODEL

Where: h_1 = .062 inches, K = 0.022 Cal/Sec-cm-^oC (see Table A2),
W = 0.10 inches, L = 0.031 inches. Similarly,

$$\theta_2 = (0.188)(2.39)(10^{-3})/(0.0254)(0.010)(0.100)(0.94) = 18.8\ ^\circ C/W$$

Where: h_2 = 0.188 inches, K = 0.94 Cal/Sec-cm-°C (see Table A2) and

$$\theta_3 = 14.1\ ^\circ C/W, \text{ Thus;}$$

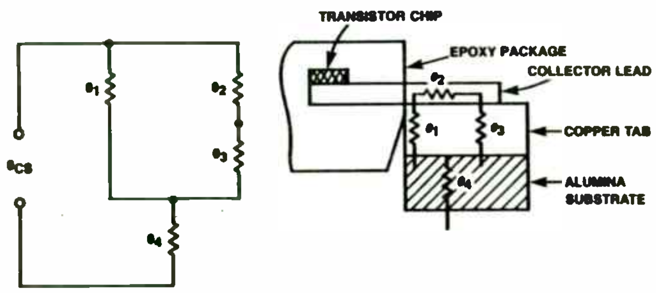
$$\theta_{CS} = 85.6 \,\|\, 32.9 = 23.8\ ^\circ C/W.$$

This value for thermal resistance is somewhat higher than the value calculated in the IR Scan data. The percentage difference is 50.6%.

In the IR Scan II, the thermal paths and the model are detailed as in IR Scan I; thus, the following thermal paths are considered:

1. θ_1, from the lead through the copper tab to the alumina substrate (assume 45 degree spreading)
2. θ_2 and θ_3, along the lead and through the copper tab to the substrate
3. θ_4, through the alumina substrate to the heat sink.

The following illustration in Figure 11 shows the thermal paths and the resulting model in detail:



$$\theta_{CS} = (\theta_1 \,\|\, (\theta_2 + \theta_3)) + \theta_4$$

FIGURE 11 — ALUMINA/COPPER SOCKET MOUNTING THERMAL MODEL

MOUNTING CONSIDERATIONS FOR RF LOW POWER PLASTIC PACKAGES
HARRY J. SWANSON MOTOROLA, INC. 1/85

$$\theta_1 = (0.072)(2.39)(10^{-3})/(0.0254)(0.036)(0.072)(0.94) = 2.0\ ^\circ C/W$$

Where: h_1 = .072 inches, W = 0.036 inches, and L = 0.100 inches.

$$\theta_2 = (0.178)(2.39)(10^{-3})/(0.0254)(0.072)(0.100)(0.94) = 4.95\ ^\circ C/W$$

Where: h_2 = .178 inches, W = 0.036 inches, L = 0.100 inches.

$$\theta 3 = (.036)(2.39)(10^{-3})/(0.0254)(0.100)(0.178)(0.94) = 0.2\ ^\circ C/W$$

Where: h_3 = 0.036 inches, W = 0.100inches, L = 0.178 inches.

$$\theta_4 = (0.028)(2.39)(10^{-3})/(0.0254)(0.125)(0.250)(0.04) = 2.1\ ^\circ C/W$$

Where: h_4 = 0.028 inches, W = 0.125 inches, L = 0.250 inches,
K = 0.04 Cal/Sec-cm-°C (see Table A2).

Thus, $\theta_{CS} = 2.0 \,\|\, (4.95 + 0.2) + 2.1 = 1.44 + 2.1 = 3.54\ ^\circ C/W.$

The value of the case to heat sink thermal resistance, θ_{CS} is somewhat lower than the value calculated from the IR scan data. The percentage difference is 43.6%.

In both cases the percentage differences were quite high, which certainly indicates that there are discrepancies between the model and the IR scan data. Possible items which could contribute to this are:

1. The model is not exact enough;
2. The resolution and calibration of the IR scan system over a broad temperature.

The model could be enhanced if convection and radiation were considered. Several papers discuss three possible means of heat transfer - radiation, convection, and conduction[2]. All three means are present in relative effectiveness in a given heat transfer system. The percentage of the heat removed by each means is dependent on the mounting and package design. Since the model presented in this paper considered only conduction through the collector lead to the heat sink, it is probable that this explains a portion of the difference in θ_{CS} between the model and the experimental results.

In AN938, the IR scan of an unmolded PowerMacro package demonstrates the loss of the heat spreading effects of the epoxy mold compound from the collector lead to the other leads and to the surface of the package.

SUMMARY

Some applications in which the power dissipation is relatively low do not utilize a circuit board mounted to a heat sink. In this case the designer relies on convection heat transfer from the surface of the transistor package and from the circuit board to provide most of the cooling for the device. However, many RF low power transistors are used in applications in which the device is mounted to a circuit board substrate; this circuit board is then mounted to a heat sink of reasonable size and capacity to provide an adequate heat sink for the collector lead. Some additional heat transfer is provided by convection and radiation from the surface of the package and heat sink.

For today's higher power devices, the power density of the package surface and the collector lead require that consideration be given to the heat sink of the collector lead. The thermal model and the experimental data have demonstrated a way to handle this task. Good values of θ_{JC} and θ_{CS} were found and are shown to be dependent on the circuit board material and mounting configuration used. The copper/alumina socket mounted directly to the heat sink offers an excellent solution.

MOUNTING CONSIDERATIONS FOR RF LOW POWER PLASTIC PACKAGES
HARRY J. SWANSON MOTOROLA, INC. 1/85

APPENDIX

Table A1 lists the IR scan results of the MRF553 PowerMacro transistor comparing two circuit board materials. The mounting and RF circuit techniques are shown in Figures A1, A2, A3 and A4.

Mounting Technique	P_{out} (W) f = 170 MHz	P_{in} (mW) / P_R (mW)	V_{CC} (Vdc) / I_C (A)	P_D (W)	Die Temp. T_J(°C) Hot Spot	Die Temp. T_J(°C) Aver.	Collector Lead Temp. T_L(°C)	θ_{JL} Aver. (°C/W)	θ_{JL} Hot Spot (°C/W)	Ckt. Heatsink Temp. T_S(°C)	θ_{JS} Aver. (°C/W)	θ_{JS} Hot Spot (°C/W)	T_A(°C) Ambient Temp.
IR Scan I:	1.5	60 / 3.0	12.5 / 0.30	1.88	96.06	93.5	69.2	15.2	30.06	30	40.2	41.7	25
G-10 PC Board	1.5	65 / 6.4	12.5 / 0.28	1.70	101.40	96.1	60.7	23.2	34.5	32	39.5	40.5	25
Thin Epoxy Case	2.0	140 / 14	12.5 / 0.30	1.80	103.0	100.2	62.7	23.2	34.5	33	39.96	40.4	25
Material	2.5	222 / 21.5	12.5 / 0.30	2.00	130	110.0	67.35	23.0	31.3	30	39.4	40.9	29
	3.0	300 / 40	12.5 / 0.41	2.40	130.35	124.0	74.2	34.0	26.1	37	30.75	41.2	25
IR Scan II:	1.5	75 / 2.7	12.5 / 0.300	1.07	66	60.7	33.0	34.0	35.7	30	38.2	37.4	25
Alumina/ Copper	2.0	130 / 13	12.5 / 0.262	1.27	73.0	71.0	38.3	34.2	28.5	30	33.8	36.3	25
Mounting	2.5	300 / 60	12.5 / 0.33	1.92	80	66	41.1	24.4	28.5	20	30.7	31.8	20
Thin Epoxy Case	3.0	710 / 106	12.5 / 0.430	3.0	120.5	117.7	48.5	23.4	24.3	20	29.7	30.7	25
Material	3.2	1000 / 300	12.5 / 0.473	3.4	120.8	121.2	40	24.2	25.8	20	30.1	31.7	25
IR Scan III:	DC		5.0 / 0.10	0.5	70.5	70.4	67.5	41.5	44.8	20	100.0	103.2	25
G-10 PC Board With	Only		5.0 / 0.15	0.75	130.4	123.5	100.0	44.0	40.5	20	130.3	143.2	25
No Heat Sink			5.0 / 0.30	1.5	171.5	166.1	137.0	41.2	43.4	20	140.1	142.3	25
Thin Epoxy Case Material													

TABLE A1 - IR SCAN RESULTS FOR MRF553 POWERMACRO

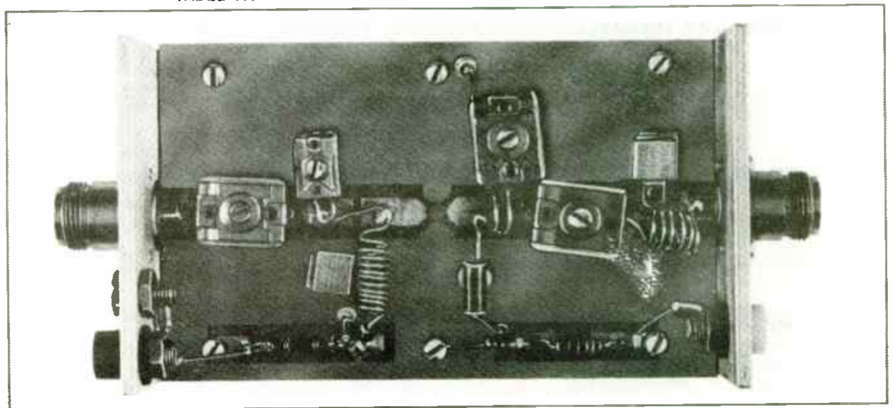


FIGURE A1 - CIRCUIT USING G-10 PC BOARD

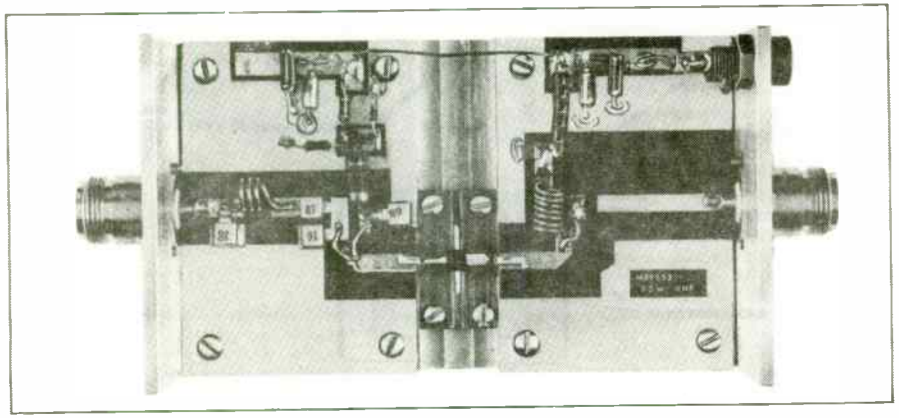


FIGURE A2 – CIRCUIT USING ALUMINA/ COPPER SOCKET

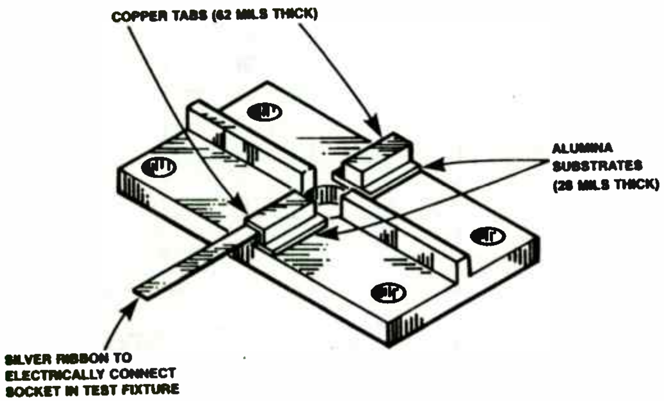


FIGURE A3 – ALUMINA/COPPER SOCKET

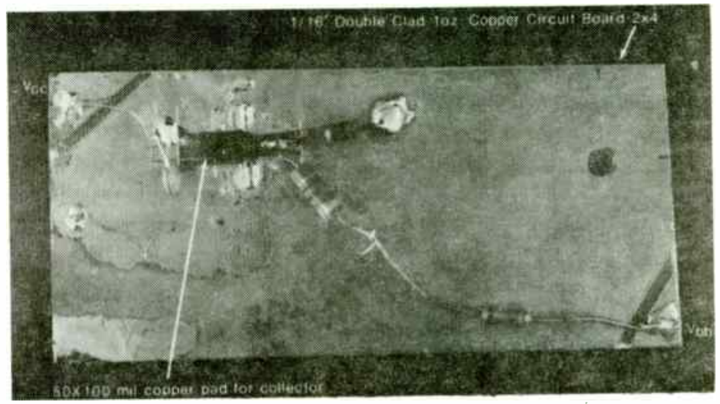


Figure A4 – Circuit Using G-10 Board with No Heat Sink

In order to aid in heat sinking and mounting designs, a table of thermal properties of common materials is presented. Three important thermal properties of common heat sink materials are given in Table A2. These properties should be considered in order to properly evaluate the choice of materials used in heat sinking /mounting of an RF plastic transistor for a given application.

Thermal Conductivity is a measure of the ability of a material of known cross sectional area to transfer heat a given distance in a given time with a given temperature difference. Generally metals are excellent thermal conductors.

Specific Heat is a measure of the amount of heat a material can accept for a given rise in temperature. The scale is normalized to the heat capacity of water ($H_2O = 1$).

Mass Density is simply the mass per unit volume of a material. This parameter is important in heat sink design in as much as large heat sinks of dense materials are undesirable.

TABLE A2. TYPICAL THERMAL PROPERTIES OF MATERIALS

Material	Thermal Conductivity K (Cal/Sec-cm-°C)*	Specific Heat S (Cal/gm-°C)	Mass Density, P (gm/cm-°C)
Copper	0.94	0.093	8.9
Beryllia Ceramic	0.56	0.31	2.8
Aluminum	0.49	0.22	2.7
Brass	0.26	0.094	8.6
Silicon	0.20	0.18	2.4
Steel	0.12	0.12	7.8
Solder	0.09	0.04	8.7
Kovar	0.046	0.11	8.2
Alumina Ceramic	0.04	0.21	3.7
Epoxy-glass (2 oz. Copper)	0.042	0.2	2.0
Epoxy-glass (1 oz. Copper)	0.022	0.2	2.0
Epoxy-glass (No Copper)	0.0007	0.2	2.0
Plastic Epoxy	0.0026	0.2	2.0
Glass	0.0026	0.2	2.0
Mica	0.0016	0.2	3.2
Teflon	0.00056	0.25	2.2
Heatsink Compound	0.0018	—	—

*Conversion Factor: 1 watt/cm = 2.39×10^{-1} Cal/Sec. cm. The thermochemical calorie = 4.184 joules. The absolute joule per second or watt is thus related in terms of calorie per second.

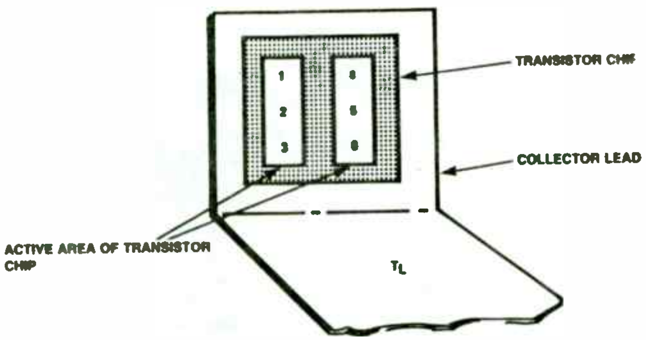


FIGURE A5 — IR SCAN MAP

MOUNTING CONSIDERATIONS FOR RF LOW POWER PLASTIC PACKAGES
HARRY J. SWANSON MOTOROLA, INC. 1/85

The IR scans were made using a Barnes radiometric scope (Model No. RM2). The transistor's active area was IR measured at 6 points to adequately map the junction temperature. Also, the collector lead was IR measured immediately adjacent to the body of the package to obtain the case temperature, TL of the device (see Figure A5).

Each operating condition was allowed to reach steady state before the IR scan measurements were made.

The devices were decapsulated using a machine called a "Jet Etch". This machine is manufactured by:

B & G Enterprises
628 Hanger Way
Watsonville, California 95076-2486

The jet etch machine uses hot sulfuric acid to decapsulate the molded device. The device can be decapsulated so that there is no mechanical damage, no corrosive damage, and no loosening of the external leads. Thus, the device is fully RF functional.

REFERENCES

1. Harry J. Swanson, AN938 - Mounting Techniques For PowerMacro Transistor, Motorola Inc., 1984.

2. Robert T. Howard, project coordinator, Thermal Management Concepts in Microelectronic Packaging, Section 3, Chapter 2: "Circuit Board Material/ Construction and Its Effect on Thermal Management" by Guy R. Wagner; Section 3, Chapter 3: "The Impact of Surface Mounted Chip Carrier Packaging on Thermal Management in Hybrid Microcircuits" by R.T. Bilson, M.R. Hepher, and J.P. Mc Carthy; Section 6, Chapter 1: "Thermal Performance Modeling for Guiding Packing Decisions" by Edward A. Wilson, Ph.D. (Silver Spring, Md.: International Society for Hybrid Microelectrics, c.1984, ISHM Technical Monograph Series 6984-003).

Application Note

Using Avantek Modamp^{TM} MMICs in Broad- and Narrow-Band Filter Design

Avantek MODAMP^{TM} MMICs are a family of state-of-the-art silicon bipolar Monolithic Microwave Integrated Circuit Amplifiers that are available in a variety of ceramic and low cost plastic packages. They are fabricated using nitride self-alignment, ion-implantation for precise control of doping and nitride passivation for high reliability. They use series and shunt feedback and exhibit very high repeatability from amplifier to amplifier. Typical applications include narrow and broadband IF and RF amplifiers in military and commercial mobile, airborne and land based systems. The series provides basic 50 ohm building blocks for the realization of simple high performance amplifier systems over the DC to 3 GHz frequency range. The series includes devices with up to 18 dB of gain and frequency response to 3 GHz or more in some applications. With a frequency response from DC to 3 GHz, 50 ohm input and output matching, flat gain vs. frequency and unconditional stability, the Avantek MODAMP MMICs are a natural choice for amplifiers or other circuits such as amplifier/filters requiring either wide or narrow bandwidths.

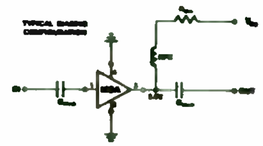
In order to design amplifier/filter cascades of finite bandwidth using filter techniques, two important criteria must be met:

1. The source and load impedances presented to the filter must be correct for the filter design and
2. The amplifier stages must be stable at all frequencies when terminated by the filter impedances.

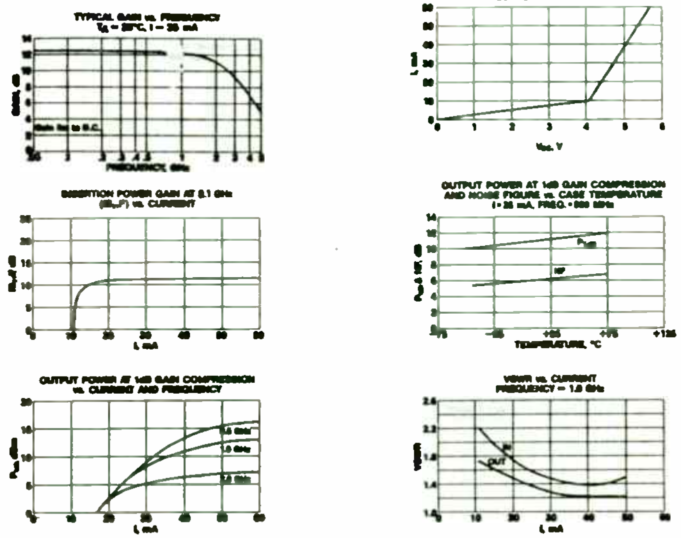
By simply cascading MODAMP MMICs as necessary to acheive a required gain, and by incorporating 50 ohm interstage filter networks, complete amplifiers with selected bandwidth/gain may be realized.

The two examples that will be presented here use a typical Avantek MODAMP MMIC, Model MSA-0335-21, a general purpose feedback amplifier having a P_{1dB} of +10 dBm. Figure 1 gives the "S" Parameters and shows some typical performance curves for this device. It can be seen from the curves in figure 1 that the device is

RF Expo - Larry Leighton - Disk D10 - Oct., 1984



unconditionally stable at all frequencies and presents exceptionally good input and output VSWR over the entire operating frequency range.



TYPICAL SCATTERING PARAMETERS*

MSA-0336-21

I = 25 mA (V = 5.0V Typ.)

Freq. MHz	S_{11} Mag	S_{11} Ang	S_{21} dB	S_{21} Ang	S_{12} Mag	S_{12} Ang	S_{22} Mag	S_{22} Ang
100	.09	179.0	12.30	177.0	.124	2.0	.13	-6.0
200	.08	176.0	12.30	174.0	.126	4.0	.13	-13.0
400	.08	175.0	12.30	166.0	.121	8.0	.14	-27.0
600	.07	172.0	12.10	159.0	.128	13.0	.14	-40.0
800	.07	172.0	12.10	155.0	.131	17.0	.14	-52.0
1000	.08	-179.0	12.00	146.0	.129	19.0	.15	-64.0
1200	.08	-170.0	11.90	141.0	.137	24.0	.16	-76.0
1400	.08	-157.0	11.70	134.0	.144	28.0	.16	-88.0
1600	.07	-142.0	11.60	128.0	.146	28.0	.17	-98.0
1800	.09	-136.0	11.30	122.0	.154	32.0	.18	-108.0
2000	.12	-137.0	11.00	115.0	.161	36.0	.19	-117.0
3000	.25	-145.0	8.20	88.0	.186	41.0	.22	-148.0
4000	.37	-163.0	6.80	67.0	.206	44.0	.20	-153.0
MSA-0335-22								
100	.07	172.7	12.16	175.0	.124	.3	.15	-8.8
200	.07	165.9	12.09	170.0	.124	1.1	.15	-18.7
400	.08	157.5	11.97	160.0	.125	2.7	.15	-29.6
500	.08	150.0	11.92	155.0	.128	3.4	.15	-38.1
750	.08	105.6	11.80	137.0	.134	-6.6	.13	-56.5
1000	.08	132.4	11.80	131.1	.164	-6.2	.17	-70.9
1500	.03	-175.2	11.41	105.4	.162	8.2	.17	-107.6
2000	.08	-158.8	10.90	81.9	.172	6.7	.19	-137.6
2500	.15	-175.2	8.83	58.5	.191	-.2	.20	-164.3
3000	.22	167.6	8.82	38.6	.203	-6.3	.21	174.5
3500	.29	150.3	7.03	17.3	.208	-11.7	.21	158.2
4000	.34	134.2	6.30	1.1	.215	-16.0	.22	140.7

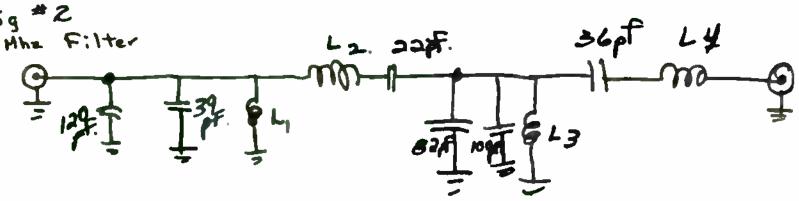
Figure 1: Typical Performance Curves and Scattering Parameters - MSA-0335-21

The first filter to be considered is a 70 MHz filter with the following specifications:

Center Frequency	67.7 MHz
Maximum in-band ripple	0.5 dB
Design Bandwidth	36 MHz
Bandwidth at -15 dB	50 MHz
Input/output impedance	50 ohms

Figure 2 shows the circuit schematic for this filter. It is a straightforward 8 pole network built entirely of discrete components. In order to accommodate component tolerance, some tuning will be required. In the prototype, this was done by spreading the turns of the inductors as necessary.

Fig #2
70 Mhz Filter



L1	3	turns	#22	space wound	1/8"	I.D.
L2	7	"	"	close	"	1/4	"
L3	2	"	"	space	"	1/8"	"
L4	4½	"	"	close	"	1/4	"

Figure 2: 70 MHz circuit Schematic

Figure 3A shows the calculated insertion gain and figure 3B the calculated output reflection coefficient (Γ_{out}) of the filter shown in figure 2. The Γ_{out} is shown without amplification as well as with a MSA-0335-21 MODAMP MMIC cascaded both in front of and behind the filter. Note that the effect on Γ_{out} is minimal. Because data is not calculated at all frequencies, Γ_{out} would appear to never have an excellent match, when in fact, at some frequencies the match would be good. The perterbations in the passband are a result of the desired design ripple.

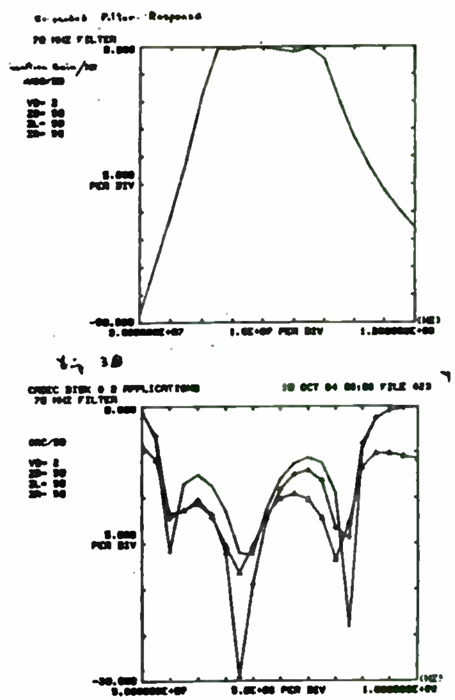
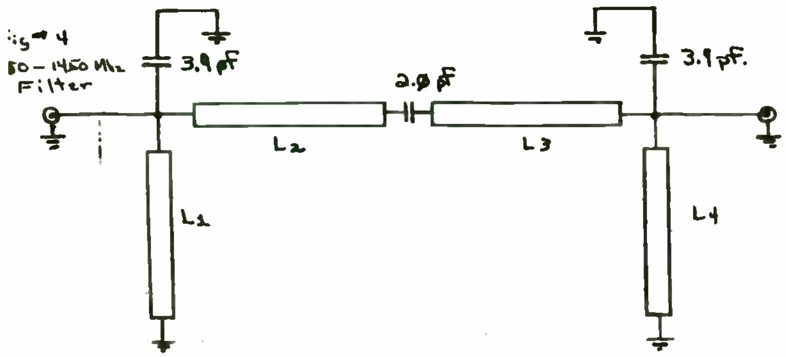


Figure 3: Calculated response characteristics of the 70 MHz filter circuit shown in Figure 2 by itself and cascaded both before and after the MODAMP MMIC.

The next filter to be considered is a 950 to 1450 MHz circuit with these specifications:

Center frequency	1200 MHz
Maximum inband ripple	1.0 dB
Design bandwidth	500 MHz
Bandwidth at -15 dB	1300 MHz
Input/output impedance	50 Ohms

The schematic for this filter is shown in Figure 4. The intent of picking this filter for an example is to show that MODAMP MMICs can be used at any frequency within their passband with little effect on other elements in the circuit.



Printed Circuit Board Material
.031 inches thick
$E_r = 2.2$
All lines are .0286 inches wide. (50Ω)

L1, L4 ~.397 inches long
L2, L3 ~.428 inches long

Figure 4: Circuit schematic for a 950 to 1450 MHz filter

RF Expo - Larry Leighton - Disk D10 - Oct., 1984

Figure 5A shows the calculated insertion gain of the filter shown in figure 4, without amplification while figure 5B shows the calculated overall effects on S_{21} of the 950-1450 MHz filter as used with and without the MSA-0335-21. Again, performance is compared with the straight filter, and with the MODAMP MMIC cascaded both in front of and behind the filter. Note that the maximum deviation between configurations is 0.4 dB with respect to S_{21}. The filter by itself is shown to have gain when in fact it doesn't. A unilateral 50 ohm amplifier was cascaded with the filter so that the effects on S_{21} could be plotted conveniently.

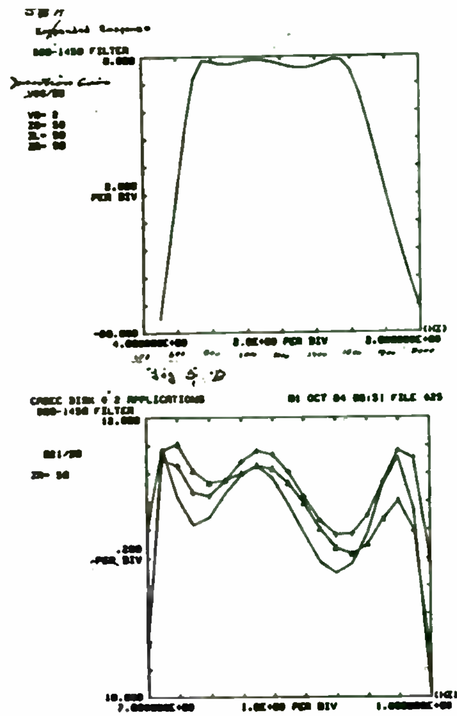


Figure 5: Calculated response characteristics of the 950-1450 MHz filter circuit shown in Figure 4 by itself and cascaded with a MSA-0335-21 MODAMP MMIC.

RF Expo - Larry Leighton - Disk D10 - Oct., 1984

Figure 6 demonstrates the ability of the MODAMP MMIC to provide isolation between the 950 to 1450 MHz filter and elements in front of the filter. Note that with the MODAMP MMIC in front of the filter that the system input reflection coefficient never exceeds -6 dB whereas the other configurations show 0 dB outside the passband.

The two filters were constructed and performance compared to that calculated. Performance deviation from that calculated was could easily be explained by component tolerances and lossy elements. Although the MODAMP MMIC amplifier is touted as a 50 ohm broadband device, this article has demonstrated the flexibility of the device as an element of a frequency sensitive circuit providing both gain and isolation.

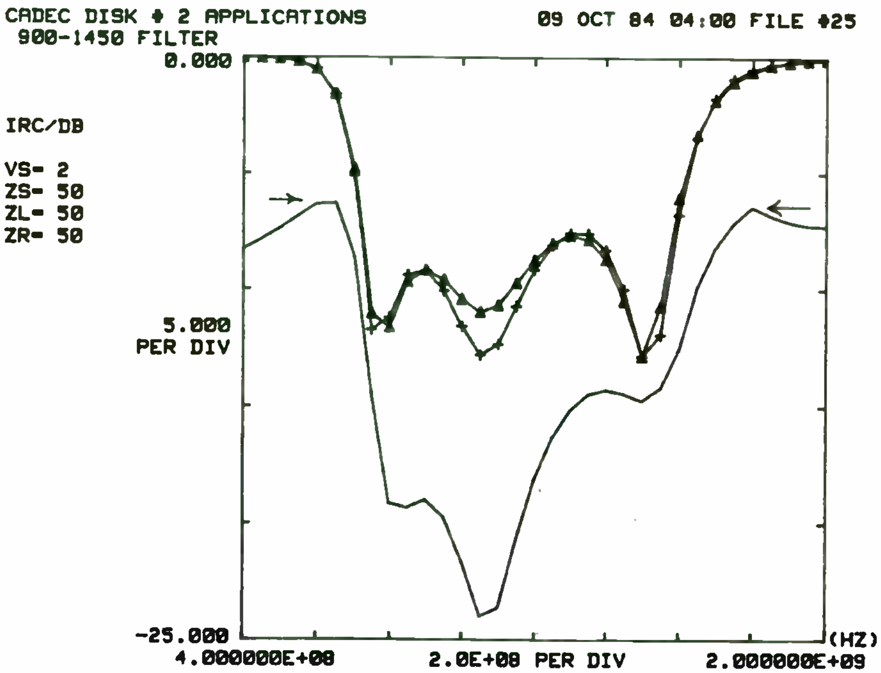
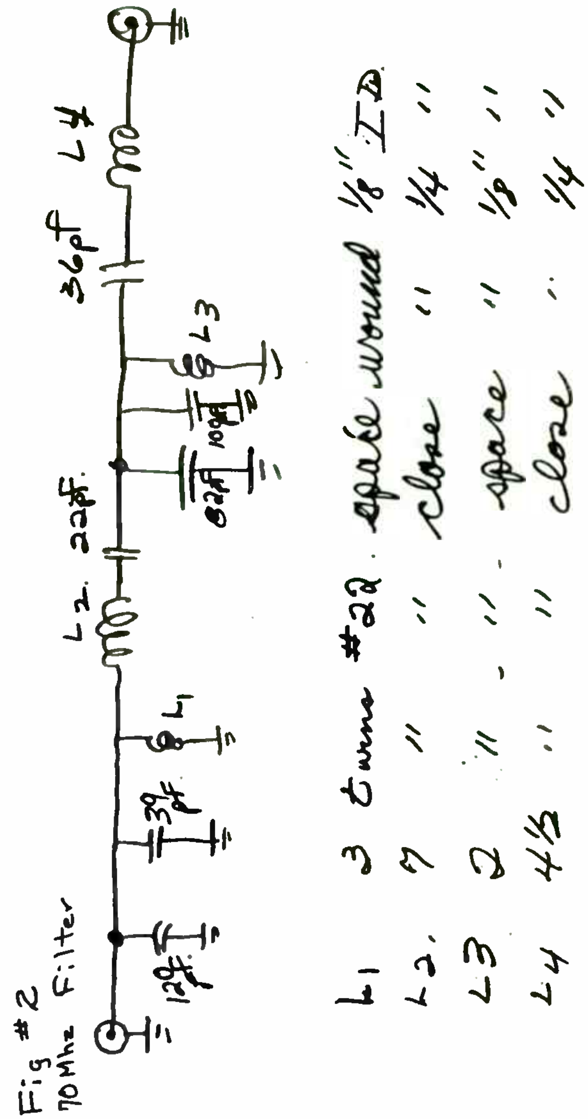
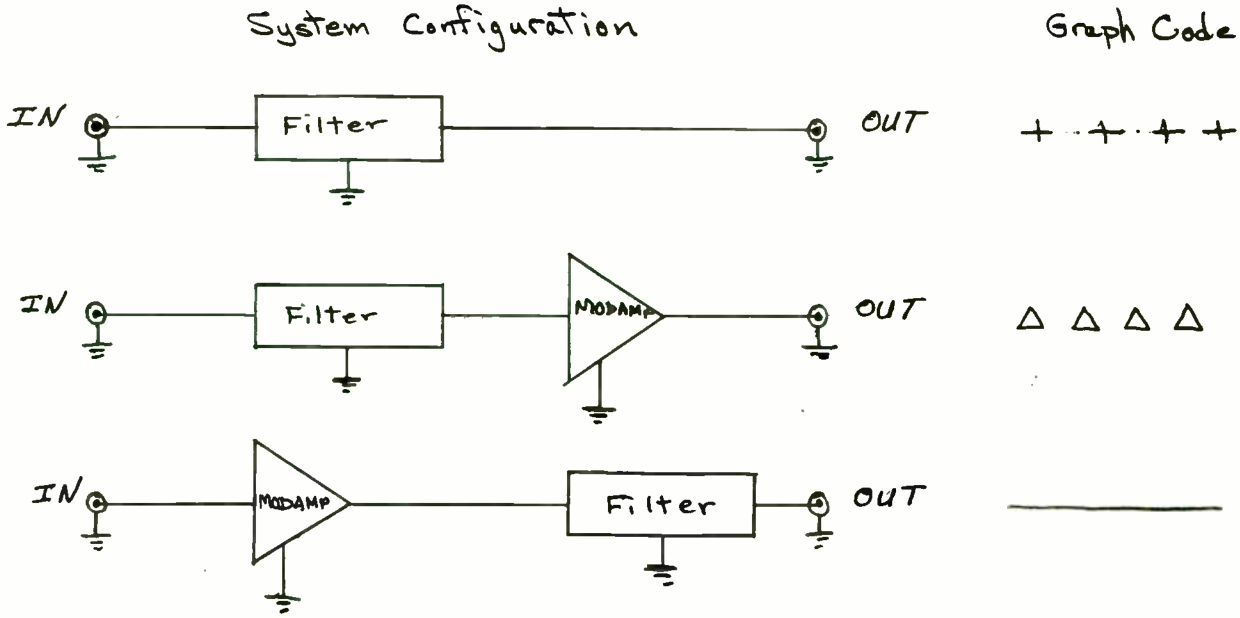


Figure 6: Input reflection coefficient of the 950-1450 filter





377

Fig. 3A

Expanded Filter Response

70 MHZ FILTER

Insertion Gain/DB

DB

VS= 2
ZS= 50
ZL= 50
ZR= 50

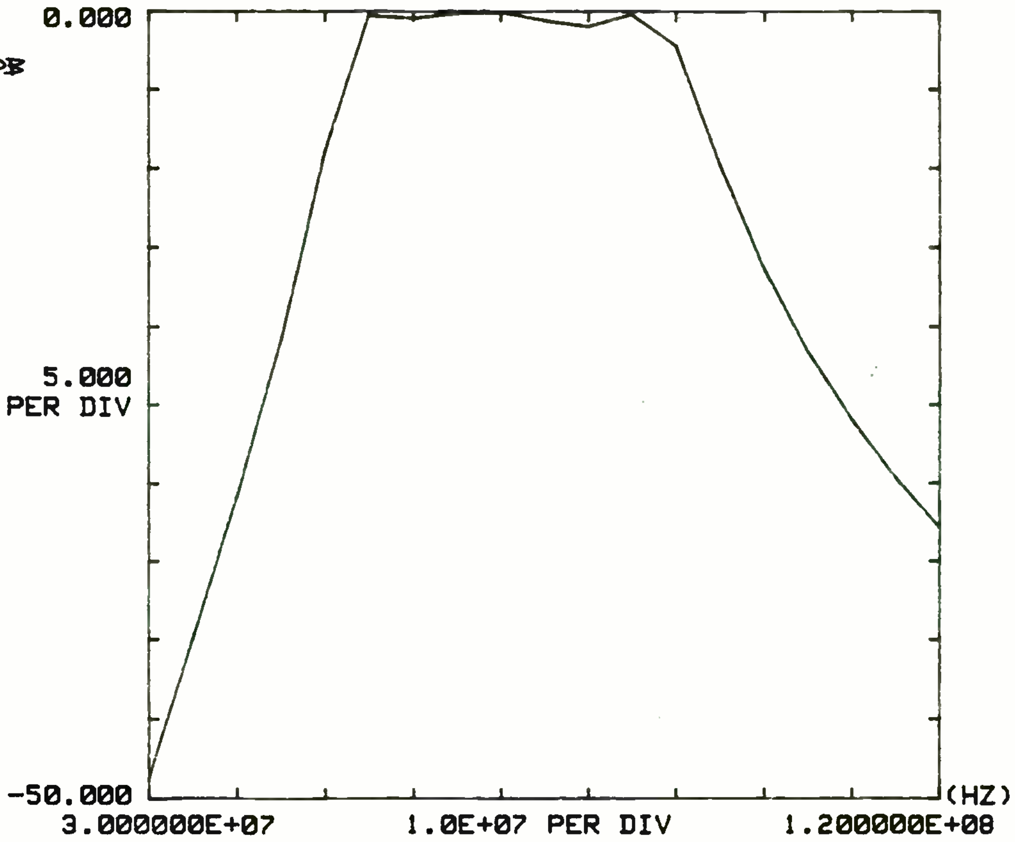
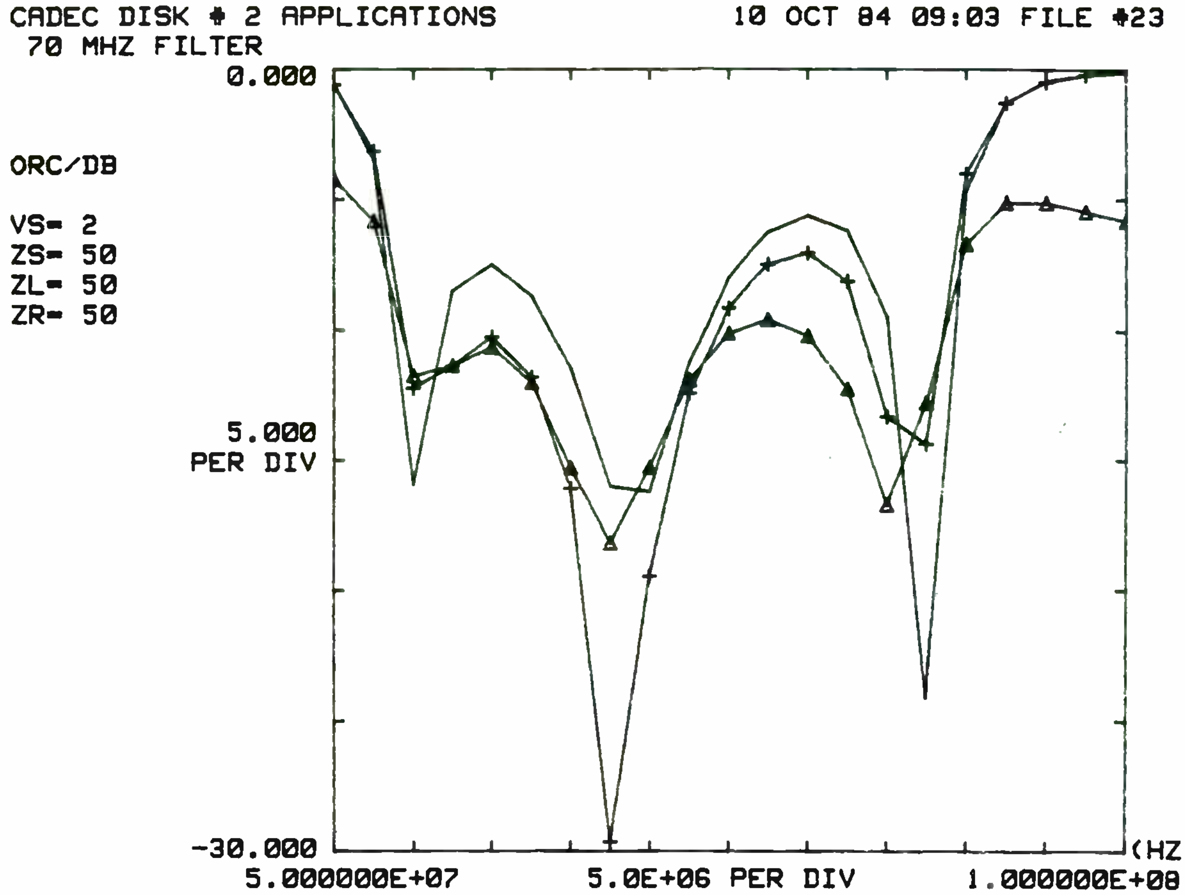
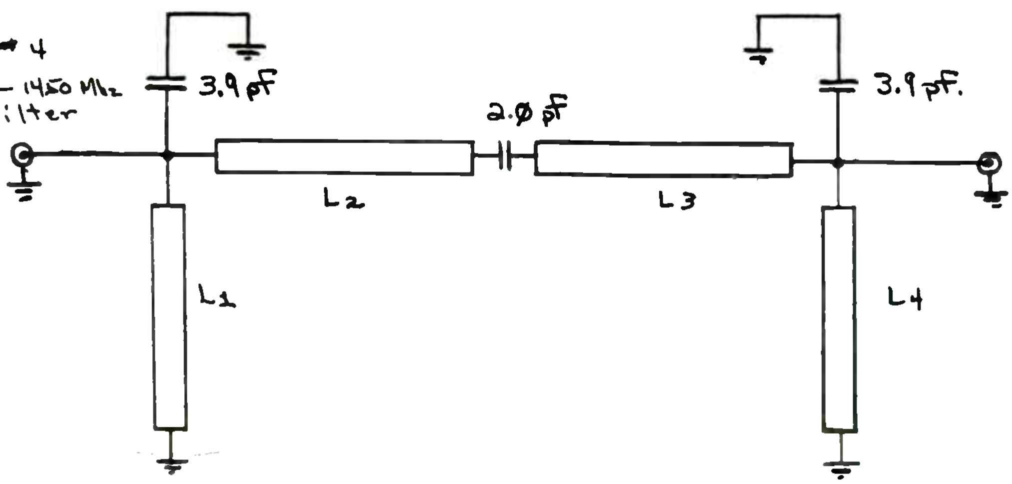


Fig 3B



378

Fig # 4
950 – 1450 MHz
Filter



Printed Circuit Board Material
.031 inches thick
$E_r = 2.2$
All lines are .0286 inches wide. (50Ω)

L_1, L_4 ~ .397 inches long
L_2, L_3 ~ .428 inches long

5~~B~~ A

~~Expanded~~ Response

900-1450 FILTER

Insertion Gain

~~VOG~~/DB

VS= 2
ZS= 50
ZL= 50
ZR= 50

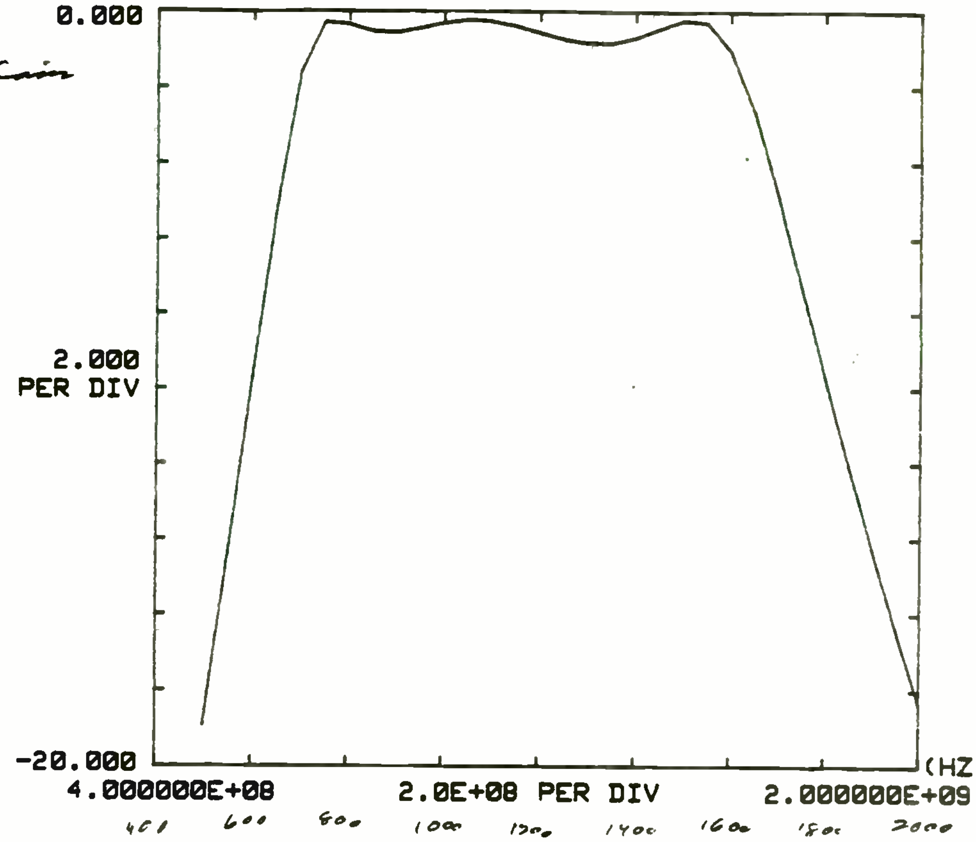


Fig 5~~B~~B

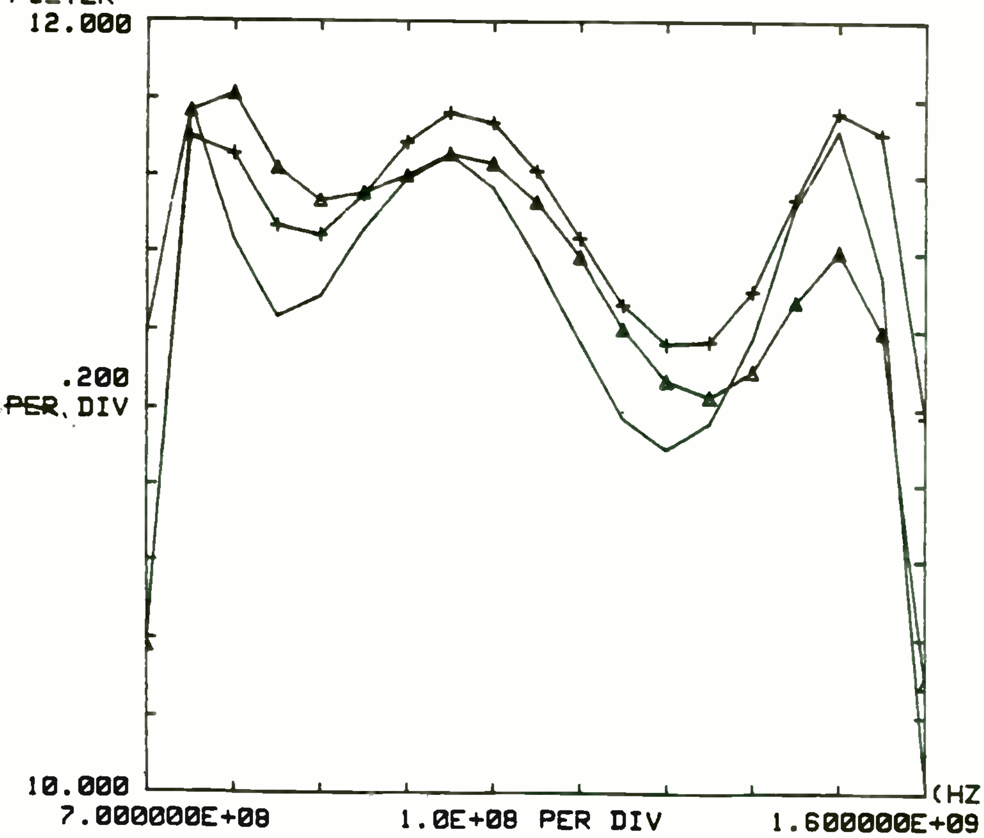
CADEC DISK # 2 APPLICATIONS

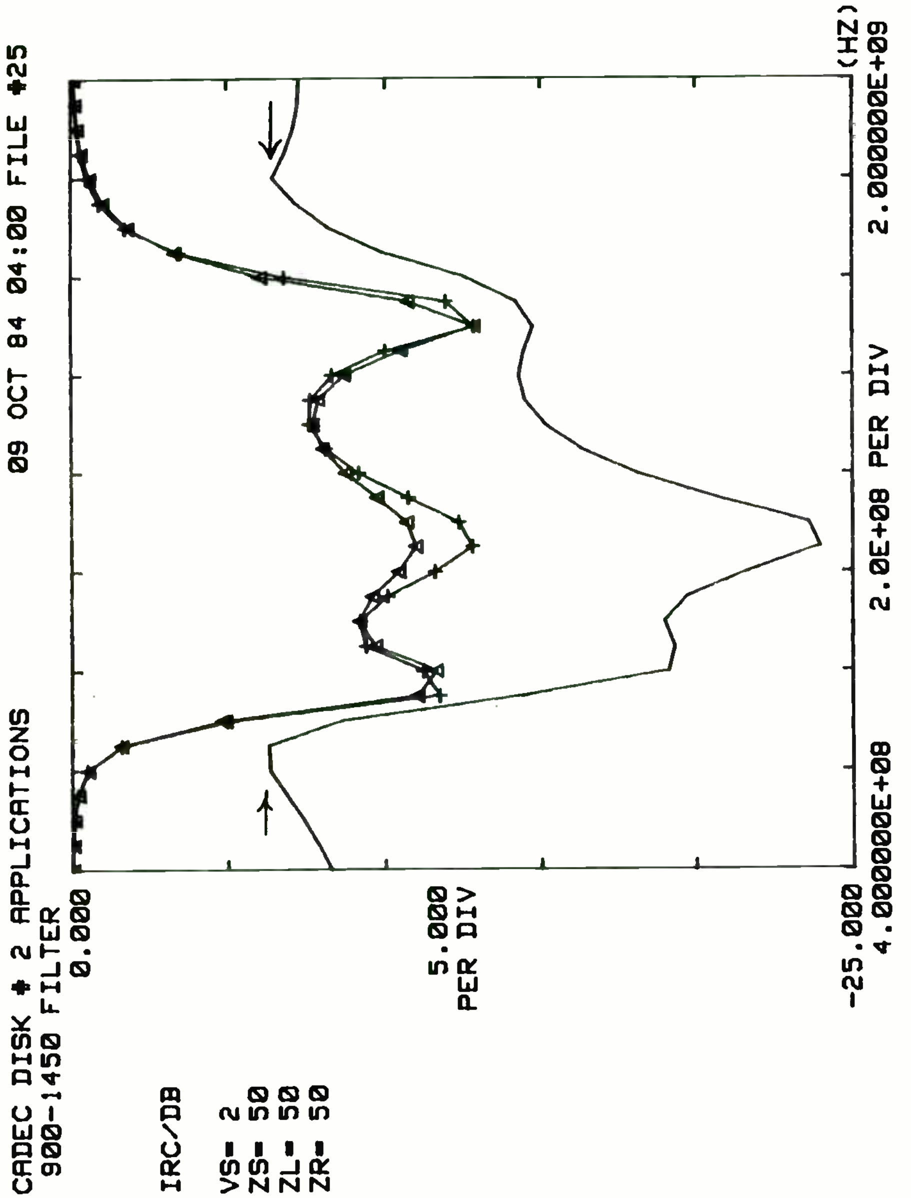
01 OCT 84 09:51 FILE #25

900-1450 FILTER

S21/DB

ZR= 50





IMFET: $V_{DS} = 9$ V, $I_{DS} = 2$ A

Conditions: $\Delta F = 10$ MHz

$P_{out} = 30.5$ dBm each tone

Lot #	Nbr of IMFET Tested	F = 5900 MHz				F_0 = 6150 MHz				F_1 = 6400 MHz			
		IMD_3		IMD_5		IMD_3		IMD_5		IMD_3		IMD_5	
		Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
I 84 3709	2	-42	-41	-55	-55	-42	-36	-50	-46	-34	-32	-46	-46
I 84 4117	2	-55	-44	-55	-55	45	-45	-52	-51	-37	-35	-49	-45
I 84 4207		-37	-36	-55	-54								
~~I 84 4209~~	2	-36	-35	-54	-53	-40	-38	-55	-55	-43	-42	-55	-55
I 84 4210	2	-35	-35	-55	-55	-39	-37	-55	-53	-36	-35	-45	-45
I 84 4501	2	-37	-34	-55	-50	-35	-34	-53	-47	-32	-31	-47	-47

Worse IMD_3 = -31 dB IMD = -44 dBc

Worse IMD_5 = -45 dB IMD = -55 dBc

IMFET: 3742-16 $V_{DS} = 9$ V, $I_{DS} = 2$ A

Conditions: $\Delta F = 10$ MHz

$P_{out} = 30.5$ dB_m each tone

Lot #	Nbr of IMFET Tested	F = 3775 MHz				F_0 = 3950 MHz				F_1 = 4200 MHz			
		IMD3		IMD5		IMD3		IMD5		IMD3		IMD5	
		Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
I842923A	4	-37	-35	-52	-47	-37	-34	-51	-48	-40	-34	-52	-49
I842924A	4	-37	-30	-53	-45	-35	-34	-54	-46	-39	-32	-53	-48
I842925A	4	-36	-33	-55	-48	-37	-34	-54	-47	-36	-34	-53	-49
I842926A	4	-38	-34	-56	-50	-39	-35	-57	-60	-40	-37	-57	-50
I843608A	4	-41	-41	-52	-46	-44	-41	-47	-42	-37	-34	-48	-43
I843704A	4	-41	-36	-50	-49	-41	-35	-49	-48	-43	-35	-50	-45
I843705A	4	-41	-40	-50	-43	-42	-41	-49	-46	-44	-42	-51	-44

Worse IMD3 = -30 dB

Worse IMD5 = -42 dB

LINEAR MICROWAVE POWER AMPLIFIER DESIGN

WILLIAM J. THOMPSON
and
Dr. KRISHNA K. AGARWAL

Rockwell International
Richardson, Texas

INTRODUCTION:

This paper on linear power amplifier design covers the performance objectives and how this performance was realized. Also covered are the reasons why linearity is required, how to design for linearity, the important bias circuit design considerations, the importance of verifying the relationship between experimental measurements and computer calculation of circuit performance. Determination of the FET source and load circuit impedances for best performance is also covered. As an example, a step-by-step design of an NE800495-4 gate circuit is described with a description of the finished circuit and its performance. The measured performance of the overall multistage amplifier is discussed.

This linear amplifier is intended to amplify a 64 symbol quadrature amplitude modulated (QAM) digital signal which has substantial amplitude modulation. Its peak signal is 7.2 dB higher than the average signal. Another possible application of this amplifier would be in single sideband (SSB) radios where the ratio of peak to average power is much higher. These signals are sensitive to both AM to AM and AM to PM distortion.

Amplifier Design Considerations:

The design objectives for the amplifier are shown in Table 1. Two versions of the amplifier were designed, one with 3 watts and another with 6 watts of output power at the 1 dB gain compression point. In the design of this amplifier the objective was to obtain the best balance between linearity, DC power consumption and cost. This is achieved by designing each stage for maximum power output at a given distortion level. A simple method of reducing nonlinear amplifier distortion is to operate the amplifier well below its maximum power capability. This is called the back-off method of distortion reduction. The lowest distortion in a multistage amplifier would be obtained by employing the highest power transistor available in every stage. This approach would be very expensive both in transistor cost and in DC power consumption. The lowest cost and DC power consumption in a multistage amplifier would be obtained by operating every stage at or near its maximum output power capability. The near equal distortion of every stage would tend to add and result in relatively high overall amplifier distortion. The compromise distortion reduction strategy employed in this amplifier was to operate the early stages of the amplifier far enough below their power capability that their contribution to the distortion of the overall amplifier was balanced against the cost of further improvement. As a result of this compromise strategy, the earliest stages of the amplifier have the greatest power backoff. Other more complicated methods of distortion reduction such as feed-forward and predistortion were not employed within this amplifier but this amplifier could be used in a system employing these techniques.

When designing for best linearity in an amplifier, it is necessary to design the individual stages considering not just gain but also power and distortion. Better results are obtained by designing the stage source and load impedances for maximum power output at the 1 dB gain compression point rather than designing the stage for maximum small signal gain. Design for maximum small signal gain, results in substantially lower output power at the 1 dB gain compression point. The measured data in table 2 shows that tuning at the one dB gain compression point results in 4.25 dB more output power at the one dB gain compression point and 2.35 dB less gain than the same amplifier tuned for maximum small signal gain. At equal output powers the measured two tone third order distortion levels of the amplifier tuned at the one dB gain compression point are 6 to 16 dB lower than those of the amplifier tuned for maximum small signal gain. The cost and power consumption required to make up the lost gain with low level input stages is less than that required to raise the power capability of the high level stages. For these reasons, the individual stages in this amplifier are designed to operate at the load and source impedances that result in maximum output power at the one dB gain compression point.

In this multistage amplifier, the last stage causes most of the amplifier distortion. The earlier stages of the amplifier are operated at progressively greater amounts of power back-off. This results in the best trade-off between overall amplifier distortion, power consumption, and cost. Table 3 shows the amplifier levels and stage gains for the 6-watt version of the amplifier.

One of the goals in the design of this amplifier was to establish a strong correlation between computed and measured circuit performance. A good connector-interface between the automatic network analyzer (ANA) and the circuit substrate simplifies this process. Figure 1 shows the connector type used. It provides a return loss of 20 dB or better up to 12 GHz.

Bias Network Design:

The design of a bias network which delivers the DC power to the FET is important. In the passband, the bias circuit should be essentially transparent to the RF signal passing from the input-/output connection to the FET. This requires that at the point where it connects to the RF path the bias network should present a high impedance to the RF circuit. Figure 2 shows the construction and measured performance of the bias network. From the FET at port 1 to the RF input at port 2, the insertion loss in the 3.7 to 4.2 GHz amplifier passband is approximately 0.3 dB. The bias network should also isolate the RF signal from the DC connection. Figure 3 shows the measured passband transmission loss from the FET to the DC connection point. A feed-through filter provides additional attenuation, but it is important to achieve substantial attenuation on the substrate itself. It is desirable to have some dissipation in the bias circuit in order to reduce the tendency towards spurious oscillations. For this reason, the bias network incorporates a resistor in the RF portion of the circuit. Figure 4 compares the calculated and measured performance of the bias network. Excellent agreement between the measured and calculated performance was obtained. A relatively small 200 MHz offset is attributed to junction discontinuity effects , which were not fully accounted for in the computer model.

Device Characterisation:

The next step in the design process is to determine the load and source impedances required for maximum FET output power at one dB gain compression. In the absence of an accurate analytic method, these impedances were determined experimentally. The opened up transistor test fixture with attached multiscrew tuners is shown in Figure 7. The internal construction of the multi-screw tuners is shown in Figure 6. This fixture holds the transistor in place by mechanical means so that it is not necessary to solder the FET into the fixture. DC bias and RF drive are applied so that the transistor is operating at the 1 dB-gain compression point. The tuners are adjusted for maximum power output. The fixture is then broken apart so that connectors can be attached to the FET input and output interfaces. With the tuners still attached, the input and output networks are measured on an ANA to determine the impedances presented to the FET at the desired operating point. This is done at 5 to 7 frequencies in and around the amplifier passband. The impedances measured by the ANA apply to the ends of the coaxial connectors so it is necessary to transform these impedances to the impedances existing at the end of the substrate to which the FET had been attached. To do this, an internal program which operates as shown in Figure 8 is employed. The program calculates the two port ABCD matrix of the connector equivalent circuit and then forms the inverse of the connectors ABCD matrix. Multiplying the ABCD matrix of the measured (with connector) network by the inverse of the connectors ABCD matrix effectively removes the connector from the rest of the circuit.

Matching Circuit Design (Lumped Element):

The following gives as an example the design of the gate-matching network of the NE800495-4 FET which has a nominal output power of 2 watts at 4.0 GHz. Figure 9 shows the measured gate-source impedance after correction for the connector. The matching networks could be designed directly from the measured impedances, but much understanding can be obtained if equivalent circuits are used in the design procedure. The simple equivalent circuits shown in Figure 10 can provide a good approximation to the FET's input and output impedances. The first step in the design procedure is to determine the equivalent circuit that best fits the measured impedance data. Figure 11-A shows the first estimate of the gate equivalent circuit and is based on the Smith chart impedance plot and previously published data. This was arrived at by using the following reasonings and rules: The larger the FET, the larger the series capacity of the gate. For this FET the estimate is 8 pF. At 4 GHz, the impedance of the series resistance would normally be less than that of the series capacitor, (40% of Xc = approximately 2 ohms). The very fact that the measured impedance, as plotted on a Smith chart, looks more like a parallel resonant circuit than a series resonant circuit shows that the shunt capacity must be significant so 1 pF was selected as a starting value. The measured impedance passes through 50 ohms real at approximately 4.4 GHz. Using the computer to analyze the circuit of Figure 11-A and viewing the computed Rp and Xp input impedance of that circuit, the value of the series inductor was varied until the value of Rp equalled 50 ohms at 4.4 GHz. The value of the shunt capacitor was varied until the impedance was real at 4.4 GHz. This resulted in the circuit of Figure 11-C. With this estimate of the gate equivalent circuit as a starting point, the computer optimizer is used to match the gate equivalent circuit to the measured gate impedances. An internal program employing a gradient search method was used for this optimization; however, there are several commercial programs (S-COMPACT,EESOFT ect.) that can do the same job. Figure 12 shows the equivalent circuit after optimization.

The initial matching circuit design employs a lumped constant matching circuit to get a feel for the matching bandwidth possibilities for the circuit. Since, in theory, it is possible to match real impedances to real impedances over an infinite bandwidth, the resistive matching problem was disregarded to concentrate on the reactive match. Figure 13-A shows a single-tuned matching circuit for the gate equivalence circuit. The 0.8 pF capacitor at the input to the FET is sufficient to make the input impedance of the entire network real (resistive) at 4 GHz. The performance of this single-tuned circuit is tabulated in Figure 13-A. The 0.9 dB roll-off at 3.6 and 4.4 GHz was considered excessive, and it was necessary to use a double-tuned matching circuit to obtain the desired flat bandwidth. The final circuit will require the use of some distributed circuit elements. In general, the bandwidth of distributed matching circuits is narrower than that of lumped element matching circuits, so the lumped element circuit should have some bandwidth margin to account for this effect.

The classic double-tuned bandpass filter circuit consists of a series-tuned circuit connected to a shunt-tuned circuit. One design criterion for a double-tuned circuit is to have a real impedance looking both ways at the junction between the series resonant circuit and the shunt resonant circuit at center frequency. Figure 13-B shows a trial double-tuned matching circuit for this gate. The 0.8 pF shunt capacitor added to the FET equivalent circuit makes the input impedence of this circuit look like a shunt resonant circuit. The 5 nH inductor in series with the 0.32 pf capacitor represents a series-tuned circuit that is also resonant at midband. If the double-tuned circuit is to have the traditional flat topped response, then the loaded Q of the series and shunt resonant circuits must be approximately the same. The starting point for choosing the 5 nH inductor was the observation that the 3 dB bandwidth of the single-tuned shunt circuit was 1.7 GHz. The 5 nH inductor was selected so that its impedance at 1.7 GHz was approximately equal to the 56.6 ohm source impedance, thus making the Q of the input series tank approximately the same as the Q of the output shunt resonant circuit. Starting with the circuit of Figure 13-B, the optimizer designed the circuit of Figure 14 which has the classic double-tuned circuit response and the desired passband flatness of 0.13 dB peak to peak over the 3.6 to 4.4 GHz range. Even though lumped constant circuits at microwave frequencies are not practical, this does demonstrate that there are no intrinsic limitations to realizing the desired circuit bandwidth.

Distributed Circuit:

The next step is to devise a microwave circuit using realizeable distributed circuit elements to replace the series resonant circuit shown in Figure 15. Figure 15 also shows a computer-generated Smith chart with an impedance plot of this series resonant circuit over the 2 to 6 GHz range. Figure 16 shows an impedance plot of one of the many possible microstrip transmission line circuit configurations that could be used to approximate the desired impedance. The half wavelength long open-circuited stub acts like a shunt-resonant circuit (high impedance at center frequency). The quarter wavelength long series line acts as an impedance inverter making the shunt resonant circuit look like a series resonant circuit at its input. The impedence of this circuit is substantially different from the prototype series tuned circuit at 2 and 6 GHz; but it is a reasonable approximation to the desired impedance over the 3.6 to 4.4 GHz range. The 0.8 pf shunt capacitor can be approximated by a short length of low impedance line. DC bias considerations require that a blocking capacitor be placed on the input line between the FET and the input connector. This blocking capacitor can either be relatively transparent (10 pf) or can be smaller and become part of the matching network.

With the basic conceptual matching circuit design complete, the optimizer can be used without fear that it will fail. The final optimization was done using the measured impedances rather than the gate-equivalent circuit in order to avoid having an approximation to an approximation. Figure 17 shows the final optimized gate-matching circuit. The 0.080 long length of the first 28.62 ohm line was predetermined as the minimum length required to accept the FET gate lead. The 1.4 pf series capacitor optimized close to this value but was forced to the exact 1.4 pf standard value in the final optimization. The circuit of Figure 17 is both realizable and effective. Figure 18 shows the computer-generated Smith chart impedance plot of the final matching network. Referencing back to the measured gate source impedance of Figure 9 will show that a reasonably good fit has been obtained over the 3.7 to 4.2 GHz band.

Amplifier Fabrication and Results:

Figure 20 is a photograph of the completed NE800495-4 stage showing the network as designed. The bias network has negligible effect on the matching network since it was designed to be essentially transparent; however, for best results, the bias network should be connected to a low impedance point in the matching network. The location is determined by using the computer to analyze the matching network at various points and then attaching the bias network to the lowest impedance/lowest voltage point. In this particular case this point is right next to the FET gate input. Figure 22-A shows the measured passband performance of the amplifier built in accordance with the computer calculated dimensions, in the unadjusted, untuned (no chips) condition. The small signal gain has a peak to peak passband nonflatness of 0.8 dB over the 3.6 to 4.3 GHz band which tends to go flat at the 1 dB gain compression point. Figure 22-B shows the same amplifier tuned with chips for a flatter (0.2 dB peak to peak) small signal gain over the 3.7 to 4.2 GHz band. Some tuning chips are apparent on these circuits but their effect is small.

In the final amplifier the NE800296 and the NE800495-4 stages are joined by an isolator into a single module. Figure 23 shows the untuned and tuned performance of three of these amplifier modules. The three untuned modules have a worst case small signal gain non-flatness of 0.6 dB peak to peak over the 3.7 to 4.2 GHz band. With slight tuning, a peak to peak gain flatness of 0.3 dB peak to peak was obtained. Figure 24 shows the overall passband of the last two modules of the amplifier consisting of three stages.

Figure 25 is a photograph of the completed amplifier. The amplifier is intended to be mounted to a heat sink to keep the FET's operating within the manufacturers specified temperature range. A PIN diode attenuator in combination with a temperature sensing control circuit compensates for temperature induced gain variations. Table 4 tabulates the overall measured performance of the amplifier. A two-tone third order distortion intercept of 49.5 dBm was measured. The system tests performed indicated adequate linearity for a 64 QAM digital system. the amplifier meets all the gain, linearity, output power and DC power objectives. The construction is simple and cost effective.

Conclusion:

In conclusion, a computer-aided design process leading to amplifiers with good passbands with a minimum of experimental tuning has been described. It is important to have a good circuit design plan that leads to approximate and rational circuit solutions before computer optimization is started. If this is not done, computer optimization may lead to undesirable local minimums and/or impractical circuit realizations. It is also important to verify that the calculated and measured circuit performance is in agreement and to resolve any sources of significant discrepancy. This should be done on small circuit portions so that the cause of any discrepancy can be easily located.

LINEAR AMPLIFIER DESIGN OBJECTIVES

FREQUENCY RANGE 3.7   4.2 GHz

GAIN 45 dB MINIMUM FLAT WITHIN 2 dB PEAK-TO-PEAK GAIN FLATNESS OVER 40 MHz INCREMENTAL BANDWIDTH +/-0.25 dB

POWER OUTPUT 3/6 WATTS

LINEARITY: +47 dBm MINIMUM THIRD ORDER DISTORTION TWO TONE INTERCEPT

PROTECTED AGAINST LOSS OF GATE BIAS

PROTECTED AGAINST ACCIDENTAL POWER SUPPLY REVERSAL

D.C. POWER CONSUMPTION: 35 WATTS MAXIMUM

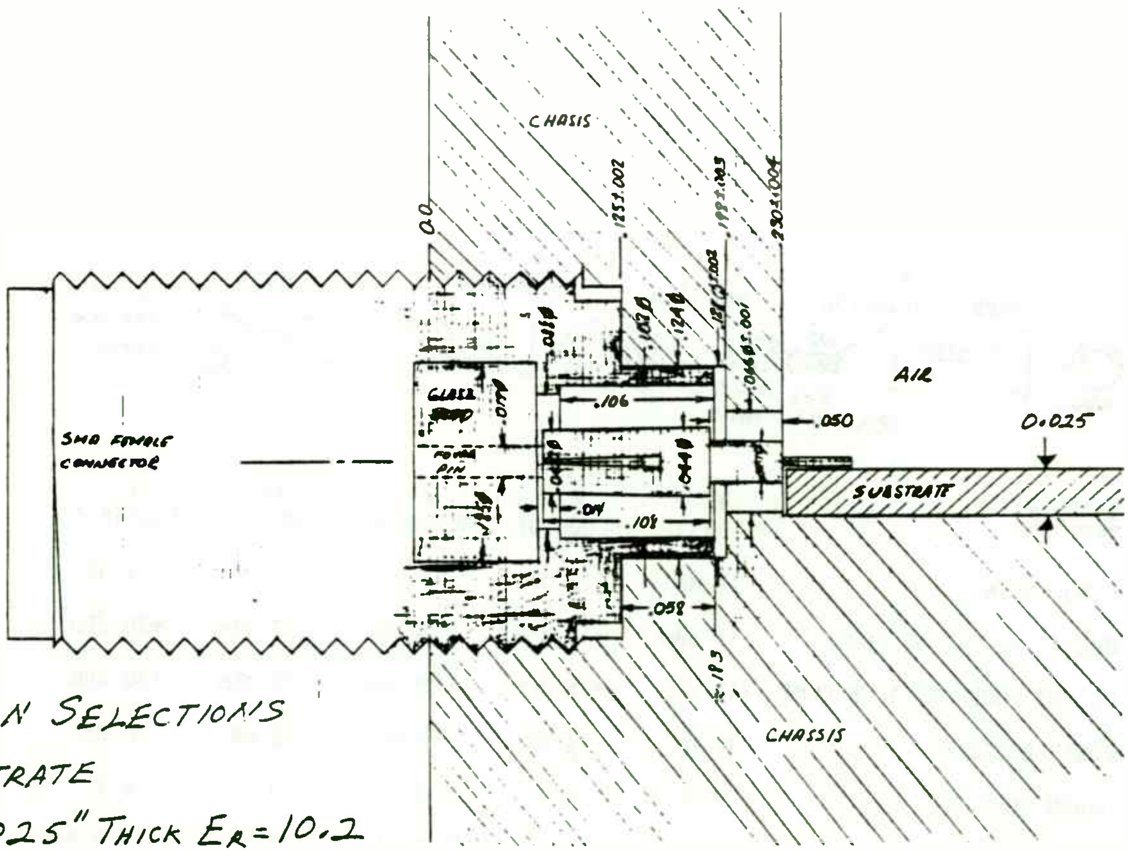
PROVISION FOR AUTOMATIC OUTPUT LEVEL CONTROL AND/OR TEMPERATURE COMPENSATED GAIN (BUILT IN PIN DIODE ATTENUATOR)

TABLE 1

ACHIEVED LINEAR POWER AMPLIFIER PERFORMANCE

POWER OUTPUT	38.0 dBW
GAIN	50 dB
THIRD ORDER INTERCEPT	+49.5 dBm
POWER CONSUMPTION	
3 AMPS AT	+9.5 V DC
0.05 AMPS AT	-12.0 V DC

TABLE 4



DESIGN SELECTIONS
SUBSTRATE
0.025" THICK $E_R = 10.2$

CONNECTOR INTERFACE

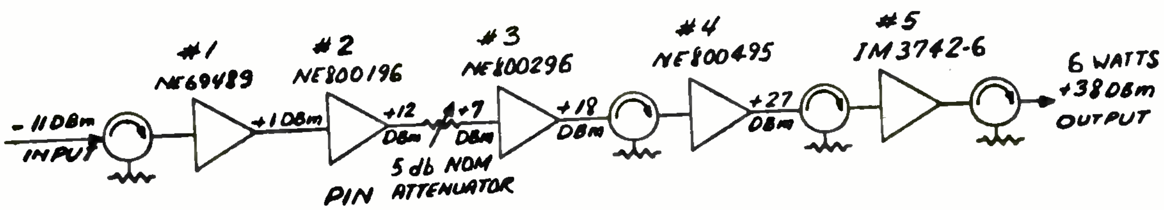
FIGURE 1

JUSTIFICATION FOR CHOOSING 1 dB GAIN COMPRESSION LOAD/SOURCE IMPEDANCES

IM3742-3 AT VD=9V ID=1.0A	SMALL SIGNAL GAIN dB	POWER OUTPUT AT 1 dB GAIN COMPRESSION dBm	TWO TONE THIRD ORDER OISTORTION LEVELS (DBC) AT DIFFERENT OUTPUT POWER LEVELS (dBm)			
			31.9dBm	30.2 dBm	26.8 dBm	24.8 dBm
TUNE FOR MAX POWER OUTPUT AT 1.0 dB GAIN COMPRESSION	11.5	+36.5	32	36	45	50.0
TUNE FOR MAXIMUM SMALL SIGNAL GAIN	13.85	+32.25	16	22	36.5	44.0
DIFFERENCE dB	-2.35	+4.25	16.0	14.0	8.5	6.0

CONCLUSION 4.25 dB MORE POWER IS MORE IMPORTANT THAN 2.35 dB LESS GAIN

TABLE 2



STAGE # / FET TYPE	#1 NE69489	#2 NE800196	#3 NE800296	#4 NE800495	#5 IM3742-6
STAGE GAIN	12 dB	11 dB	11 dB	9 dB	11 dB
OPERATING OUTPUT POWER	+1 dBm	+12 dBm	+18 dBm	+27 dBm	+38 dBm
OUTPUT POWER AT 1 dB COMP	+18 dBm	+26 dBm	+28 dBm	+32 dBm	+38 dBm
POWER BACKOFF	17 dB	14 dB	10 dB	5 dB	0 dB
DRAIN VOLTAGE	7 V	9 V	9 V	9 V	9 V
DRAIN CURRENT	30 mA	150 mA	235 mA	470 mA	2.0 Amp
DC POWER WATTS	0.21 W	1.35 W	2.11 W	4.23 W	18 W

AMPLIFIER LEVELS AND GAINS

TABLE 3

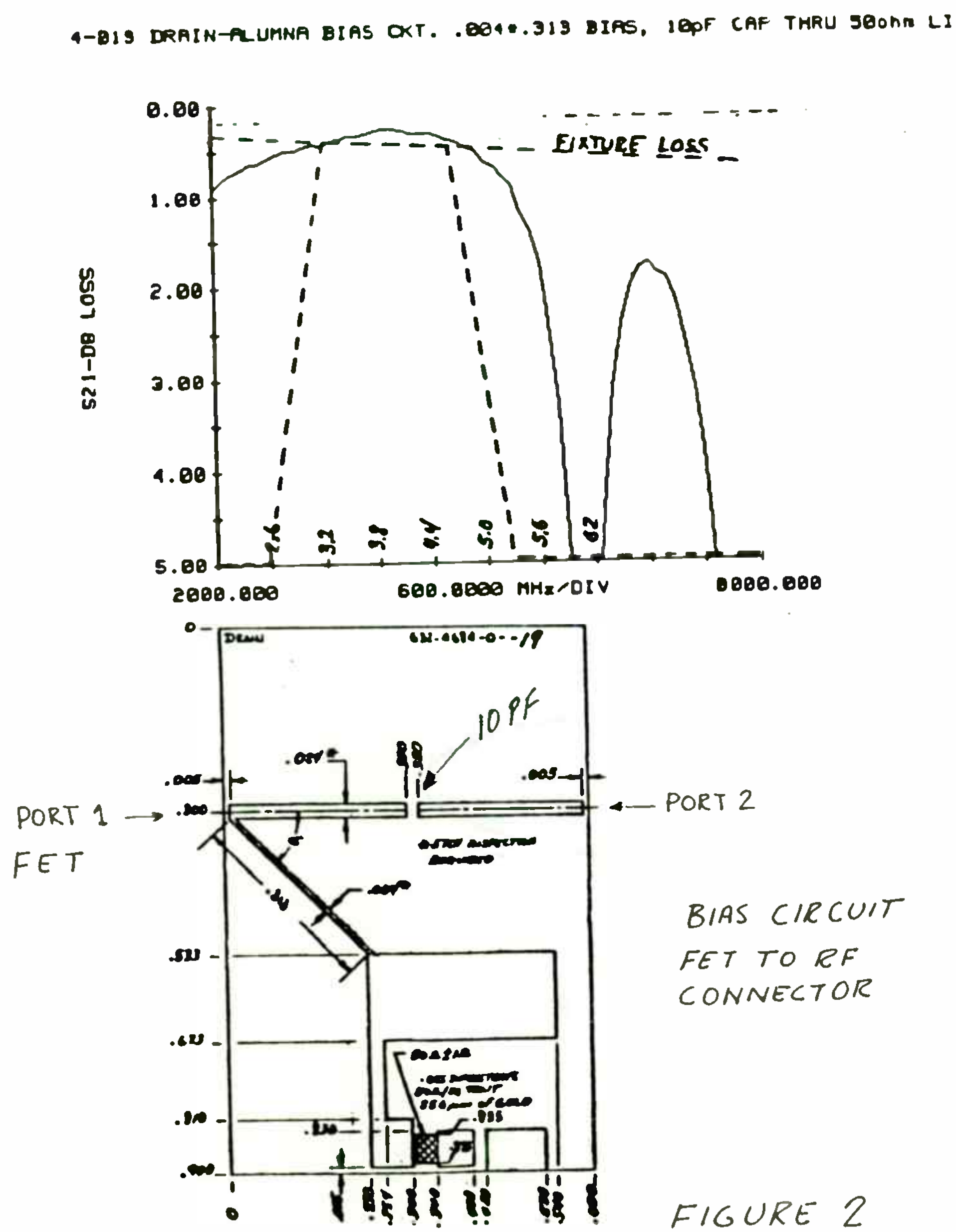


FIGURE 2

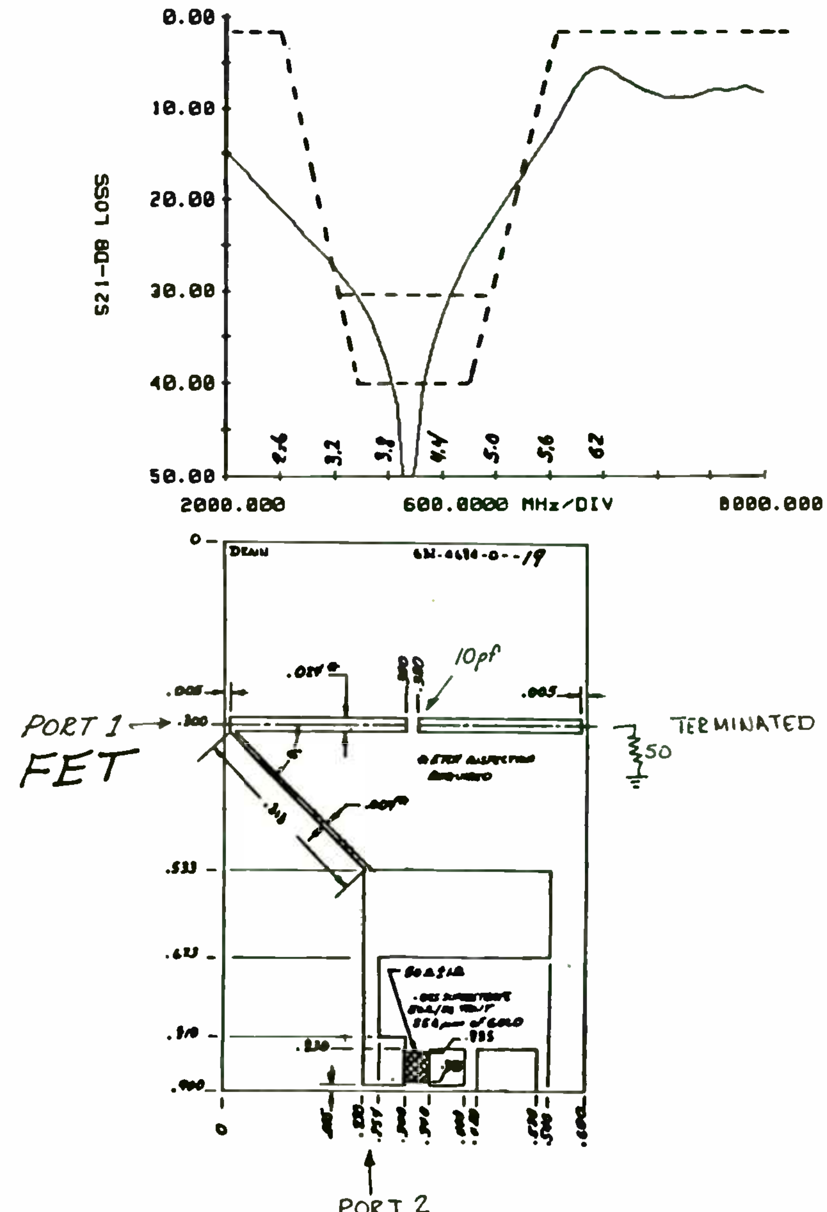


FIGURE 3

DC CONNECTION

CALCULATED AND MEASURED BIAS NETWORK - THRU TRANSMISSION (S_{21})

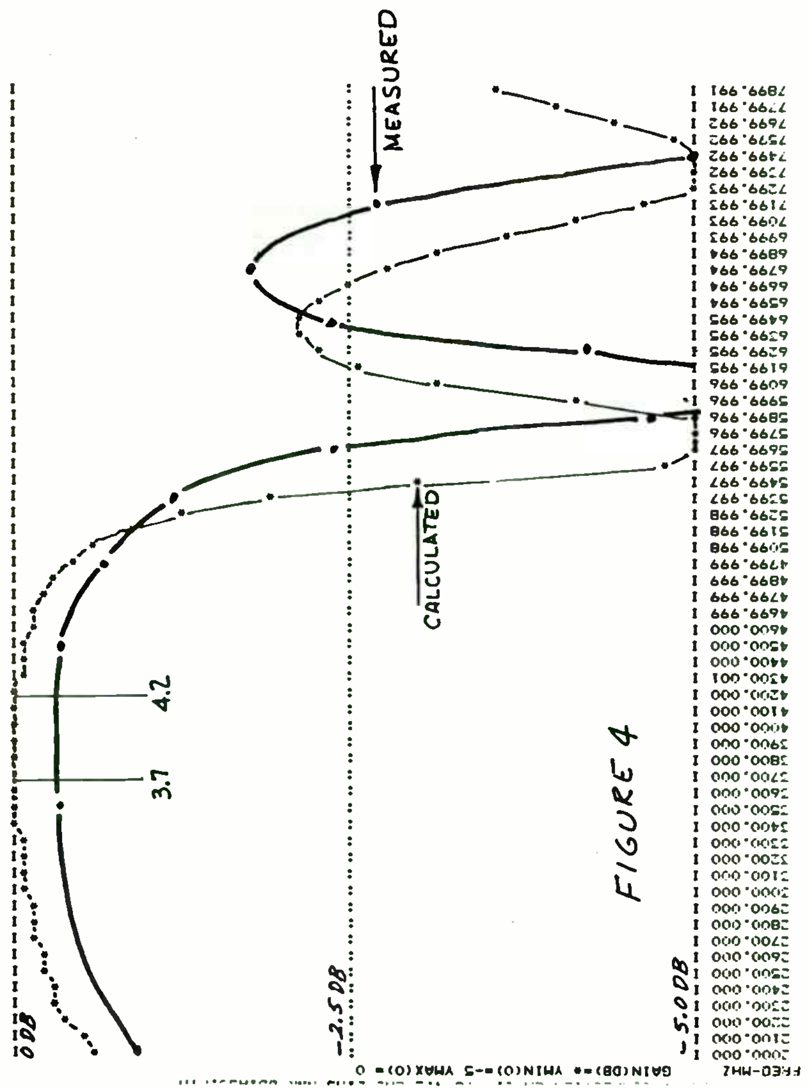
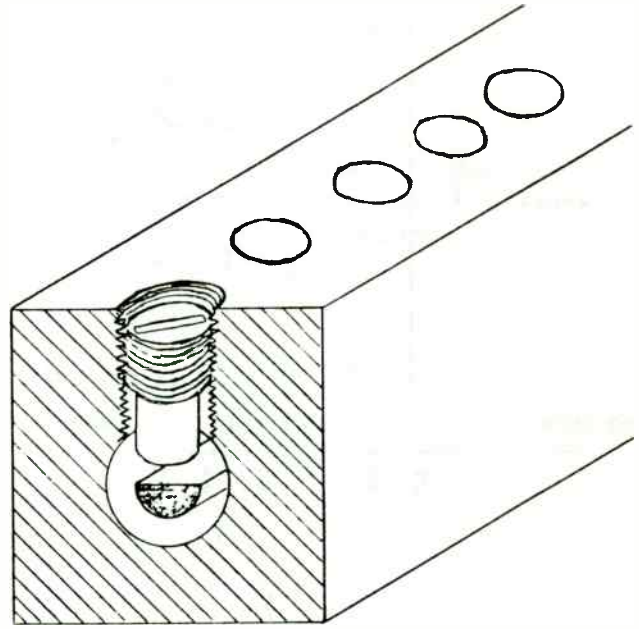


FIGURE 4



CROSS SECTION OF MULTIPLE SCREW TUNER

FIGURE 6

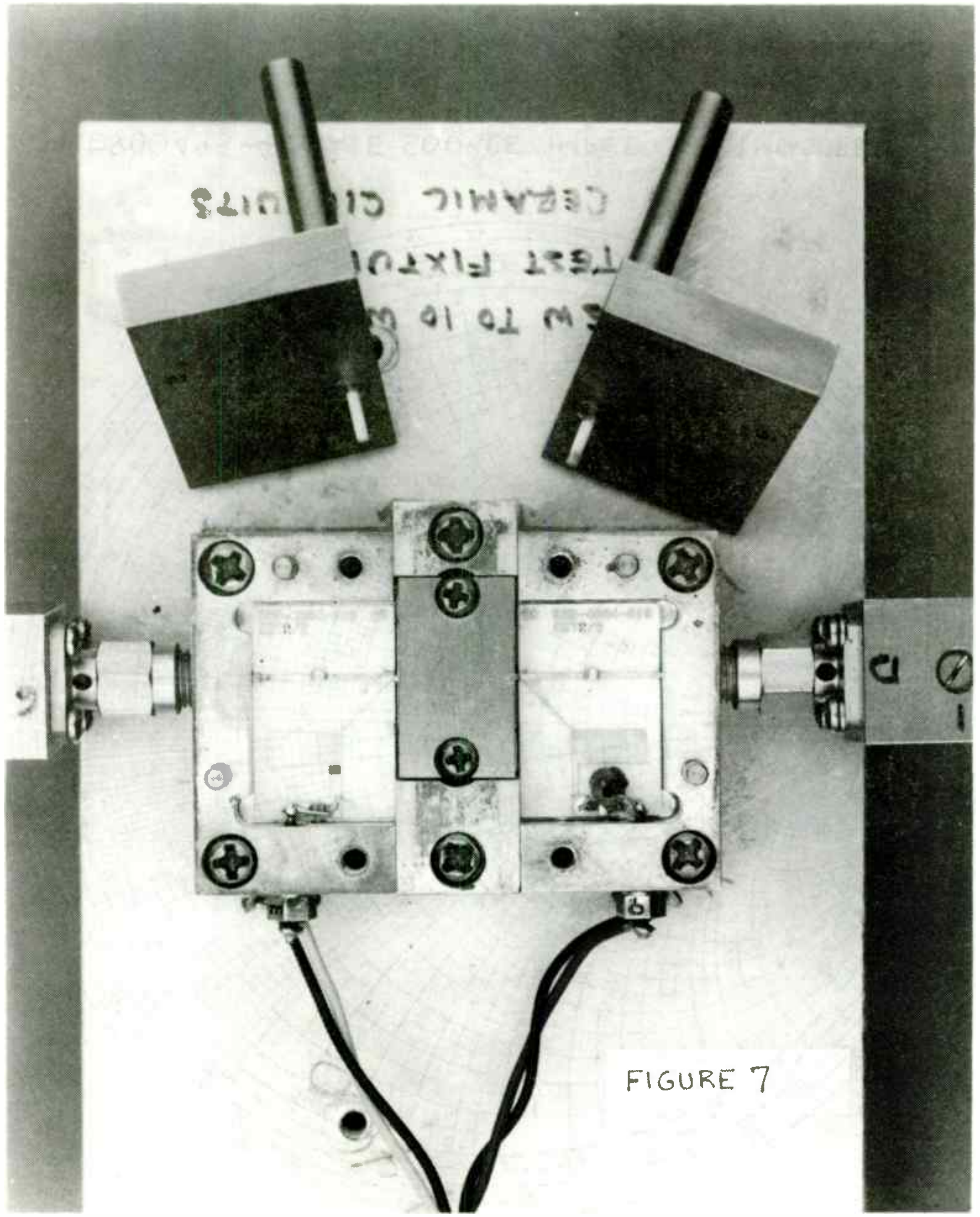
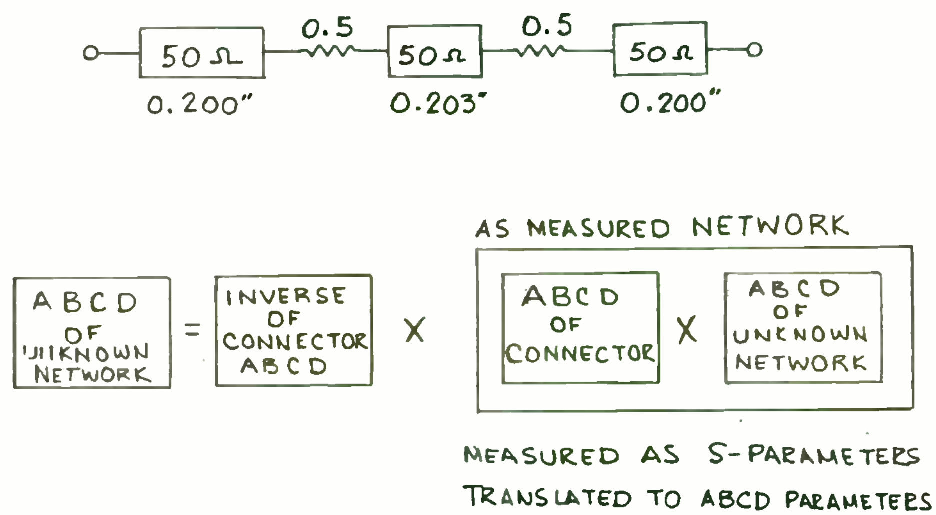


FIGURE 7

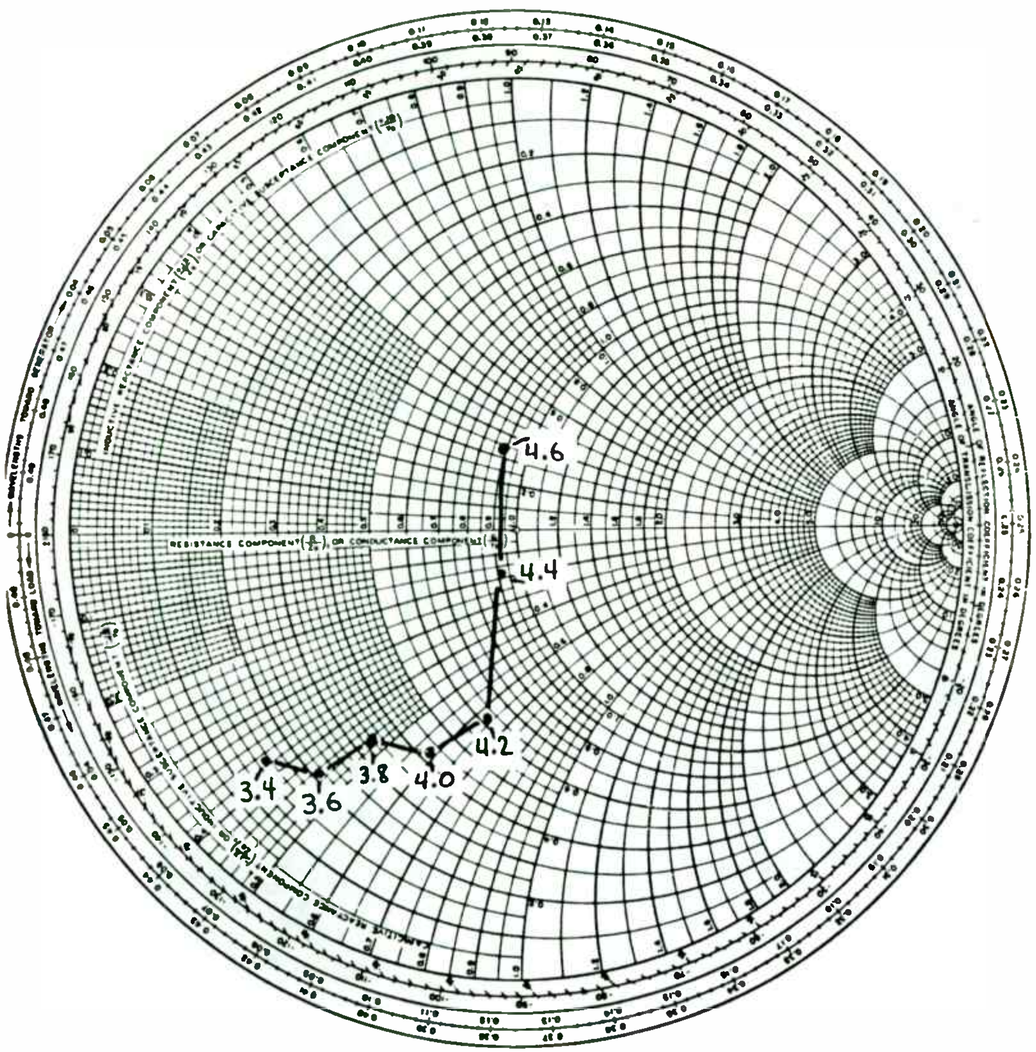
CONNECTOR EQUIVLENT CIRCUIT



OPERATION OF PROGRAM FTX33

FIGURE 8

IMPEDANCE COORDINATES



NE800495-4 GATE SOURCE IMPEDENCE (MEASURED)

FIGURE 9

FET GATE EQUIVLENT CIRCUIT

FET DRAIN EQUIVLENT CIRCUIT

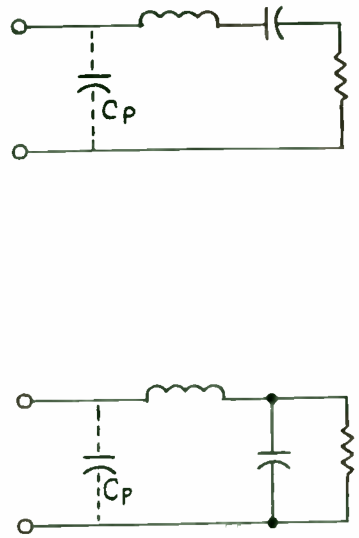
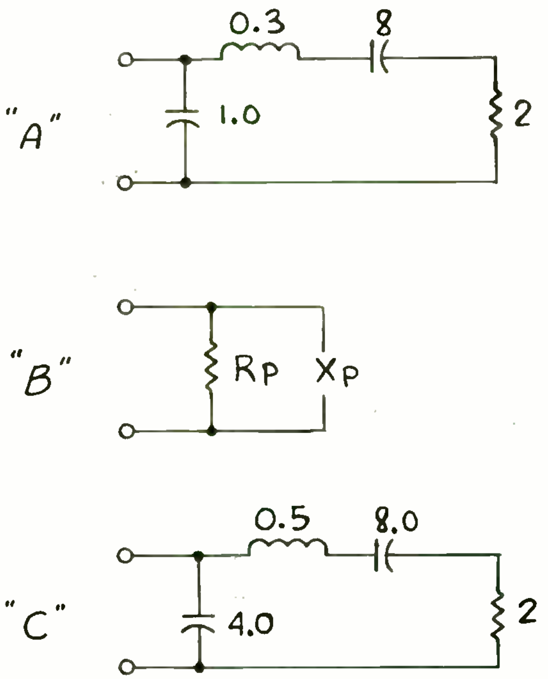


FIGURE 10

NE800495-4 GATE MATCHING DESIGN

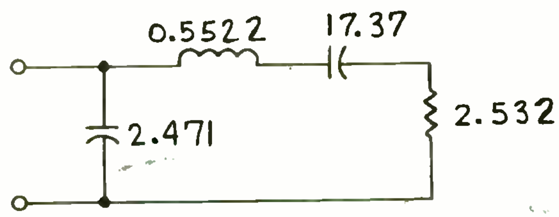


FIRST ESTIMATE OF GATE EQUIV CIRCUIT

SECOND ESTIMATE
NH, PF, OHMS
N4954G1. FRM

FIGURE II

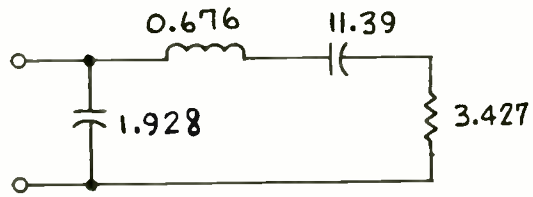
OPTIMIZED 3.4 → 4.6 GHz
USING FTRM



"A"

NH, PF, OHMS

OPTIMIZED 3.6 → 4.4 GHz

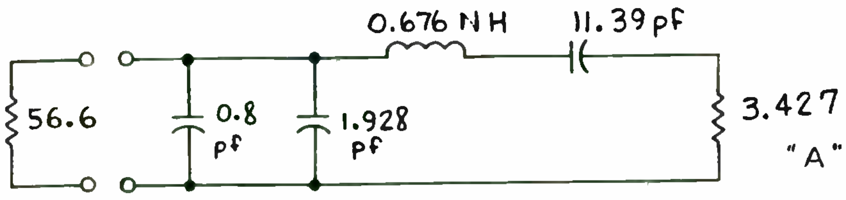


"B"

DESCRETE EQUIVLENT TO MEASURED
NE800495-4 GATE IMPEDANCE

FIGURE I

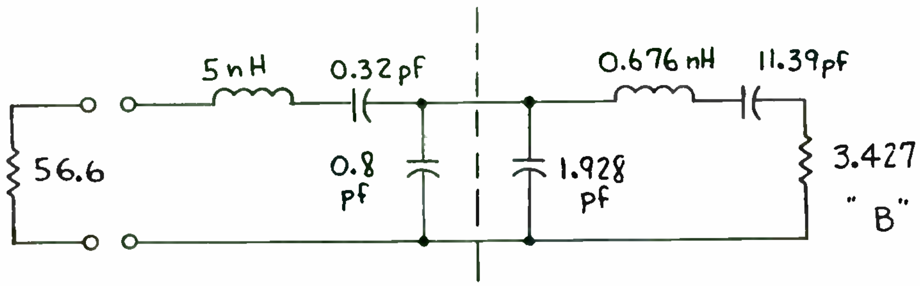
NE800495-4 GATE MATCHING DESIGN



3.6 GHz -0.90 DB   4.0 GHz 0.00 DB   4.4 GHz -0.87 DB

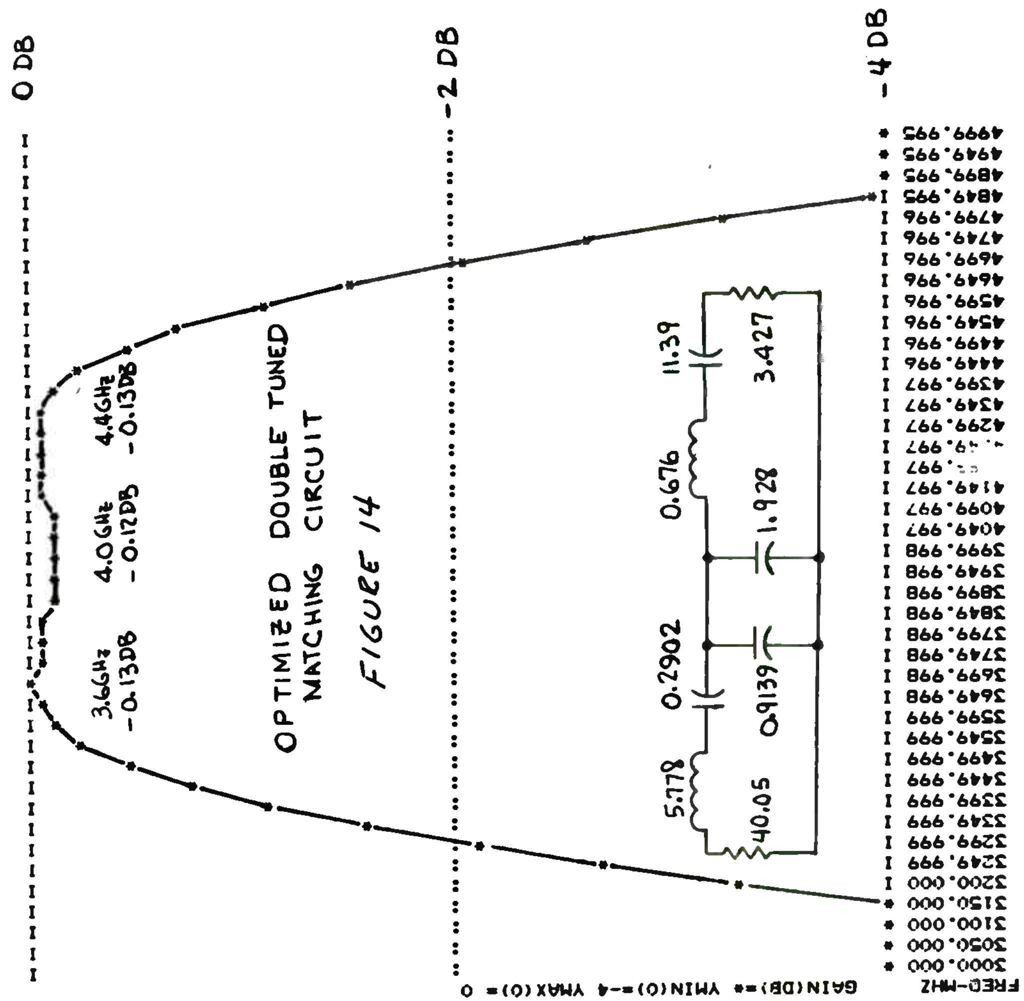
SINGLE TUNED MATCHING CIRCUIT

1.7 GHz 3 DB BW (5 NH ≈ 50 Ω AT 1.7 GHz)



TRIAL DOUBLE TUNED MATCHING CIRCUIT

FIGURE 13



OPTIMIZED DOUBLE TUNED MATCHING CIRCUIT

FIGURE 14

SMITH/POLAR CHART REFLECTION COEF/MAGNITUDE OF OUTER CIRCLE = 1
RESISTANCE VARIABLE= ZIN(R) REACTANCE VARIABLE= ZIN(I)
SMITH CHART NOMINAL IMPEDENCE= 40.05

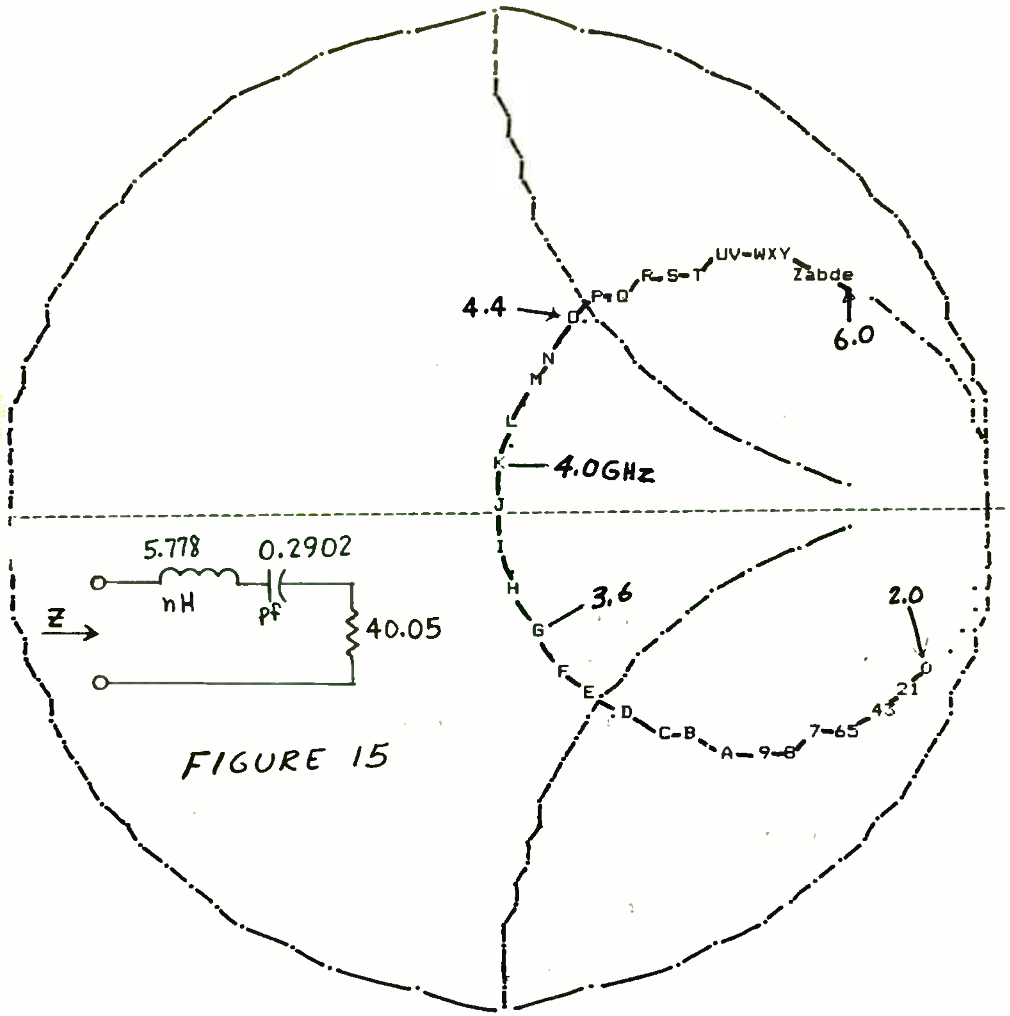


FIGURE 15

SMITH/POLAR CHART REFLECTION COEF/MAGNITUDE OF OUTER CIRCLE = 1
RESISTANCE VARIABLE= ZIN(R) REACTANCE VARIABLE= ZIN(I)
SMITH CHART NOMINAL IMPEDENCE= 50

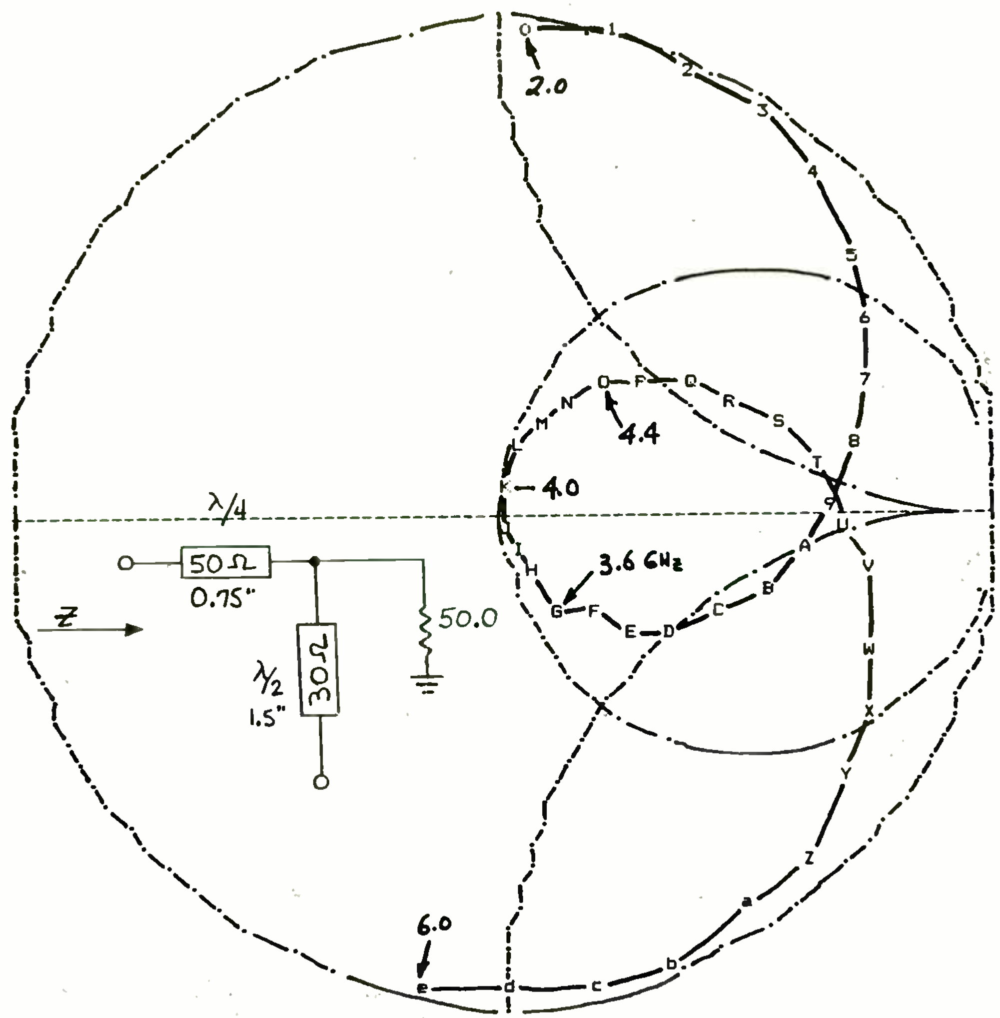
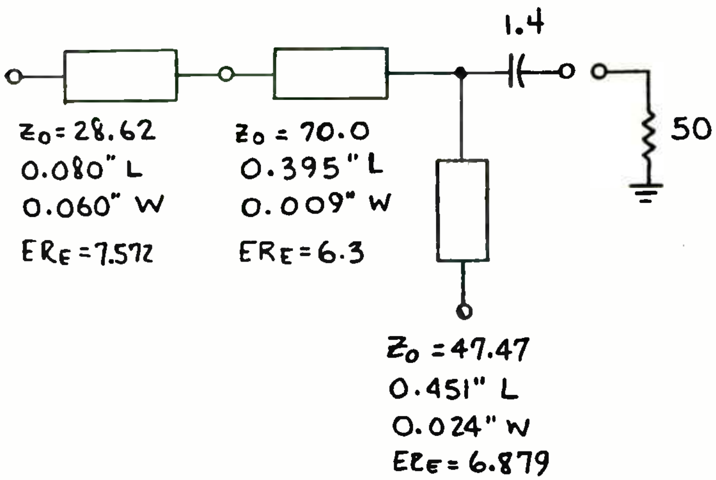


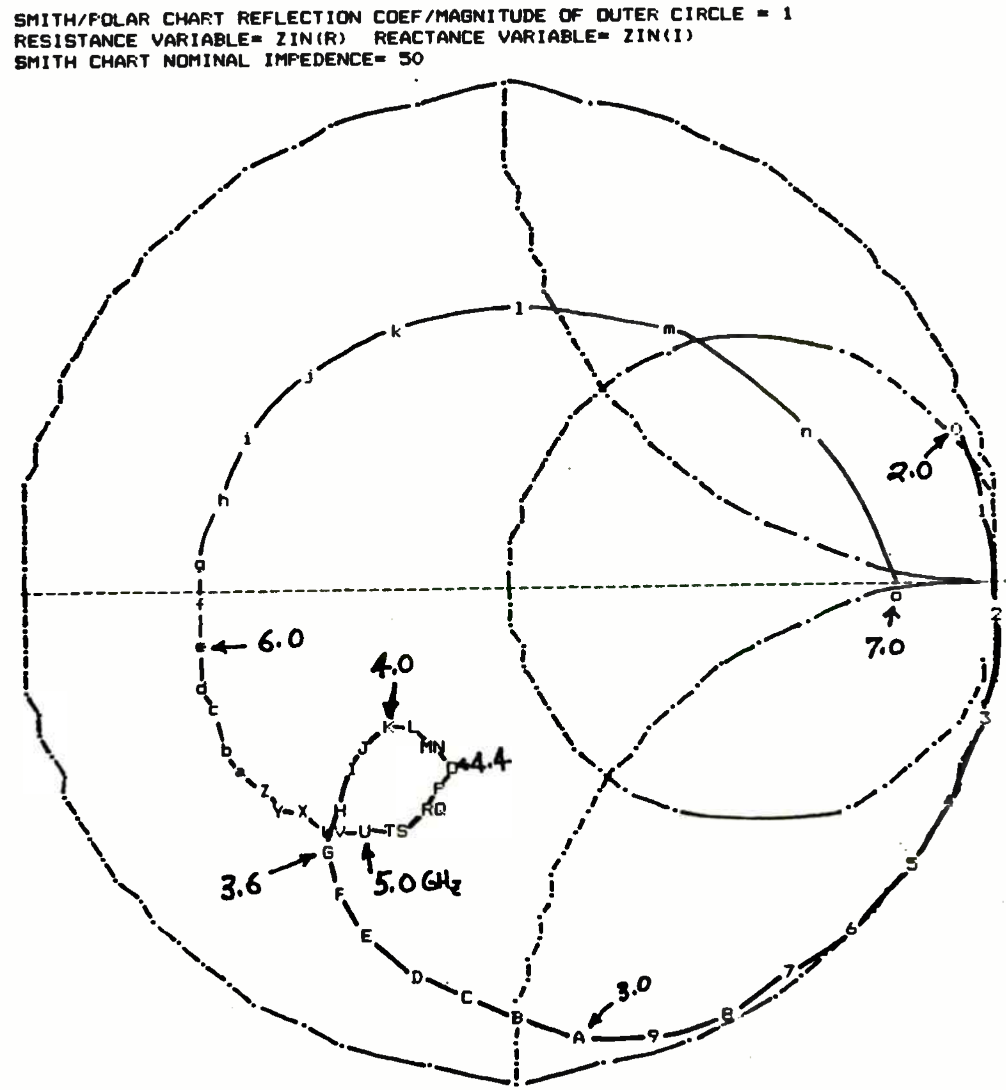
FIGURE 16



FINAL NE800495-4 GATE MATCHING NETWORK

BUILT ON 0.025" THICK E_R = 10.2

FIGURE 17



FINAL NE800495-4 GATE MATCHING NETWORK

FIGURE 18

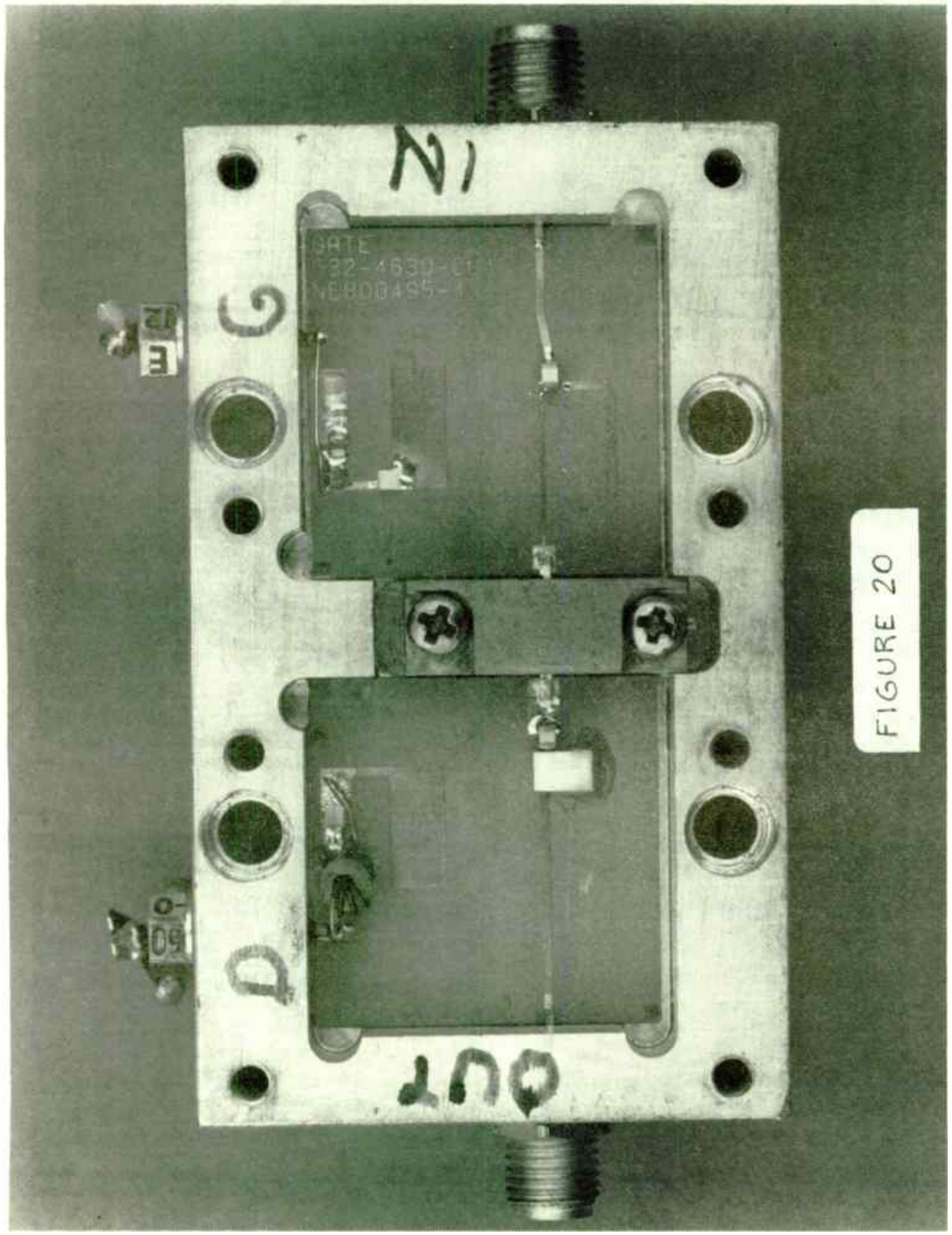
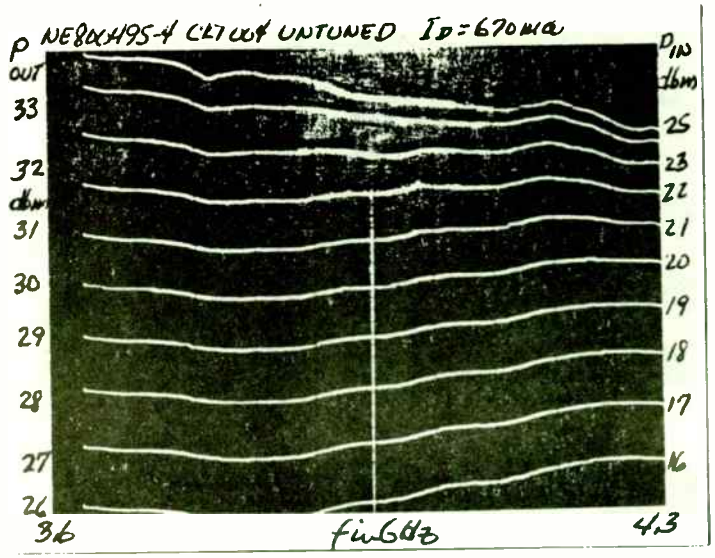
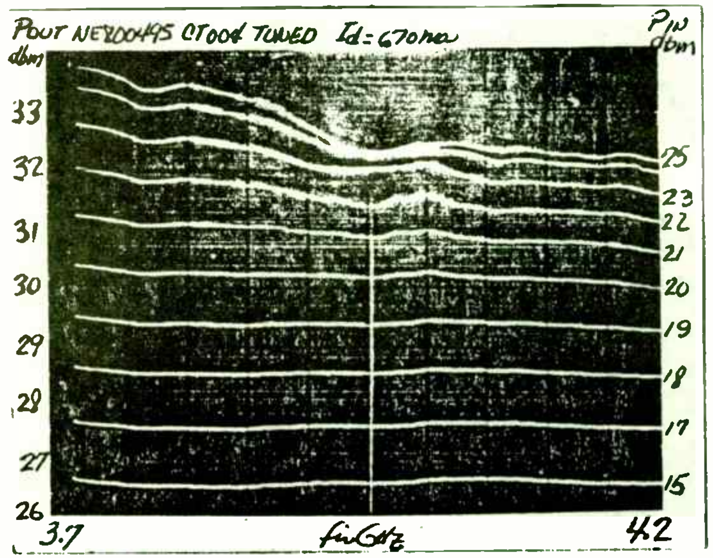


FIGURE 20

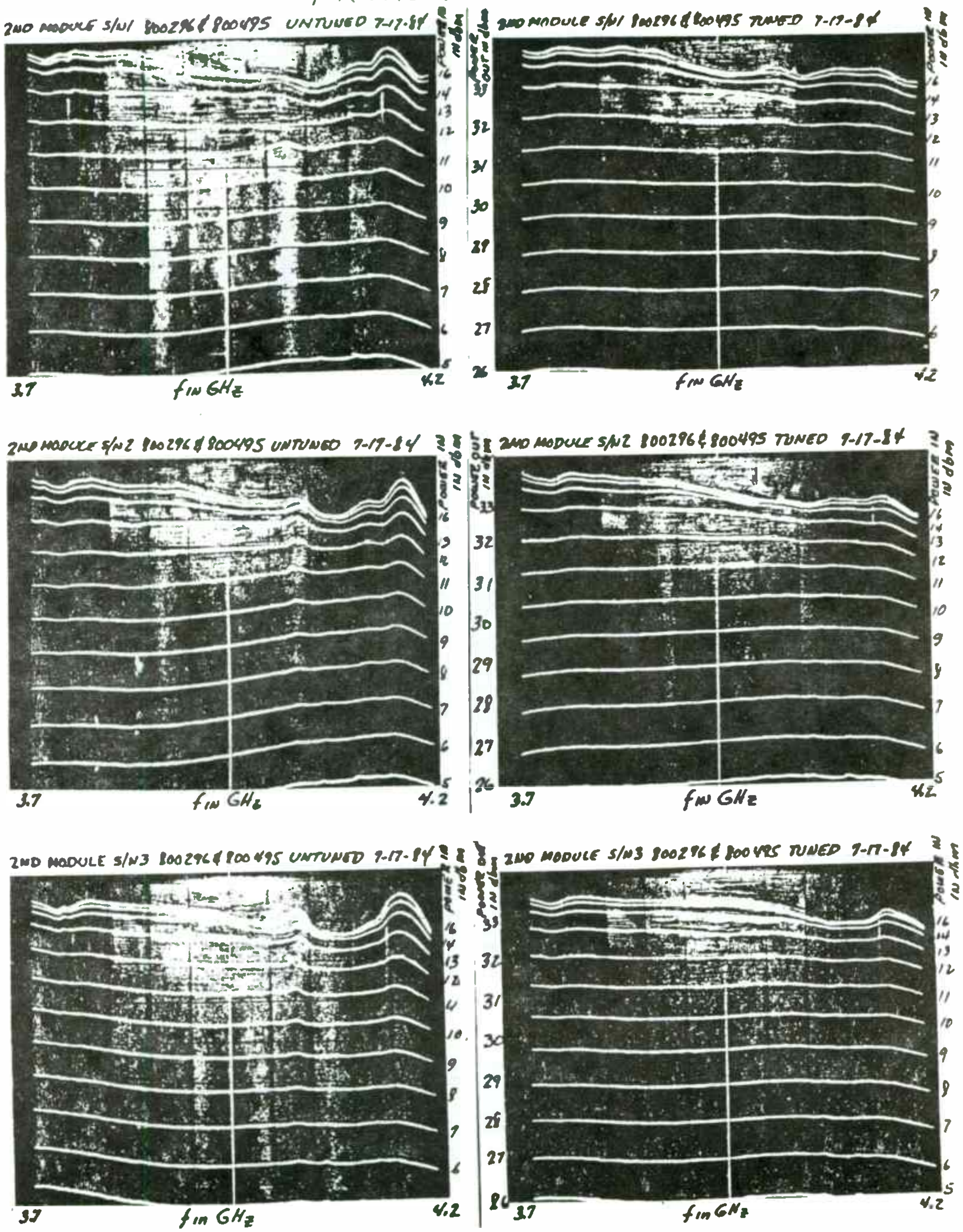


NE800495-4 CIRCUIT-004 UNTUNED

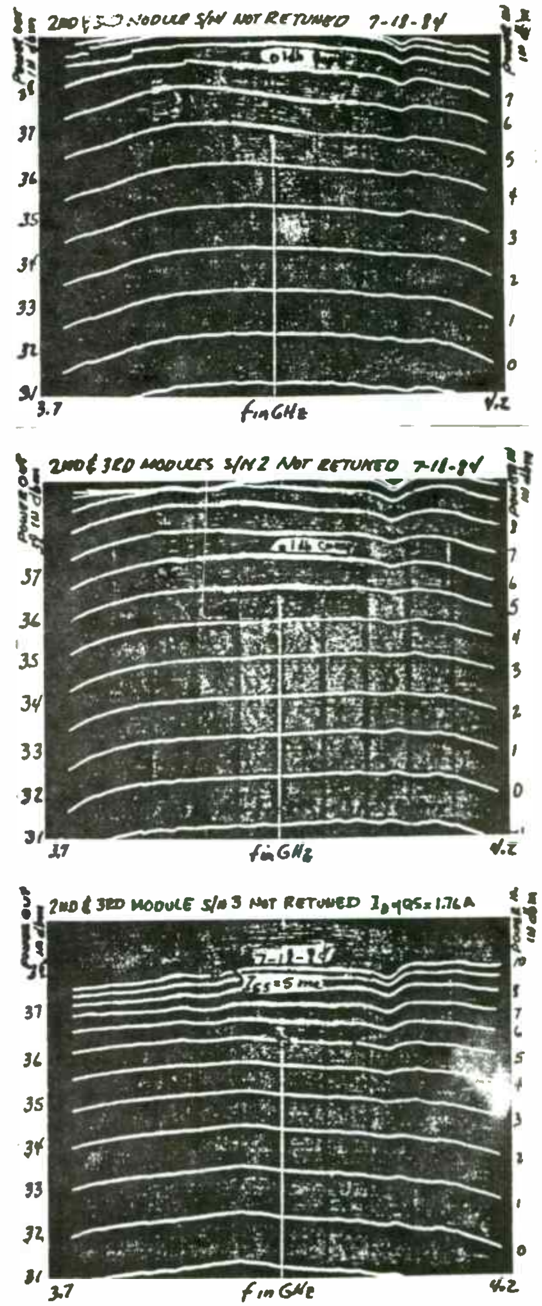


NE800495-4 CIRCUIT-004 TUNED

FIGURE 22



MEASURED PERFORMANCE OF NE800296 → NE800495-4 → JM3742-6 CASCADE



I_S = 2.0 A
I_Y = 300 ma
I_3 = 155 ma
@ +38 dbm OUT
& +7 dbm IN

I_S = 1.85 A
I_Y = 310 ma
I_3 = 235 ma
@ +38 dbm OUT
& +7 dbm IN

I_S = 2.0 A
I_Y = 477 ma
I_3 = 280 ma
@ +37 dbm OUT
& +7 dbm IN

FIG. 24

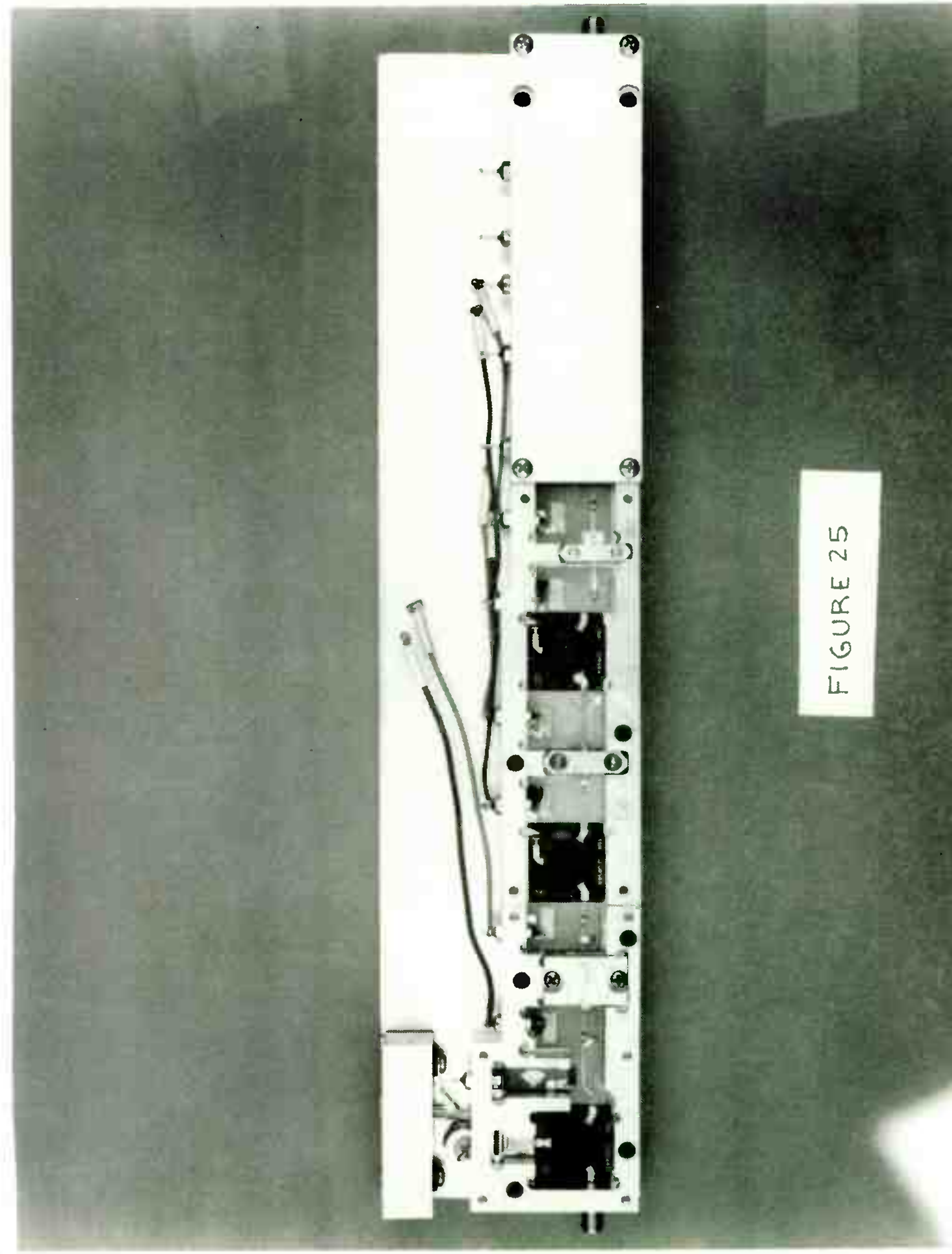


FIGURE 25

JON P. GROSJEAN
CONSULTING ENGINEERS
PULPIT ROCK ROAD
HCR BOX 110
SOUTH WOODSTOCK, CT 06267
203-974-2035

A NEW APPROACH TO PSK DEMODULATION

Introduction

Transmission of data using phase-shift-keyed (PSK) modulation offers the possibility of lower data error rates for a given signal-to-noise ratio than either AM or FSK modulation without the disadvantage of the large occupied bandwidth of FSK systems.(1) If crystal sontrol of frequencies is desired, it is also much easier to generate AM or two phase PSK signals than FSK signals.

Unfortunately, demodulation of PSK signals is not very simple and the phase ambiguity problem always exists. For example, if the PSK signal consists of either 0° or 180°, it is easy for a phase-locked-loop demodulator to lock on to the 180° phase and to produce 1s for the 0° phase when the complement of this is desired. This problem can be resolved in the digital portion of the receiver, but this can increase acquisition times and thus reduce the effective data rate as well as increase the complexity of the digital portions of the modem. In polled systems where many remote units are accessed by a central transmitter such as in CATV data transmission systems, the cost and acquisition time become critical, so a simple PSK receiver would be even more advantageous in these cases.

Current PSK Demodulators

The majority of PSK demodulators use a phase-locked-loop to lock on to the carrier, and a separate in-phase detector to do the demodulation. In its simplest form, we could have the arrangement in figure 1.

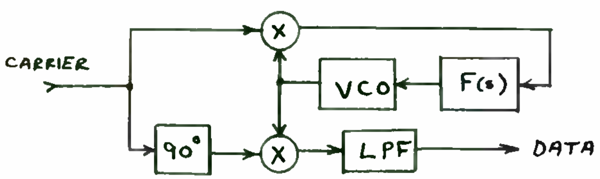


figure 1

Unfortunately, when the carrier changes phase 180°, the loop will follow it at a rate depending on the loop bandwidth. For example, if the loop is second order with a damping factor of 1, it will change from a phase error of 0° to one of 90° with a 180 degree input phase change in $\omega_n t$=.3 where ω_n is the loop natural frequency in radians per second.(2)Thus, if the data rate were 38,000 BPS and the loop had to be able to demodulate as many as 8 bits or 210 uS, we would need to make $\omega_n t$=.1. This gives:

t=210 x 10
ω_n =476
f_n = 75 Hz

If the loop bandwidth were this small, the acquisition time would be very long for any appreciable frequency error. The pull-in time can be given by the equation:(2)

$$T_p = \frac{(\Delta\omega)^2}{2\zeta\,\omega_n^3}$$

ζ = loop damping factor
$\Delta\omega$ = frequency error
ω_n = loop natural frequency

If Δf = 1.4 KHz ζ = .7
ω_n = 476
Tp= 512 mS

This is about ½ second, so unless long pull-in times can be tolerated, something must be done to make it more useful.

Almost all current approaches to PSK demodulation avoid this problem by changing the input phase changes to some multiple of 360 degrees. In a 2 phase system, this means doubling the phase shift. The simplest form uses a frequency doubler at the input, locks the VCO to 2xFin, and then divides the output by 2. Unfortunately, the output phase is ambiguous, and digital signal processing is needed to resolve this.(2) This also requires additional time and uses up data slots.

Three other types of loops, the remodulator, inverse modulator, and Costas loop operate by changing the signal applied to the VCO so that it is in the form of sin $2(\theta_i - \theta_o)$ instead of sin $(\theta_i - \theta_o)$ where θ_i is the input phase and θ_o is the oscillator phase. The most popular of these and hesiest to understand is the Costas loop shown in figure 3.

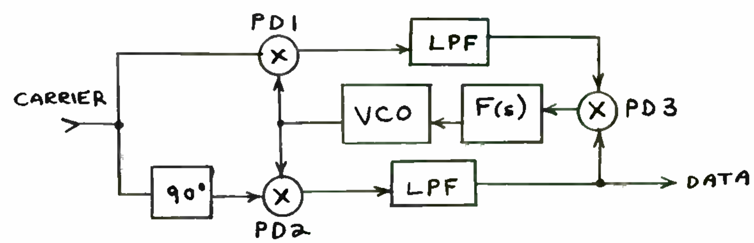


figure 3

The third multiplier or phase detector, PD3, simply inverts the phase of the VCO control signal when the data changes phase. Thus, the VCO control voltage remains constant with 180 degree changes in the input signal. Note also that the LPF is now inside the loop and must be included in the loop calculations unless its cutoff frequency is approximately 10 times the loop bandwidth. In systems having relatively low data rates and rapid acquisition, this is usually not the case. A more detailed description of the operation of these PSK demodulators is given in references 2,3, and 4.

Polled Systems

Polled systems are systems in which remote receiver-transmitters each have an identification or address and are polled by a central unit. An example of this might be in a CATV system where homes have security or meter reading devices installed. Each one would receive a signal addressed to it and send back a message when requested. Since most of the available frequencies are used for pictures, the those available for data are limited, and PSK or AM (OOK) modulation are more desirable.(5) PSK is often used in the return path because the signal-to-noise ratio is often poor and the variation in received levels requires limiting in the receiver.

Important characteristics or requirements of this type of system are:

1. rapid acquisition
2. short messages
3. good interference rejection

The rapid acquisition requires either a relatively large loop bandwidth or a small VCO frequency error. However, since these units must be low cost, it must be assumed that the frequency error will be .005% at best at each end, so at 14 MHz, this comes out to be 1400 Hz. This is quite large, so the loop bandwidth must be large.

The short messages almost preclude using some synchronizing sequence in the digital sections of the PSK receiver because it would greatly increase the response time of a large number of units.

Good interference rejection is required for two reasons: If there are many data channels occupying, for example, one TV channel, they might all be active. If the data rate were 38,000 BPS, the occupied bandwidth is about 76 KHz. Data channels could be spaced at 100 KHz and 60 data channels could fit into one TV channel. Now, if the PSK demodulator is very susceptible to spurious signals from the adjacent channels, the receiver IF filtering problem becomes more difficult. Thus, the ability of the PLL to reject spurious signals makes the receiver design simpler and reduces the possibility of errors.

The New Way

A new approach to demodulating 2 phase PSK signals starts with a simple technique: Disconnect the VCO from the loop when the phase changes 180 degrees. This, of course, puts some limits on the type of message which can be sent such as the maximum length of time 180° can exist, but since most polled systems are tailored to do a specific task, the digital designer can set up the software to avoid the limitations. The basic system is shown in figure 4:

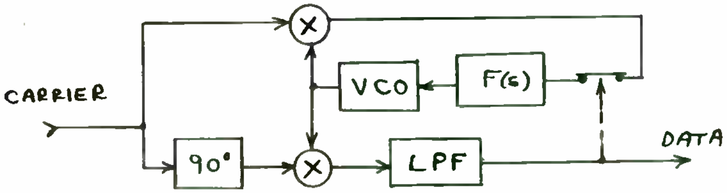


figure 4

The reference phase is the phase the loop first locks on to. Synchronization requires only the transmission of a start pulse long enough for the loop to lock. After locking, all data pulses are correct, and it is relatively easy for the system to demodulate data pulses where the 180° state lasts for ten 0° states. If synchronization is lost for some reason, it can be re-acquired by removing and re-applying the carrier or by sending a long reference pulse. Additionally, once locked, the demodulator appears to have much better spurious signal rejection than the Costas Loop. Measurements indicate that the rejection is equivalent to 15 to 20 dB more IF selectivity. Additionally, the loop itself remains simple, and only F(s) is the loop filter.

A schematic diagram of a 4.5MHz PSK demodulator is shown in figure 5. The PLL section consists of IC1, Q2, Q3 and associated components. Q3 and D1 form the VCO and Q2 improves the temperature stability of the oscillator. Q1 is used as a gate to disconnect the VCO during data pulses. T1 is used to shift the phase of the incoming signal 90 degrees, and IC2 is the in-phase or data detector. The no-signal output voltages on pins 6 and 12 of IC2 are approximately 9.1 volts. When a carrier is acquired, pin 6 of IC2 goes more positive and pin 12 more negative. When a 180° data pulse occurrs, the opposite happens. IC3, a dual comparator, is biased so that pin 7 goes high when a carrier is detected. Pin 1 is high and goes low when data pulses are received. Only data pulses appear at pin 1, and these are used to turn off Q1 as well as for the data out signal. Waveforms for a single data pulse are shown in figure 6.

This loop works well as shown but will sometimes have a longer than normal acquisition time if a carrier occurrs when the VCO is 180 degrees from where it should be to lock. This causes the gate to open and several beat cycles occurr before lock is achieved. A solution to this is to prevent the gate (Q1) from operating until the loop has locked. To do this, a carrier detect signal without data is necessary. (It should be noted that the carrier output pin in figure 5 also contains data pulses.)
If we let:

C = carrier
D = data
CD = carrier and data

In figure 5, the output from pin 7 of IC3 is CD, and from pin 1, it is D. Using logic gates, it is relatively easy to produce a signal containing only C. This is shown in figure 7:

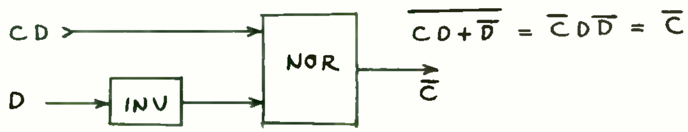


figure 7

$\overline{C}$ can now be delayed to produce a D signal which cannot occur until $\overline{C}$ has been present long enough for the loop to lock. Figure 8 shows the arrangement used using a CD4001BE quad NOR gate. Note that the output of the 4th NOR gate is labeled Dq for quiet data. When no signal is being received and the receiver has a high gain limiter, both CD and D will be noisy, but Dq has no output until lock is achieved. Thus, it is noise-free data. Another advantage of this arrangement is the signal labeled $\overline{C} + \tau$ across C1. This is a very convenient point for triggering an oscilloscope to observe a received signal. Without this, it is very difficult to see how the receiver is working because all the signals have noise levels as large as signals until the loop locks. Dq occurrs too late and varies in position depending on the data message. A summary of the received signals is shown in figure 9.

Now that $\overline{C} + \tau$ is available, it can be used to change the loop bandwidth so acquisition is faster and longer data pulses can be received. Figure 10 shows a block diagram of one possible implementation. The ULN2241A is an AM-FM radio IC and is being used here in the FM mode. The internal FM detector is used for the loop phase detector by driving the detector input with the VCO instead of a quadrature coil.(6) For this phase detector, Kd = 3.8V/radian. The VCO is the same as the one in figure 5, and a CD4066B CMOS gate is used instead of a JFET. The parameters for the loop are:

$K_o = 577 \times 10^3$ rad/sec/V
$K_dK_o = 2.2 \times 10^6$

In the wide band mode: R1=1.9K R2=470 C=.47 μF
Fn= 7.07 KHz ζ = 5.41

For Tp=1 mS (pull in time), the frequency error can be 155 KHz.

In the narrow band mode: R1=100K R2=470 C= .47μF
Fn= 1.09 KHz ζ = .76

The S042P IC shown is a balanced mixer in which the lower transistors cam be made into an oscillator. It is used to keep the LC VCO locked to 4.5 MHz when no signal is being received. Its loop gain is low, and when a carrier is received, it is disconnected from the main loop by the $\overline{C} + \tau$ signal opening S1. This was found to be easier than building a crystal VCO with enough range required for the system.

The PSK demodulator was designed for a 38,000 BPS data rate. The lock up time, τ, is 1 mS, and the maximum 180 time is over 2 mS corresponding to 76 bit times. Field testing on CATV systems indicates that the overall system bit-error-rate is less than 1 in 10^8.

References

1. Garner, William J. "Bit Error Probabilities Relate to Data Link S/N " Microwaves, Nov. 1978 pp 101-105

2. Gardner, Floyd M. " Phaselock Techniques " Second Edition John Wiley & Sons 1979

3. Steber, Mark J. PSK Demodulation Techniques Provide Lowest Probability of Error" Microwave Systems News June 1984 pp. 150-176

4. Steber, Mark J. " Understanding PSK Demodulation Techniques" Microwaves & RF, March 1984 pp. 137-145

5. GrosJean, Jon P. " Selection of an Optimum Modulation Scheme For CATV Data Transmission" NCTA 1982 Convention Technical Papers pp. 69-71

6. GrosJean, Jon P. " Phase Locked Loops Using Quadrature Detector Integrated Circuits" IEEE Transactions on Consumer Electronics, Feb. 1976 Vol. 22 pp. 95-98

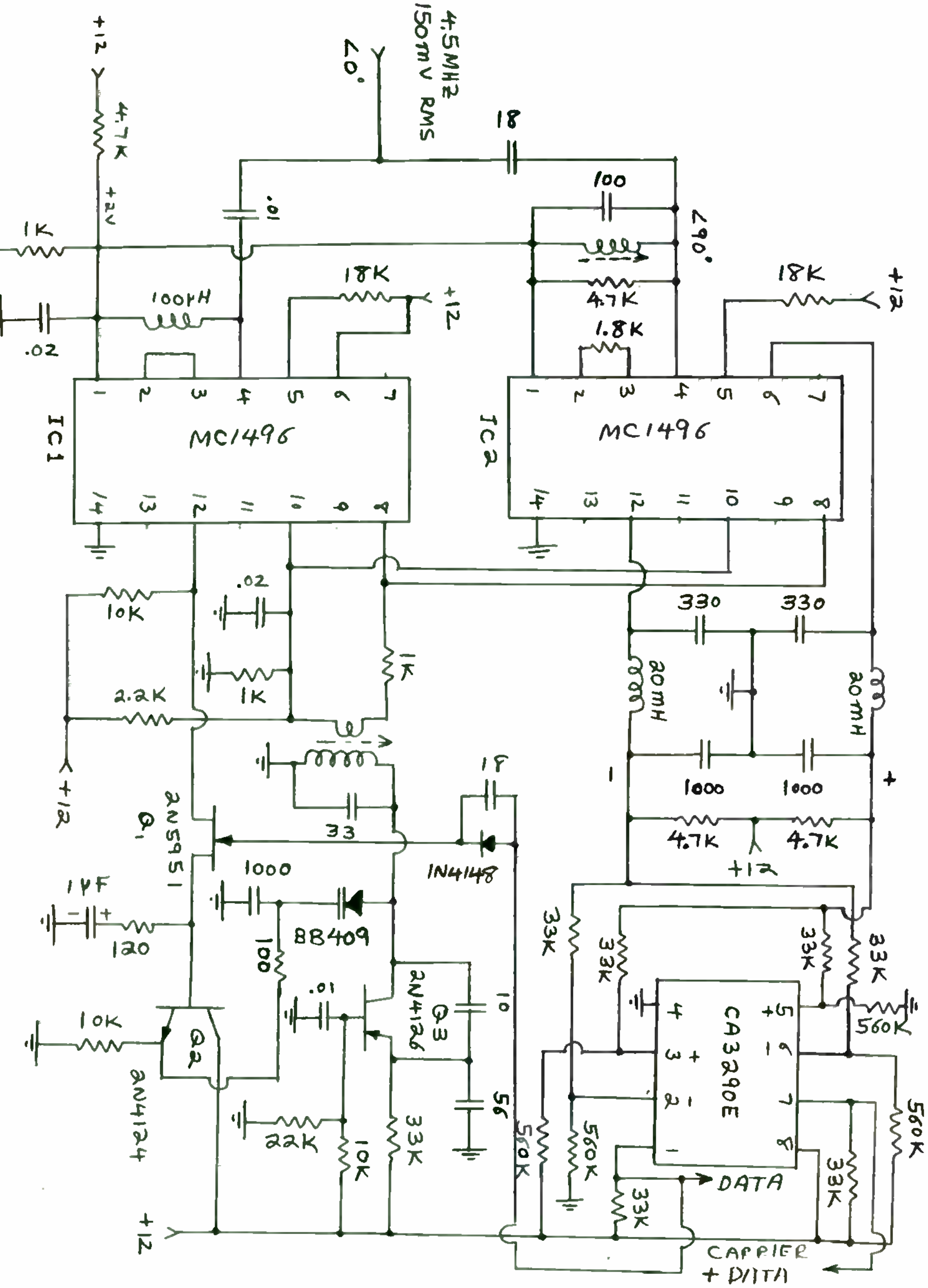


FIGURE 5

PSK DETECTOR

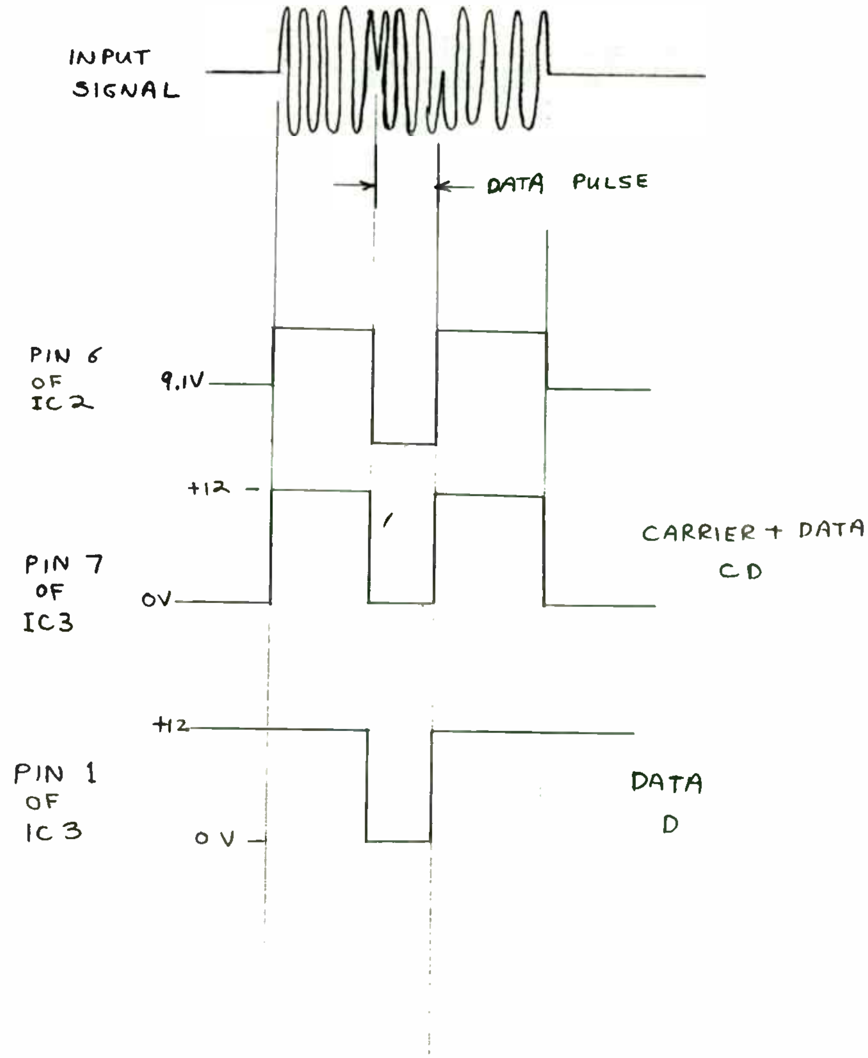


FIGURE 6

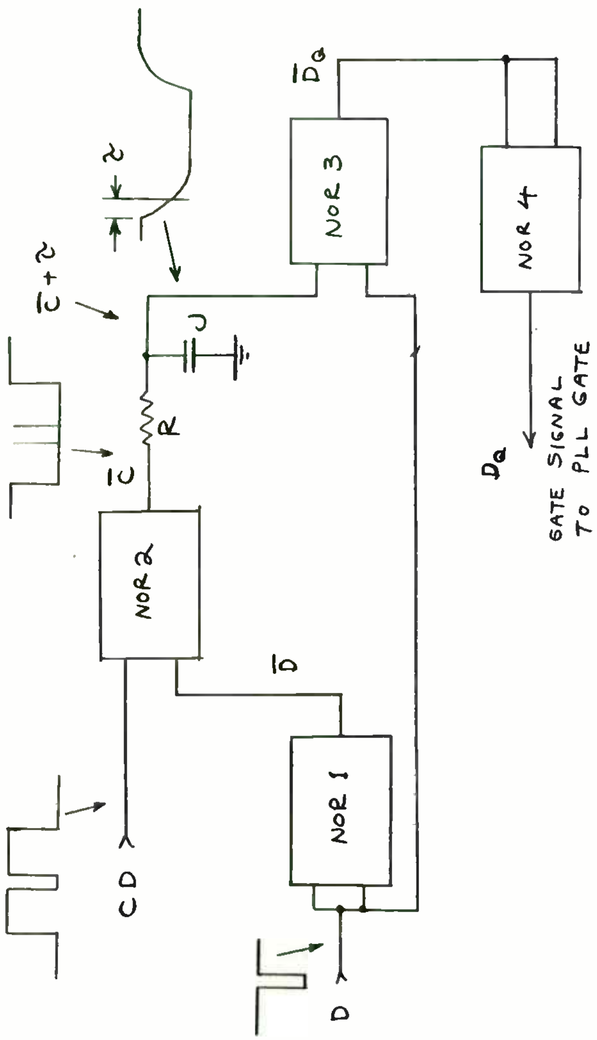


FIGURE 8

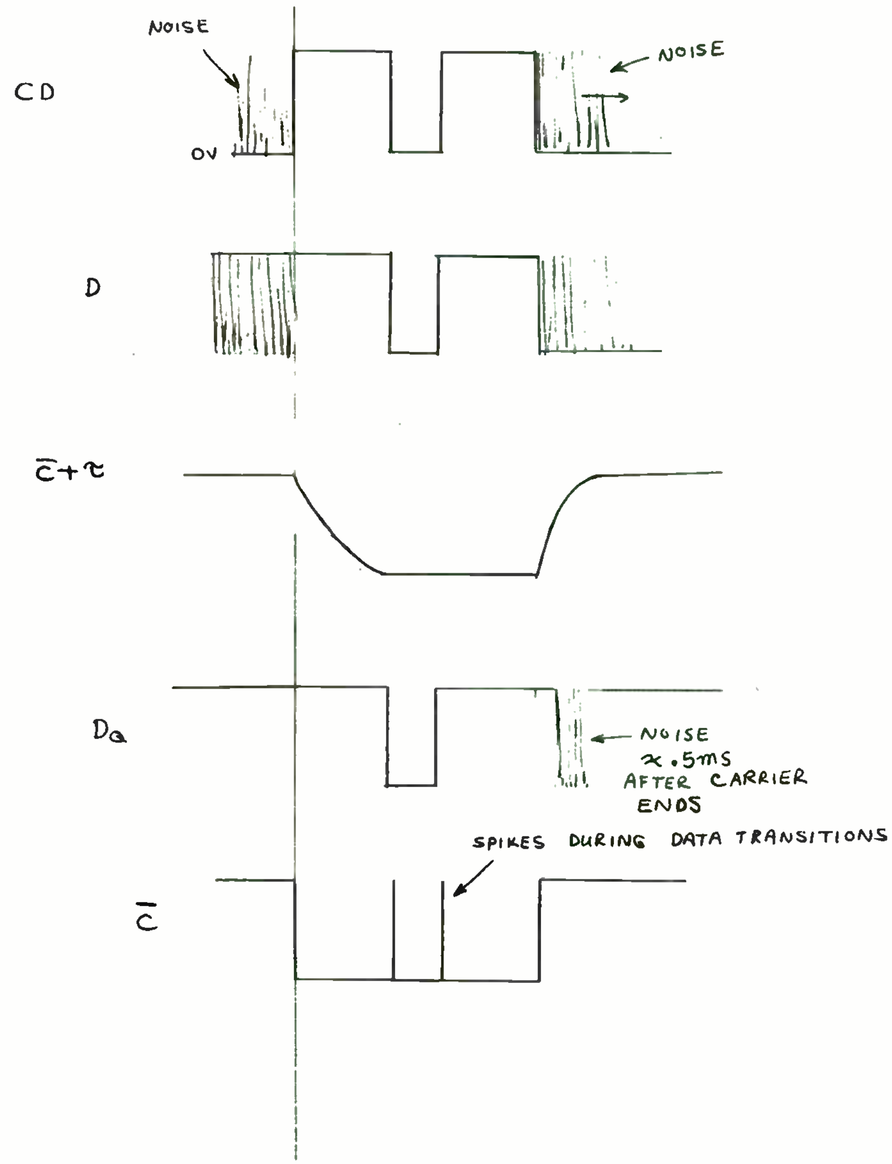
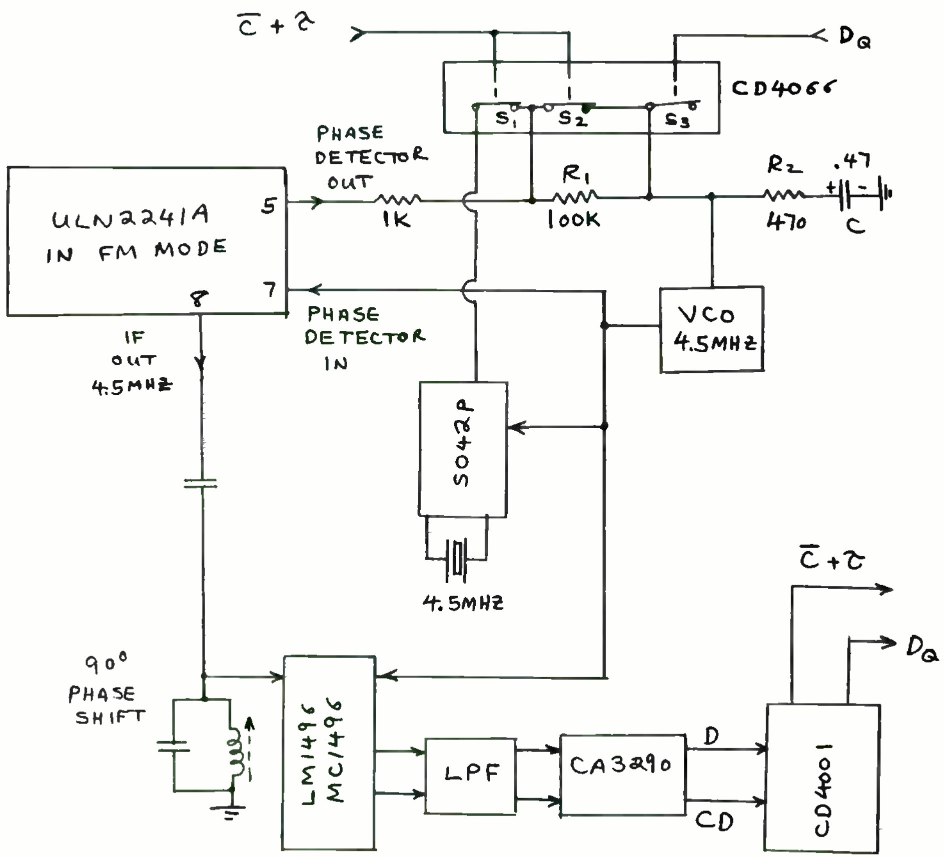


FIGURE 9



$\overline{C}+\tau$
D_Q
CD4066
S_1
S_2
S_3
PHASE DETECTOR OUT
ULN2241A IN FM MODE
5
7
8
1K
R_1
100K
R_2
470
.47
C
PHASE DETECTOR IN
VCO 4.5MHZ
IF OUT 4.5MHZ
SO42P
4.5MHZ
90° PHASE SHIFT
LM1496 MC1496
LPF
CA3290
D
CD
CD4001
$\overline{C}+\tau$
D_Q

FIGURE 10

Performing EMI/EMC Evaluation of Electronic Equipment Using TEM Cells

M. L. Crawford
National Bureau of Standards
Boulder, Colorado 80303

Abstract

This paper summarizes the basic physical and electrical properties which influence the design, operation and use of a transverse electromagnetic (TEM) cell. Guidelines are given for using a TEM cell for both radiated susceptibility and emissions measurements of electric components and equipment. The paper describes the test setups and outlines the test procedures, step by step, indicating precautions to observe to improve the repeatability and the limitations inherent in using TEM cells. Examples of some applications are then given.

Key words: electromagnetic compatibility measurements; radiated susceptibility and emissions; TEM cells.

1. Introduction

The use of electrical and electronic devices is having an increasing impact upon 20th century life. The broad use of such devices can be seen in nearly every aspect of life, from processing and development of raw materials to their use as consumer products. With this explosion in high technology has come the significant challenge of ensuring reliability. Nearly all types of electronic or electromechanical equipment emit and/or are susceptible to electromagnetic (EM) radiation. Compatible operation of such equipment in the presence of EM interference (EMI) is dependent on the ability to accurately measure and characterize its EM compatibility (EMC) profile and then to effectively control it, and/or to shield against it.

A number of techniques have been developed to measure levels of EM emanations or to determine the susceptibility or immunity of equipment to EMI. These measurements are typically made either on open-field sites where residual noise levels are low or inside shielded enclosures where electrical isolation or shielding is obtained. Presumably, open-field sites provide both a reference and a perturbation free environment in which actual operating conditions of the equipment can be simulated. However, open-field sites do not provide isolation from the environment and hence can only be used under special circumstances. If shielded enclosures are used, serious measurement errors (i.e., as large as ±40 dB [1]) can occur. These errors are due to the high conductivity and reflectivity of the enclosure walls which set up standing waves which interfere with the signal being measured and/or cause large gradients in a susceptibility test field. In addition, the radiating or receiving characteristics of the antennas used in making measurements inside shielded enclosures are altered by the close proximity to the equipment under test (EUT) and the confining metallic enclosure.

Another limitation inherent in using antennas for EMC/EMI tests is their limited bandwidth. Even so-called broadband antennas with reasonably flat amplitude response typically do not have a linear phase response versus frequency. Hence they are useful primarily for frequency domain measurements and have very limited application for transient impulsive EMI testing and evaluation. In addition, for accurate measurements, the separation distance between the antenna and the EUT should be sufficiently large to ensure far-field conditions. This is not always possible, especially in confined chambers or enclosures or at low frequencies.

Some of these limitations and problems can be eliminated or minimized by the use of a TEM cell. A cut-away view of a TEM cell is shown in figure 1. The cell is essentially a 50 ohm triplate transmission line with the sides closed in to provide electrical isolation from the surrounding environment. The line is tapered at each end to maintain a 50 ohm impedance all the way to standard 50 ohm connectors at the two ports of the cell. One of the ports is usually terminated with a 50 ohm load, while the other port is connected to either a rf source or a receiver depending on whether the cell is used for radiated susceptibility or emission testing. The cell then serves as the transducer for either establishing the test field between its center and outer conductors or for detecting the radiated fields from the EUT, thus eliminating the use of antennas.

To support a TEM mode, the cell is necessarily a two-conductor system with the region between the inner and outer conductors (either upper or lower sections) used as the test zone. Typically, however, the EUT is placed inside the cell in the lower section centered either near the floor or midway between the center plate and the floor as shown in figure 1.

The TEM cell has shown considerable potential for performing EMC evaluation and calibration of electrically small equipment and devices [2]. They are portable, simple to build [3], useful for broadband swept frequency measurements, and capable of providing test field strengths simulating planar fields from a few μ V/m to a few hundred V/m. They have proven to be useful for EMC evaluation for electronic components [4] in the frequency range from a few kHz to a few hundred MHz. Cost to build a TEM cell is typically much lower than that for conventional facilities such as anechoic chambers and shielded enclosures.

This paper summarizes some of the basic TEM cell properties and design. It discusses TEM cell EMI/EMC measurement procedures and outlines the steps necessary to obtain meaningful, repeatable results. Some applications are suggested and limitations of its use are given. It is important to note that TEM cells are intended primarily for use in diagnostic testing to determine, for example, frequencies at which the EUT is susceptible, or, frequencies at which the EUT emits radiations. Test results also should give some indication of how EMI is coupled into or from the EUT, and the relative improvement in EMC characteristics that may result from efforts to improve an EUT's immunity or to reduce its emissions. It is not intended for use, (except under some limited conditions), in determing EUT susceptibility to absolute field levels or the absolute amplitude of the EUT's radiated emissions, particularly if the EUT includes long wiring harnesses that must be exposed, polarization matched, to the test field.

2. TEM Cell Properties and Design

TEM cells have been designed with different cross sections depending upon their intended use. A square cross section is used to maximize the test area at a sacrifice of test field uniformity. Increasing the width to height ratio improves the test field uniformity but with reduced vertical test area for the same upper useful test frequency.

The design of a rectangular TEM cell is shown in figure 2. Figure 3 gives the pertinent dimensional relationship for constructing 50 ohm TEM cells of various cross sections. The characteristic impedance, Z_0, of a rectangular TEM cell is given approximately as [5]

$$Z_0 = \frac{120\pi}{4\left[\frac{a}{b} + \frac{2}{\pi}\ln\left(\sinh\frac{\pi g}{2b}\right)\right] - \frac{\Delta C}{\epsilon_0}} \qquad (1)$$

where the cell parameters a, b, and g, all in meters are indicated in figures 2 and 3. ΔC is the fringe capacitance between the edges of the center plate and the side walls of the TEM cell. Under practical conditions, $a/b>1$, $\Delta C/\epsilon_0$ is negligible.

The TEM cells at the National Bureau of Standards were designed using expression 1 to obtain, to a first order approximation, matched 50 Ω transmission lines. A time domain reflectometer was then used to measure the distributed impedance of a particular TEM cell and to make necessary adjustments to obtain 50 ohms along the length of the cell's transmission line. Typical voltage standing wave ratios (VSWR) as measured at the cell's ports should be less than 1.2/1.0 for single TEM mode propagation through the cell.

The electric field distribution inside an empty TEM cell operating in a TEM mode can be determined from Jacobian elliptical functions [6]. Numerical results for the electric field in two typical symmetric cells are given in figures 4 and 5 in terms of x- and y- components and for amplitude and polarization angle. These results have been normalized with respect to V/b where V is the voltage across the center plate to the outer wall, and b is as defined earlier. V/b represents the electric field at the center of the test zone ($x=0$, $y=b/2$) and can be obtained in volts from the expression:

$$V = (P_n Z_0)^{1/2} \qquad (2)$$

where P_n is the net power in watts flowing through the cell, and Z_0 is the characteristic impedance in ohms given in (1).

The application of TEM cells for EMC testing has some obvious limitations. The most significant is the restriction on the upper useful frequency caused by multimodes and resonances [6,7,8] given in figure 6. The volume available for test purposes is inversely proportional to this upper frequency limit. The frequencies, f_{cmn}, given for the first few modes can be determined using figure 7. The first order resonant frequencies, $fres_{mn}$, associated with these modes can be found from the following expression:

$$fres_{mn} = \sqrt{(fc_{mn})^2 + (\frac{c}{2\ell})^2} \tag{3}$$

where c is the wave propagation velocity (3.0×10^8 m/sec), ℓ is the resonant length of the cell and m and n are integers corresponding to the particular waveguide mode. Note that the resonant length of the cell is dependent upon the particular mode and the cross sectional geometry of the cell. For the TE_{01} mode in a square cross section cell, it corresponds approximately with the total length of the cell. For the TE_{10} mode, however, it is only slightly longer than the cell's main body length, L, for the first resonance. It is also important to note (i) that the influence of the first order TE modes does not become significant until approaching their resonant frequencies; and (ii) since the septum (center conductor) of the cell is centered symmetrically, the odd order TE modes are not excited in an empty cell. Placing an EUT inside the cell, however will excite these modes. Thus, the recommended upper frequencies exceed the multimode cutoff frequency of the first higher-order mode (TE_{01}) but are less than this mode's resonant frequency.

Efforts have been made to extend the use of TEM cells to frequencies above cutoff [8,9]. The use of rf absorber helps to lower the Q of the cell and suppresses resonance effects associated with multimodes. However it also has some effect on the fundamental TEM mode. Thus, care must be exercised when considering the placement of absorbing materials inside a TEM cell.

Another limitation closely related to the upper useful frequency restriction is the need to keep the size of the EUT small relative to the test volume. A reasonable criteria that has been established is to limit the EUT size to less than L/3 x 2w/3 x b/3. These dimensions are considered a maximum to prevent excessive impedance loading and test-field perturbation when inserting the EUT into the cell. EUT's that exceed the 1/3 linear dimension criterion can be tested in the cell, bearing in mind that excessive loading reduces accuracy in determining the test field. Placing the EUT in the cell tends to short out the test field in the region between the plates, increasing the vertically polarized test field. The error can be partially corrected by measuring the field in the region above and below the EUT with miniature E-field probes and making appropriate corrections or by using a technique outlined in Appendix A of NBS Tech Note 1013 [4].

3. Performing Radiated Susceptibility Tests

As alluded to in the introduction, the TEM cell was developed as an alternative to the conventional shielded enclosure for EMC/EMI testing of electronic equipment. The main purpose of radiated susceptibility testing is to determine if and how EM energy is coupled into the EUT to cause possible degradation to the equipment's performance. Thus a criterion for what constitutes degradation (susceptibility) of the EUT and how this is translated into measurable parameters is normally established first by the user.

The following steps are suggested as a systematic approach for making the radiated susceptibility evaluation [10].

Step 1. Place the EUT inside the cell.

The first step is to place the EUT in a TEM cell, centered in the lower half space below the septum. The first position (position A) as shown in figure 8a is near the floor but insulated from the floor with approximately 2 cm of foam dielectric. Plastic foams with dielectric constants of 1.04 and 1.08 are readily available, are almost invisible electrically, and make good supporting material. If grounding of the equipment case is desired, the EUT would then be placed on the floor. This position (position A) is used to minimize exposure of the EUT's input/output leads to the test field as explained in step 2. Another common EUT position (position B) for testing, as shown in figure 8b, is midway between the septum and the floor. Again, the EUT is supported on a low dielectric foam material. This position increases the exposure of leads to the test field because an increased portion of the lead is oriented polarization matched with the cell's vertically polarized E-field. A comparison of the test results to be taken later for both positions A and B should give some indication of how energy is coupled into the EUT. After placing the EUT in positions A and

B, the EUT may be reoriented as desired, relative to the cell's field polarization. Typically, the first orientation is with the EUT lying flat as in normal use. Care must be taken to record the placement location and how this is done so that it can be repeated if necessary. It may be helpful to mark the bottom of the cell with a uniform array of scribe marks to assist in determining placement locations precisely.

Step 2. Access the EUT as required for operation and performance monitoring.

The EUT input/output and ac power cables should approximate those anticipated for use. Cables should be the same length if possible, be terminated into their equivalent operational impedances so as to simulate the EUT in its operational configuration, and be carefully routed inside the cell to minimize field perturbation. Dielectric guides or holders may be installed in the cell to assure repeatability of the placement location of the cables. These may be placed on the floor to allow the cables to be covered with conductive tape (minimum exposure) and/or on dielectric standoffs to provide coupling of the test field to the leads. If required, any excess portion of the EUT's leads (wiring harness) may be carefully coiled and covered with conductive tape on the floor of the cell. When the leads are bundled together, it may be helpful to twist the input/output monitor leads as separate conductor pairs or use shielded cables to minimize cross coupling between them. It may also be necessary to space the windings in the coil to avoid introducing resonances associated with the coil inductance and distributed capacitance. If braided rf shielding is used, it should be placed in electrical contact with the cell floor, and not in contact with the case of the EUT unless a common ground between the EUT and cell is required. Grounding the two together will influence the results of the susceptibility measurements. The input and output leads, after being connected to the appropriate feedthroughs for accessing and operating the EUT, should also be filtered to prevent rf leakage from the cell, otherwise the shielding integrity of the measurement system will suffer. Care must be exercised in selecting these filters so they do not significantly affect the measured results. The monitor leads used for sensing and telemetering the performance of the EUT may require special high-resistance lines made of carbon-impregnated plastic or fiber optic lines to prevent perturbation of, or interaction with, the test environment. Dc signals or signals with frequency components below 1 kHz may be monitored via the high-resistance lines. Radio frequency signals should be monitored via fiber optic lines.

If the monitor signal is at a frequency or frequencies sufficiently different from the susceptibility test frequency or frequencies, metallic leads may be used with appropriate filtering (high-pass, low-pass, band-pass, etc.) at the bulkhead. Such leads, however, will cause some perturbation of the test field; thus, their placement location must be carefully defined for future reference. Note that a separate, shielded filter compartment should be provided on the outside of the cell for housing the filters, as shown in figures 8a and 8b.

Step 3. Connect the measurement system as shown in figures 9 and 10.

Figures 9 and 10 show the block diagrams of systems using the TEM cell for susceptibility measurements. These figures are used for frequencies from approximately 10 MHz to the recommended upper frequency for the particular cell used. At frequencies below 10 MHz, the dual directional coupler and power meters are replaced by a voltage monitor tee and rf voltmeters. Figure 9 is a diagram of essentially a discrete (manually operated) system or can be used for swept frequency testing. Figure 10 is is a diagram of a system for automated testing under computer control which allows the test field level in the cell to be carefully controlled and progressively increased over selected frequency ranges and intervals while monitoring the EUT performance. If degradation occurs as determined from a pre-established threshold limit and as evidenced by the EUT monitors, the computer can respond interactively with the EUT, thus limiting the test field level and preventing damage to the EUT. The computer can also be used to store the raw data, process the data incorporating correction factors as needed and output the results to printers or plotters according to the software instructions and format.

Step 4. Initialize the measurement system.

This includes zeroing the appropriate instrumentation and measuring the residual offset values of the EUT monitors with the rf source turned off and the EUT turned on in the desired operation mode. These values are then recorded for future reference.

Step 5. Establish the test field and determine the EUT's response.

After initialization of the measurement system, the rf source is then turned on at the desired test frequency, modulation rate, test wave form, etc., and its output level is increased gradually until the maximum required test level is reached or the EUT response monitors indicate vulnerability. Care must be exercised to ensure that sufficient time is spent at each frequency and field level to allow the EUT to respond. The EUT's susceptibility profile is then determined for each position (A or B as shown in figure 8a and 8b) and orientation. It may be necessary to test all three orthogonal orientations of the EUT inside the cell. This is required if all surfaces of the EUT to be tested are to be polarization matched to the TEM field of the cell.

If the test frequency is below 10 MHz, the electric field level in V/m generated inside the cell is determined by the rf voltmeter reading, V_{rf} in volts, in accordance with V_{rf}/b, where b is the separation in meters between the septum and the floor. When the test frequency is 10 MHz or above, where the electric length of the cell is significant, the electric field level is determined by V/b, where V is given in (2) and the net power may be determined by

$$P_n = C_f P_i - C_r P_r \quad \text{(watts)} \quad , \tag{4}$$

with C_f and C_r as the respective forward and reverse coupling ratios of a calibrated bi-directional coupler, and P_i and P_r as the indicated incident and reflected coupler sidearm power readings in watts. Note that the absolute level of the test E-field inside the cell is a function of the location of the EUT in the test zone. An appropriate correction can be made based upon the particular cell's cross section from the data given in figures 4 or 5. Note also, as already mentioned earlier, that the size of the EUT relative to the test volume can influence the determination of the amplitude of the test field.

If the objective of the measurement program is simply to reduce the vulnerability of the EUT to EMI without the additional requirement of determining worst-case susceptibility as a function of absolute exposure field level, one EUT orientation with input/output lead configuration may be tested in one particular operational mode under a pre-selected susceptibility test-field waveform and level. Similar tests may then be duplicated at the same test position with the same lead configuration and test-field waveform and level, after the corrective measures such as providing additional shielding, etc. are made to the EUT. These testing results are then compared to determine the degree of improvement.

Sometimes, it is desirable to monitor the field distribution inside the cell using small calibrated electric and/or magnetic probes, while an EUT is in position. If this is the case, one must be careful in interpreting these monitored results, because the results are a combination of the incident TEM field launched inside the cell and the scattered fields from the EUT and its leads in the near-field. The field so monitored can be quite different from the unperturbed test field, leading to potentially erroneous conclusions. Whenever possible, it is preferable to mount the field monitoring probes in the other half space of the cell in the mirror image location of the EUT.

4. Performing Radiated Emission Tests

Electronic or electromechanical equipment or components may emit energy which interferes or interacts with the normal operation of either the system and/or other receptors. To ensure the electromagnetic compatibility (EMC) of such systems, it is important to determine the amplitude levels of these emanations and to characterize their waveform, polarization, etc. This is apparent since equipment performance degradation or failure is often dependent upon the interfering signal waveform and amplitude.

A TEM cell is especially useful for emitted signal waveform characterization because of its characteristics as a TEM transducer which permits detection of the signal with little or no distortion in the signal waveform. TEM cells are reciprocal devices (i.e., can receive or detect radiated fields from equipment as well as establishing fields for testing). Thus, energy radiated from an EUT placed inside the cell will couple via the TEM mode to the cell's ports.

Two procedures have been developed for performing radiated emissions measurements using a TEM cell. The first only provides information for determining the equivalent, free-space radiated electric field strength for a single orientation of the EUT inside the TEM cell. The details for performing these measurements are contained in NBS Technical Note 1013 [4]. The second procedure is more complicated but yields (assuming necessary conditions are met) significantly more complete detailed results. For example, detailed radiation patterns and total power radiated by the EUT in free space can be computed. Complete details for performing these measurements are contained in NBS Technical Note 1059 [11].

Steps 1 and 2. Place the EUT in the cell and access it as required for operation.

The first procedure outlined in NBS Technical Note 1013 consists of placing the EUT inside the TEM cell in the desired orientation and test configurations. The procedures outlined in steps 1 and 2 for susceptibility measurements apply and should be followed as steps 1 and 2 for emissions testing.

Step 3. Connect the measurement system as shown in the figures 11 or 12.

Energy emitted from the EUT is coupled via the TEM mode of the cell to a spectrum analyzer, receiver or oscilloscope connected to one port of the cell. The other cell port is terminated in a 50 ohm load.

Step 4. Turn the EUT on and measure and record the detected emissions.

The voltage measured at the port of the cell can then be used to determine the equivalent free-space radiated field as follows: The equivalent free-space radiated field, E_R, at a distance, d, from the EUT is then given approximately as:

$$E_R \approx \frac{6.2\ bV_R}{d\lambda \tilde{E}\ \cos\theta}\sqrt{G} \quad \text{(volts/meter)} \qquad (5)$$

where b is as defined earlier, V_R is the RMS voltage amplitude in volts measured at one end of the cell, λ is the wavelength of the radiated signal in meters, $\tilde{E}$ is the normalized electric field inside the cell relative to the field strength at the center of the cell test region, $\cos\theta$ corresponds to the polarization match between the radiated field from the EUT and the TEM field of the cell, and G is the gain of the EUT as a radiator. Limitations on the upper frequency range and size of the EUT are similar to those that apply for susceptibility testing, (i.e., the EUT is assumed to be electrically small).

If the time domain analysis or emanation characteristics are required, use the block diagram shown in figure 12 with the oscilloscope either connected directly to the cell measurement port, or with the oscilloscope connected to the predetection or postdetection outputs of the receiving instrument. The second arrangement using a receiver with the oscilloscope provides greater measurement sensitivity. In either arrangement the oscilloscope must be synchronized with either the periodic detected signal from the cell or with an EUT-monitored periodic signal represented by the dashed line from the EUT. The measurement results can then be recorded by photographing the oscilloscope display. If the emanation is random, the oscilloscope cannot be synchronized properly and the detected signal must be either: 1) recorded with a video disk or tape recorder and played back frame by frame to analyze the emanation; or 2) analyzed statistically using amplitude probability distribution analyzers, etc.

Similar tests can be made for different orientations of the EUT in the cell, for different arrangements of the EUT's leads, and/or for different EUT operating modes as required to evaluate various test conditions.

The second procedure outlined in NBS Technical Note 1059 can be explained briefly as follows: Since the EUT must of necessity be electrically small, (to avoid multimoding and excessive loading of the cell), it can be modeled as equivalent short electric and magnetic dipole sources. These dipole sources may then be combined vectorially to form a composite equivalent source consisting of three orthogonal electric and three orthogonal magnetic dipoles as shown in figure 13.

When an unknown source object (EUT) is placed at the center of a TEM cell, its emission couples into the fundamental TEM mode and propagates toward the two ports of the cell. With a hybrid junction inserted into a loop connecting the cell output as shown in figure 14, it is possible to measure the sum and difference of the powers and the relative phase between the sum and difference outputs. This way of measuring the relative phase is very advantageous because it avoids the complication of having to establish an absolute phase reference physically connected to the EUT. Systematic measurements of the powers and relative phases at six different EUT orientations are sufficient to determine the amplitude and phases of the unknown equivalent component dipole moments as depicted in figure 13. From these data the corresponding detailed radiation pattern and total power radiated by the unknown source in free space can then be computed.

Details for specifying the six different EUT orientations and the analytical expressions used to calculate the above values of power and phase are contained in NBS Technical Note 1059.

5. Examples of Some TEM Cell Applications

A number of applications have been identified for using TEM cells. Initial work involved the development of a cell for use to establish high level fields for biological effects research [12]. Since then a number of cells have been designed and are in current use as exposure chambers. Early in the development of the TEM cell, it was realized that its broadband frequency response characteristics made it a prime candidate for use in TEMPEST testing of EUT [13] and as an EM pulse (EMP) simulator [14]. The largest known TEM cell in existence (3m x 20m x 24m) is located at Sandia Labs, in Albuquerque, New Mexico, and it is used as a dual purpose facility for both EMP testing and CW susceptibility testing of whole DoD weapon systems. The National Bureau of Standards (NBS) recognized the TEM cell's potential for use as a calibration standard for establishing standard TEM fields for use, in calibrating electrically small probes and rf hazard meters. Considerable development and analysis work was done at NBS to carefully evaluate various TEM cells for this use [2,15]. TEM cells are also used extensively for radiated susceptibility testing of components by the Automotive industry. The Society of Aumotive Engineers (SAE) has adopted the use of TEM cells for evaluating the susceptibility of automotive vehicle components in the frequency range 14 kHz to 200 MHz [16]. In addition, one motor vehicle manufacturer has constructed a very large cell with a test region 2m x 5m x 7m for whole vehicle testing [17]. Another large cell, 2.8m x 2.8m x 5.6m, is in use at AT&T Information Systems for measuring both susceptibility and emissions of communication equipment [18,19]. Considerable work was done to evaluate and compare the use of this cell to similar measurements performed on a 3-meter open-field site and in an anechoic chamber. More recently, 2m x 2m x 4m TEM cells have been evaluated and proposed for use by the Electronics Industries Association (EIA) to measure TV/VCR immunity to EMI [20,21]. Finally, a recent application is the use of a pair of cells, one on top of the other, with a common aperture cut between them, to evaluate the shielding effectiveness of materials [22].

6. References

[1] W.R. Free and C.W. Stuckey, "Electromagnetic interference methodology communication equipment," Final Report ECOM-0189-F, October 1969.

[2] M.L. Crawford, "Generation of standard EM fields using TEM transmission cells," IEEE Trans. on EMC, Vol. EMC-16, No. 4, November 1974.

[3] W.F. Decker, W.A. Wilson, and M.L. Crawford, "Construction of a large transverse electromagnetic cell," NBS Tech. Note 1011, February 1979.

[4] M.L. Crawford and J.L. Workman, "Using a TEM cell for EMC measurements of electronic equipment," NBS Tech. Note 1013, Revised July 1981.

[5] J.C. Tippet and D.C. Chang, "Radiation characteristics of dipole sources located inside a rectangular, coaxial transmission line," NBSIR 75-829, January 1976.

[6] J.C. Tippet, "Modal characteristics of rectangular coaxial transmission line. Thesis submitted June 1978 for degree of Doctor of Philosophy to University of Colorado, Electrical Engineering Dept., Boulder, Colorado.

[7] C.M. Weil, W.T. Joines, and J.B. Kinn, "Frequency range of large-scale TEM mode rectangular strip lines," Microwave Journal, November 1981, pp. 93-113.

[8] D.A. Hill, "Bandwidth limitations of TEM cells due to resonances," J. Microwave Power, Vol. 18, No. 2, pp. 181-195, 1983.

[9] M.L. Crawford, J.L. Workman, and C.L. Thomas, "Expanding the bandwidth of TEM cells for EMC measurements," IEEE Trans. on EMC, Vol. EMC-20, No. 3, pp. 368-375, August 1978.

[10] M.L. Crawford, "Improving the repeatability of EM susceptibility measurements of electronic components when using TEM cells," SAE Technical Paper Series 830607, SAE International Congress and Exposition, Detroit, Michigan, Feb. 28 - Mar. 4, 1983.

[11] M.T. Ma and G.H. Koepke, "A method to quantify the radiation characteristics of an unknown interference source," NBS Tech Note 1059, October 1982.

[12] G.A. Skaggs, "High frequency exposure chamber for radiobiological research," NLR Memo, Rep. 2218, Feb. 1971.

[13] M.L. Crawford, "Measurement of electromagnetic radiation from electronic equipment using TEM transmission cells," NBSIR 73-306, Feb. 1973.

[14] N. Pollard, "A broadband electromagnetic environments simulator (EMES)," IEEE 1977 Inter. Symp. on EMC., Seattle, Washington, August 2-4, 1977.

[15] M.L. Crawford, "Generation of standard EM fields for calibration of power density meters, 20 kHz to 1000 MHz. NBSIR 75-804, Jan. 1975, 40 pages.

[16] Society of Automotive Engineers (SAE) Recommended Practice, Electromagnetic susceptibility procedures for vehicle components (except aircraft) -SAE J1113 June 84.

[17] G.F.E. Vrooman, An indoor 60 Hz to 40 GHz facility for total vehicle EMC testing," SAE technical paper series No. 831001 SAE passenger car meeting, Dearborn, MI., June 6-9, 1983.

[18] D.N. Heirman," Automated Immunity Measurements," 6th EMC Sym p & Tech Expo. Zurich, Mar 5-7, 1985.

[19] M.L. Crawford, "Comparison of open-field, anechoic chamber and TEM cell facilities/techniques for performing electromagnetic radiated emission measurements," IEEE 1983 Inter. Symp. on EMC, Arlington, VA; Aug. 23-15, 1983.

[20] G.C. Hermeling, "Method for immunity measurements on home entertainment productes using the 2 meter TEM cell," IEEE 1983 Inter. Symp. on EMC. Arlington, VA. Aug. 23-25, 1983.

[21] EIA Interim Standard No. 32 (proposed), "Immunity of television receivers and video cassette recorders (VCR's) to direct radiation from radio transmissions, 0.5 to 30 MHz," revised Sept. 6, 1984.

[22] A.N. Faught, Jr., J.T. Dowell, and R.D. Scheps, "Shielding material insertion loss measurements using a dual TEM cell system," IEEE 1983 Inter. Symp. on EMC, Arlington, VA, Aug. 23-25, 1983.

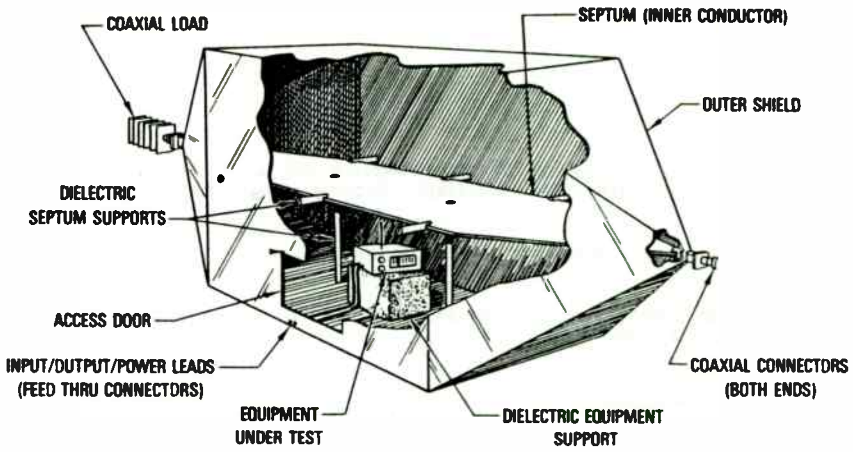


Figure 1. Cut-away view of TEM cell showing placement of EUT.

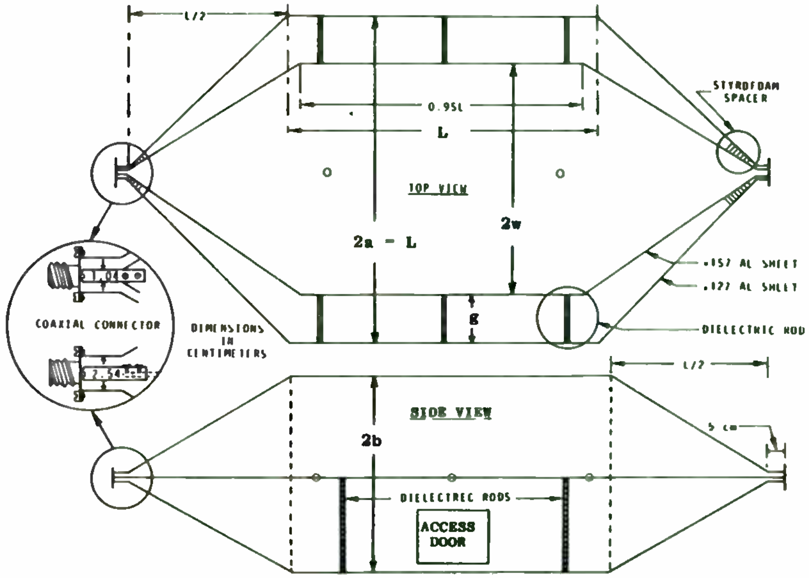


Figure 2. Design for rectangular TEM transmission cell.

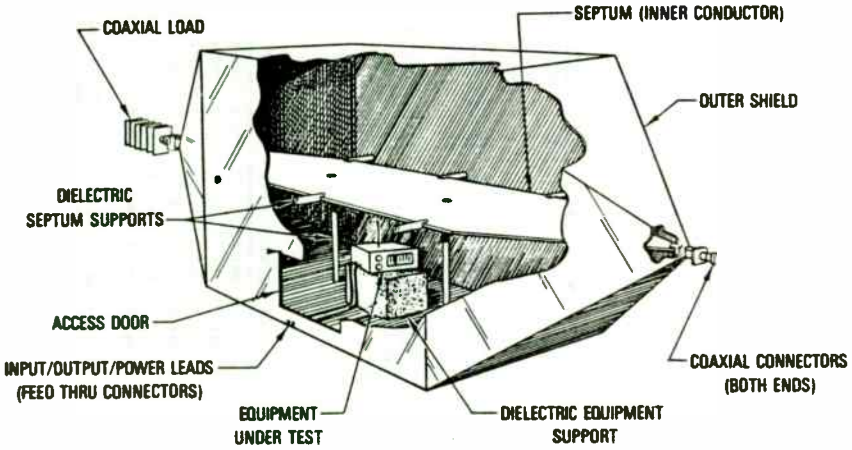


Figure 1. Cut-away view of TEM cell showing placement of EUT.

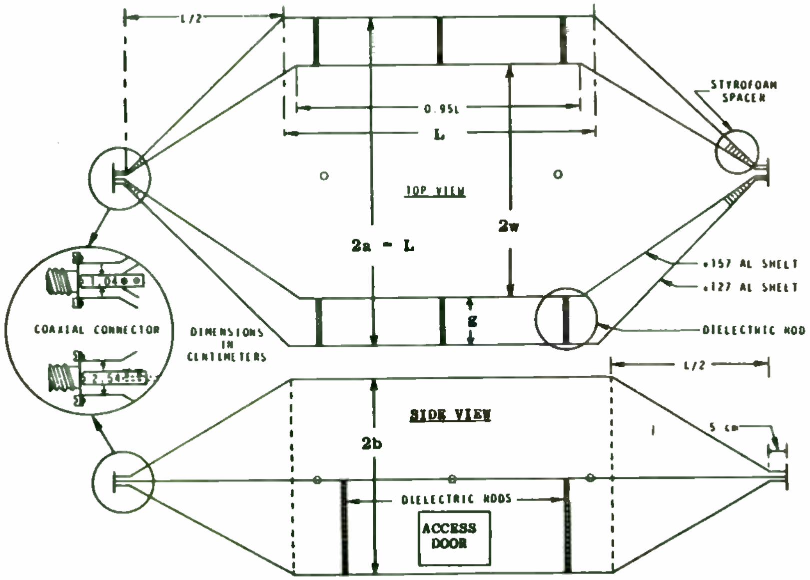
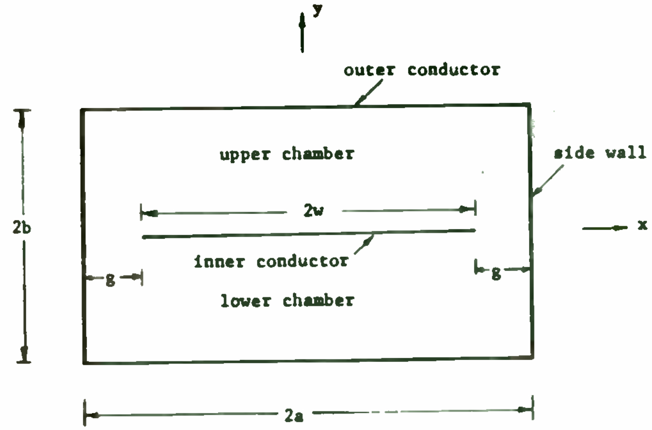


Figure 2. Design for rectangular TEM transmission cell.



Cross section of a TEM cell.

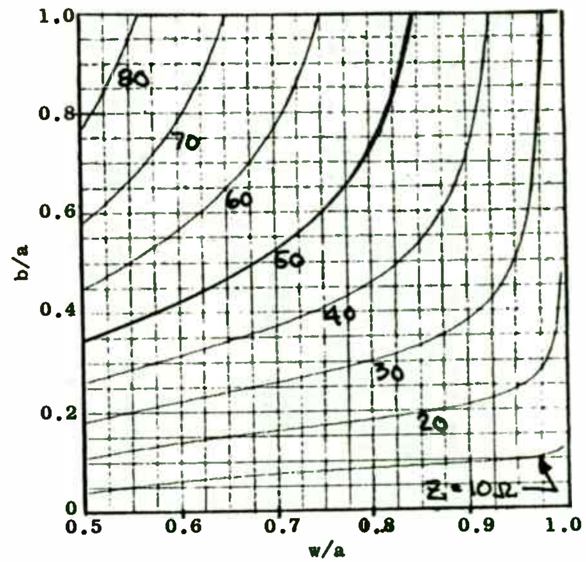


Figure 3. TEM cell design curves showing ratio of cell outer conductor and center conductor dimensions for given characteristic impedances.

Table 1.

x-component of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w=0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.000	0.000	0.000	0.000	0.000	0.000
b/5	0.000	0.060	0.129	0.208	0.278	0.307
2b/5	0.000	0.108	0.245	0.422	0.600	0.680
3b/5	0.000	0.127	0.311	0.620	1.029	1.237
4b/5	0.000	0.090	0.248	0.647	1.684	2.285
Septum	0.000	0.000	0.000	0.000	0.000	3.603

center of TEM cell

Table 3.

Magnitude of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w = 0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.824	0.793	0.698	0.530	0.289	0.000
b/5	0.853	0.827	0.747	0.605	0.420	0.307
2b/5	0.935	0.924	0.886	0.817	0.727	0.680
3b/5	1.049	1.060	1.096	1.157	1.218	1.237
4b/5	1.153	1.189	1.321	1.633	2.154	2.285
Septum	1.196	1.245	1.431	1.986	6.640	3.603

center of TEM cell

Table 2.

y-component of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w=0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.824	0.793	0.698	0.530	0.289	0.000
b/5	0.853	0.825	0.736	0.568	0.315	0.000
2b/5	0.935	0.917	0.852	0.699	0.410	0.000
3b/5	1.049	1.052	1.051	0.977	0.652	0.000
4b/5	1.153	1.186	1.298	1.499	1.343	0.000
Septum	1.196	1.245	1.431	1.986	6.640	0.000

center of TEM cell

Table 4.

Polarization angle of the electric field in degrees in a square, symmetric TEM cell.

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	90.00	90.00	90.00	90.00	90.00	--
b/5	90.00	85.86	80.05	69.89	48.54	00.00
2b/5	90.00	83.27	73.97	58.89	34.35	00.00
3b/5	90.00	83.14	73.50	57.60	32.36	00.00
4b/5	90.00	85.64	79.20	66.67	38.56	00.00
Septum	90.00	90.00	90.00	90.00	90.00	00.00

center of TEM cell

Figure 4. Distribution of the components of the normalized electric-field inside a square, symetrical TEM cell, b/a = 1, w/a = 0.83.

Table 1.

x-component of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w=0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.000	0.000	0.000	0.000	0.000	0.000
b/5	0.000	0.060	0.129	0.208	0.278	0.307
2b/5	0.000	0.108	0.245	0.422	0.600	0.680
3b/5	0.000	0.127	0.311	0.620	1.029	1.237
4b/5	0.000	0.090	0.248	0.647	1.684	2.285
Septum	0.000	0.000	0.000	0.000	0.000	3.603

center of TEM cell

Table 3.

Magnitude of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w = 0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.824	0.793	0.698	0.530	0.289	0.000
b/5	0.853	0.827	0.747	0.605	0.420	0.307
2b/5	0.935	0.924	0.886	0.817	0.727	0.680
3b/5	1.049	1.060	1.096	1.157	1.218	1.237
4b/5	1.153	1.189	1.321	1.633	2.154	2.285
Septum	1.196	1.245	1.431	1.986	6.640	3.603

center of TEM cell

Table 2.

y-component of the electric field in a square symmetric TEM cell. (fig. 9, a=b, w=0.83a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.824	0.793	0.698	0.530	0.289	0.000
b/5	0.853	0.825	0.736	0.568	0.315	0.000
2b/5	0.935	0.917	0.852	0.699	0.410	0.000
3b/5	1.049	1.052	1.051	0.977	0.652	0.000
4b/5	1.153	1.186	1.298	1.499	1.343	0.000
Septum	1.196	1.245	1.431	1.986	6.640	0.000

center of TEM cell

Table 4.

Polarization angle of the electric field in degrees in a square, symmetric TEM cell.

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	90.00	90.00	90.00	90.00	90.00	--
b/5	90.00	85.86	80.05	69.89	48.54	00.00
2b/5	90.00	83.27	73.97	58.89	34.35	00.00
3b/5	90.00	83.14	73.50	57.60	32.36	00.00
4b/5	90.00	85.64	79.20	66.67	38.56	00.00
Septum	90.00	90.00	90.00	90.00	90.00	00.00

center of TEM cell

Figure 4. Distribution of the components of the normalized electric-field inside a square, symetrical TEM cell, b/a = 1, w/a = 0.83.

Table 5.

x-component of the electric field in a rectangular symmetric TEM cell. (fig. 10, 0.6a = b, w = 0.72a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.000	0.000	0.000	0.000	0.000	0.000
b/5	0.000	0.024	0.067	0.143	0.220	0.249
2b/5	0.000	0.040	0.121	0.284	0.462	0.517
3b/5	0.000	0.043	0.141	0.410	0.763	0.817
4b/5	0.000	0.028	0.101	0.440	1.247	1.112
Septum	0.000	0.000	0.000	0.000	1.969	1.254

center of TEM cell

Table 7.

Magnitude of the electric field in a rectangular symmetric TEM cell. (fig. 10, 0.6a = b, w = 0.72a

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.966	0.946	0.872	0.698	0.394	0.000
b/5	0.972	0.956	0.892	0.738	0.466	0.249
2b/5	0.989	0.982	0.951	0.856	0.655	0.517
3b/5	1.010	1.016	1.038	1.062	0.945	0.817
4b/5	1.028	1.046	1.125	1.383	1.404	1.112
Septum	1.035	1.058	1.164	1.664	1.969	1.254

center of TEM cell

Table 6.

y-component of the electric field in a rectangular symmetric TEM cell. (fig. 10, 0.6a = b, w = 0.72a)

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	0.966	0.946	0.872	0.698	0.394	0.000
b/5	0.972	0.956	0.890	0.724	0.411	0.000
2b/5	0.989	0.981	0.944	0.807	0.464	0.000
3b/5	1.010	1.015	1.028	0.979	0.557	0.000
4b/5	1.028	1.046	1.120	1.311	0.645	0.000
Septum	1.035	1.058	1.164	1.664	0.000	0.000

center of TEM cell

Table 8.

Polarization angle of the electric field in degrees in a rectangular, symmetric TEM cell.

	center	a/5	2a/5	3a/5	4a/5	wall
Floor	90.00	90.00	90.00	90.00	90.00	..
b/5	90.00	88.58	85.68	78.80	61.76	00.00
2b/5	90.00	87.66	82.70	70.61	45.12	00.00
3b/5	90.00	87.60	82.20	67.26	36.15	00.00
4b/5	90.00	88.48	84.82	71.43	27.36	00.00
Septum	90.00	90.00	90.00	90.00	90.00	00.00

center of TEM cell

Figure 5. Distribution of the components of the normalized electric-field inside a rectangular, symetrical TEM cell, b/a = 0.6, w/a = 0.72.

Cell #	Recommended upper frequency (MHz)	Cell form factor a/b	Plate separation b (cm)	Center septum w (cm)	TE01 cutoff/multimode frequency Propagation (MHz)	TE01 cutoff/multimode frequency Resonance (MHz)
1	40	1.0	150	124.5	29	44
2	60	1.0	100	83	43	66
3	80	1.5	60	68	66	86
4	100	1.0	60	50	72	110
5	120	1.5	40	45.6	100	129
6	150	1.67	30	36.0	128	162
7	160	1.5	30	34.2	134	172
8	200	1.0	30	24.9	143	220
9	240	1.5	20	22.8	204	260
10	300	1.67	15	18.0	256	312
11	350	1.5	14	16.0	272	371
12	400	1.0	15	12.5	287	440
13	500	1.5	10	11.4	408	520
14	600	1.5	8	9.1	510	650
15	1000	1.0	6	5.0	718	1100

Figure 6. Recommended upper frequencies for various TEM cell sizes and cross-sections.

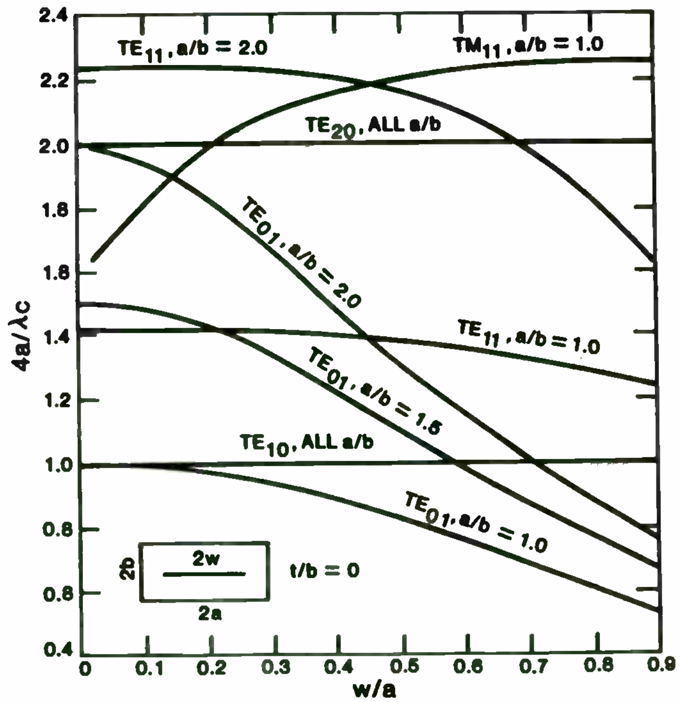


Figure 7. Normalized cut-off frequency versus w/a for first order modes in rectangular strip line with b/a = 1.0, 0.67, and 0.5.

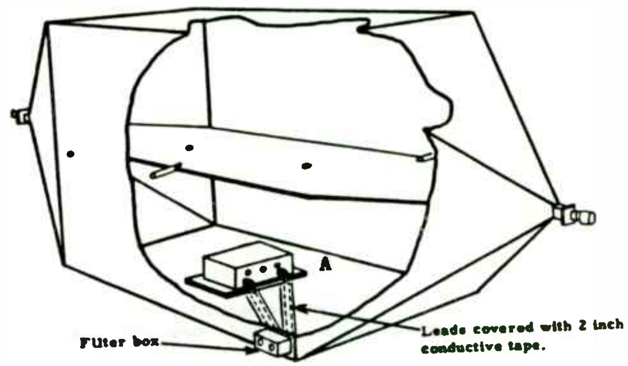


Figure 8a. EUT near floor of TEM cell for minimum exposure of leads to test field.

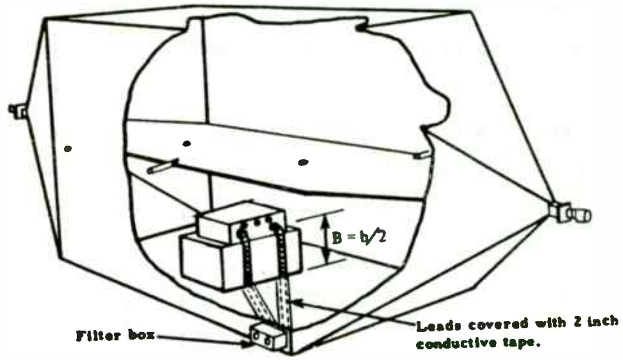


Figure 8b. EUT centered in test zone midway between septum and floor.

Figure 8. Placement of EUT inside TEM cell showing routing of leads.

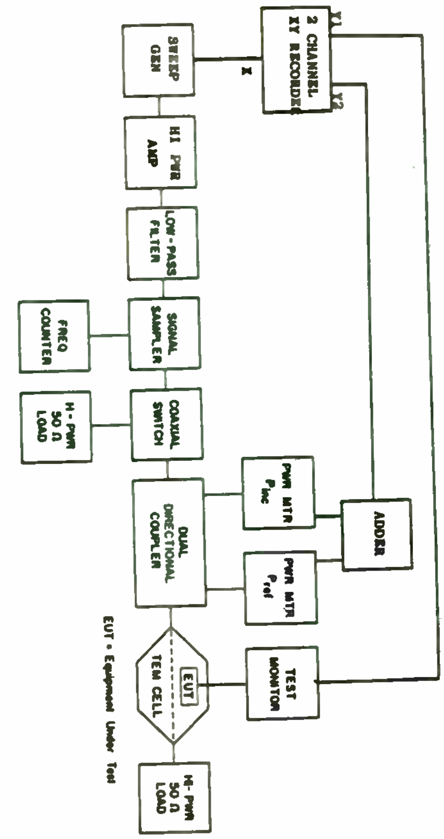


Figure 9. Block diagram of system for susceptibility testing of equipment at frequencies ≳ 10 MHz.

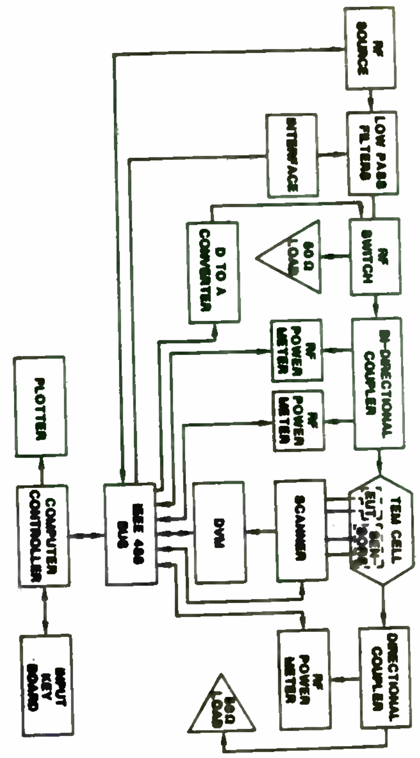


Figure 10. Block diagram of automated TEM cell susceptibility measurement system.

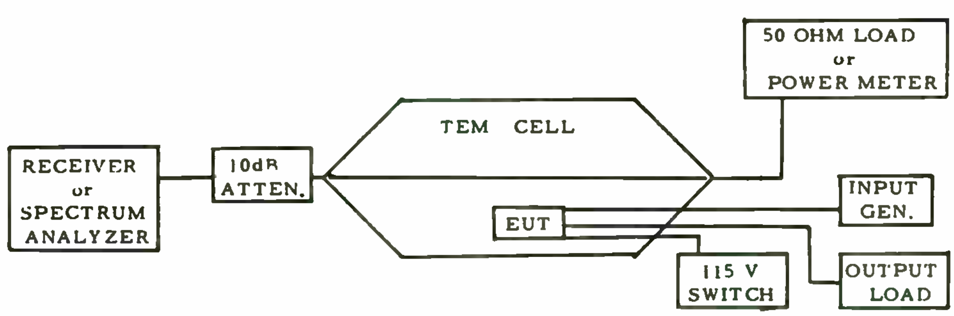


Figure 11. Block diagram of measurement system for frequency domain analysis of radiated emissions from EUT using a TEM cell.

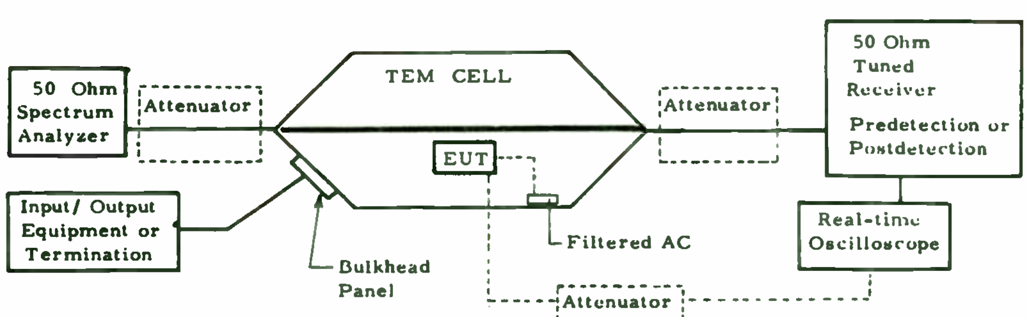


Figure 12. Block diagram of measurement system for time domain analysis of radiated emissions from EUT using a TEM cell. (Attenuator required for hard line sync of some EUT.)

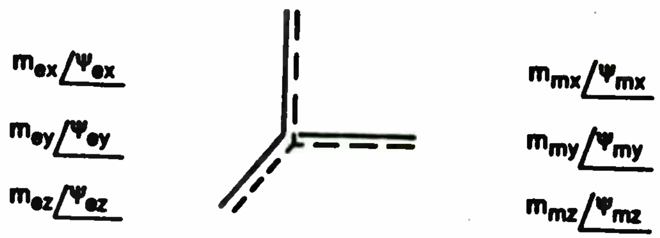


Figure 13. Unknown electrically small source representation in terms of three orthogonal electric and magnetic dipoles.

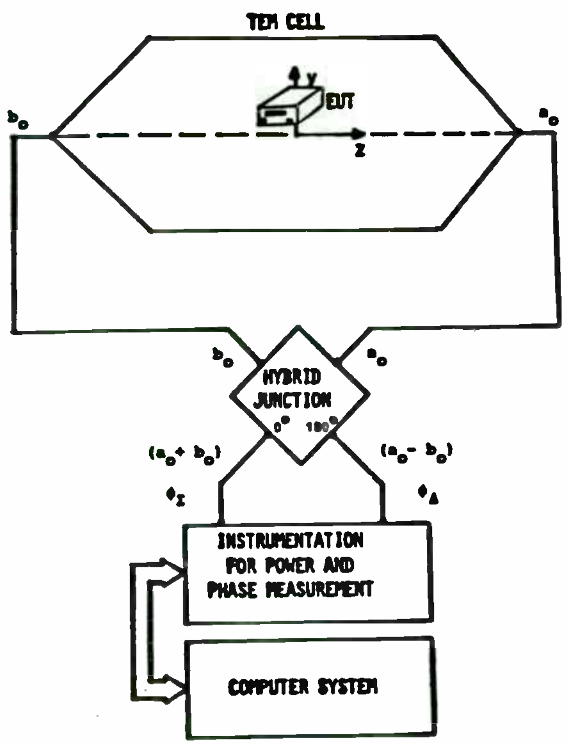


Figure 14. Block diagram of radiated emission measurement system using TEM cell for complete pattern and total radiated power determination.

AMATEUR SATELLITE COMMUNICATION LINKS

Presented by

Marlin E. Greer

Electrical Program Coordinator

Murray State University

Co-authors

William Call

Associate Professor

Murray State University

Ken McGary

Engineer

General Dynamics

AMATEUR SATELLITE COMMUNICATION LINKS

Amateur radio operators through the AMSAT organization have placed a series of satellites in orbit. These orbital vehicles have increased in capability and sophistication through the Phase I, Phase II and Phase III progressions. The AMSAT-OSCAR 10 (AO-10), formerly designated Phase IIIB, is placed in a highly elliptical, inclined orbit. The orbit was intended to have an eccentricity of 0.685 associated with an apogee of 36000 km above the northern hemisphere and a perigee of 1500 km and to have an inclination of 57 degrees. This Molniya style orbit allows a single satellite to service the entire globe and permits northern hemisphere "hams", who comprise 90 percent of the potential uaers, to make contacts and carry out propagation experiments for up to fifteen hours a day on favorable passes. The actual orbital parameters turned out to be a .609 eccentricity, a 35600 km apogee, a 3800 km perigee and a 25.6 degree inclination which caused a minor reduction in the desirability of the availability pattern. The satellite's position can be tracked with a personal computer using updated Keplerian orbital elements.

Phase III uses an inverting linear translator called a transponder that is a high gain, relatively wide bandwidth AGC amplifier and frequency translator. The 70 cm/2m transponder, designated mode B, has a 150 kHz bandwidth and operates over an uplink range of 435.025-435.175 MHz and a down link range of 145.975-145.825 MHz. The inversion of the transponder causes a low end 435.04 MHz transmission to become a high end 145.96 MHz reception and upper side band uplinks change to a lower side band downlinks requiring the reinserted carrier to be on the opposite side of the 3 kHz passband at the receiver. Inversion has the advantage that Doppler shifts on the up and down links will partially cancel. Mixing the ground based transmission with a local oscillator (LO) that is above the uplink frequency and filtering to the difference produces the inversion. For example, a 435 MHz low-end incoming signal mixed with a 581 MHz LO produces an upper-end 146 MHz two-meter output. A 435.003 MHz upper side band input is transposed to a 145.997 MHz lower side band at the output.

The antenna interface on the satellite includes a high-gain apogee and an omnidirectional perigee antenna on the transmission link to help minimize signal strength cycles.

The particular satellite ground station being described takes advantage of an existing ten-meter system. Adapting the rig to AO-10, Mode B requirements entails the use of a 10 m/70 cm transmitting converter (transverter) on the uplink. The tranverter samples the output of the ten meter exciter since a low level input is required. The VHF/UHF converter used is a linear translator and provides a three-fourths watt output for a one-half watt input. This -2.9 dBW level SSB signal is increased in two stages. The power amplifier provides 18 dB of power gain, with a resulting signal level of 32 watts. A further increase is obtained by the gain of the antenna which consists of a multi-element Yagi-Uda parasitic array rated at a gain of 12 dBd. This is equivalent to 14.15 dBi. If the coax line loss is assumed to be 2 dB, the net radiated power comes to 27 dBW which is equal to 500 watts.

A minimum of about 60 watts EIRP (effective isotropic radiated power) is needed to establish a dependable uplink. However, the ground based transmitter should not radiate more than 500 watts (+27 dBW) EIRP, otherwise desensing of the satellite transponder can occur through AGC action.

The returning 2m signal from the satellite is also received by a multi-element parasitic array. The signal is then amplified by a low-noise system and translated to 10m by the receiving converter.

On board computer-controlled pulses to an electromagnet that works against the earth's magnetic field is used to maintain a satellite spin of approximately one revolution per second in order to stabilize the craft. The side effect is a periodic pulsation of the transmitted signal amplitude when received by a dipole called spin modulation. The use of linear antennas creates signal fades of an annoying degree. Circularly polarized antennas of the proper sense minimize this effect and allow communication with little disturbance. AO-10 transmits a circularly polarized signal with a clock-wise (right-hand) sense. However, due to the existence of Faraday rotation of signals in the 100 MHz to 1 GHz range it is advisable to provide for either polarization since cross polarization can result in as much as 30 dB attenuation. We plan to do additional study on the effects of Faraday rotation on this type of satellite orbit.

Cross-coupled Yagi antennas were used with remote controlled coaxial relays to permit selection of vertical, horizontal or clockwise or counter-clockwise circular polarization. Two helical antennas, one wound for RHCP, the other for LHCP, could have also been used but these would not have allowed for horizontal or vertical polarization.

In order to impedance match the 20-25 ohm of a dipole with parasitic elements to the 52 ohm feedline, the gamma approach was used. As an unsplit dipole is center-grounded and driven off-center, the impedance increases, but also becomes more inductively reactive, as the tap point is moved out on the element. The positive reactance is cancelled by an adjustable series capacitor placed in the drive line. Since matching is also influenced by gamma rod length and diameter and by gamma-rod-to-driven-element spacing, the actual tapping point and capacitance setting was determined by adjusting for minimum reflection using a directional coupler.

The cross-coupling of these two 52 ohm systems creates two problems: 1) the impedance would be reduced to 26 ohms if directly paralleled and 2) circular polarization requires a 90 degree phase adjustment between the two lines to correspond to the 90 degree geometry of the vertical and horizontal elements.

The impedance of each line is brought to the target level of 100 ohms by using a quarter-wave transformer which is a narrow band technique. The impedance looking into a loaded $\lambda/4$ section is:

$$Z_i = Zo^2/Z_L$$

where Zo is the impedance of the quarter-wave section. The length of a quarter-wave section is determined by the $\lambda = c/f$ relationship where c is the speed of light in the medium. Coax cables have published velocity factors permitting determination of their electromagentic wave propogation velocities. For example, a quarter-wave length for RG8/U is given by $.66\ (3 \times 10^8)/4\ (145.9 \times 10^6)$ at the 2m center frequency of 145.9 MHz. This length is 0.339m. With this matching arrangement there will be no standing waves on the primary feeder line but there is a mismatch between the quarter-wave element and the load and hence the transformer section has significant standing waves. This energy loss can be tolerated because of the short length of the matching section. Using coax with a characteristic impedance of 73 ohms on the quarter-section causes a 52 ohm load to be transformed to 102 ohms.

To meet the 90 degree phasing requirement, the two near 100 ohm antenna lines are connected together by a quarter-wave section of near-100 ohm coax (93 ohm). By driving the interconnected system at either end of the 93 ohm quarter-wave interconnect, either side can have a 90 degree delay inserted allowing both RHCP and LHCP.

Impedance matches were verified by directional coupler techniques. Measured VSWR's were well under 1.2:1. This compares to the worst-case mismatch in the system of 1.1:1 between the 93 ohm quarter-wave section and the 102 ohm impedance looking into the 73 ohm transformer.

VSWR = Zo/ZL or ZL/Zo

Circularity was verified by monitoring calibrated reception with the completed cross-coupled antenna in the presence of a rotating polarization test signed. This field was created by turning a transmitting linear dipole antenna. The completed receiving antenna had accurate circularity to within about 2 dB. When operating the antenna in the linear mode, a loss of about 20 dB is noted when the transmitting and receiving antennas are 90° apart in polarization. We had to install the antenna with a non-metallic support mast to obtain this accuracy. A metallic support mast in-line with one section's elements causes about 10 dB of loss to that sense, resulting in an elliptical rather than circular polarization.

The system is designed such that polarization can be under manual or computer control. Push button manual instructions are debounced and digitized using a monostable strobe to D-type latches. The drives from the manual switches or from software to the VMOS relay controllers are multiplexed by a data selector. The D/A, A/D computer interface card and a clock card allows for polarization studies by automatic recording of timed signal strength records associated with the various polarization modes and the orbital tracking program permits automatic antenna positioning through the interface board.

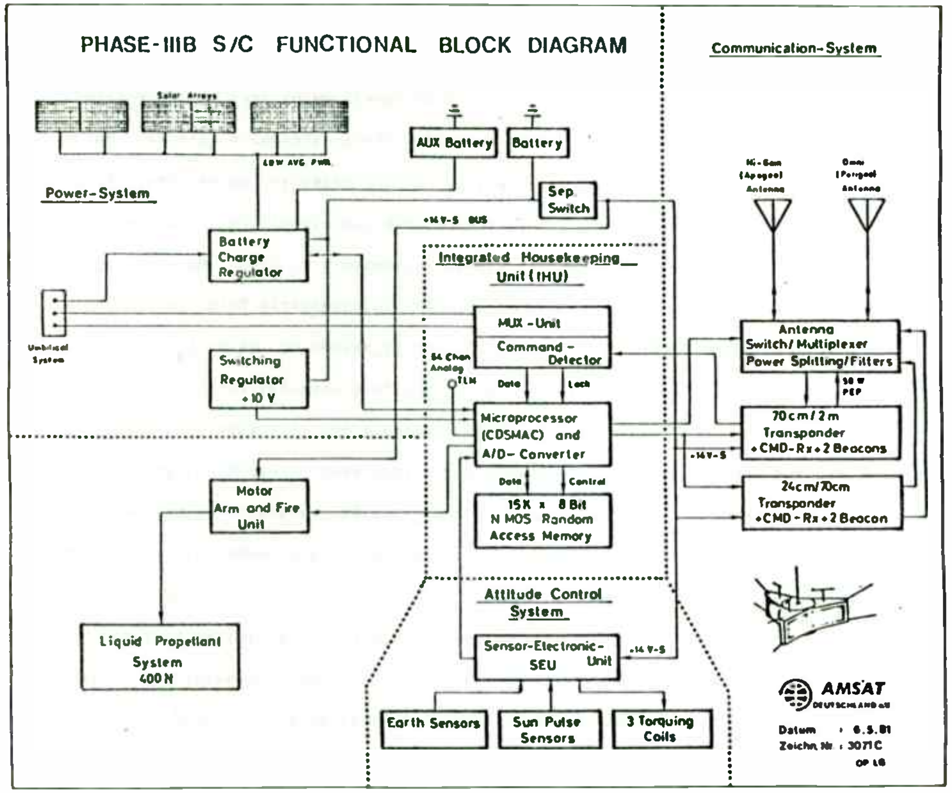
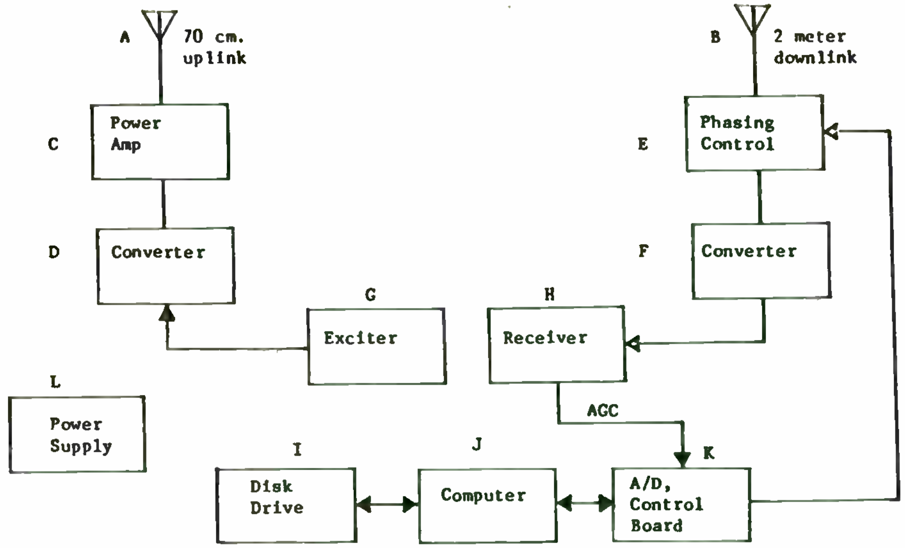


Figure 1: Phase III Functional Diagram



A. KLM 42045018c 70 cm. Antenna

B. KLM 14415016C 2M. Antenna

C. Hamtronics LP 430 70 cm. 30 Watt Amplifier

D. Hamtronics XV4 UHF Transmitting Converter (10m to 70 cm)

E. Antenna Relay and Phasing Harnesses

F. Microwave Modules MMc 144 Receiving Converter (2m to 10m)

G. HF Exciter

H. HF Receiver

I. Disk Drive

J. Microcomputer

K. Analog/Digital Board

L. Power Supply (12v, 12a)

Figure 2: Satellite Ground Station

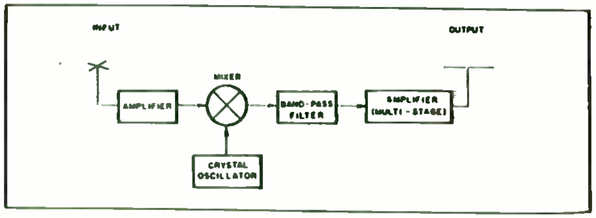


Figure 2: Block Diagram of a Linear Transponder

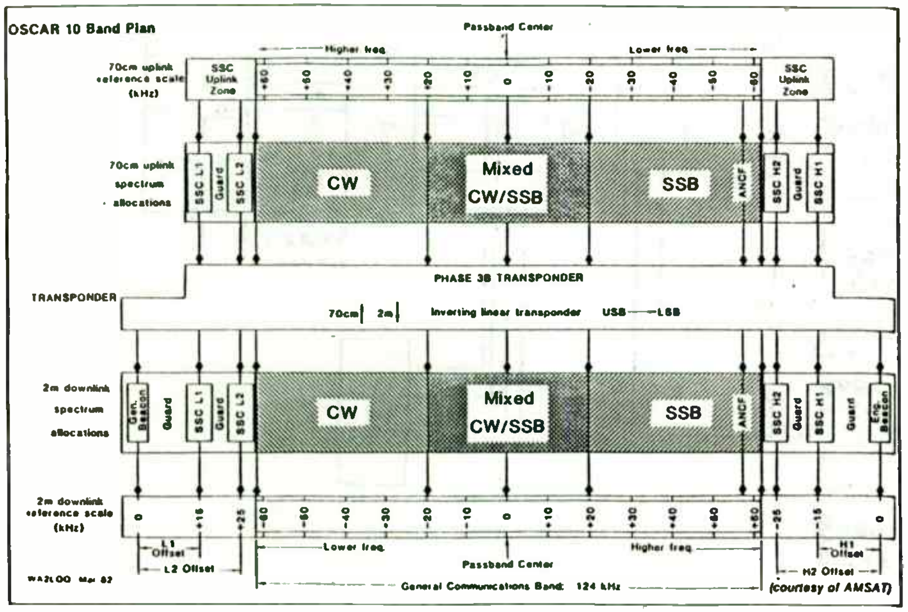
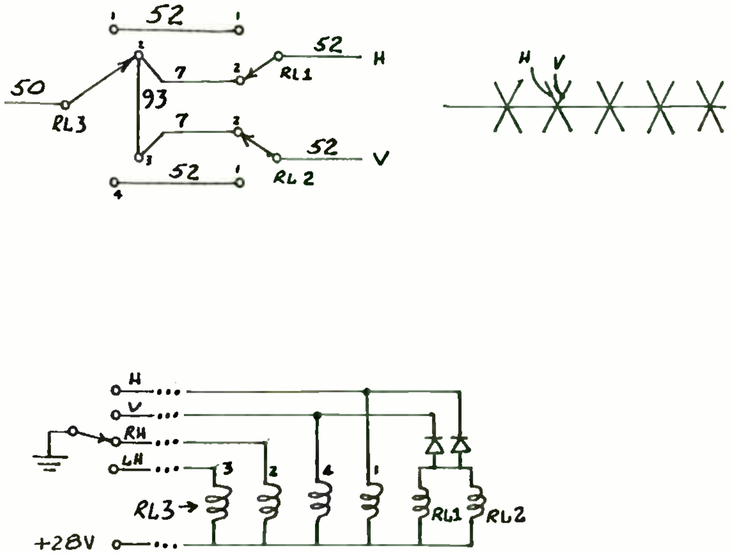


Figure 3: OSCAR Bandplan

William Call



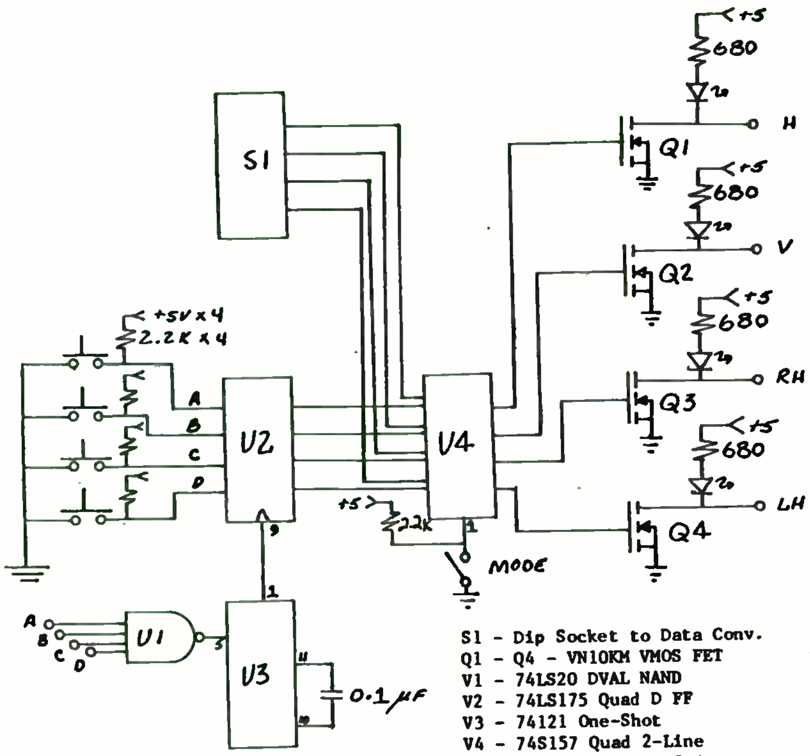
Relay Switching

Mode	RL1	RL2	RL3
HOR.	1	1	1
VERT.	1	1	4
RHCP (CW)	2	2	2
LHCP (CCW)	2	2	3

Notes

1. Lengths of 50 ohm/52 ohm cable uncritical; lengths of phasing/matching sections are $\lambda/4$.
2. Switching shown for RHCP.
3. 52 ohm coax is RG8/U, velocity factor of 0.66.
4. 73 ohm coax is RG59/U, velocity factor of 0.66.
5. 93 ohm coax is RG62/V, velocity factor of 0.86.
6. 50 ohm hardline is foam dielectric, velocity factor of 0.81.

Figure 5: Crossed-Yagi Feed Arrangement



S1 - Dip Socket to Data Conv.
Q1 - Q4 - VN10KM VMOS FET
V1 - 74LS20 DVAL NAND
V2 - 74LS175 Quad D FF
V3 - 74121 One-Shot
V4 - 74S157 Quad 2-Line
to 1-Line Data Selector

Figure 6: Phase Controller

Precise Phase Noise Measurements of Oscillators and Other Devices from 1 MHz to 20 GHZ

Fred L. Walls
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

Abstract

In this talk the commonly used measures of phase noise are briefly defined and their relationships explained. Techniques for making precise measurements of phase noise in oscillators, multipliers, dividers, amplifiers, and other components are discussed. Particular attention is given to methods of calibration which permit accuracies of 1 dB or better to be achieved. Common pitfalls to avoid are also covered. It is shown that the double balanced mixer approach is the most versatile of these techniques. Phase noise floors (precisions) in excess of −170 dB relative to 1 radian^2 per hertz are achievable for carrier frequencies from well below 1 MHz to the GHz range. The disadvantage for precise source measurements is the need for a reference source of comparable or better performance. This limitation does not apply to the measurement of amplifiers, multipliers, dividers, etc. Other techniques avoid this requirement by using a delay line or cavity to generate a pseudo reference generally with some sacrifice in noise floor near the carrier. Analogues of these techniques are used for carrier frequencies from a few Hz to 10^{15} Hz.

I. Introduction

The output of an oscillator can be expressed as

$$V(t) = [V_o + \varepsilon(t)] \sin(2\pi\nu_o t + \phi(t)) \qquad (1$$

where V_o is the nominal peak output voltage, and ν_o is the nominal frequency of the oscillator. The time variations of amplitude have been incorporated into $\varepsilon(t)$ and the time variations of the actual frequency, $\nu(t)$, have been incorporated into $\phi(t)$. The actual frequency can now be written as

$$\nu(t) = \nu_o + \frac{d[\phi(t)]}{2\pi dt} \qquad (2$$

The fractional frequency deviation is defined as

$$y(t) = \frac{\nu(t) - \nu_o}{\nu_o} = \frac{d[\phi(t)]}{2\pi\nu_o dt} \qquad (3$$

Power spectral analysis of the output signal V(t) combines the power in the carrier ν_o with the power in $\varepsilon(t)$ and $\phi(t)$ and therefore is not a good method to characterize $\varepsilon(t)$ or $\phi(t)$.

Since in many precision sources understanding the variations in $\phi(t)$ or y(t) are of primary importance, we will confine the following discussion to frequency-domain measures of y(t), neglecting $\varepsilon(t)$ except in cases where it sets limits on the measurement of y(t). The amplitude fluctuations, $\varepsilon(t)$, can be reduced using limiters whereas $\phi(t)$ can be reduced in some cases by the use of narrow band filters.

Spectral (Fourier) analysis of y(t) is often expressed in terms of $S_\phi(f)$, the spectral density of phase fluctuations in units of radians squared per Hz bandwidth at Fourier frequency (f) from the carrier ν_o, or $S_y(f)$, the spectral density of fractional frequency fluctuations in a 1 Hz bandwidth at Fourier frequency f from the carrier ν_o [1]. These are related as

$$S_\phi(f) = \frac{\nu_o^2}{f^2} S_y(f) \quad \text{rad}^2/\text{Hz} \quad 0<f<\infty \qquad (4$$

It should be noted that these are single-sided spectral density measures containing the phase or frequency fluctuations from both sides of the carrier.

Other measures sometimes encountered are $\mathcal{L}(f)$, dBC/Hz, and $S_{\Delta\nu}(f)$. These are related by [1,2]

$$S_{\Delta\nu}(f) = \nu_o^2 S_y(f) \quad \text{Hz}^2/\text{Hz}$$

$$\mathcal{L}(f) = (1/2)S_\phi(f) \quad f_1<|f|<\infty \qquad (5$$

$$\text{for} \int_{f_1}^{\infty} S_\phi(f)df < 1 \text{ rad}^2$$

$$\text{dBC/Hz} = 10 \log \mathcal{L}(f)$$

$\mathcal{L}(f)$ and dBC/Hz are single sideband measures of phase noise which are not defined for large phase excursions and are therefore measurement system dependent. Because of this an IEEE subcommittee on frequency stability recommended the use of $S_\phi(f)$ which is well defined independent of the phase excursion [1]. This distinction is becoming increasingly important as users require the specification of phase noise near the carrier where the phase excursions are large compared to 1 radian. Single sideband phase noise can now be specified as $(1/2)S_\phi(f)$.

The above measures provide the most powerful (and detailed) analysis for evaluating types and levels of fundamental noise and spectral density structure in precision oscillators and signal handling equipment as it allows one to examine individual Fourier components of residual phase (or frequency) modulation. On the other hand this analysis is extremely detailed and one often needs an analysis of the long-term average performance.

II. Methods of Measuring Phase Noise

Figure 1 shows the block diagram for a typical scheme used to measure the phase noise of a precision source using a double balanced mixer and a reference source. Fig. 2 illustrates a similar technique for measuring only the added phase noise of multipliers, dividers, amplifiers, and passive components. The output voltage of the mixer as a function of phase deviation, $\Delta\phi$, between the two inputs is normally given by

$$V_{out} = K \cos \Delta\phi \qquad (6$$

Near quadrature this can be approximated by

$$V_{out} = K_d \delta\phi, \text{ where } \delta\phi \equiv [\Delta\phi - \frac{2n-1}{(2)}\pi] < .1 \qquad (7$$

where n is the integer to make $\delta\phi \sim 0$. The phase to voltage conversion ratio sensitivity, K_d, is dependent on the frequency, the drive level, and impedance of both input signals, and the IF termination of the mixer [7]. The combined spectral density of phase noise of both input signals including the mixer and amplitude noise from the IF amplifiers is given by

$$S_\phi(f) = \left(\frac{V_n}{G(f)K_d}\right)^2 \frac{1}{BW} \qquad (8$$

where V_n is the RMS noise voltage at Fourier frequency f from the carrier measured after IF gain G(f) in a noise bandwidth BW. Obviously BW must be small compared to f. This is very important where $S_\phi(f)$ is changing rapidly with f, e.g., $S_\phi(f)$ often varies as f^{-3} near the carrier. In Fig. 1, the output of the second amplifier following the mixer contains contributions from the phase noise of the oscillators, the mixers, and the post amplifiers for Fourier frequencies much larger than the phase-lock loop bandwidth. In Figure 2, the phase noise of the oscillator cancels out to a high degree (often more than 20 dB). Termination of the mixer IF port with 50 ohms maximizes the IF bandwidth, however, termination with reactive loads can reduce the mixer noise by ~ 6 dB, and increase K_d by 3 to 6 dB as shown in Fig. 3. [3] Accurate determination of K_d can be achieved by allowing the two oscillators to slowly beat and measuring the slope of the zero crossing in volts/radian with an oscilloscope or other recording device. The time axis is easily calibrated since one beat period equals 2π radians. Estimates of K_d obtained from measurements of the peak to peak output voltage sometimes introduce errors as large as 6 dB in $S_\phi(f)$ [3]. By comparing the level of an IF signal (a pure tone is best), on the spectrum analyzer used to measure V_n with the level recording device used to measure K_d, the accuracy of $S_\phi(f)$ can be made independent of the accuracy of the spectrum analyzer voltage reference. Some care is necessary to assure that the spectrum analyzer is not saturated by spurious signals such as the line frequency and its multiples. Sometimes aliasing in the spectrum analyzer is a problem. Typical best performance is shown in Fig. 4. This measurement system exceeds the performance of almost all available oscillators from 0.1 MHz to 10 GHz and is generally the technique of first choice because of its versatility and simplicity. The use of specialized mixers with multiple diodes per leg increases the phase to voltage conversion sensitivity, K_d and therefore reduces the contribution of IF amplifier noise [4] as shown in Fig. 4. The resolution of the above systems can be greatly enhanced (typically 20 dB) using correlation techniques to separate the phase noise from the device under test from the noise in the mixer and IF amplifier [4].

For example consider the scheme illustrated in Figure 5. At the output of each double balanced mixer there is a signal which is proportional to the phase difference, $\Delta\phi$, between the two oscillators and a noise term, V_N, due to contributions from the mixer and amplifier. The voltages at the input of each bandpass filter are

$$V_1(\text{BP filter input}) = G_1\, \Delta\phi(t) + C_1 V_{N1}(t) \qquad (9$$

$$V_2(\text{BP filter input}) = G_2\, \Delta\phi(t) + C_2 V_{N2}(t)$$

where $V_{N1}(t)$ and $V_{N2}(t)$ are substantially uncorrelated. Each bandpass filter produces a narrow band noise function around its center frequency f:

$$V_1(\text{BP filter output}) = G_1[S_\phi(f)]^{1/2} B_1{}^{1/2} \cos[2\pi ft + \psi(t)]$$

$$+ C_1[S_{VN1}(f)]^{1/2} B_1{}^{1/2} \cos[2\pi ft + n_1(t)] \qquad (10$$

$$V_2(\text{BP filter output}) = G_2[S_\phi(f)]^{1/2} B_2{}^{1/2} \cos[2\pi ft + \psi(t)]$$

$$+ C_2[S_{VN2}(f)]^{1/2} B_2{}^{1/2} \cos[2\pi ft + n_1(t)]$$

where B_1 and B_2 are the equivalent noise bandwidths of filters 1 and 2 respectively. Both channels are bandpass filtered in order to help eliminate aliasing and dynamic range problems. The phases $\psi(t)$, $n_1(t)$ and $n_2(t)$ take on all values between 0 and 2π with equal likelihood. They vary slowly compared to 1/f and are substantially uncorrelated. When these two voltages are multiplied together and low pass filtered only one term has finite average value. The output voltage is

$$V^2_{out} = 1/2\, G_1 G_2 S_\phi(f)\, B_1{}^{1/2} B_2{}^{1/2} + D_1\langle\cos[7(t)]\rangle \qquad (11$$

$$+ D_2\cos[\psi(t) - n_2(t)]\rangle + D_3\langle\cos[n_1(t)-n_2(t)]\rangle$$ so that $S_\phi(f)$ is given by

$$S_\phi(f) = \frac{(2)V_N^2}{G_1G_2\sqrt{B_1B_2}} \qquad (12$$

For times long compared to $B_1^{-1/2}B_2^{-1/2}$ the noise terms D_1, D_2 and D_3 tend towards zero as $\sqrt{t}$. Limits in the reduction of these terms are usually associated with harmonics of 60 Hz pickup, dc offset drifts, and nonlinearities in the multiplier. Also if the isolation amplifiers have input current noise then they will pump current through the source resistance. The resulting noise voltage will appear coherently on both channels and can't be distinguished from real phase noise between the two oscillators. One half of the noise power appears in amplitude and one half in phase modulation.

Obviously the simple single frequency correlator used in this illustration can be replaced by a fast digital system which simultaneously computes the correlated phase noise for a large band of Fourier frequencies. Typical results show a reduction in noise floor of order 20 dB over the noise floor of a single channel (See Fig. 4). The great power of this technique is that it can be applied at any carrier frequency where one can obtain double balanced mixers. The primary limitations come from the bandwidth and nonlinearities in the cross correlator.

Another method of determining $S_\phi(f)$ uses phase modulation of the reference oscillator by a known amount. The ratio of the reference phase modulation to the rest of the spectrum then can be used for a relative calibration. This approach can be very useful for measurements which are repeated a great many times.

It is sometimes convenient to use a high-Q resonance directly as a frequency discriminator as shown in Fig. 6.

The oscillator is typically tuned 1/2 linewidth ($\nu_o/2Q$) away from line center yielding a detected signal of the form

$$V_{out} = G(f)kQdy(f)\,[V + \epsilon(t)] \qquad (13$$

Note that this approach mixes frequency fluctuations between the oscillator and reference cavity with the amplitude noise of the transmitted signal. By using amplitude control (e.g. by processing to normalize the data), one can reduce the effect of amplitude noise. [5] The measured noise at the detector is then related to the oscillator reference cavity phase fluctuation by

$$S_\phi(f) = \left(\frac{\nu_o V_N}{f\nu_o QkG(f)}\right)^2 \frac{1}{BW} \qquad (14$$

This approach has the limitations that $\Delta\nu$ must be small compared to the linewidth of the cavity, and removing the effect of residual amplitude noise is difficult; however, no reference source is needed.

Differential techniques can be used to measure the inherent frequency (phase) fluctuations of two High-Q resonators as shown in Fig. 7 [6]. The output voltage is of the form $V_{out} = 2QK_d\, dy(f)$. The phase noise spectrum of the resonators is then obtained using equation 4.

$$S_\phi(f) = \left(\frac{\nu_o V_N}{2QfK_d}\right)^2 \frac{1}{BW} \qquad (15$$

The phase noise in the source can cancel out by 20 to 40 dB depending on the similarity of resonate frequencies Q's and the transmission properties of the two resonators. This approach was first used to demonstrate that the inherent frequency stability of precision quartz resonators exceeds the performance of most quartz crystal controlled oscillators [6].

A still different approach uses a delay line to make a pseudo reference which is retarded relative to the incoming signal [7-10] as shown in Fig. 8.

The mixer output is of the form

$$V_{out} = 2\pi t_d K_d \nu_o dy \qquad (16$$

and the oscillator phase noise is given by

$$S_\phi(f) = \left(\frac{V_n}{2\pi f\tau_d G(f)\, K_d}\right)^2 \frac{1}{BW} \qquad f < \frac{1}{\tau_d} \qquad (17$$

This approach is often used at microwave frequencies when only one oscillator is available. However the phase noise close to the carrier becomes virtually unresolvable for a finite delay line. For example if $f = 1$ Hz and $\tau_d = 500$ ns, then, $(2\pi f\tau_d)^2 \sim 10^{-11}$. The noise floor of this technique is 110 dB higher at $f = 1$ Hz than that of the two oscillator method and it decreases as $1/f^2$. The noise floor can be reduced by ~ 20 to 40 dB using the correlation techniques described above. [10]

The use of frequency multipliers (or dividers) between the oscillators and the double balanced mixer increases (decreases) the phase noise level [11] as

$$S_{\phi\,\nu_2}(f) = \left(\frac{\nu_2}{\nu_1}\right)^2 S_{\phi\,\nu_1}(f) \qquad (18$$

Figure 4 shows the noise of a specialized 5 to 25 MHz multiplier referred to the 5 MHz input. A potential problem with the use of the multiplier approach comes from exceeding the dynamic range of the mixer. Once the phase excursion, $\Delta\phi$, exceeds about 0.1 radian, nonlinearities start to become important and at $\Delta\phi$ - 1 radian, the measurement is no longer valid [11].

References

1. J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, Jr., W. L. Smith R. L. Sydnor, R. F. C. Vessot, G. M. Winkler, Characterization of Frequency Stability, Proc. IEEE Trans. on I & M 20, 105-120 (1971).

2. J. H. Shoaf, D. Halford, and A. S. Risley, Frequency Stability Specifications and Measurement, NBS Technical Note 632, (1973). Document available from US Government printing office. Order SD at #C13.46:632.

3. F. L. Walls, and S. R. Stein, Accurate Measurements of Spectral Density of Phase Noise in Devices, Proc. of 31st SFC, 335-343, (1977). (National Technical Information Service, Sills Building, 5825 Port Royal Road, Springfield, VA 22161).

4. F. L. Walls, S. R. Stein, J. E. Gray, and D. J. Glaze, Design Considerations in State-of-the-Art Signal Processing and Phase Noise Measurement Systems, Proc. 30th Ann. SFC, 269-274 (1976). (National Technical Information Service, Sills Building, 5285 Port Royal Raod, Springfield, VA 22161).

5. R. L.. Barger, M. S. Soren, and J. L. Hall, Frequency Stabilization of a cw Dye Laser, Appl. Phys. Lett. 22, 573 (1973).

6. F. L. Walls and A. E. Wainwright, Measurement of the Short-Term Stability of Quartz Crystal Resonators and the Implications for Crystal Oscillator Design and Applications, IEEE Trans. on I & M 24, 15-20 (1975).

7. A. S. Risley, J. H. Shoaf, and J. R. Ashley, Frequency Stabilization of X-Band Sources for Use in Frequency Synthesis into the Infrared, IEEE Trans. on I & M, 23, 187-195 (1974).

8. J. R. Ashley, T. A. Barley, and G. J. Rast, The Measurement of Noise in Microwave Transmitters, IEEE Trans. on Microwave Theory and Techniques, Special Issue on Low Noise Technology, (1977).

9. A. L. Lance, W. D. Seal, F. G. Mendoza, and N. W. Hudson, Automating Phase Noise Measurements in the Frequency Domain, Proc. 31st Ann. Symp. on Freq. Control, 347-358 (1977).

10. A. L. Lance and W. D. Seal, Phase Noise and AM Noise Measurements in the Frequency Domain at Millimeter Wave Frequencies, from Infrared and Millimeter Waves, Ken Button Ed., Academic Press, NY 1985.

11. F. L. Walls and A. DeMarchi, RF Spectrum of a Signal After Frequency Multiplication Measurement and Comparison with a Simple Calculation, IEEE Trans. on I & M 24, 210-217 (1975).

Figure Captions

Fig. 1. Precision phase measurement system using a spectrum analyzer. Calibration requires a recording device to measure the slope at the zero crossing. The accuracy is better than 0.2 dB from dc to 0.1 ν_o Fourier frequency offset from the carrier ν_o. Carrier frequencies from a few Hz to 10^{10} Hz can be accommodated with this type of measurement system. [3]

Fig. 2. Precision phase measurement system featuring self calibration to 0.2 dB accuracy from dc to 0.1 ν_o Fourier frequency offset from carrier. This system is suitable for measuring signal handling equipment, multipliers, dividers, frequency synthesizers, as well as passive components. [3]

Fig. 3. Double-balanced mixer phase sensitivity at 5 MHz as a function of Fourier frequency for various output terminations. The curves on the left were obtained with 10 mW drive while those on the right were obtained with 2 mW drive. The data demonstrate a clear choice between constant, but low sensitivity or much higher, but frequency dependent sensitivity. [3]

Fig. 4.

Curve A. The noise floor $S_\phi(f)$ (resolution) of typical double balanced mixer systems (e.g. Fig. 1 and Fig. 2) at carrier frequencies from 1 to 100 MHz. Similar performance possible to 20 GHz. [4]

Curve B. The noise floor, $S_\phi(f)$, for a high level mixer. [4]

Curve C. The correlated component of $S_\phi(f)$ between two channels using high level mixers. [4]

Curve D. The equivalent noise floor $S_\phi(f)$ of a 5 to 25 MHz frequency multiplier.

Curve E. Approximate phase noise floor of Figure 8 using a 500 ns delay line.

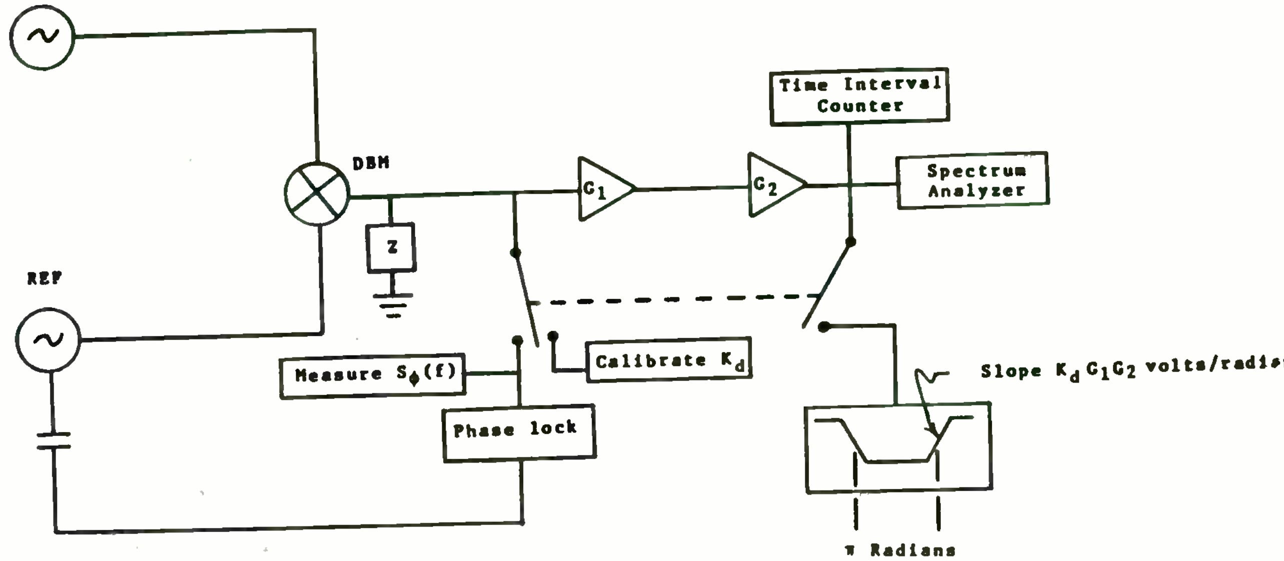
Fig. 5. Correlation phase noise measurement system.

Fig. 6. High-Q resonance used as a frequency discriminator.

Fig. 7. Differential cavity frequency discriminator.

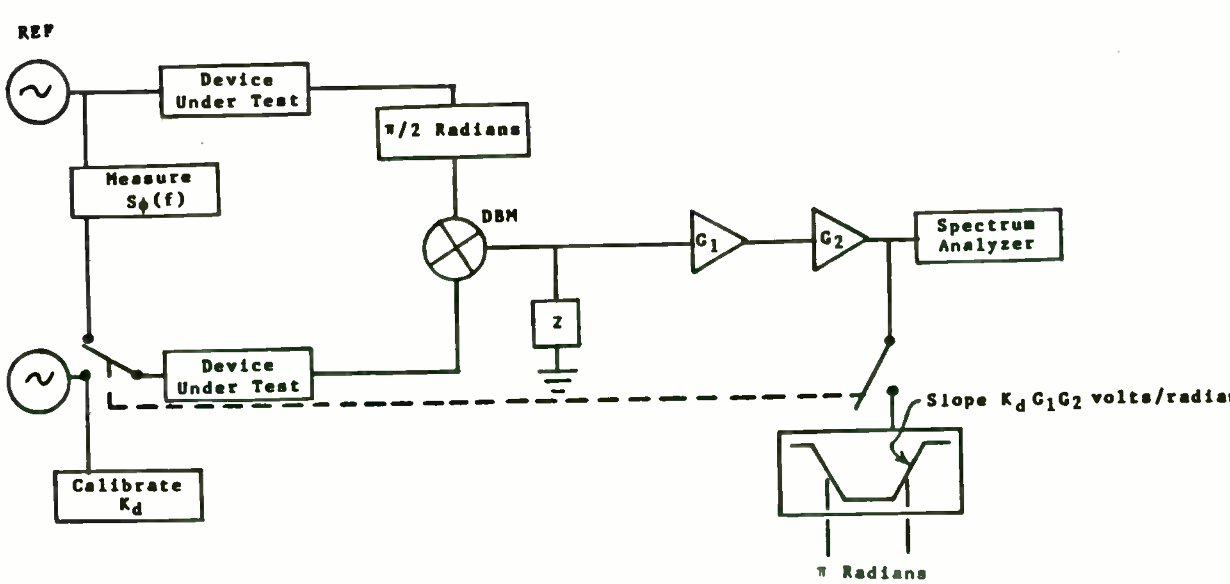
Fig. 8. Delay line frequency discriminator.

Fig 1

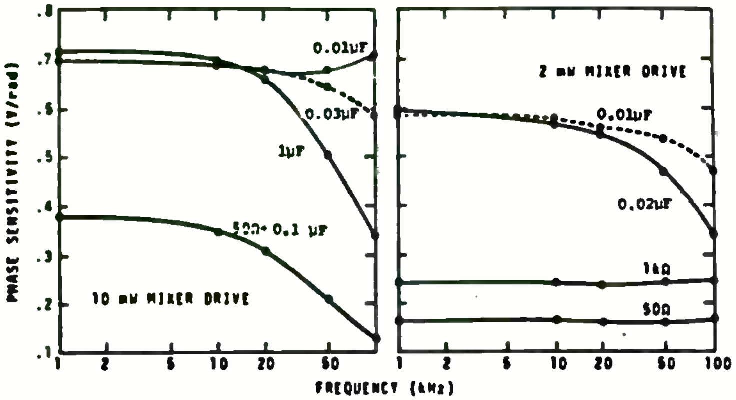


DBM
REF
Z
G_1
G_2
Time Interval Counter
Spectrum Analyzer
Measure $S_\phi(f)$
Calibrate K_d
Phase lock
Slope $K_d G_1 G_2$ volts/radian
π Radians

Fig 2



REF
Device Under Test
$\pi/2$ Radians
Measure $S_\phi(f)$
DBM
Z
G_1
G_2
Spectrum Analyzer
Device Under Test
Calibrate K_d
Slope $K_d G_1 G_2$ volts/radian
π Radians


PHASE SENSITIVITY (V/rad)
.8
.7
.6
.5
.4
.3
.2
.1
0.01μF
0.03μF
1μF
50Ω+0.1 μF
10 mW MIXER DRIVE
2 mW MIXER DRIVE
0.01μF
0.02μF
1kΩ
50Ω
1
2
5
10
20
50
100
FREQUENCY (kHz)

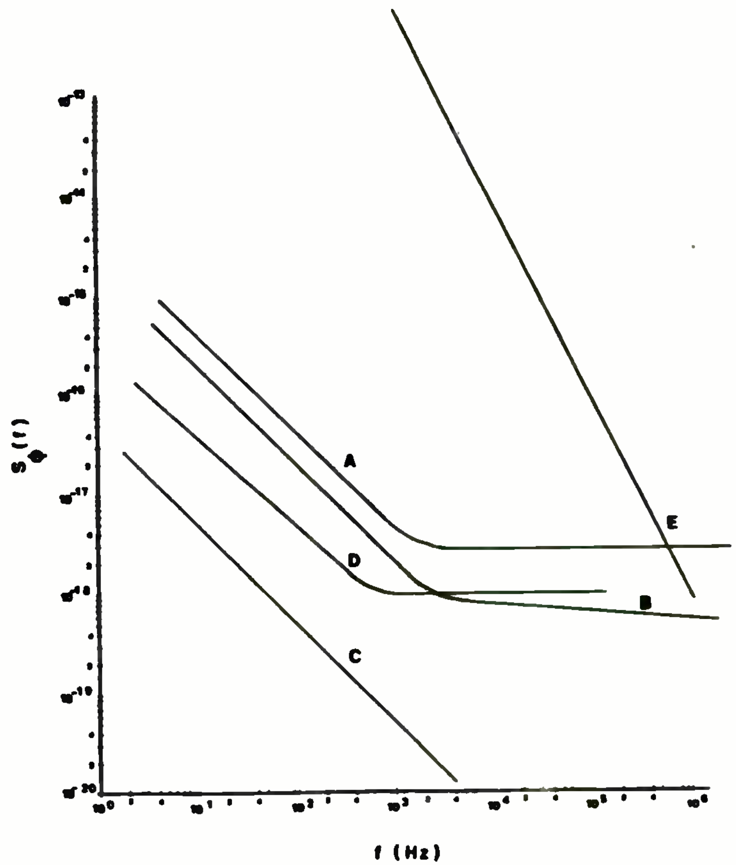

S_φ(f)
A
B
C
D
E
f (Hz)

Fig 5

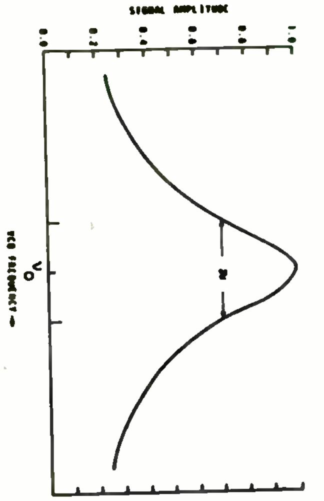
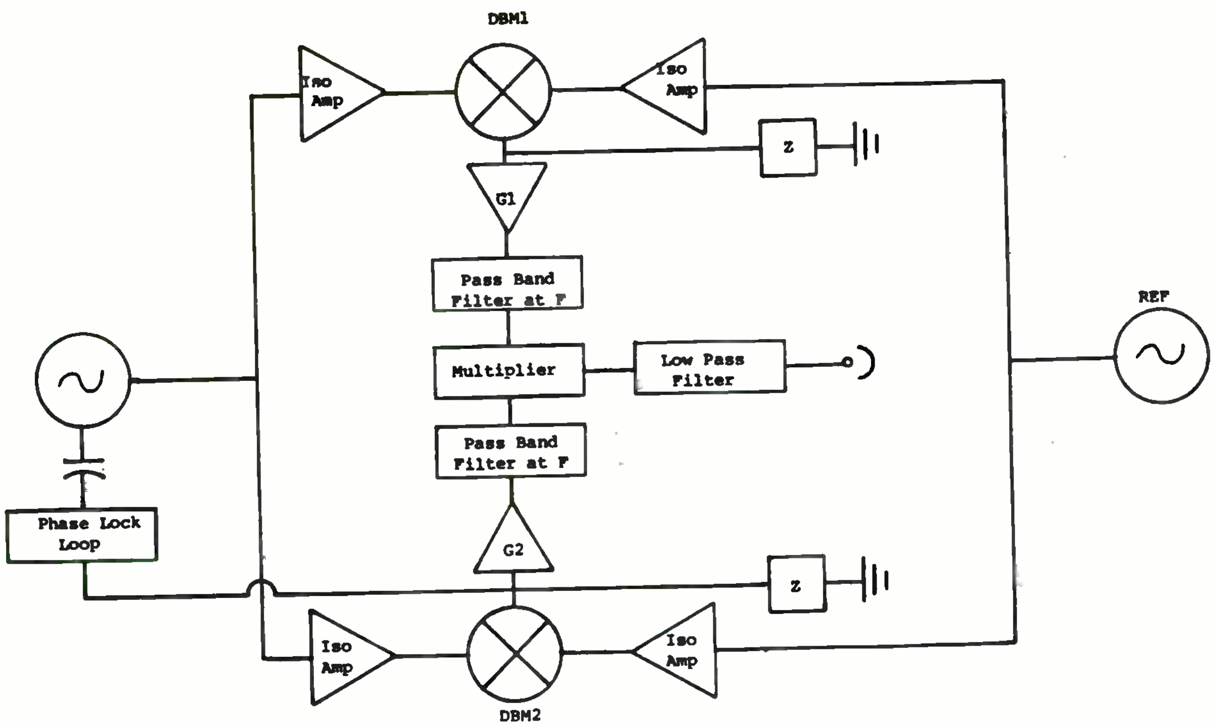


Fig 6


DBM1
Iso Amp
Iso Amp
Z
G1
Pass Band Filter at F
Multiplier
Low Pass Filter
Pass Band Filter at F
G2
Z
Iso Amp
Iso Amp
DBM2
REF
Phase Lock Loop

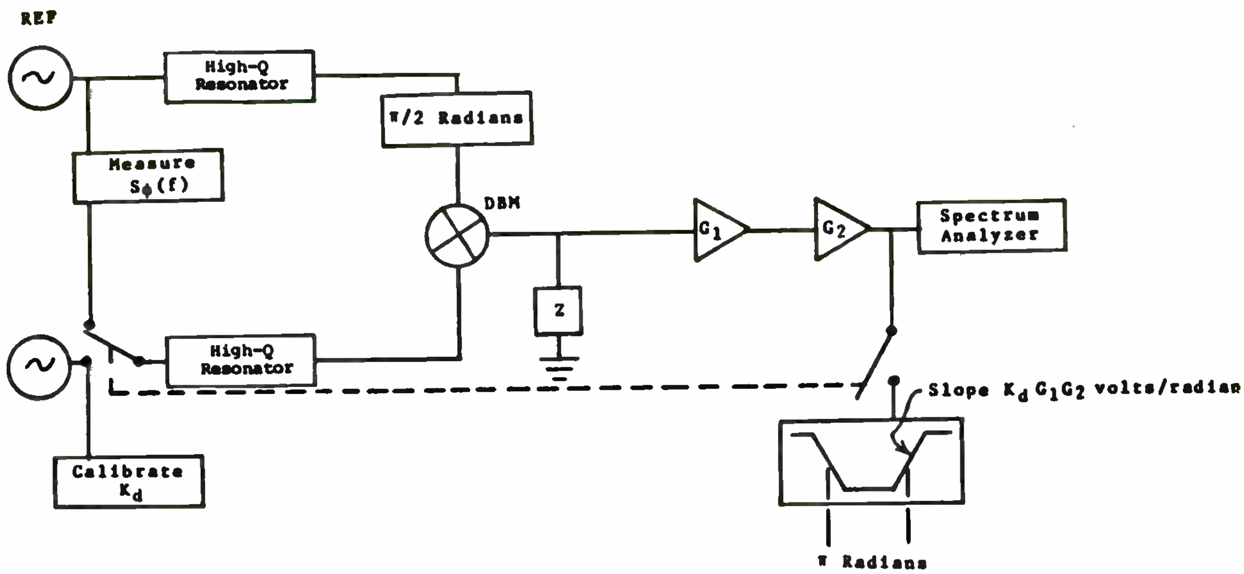
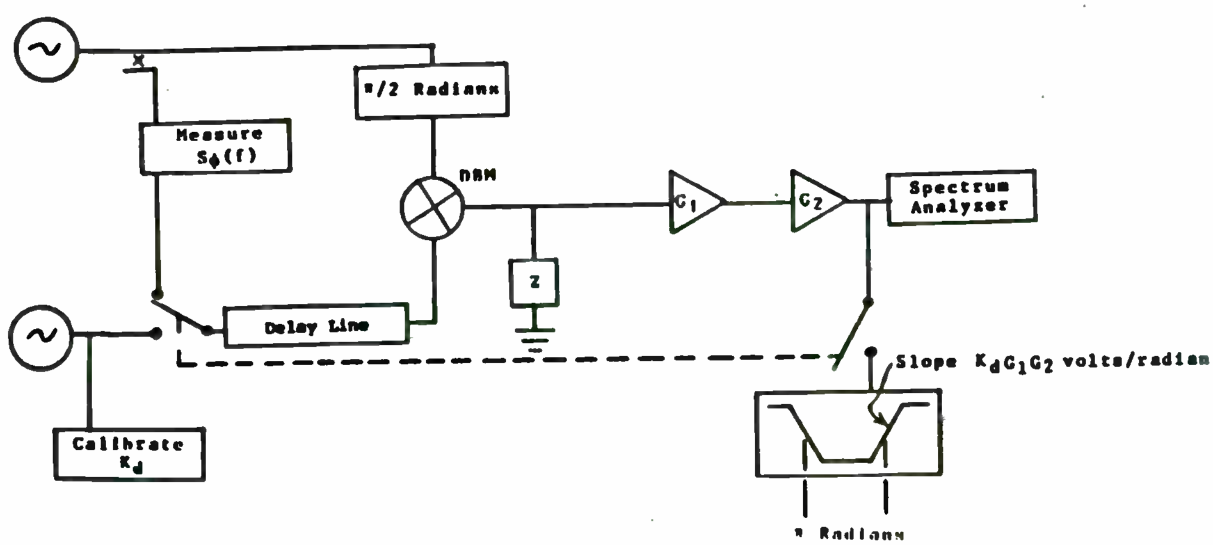

REF
High-Q Resonator
π/2 Radians
Measure Sϕ(f)
DBM
G1
G2
Spectrum Analyzer
Z
High-Q Resonator
Slope Kd G1G2 volts/radian
Calibrate Kd
π Radians

Fig 8


π/2 Radians
Measure Sϕ(f)
DBM
G1
G2
Spectrum Analyzer
Z
Delay Line
Slope KdG1G2 volts/radian
Calibrate Kd
π Radians

APPLICATIONS OF DIGITAL SIGNAL PROCESSING

by

Tom Callaghan

WATKINS-JOHNSON COMPANY
CEI Division
700 Quince Orchard Road
Gaithersburg, Maryland 20878

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

When one first looks into digital signal processing, (DSP), one finds that the literature and textbooks abound with numerous theories and mathematical equations. Many of these concepts are difficult to understand when first encountered, the reasons being that DSP is a new area to many; and there is no base knowledge upon which to build. Therefore, it is the purpose of this paper to provide that base. The paper will treat DSP in a practical manner, comparing it with analog applications whenever possible. The paper will include enough of the theory to maintain accuracy and understanding, but shall attempt to keep towards the applications side of DSP.

The paper starts with the input signal itself and discusses how the signal is sampled including filter considerations and sampling rates. Next, the paper covers how the sampled signal is quantized, or analog/digital converted. This section includes different quantization rules, dynamic range, and bit size of the digital word. To complete the overview of DSP, the digital processor itself is discussed. An overview of digital processors is presented, including architectures best suited for signal processing, parallel versus sequential structures, and some of the more useful functions for DSP.

Next, the paper covers two of the more common applications of DSP; namely, digital filters and Fast Fourier Transforms (FFT). In this part of the paper, two types of filters are discussed: the infinite impulse response (IIR) filter and the finite impulse response (FIR) filter. In the discussion on FFT's, two different algorithms are presented, the base 2 decimation in time and the base 2 decimation in frequency. The use of window functions for FFT's is also covered. Finally, the paper concludes with the advantages of using digital signal processing over analog techniques.

DIGITAL SIGNAL PROCESSING OVERVIEW

All aspects of DSP fall into three main categories:

1. Conversion of the input signal
2. The digital processor
3. Information extraction

The Input Signal

The input signal, which is analog in nature, can be classified as a continuous time, continuous amplitude signal. To be of a form useful to a digital processor, the signal must be converted to a discrete time, discrete amplitude signal.

Sampling

Sampling is the process by which the continuous time, continuous amplitude signal is converted to a discrete time, continuous amplitude signal. This is done by periodically taking minute time chunks out of the analog signal. The output of a sampler is a series of pulses whose envelope approximates the original signal. The process can be represented by multiplying the input signal by a pulse train of uniform amplitude and equal spacing, see **Figure 1.**

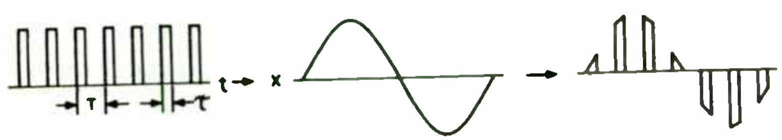


Figure 1. Characteristic Signals In Sampling

The pulse spacing, T, that is the sampling frequency, 1/T, cannot be arbitrarily chosen. According to the Sampling Theorem, "If a continuous time function contains only frequency components below F cycles per second, 2F samples per second suffice to represent it perfectly and permit perfect recovery." The reason for this is more readily seen in the frequency domain.

Referring to **Figure 2**, one can see that one of the byproducts of sampling is the duplication of the signal every Fs, or the sample frequency. If the sample frequency is kept above or equal to 2F, **Figure 2b**, the sampled frequency spectrum produces no overlap; and the original signal can be retrieved. If the sample frequency drops below 2F, the sample frequency spectrum starts to overlap. Any of the frequencies whose spectra are contained in this overlap tend to fold over about Fs/2 and into the actual frequency band, as indicated in **Figure 2c**. This foldover of frequency is called aliasing.

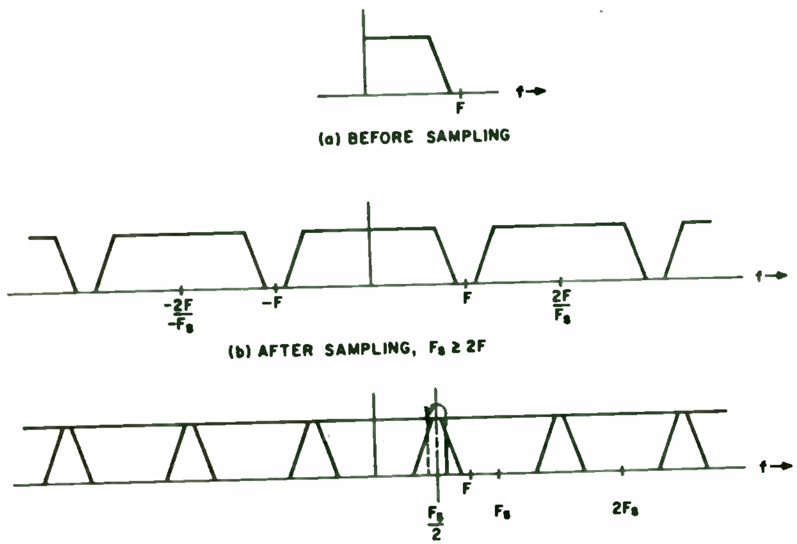


Figure 2. Frequency Domain

Examples of frequency aliasing are witnessed daily by many unknowing observers. For example, in the old "Westerns", the stage coach rides out of town; and its wheels appear to speed up, stop, and then go backwards. The backwards revolution is caused by the movie camera not taking sufficient frames per second of the wheel to dependably capture its revolution. Another example is when a television or computer terminal is shown in a movie, with black bars across the screen. Again, the movie camera is not updating the screen fast enough. In this case, the camera moves at 24 frames per second while the television scans a new frame at 60 times per second.

The sampling theory is all well and good on paper, but in reality it is difficult to remove all unwanted frequencies above the frequency of interest. The most common solution is to place a filter ahead of the sampler to remove unwanted frequencies. The filter, however, has two disadvantages: it alters the signal of interest and, unless it has very steep skirts, certain unwanted frequencies are passed. A second solution is to sample at a faster rate than 2F. Here again, how much faster is open to debate. Sample too fast and one puts an undue burden on the digital processor.

Another consideration in determining sampling frequency is interpolation errors. Since one is dealing in discrete time, it is not known what has happened to the input signal between samples. Naturally, the faster the sample rate the more information known about the signal. When processing is finished, it may be desired to convert back to an analog signal. The conversion is usually done with a D/A converter and some type of low-pass filter to smooth, or interpolate, the discrete points.

So how does one select the sample rate? Gardenhire in [1] presents a good approach to the problem. In his work, the sampling rate is based upon the amount of interpolation error tolerable from the output filter given a specified input filter roll-off characteristic. The input filter roll-off determines the order of the system. Given a certain roll-off, m, the sample rate can be determined from the acceptable interpolation error. The rate is given for a 5% error in Table 1 and a 0.2% error in Table 2.

Table 1. Systems Resulting in a 5% Interpolation Error [1]

Normalized sample frequency (F_s/f_1)

Interpolation Method		m=1	m=2	m=3	m=4	m=5	m=
Wiener Optimum Filter		640	11	5.1	3.8	2.6	2.0
Butterworth	n=4	-	16	8.3	5.5	5.5	5.5
	n=3	-	18	9.2	6.7	6.7	6.7
	n=2	$1.2x10^3$	29	17.0	11	11	11
RC Filter	n=1	$1.2x10^4$	220	130	91	91	91
2 Point and Linear Interpolation		640	13	8.3	5.9	5.9	5.9
3 Point Linear Interp.		640	12	6.2	5.2	-	4.0
4 Point Linear Interp.		640	12	5.7	4.3	-	3.3

In these two tables, the filter methods employ analog interpolation after D/A conversion. The 2, 3, or 4 point linear interpolation methods are done via a computer; and the output is still in a digital format. It is interesting to note how the sample rate approaches the theoretical rate of 2F as the order of the filter increases. It is also worth noting how the sample rate increases drastically as one tries to follow the input signal more closely with a minimum of error. Figure 3 shows the non-idealized filter frequency response for various valves of m.

For communications applications, however, one is dealing primarily with input filters whose order does approach infinity. The output filters are also optimized, or of Butterworth types. For these reasons a sample rate in the range of 2 to 5 times F should prove satisfactory.

Table 2. Systems Resulting in a 0.2% Interpolation Error [1]

Normalized sample frequency (F_s/f_1)

Interpolation Method		m=1	m=2	m=3	m=4	m=5	m=
Wiener Optimum Filter		$4.1x10^5$	93	19	10	3.4	2.0
Butterworth Filter,	n=4	-	290	61	26	26	26
	n=3	-	430	100	46	46	46
	n=2	-	$1.1x10^3$	300	160	160	160
Linear Phase Filter		-	$1.5x10^3$	610	400	400	400
RC Filter,	n=1	-	$2.0x10^3$	850	540	540	540
2 Point Linear Interp.		$4.1x10^5$	120	42	29	29	29
3 Point Linear Interp.		$4.1x10^5$	105	25	17	-	12
4 Point Linear Interp.		$4.1x10^5$	100	21	13	-	7.8

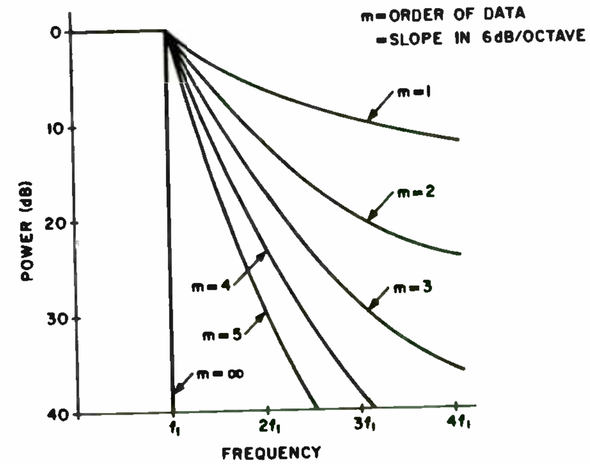


Figure 3. Non-Idealized Frequency Response [1]

Another error that appears during sampling is known as aperture error. This error is caused by the sample pulses taking a small amount of time to capture the analog signal. During this time, the input voltage can change drastically as the signal goes through zero. This is indicated in **Figure 1** by the triangular shapes of some of the sampled data pulses. The aperture error can be expressed mathematically [2] for a sinusoidal input by:

$$\epsilon = \frac{\Delta V}{V_{FS}} = 2\pi\tau,$$ where V_{FS} is the full scale voltage and τ is the duration of the pulse.

Expressed another way, the acceptable pulse duration can be calculated as

$$\tau = (2\pi \times F \times 2^n)^{-1}$$ where n is the number of bits of resolution of the A/D converter

for an A/D converter with a 1LSB error (the 1LSB error is explained in the next section). For instance, if the input frequency is 10 kHz and eight bits of resolution are used, the allowable pulse duration = 63.5 ns.

To alleviate aperture error, a track and hold amplifier can be inserted before the A/D converter. This device's output will follow the input signal while its track digital level is active. When the track input switches, the output is held at its current level and will not vary over the duration of the pulse.

Quantization

After the discrete time, continuous amplitude signal is obtained, it must then be converted to a discrete time, discrete amplitude signal for the digital processor. This process is formally called quantization, but it is more commonly known as analog to digital (A/D) conversion. A quantizer takes a specific amplitude range and divides it into a series of discrete steps, Q. A digital number is then assigned to each Q.

The number of bits in the digital word determines the number of steps that can be achieved. For n bits, the number of steps would be 2^n. Each step Q, or for each change of 1LSB, would be

$Q = V_{FS}/2^n$, where V_{FS} is the full scale magnitude of the allowable input voltage.

If the input signal falls between steps then the digital number assigned to it depends on whether the quantizer rounds the samples or truncates them.

Figure 4a shows the quantizer characteristics with rounding. As its name implies, the quantizer rounds off the analog input to the nearest quantizer level, Q. In truncation, **Figure 4b**, the signal is represented by the highest Q level that is not greater than the signal. Thus, for each LSB (least significant bit) step of the digital output word, the error from the original signal would be ±1/2 LSB for rounding and a 0 or +1LSB for truncation. An error of ±1/2 LSB yields a mean error of zero, whereas a 1LSB error yields a mean error of + 1/2 LSB. For this reason, rounding is preferred in most practical considerations.

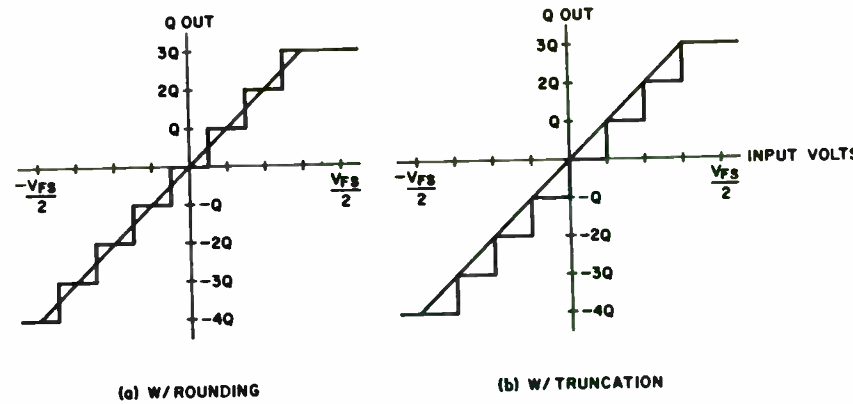


Figure 4. Quantizer Characteristics

How the digital word is represented is another area of consideration. Binary representation varies widely for positive and negative numbers. Again, for most practical considerations, two's complement representation is chosen because most processors use this type of representation. Also, for other applications, only positive numbers are expected. Many A/D converters allow either; that is, a range of 0 to V_{FS} or $-V_{FS}/2$ to $+V_{FS}/2$, as well as other representations. **Table 3** shows some typical outputs for different types of input format with 8 bit resolution.

Finally, once the quantizer or A/D converter is chosen, the number of bits of resolution must be decided. It is important to point out here that while a large number of bits will represent an analog signal more accurately, they will not represent a cleaner analog signal. In every analog signal, there exists some inherent noise that is some small portion of that signal. With more bits of resolution, the smaller the step size in the quantizer. The smaller the step, the more the digital word is affected by noise, so that the lower significant bits only serve to give a good representation of this noise.

Table 3a **Digital Word for Bipolar Inputs, $\pm V_{PS}/2$ $V_{PS} = 10V$**

INPUT	OUTPUT MSB LSB
-5V	10000000
-4.961	10000001
-4.922	10000010
.	.
.	.
.	.
-0.078	11111110
-0.039	11111111
0.000	00000000
0.039	00000001
.	.
.	.
.	.
4.922	01111110
4.961	01111111

Table 3b **Digital Word for Unipolar Inputs, $V_{PS} = 10V$**

INPUT	OUTPUT MSB LSB
0.000	00000000
0.039	00000001
0.078	00000010
.	.
.	.
.	.
5.000	10000000
.	.
.	.
.	.
9.961	11111111

Rather, it would be better to base the number of bits of resolution on the dynamic range desired in the system. A formula that relates the number of bits of resolution, n, to dynamic range is:

$$\text{Range (dB)} = \left| -20 \log (2^n) \right|$$

or rewriting to solve for n

$$n = \text{Range (dB)} / 6$$

There also exists a certain amount of quantization noise due to the ±1/2 LSB error. To allow for this, it can be shown that [2] another way of expressing dynamic range is as a signal to noise ratio

$$\text{SNR} = 6n - 1.24 \text{ dB}$$

Digital Processor

A digital processor refers to any form of hardware and/or software that has been built and/or programmed to perform signal processing algorithms on the data output from the sampler/quantizer. This definition encompasses a broad range of computing power. For the purposes of this paper, we shall concentrate on those processor architectures that lend themselves to signal processing and on the basic structures used in determining the processor(s) configuration.

Basically a digital processor should be modular in design; i.e., the processor, by itself, is capable of such rudimentary tasks as addition, subtraction, multiplication, memory transfers, and I/O transfers. Additional support such as direct interface to A/D and D/A converters, and the capability of working with blocks of data would also be helpful. Although it is not expected to have all of these capabilities, the processor should still perform whatever capabilities it does have without outside support. In this way, when a simple task needs to be executed, only one module is used. As the task becomes more complicated, processor modules can be added to handle the extra load. Indeed, most applications in signal processing lend themselves readily to this modularity. Each application can be broken down into several smaller algorithms which can be further decomposed into simpler tasks.

Multiple processors can be implemented using two basic approaches, a parallel structure or a sequential structure. A parallel structure is defined by Tewksbury, et al, in 5 as being the concurrent execution of several functional operations using a number of distinct functional modules under the coordination of a common control structure. A sequential structure, on the other hand, sequentially executes all the functional operations of the algorithm in a single functional module. One could go on to say that a sequential structure could have several modules, but one module cannot begin processing until the one preceding it is finished.

Parallel Structures

A parallel structure is shown in **Figure 5.** Each processor receives its data from a shared input source. Each processor then sends its processed data to a shared output. Using this type of structure, the functional operation speed can be increased by dividing the data processing over a greater number of processing modules. However, there is a point of diminishing returns for additional processors due to the increased complexity of the control structure.

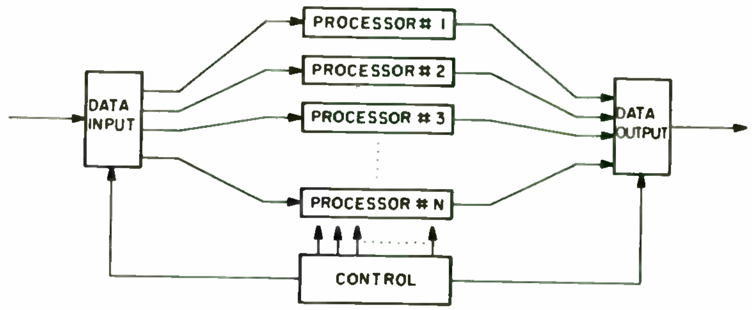


Figure 5. Parallel Processor Structure

Another advantage of a parallel structure is that each processor module can be composed of an efficient functional unit with limited capabilities. This does, however, deviate from the strict definition of modularity and also does allow fewer common tasks, such as multiplication, to be handled by processors especially designed for the functional task. Other general purpose processors can then handle some of the more common tasks; i.e., overflow detection, as well as be reprogrammed to handle different tasks. In this way overall data processing speed can be increased for the structure by allowing each processor to handle the tasks for which it is best suited.

One of the disadvantages of parallel structures is that their control structures can be quite complex. Since each module is operating at a different throughput, timing of all the functions becomes another critical area. For both of these reasons, hardware design and consequently, programming become difficult.

Sequential Structures

A sequential processor is shown in **Figure 6**. The first processor in the sequence receives the incoming data. The first processor performs its assigned tasks on the data and passes along this new data to the next stage. The process continues as each processor performs its specific tasks until the final data output is reached. It is not necessary that each processor work on all the incoming data. It could just pass some of the data through to the next stage. The next stage's processor, however, would wait until it received all of the data from the previous stage before it started processing.

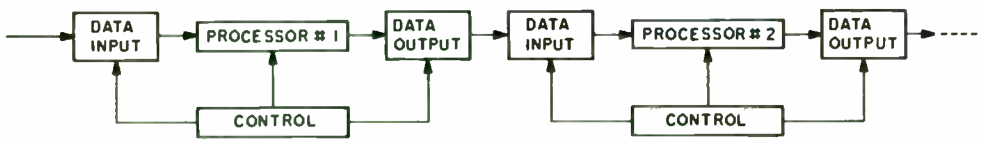


Figure 6. Sequential Processor Structure

The added delay caused by each stage waiting for the previous stage to finish processing is one of the sequential structures disadvantages. This same delay also produces one of the advantages; it simplifies controller design. The controller does not have to keep track of data that is earmarked for a particular stage. At each stage, all of the data is passed along at once. A simpler controller allows more efficient, more straightforward hardware designs. Consequently, sequential structures are easier to program.

Parallel vs Sequential

Ultimately, the choice between parallel or sequential structuring depends greatly on how well a particular signal processing algorithm lends itself to one or the other. Indeed, some algorithms may incorporate both. Another design consideration would involve a trade-off of hardware/software cost and complexity to that of speed. If speed is not a prime consideration, then the less expensive, easier hardware designs of the sequential processor would be preferred.

Information Extraction

The final end product of any form of signal processing, analog or digital, is the extraction of information from the input signal. The information can be a breakdown of the signal into its frequency components for signal analysis. The information could also be the altering of the signal's spectrum as in filtering. It can be the intelligence content of the signal as in detection.

Whatever the reason for signal processing, the end product is as varied as the system designers who try to make use of it. This paper will now continue with two of the more common applications of digital signal processing: Filters and Fast Fourier Transforms.

DIGITAL FILTERS

Signal Flow Notation

By the very nature of digital signal processing, one is limited to only three types of operations in a processor. These are shown in **Figure 7a**. The common form of gain or multiplication, summation, and delay for DSP is in signal flow graph notation as illustrated in **Figure 7b**.

Difference Equations

In analog systems, filters are designed by using differential equations. The filter transfer function H(t), such as

$$x(t) \longrightarrow \boxed{H(t)} \longrightarrow y(t) = X(t) * H(t)$$

is defined in terms of differential equation of y and x. For digital systems, these equations are replaced with difference equations. There exist two forms most commonly used to describe realizable digital filters. One form is the recursive form and the other is the non-recursive form.

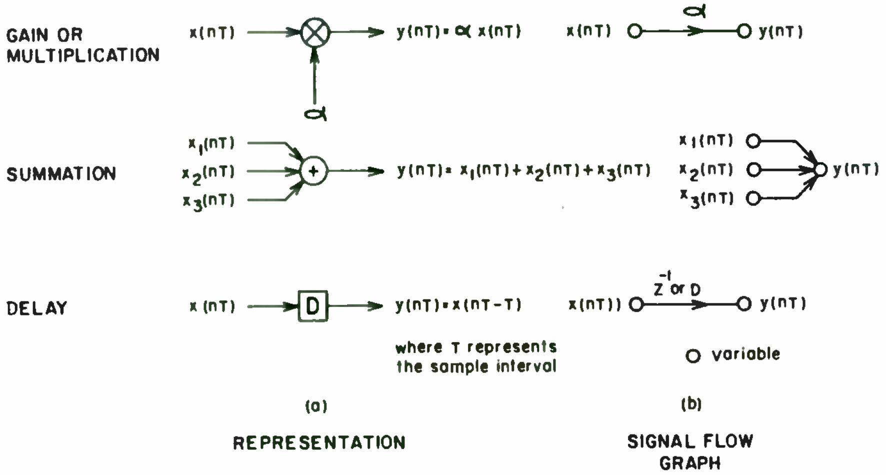


Figure 7. Three Basic DSP Operations

The recursive equation is given as

$$y(nT) = \sum_{i=0}^{N} a_i\, x(nT - iT) - \sum_{j=1}^{M} b_j y(nT - jT),\ \text{where T is the sample interval.}$$

This form is called recursive because the current output, y(nT), is not only a function of present and past inputs x(nT-jT), but also a function of all past outputs, y(nT-iT). The recursive difference equation is used to describe an IIR filter. The reason why shall be shown later in the section on IIR filters.

The non-recursive equation is given as

$$y(nT) = \sum_{i=0}^{N} a_i\, x(nT - iT).$$

For these equations the output y(nT) is defined only in terms of the current and past inputs. Non-recursive difference equations are used to describe FIR filters. More on this is provided in the FIR filter section.

Z Transform

The Laplace transform is used to simplify the design of analog filters. The differential equations that describe the filters may be solved using algebraic techniques after they have been transformed. Likewise, the Z transform allows the use of algebraic techniques for solving difference equations.

The Z transform, X(z), of a series, x(nT) is given as

$$X(z) = \sum_{n=-\infty}^{\infty} x(nT)\, z^{-n}, \quad \text{two sided or bilateral}$$

$$X(z) = \sum_{n=0}^{\infty} x(nT)\, z^{-n}, \quad \text{one sided transform.}$$

For most applications in linear, time invariant digital filters, only the one sided transform will be used. The two sided transform finds use in image and video processing where there may be two, three or more dimensions and functions that are defined for values of n 0.

By way of example, it would be helpful to transform some of the discrete counterparts of analog functions as shown below:

Unit impulse, or unit sample function $\delta(nT)$

$$\delta(nT) = \begin{cases} 1 & n = 0 \\ 0 & n < 0 \end{cases}$$

$$Z\{\delta(nT)\} = \sum_{n=0}^{\infty} \delta(nT)\, z^{-n}$$

$$= \delta(0)z^{-0} + \delta(T)z^{-1} + \delta(2T)z^{-2} + \dots$$

$$= 1 + 0 + 0 + \dots$$

$$Z\{\delta(nT)\} = 1$$

Unit Step **u(nT)**

u(nT)=1, for n 0

$$Z\ u(nT) = \sum_{n=0}^{\infty} u(nT)\, z^{-n}$$

$$= z^{-0} + z^{-1} + z^{-2} + \dots$$

$$Z\ u(nT) = \frac{z}{z-1},\ |z|>1$$

Exponential Ka^n

$$Z\{Ka^n\} = \sum_{n=0}^{\infty} Ka^n z^{-n} = K\sum_{n=0}^{\infty} (a^{-1} z)^{-n}$$

$$= K(a^{-1}z)^0 + K(a^{-1}z)^{-1} + K(a^{-1}z)^{-2} + \ldots$$

$$Z\{Ka^n\} = \frac{Kz}{z-1}, \ |z|>1$$

Delay

$y(nT) = X(nT-T)$

$$Z\{y(nT)\} = Z\{x(nT-T)\} = \sum_{n=0}^{\infty} x(nT-T)\, z^{-n}$$

$$Y(z) = X(-T)z^0 + X(0)z^{-1} + X(1)z^{-2}$$

since $X(nT) = 0$ for $n < 0$, $x(-T) = 0$

then factoring by Z^{-1}, the equation becomes

$$Y(z) = z^{-1} (x(0)z^0 \ x(1)z^{-1} + x(2)z^{-1} + \ldots)$$

$$= z^{-1} X(z)$$

thus z^{-1} can represent a T delay of the function $x(nT)$.

The two difference filter equations will next be transformed.

Recursive

$$y(nT) = \sum_{i=0}^{N} a_i\, x\,(nT - iT) - \sum_{j=1}^{M} b_j\, y(nT - jT)$$

$$Z\{y(nT)\} = Z\left\{\sum_{i=0}^{N} a_i\, x\,(nT - iT) - \sum_{i=1}^{M} b_j\, y(nT - jT)\right\}$$

since these equations are by definition linear, the above becomes

$$Y(z) = \sum_{i=0}^{N} a_i\, Z\{(nT - iT\} - \sum_{j=1}^{M} b_j\, Z\{y(nT - jT)\}$$

which was just solved, thus

$$Y(z) = \sum_{i=0}^{N} a_i\, z^{-i}\, X(z) - \sum_{j=1}^{M} b_j\, z^{-j}\, Y(z)$$

rearranging the terms yields

$$\frac{\sum_{i=0}^{N} a_i\, z^{-i}}{1 + \sum_{j=0}^{M} b_j\, z^{-j}} = \frac{Y(z)}{X(z)} = H(z), \text{ where } b_0 = 1$$

Non-recursive

$$y(nT) = \sum_{i=0}^{N} a_i\, x(nT - iT)$$

$$Z\ y(nT) = \sum_{i=0}^{N} a_i\, x(nT - iT)$$

$$Y(z) = \sum_{i=0}^{N} a_i\, Z\ x(nT - iT)$$

$$Y(z) = \sum_{i=0}^{N} a_i\, z^{-i}\, X(z)$$

$$\sum_{i=0}^{N} a_i\, z^{-i} = \frac{Y(z)}{X(z)} = H(z)$$

IIR Filters

In the previous section it was shown that the general equation for IIR filters was the recursive form. This is because, just like their analog counterparts, discrete filters are described by their response to classic input functions. The most typical of these functions is the impulse function or unit sample function for discrete. The response of a recursive filter, or IIR filter, to the unit sample function is given below.

Let $X(nT) = \delta(nT)$

$$h(nT) = 0, \quad n<0$$

$$h(nT) = \sum_{i=0}^{N} a_i\, x(nT - iT) - \sum_{j=1}^{M} b_j\, h(nT - jT)$$

letting M = N = 1 for a first order equation,

$$h(nT) = a_0\, x(nT) + a_1\, x(nT-T) + (-b_1)\, h(nT-T)$$

$$h(0) = a_0\, x(0) + a_1\, x(-T) + (-b_1)\, h(-T) = a_0$$

$$h(T) = a_0\, x(T) + a_1\, x(0) + (-b_1)\, h(0) = (-b_1)a_0 + a_1$$

$$h(2T) = a_0\, x(2T) + a_1\, x(T) + (-b_1)\, h(T) = a_0(-b_1)^2 + a_1(-b_1)$$

$$\vdots$$

$$h(NT) = a_0\, x(NT) + a_1\, x(NT-T) + (-b_1)\, h(NT-T) = a_0(-b_1)^N + a_1(-b_1)$$

$$h(n) = \left[a_0(-b_1)^n + a_1(-b_1)^{n-1}\right] u(n)$$

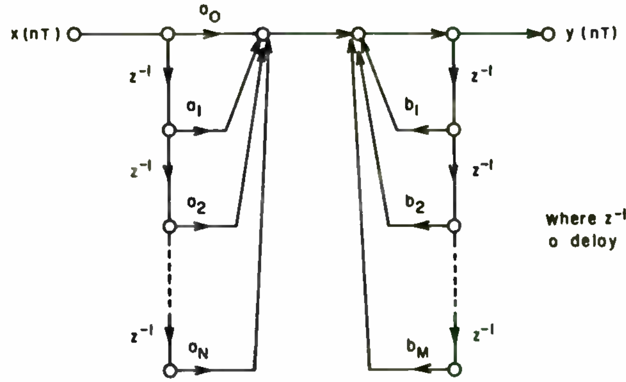
The response h(n) to an excitation x(n) = (n) goes to infinity as n . Thus the recursive difference equation has an infinite impulse response and hence its name of IIR.

A signal flow graph for a general IIR filter is shown in **Figure 8**. This form is known as the direct form 1. Its implementation however, is very inefficient. It can be shown [4] that another equivalent form of the direct form 1 is that shown in **Figure 9**. This equivalent form is known as the direct form 2.

Note that in both **Figures 8** and **9** N=M. For this configuration the order of the filter can be given as either M or N. If N≠M then the order of the filter would be whichever is greater. To obtain maximum efficiency, there should be as many delay stages as the order of the filter. In **Figure 8**, there are M+N delay stages, or 2N. Whereas, in **Figure 9**, the delay stages were reduced to N.

Through various mathematical manipulations, the IIR filter may be configured in many forms. Each form will yield a slightly different flow graph; however, all forms are equivalent and produce the same type of filtering.

$$y(nT) = a_0\, x(nT) + a_1\, x(nT-T) + a_2\, x(nT-2T) + \cdots + a_N\, x(nT-NT) + b_1\, y(nT-T) + b_2\, y(nT-2T) + \cdots + b_M\, y(nT-MT)$$



where z^{-1} represents a delay of T seconds

Figure 8. Direct Form I IIR Filter

$$p(nT) = x(nT) + b_1\, p(nT-T) + \cdots + p(nT-MT)$$

$$y(nT) = a_0\, p(nT) + a_1\, p(nT-T) + \cdots + p(nT-NT)$$

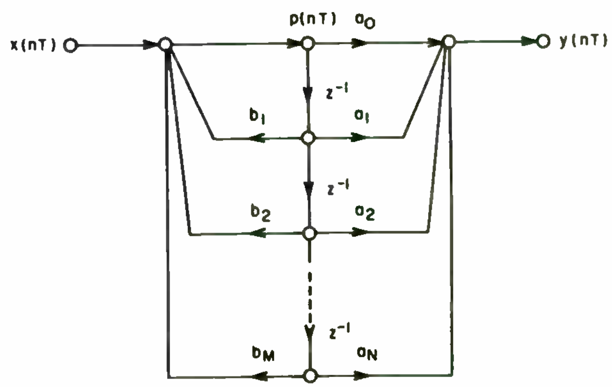


Figure 9. Direct Form II IIR Filter

FIR Filters

FIR filters are represented by non-recursive difference equations. Their name results from the fact that a unit sample input produces a finite impulse response. For instance

$$\text{let } x(nT) = \delta(nT)$$

$$h(nT) = 0 \quad , n < 0$$

$$h(nT) = \sum_{i=0}^{1} a_i\, x(nT - iT), \text{ letting } N=1$$

$$h(nT) = a_0\, x(nT) + a_1\, x(nT-T)$$

$$h(0) = a_0\, x(0) + a_1\, x(-T) = a_0$$

$$h(T) = a_0\, x(T) + a_1\, x(0) = a_1$$

$$h(2T) = a_0\, x(2T) + a_1\, x(T) = 0$$

$$\vdots$$

$$h(NT) = 0$$

Thus, for a first order difference equation, an FIR filter only has an output for two sample periods. The signal flow graph for an FIR filter is illustrated in **Figure 10.**

$$h(n) = \sum_{i=0}^{N-1} a_i\, x(nT - T),$$ the upper value of i is changed to N-1 so that h(n) is defined over N samples

$$y(nT) = a_0\, x(nT) + a_1\, x(nT-T) + \cdots + a_{N-1}\, x(nT-(N-1)T)$$

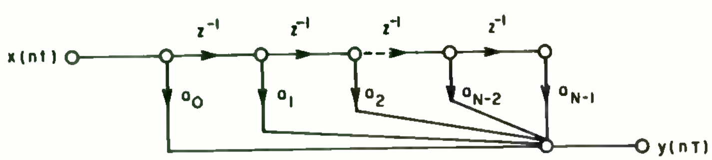


Figure 10. General Form of FIR Filter

As in IIR filters, the order of the filter (N-1) determines the number of delay stages used.

IIR and FIR Designs

The final step before going to hardware is choosing the filter coefficients a_i and b_j. The values of these filter coefficients determine the type of filter; i.e., lowpass, bandpass, highpass.

The actual coefficient calculations are somewhat involved. For some applications, one only need to put the filter specifications into the proper computer program. The details of which will not be presented in this paper, but references will be listed at the end of this paper.

Some other considerations during the hardware implementation are as follows [8]:

1. Choose filter structure.
2. Choose between fixed point and floating point arithmetic (floating point allows greater dynamic range but could have greater inherent noise).
3. Choose between serial and parallel processing.
4. Choose arithmetic devices.

As in many of the previous design structures, the filter structure is basically a trade off of desired performances. For instance, if linear phase is desired, then the FIR filter is one's only choice. If one desires to implement an existing analog filter, and IIR structure could be designed using existing transforms [4] that allow direct calculations of filter coefficients. Other trade offs are listed in **Table 4.**

Table 4. Digital Filter Trade Offs [2]

	IIR	FIR
Coefficient sensitivity	HIGH	LOW
Data word size growth	HIGH	SLIGHT, only in adder
Number of multiplications	LEAST	MOST
Required memory	LEAST	MOST
Ease of design	MODERATE	EASY
Filter control complexity	MODERATE	EASY
Linear phase	NO	YES
Stability	DIFFICULT	UNCONDITIONALLY
Adaptive possible	NO	YES

THE FAST FOURIER TRANSFORM

One of the most common uses of DSP is frequency domain analysis. The wide spread use of this analysis tool has only recently been made available to small processors through advances in power and speed of the processors and by the Fast Fourier Transform, or FFT. The FFT is a very efficient means of calculating the discrete Fourier Transform, or DFT. The DFT is the discrete form of the Fourier Transform. The Fourier Transform is a means of transforming time-based signals into their frequency based components.

Recall that the Fourier transform, C(f), of a continuous time function , x(T), is given as

$$C(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\ ft}\, dt$$

Notice its similarities to the DFT of a finite duration sequence, x(nT), over the interval $0 \quad n \quad N-1$

$$C(kF) = \sum_{n=0}^{N-1} x(nT)\, e^{-j2\pi kFnT}, \qquad k = 0, 1, \ldots, N-1$$

where T is the sample interval and F is the resulting frequency interval.

To simplify notation, kF and nT are replaced by just k and n, resp. and the substitutions are made of $F = 1/NT$ and $W = \exp(-j2\pi/N)$

$$C(k) = \sum_{n=0}^{N-1} x(n)\, W^{nk}$$

It shall be shown later that W is periodic in N. To show this periodicity, W is often written as W_N.

To show how much computation time the FFT is saving, we must first see what it takes to solve the DFT. The direct evaluation of the complex DFT can be expanded as

$$\begin{aligned} C(0) = {} & \mathrm{Re}\{x(0)\} \cdot \mathrm{Re}\{(W_N)\} - \mathrm{Im}\{x(0)\} \cdot \mathrm{Im}\{W_N^{0\ 0}\} \\ & + j[\mathrm{Re}\{x(0)\} \cdot \mathrm{Im}\{W_N^{0\ 0}\} + \mathrm{Im}\{x(0)\} \cdot \mathrm{Re}\{W_N^{0\ 0}\}] \\ & + \ldots\ldots + \\ & + \mathrm{Re}\{x(N-1)\} \cdot \mathrm{Re}\{W_N^{(N-1)0}\} - \mathrm{Im}\{x(N-1)\} \cdot \mathrm{Im}\{W_N^{(N-1)0}\} \\ & + j[\mathrm{Re}\{x(N-1)\} \cdot \mathrm{Im}\{W_N^{(N-1)0}\} + \mathrm{Im}\{x(N-1)\} \cdot \mathrm{Re}\{W_N^{(N-1)0}\}] \end{aligned}$$

.

.

.

.

$$\begin{aligned} C(N-1) = {} & \mathrm{Re}\{x(0)\} \cdot \mathrm{Re}\{W_N^{0(N-1)}\} - \mathrm{Im}\{x(0)\} \cdot \mathrm{IM}\{W_N^{0(N-1)}\} \\ & + j[\mathrm{Re}\{x(0)\} \cdot \mathrm{Im}\{W_N^{0(N-1)}\} + \mathrm{Im}\{x(0)\} \cdot \mathrm{Re}\{W_N^{0(N-1)}\}] \\ & + \ldots\ldots + \\ & + \mathrm{Re}\{x(N-1)\} \cdot \mathrm{Re}\{W_N^{(N-1)(N-1)}\} - \mathrm{Im}\{x(N-1)\} \cdot \mathrm{Im}\{W_N^{(N-1)(N-1)}\} \\ & + j[\mathrm{Re}\{x(N-1)\} \cdot \mathrm{Im}\{W_N^{(N-1)(N-1)}\} + \mathrm{Im}\{x(N-1)\} \cdot \mathrm{Re}\{W_N^{(N-1)(N-1)}\}] \end{aligned}$$

As can be seen, the direct evaluation of the DFT can get rather messy. For each k; there are four real multipliers for each value of n, with N values of n this comes to 4N multipliers. Since there are N values of k, the total number of real multipliers comes to $4N^2$. For each k; there are also 4n-2 real additions. Again since there are N values of k, the total number of additions is N(4N-2). Put another way, the DFT takes N^2 complex multiplications and N(N-1) complex additions to perform. By using the FFT, it will be shown that these calculations can be reduced to $N \log_2(N)$ each for multiplication and addition.

To simplify the notations, for the rest of the FFT discussions it will be assumed that N=8, although it should be easy enough for the reader to expand the discussions to larger values of N. Now with N=8, the time index n can take on value from 0 n 7 and likewise for the frequency index k, 0 k 7. Thus both of these indices can be represented as three bit binary words. Rewriting n, k, and the DFT equations as

$$n = 2^2 n_2 + 2^1 n_1 + 2^0 n_0$$

OR

$$n = 4n_2 + 2n_1 + n_0$$
$$k = 4k_2 + 2k_1 + k_0$$

$$C(k_2, k_1, k_0) = \sum_{n_2=0}^{1} \sum_{n_1=0}^{1} \sum_{n_0=0}^{1} x(n_2, n_1, n_0)\, W^{(4n_2 + 2n_1 + n_0)(4k_2 + 2k_1 + k_0)}$$

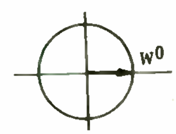
To reduce the DFT equation further, either the time index n or the frequency index k can be factored. In doing so, there is a division or "decimation in time" or a decimation in frequency".

Decimation in Time (DIT)

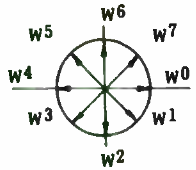
The decimation in time (DIT) algorithm starts by factoring the time index n of the above DFT equation.

$$C(k_2, k_1, k_0) = \sum_{n_2=0}^{1} \sum_{n_1=0}^{1} \sum_{n_0=0}^{1} x(n_2, n_1, n_0)\, W^{4n_2(4k_2 + 2k_1 + k_0)} \quad \ldots$$

It is here where W's periodicity in N begins to be exploited. Remember that $W = \exp(-j2\pi/N) = \exp(-j2\pi/8)$. W^{nk} can be represented by a vector on a unit circle in the s-plane as



Since N=8, there are eight possible vectors



Now, it should be clear how W is periodic in N. For N=8, $W^8 = W^0$, $W^9 = W^1$, etc. Since $W^0 = 1$ and any integer multiple of 8 will yield $W^{8nk} = 1$ thus

$$W^{2n_2(4k_2 + 2k_1 + k_0)} = W^{16n_2k_2} W^{8n_2k_1} W^{4n_2k_0} = W^{4n_2k_0}$$

$$W^{2n_1(4k_2 + 2k_1 + k_0)} = W^{8n_1k_2} W^{2n_1(2k_1 + k_0)} = W^{2n_1(2k_1 + k_0)}$$

so that

$$C(k_2, k_1, k_0) = \sum_{n_2=0}^{1} \sum_{n_1=0}^{1} \sum_{n_0=0}^{1} x(n_2, n_1, n_0) W^{4n_2k_0} W^{2n_1(2k_1+k_0)} W^{n_0(4k_2 + 2k_1 + k_0)}$$

The inner summation can be represented as a separate equation by replacing n_2 by k_0 in $x(n_2, n_1, n_0)$ and summing over n_2

$$X_1(k_0, n_1, n_0) = \sum_{n_2=0}^{1} X(n_2, n_1, n_0) W^{4n_2k_0}$$

The next summation can be similarly rewritten in k_1 and n_1 as

$$X_2(k_0, k_1, n_0) = \sum_{n_1=0}^{1} X_1(k_0, n_1, n_0) W^{2n_1(2k_1 + n_0)}$$

and finally

$$X_3(k_0, k_1, k_2) = \sum_{n_0=0}^{1} X_2(k_0, k_1, n_0) W^{n_0(4k_2 + 2k_1 + k_0)}$$

The DFT can now be expressed as

$$C(k_2, k_1, k_0) = X_3(k_0, k_1, k_2).$$

Note that three sets of equations were developed, X_i, i=1,2,3. This is because N=8. If N were larger then the number equations would be equal to log2(N), N is assumed to be a power of 2. Also note that the ordering of the binary digits (bits) is reversed between the final result, $C(k_2, k_1, K_0)$ and the last equation $X_3(k_0, k_1, k_2)$. This reversed is known as bit reversal mapping and shall be further discussed later.

Expanding the three equations

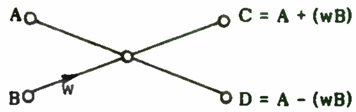
STAGE 1

$$X_1(k_{01}\, n_1, n_0) = \sum_{n_2=0}^{1} X(n_2, n_1, n_0) W^{4n_2k_0}$$

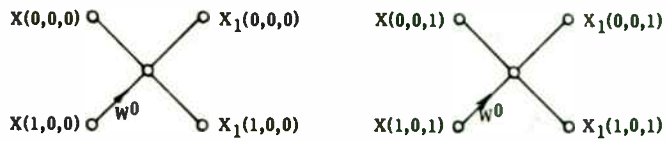
$$X_1(0,0,0) = X(0,0,0) + X(1,0,0)\, W^{4*1*0}$$
$$X_1(0,0,1) = X(0,0,1) + X(1,0,1)\, W^{4*1*0}$$
$$X_1(0,1,0) = X(0,1,0) + X(1,1,0)\, W^{4*1*0}$$
$$X_1(0,1,1) = X(0,1,1) + X(1,1,1)\, W^{4*1*0}$$
$$X_1(1,0,0) = X(0,0,0) + X(1,0,0)\, W^{4*1*1}$$
$$X_1(1,0,1) = X(0,0,1) + X(1,0,1)\, W^{4*1*1}$$
$$X_1(1,1,0) = X(0,1,0) + X(1,1,0)\, W^{4*1*1}$$
$$X_1(1,1,1) = X(0,1,1) + X(1,1,1)\, W^{4*1*1}$$

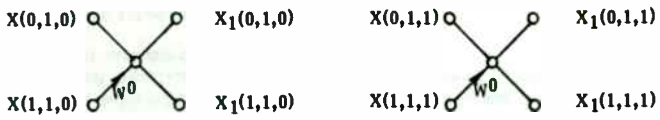
8 complex additions, 8 complex multiplications

In these equations only two powers of W are used, W^0 and W^4. Relating to the periodicity of N, W^4 is nothing more than the negative of W^0, that in $W^4 = W^0$. The above equations can be more easily written in their signal flow graph notation. To remove some of the clutter of the graph a special notation known as the butterfly is used.



As can be seen, the top right node is the sum of the two left nodes and the bottom right node is the difference of the two left nodes. With that in mind, compare there flow graphs with the preceding equations, using the fact that $W^4 = -W^0$





STAGE 2

$$X_2(k_0, k_1, n_0) = \sum_{n_1=0}^{1} X_1(k_0, n_1, n_0)\, W^{2n_1(2k_1, +k_0)}$$

$X_2(0,0,0) = X_1(0,0,0) \quad + \quad X_1(0,1,0)\, W^{2*1(0+0)}$

$X_2(0,0,1) = X_1(0,0,1) \quad + \quad X_1(0,1,1)\, W^{2*1(0+0)}$

$X_2(0,1,0) = X_1(0,0,0) \quad + \quad X_1(0,1,0)\, W^{2*1(2+0)}$

$X_2(0,1,1) = X_1(0,0,1) \quad + \quad X_1(0,1,1)\, W^{2*1(2+0)}$

$X_2(1,0,0) = X_1(1,0,0) \quad + \quad X_1(1,1,0)\, W^{2*1(0+1)}$

$X_2(1,0,1) = X_1(1,0,1) \quad + \quad X_1(1,1,1)\, W^{2*1(0+1)}$

$X_2(1,1,0) = X_1(1,0,0) \quad + \quad X_1(1,1,0)\, W^{2*1(2+1)}$

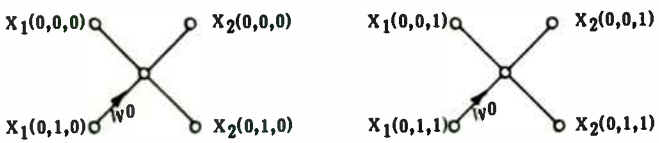
$X_2(1,1,1) = X_1(1,0,1) \quad + \quad X_1(1,1,1)\, W^{2*1(2+1)}$

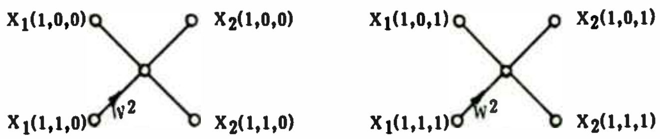
8 complex additions, 8 complex multiplications

powers of W: W^0, W^2, W^4, W^6

equalities : $W^4 = -W^0,\ W^6 = -W^4$

Signal Flow Graphs





STAGE 3

$$X_3(k_0, k_1, k_2) = \sum_{n_1=0}^{1} X_2(k_0, k_1, n_0)\, W^{n_0(4k_2 + 2k_1 + k_0)}$$

$X_3(0,0,0) = X2(0,0,0) \quad + \quad X_2(0,0,1)\, W^{(0+0+0)}$

$X_3(0,0,1) = X2(0,0,0) \quad + \quad X_2(0,0,1)\, W^{(4+0+0)}$

$X_3(0,1,0) = X2(0,1,0) \quad + \quad X_2(0,1,1)\, W^{(0+2+0)}$

$X_3(0,1,1) = X2(0,1,0) \quad + \quad X_2(0,1,1)\, W^{(4+2+0)}$

$X_3(1,0,0) = X2(1,0,0) \quad + \quad X_2(1,0,1)\, W^{(0+0+1)}$

$X_3(1,0,1) = X2(1,0,0) \quad + \quad X_2(1,0,1)\, W^{(4+0+1)}$

$X_3(1,1,0) = X2(1,1,0) \quad + \quad X_2(1,1,1)\, W^{(0+2+1)}$

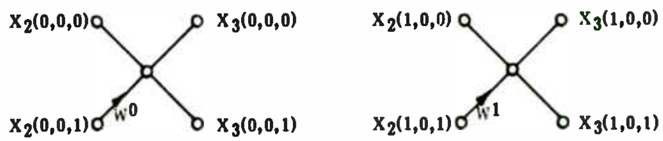
$X_3(1,1,1) = X2(1,1,0) \quad + \quad X_2(1,1,1)\, W^{(4+2+1)}$

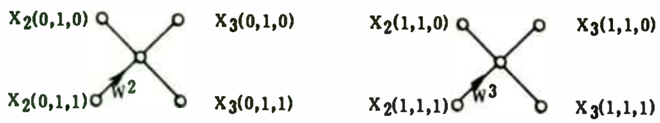
8 complex additions, 8 complex multiplications

powers of W: $W^0, W^1, W^2, W^3, W^4, W^5, W^6, W^7$

equalities : $W^7 = -W^3,\ W^6 = -W^2,\ W^5 = -W^1,\ W^4 = -W^0$

Signal Flow Graphs





Advantages of DIT Algorithm

One may note, that for each stage of the FFT the results could be placed back into the same location that the input data came from. For example $X_{i+1}(0,0,0)$ could be put into $X_i(0,0,0)$. This is what is known as an in place calculation. The advantage of an in place algorithm is that, except for a few memory locations to store intermediate result, only enough memory is needed to hold the data. As opposed to other algorithms that need twice as much memory as data, one set of memory for data and an equivalent set for results.

By using the DIT algorithm, the calculations to perform the DFT are greatly reduced. For each stage it only took 8 complex multiplications and 8 complex additions. There were three stages so only 24 operations of each were needed. For the original DFT 64 operations were required. In general the operations in the FFT is represented by $N \log 2(N)$, whereas the DFT is N^2. **Table 5** shows some representative tuning for some of the more common values of N.

Table 5. DFT vs. FFT Timing

N	N^2	$N \log_2 N$
32	1024	160
64	4096	384
512	262144	4608
1024	1048576	10240
2048	4194304	22528

Bit Reversal Mapping

One disadvantage of the DIT algorithm is that the data arrays get shuffled during calculation operations. Even though the calculations are performed in place, the order of the data in the array is shuffled. How the data is shuffled is known, however, so it is easy to put the arrays back in order.

As seen earlier, the output frequency array, C(k), was equal to the last stage output array, $X_3(k_0, k_1, k_2)$ by the mapping equation of:

$$C(k) = C(k_2, k_1, k_0) = X_3(k_0, k_1, k_2), \text{ where } k = 4k_2 + 2k_1 + k_0$$

The only difference in the indices is that ordering of the k bits has been reversed. Thus, to restore the data ordering of the array, reverse the ordering of the bits and map the data back into its proper location in the frequency array C(k).

$$X(n) = X(4n_2 + 2n_1 + n_0) \Rightarrow X(n_2, n_1, n_0) \rightarrow X_3(k_0, k_1, k_2)$$
$$\Rightarrow C(k_2, k_1, k_0) = C(4k_2 + 2k_1 + k_0) = C(k)$$

$$X(0) \Rightarrow X(0,0,0) \rightarrow X_3(0,0,0) \Rightarrow C(0,0,0) = C(0)$$
$$X(1) \Rightarrow X(0,0,1) \rightarrow X_3(0,0,1) \Rightarrow C(1,0,0) = C(4)$$
$$X(2) \Rightarrow X(0,1,0) \rightarrow X_3(0,1,0) \Rightarrow C(0,1,0) = C(2)$$
$$X(3) \Rightarrow X(0,1,1) \rightarrow X_3(0,1,1) \Rightarrow C(1,1,0) = C(6)$$
$$X(4) \Rightarrow X(1,0,0) \rightarrow X_3(1,0,0) \Rightarrow C(0,0,1) = C(1)$$
$$X(5) \Rightarrow X(1,0,1) \rightarrow X_3(1,0,1) \Rightarrow C(1,0,1) = C(5)$$
$$X(6) \Rightarrow X(1,1,0) \rightarrow X_3(1,1,0) \Rightarrow C(0,1,1) = C(3)$$
$$X(7) \Rightarrow X(1,1,1) \rightarrow X_3(1,1,1) \Rightarrow C(1,1,1) = C(7)$$

The above sets of equations are trying to show that X(n) is mapped, ($\Rightarrow$), into $X(n_2, n_1, n_0)$. The results of the calculation from X end up in $X_3(k_0, k_1, k_2)$. Then X_3 is mapped back into $C(k_2, k_1, k_0)$ which yields C(k). Because of the symmetry of the DIT algorithm, it does not matter if the bit reversed mapping is done before to the time input array so that the frequency comes out in order, or after to reorder the frequency array. **Figure** 11 shows the entire flow diagram for the DIT FFT where the input array has been bit reverse mapped, and the output C(k) is in order.

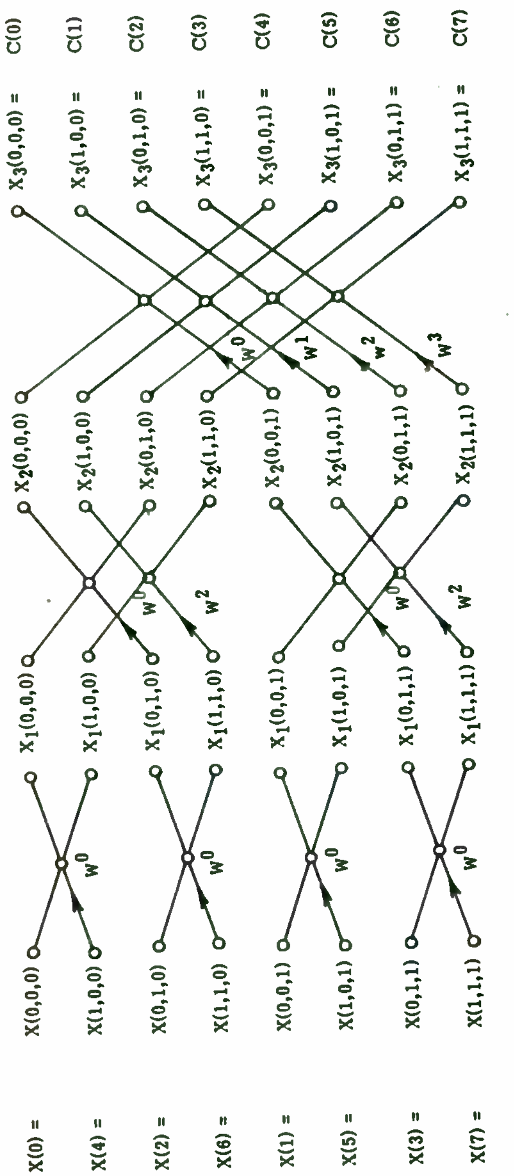


Figure 11. Decimation in Time, FFT

Decimation in Frequency (DIF)

The decimation in frequency (DIF) algorithm follows exactly as the DIT algorithm, except that the frequency indices (and hence its name) are factored. Starting with the original DFT equation

$$C(k) = C(k_2, k_1, k_0) = \sum_{n_2=0}^{1} \sum_{n_1=0}^{1} \sum_{n_0=0}^{1} X(n_2, n_1, n_0)\, W^{(4n_2 + 2n_1 + n_0)} * W^{(4k_2 + 2k_1 + k_0)}$$

Reordering the equations to get the desired factor:

$$C(k_2, k_1, k_0) = \sum_{n_0=0}^{1} \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} X(n_2, n_1, n_0)\, W^{k_0(4n_2 + 2n_1 + n_0)}\, W^{2k_1(2n_1 + n_0)} * W^{4k_2 n_0}$$

so that

$$X_1(k_0, n_1, n_0) = \sum_{n_2=0}^{1} X(n_2, n_1, n_0)\, W^{k_0(4n_2 + 2n_1 + n_0)}$$

$$X_2(k_0, k_1, n_0) = \sum_{n_1=0}^{1} X_1(k_0, n_1, n_0)\, W^{2k_1(2n_1 + n_0)}$$

$$X_3(k_0, k_1, k_2) = \sum_{n_0=0}^{1} X_2(k_0, k_1, n_0)\, W^{4k_2 N_0}$$

$$C(k_2, k_1, k_0) = X_3(k_0, k_1, k_2)$$

STAGE 1

$$X_1(0,0,0) = X(0,0,0) + X(1,0,0)$$

$$X_1(0,0,1) = X(0,0,1) + X(1,0,1)$$

$$X_1(0,1,0) = X(0,1,0) + X(1,1,0)$$

$$X_1(0,1,1) = X(0,1,1) + X(1,1,1)$$

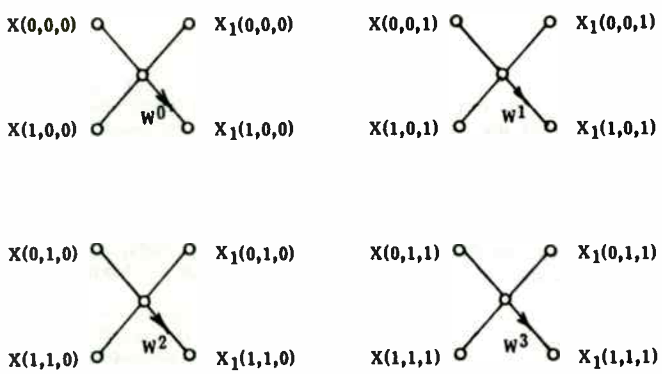
$$X_1(1,0,0) = X(0,0,0)W^{1(0+0+0)} + X(1,0,0)W^{1(4+0+0)} = [X(0,0,0) - X(1,0,0)]\, W^0$$

$$X_1(1,0,1) = X(0,0,1)W^{1(0+0+1)} + X(1,0,1)W^{1(4+0+1)} = [X(0,0,1) - X(1,0,1)]\, W^1$$

$$X_1(1,1,0) = X(0,1,0)W^{1(0+2+0)} + X(1,1,0)W^{1(4+2+0)} = [X(0,1,0) - X(1,1,0)]\, W^2$$

$$X_1(1,1,1) = X(0,1,1)W^{1(0+2+1)} + X(1,1,1)W^{1(4+2+1)} = [X(0,1,1) - X(1,1,1)]\, W^3$$

Signal Flow Graph



By following through the patterns developed in the DIT section, comparing them to the stage 1 DIF development, one can deduce the complete signal flow graph for the DIF algorithm as in Figure 12.

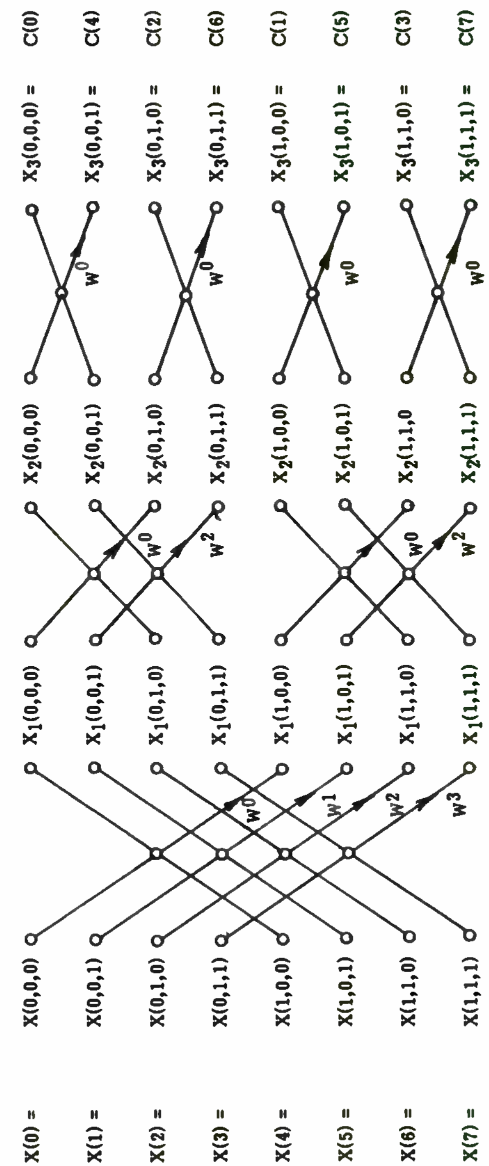


Figure 12. Decimation in Frequency, FFT

DIT vs. DIF Algorithms

Mathematically, there is very little difference between the DIT and the DIF algorithm; both algorithms require the same $N\log_2 N$ operations to complete, both are done in place, and both require bit reversal mapping to either the input or output. The advantage of the DIF algorithm is that its complex multiplication is done after the complex addition. This could save time in the hardware implementation because the sum result can be stored away while the processor is waiting for the multiplication to be finished. During this same time, the next pairs of data can be retrieved to start the next operation.

Other Algorithms

There are many other algorithms that can be used to compute the FFT. They are all basically modifications of the same theme, reducing an N point DFT into smaller and smaller sequences. The DIT and DIF do this by successively halving the N point DFT. For this reason, they are known as radix 2 algorithms. Other algorithms are based on a radix 4, where the N point DFT is reduced by 1/4 for each pass. Other algorithms work when N is any integer at all. These are referred to as prime factor algorithms. Still other algorithms use double memory by storing the output of each pass in a different location than the input, thus saving memory access time.

In the final analysis, the choice of algorithm comes down to the hardware design. How much memory can be made available? How complicated does one want or is able to make the control structure? How fast or slow is the support hardware; i.e., memories, multipliers, accumulators?

Windows For Use With FFT's

When using the Fourier Transform, the input signal is defined over time from $-\infty$ to $+\infty$ and it is assumed that the signal is continuous and periodic. When using the Discrete Fourier Transform, the input signal is also assumed to be continuous and periodic; however, the input signal cannot be defined over infinity and must in fact be defined over a time interval NT. The sampling over this interval can be thought of as the application of a rectangular window to the continuous data as shown in Figure 14.

The rectangular window tends to corrupt the data, especially if NT is not some integer multiple of the input signal. If one recalls, after sampling, the input signal is assumed to be periodic, so the rectangular window can be placed next to itself over time. When this happens, sharp discontinuities result as shown in Figure 15. These discontinuities produce sidelobes of the main spectra in the frequency domain, referred to as spectral leakage.

Another way of looking at this phenomenon is to think of sampling as multiplying the continuous time signal by a rectangular pulse to produce the finite duration sequence for processing. As is well known, multiplication in the time domain is the same as convolation in the frequency domain. This process is illustrated in Figure 16.

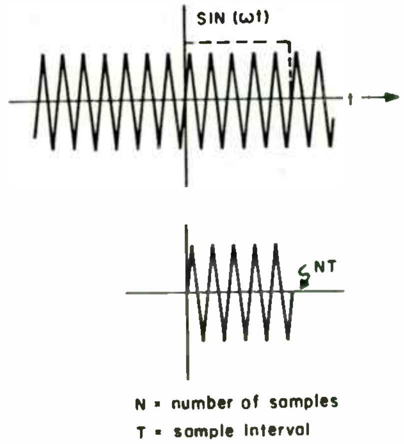


Figure 13. Windowing Effect of Sampling

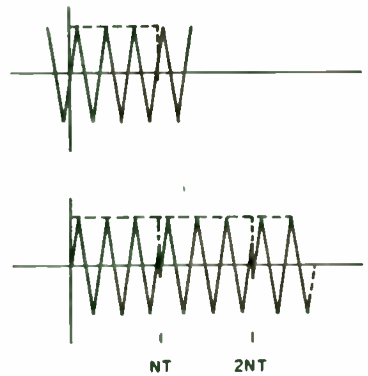


Figure 14. NT As Non-Integer Multiple of Signal

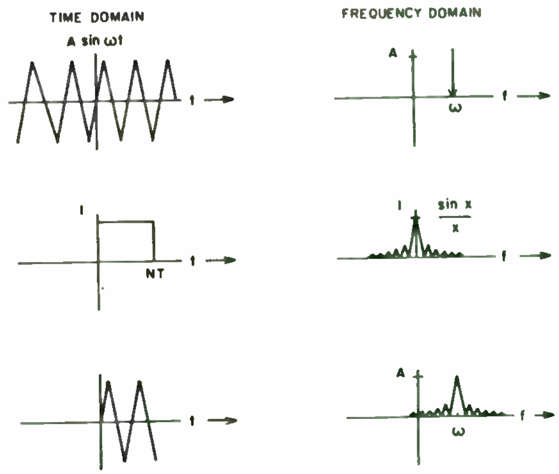


Figure 15. Spectra Distortion Caused By Rectangular Window

To correct for this spectral leakage, a window weighting function is first applied to the timing data. Many useful window functions exist. Harris in [9] presents a good overall tutorial of many of these functions. Only three window functions will be examined here:

1. Triangle
2. Hamming
3. Blackman Harris

A sample FFT done with each of these windows appears at the end of this section.

Triangle

The triangle window is by far the easiest to calculate on any processor. The function is given as:

$$w_T(n) \begin{cases} = \dfrac{n}{N/2} & 0 \leq n < N/2 \\[2ex] = \dfrac{N-n}{N/2} & N/2 \leq n < N-1 \end{cases}$$

The spectral improvement, however, over the rectangle is slight and yields only an approximate 10 dB drop in the sidelobes. Also, because the input signal now contains only half of its original energy, there is a 6 dB loss in spectral magnitude.

Hamming

The hamming window function is described by the equation

$$W_H(n) = 25/46 - (1-25/46) \times \cos(2\ \ n/N),$$

$$n = 0, 1, 2, \ldots N-1$$

This window has the advantage of attenuating sidelobes by 40 dB while keeping losses at 5.3 dB.

Blackman-Harris, Three Term

The Blackman-Harris window is defined as

$$W(n) = 0.42323 - 0.49755 \cos(2\ \ n/N) + 0.07922 \cos(4\ \ n/N)$$

This function attenuates sidelobes by 67 dB with a window loss of 7.5 dB.

Final Notes On The FFT

Each sample, n, is often referred to as a bin. The magnitude of the complex output X(n) represents the spectral amplitude of the components of that bin. If a frequency component falls between bins, its energy is smeared over many bins. Again, this is known as spectral leakage.

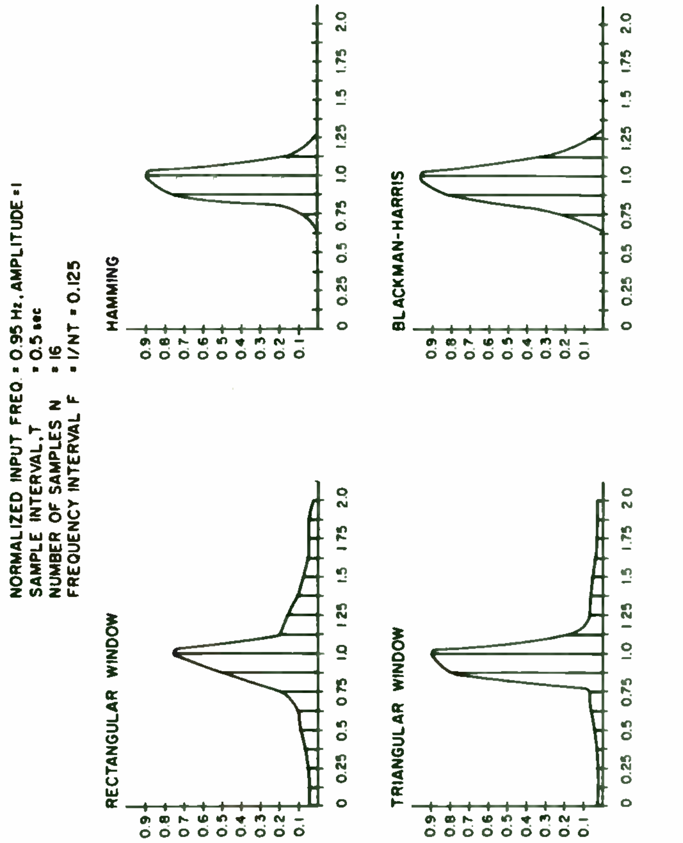


Figure 16. Window Function Example

The frequency step is given as F = 1/(NT) = fs/N where fs is the sample frequency. The frequency component for the bins is simply

BIN # x F, i.e., for n=0, 1, 2, . . . 7

X(0) , dc component
X(1) , frequency component at F
X(2) , frequency component at 2F
.
.
.
X(7) , frequency component at 7F

Also note that when applying windows, a frequency smearing occurs. Without a window, frequency components need only be separated by a single bin to be recognized as separate frequencies. With a window function applied, this minimum bin distance widens. For the windows described, the minimum bin spacing is as follows:

Triangle	=	1.33
Hamming	=	1.36
Blackman-Harris	=	1.7

WHY USE DSP

The number one force driving the push into DSP is cost. Using FFT's, signal analysis can now be performed at a substantially lower cost than encountered when using swept banks of analog filters. Once digitized, the power of analysis begins. Algorithms are being developed that can determine the type of modulation being used, phase detection, direction finding and digital demodulation.

The number one force preventing the acceptance of DSP is also cost. The cost of a couple of op-amps, some resistors and capacitors for a simple two pole active analog filter is only a few dollars, compared to a few hundred dollars for digital multipliers, memories and support processing for a digital filter. Of course, if the processor can be time-shared among many modules, the cost factor depreciates. For one two-pole filter, there is a big difference in cost; for a hundred two-pole filters, the analog version increases one hundred fold, whereas the digital increase is negligable. The digital version uses the same hardware plus a few other components to muliplex the data and processor.

Another consideration in DSP is the stability of digital parts. Once coefficients are computed for a filter, they will not drift over time nor change with temperature. This allows sharp cutoffs in filters to be realized using digital processing. Furthermore, provided the hardware is set up properly, a simple change in software will turn an FFT unit into a bandpass IIR filter. Another change will result in a lowpass filter. Thus digital components are very versatile.

Perhaps the greatest advantage of DSP is performance. A digital signal monitor can achieve a much finer resolution at a faster scan rate than analog version. The cutoff frequency of filters can be more precisely tuned and show a steeper roll-off than analog filters.

Finally, many communications are being done in a digital format; for example, T1 standards of the Bell system. With this digital format, analog signal processing becomes the expensive option. As the trend towards digital circuitry increases, most types of signal processing will be made simpler and more powerful via DSP.

CONCLUSION

Digital signal processing is not an easy topic to comprehend. It combines theories of both the analog and the digital disciplines. Many times it requires several readings of the same concept, presented by several authors, before the concept can take hold. To ease this understanding, one should remember that all aspects of DSP can be divided into three areas:

1) Input Signal
2) Digital Processor
3) Information Extraction

The input signal must be sampled and quantized before it can be fed to the digital processor. The digital processor can only perform the operations of addition, multiplication and delay (storage) on the signal data. The end result of the digital processor is to extract information from the signal or modify it in some way.

REFERENCES

GENERAL

1 L. W. Gardenhire, Selecting Sampling Rates, Instrumentation Society of America, April 1964.

2 Willard Bucklen, Dr. John Eldon, Louis Schirm, and Fred Williams, "Designers Guide to DSP," EDN, Magazine.

Part 1, "Understanding the Fundamentals Eases Signal Processing Task," March 18, 1981.

Part 2, "Single Chip Digital Multipliers Form Basic DSP Building Blocks," April 1, 1981.

Part 3, "Digital Processing Facilitates Signal Analysis," April 15, 1981.

Part 4, "DAC's and Multifunction Chips Enhance Your DSP Options," April 29, 1981.

Part 5, "Digital Correlators Illustrate DSP Trends," May 13, 1981.

3 L. R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Prentice Hall, 1975

4 A. V. Oppenheim and R. Shafer, Digital Signal Processing, Prentice Hall, 1975

5 S. K. Tewksbury, R. B. Kieburtz, J. S. Thompson, and S.P. Verma, "Tutorials on Signal Processing for Communications, Part II, Digital Signal Processing Architecture," IEEE Communications Society Magazine, January 1978.

6 Electronic Design, Magazine, Hayden Publishers, Vol. 32, No. 10, May 17, 1984, entire issue.

7 Fred Mintzer, "Parallel and Cascade Microprocessor Implementations for DSP," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 5, October 1981.

8 Andreas Antoniou, Digital Filters: Analysis and Design, McGraw-Hill, 1979.

9 Fredric J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform:, Proceedings of the IEEE, Vol. 66, No. 1, January 1978.

10 Robert Ramirez, "The Fast Fourier Transform's Errors are Predictable Therefore Manageable," Electronics Magazine, McGraw-Hill, June 13, 1974.

DIGITAL FILTER PROGRAMS

a) See references 3 and 8 .

b) "Programs for Digital Signal Processing," New York: IEEE Press, 1979.

HARDWARE FOR DSP

See references 3, 6, and 8

ACSB - AN OVERVIEW OF AMPLITUDE COMPANDORED SIDEBAND TECHNOLOGY

James Eagleson
IDX, Inc.

January 1985

Early voice communications by radio was accomplished by Amplitude Modulating the Carrier Wave. This was the simplest and most obvious way to put audio information on a radio wave.

Later, Frequency Modulation was proposed as a potentially superior method of audio transmission and Armstrong and others developed practical FM systems.

In the 1940's and even into the 50's one could find AM communications throughout the radio spectrum. The LA Police Department, for example, could be monitored just above the AM Broadcast Band and World War II "Walkie-Talkies" were on a variety of Medium Wave, Short Wave, and even low VHF frequencies...predominately using AM.

About this time there was the opening of the Low and High VHF frequencies for Land Mobile. After theoretical and actual studies of various modes available, Buesing(1) and others determined that FM was the most effective mode for the VHF mobile environment. For a given carrier power, FM had better signal-to-noise performance, mobile flutter control, and provided hands-off operation...an important consideration in that marketplace.

FM VERSUS AM VERSUS SSB

Using the technology of that day there was an obvious advantage of FM over either SSB or AM. For example, AM required one-and-one-half times the power of FM due to the need for a modulator. SSB, on the ohter hand, required stabilities we are just now able to achieve at these frequencies. The capture effect of FM provided superior quieting performance... an AM transmitter would need to be **many** times more powerful for equal performance at moderate signal levels.

Sideband, of course, had a **6-9 dB advantage** over either AM or FM at **very weak signal levels**, but lost any advantage once the signal reached FM's SNR enhancement threshold. The predominate means of achieving SSB's advantage at low SNR is the **elimination of the carrier** present in both FM and AM coupled with **elimination of one sideband** which allows a much narrower receiver bandwidth and, thus, noise bandwidth limiting.

Unfortunately, SSB requires very tight frequency stability, a **slow** AGC making flutter control at VHF impossible, and was much less convenient to generate with the technology of the 40's and 50's.

THE BIG CRUNCH

What few envisioned was that VHF and UHF Land Mobile would become so popular, even **vital** to American Businesses. It did not take very long to completely fill all of the 50 KHz channels on VHF requiring opening of UHF frequencies and splitting both VHF and UHF channels into 25, 20, even 15 KHz spacings. In fact, in much of Europe the channels are now 12.5 KHz spaced!

SOME PROPOSED SOLUTIONS

There has come a recognition in the Land Mobile industry that we have a real problem here.

Many solutions have been proposed including Spread Spectrum, Cellular and Trunking radio systems, even Single-Sideband FM...a technique requiring half the spectrum for performance equivalent to normal FM. Each of these has its good points and bad points.

A more recent proposal was to take a second look at SSB on VHF. The result: **Amplitude Compandored Sideband (ACSB).**

The term ACSB was first used in a paper by Dr. Bruce Lusignan in a report prepared for the FCC's UHF TASK FORCE study of **Spectrum Efficient Technologies.(2)**

In essence, Dr. Lusignan, Dr. Fred Cleveland and others prepared SSB equipment to study what, if anything, could be done to make Single Sideband, well known for its spectrum efficiency, compatible with the needs of VHF and UHF Land Mobile users. While SSB can be placed at 5 KHz channel spacings (three to five times as efficient as FM), the elimination of mobile flutter, improvement of signal-to-noise ratio, hands-free operation, squelch, and tone signalling capabilities of FM are all features that Land Mobile users have come to expect. Providing equivalent performance in these areas using SSB is much more difficult!

DIFFICULT BUT NOT IMPOSSIBLE

The Lusignan study concluded that all of the desired characteristics of FM for Land Mobile service could be achieved using ACSB...a much processed form of SSB. To be sure, these conclusions did not go unchallenged by the competition and were published in IEEE Spectrum with an unprecedented "disclaimer" by the FCC TASK FORCE. On the other hand, of the numerous **technical** objections raised, an equal number of **suggested answers** were brought forth by various members of the communications industry(3).

There is a long way to go before ACSB radios will achieve the same level of sophistication that 30 years of FM development has brought to that mode. Then again, current ACSB radios outperform FM for several applications and certainly outperforms the FM equipment of 20-30 years ago. While it is not the purpose of this paper to argue these issues (which seem to follow **any** new technology), I personally believe that given no **political** roadblocks, ACSB will prove a valueable resource to meet many of the communications needs of the industry and the only practical technology for **some** of those needs.

WHAT IS AMPLITUDE COMPANDORED SIDEBAND?

In actuality, this is a difficult question to answer. ACSB is developing in many forms to meet the needs of many kinds of communications networks.

In essence, ACSB has four or five key elements:

1) **Amplitude Compandoring**
2) **Spectrum Equalization**
3) **Transmitted Level Reference**
4) **Transmitted Frequency Reference**
5) **Positive Acting Squelch**

In most systems the last three of these are built up around some form of **Pilot Tone** or **Carrier**. Two forms of Pilot Tone are used: **Above Band** and **In-Band (TAB** and **TIB)**.

Let's take a look at each element in turn.

AMPLITUDE COMPANDORING

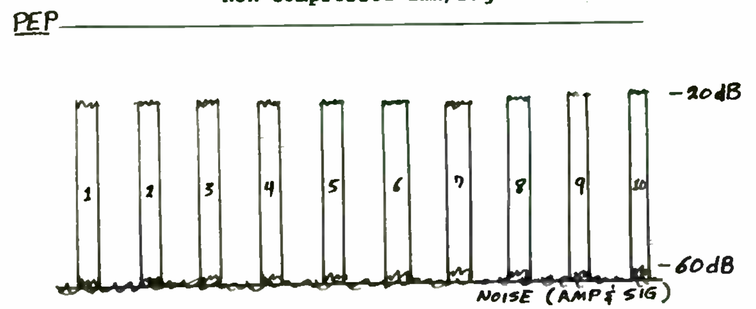
To those working in the telephone industry, amplitude compandoring is not new at all. Most long-line transmissions have been using this technique for some years now.

In that linear amplifiers have a limited peak power capability, ultilization of a "common carrier" for multiple signals is limited by the power level and number of those signals.

In order to add more signals to the system, the level of each signal must be reduced accordingly. At some point this reduction will degrade the signal-to-noise ratio of each of the signals to an unacceptable level.

For example, given a line amplifier with an output capability of 100 mW and ten signals of 1 mW each. The Peak Envelope Power (PEP) of those ten signals would be N^2 or 10^2 = 100mW. Thus we can only use ten signals with this amplifier regardless of the bandwidth available or the number of signals desired. (I'm making a simplified case here in that various other factors change the formula in the real world).

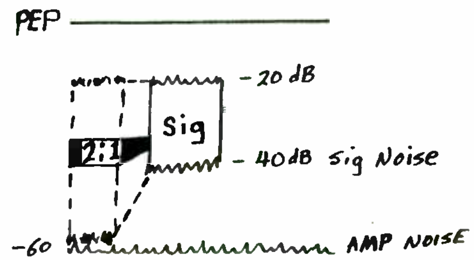
Non-Compressed 1mW/Signals



Assuming that the amplifier has a 60 dB Dynamic Range, that is 60 dB PEP output level to amplifier noise output level, each signal would be 20 dB down from the PEP limits of the amplifier which is 40 dB up from the amplifier's noise floor. Adding any additional signals would begin to degrade the desired 40 dB signal-to-noise ratio for each signal because we would need to lower the level of each signal to maintain our PEP output below the maximum level.

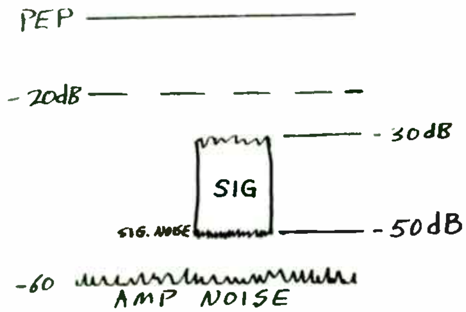
However, if we were to compress each signal at 2:1 input to output ratio, the dynamics of each signal would then be reduced to only 20dB. Given that each signal maintained the same peak output (1mW), our PEP for the system would remain the same (100 mW) but the lowest level portions of each signal would clear the amplifier's noise floor by about 20 dB.

Compressed 1mW/Signal



If we then **reduce** each signal's level by **20 dB**, we will still maintain all desired audio above the amplifier's noise floor but will be able to add **10 times as many signals** to the system (100 signals) while keeping our 100 mW PEP power limit. Now the bandwidth becomes the limiting factor, not the PEP capability of the amplifier. In actuality, a reduction of about than 10 dB is typical yielding a 3 times expansion in the number of signals (to 30 instead of 10) for the same PEP output.

Compressed 0.1 mW/Signal



In reality, of course, several other factors enter into the formula when loading becomes this high. Intermodulation and other things become the limiting factors in addition to amplifier dynamic range.

EXPANSION

We have one other problem remaining in our simplified example.

Compression is fine, but...

We left our signal with all audio down to 40 dB or so compressed into 20 dB of dynamics. This is fine in a broadcast studio or a quiet location but anywhere else we will have objectionable background noise... especially in a mobile environment. These noises and reverberations are annoying, even distracting.

This is clearly unacceptable in the commercial communications marketplace though it might be beneficial on an already noisey Shortwave communications circuit.

We have another option, however. If we set up an amplifier on the receiving end of the circuit with the exact opposite transfer function, that is 2 dB of output for every 1 dB of input change, we will then restore the 20 dB dynamics of the incoming signal to its original 40 dB range. In other words, when the signal is at 0 reference coming in, it will be at 0 reference going out. As it drops to -20 dB coming in, it will drop to -40 dB going out of the **Expansion Amplifier**.

So now we have achieved our goal of increasing the number of signals put on a given circuit while maintaining our desired signal-to-noise dynamic range. Fidelity depends largely on the accuracy with which the compressor and expandor amplifiers do their job. That this should be acceptable for communications circuits can be seen by its use in the more critical **Dolby** and **dBX** Hi Fidelity recording systems. It is also used on almost all telephone long-line circuits.(4)

ACSB uses the technique for a slightly different purpose, however. Here we are trying to improve a **poor** signal-to-noise mobile radio circuit rather than trying to maintain an already fairly good signal-to-noise ratio (as is the case in Telco and Hi Fi systems).

Therefore we come to utilization of another technique which also is not new but is used in a slightly different manner... **Spectrum Equalization.**

SPECTRUM EQUALIZATION

Early in the development of FM it was discovered that due to the wide receiver bandwidth required and the nature of the detector used, high frequency baseband noise was quite pronounced on weaker signals. In fact, most FM squelch circuits look for the reduction of this noise with the increasing strength of a signal as their reference point in determining whether to let audio go to the speaker.

It was also noted that both human speech and music tend to concentrate most energy in the lower portion of the audio range. The energy above 1000 Hz falls off rapidly compared to the audio in the 100-700 Hz region (5).

It was apparent, then, that one could **boost** the high frequency response of one's modulation without causing overmodulation since levels at these frequencies would not be nearly as high as they are at lower frequencies. That, in turn, meant that you could **reduce** the high frequency response of the **receiver's** audio effectively limiting the receiver and detector's high frequency noise content and thus improve the overall system signal-to-noise ratio by several dB.

We, of course, call this process **"pre-emphasis"** and **"de-emphasis"** and it is used in all communications and broadcast FM circuits.

For reasons that are not clear, this technique was never appplied to SSB. Perhaps it was because SSB has such a narrow bandwidth so that the noise generated in the receiver is much "softer" sounding than the wideband "white noise" coming out of a non de-emphasized FM receiver. Additionally, the 1.5-3 dB advantage offered by de-emphasis in a 3 KHz SSB circuit is hardly as dramatic an improvement as the 10-12 dB noise bandwidth reduction it has in a 15-20 KHz circuit.

In any case, it is not generally used with normal SSB.

On the other hand, when we are talking about VHF SSB using 4:1 amplitude compandoring (in the case of ACSB), this 1.5-3 dB SNR improvement becomes much more significant!

Expanded four times this improvement becomes 6-12 dB on moderate level signals and also reduces susceptibility to impulse noise considerably (6).

It also helps remove any above channel interference from the audio channel (assuming upper sideband transmission) and reduces the "puffy" sounding noise associated with expansion of signals at poor signal-to-noise ratios (7).

PILOT TONE OR CARRIER

While it is easy to see how amplitude compandoring helps on a telephone line where signals are all about the same level and circuit gain remains nearly constant, the VHF mobile environment is much more severe. We find widely different signal levels and the desired signal level also fluctuates wildly.

Common SSB AGC circuits cannot handle these fluctuations.(8)

The solution is to provide a transmitted reference level with which the receiver can determine the proper audio levels independant of the path loss variations due to multipath cancellations and enhancements.

One can transmit full carrier with the sideband signal, but this is wasteful of energy. Transmitting a **reduced carrier** (say **-10 db**) is less wasteful but is not the ideal solution since it must be **processed at IF frequencies**...not generally considered the most economical approach.

This approach does allow totally synchronous carrier "re-insertion" by various means, however. (9)

Tone above-band as used by Dr. Lusignan and others places a 3.1 KHz tone just above the audio range (which is Low Pass Filtered at 2.5 KHz or so). This reference tone is accurately generated by frequency dividing a crystal oscillator in the 3 MHz region.

The tone can be processed at baseband frequencies for both AGC purposes (3 KHz is high enough to provide a good fade rate to carrier frequency ratio), and as part of a phase-locked tuning loop. Tuning accuracies of a few Hz are possible.

Tone in-band and **Transparent Tone in-band** have also been explored as alternatives to tone above-band. This is to get rid of the problem of working on the edge of the receiver's passband with a phase locked loop. The edge of most crystal filters has **significant phase shift**, of course, and there is also the problem of having the pilot **drift outside the passband** of the filter making signal lock up difficult.

Additionally, it has been found that at UHF frequencies in particular (and less so at VHF), the correlation between fades at the bottom edge of the passband and the top edge is much less than the correlation of fades at the middle of the passband relative to either edge. On HF this is called "selective fading". **Tone in-band** reduces the de-correlation.

Transparent Tone in-band takes another step. The audio above the chosen Pilot Tone is selected out and mixed to shift it up about 300 Hz in the spectrum so that the Pilot can then be placed in the gap that is left. This prevents placing a "hole" in the middle of the audio spectrum when the pilot tone is filtered out upon reception such as occurs with normal tone in-band techniques (10).

The shifted audio is then re-shifted to its original spectrum to provide normal audio.

Another advantage of Tone in-band (or TIB) techniques is that the tone selected can be the lower of the two tones used for normal data communications. Thus our tone can serve a dual purpose.

AUTOMATIC GAIN CONTROL

Whether carrier, tone above-band, or tone in-band, the Pilot Tone provides a fixed, **constant reference for AGC control**. One problem of normal SSB is that with no carrier present, there is no way to provide an AGC intelligent enough to tell whether someone has quit talking or whether the signal has faded.

The usual approach is to make the AGC have as fast an attack time as possible, but to have a relatively slow release time. Thus a sudden loud signal will be brought under control quickly (typically under 5 mS with audio derived AGC's and perhaps 2 mS with IF derived systems). The gain will then stay at this level for at least 400 mS and sometimes as much as 2 Seconds before full receiver gain is restored.

Another common approach uses a long decay time constant, say 3-5 seconds, but has a secondary detector which constantly looks for the presence of audio. If no audio is detected for more than 100 to 500 milliseconds (depending on the time constant chosen), this detector turns on a switch which discharges the main AGC time constant capacitor until either the audio signal is once again detected or full receiver gain is re-established.

This "hang" AGC produces very little gain change while someone is talking unless 1) he pauses for more than the reset period, or 2) the signal fades for more than the reset period, or 3) the signal goes away. This prevents the moderate to heavy integration of desired signal with background noise or interference common to simpler AGC's **without** having to have a time constant that is so long that strong noise pulses can totally blank out the desired signal while the AGC recovers. With hang AGC the receiver gain recovers after only one reset period...200 milliseconds rather than 2-3 seconds!

Unfortunately, neither system just described can handle the flutter rates experienced at VHF.

CARRIER DERIVED AGC

It became evident early on in VHF experimentation that **NO ENVELOPE DERIVED AGC IS ADEQUATE AT VHF for mobile.**

While the fade rates at low Band (30-50 MHz) could be handled with fairly conventional Carrier or Pilot Tone derived AGC's, High Band and UHF systems needed something faster.

Certainly carrier derived systems are moderately fast compared to standard envelope detection as normally used on HF SSB systems. Even older AM AGC's were able to use time constants on the order of 200-500 milliseconds without serious degradation of the modulation envelope.

Pilot Tone or Carrier based AGC's in VHF SSB experiments are able to achieve fade correction rates to somewhere between 5-15 Hz without too much complexity or instability.

The problem is mainly one of separating the pilot or carrier from the modulation. If any of the modulation is seen by by AGC detector, it will be impressed negatively on the audio output signal causing distortion or signal reduction.

One could use an IF filter at the Pilot frequency (whether carrier or tone pilot doesn't matter), but most designers feel that this is not economical. The cost of another, very narrow IF filter is far more expensive than **any** form of baseband (audio/video) processing...especially if one looks towards Large Scale Integration of the circuitry into chip form in the future. A few cents in an LSI chip is more attractive than tens of dollars for a filter.

This also explains one reason that merely transmitting SSBRC, or Single Sideband with Reduced Carrier is not as desireable as TIB or TAB Audio Pilot Tones. Transmitted carrier essentially becomes zero frequency when tuned in correctly thus making baseband processing of the AGC (and Phase Locking, for that matter) impossible.

FBAGC VERSUS FFAGC

Use of SSB at VHF brought back another technique that is not totally new but is finding new application.

We are all familiar with **Feed Back AGC/(FBAGC)** where a detected level reference is filtered to eliminate modulation components then sent back to the input to set the RF and IF stages to a gain that won't overload the detector and will keep the output audio relatively constant.

As previously mentioned, it does have a speed limitation even when acting only on the carrier due to the propagation delay through the RF and IF stages and AGC detector/filter before the control voltage is sent back to the controlled stages. Using baseband techniques at least, the control speed seems to become unstable in the 5-15 Hz region.

Feed Forward AGC(FFAGC), on the other hand, has been successfully used to control flutter even at UHF frequencies where flutter rates approach **100-300 Hz!**

This technique was actually developed for AM UHF systems and essentially low passes the audio output of an AM detector at a frequency below the lowest modulation frequency being transmitted. It is presumed that any output lower than the lowest frequency transmitted is being caused by signal level variation coming into the receiver, not transmitted audio.

This signal is then passed **forward** to a variable audio amplifier whose gain is controlled by the **FFAGC** voltage so that the detected modulation audio passing through the amplifier is increased or decreased in inverse proportion to the **FFAGC** voltage. If the **FFAGC** voltage goes up (indicating increased input signal strength), the audio gain is reduced accordingly. If the **FFAGC** voltage goes down, the audio gain is increased to compensate.

Additionally, since the controlled stage **follows** the AGC detector, the detected modulation audio can be **delayed** to **compensate for the FFAGC filter phase delay** so that the control voltage and audio signal arrive at the same moment. This prevents the loop oscillation and propagation delay inherent in **FBAGC** circuits and allows **FFAGC** to operate with much greater speed.

Futhermore, since we are dealing with a pilot carrier rather than an AM envelope detected signal (which includes the modulation component), the **bandpassed** Pilot Tone can actually work about **two times faster than the low pass variety** used previously in AM systems.

British studies at 457 MHz using both AM and SSB systems show very good results using this system.

SELECTIVE FADING CORRELATION

One last item to be consider is the effect of Carrier, TIB, and TAB derived AGC systems relative to **selective fading.**

In that the fading of VHF and, particularly, UHF signals is largely due to Dopper differences between direct, in front, and following reflections relative to a vehicle's motion, fading at the carrier frequency will occur at a slightly different time than fading at the frequency of the highest modulation sideband. With FM the sideband components can be

6-8 KHz removed from the carrier and 12-16 kHz removed from each other. There is a great deal of de-correlation of fades separated this widely in frequency.

With SSB the maximum frequency difference from carrier to highest sideband component will be equal to the modulating frequency...typically about 3 KHz. The de-correlation between fades separated by 3 KHz are relatively minor at VHF but can be significant at UHF.

Obviously, either Carrier derived or Tone Above-Band (TAB) AGC systems will have more de-correlation between audio channel fading and the pilot channel than Tone In-Band (TIB) systems but this is not too significant at VHF.

On the other hand, TAB systems will have another problem (and, depending on techniques used, carrier systems as well). Since the tone is at the side of the IF filter phase variations normal to the edges of very selective filters will create further AGC anomalies. (If carrier is used with its own IF filter, this will not be a problem in that the carrier will be in the center of the filter.)

AGC SUMMARY (11)

1) **Feed back AGC (FBAGC)** is useful to establish gain control of RF and IF stages but is limited in speed and is insufficient by itself for VHF fade rates.

2) **Feed Forward AGC (FFAGC)** is easily implemented in Audio Pilot AGC systems and provides good AGC control at VHF frequencies with adequate control even at UHF frequencies.

3) **Carrier derived AGC** must be processed at IF frequencies and is therefore not the favored approach.

4) **Tone In-Band (TIB)** provides the best selective fading correlation but requires filtering out part of the audio spectrum to eliminate the tone.

5) **Tone Above-band (TAB)** provides reasonable correlation with selective fading at VHF but has less correlation at UHF. It also has problems associated with being located at the edge of a selective filter.

6) **Transparent Tone In-Band (TTIB)** shifts a portion of the audio spectrum 200-300 Hz upward, inserts the pilot tone in the gap, then re-assembles the original spectrum upon reception. This is a more complicated system to implement using discrete parts, but may be very practical with LSI techniques.

FREQUENCY TUNING

One of the main arguments in the early days of Land Mobile against use of SSB (and brought up again in recent times) is that it requires "extreme frequency accuracy".

Of course, this is true. Any SSB technique needs tuning precision measured in tens of Hz if natural sounding audio acceptable to commercial users is to be achieved.

Several things have modified the practicality of VHF SSB over the years. It is not at all unusual to maintain VHF oscillator accuracies within a few hundred Hertz. Higher accuracies are possible with state-of-the-art circuitry recently developed.

Then, again, "a few hundred Hertz" is not adequate for VHF SSB reception, but it provides accuracy close enough for use of a second technique that is now definitely in the "commercially viable" domain...the phase locked loop.

While the **Phase-Locked Loop** is not new, development to cost effective utilization in the commercial marketplace had to wait for the CB Radio boom of the early 70's. Prior to this almost all use was in military equipment or in test equipment, neither of which are generally low budget items. CB Radio was much more cost-conscious and had sufficient volume to interest the chip makers in Japan and the US in using the new IC technologies to develop PLL systems that could be sold in $75 radios at **a profit!**

We even have handheld, synthesized Walkie-Talkies and Monitor receivers that are in the $300 price range...the cost of the synthesizer alone not too many years ago.

Obviously, then, by using carrier or pilot tones we can create a PLL that will easily pull in a signal that is already within a few hundred Hertz and lock it within a few cycles of the original signal. Additionally, if we are smart in our design, we can reduce or eliminate much of the Doppler frequency shifting present on mobile signals at VHF and UHF.

Several problems remain, however. Some means of identifying the desired carrier from among several possible co-channel stations would be useful. Attention to loop design to minimize phase noise and sideband generation must be used if adjacent channel selectivity is to be maintained. Loop design trade-offs must be balanced between fast lock up, tight frequency control, phase noise/sideband considerations, and pull-in range.

The references yield several papers on these topics that will go into these areas in more detail (12).

Most current systems are able to pull in signals as much as plus and minus 800 Hertz from the receiver's frequency. Typical loops have settling times in the 100-200 millisecond region. STI radios ensure lock up by transmitting only the pilot carrier at full output for about 100 milliseconds before the audio channel of the transmitter is activated.

THEORETICAL PERFORMANCE DIFFERENCES - SSB/FM

In theory, FM provides 20 dB Quieting and about 24 dB SNR at a Carrier-to-Noise Ratio (CNR) input of about 12dB. This will vary with modulation index, receiver bandwidth, clipping levels used and so forth but is an adequate model for comparative purposes.

Given a receiver with a sensitivity of -117 dBm (about 0.3 uV) for 20 dB quieting, we have a noise floor of -129 dBm (12 dB lower).

Bandwidth for most VHF SSB systems, on the other hand, require 3.2 KHz. This is a 10 Log (18KHz/3.2KHz) dB improvement in noise floor, about 7.5 dB or -136.5 dBm. Thus, the CNR and SNR for non-compandored SSB at an input level of -117 dBm (.3uV) will be only 18.5 dB (CNR=SNR for non-compandored SSB).

The same bandwidth factor would apply to any external noises entering the receiver. The SSB selectivity would reduce the incoming noise by a factor of 7.5 dB. Unfortunately, it also tends to "stretch" the pulses somewhat so that part of this improvement is lost. Additionally, FM, once firmly into a limited state, will rapidly depress the AM (pulse) components well below the desired modulation.

Use of Amplitude Compandoring on the SSB signal provides about a 4:1 effective improvement in SNR for a given CNR at levels at least 10 dB above the noise. Some improvement in SNR is evident down to about 5 dB CNR...making ACSB useful somewhat below FM's threshold. In our example, then, 4:1 compandored ACSB would yield an effective SNR with an input CNR of 18 dB equivalent to "Full Quieting" (greater than 40-50 dB).

However, expansion does not really improve SNR...a **tone** in the receiver would still have only 18 dB SNR while present. What the expandor does is to **change the gain proportionate to the signal level so that the ear/brain combination does not "hear" the noise since it "goes away" when the modulation goes away.** This provides the **effect** of a signal with up to four times the SNR as long as the incoming CNR is better than 8-10 dB above the noise (that is, the incoming signal has only moderate noise riding in on it).

DATA COMMUNICATIONS

Of course, this brings up another important problem with ACSB...data communications. In that the actual CNR under field conditions typically falls into the 5-20 dB area, use of various tone signalling and data communications must be made to match the mode being used. Techniques quite adequate using FM equipment will be marginal using ACSB.

However, PSK, QPSK and other techniques have been implemented to provide data transfer with good speed and accuracy. One technique called Rectangular Spectrum Modulation (or RSM) provides for a 10-6 bit error rate at an SNR of 11.2 dB in 2500 Hz bandwidth. Baud rate is 2400 baud. This was developed for the **Skylink Land Mobile Satellite system.**

OTHER COMPARISIONS - FM/ACSB

ACSB radios tend to be somewhat quieter than FM radios when used at poor signal-to-noise conditions (though similar effects could be had using compandoring with FM except for the noise bursting common on that mode under weak signal conditions).

Range of an ACSB radio, all other conditions equal, should show up to 1.5 to 2 times the range of FM. However, without Noise Blanking and because the average power of ACSB is 6 dB lower than FM, this is not generally seen in mobile applications. Some report 10-20 % range increase under many operating conditions but this is variable.

FM, on the other hand, should see about 3-6 dB advantage when using DTMF or FSK data systems due to its actual SNR (compandoring does not work on tones, obviously). Then again, ACSB with PSK-based data systems should have parity with FM FSK-based systems if designed properly.

FM, it is generally conceded, has better "transparency" than ACSB. Both systems, however, have about the same "clarity" with ACSB slightly better in this regard at poor SNR conditions and FM "sounding better" at higher SNR levels.

The best one can say is that the systems sound different. Everything else seems to be a matter of taste, not effectiveness.

USER COMMENTS

Several users of ACSB have done direct comparisons with their FM systems.

One Rochester, New York user (in STI's home town), finds the ACSB system "equal to or better" than the FM system . He notes that it has better range.

A user in Santa Cruz, California says that his ACSB repeater has superior range with acceptable quality in this difficult, hilly service area.

A participant (guest) in some of the FCC testing between ACSB and FM in the Washington, D.C. area reported that the ACSB would still be readable after driving into an area where the FM receiver would completely fade to "White noise". This was using the same antenna for both receivers.

Furthermore the FM signal would have considerable "chop" driving fast through this area where the ACSB signal became a bit "fuzzy" or "distorted" but was perfectly readable.

The same observer was somewhat surprized when the FCC report on these tests came out as less than enthusiastic about the performance of ACSB!

Then again, another experimenter with ACSB at Cal Poly concluded that FM was clearly superior in almost all cases. I have not been able to follow up on this report or whether "preference" or actual test results were the basis of this conclusion, however.

It is well established that "FM sounds better" when A/B testing is done. This is largely due to its "transparency" and typically wider frequency response since many ACSB radios cut off frequencies above 2500 Hz.

On the other hand, most prefer ACSB under weak signal conditions even though FM sounds better at moderate to high SNR conditions.

E.F. Johnson, long in the FM Land Mobile business, has suggested that many customers prefer their ACSB based Mobile Telephone performance to FM-based systems. One would presume that it is the lack of signal loss noise bursts and the generally quieter background (FM frequently has residual high frequency hiss) that prompts this evaluation. One is less tolerant of these things when he has a phone set directly on his ear!

Conclusion? FM has a place. ACSB has a place. Each has something to offer.

OTHER APPLICATIONS OF ACSB TECHNIQUES

As previously mentioned, some systems are well suited for ACSB techniques either by bandwidth requirements or SNR threshold requirements or both.

Quite a number of satellite communications systems have switched to SSB techniques which are close enough in implementation to be considered ACSB. They have compandoring, frequency equalization, a pilot reference, and so forth.

California Microwave has converted to SSB techniques for their satellite transponders and actually loads many channels with noise to simulate fully loaded transponder conditions. So far they haven't been able to fully utilize all of the **6,000 voice channels possible on their transponder using SSB!**

ATT Long Lines continues to replace much of their land based microwave system with SSB based equipment.

MCI recently converted their Southwestern system to SSB.

SKYLINK

An interesting proposal with wide implications is that recently approved for experimental development by the FCC as proposed by Skylink, Inc.

This system uses a HELAPS transponder (High Efficiency Linear Amplifier by Parametric Synthesis) to provide DC to RF conversion efficiencies in the 35-45% range instead of the usual 10-15% associated with multiple signal linear amplifiers. Thus, the spacecraft DC load is reduced from about 2.5KW to only 630 Watts. One can appreciate the savings in power budget this has on a solar powered satellite.

We've already discussed the spectrum savings over an FM satellite system, which is 5 or 6 to 1 but the power budget savings compared to an FM system is also a factor of 1-6 dB, depending on the characteristic of the FM system used.

Further, the ground station requirements are very reasonable...we are talking about 6 dBi omni-directional mobile antennas fed with 5 Watts from the ACSB transmitter or 14 dBi antennas with a 1 Watt Portable system.

As illustrated, the transponders on board this Space Craft will each support up to 50 channels at 5KHz spacings in each 150 KHz transponder channel. The Space Craft will have 16 of these transponders, half of which are Left Hand Circular and half of which are Right Hand Circular polarization (LHCP, RHCP). A total spectrum of 4 MHz, then, will support 50 x 16 channels or a total of 800 channels. Obviously, the use of NBFM in the system would increase the spectrum requirement at least four-fold.

Utilization will be a mix of ACSB Voice Channels, Linear Predictive Coding (LPC) Secure Voice Channels using RSM modulation, one 2400 Baud Data Channel, or N times 2400/N Baud TDMA Data Channels.

Concerning SSB Skylink says, "The use of SSB has many advantages...in SCPC satellite networks."

"..channel density will be six times that of cellular NBFM"

"SSB requires only one tenth the average satellite transponder power for link performance comparable to NBFM."

"The lack of a "threshold phenomenon"..is a distinct advantage..rapid fading just below the nominal link margin is more user acceptable with SSB's characteristic graceful degradation"

"Yet another advantage of SSB is..reuse of the same linear channel for both inbound and outbound links"

Concerning their basic system design they say,

" Skylink proposes to provide the nation's first Thin-Route/Mobile Satellite Network, providing mobile, portable, and limited fixed services tailored to fit the most pressing needs in Rural/Remote Areas."

" The Geographically diverse and frequently mobile rural market has been difficult, expensive, or impossible to serve using terrestrial technology better suited for urban markets."

" Conventional satellite systems requiring large, expensive earth terminals provide little or no help to the rural communications industry; such systems are also optimized for the inherently high capacity needs of inter-urban networks."

A close study of the application made by Skylink to the FCC to establish a Developmental Land Mobile satellite Service in the 800 MHz region clearly shows that the practicality of such a system depends very much on the spectrum and transponder efficiencies to be had only through ACSB, RSM, and other narrowband techniques.

Summary

ACSB technology has much to offer designers of VHF and UHF communications networks.

Many earlier reasons for not using SSB on VHF and UHF have been made obsolete by various older and newer technologies when they are conscientiously applied to the problem.

For multi-channel and/or weak signal (long range) applications, ACSB has several advantages over NBFM. It provides very attractive spectrum savings and range extension through threshold extension.

If ACSB is allowed free access to the marketplace by FCC action to make its status permanent rather than developmental, it is sure to establish itself as an answer to several problems now facing the Land Mobile industry. It will, in fact, open up new areas such as proposed by Skylink.

For the RF and Analog engineer (and even the digital engineer), ACSB provides numerous systems and applications challenges that require unique and creative solutions.

-8-

ACSB NOTES:

1. Buesing biblio #15
2. FCC UHF TASK FORCE biblio #1
3. Lusignan, Bruce IEEE Spectrum biblio #2, 38
4. Application notes - NE 570
5. Fred Cleveland - West Coast VHF/UHF Convention 1979
6. Ray Wells - biblio #6
7. Ap notes - NE 570 Signetics
8. AGC biblio # 19, 20, 21, 22
9. Ray Wells - biblio # 6
10. Mc Geehan biblio # 22
11. biblio # 19-22
12. biblio # 4, 13, 38
13. biblio # 32

ACSB BIBLIOGRAPHY

1) Federal Communications Commission UHF Task Force Report "Spectrum-Efficient Technology for Voice Communications" February 1978 Office of Plans and Policy NTIS, 5285 Port Royal Road Springfield, VA 22161

2) Lusignan, Bruce Single Sideband Transmission for Land Mobile Radio IEEE Spectrum - July 1978

3) P.R. Docket 80-440 FCC, Washington, D.C.

4) McGeehan and Sladen Elimination of False Locking in Long Loop Phase-Locked Receivers IEE 0090-6778/82/1000-239100.75

5) McGeehan Design and Characterization of a Phase-Locked VHF Mobile Receiver IERE Conference - Civil Land Mobile Radio London, England November 1975, pp 37-46

6) R. Wells "SSB for VHF Mobile radios at 5KHz Channel Spacings" IERE Conference Radio Receivers, Associated Systems London, England 1978, pp 29-36

7) J.R. Ball and D. Holmes "An SSB with Pilot Receiver for Mobile Radio" Proceeding IERE Clerk Maxwell Commemorative Conf. Radio Receivers, Associated Systems Leeds, England July 1981, pp 429-435

8) W. Gosling, J.P. McGeehan, and P.G. Holland "Receivers for the Wolfson Single Sideband VHF Land Mobile Radio system" Radio Electronics Engineering Vol. 49, pp 321-325, 1979

9) "A Digital Technique for Temperature Compensation of Crystal Oscillations" Rec. and Assoc. Systems. Proc. of Conf. Radio IERE, London, 1978 (see also # 17)

10) B.N. Biswas and P. Banerjee, "Range Extension of a Phase-Locked Loop", IEEE Trans. Commun., Volume COM-21, pp 293-296, 1973

11) J.P. McGeehan, G. Lightfoot, A. Lymer, and W. Gosling "Optimisation of the Wolfson SSB Radio Receiver" Proceedings- IERE- Conf. Radio Receivers Leeds, England, July 1981, pp 417-428

12) F.M. Gardener Phaselock Techniques New York: Wiley, 1979 (Standard Reference on PLLs)

13) J.P. McGeehan "Comparative Adjacent-Channel Selectivity Performance of Phase-Locked Pilot-tone SSB Mobile Radio Receivers with Particular Reference to the Long-, Short-, and Split-Loop Configurations" IEE Proceedings, Vol 129, Pt. F, No. 6, December 1982

14) Magnuski, H. "Improving Spectrum Utilization in Mobile Mobile Communication" 26th Annual Conf. IEEE Vehicular Technology Group pp 105-110, March 1976

15) Buesing, R.T. "Modulation Methods and Channel Separation in the Land Mobile Service" IEEE Trans. Vehic. Tech., Vol VT19, p 187, May 1970

16) Rawling, A.J., McGeehan, J.P., and Gosling, W. "Forward Feeding AGC with Extended Signal Delays" Proc. of Conf. "Radio Receivers and Assoc. systems" IERE, London, 1978 Warwich, G.A., Gosling, W., and Prescott, A.J.

17) Warwich et al "Digital Compensation Technique in Crystals Yields Ultrastability" Electronics, Vol 51, No. 19, pp 67-68, Sept. 14, 1978

18) Gosling, W. and Weston, D. "A Quasi-Synchronous VHF SSB System" Proc. of Conf. Civil Land Mobile Radio, 1975 IERE, London, 1975

19) McGeehan, J.P. and Bateman, A.J. "Theoretical and Experimental Investigation of Feedforward Signal Regeneration as a Means of Combating Multipath Propagation Effects in Pilot-based SSB Mobile Radio Systems" IEEE Transactions on Vehicular Tech. Vol VT-32, No. 1, Feb 1983 0018-9505/83/0200-010601.00 1983 IEEE

20) McGeehan, J.P. and Burrows, D.F. "Performance Limits of Feedforward AGC in Mobile Radio Receivers" Proc. IEE, part F, no 6, pp 385-392, 1981

21) K.W. Leland and N.R. Sollenberger "Impairment Mechanisms for SSB Mobile Communication at UHF with Pilot-based Doppler/fading Correction" Bell System tech. J., vol 59, pp 1923-1942,1980

22) J.P. McGeehan, D.F. Burrows, and A.J. Bateman, "The use of Transparent Tone-in band(TTIB) and Feedforward Signal Regeneration (FFSR) in Single Sideband Mobile Communications Systems" Proc. Inst. Elec. Eng. Conf. Communications 82 pp 121-126, 1982

23) J.P. McGeehaan and A.J. lBateman, "Subjective performance of Amplitude Companding in SSB Mobile Radio Systems" Electron. Lett., vol 17, pp 859-860, Oct. 1981

24) H. Magnuski and w. Firestone "Comparison of SSB and FM for VHF Mobile Strength" Proc. IRE, vol 44, pp 1834-1848

25) W.B. Bruene "Distortion Reducing Means for Single-Sideband Transmitters" Proc. IRE< vol 44, pp 1760-1765, Dec. 1956

26) Belrose, J. " On SSB Radio Communications." QEX, July, pp. 3-4.

27) Bishop, D. "Mobile Satellite Experiment Tests ACSB, Data Techniques." Mobile Radio Technology, June 1984, p. 43.

28) Childs, F. "ACSB--Filling the Gaps Between FM Channels." Mobile Radio Technology, January/ February 1983, pp 60-66.

29) Daniel, J. "Additional Channels for RCCs." Telecourier, June 1983, p. 12.

30) Daniel, J. "New Channels." Two Way Radio Dealer, July 1983.

31) Fahnestock, J. "MRT Previews STI ACSB Handhelds." Mobile Radio Technology, March 1984, p.20

32) Haller, R. and H. Van Deursen, Amplitude Compandored Sideband Compared to Conventional Frequency Modulation for VHF Mobile Radio--Laboratory and Field Testing Results. FCC OST Technical Memorandum, FCC/OST TM83-7, October 1983.

33) Mang, B. "ACSB: Solution for Spectrum

Starvation." Radio Communications Report, July 15, 1982.

34) Manz, B. "SSB Technology Fights its Way into the Land-Mobile Market." Microwaves and RF, August 1983, p. 72

35) Snyder, B. "What ACSB Means to You." Business Radio Action, June 1982, pp. 13-17. (Volume XVIII no. 3)

36) Technical Services Dept., Stevens Engineering Associates, Inc., "ACSB Presents New Challenge to FM Service Technicians." Mobile Radio Technology, February 1984, p. 28.

37) Woodward, G., C. Hutchinson and P. Rinaldo, eds. The Radio Amateur's Handbook (61st edition). Newington: The American Radio Relay League, Inc., 1984

38) "The Problem of Speech Pulling and its Implications for the Design of Phase-Locked SSB Radio Systems" Proceedings-Inst. Elec. Engineering, Volume 128, part F. no. 6, pp 361-369, 1981

ELECTRICAL CHARACTERIZATION OF QUARTZ CRYSTAL UNITS
THROUGH HIGH PERFORMANCE VECTOR NETWORK ANALYSIS

Kenneth M. Voelker, Product Manager

Hewlett-Packard Company
P.O. Box 69
Marysville, Washington, 98270

ABSTRACT

Measurement of the electrical parameters of quartz crystal units is simpler and more accurate than ever before, as a result of recently introduced instrumentation. A 200 MHz vector network analyzer and companion s-parameter test set can rapidly determine the equivalent circuit values and resonant frequencies for even high performance crystal units. Internal calibration firmware, which compensates the effects of both test fixturing and residual instrument errors, makes excellent correlation of test results possible.

For years, the ability of quartz crystal manufacturers to build products with precisely controlled characteristics has rivalled or exceeded their ability to measure them. Non-standard testing techniques and inadequate test instrumentation have given these devices a reputation for being nearly impossible to characterize accurately. Crystal manufacturers and users alike can attest to the difficulties of obtaining correlation between measurements made by vendor and customer, or even among locations within the same company.

Historically, when highest accuracy was required, the instruments called upon were members of the network analyzer family, including vector voltmeters, gain/phase meters, impedance analyzers and network analyzers. Fundamentally capable of making the required measurements, these instruments required a great deal of test engineering expertise to implement a complete solution. Complex test fixturing, hundreds of individual calibration and measurement points, and a monumental data reduction effort were required for measurements of even minimal accuracy.

Recent developments in test instrumentation have produced products that are not only more accurate, but which can do a great deal of data collection, correction and manipulation with little or no operator intervention. One such product is the recently introduced Hewlett-Packard 3577A 200 MHz Vector Network Analyzer. With proper implementation, its capabilities for stand-alone crystal measurement equal or exceed those previously found in even fully automated systems.

This paper will review the essentials of both crystal parametric measurements and network analysis, and will give practical hints for facilitating these difficult measurements.

The Measurement Problem - Crystal Parameters

Proper characterization of a quartz crystal resonator involves finding, as a minimum, the electrical parameters listed in Table 1. The first four are the lumped elements from the familiar equivalent circuit of Figure 1.

Co	shunt capacitance
C1	motional capacitance
R1	motional resistance
L1	motional inductance
fs	series resonant frequency
fp	parallel (or anti-) resonant frequency
Q	figure of merit

Table 1: Required Electrical Parameters for Quartz Crystal Units

Accurate determination of these values assists design engineers by enabling them to closely predict crystal behavior under actual circuit conditions. Repeatable measurement results allow the crystal manufacturer to precisely describe the characteristics of his product to customers. All of these parameters (and more) can be derived from the frequency-dependant complex impedance of the crystal unit. Several important mathematical relationships are involved, which will be reviewed in the following paragraphs.

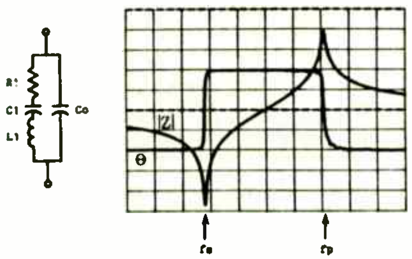


Figure 1: Resonator Impedance Variations with Frequency

Motional Parameters L1, C1 and R1

As suggested by the equivalent circuit, the crystal unit exhibits a series resonance at a frequency

$$fs = \frac{1}{\sqrt{4\pi^2 \; L1\,C1}} \qquad (1)$$

As shown in Figure 1, fs is identifiable (to a first approximation) as the frequency of minimum impedance and zero reactance (zero phase angle). At fs, the inductive and capacitive motional reactances cancel and the only significant component to the impedance is R1, which can be read directly from the plot.

L1, the motional inductance, is directly related to the rate of change of reactance with frequency at fs, in accordance with

$$L1 = \frac{1}{4\pi}\,\frac{dX}{df} \qquad (2)$$

Finally, C1, the motional capacitance, is calculated from fs and L1.

$$C1 = \frac{1}{L1\,(2\pi\,fs)^2} \qquad (3)$$

Shunt Capacitance Co

While the equivalent circuit for a crystal resonator assumes a shunt capacitance Co, the preceding discussion has ignored its influence. Practically speaking, this can be quite valid, especially in cases where resonant frequency is low and/or Q is high (or where accuracy is not a primary concern). Co causes the following:

1. Creation of a parallel resonance Often referred to as "anti-resonance", this is a second frequency of zero phase, but with impedance at a maximum value. This occurs at a frequency of

$$fp = \frac{1}{2\pi}\sqrt{\frac{1}{L1\,C1} + \frac{1}{L1\,Co}} \qquad (4)$$

2. Spreading of critical frequencies The addition of a parallel reactance complicates the impedance plot around fs (see Figure 2). No longer

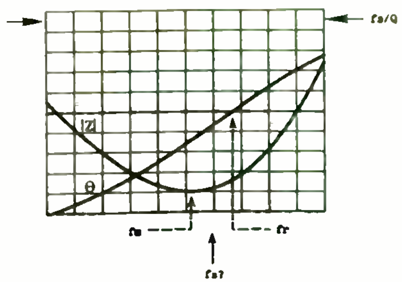


Figure 2: Crystal Impedance near Series Resonance, Showing Influence of Co

do the frequencies of zero phase (fr), minimum impedance (fm) and series resonance (fs) coincide. In fact, fs (critical for calculating R1, C1 and L1) no longer corresponds to any identifiable point on the impedance plot.

A further difficulty arises from the fact that any stray capacitance in the measurement setup or test instrumentation will appear in parallel with Co, adding directly to it. Without the proper measurement, calibration and data manipulation techniques, this causes a twofold problem: 1) the shunt capacitances will change the apparent values of all crystal parameters, and 2) this influence will vary dramatically as a function of test setup. This is a major reason why crystal measurements performed at different locations often correlate poorly.

Co can be calculated from the impedance of the crystal unit at a frequency a few percent away from resonance. At such frequencies, the impedance of the motional arm (L1, C1 and R1) will be high enough to have negligible effect on the overall device impedance, allowing use of the equation

$$Co = \frac{1}{2\pi\,f\,Xo} \qquad (5)$$

A later section of this paper will present a means for determining Co analytically from the series resonance impedance plot.

Resonator Q

Several methods are available for finding the Q of the crystal resonator, the simplest being to calculate it from the motional parameters already obtained, using equations such as

$$Q = \frac{1}{2\pi\,fs\,C1\,R1} \qquad (6)$$

or

$$Q = \frac{2\pi\,fs\,L1}{R1} \qquad (7)$$

Another widely used technique determines Q from the rate of change of phase with frequency (phase "slope"). In the vicinity of fs,

$$Q = \frac{\pi f}{360}\,\frac{d\theta}{df} \qquad (8)$$

Fortunately, many network analyzers include the ability to measure phase slope directly as "group delay", as defined below. This simplifies the above equation to

$$Q = -tg\,\pi\,f \qquad (9)$$

where

$$tg = -\frac{d\theta}{360\,df} \qquad (10)$$

The Measurement Solution - Network Analysis

A state-of-the-art network analyzer can perform all of these measurements with ease and accuracy. A network analyzer is a complete, self-contained stimulus-response test system, as shown in the block diagram of Figure 3. It performs its measurements by furnishing a test device with a known stimulus and characterizing the response. To aid in understanding the functions of a network analyzer, its three key blocks are now discussed in more detail.

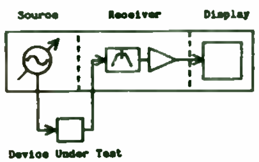


Figure 3: Functional Block Diagram, Network Analyzer

The Source

The source furnishes sinewave energy to the device under test, and determines the measurement frequency. The following characteristics are very desirable:

Frequency synthesis The operating frequency of the source should be derived from and locked to a stable reference oscillator. This is especially critical when measuring crystals or other (hi-Q) devices whose parameters change radically with very small frequency changes. A suitable synthesizer would include its own TCXO reference and be settable in frequency increments of .01 Hz or less.

Frequency sweep Most devices will be measured over a span of frequencies, so the source must have the ability to sweep. For many synthesizers, this is an unwelcome complication, as their frequencies can only be changed in abrupt, discontinuous steps. The resulting transients tend to cause "ringing" in hi-Q devices, and a time delay at each frequency must be provided to allow the measured data to settle to a stable value. The synthesizer used in the Hewlett-Packard 3577A network analyzer changes frequency with no phase discontinuities, and provides a quick, smooth "analog-like" sweep.

Adjustable amplitude Output level must be variable over a range sufficient to accommodate the devices to be tested. Good source flatness (i.e. constant output voltage) as a function of frequency is also desireable, but with proper measurement setup and today's computerized error-correction techniques, it can be made less critical.

The Receiver

The receiver is the measurement portion of the network analysis system, quantifying the response of the test device to the stimulus. Although every network analyzer will have at least one receiver, having two or three will simplify and speed up the measurement task. The key technical requirements for this functional block include:

Vector measurements A major distinction between network analyzers and other types of analyzers is their ability to measure phase as well as magnitude. A suitable receiver must be able to measure both the real and imaginary components of its input signal.

Narrowband response A narrowband receiver optimizes dynamic range, allowing the measurement of higher amounts of gain and attenuation and wider ranges of impedance. Each decade decrease in receiver bandwidth lowers the noise floor by 10 dB (although at the expense of measurement time). Most network analyzers provide a selection of bandwidths, ranging from as narrow as 1 Hertz, for very noise free (but slow) measurements, to rapidly responding bandwidths of a kilohertz or more.

Source tracking Because the frequency of the input signal is constantly sweeping, the tuned frequency of the receiver must also sweep. Use of the same synthesized local oscillator for both source and receiver enables them to track perfectly.

The Display

The raw data provided by the receivers will rarely be in a form convenient for interpreting the measurement. Instead, mathematical operations on the data will be necessary, as well as a means for graphically presenting the results. The display section handles these functions.

Computation Even the most basic measurements can require complex manipulation of the raw data. Phase, for example, is obtained by dividing the imaginary portion of the received signal by the real portion and taking the arctangent. Further, those measurements that are inherently ratios (gain, for example) require the simultaneous handling of data from two (or more) receivers or memory registers.

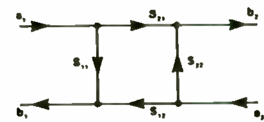
Graphics Finally, a high quality display such as a CRT should be used to plot the results, nominally with frequency on the horizontal axis and the measured value on the vertical. A more useful display would allow two or more plots to be shown at once, or might implement more complicated displays such as polar plots, Smith charts, etc.

Network Measurements

The following sections will discuss terms and techniques common to network analysis measurements.

S-Parameters

Various schemes exist for mathematically describing the characteristics of an electrical network. Among the most common are H-, Y- or Z-parameters, which variously define networks in terms of port-to-port impedance, conductance, etc. Those familiar with these parameters will recall, however, that they can be measured only under certain conditions, generally with specific ports open or shorted. Because terminations of this sort are incompatible with most RF devices, a different set of parameters is used, called S- ("scattering") parameters.



S11	$\left.\frac{b1}{a1}\right\vert_{a2\,=\,0}$	Input reflection coefficient
S21	$\left.\frac{b2}{a1}\right\vert_{a2\,=\,0}$	Forward gain
S12	$\left.\frac{b1}{a2}\right\vert_{a1\,=\,0}$	Reverse gain
S22	$\left.\frac{b2}{a2}\right\vert_{a1\,=\,0}$	Output reflection coefficient

Table 2: Scattering Parameters, Definitions

As shown in Table 2, s-parameters are defined in terms of power transferred to or from a network under conditions of proper source and load impedance match. Not only are these measurements and their requisite conditions easily furnished by a network analyzer, the results themselves provide valuable network data with no further calculations. S11, for example, is simply input reflection coefficient, while S21 is forward power gain. For these reasons, s-parameters are a natural choice for use in making network measurements, and are used as the starting point characterizing quartz crystals.

Basic Transfer Function Measurements

Figure 4 shows the necessary connections between a typical two port test device and a state-of-the-art network analyzer such as the Hewlett-Packard 3577A. In accordance with the definition for transfer function G = S21 = Vo/Vi, both the input and output signals to the device under test must be measured. Vo is measured by connecting the device output directly to the input of receiver B, which also furnishes a 50 ohm termination. Vi is derived from the power splitter, which provides the analyzer's reference input (receiver R) with a signal identical to that supplied to the test device.

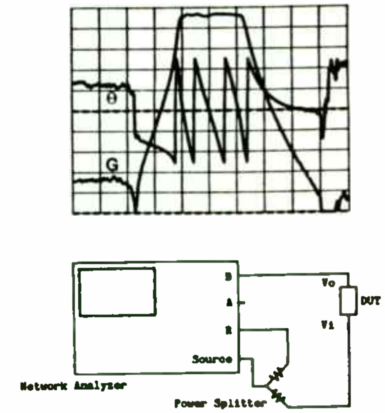


Figure 4: Test Setup and Display for Transfer Function Measurement

In the course of an actual measurement, receivers R and B each produce two voltage values (one real, one imaginary) for each of 400 frequency points across the display screen. For a transfer function measurement, the analyzer must calculate and display "B/R" at each point, in accordance with the above definition. Because the values for both B and R are complex, the quantity B/R has both a magnitude and phase associated with it, which the analyzer can display as shown in Figure 4.

Use of the analyzer's reference input (receiver R) to monitor the actual signal that is supplied to the test device makes it unnecessary to assume a perfectly flat source. The measurement error removed in this way can amount to ± 1 dB or more.

Reflection and Impedance Measurements

As useful as transfer function measurements are, crystal electrical parameters are more easily derived from impedance measurements which are, in turn, based on reflection measurements.

Reflections occur whenever there is an impedance mismatch between a source and its load. Relative to the magnitude of the incident power, some of the incident power is not absorbed by the load, but is reflected back toward the source. The vector ratio of reflected to incident power for a test device is known as its reflection coefficient, or S11.

The addition of a directional coupler to the test setup as shown in Figure 5 allows the analyzer to separately measure incident (receiver R) and reflected (receiver A) power. S11 is then calculated as the ratio A/R. Reflection measurements require only one connection to the analyzer setup, and characterize a single device port. For this reason, they are sometimes referred to as "single port" measurements.

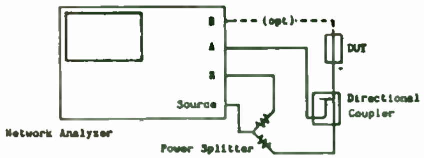


Figure 5: Test Setup for Reflection or Impedance Measurement

The equation

$$Z = Zo \frac{1 + S11}{1 - S11} \quad (11)$$

where:

$S11 = A/R$

Zo = system characteristic impedance

converts any S11 vector to a unique value of complex impedance. The Hewlett-Packard 3577A will perform this calculation automatically, displaying the results on the CRT with no intervening steps. Further, the impedance can be shown in either magnitude/phase format, simplifying the search for impedance maxima and minima, or in real/imaginary format, separately showing the resistive and reactive components of the impedance.

A subtle concern in calculating impedance from reflection coefficient relates to the range of impedances that are measureable. As the impedance of the unknown becomes very different from the charscteristic impedance of the network analyzer, S11 approaches a value of unity. Examining the equation above, it is seen that when this occurs, the calculated value of Z will be very sensitive to minute variations in S11 such as are caused by noise, etc. As a rule of thumb, reasonably stable ($\pm$ 1 %) impedance measurements are limited to a range of about $\pm$ 2 orders of magnitude around the characteristic impedance of the measurement system. This range is quite suitable for the crystal measurements described herein.

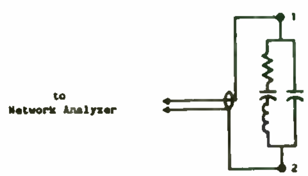


Figure 6: The Crystal as a One Port Device

Practical Measurement Techniques

Test Setup

As was demonstrated at the outset, the frequency-dependant impedance of a crystsl unit is used to determine its equivalent circuit parameters. Impedance is readily obtained by treating the crystal as a single port network, measuring its reflection coefficient (S11) and performing a mathematical conversion. Figure 6 shows a suitable measurement setup.

It should be noted that crystals can also be measured as two port networks, as shown in Figure 7. This provides certain additional information about the crystal, most notably the crystal-to-case capacitances C11 and C22. The cost of this data will be the need to make four s-parameter measurements instead of one, as well as the need for somewhat more complicated calculations.

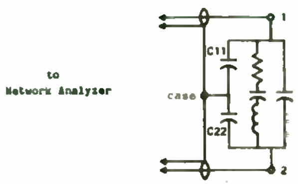


Figure 7: The Crystal as a Two Port Device

The remaining sections of this paper will give practical guidance for making crystal measurements, using the one-port method.

S-Parameter Test Set The various items shown in Figure 5 as being adjunct to the network analyzer may be obtained in a single unit called an s-parameter test set. This convenient accessory, connected between the analyzer and the device under test, integrates the RF splitters, directional couplers, terminations and other devices required for a complete set of s-parameter measurements.

Fixturing Some sort of test fixture will be needed to interface the crystal unit to the analyzer or test set. The physical adaption of connector types is, however, only part of the problem.

Repeatability is a major concern. With proper calibration, the parasitic impedances presented by the fixture will not affect the measured results, provided they remain constant from measurement to measurement. Within limits, it is more beneficial to insure the stability of these values than to attempt to minimize them. A solid mechanical structure will aid in satisfying this requirement.

Compatibility with calibration standards is a second consideration. The calibration process requires that accurate short, open and fifty ohm terminations be applied directly to the crystal socket. Some users will want to construct these in actual crystal cans, while those with the most stringent accuracy requirements will have to adapt directly to their secondary or tertiary impedance standards.

Measurement Calibration

As accurate as a network analyzer may be, most of the factors influencing overall measurement accuracy will be found external to the instrument. Consider a device connected to the analyzer through coaxial cables. The measurement results will include the attenuation and phase shift of the cables superimposed on the transfer function of the test device. Most measurements require not only cables, but also power splitters, directional couplers and test fixtures. These other devices can introduce insertion losses of 10 dB or more and flatness errors exceeding 1 dB. A modern network analyzer will include some means for modelling, measuring and removing these errors.

The Hewlett-Packard 3577A implements this via built-in calibration routines, activated by front panel commands. The three term error model used for reflection measurements requires the user to supply standard open, short and fifty ohm terminations to the measurement port. After each has been measured and stored, this data is used to mathematically correct further measurements.

The physical point at which the calibration standards are applied becomes the measurement's "reference plane". After calibration, all measured gains, losses and phase angles are displayed relative to the values that exist at this point. Errors that occur inside the reference plane, such as fixture parasitics, no longer affect the measurement; a device connected directly to the reference plane is measured as if there were no intervening fixtures, devices or cables.

Figure 8 illustrates the benefits of measurement calibration. In Figure 8(a), no calibration has occurred, and the reference plane remains at the instrument's input terminals. The accompanying plot shows how the crystal's parallel resonance frequency has been "pulled" (lowered) by the parasitic impedances of the fixture.

Figure 8(b) shows the effect of calibrating at the fixture socket itself. The fixture is now within the reference plane, and has no influence on the crystal's resonant characteristics.

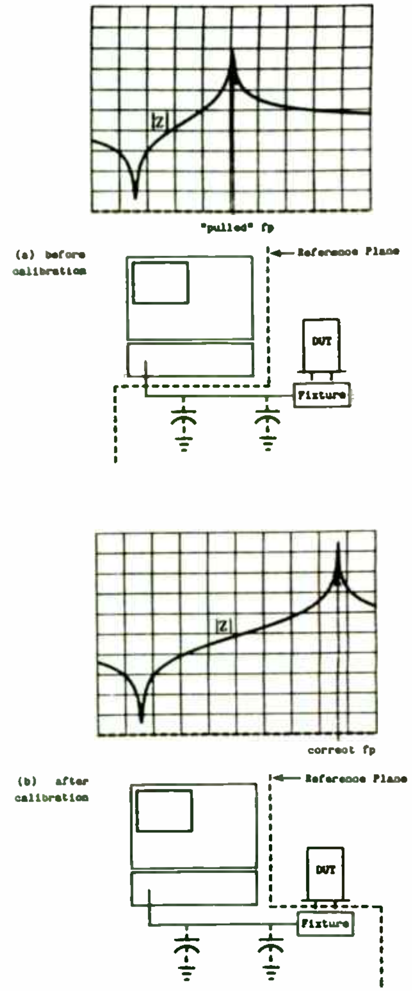


Figure 8: Effects of Calibration on Crystal Impedance Measurement

<u>General Measurement Sequence</u>

The following procedure incorporates the foregoing information into a single, unified test sequence. While based on a Hewlett-Packard 3577A Network Analyzer and 35677A S-Parameter Test Set, the basic principles are applicable to any network analyzer.

1. Initial instrument settings Setting up the analyzer requires four front panel selections:

INPUT instruct the analyzer to display impedance (calculated from S11) by selecting user-defined function "F4".

DISPLAY FUNCT display linear magnitude (in ohms) on trace 1, phase on trace 2.

FREQUENCY select a center frequency equal to fs and a sweep span equal to to approximately 5 fs/Q.

AMPLITUDE select a source amplitude suitable for the device under test.

2. Calibrate Select "One Port Full Cal" and follow the instructions provided on the CRT display.

3. Refine the instrument settings The following instrument parameters should next be adjusted as required to optimize display resolution and noise immunity.

Frequency: center the display on the impedance minimum, choosing a sweep span that covers the region between about minus twenty and plus twenty degrees of phase. Re-calibrate after the final sweep parameters have been chosen.

Display Scale: while display scaling affects neither the accuracy nor the resolution of the measured data, an expanded vertical scale may aid in visually determining frequencies of minimum impedance, zero phase and other characteristics.

Resolution Bandwidth: noise variations of the measured data can be reduced by selecting a narrower receiver bandwidth. This will necessitate a proportionally slower sweep time.

Averaging: noise can be statistically reduced by averaging together a number of measurement sweeps.

4. Removal of Co The display now plots the pin-to-pin impedance of the crystal unit in the vicinity of series resonance. As was previously mentioned, the presence of shunt capacitance Co can distort this plot, especially if Co is large relative to C1. This will obscure the correct value for fs and motional components L1, C1 and R1, and cause the frequencies of series resonance, minimum impedance (fm) and zero phase (fr) to be unequal.

When this occurs, the shunt reactance of Co must be mathematically removed from the impedance plot before the motional components are measured. The 3577A's internal computational capabilities allow it to simultaneously measure and remove Co, using the following procedure:

A. Returning to the instrument's INPUT menu, instruct the instrument to display the following "User-Defined Input":

K3*F4/(K3-F4) (12)

This equation calculates the effect of removing a parallel impedance, represented by complex constant K3, from the overall device impedance, calculated by function F4.

B. On the DEFINE MATH menu, enter an initial value of 0 - j1000 for K3, and manually adjust its imaginary (reactive) portion, using the keypad or knob. As K3 approaches the actual value of XCo, fr and fm will begin to move toward each other. When they coincide, Co can be calculated from

$$Co = \frac{1}{2 \pi f \, XCo} \qquad (13)$$

For certain hi-Q and low frequency crystals, fr and fm will initially coincide, and the above steps will not be necessary. For these cases, the reactance of Co will have to be measured separately, at a frequency away from resonance.

5. Determination of Crystal Parameters Once fm and fr are equal, they are also equal to fs, the the series resonant frequency. By moving the display marker to this frequency, R1 can be read directly from the screen. Selecting DISPLAY FUNCTION = "Imaginary" displays the reactive portion of the impedance, used to calculate L1. Selecting DISPLAY FUNCTION = "Delay" allows the phase slope to be measured for Q calculations. C1, Co and fp can then be determined mathematically using equations previously shown.

Conclusions

While a rigorous error analysis of these techniques is beyond the scope of this work, empirical findings are quite encouraging. Repeatability of one percent or better is readily obtainable for component values, as well as a few parts in 10^8 for frequency values. With traceable calibration standards, these measurements will correlate quite well from location to location. In addition, they will also be sufficiently accurate to calibrate simpler, less accurate measurement processes, such as those used in high volume manufacturing situations.

Precise quartz crystal measurements are no longer limited to metrology laboratories and elaborate, computer-controlled test setups, thanks to the capabilities of modern network analyzers. A stand-alone instrument such as the Hewlett-Packard 3577A brings these measurements to all portions of the manufacturing process, from the R&D bench to final QA.

Application of the Digital RF Memory to Communication Jamming

William J. Schneider
formerly
Communication System Engineer
ASAS Project
Communications Research Section
Jet Propulsion Laboratory
Pasadena, California

This work was sponsored by:
Paradigm Technologies
2680 Monterey Road
San Marino, California
818 799 2780

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

Application of the Digital RF Memory to Communication Jamming

Part I
Military Communications

Introduction

The purpose of this paper is to describe the potential impact of Digital RF Memory technology on communications jamming. The first step in doing this is to establish the background, and then to follow the somewhat see-saw course of Communication vs ECM vs ECCM.

It is first necessary to state say what Army communications is made of today and then to show how those communication links are vulnernable to certain jamming techniques. Next it is necessary to set out the on going developments in Army communications those developments which will lead to jam resistance communications.

Having discussed the background of present and future communications systems is set it is appropriate to attempt to show how even these advanced systems might be vulnerable to the replica jamming as implemented using Digital RF Memories. This last step will bring us to the discussion of the DRFM its implementation and properties.

Before proceeding a disclaimer of sorts is required. One of the first things one discovers when investigating military communications is the fact that the Army and DOD have a great number of developments in process. The systems mentioned here are representative of those which have been described in the open literature. There are other advanced developments which cannot be discussed here. There is therefore a substantial possibility of omission even though a substantial effort has been made to make this presentation complete.

Army Battle Field Communications

This paper deals with Army battle field communications. Because of the complexity and the variability of the topic and the limited time available for this discussion, it is necessary to narrow the discussion to a particular hypothetical situation. In this hypothetical situation a Brigade, located near the FLOT (Forward Line Own Troops), has a requirement to communicate with a command element, the Division Tactical Operation Center (DTOC) situated some distance behind the FLOT. These communications will include reports of various observations and the receipt of orders.

The Brigade will have with it radios from the Division inventory. These radios are smaller, lighter, and more reliable then their predecessors but are basically the same as those used for the last 20 years. They operate on the same frequencies, HF, VHF, and UHF. The primary mode of communication is voice reports.

There have been some advances in technology, recent developments in HF have resulted in means of measuring and predicting HF propagation in the 2 to 30 MHz band. In addition some of the newer HF radios incorporate automatic antenna matching so that operation with relatively short whip antennas is practical, a factor greatly effecting mobility.

There have emerged a new generation of HF modems which allow digital transmission over HF radio despite the variations in amplitude and group delay across the typical HF channel. Even so the rates are not high, 1200 bps being about the best that can be done.

VHF radios, 30 to 80 MHz, and UHF radios, 200 to 400 MHz, now offer X-Mode data transmission to 16 kbps. This capability allows the transmission of digitized voice. Since only digital signals can be encrypted to the highest levels these probable radios now offer secure communications.

In addition to single channel radios, those carrying a single full duplex channel, the Army has developed ATACS, the Army Tactical Communication System. This system provides dial/DTMF telephone service to areas quite near to the FLOT and in theory would allow communication from those forward areas back to higher echelons, even to the continental US. This Army version of Ma Bell is implemented by VHF/UHF multi-channel radios carrying broad band TDM signals between switching centers. Typical parameters of equipment are summarized in Table 1.

Typical Military Radio Equipment
Table 1

Type	Mode	Bit Rate/BW	Power(W)	Antenna
HF Sing. Ch.	AM, SSB	4kHz Voice Data Modem 1.2kbps	10-1000	0 dbi Whip
VHF Sing. Ch.	Binary FSK	Voice/Data 16kbps/25 kHz BW	25	0 dbi Whip
UHF Sing. Ch.	Binary FSK	Voice/Data 16kbps/25 kHz BW	25-100	0 dbi Whip
UHF Mult. Ch.	Binary FSK	Voice 576kbps/750 MHz	15 Watts	9-15 dBi Refltr

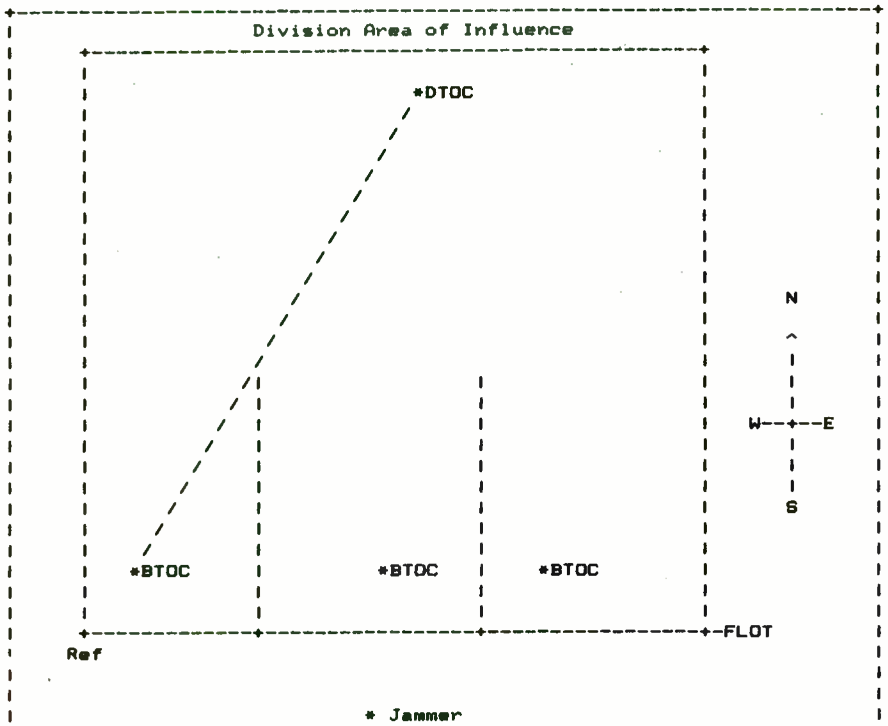
R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

Single channel radios are much preferred since their simplicity of operation allows them to be owner operated. The owner operated concept refers to a radio so simple to operate and so reliable that communication to be established without the aid of Signal Brigade personnel. This eliminates a requirement for a communication specialist and makes another fighting man is available.

Hypothetical Situation

In this situation a communication link between the Brigade Tactical Operations Center (BTOC) a forward area command post 5 km behind the FLOT and the DTOC located 60 km behind the FLOT. We will now explore what is necessary to jam this link.

The Division area of influence is shown in Figure 1. The actual dimenisions of the area will vary with the geographical situation but an area 100 by 100 km will provide a reasonable basis for discussion.



Hypothetical Situation
Figure 1

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

The parameters for each link: distances, Line Of Sight (LOS) path loss and and propagation delay times are given in Table 1. The initial calculations assume that VHF radios operating at mid-band with omni antennas are used and that LOS is available.

Link Parameters
Table 1

Unit	Coord Ref	Separation km	LOS Loss* dB	Prop Delay microsec
DTOC	60N,50W			
		71.7	104	239
BTOC 1	5N, 4W			
		39	99	130
Jammer	10S,40W			
		71	104	236
DTOC	60N,50W			

* LOS loss = 20*log(km) + 20*log(MHz) + 32.45

The Barrage Jammer

For the purpose of the base line situation, a barrage jammer is presumed to be employed. This jammer is a set-on receiver/noise transmitter type. This jamming techniques involves a four step process:

Intercept
: Search the band of interest for signals of opportunity.

Analysis
: Examine the content of the intercepted transmission to determine what action to take.

Set-On
: Tune the jamming transmitter to the intercept frequency.

Noise Transmission
: Transmit on the intercept frequency with noise modulation.

In the intercept mode the receiver is tuned across the band in a search for communication links, this may be done manually or automatically and will take several seconds. When a target is discovered the jamming transmitter will be tuned to that frequency and noise transmission begun.

Jamming will be discontinued periodically for 'look-throughs' to verify that the link is still covered. Most VHF radios have the capability of off-setting transmissions up to 15 kHz from the channel center to avoid narrow band jammers.

These factors, plus the possibility of equipment drifts, make it necessary for the jammer to transmit a noise spectrum wider than the communication spectrum. The required jammer band width may be determined as shown in Table 3.

Jammer Noise Band Width Calculations
Table 3

Communication Band Width	25 kHz
Link Offset Capability	+/- 15 kHz
Equipment Drift Allowance	+/- 5 kHz
Jammer Band Width	55 kHz
Jammer Disadvantage	3.4 dB

The jammer must maintain a sufficient J/S (jamming power to signal power) ratio over the channel to disrupt communication. The required J/S will vary with the type of transmission and with the modulation. For the example a J/S of at least 10 dB over the 25 kHz of a digital link will be considered to be sufficient.

Since the 25 kHz communication channel may be any where in the 65 kHz band the jammer must have a noise band width of 55 kHz. This gives the jammer a band width disadvantage of about 4 dB compared to the communication transmitter.

Replica Jammer Application.....The DRFM has a lot to offer in this initial situation.

Look through time....Since it is simply a memory device the set on time is equal to the duration of the wave form stored, the instruction pulse width. The operating band width of the DRFM can easily cover the VHF band, reducing the look through time from seconds to microseconds. The initial acquisition time may not be shortened since no signal received would be jammed with out manual analysis.

Band Width Disadvantage.....One outstanding feature of the DRFM is frequency accuracy. In the case of a binary FSK signal the spectrum width is determined by the chip frequency and the deviation. A DRFM can store a portion of the transmission containing the FSK frequencies and then retransmit optionally adding noise modulation. The result is a spectrum very similar to the original spectrum in width and modulation. This application is illustrated in Figure 2.

The ability to create a jamming signal having the same band width as the communication signal essentially eliminates the band width disadvantage. If longer instruction pulses are stored the DRFM will capture entire code sequences, words or packets of information.

DRFM Application as Narrow Band Noise Jammer
Figure 2

```
                          | | |
                          +-+-+
                            |
                            |
                        +---+---+
     +------------------+Switch +------------------+
     |                  +-------+                  |
     |                                             |
     |                                             |
     |   +-------+   +--------------+   +-----+    |
     +---+ DRFM  +---+ SSB Modulator +---+ AMP +---+
         +-------+   +------+-------+   +-----+
                            |
                     +------+-------+
                     |Noise Generator|
                     +--------------+

Switch      RRRRR
To DRFM     XXXXX
From DRFM         XXXXX XXXXX XXXXX XXXXX XXXXX
```

The each DRFM output is a replica of the original but the repetition causes a preferred mode spectrum at 1/pw. Generally considered to be non-coherent. but has the same envelope as the original sample. After noise is applied it is a narrow band noise centered exactly at the received frequency.

The required jammer power output can now be approximated:

I BTOC communication signal power received at DTOC

Comm Transmitter	+44 dBm	(25 watts)
Comm Rx Ant. Gain	0 dBi	
Path Loss	-71 db	
Comm Tx Ant Gain	0 dbi	
	-27 dbm	

II BTOC communication signal power received by the Jammer

Comm Transmitter	+44 dBm
Comm Rx Ant Gain	0 dBi
Path Loss	-39 db
Jammer Tx Ant Gain	0 dbi
	+ 5 dbm

III Jammer transmitter power necessary to provide 10 dB J/S at DTOC communication receiver

Jam Pwr at DTOC	-17 dbm	(-27+10)
Comm Ant. Gain	0 dBi	
Path Loss	-(-72 db)	
Jammer Ant Gain	0 dbi	
BW Disadvantage	-(- 3 db)	
Jammer Pwr	+58 dbm	(631 w)

It appears as though it would be quite simple to jam such a communication link. Clearly some ECCM is will be required to maintain communications.

ECCM Techniques

Directional Antennas......Directional transmit and receive antennas would be an excellent ECCM technique. There use would improve the communication signal power while attenuating the jammer signal power.

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

A problem arises because of the size of directional antennas at VHF and because of the time required for alignment of transmit an receive antennas. The need for rapid deployment and mobility will mitigate against any antenna which requires link alignment. Where directional antennas are used they are designed as a compromise between signal improvement and alignment, 30 degrees 3 dB beam with a 10 to 12 dB gain is typical.

It is quite possible that the jammer will also use directional antennas. If so this would diminish the advantage gained by directional communication antennas.

Interference Cancellation Equipment.....Interference cancellation equipment (ICE), uses the signals from two antennas and balances them in phase and amplitude to create a notch for signals of a particular frequency and angle of incidence. The accuracy of the balance require for various null depths is shown in Table III.

Null Depth
as a Function of
Phase and Amplitude Balance
Table 5

Null Depth dB	Amplitude error dB	Phase error Deg
20	1.7	11.4
25	1.	6.4
30	.55	3.6
35	.3	2.
40	.17	1.1

The equipment is separate and functionally independent of the radio. Typical equipment uses a two step operation: 1) tune to the transmitter frequency and 2) (with transmitter off) Null out the strongest signal present in the pass band. The total operation probably takes less than 0.5 seconds.

These devices claim to provide up to a 40 dB null. In view of the .17 dB or 1.1 degree nulling requirement is seems optimistic to expect this level of performance routinely. A null of 20 dB is more reasonable. Use of the ICE capable of a 20 dB null would require an increase in jammer power of about 20 db, from 320 watts to 32,000 watts.

Frequency Hopping......One long existing means of ECCM is to merely change link frequencies if the link is jammed; moving the link "out from under" the jammer. Frequency hopping carries this technique to the logical extreme by subsequently moving the link to a series of prearranged frequencies distributed over some portion of the band.

The value of frequency hopping is dependent on the speed at which it is accomplished. To be of value it must be substantially faster then the "look through" period of the jammer. Frequency hopping system are generally classified according to hop speed as follows:

Designation	Hop Rate	Dwell Time
Slow (SFH)	100 hops/sec	10,000 microsec
Fast (FFH)	1000 hops/sec	1,000 microsec.

Since the look through and tune time of barrage is typically several seconds even a SFH is a very effective ECCM.

A frequency hopping communication link can be attacked using a barrage jammer if the jammer noise band is expanded to cover all link frequencies. The jammer must either know all of the frequencies in the sequence and designate spot noise jammers to each of them or it must assume that the jammer can appear any where and jam all possible channels all of the time. The jammer power requirement increases in proportion;

Jammer power increase = 10*Log(number of frequencies).

If it is assumed that the communication link can appear any where in the VHF band; the corresponding band width disadvantage to the jammer would be:

$$10 \cdot \text{Log}((88000-30000)/25) = 47.6 \text{ dB},$$

bringing the jammer power requirement to 103 dBm (20 MW). This may be possible but there are at least two problems, 1) such a jammer would block all VHF communication, friend and foe and 2) the jammer would be a target for an anti-radiation missile.

Replica Jammer Application

Multi-Frequency Spot Noise.....Using a DRFM it is possible to receive and store a sample of each frequency used by the communication link. A jamming signal comprised of those stored frequencies could be read from memory continuously and transmitted by a single transmitter; adding noise modulation if desirable. This mode of operation is indicated in Figure 6.

The each DRFM output cycles is a collection of replicas of the communciation link frequencies time multiplexed together. The DRFM cycle time is less then the hop dwell so that jamming is present part of the dwell time on each hop even though the DRFM cannon predict which frequency will occur next. Narrow band noise may be added to the DRFM transmission.


```
                        | | |
                        +-+-+
                          |
                          |
                      +---+---+
     +----------------+Switch +-------------------+
     |                +-------+                   |
     |                                            |
     |                                            |
     |          +---------+      +-----+          |
     +----------+ DRFM    +------+ AMP +----------+
                +---+-----+      +-----+
                    |
           +--------+-------+
           |Memory Control  |
           +----------------+

 Comm Freq     AAAAAAAABBBBBBBBBCCCCCCCCCDDDDDDDDDEEEEEEEEEAAAA
 DRFM Storage  A       B        C         D        E
 DRFM Output   A    A    AB   AB   ABC  ABCD ABCD  ABCDEABCDEAB
 Jamming       -    -    -    -    -    -    -     -         -
  on Freq
```

Multi-frequency Spot Noise Against a Frequency Hopper
Figure 2

The effect of being able transmit on the precise frequencies that will be use by the frequency hopper is to reduce jammer band width disadvantage and transmitter power.

$$\text{Band width disadvantage} = 10 \cdot \log(\text{number of frequencies}) = 10 \cdot \text{Log}(100) = 20 \text{ dB}$$

Bringing the Jammer power requirement to 32. kW. Note that the tuning time required for the ICE system precludes the combination of ICE and Hop.

A second possible application of the replica jammer is possible if the hop rate is slow enough to allow a frequency by frequency intercept and replica generation.

In the hypothetical situation the transit time for the communication link is 239 microseconds while that for the jammer (communication transmitter to jammer plus jammer to communication receiver is 366 microseconds. If the frequency dwell time is greater than 127 Microseconds plus the DRFM response time it will be possible for the jammer to keep up with the frequency hopping communication link, and respond with a jamming pulse exactly on frequency. This reduces the band width disadvantage to 0 dB.

Because of path length differences the communication gets from BTOC to DTOC before the jamming signal. Until the jamming signal arrives, the communication link is jam free. If the hopper is fast enough it can change frequencies to F2, abandoning F1, before the F1 jammer signal arrives. In this case an 8 kHz hop rate would be required.

Developmental Systems

The ECCM techniques mentioned above are currently available as MIL qualified equipment. Of the systems now in development spread spectrum is the most promising.

Spread Spectrum

In spread spectrum, each bit of a digital transmission is represented by binary code of length n. The receiver is responsive only to predetermined set of binary codes and rejects others by the processing gain of 20*Log(n). Spread spectrum has the disadvantage of requiring n times the band width of the simple binary phase shift system transmitting the same data rate. Transmission of a 16 kbps data over such a system using 15 bit code sequences would require a band width of 300 kHz and would provide an additional margin against jamming equal to the processing gain of 24 dB.

Replica Jamming Application

Phase Coherent Storage.....The DRFM has an application against spread spectrum signals because it has the ability to store and reproduce the entire code word including the phase modulation content. The retransmitted signal will contain the phase code sequences which match the receiver correlator. The likely response of the receiver will be the introduction of bit errors or loss of synchronization.

Conclusion

Each of the ECCM measures considered can be considered to add to the power requirements for effective jamming. Taken together, a communication system which incorporated a combination of ECCM techniques would create a jam proof communication link at least one beyond the capability of a conventional barrage jammer.

Application of the DRFM can reduce the effectiveness of many of these ECCM techniques. These are summarized in Table 6 below:

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

Impact of DRFM on Jammer Power Requirements

Table 6

Factor	Impact On Jammer Power Barrage	DRFM
Baseline Jammer Power (Determined by Path Loss)	55 dBm	55 dBm
Band width Disadvantage (Set on accuracy)	3 dB	0
Directional Antennas	22 dBi	22 dBi
(Counter by Jammer Ant)	-11 dBi	-11 dBi
Frequency Hopping (100 Freq in VHF band)	46 dB	20
Spread Spectrum (20 bit code)	23 dB	0
Jammer Power Required (re Comm Power)	150 dBm	86

Without the DRFM communication jamming is probably doomed by the advent of combined Frequency Hopping and Spread Spectrum communication links. The impact of the DRFM is clearly to make communication not only possible but practical for the foreseeable future.

Part II
Digital RF Memory Implementation

DRFM Implementation

The DRFM is implemented for radar ECM is shown in Figure 1. Conversion of the RF signal to inphase and quadrature base band signals is accomplished by a QIFM (quadrature IF mixer). These signals are digitized by the A/D converters and stored in the MEMORY. Replicas of the RF signal are generated by reading from the MEMORY converting back to analog in the D/A converter, equalizing and filtering the resulting signal in the EQLR/FLTR. The result is recovery of the analog in phase and quadrature base band signals. These signals are then up converted using the quadrature phase modulator QPM to recover a replica of the RF signal. Up conversion and down conversion are accomplished relative to the local oscillator LO.

```
+----------------------------------------------------------------------+
|                                                                      |
|                          +----+                                      |
|          +---------------+ LO +----------------------+               |
|          |               +----+                      |               |
|          |                                           |               |
|        +-+---+  +---+  +--------+  +---+  +-----+  +-+---+           |
|        |     +--+   |  +--+      +--+   |  +--+  +--+     |           |
| RF     |     |  |   |  |         |  |   |  |EQLR|  |     |       RF  |
|SIGNAL--+QIFM |  |A/D|  |MEMORY   |  |D/A|  |FLTR|  |QPM  +---SIGNAL  |
|INPUT   |     |  |   |  |         |  |   |  |    |  |     |    OUTPUT |
|        |     +--+   |  +--+      +--+   |  +--+  +--+     |           |
|        +-----+  +---+  +---+----+  +---+  +-----+  +-----+           |
|                            |                                          |
|                 +----------+-----------+                              |
|                 |  CONTROL CIRCUITS    |                              |
|                 +----------+-----------+                              |
|                            |                                          |
|                            |                                          |
|                     USER INTERFACE                                   |
|                                                                      |
|                  DRFM Functional Diagram                             |
|                         Figure 3                                     |
+----------------------------------------------------------------------+
```

Analog to Digital Conversion

One of the important questions that arises in DRFM discussions is the impact of Analog to Digital conversion speed and quantization levels on system performance. There are a number of options in the A/D and D/A conversion but the most popular option is comprised of a quadrature down conversion 1 bit A/D. The popularity of this combination is a result of cost effectiveness. It provides the frequency and phase accuracy required for acceptance by modern spread spectrum receivers and at the same time the D/A technique is simple, can be made to operate at the highest speeds and requires minimum storage. The waveforms in this converter are illustrated in Figure 4.

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

```
+----------------------------------------------------------------------+
|                                                                      |
|       1   2   3   4   5   6   7   8   9   10  11  12  13  14  15     |
| Clk   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |      |
|                                                                      |
|            -------            -------            ------             |
| I Sig    /       \    ------ /       \    ------ /                   |
| Limited                                                             |
|               -------          ------          ----                  |
| Q Sig        /       \        /      \        /                     |
| Limited-----          --------        -------                       |
|                                                                      |
| I Sig         _______        __________        ______               |
| Clocked______|       |______|          |______|      |____          |
|                                                                      |
| Q Sig     _______        __________        _______                  |
| Clocked__|       |______|          |______|       |______|          |
|                                                                      |
| Phase|0   |90  |180|270| 0 |90   |180   |270| 0 | 90|180|270|0  |   |
|                                                                      |
| Average Frequency = del Phase/del time = 3 sig cycle/15 clk cycle   |
|                   = 0.2 * Clock Frequency                            |
|                                                                      |
|              Analog to Digital Converter Waveforms                   |
|                            Figure 4                                  |
+----------------------------------------------------------------------+
```

Spurious Responses

The one bit A/D conversion process is functionally equivalent to a hard limiter followed by a sample and hold. The result of hard limiting is that while phase and frequency data are retained all amplitude data is lost. The hard limiting generates odd harmonics corresponding to square wave harmonics. These harmonics intermodulate with the sampling clock to generate multiple spurious responses. The Fourier transform G(f) for the output of the quadrature 1 bit DRFM is given by the relationship:

$$G(f)=\mathrm{Sin}(pi*fs/fc)/(pi*fs/fc) * \sum_{n=-inf}^{n=+inf} \{F[j(n*fs-m*fc)]\}$$

where n=1,3,5,..; +/-m=1,2,3,...; fs is base band signal frequency; and fc = clock frequency, F represents Fourier transform.

The phase relationships in a quadrature system cause the successive odd harmonics to alternate above and below the LO as shown in Figure 5.

Frequency Accuracy

The effect of alising between base band signal and sample clock on the steady state DRFM response are confined to these intermodulation products. The alising cause a "edge jitter appearance to the I and Q converter waveforms but contrary to intuition there is no resultant error in steady state frequency.

Phase Accuracy

The phase quantization levels of the quadrature DRFM are 0, 90, 180 and 270 degrees. Applying the common approximations 1) the phase errors are uniformly distributed and 2)have a peak value 0/- 45 degrees. The RMS phase error is then:

$$E(ph)=45/SQR(3)=26 \text{ degrees, rms}$$

Output Spectrum

The typical frequency spectrum of the DRFM response is shown in Figure 5.

```
+------------------------------------------------------------------+
|                                                                  |
|                                  |                               |
|                                  |   X                           |
|                                  |   X                           |
|                         X        |   X                           |
|                         X        |   X             X             |
|    X                    X        |   X             X             |
| +-X---+-----+-----X-----+-----+---X-+-----+-----+-X---+-----+    |
|   |   |                 |        |   |             |             |
|   | LO-C/2              |        LO  |             | LO+C/2      |
|   |                     |            |             |             |
|  7th                   3rd          Fund          5th            |
| -15 dBc               -9.5 dBc      0 dBc        -13 dBc         |
|                                                                  |
|                  Theoretical DRFM Output Spectrum                |
|                             Figure 5                             |
+------------------------------------------------------------------+
```

Performance

The performance parameters of a DRFM vary as a function of the implimentation. Typical DRFM parameters for a one bit quadrature system are given in Table 5 below.

R.F. Expo*Disneyland Hotel*Anaheim*California*January 23-25, 1985

Typical DRFM Performance Parameters
Table 5

Parameter	Limitation	Typical Performance
INPUT		
Center Frequency	Mixer BW Lo Tuning Range	Fixed to Octave Band
Instantaneous BW	A/D Conv. Rate	1 Bit 500 Mhz 6 bit 100 Mhz
Dynamic Range	Crystal Video	-20 to 0 dBm
Multiple Signals	A/D Conv	1 bit I/Q 6 dB Capture 6 bit I/Q Linear
MEMORY		
Memory Permanence	continuous power	indefinite
Memory Duration	RAM size, power	250 microsec
Propagation Delay RF in to RF Out Pulse to RF Out	Physical Size	 40 ns 50 ns
Delay Resolution	Mux Structure	10-50 ns
OUTPUT		
Frequency Accuracy	LO Stability Pulse Width Storage Time	10 Hz
Phase Accuracy	A/D Precision	1 Bit I/Q 26. deg rms 6 Bit I/Q 3.3 deg rms
Power Output	LO Mixer Levels	-10 dBm

Spur/Phase Performance	System Type	Spurs	Phase Tracking
	2 Phase System (90 deg. res.)	-9.5 dB	26. deg
	Six phase System (30 deg. res.)	17.3 dB	10. deg
	2 Ph/6 Bit System (2.1 deg. res.)	44.0 dB	3.3 deg

Conclusions

The emerging ECCM techniques available to communication system designers will make today's barrage jammers obsolete in The Digital RF Memory with its very short memory access times, frequency accuracy and ability to handle several frequency intercepts nearly simultaneously make it the key to needed performance improvements in communication jamming.

CHAMPAGNE MEASUREMENTS ON A BEER BUDGET

or

FCC REQUIRED MEASUREMENTS FOR PENNIES

by

Jim Weir
VP Engineering, Radio Systems Technology (RST)
MSEE Plodder, California State University, Sacramento

RADIO SYSTEMS TECHNOLOGY
13281 GRASS VALLEY AVENUE
GRASS VALLEY, CA 95945
(916) 272-2203 123.3MHz

AUTHOR'S FOREWORD

Yes, I too have been triple-integrated by the best of the hi-falutin' technical journals. Doggoned if I can figure out what all that Greek means in the real world either. I neither apologize for nor regret my choice of presenting the concepts in this article in a form decipherable by any competent History major.

A. ABSTRACT

All RF communication devices must be tested to meet FCC leakage and spurious emission limits. This paper will show how to construct laboratory and field measurement devices to test these limits using both in-house ("home-brew") and inexpensive commercial instrumentation. Construction details of an FCC-approved far-field outdoor range will be shown.

B. GENERAL COMMENTS

In the beginning was the spark gap transmitter, and in the ensuing hullabaloo about its interference with broadcast reception was begat the vacuum tube transmitter, and with its accompanying unwanted second harmonic output was begat the pi-network filter, and with its excellent performance was begat the requirement that all our electromagnetic emitters (whether they be disguised as local oscillators, clocks, or true transmitters) be purged of any and all spurious or harmonic trash that may interfere with our neighbor's quiet enjoyment of the electromagnetic spectrum. The bottom line is that the federal government (in the corpus of the Federal Communications Commission - FCC) has made demands upon the designers of RF equipment to competently design and then test that designed equipment to make sure that it will not interfere with other users of the RF spectrum

Now, this is written for those of you who believe that RF is a rather simple discipline surrounded by a plethora of complex subtleties. For those of you who believe that RF is a deep, dark art presided over by the high priests of black magic, I suggest that you dip your toes in my humble brand of simplistic design and measurement. And, for those of you who cling to the belief that RF is a design art at the fringe of science, I invite your indulgence for a few moments.

For those of us who profess the RF discipline, the end results of our efforts is one of two effects: our designs will heat or they will communicate. In either case the basic process is the same — we generate RF power and allow it to radiate into a dielectric. In the first effort the thermal dissipation into a bone or tissue is the desired effect, and in the second, we hope that at the end of a long transmission path, there is enough energy left to decode some transmitted intelligence. ("LaVerne and Shirley" notwithstanding.)

No matter whether we are designing diathermy machines or deep space data links, we have one common problem: how to keep the signals I generate for MY job to keep from messing up the machine you generate for your customer. This is not trivially simple; in my designs there are oscillators that may be spot on frequency on the RF data channel that you have chosen for your information link.

A classic example of unwanted interference used to happen when a passenger decided to play a little FM radio music on board an aircraft. Take, for example, an aircraft approaching Los Angeles where the pilot is navigating to LAX by means of the radionavigation station on 113.6 MHz. The passenger decides to tune in "Rockin' Radio 103" on frequency 102.9 MHz on his personal FM radio that he always keeps in his briefcase. What he is really doing is tuning the local oscillator of that FM radio 10.7 MHz (the IF frequency) above his desired receive frequency. Unfortunately, in this case the local oscillator will be tuned to 102.9 + 10.7 = 113.6 MHz. Oops. The passenger's little "transmitter" (local oscillator) radiated directly into the aircraft navigation antenna, the navigation needles in the cockpit went crazy, the autopilot nearly performed an inverted snap roll trying to follow the needles, and the pilot was placed in what we lovingly refer to as a "character building situation".

The Feds thought that this wasn't a particularly good idea, so some years ago the FCC and the electronics industry sat down and came up with some pretty good rules and regulations concerning how much unwanted radiation could be emitted by a radio device. If the radio device was a receiver, then the regulation required "certification", whereas a transmitter was either "type accepted" or "type approved". This paper will not treat type approval because this method of meeting the regulations require that you send a working model of your proposed device to the FCC Labs in Maryland for them to test in their facility. Certification and type acceptance, though, is done in the manufacturer's labs and only the results are sent to the FCC Labs for their perusal. It is certification and type acceptance that will be the focus of this paper.

C. A BRIEF OVERVIEW OF THE SITUATION

The FCC rules provide a fairly clear delineation of the type and style of test that must be performed on any particular piece of equipment. There have been books written telling just how and why certain tests are performed [1,2]. It is NOT the intent of this paper to show step-by-step procedures for these various tests. Instead, this paper will show how to obtain or build certain pieces of test equipment and laboratory apparatus that are common to almost all FCC tests.

1. CERTIFICATION

This procedure is usually dictated for devices that aren't supposed to radiate, or are only supposed to radiate miniscule amounts of power. The FCC regulation that covers certification is part 15 of the FCC rules (47 CFR part 15), and even the title "Incidental and Restricted Radiation Devices" gives some idea of the scope of this chapter.

In this part, you will find the requirements for measurement of superheterodyne receivers, wireless microphones, cordless telephones, and other low powered devices. In addition to part 15, FCC rules part 2 is the guide to how your radio has to be marked, how and when you can advertize it for sale, and what kinds of data must be submitted for a certification request [3,4].

2. TYPE ACCEPTANCE

This procedure is meant for transmitters -- devices that are INTENDED to be radiators. The type acceptance regulations are designed to insure that a transmitter does not spew out too much garbage and gunk on frequencies other than that for which the transmitter is designed.

The pertinent section of the FCC regulations regarding the technical specifications that your transmitter has to meet is really determined by the particular service for which your transmitter is intended. For example, if you are designing a taxicab radio, your transmitter has to meet the specifications of part 90. Part 83 covers shipboard radios, part 87 details the requirements for aircraft transmitters, and so forth. As with certification, part 2 of the FCC rules give the guidelines on format of the reported data, how to mark the transmitter, and other "housekeeping" type rules [5].

3. SITE CERTIFICATION

Either of the two testing procedures, certification and type acceptance, will require the use of a site for far-field testing. Basically, this is nothing more than an open-field "pasture" type of test, where the device to be tested is placed a known accurate distance away from a calibrated receiver and calibrated antenna, and the levels of spurious emissions from the device are measured on the calibrated setup.

The real problem comes in when this far-field "range" is first selected, as several fairly critical tests must be made on the antenna range itself before any new equipment measurements can be made. Not only that, but the site itself needs to be certified with the FCC Labs before you can use the range for certification or type acceptance testing (6,7,8).

D. EQUIPMENT NEEDED, BOTH LAB AND FAR-FIELD MEASUREMENTS

First of all, I started this paper with the title implying that you could do these measurements on a shoestring. I may have misled you, but only slightly. If you sent one piece of equipment to a lab that specialized in making FCC measurements, you can bet that the bill for a simple certification will run upwards of $1500 and up to $5K to $6K for a moderately complex transceiver. Those were estimates I received in 1976 for a small aircraft band transceiver, and while I haven't checked recently, I seriously doubt if the price has come down.

The bill for equipment to set up your own lab will be somewhere between $2K and $5K, depending upon how good a scrounger you are. This is roughly the same amount you would have to pay for an outside lab's tests, but once you have purchased the equipment, subsequent tests are literally done for pennies worth of expendables.

1. PURCHASED EQUIPMENT

There are a few pieces of "store-boughten" lab gear that are absolutely essential to this process; almost everything else can be done "homebrew" style. Here is a list of the equipment that you will need to buy and the approximate cost thereof:

a. Spectrum analyzer. The frequency range need only go up to 1000 MHz for receivers (or twice the L.O. frequency, whichever is greater), but for transmitters, you need to be able to search the spectrum up to the 10th harmonic of the output frequency. By using a communications receiver (see c, below), the analyzer only needs a low frequency response of 30 MHz, but for fast and sure searching of the low frequencies, a response down to 0.5 MHz is preferred. These analyzers show up on the used equipment market on a quite regular basis and I have seen them priced from under $1000 for a real dog up to $2000 or so for a pretty nice piece of equipment. It's not going to be an HP 8558, but for $2K you couldn't even buy the line cord for that rascal.

b. Signal Generator. No doubt about it, the used HP 608 is available, relatively cheap ($500-800 for a calibrated unit), and reliable unto the last glow of its 20-odd tubes. The best thing about this old boat anchor is that its calibrated attenuator is so accurate that you can use it as a standard by which to measure all manner and form of signal levels. In particular, you can measure the absolute strength of a signal on your spectrum analyzer to an accuracy of a dB or so, and in this game, that is accuracy enough.

c. Communications Receiver. A good general coverage receiver with an S-meter is a sort of a luxury that you ought to have if you possibly can afford one. While it is not an ABSOLUTE necessity (the spectrum analyzer can be pressed into service), it makes life a whole bunch easier. You can usually find them in the "used ham equipment" stores for less than $100, or you can splurge on one of the new nifty digital readout receivers for $300 or so from the "Hobby Shack" store.

d. Power Meter. The old HP 430 is available on the surplus market for less than $200 with bolometer mount, or for less than $50 if you want to try to rig up your own thermistors. HP431's are on the market for less than $300 with mount, and a.; more stable over temperature than the older 430 series. It is possible to make your own power met.r, but even the older HPs will go down to -20 dBm, and that is a real trick on the homebrew bench.

e. Frequency Counter. With the advent of digital circuits, the price of frequency counters has fallen to the point that even a new 1 GHz counter, with readout to 1 Hz, a temperature compensated crystal timebase, and accuracy 10 times better than required is available from several sources for less than $200. It will probably not benefit you to shop this item used or surplus as the new breed has not been around long enough to have made it to the surplus market.

f. A source of 110 vac power other than the commercial power lines. The point being, if you use commercial power for the far-field tests, the cable to your equipment from the wall socket has to be buried, you have to use a line filter, and all this has to remain absolutely constant from the time you certify your site through every measurement you make on the site. If you use a portable gas generator, or use a transistor inverter with a storage battery, none of these restrictions apply. Besides which, a 300-400' extension cord isn't all that inexpensive. You can always use the generator when you go camping, too.

g. Miscellaneous: Pads (fixed attenuators), coax cable, connectors, small bits and pieces; the whole lot should come to less than $50.

2. HOMEBREW EQUIPMENT

Now we have the store-boughten pieces all accounted for, it is time to take a look at the homebrew equipment you will need for both the lab and far-field tests:

a. The first thing to homebrew is a far-field site for the tests. The Feds say that the site should be level, with rich dark soil, and with short-clipped grass in all directions. What I've settled for is a site that has a constant slope, soil that won't grow rocks, and with native weeds as far as the eye can see. Actually, the site is the emergency runoff area of the local airport, and if you can talk the airport manager into letting you use an unused portion of your local airport, you couldn't ask for a more perfect site. First, you will probably pay no rent. Second, the site will probably stay weeds and grass as long as the airport remains. Third, the only interference will be the occasional takeoff and landing of a private airplane; the site is guaranteed to be free of most electromagnetic sources strictly because of the prohibition of antennas around an airport. EMPHASIZE to the airport manager that you are running some very SHORT tests that will involve very SMALL amounts of power and that your antennas will be up for a very SHORT period of time and that you will be glad to monitor the local airport frequency and discontinue your tests as soon as the word "EMERGENCY" is uttered. Trust me, tear down the masts, leave the prototype, and interrupt anything that is happening if that magic word is said. (See Figure 1.)

b. Antenna masts and mounts. This is the first homemade piece of gear to be made, and really the easiest technically (although the manual labor is somewhat strenuous). A Douglas fir 4x4, 16 feet in length, is a common building timber and is a relatively cheap ($15) component of the antenna mounting system. The rules say you have to get your test and reference antenna systems 3 meters apart (some procedures call for 100 meter separation) and 4 meters above the ground, so 16 foot masts will provide a good margin. If you can plane or shave the base of your 4x4 into a round form factor, it will fit quite nicely into a 3 1/2" piece of PVC pipe stuck into a concrete base in the ground. The 4x4 masts should have "shear cuts" made in the base of the mast on the off chance that an aircraft loses power during one of your tests and crashes into the antenna range. The 4x4's will shear off at the shear cuts, and will not unduly damage the aircraft or passengers. Dry and then varnish the 4x4 thoroughly to ensure that the mast will be as much an insulator as possible. (See Figure 2.)

You will also have to make two antenna mounts out of plywood. One of the mounts is meant to sit on the top of the 4 meter "equipment under test" mast, and the other one is a sleeve that will have to go from 1 meter above the ground to 4 meters above the ground. This sleeve is pulled up and down the "spectrum analyzer" mast with waxed twine and ceramic pulleys mounted at the top of the 4 meter mast. Both the dipole and loop antennas (see below) fasten to the mounts with plastic cable clamps and nylon screws and nuts. To the extent possible, fasten the plywood together with glue; use nails only as a means of keeping the glue joint tight until it dries, or use clamps. You do not want any metal in the antenna structure at all, except the antenna, balun, and coaxial cable.

The sleeve mount will only have one cable clamp holding the antenna onto the mount, as this antenna will have to rotate from horizontal to vertical polarization (and everywhere in between) in order to maximize any spurious emission. The equipment mount should have 2 clamps in order to maintain a horizontal antenna at all times.

c. The test antennas are the crux of the homebrew part of this setup. The far-field test measurements are confined to two major portions of the spectrum: 0.1-25 MHz, and 25-1000 MHz.

1. The frequency range 25-1000 MHz is concerned with the E-field measurements from 12 - 0.3 meters wavelength. This spectrum can be measured with dipole antennas and coaxial baluns. The dipole supports are made from plain old PVC thinwall water pipe and PVC fittings. (Point of information -- get the white pipe. The black pipe is made so with carbon black, and is a fairly good conductor.) The antenna elements themselves are made from copper foil tape, which is available at most office supply stores and stained glass window works. If you use the 1 cm wide tape, it will fit quite nicely on "1/2 inch" water pipe. (See Figure 3.) The balun to be used between the signal source/spectrum analyzer and the dipole is of a variety developed by W.K. Roberts of the FCC Labs in Laurel, Maryland. These baluns, if made from RG58 coaxial cable, have measured losses in the range .1 to .5 dB, and have proven to be absolutely flawless in operation. (See Figure 4.)

It so turns out that 4 baluns will cover the frequency range 20-1000 MHz, so four separate PVC antenna mounts were fabricated, and each one marked with permanent marker with a calculated resonant frequency in Megahertz, every 5 MHz. The lowest frequency antenna (covering 25-65 MHz.) is quite long, and the horizontal portion of the dipole was given support with monofilament fishing line lashed to the vertical mast. Small holes drilled in the PVC pipe allowed us to fasten the monofilament to the pipe.

The copper tape is fairly cheap, so instead of having a dipole with adjustable arms, the dipole is rigged with a full length of copper tape, and then small lengths of tape are razor-bladed off each end of the dipole to tune it to a particular frequency. The actual practice is that the device you are going to test is probed very carefully in the laboratory, with a non-resonant "whip" type of antenna, or a very small loop made from coaxial cable, either of which is placed VERY close (< 10 cm) from the device. If it is a transmitter that you are testing, then the output of the device into a 50 ohm load is also observed on the spectrum analyzer. With these two measurements, you will have almost certainly identified any spurious emissions from the device, and you will have a very good idea on the far-field tests what frequencies you should be most concerned with.

Of course, the rules say that ALL frequencies must be examined, so you should not just clip the antennas to the resonant frequencies you have identified in the close-field examinations, but a reasonable step from one antenna resonant frequency to another is quite justifiable -- say, in steps of 5 MHz at the lower frequencies and 20 MHz at the upper limits of your search. When you approach one of your pre-identified spurious frequencies, THEN you can start splitting hairs.

2. The frequency range 0.1 - 25 MHz may be done with a loop antenna, if you prefer. Measurements below 18 MHz MUST be done with a loop. The problem is that the loops you can buy today are both clumsy to use and expensive. The clumsiness comes from the fact that internal ferrite or iron core transformers are used to match the high-impedance loop to the 50 ohm input of most receivers and/or spectrum analyzers, and I strongly suspect that this is where the cost comes from, also.

The answer is to build your own loop for these measurements. The only small problem that the builder of a homebrew loop has to deal with is the conversion of the fairly high loop impedance to the 50 ohms of the analyzer. Fortunately, solid state electronics comes to the rescue.

Once again, the loop support is made of 1/2" PVC water pipe and plastic fittings. A loop made of RG8 coaxial cable is fashioned about 0.7 meter (2' or so) in diameter and run through holes drilled in the PVC. The actual diameter is not critical. The braid of the coax is broken for about 1 cm or so at the top of the loop to make a balanced, shielded, one turn loop antenna. (See Figure 5.) The center conductor of the loop is fed to a balanced transistor differential amplifier, and a 50 ohm output port is fed from an emitter follower in turn fed from one collector of the diff amp. (See Figure 6.)

What you have at this point is an uncalibrated loop, which really doesn't do you much good. What you do at this point is RENT a calibrated loop and use your spectrum analyzer or communications receiver and calibrated pads to spot-calibrate your loop against the known loop at several places in the 0.1-25 MHz. spectrum. What do you use for signal sources? The calibration procedure is to place the two loops at the same distance above the ground, one each on one of the 4 meter masts, "point" them in the same direction, then use a small SPDT switch to select one or the other of the loops, and then tune in a signal on the frequency where you want a calibration point. "Signals" may be from any source you choose -- broadcast stations, ham stations, CB stations, overseas broadcast, or any other source that is transmitting on the frequency; if the switching is done fast enough, you will not get any error from fading or atmospheric effects. Switch select between the two antennas and read the difference in dB on the spectrum analyzer, or place pads in the stronger of the two antennas in the case of the communications receiver to make the S-meter readings the same. (See Figure 7.) Store-boughten loop antennas, by the way, are about $1000-$1200; you've just saved about $990. Figure $50 for rental of the known loop and you're STILL $940 ahead.

d. The rules also say that you have to test your radio over temperature. Most requirements are to test from -30 to +50 degrees Centigrade. The classic way to run the temperature up and down is in a very well insulated temperature chamber with liquid CO2 in cylinders to run the temperature down and a resistive heating element to run the temperature up.

There are a few things that I dislike about using commercial chambers. The chamber itself (even on the used market) is a couple of $K, the liquid CO2 is not cheap, the cold valve always hangs up at the wrong time, and the temperature control requires constant fiddling (excuse me, fine tuning) to keep the temperature where you want it.

Build a 2-chamber box out of thick plywood, with regular home-type insulation covering all the walls, and a small hole in the top of the chamber that is fitted with a cork and a $5 mercury thermometer. The partition between the top and the bottom of this box has a small muffin fan bolted so that you can push the air from the bottom of this box up into the top, with a switch mounted so that you can turn the fan on and off at will. The bottom chamber of this box is insulated equally as well as the top; opening the hinged front door of this box exposes the interior of both chambers. The bottom chamber is also fitted with a light bulb socket that is wired so that it can be turned on and off from outside the box. Drill a few 2 cm (1") holes in the door and buy enough corks to fill those holes. Run your coax, power cables, etc. through holes drilled in the corks, and then seal the gap between cork holes and cables with RTV or silicone sealant. (See Figure 8.)

If you want cold, go down to your local dairy and buy enough dry ice (solid CO2) to fill the bottom chamber of the box. Put your device under test into the top chamber and run the fan until the thermometer registers the cold temperature you want. Modulate the fan switch to maintain your desired temperature for as long as you want your device to soak before taking a measurement (I usually allow half an hour per kilogram of mass). Remember, cold air sinks, so your thermometer probe ought to be sitting at the same vertical level as your device under test. I honestly don't know how cold this chamber will get, but I do know I forgot and left the fan on and when I got back from lunch, the mercury was frozen solid (mercury freezes at -40 degrees). When you are getting up from -30 to temperatures around +10 or so, you will have to remove almost all of the dry ice to make the chamber warm enough. Even without the fan, if the bottom chamber is almost full of dry ice, the chamber will take DAYS to come back to room temperature with the fan off.

If you want hot, leave the fan run continuously and modulate the on-off duty cycle of the light bulb. The bigger the light bulb, the faster it will get hot -- just remember that this is a WOODEN box and you would prefer not to charcoal your new radio inside the box. I have NOT forgotten the switch on the hot side, but I have asked for and gotten +70 degrees C with a 100 watt bulb. If you want to keep the bulb from cracking, remove it before you put the dry ice in for the cold test.

There are a few refinements that I would like to make in this setup some day when I get the time: quite a few semiconductor companies will sell an integrated circuit temperature controller, and it would probably save me as many hours in switch-flicking time as it might cost to make. One drawback to this homebrew chamber is that the fan bearings do not like very cold or very hot temperatures, and the fan life is no more than a few hundred hours. Fortunately, the "hobby shack" stores all have these fans for a few bucks, so replacement once a year or so isn't all that painful. It is also possible to substitute one of the new digital voltmeter temperature accessory probes for the glass thermometer, and attach the probe to the largest thermal mass of your radio. This would be superior to the mercury thermometer, with the additional advantage that the digital probe is not likely to freeze.

e. The rules also state that you must use some sort of equipment tables for the far field tests, and that the equipment table that supports the radio that you are testing must have a rotatable surface. This is so that you can rotate the radio at a spurious output for maximum reading.

I have made two tables. One is made from redwood 2x4s with a standard hardboard door as a table surface. I use redwood for the framing because it is very light weight and has very little moisture content, and is hence a poor reflector of radio waves. The hardboard door is also a non-specular surface, is hollow, and is relatively light weight. This table is used to hold the spectrum analyzer, the signal generator, and any other small test equipment. Of course, all joints are glued, not nailed.

The second table is made from redwood with a plywood table top. Into this plywood is drilled a 1" hole, and a second plywood top is also drilled and a 1" dowel glued into the hole. The dowel is sanded so that when it is placed into the hole in the table top, it swivels with little or no friction. The rotatable top is then drilled with small holes very close to the edge on all 4 sides. Waxed twine ("lacing cord") is tied into these small holes and the free end is run over to the equipment table. In this manner, the table is played like a puppet by the test engineer. In matter of fact, the whole proceedings are much like a Punch and Judy show; the table is being rotated by strings, the antenna is being raised and lowered by strings, and the test antenna polarization is being rotated by strings. (See Figure 9.)

E. AN ABBREVIATED DESCRIPTION OF A SITE CERTIFICATION PROCEDURE

In March, 1981, I began the construction of our permanent far-field test site at the Nevada County airpark (better known as Grass Valley Intentional Airpatch). My first stop was the county airport authority for a check of the airport master plan. When the plan showed the west end of the airport as a weeds-and-rocks runoff area well into the year 2000, I began my construction by obtaining an informal approval of my plans from the airport manager. I emphasize -- once you do these tests, you probably don't want to do them again for a while; it does not matter what site you choose as long as it is not planned for a shopping center next year.

Since our procedure called for a 3 meter antenna separation (almost all procedures these days are using 3 meters, and even the old 30 meter documents can be modified to the 3 meter specification), I chose the flattest site with the gullies, trees, fences, and buildings as far away as I could get them. If you are so fortunate as to have a pasture that you can use to put up permanent masts, that is fine, but on airport property, we needed to leave practically no upright obstacles of any kind. (See Figure 10.)

A hole some 1 meter in diameter and 1 meter deep was dug and a PVC sleeve, capped on the bottom with one drain hole drilled, was cast into concrete poured into this hole. The open end of the PVC pipe was left about 2 cm above the surface of the concrete so that a tight fitting cap could be placed over the pipe to keep out water, mice, etc.. While the concrete was still liquid, a scrap 4x4, rounded on the edges to fit into the PVC sleeve, was placed into the PVC and the pipe was made vertical with a bubble level. A second PVC sleeve was then identically cast into concrete with the facing outside walls of the pipes about 3 meters apart, plus an allowance for the thickness of the antenna mounts and the antenna thickness, so that the radiating elements are exactly 3 meters apart. The cement was then allowed to dry for a week. (See Figure 11.)

During that week, we made the masts, the antenna mounts, the antennas, and the equipment tables. If you have been successful in your electronic equipment scrounging, you may wish to purchase a power saw; what takes hours by hand takes seconds with a good radial arm saw.

That week was also spent in calibrating the spectrum analyzer and RF generator that would be used to make the site attenuation tests. The FCC doesn't really care who does the calibration, but whoever does it needs to be able to trace the ultimate calibration to the National Bureau of Standards. In our case, we had just sent the signal generator out for its routine calibration, but the FCC would have accepted a spectrum analyzer calibrated by a signa generator referenced to a power meter that has just been calibrated by a standard battery that was compared to a NBS source.

We were thus ready to perform the tests to calibrate and submit measurements for our site attenuation and site approval so we made a few elementary calculations. The equation for the loss between two antennas is fairly simple:

$$A = 20 \log D + 20 \log f - Ga - Gb + Cc - 20 \log (1+r) -27.56 \quad (dB)$$

where D is the distance in meters between the two antennas, f is the frequency being tested in Megahertz, Ga is the dB gain of antenna "a" (relative to isotropic), Gb is the dB gain of antenna "b", Cc is the loss of the coax cable connecting the signal generator to one antenna, r is the ground reflectivity as a decimal fraction, and A is the attenuation between these two antennas in dB.

Ground reflection varies from about 60% (r=0.6) for super rich soil with a lot of carbon content and tall weeds to 99% (r=0.99) for hardpan soil just short of iron ore. The average seems to be around 70-80%. This entire site certification procedure really boils down to a curve fit of your particular site to find a value of r that fits your attenuation curve. This value is then used forever more in making measurements between these same two antennas at this same location.

You will notice that the cable loss from the spectrum analyzer to the test antenna was not factored into this equation. This is because during these tests, the cable from the spectrum analyzer is disconnected from its antenna and connected directly to the signal generator to find the dB difference between direct and radiated conduction. You should note two things: one, the cable thus used on the spectrum analyzer becomes part of the calibrated equipment list for this antenna range, and two, the measurement is not an absolute measurement, but a differential measurement which is quite a bit more accurate.

Now, given dipoles (G = 2.15 dB), a 3 meter range, and a 75% reflectivity coefficient (r = 0.75), this equation reduces to:

$$A = 20 \log f - 27.2 \text{ dB}$$

Knowing the order of magnitude of what we were looking for made the calibration of the range a two hour job. After all was said and done, we found that the value of reflectivity for our particular range was 73% for best curve fit, and at that value, we had less than 1 dB variation between our measured graph of attenuation versus frequency and a straight line theoretical graph using the above equations. (See Figure 12.)

If we had noticed a large error at one frequency or a group of related frequencies, we would have been forced to either select another site or to find out the source of the error. In this part of the world, it is not at all uncommon to be digging a foundation for a house and find large boulders with high mineral content (gold? !!) or discarded metal tools from the last century.

There is no standard form for submission of these tests, nor is there any formal reply to you of the ACCEPTANCE of your measurements. However, if there is something wrong with the way you did a test, or some question, then you wil most certainly get a letter from the Feds. There is also a division between a site that you plan on using to certify someone else's equipment (in other words, if you are setting up a commercial antenna test range) and a site that is strictly yours for your own equipment. Quite honestly, if my results came out within a dB or so of theoretical, and if I had my equipment recently calibrated, I'd go ahead immediately with my equipment testing and presume the acceptance of my measurements.

One more word and we will leave the subject of the site qualification. That is, the engineer that made the tests needs to be qualified too. There are no formal requirements for education, experience, or background, but you get the feeling talking to these folks that they would rather take the data from a NDE (non-degreed engineer) that had a first 'phone, amateur advanced, and twenty years in the EMI business rather than a newly minted EE PHd with only book-learning for experience. You need to put in to the application a listing of your credentials and experience, and trust me, it is difficult in the extreme to "snow" the engineers at FCC Laurel Labs.

F. AN ABBREVIATED DESCRIPTION OF A CERTIFICATION / TYPE ACCEPTANCE TEST

With the far-field site measurements plotted, the report written and mailed, and the site thus calibrated, the first type acceptance and certification tests were scheduled on a new design, the RST-571 aircraft band transceiver. This transceiver consisted of both a receiver and a transmitter, so both certification (part 15) and type acceptance were required (part 87).

The lab tests for receiver certification involve the measurement of the amount of local oscillator fundamental and harmonics being emitted from the antenna and chassis, as well as the levels of any other oscillators, I.F. signal leakage, or other spurious generators. Part 15 specifies certain IEEE and EIA standard tests that will be accepted, but does not specify any one particular procedure to be followed. In the case of aviation receivers, there is a RTCA (Radio Technical Commission for Aeronautics) procedure that is a special modification of IEEE procedure 187, and that, too is an acceptable way of testing the spurious emissions. (8)

FCC regulation 15.63 defines the limits of spurious radiation, and it is to your advantage to measure the strength of the emissions out of the antenna jack of your receiver BEFORE going over to the pattern range. If the strength of the emission is such that it would not pass the field test, why bother going over and getting your boots muddy? Hint: the second harmonic strength of an antenna jack radiated signal will always be much less than calculated because the antenna used at the device under test is resonant at the fundamental, and therefore anti-resonant at the second harmonic. If you "sniff" around the case in the lab using a short whip or small loop and find radiation not coming from the antenna, then the case or input/output wires are conducting the trash out into the world, and it is almost impossible to tell what the test will come out like when you do a far-field measurement. Fortunately (I say it was good design -- others say the blind pig picked up another acorn) our design was some 15 dB below the limit, and it simply became a matter of going over to the airport and making the tests. (See Figure 13.)

FCC part 87 is the bible for aircraft band transmitters, and part 87 is quite detailed in the tests required, the results required, and the method required to get those results. The far-field tests required are very similar to those required for receivers, and as a matter of fact, were done at the same time that the receiver portion of the transceiver was tested. No use putting those poles up a second time if you don't have to. Again, the second harmonic of the transmitter will be much less than you expect due to the antenna being anti-resonant. We used the leftover dry ice from the temperature measurements to cool the beverages used to celebrate our new product.

G. SOME FINAL COMMENTS

I cannot emphasize too strongly that references 1 and 2 are a take-the-novice-by-the-hand kind of reference work, and make this whole subject almost easy to learn.

The papers put out by the FCC Chief Engineer are also excellent; these are the folks who are going to examine your work telling you how and what they are going to examine. There is no excuse for not having these in your library if you are doing FCC work.

I do hope that my brand of simplistic hand-waving has not offended those of you who demand rigorous equations for every step of your design, but I would rather hope that you will take what I have to say and beef it up as you require.

If the device that you are certifying is meant to be powered from the AC line instead of an automotive or aircraft battery system, then there are certain "conducted" tests that you must perform to show how much radiation you are pumping back into the wall socket. Again, reference 1 shows the manufacture of some extremely simple test equipment to perform these conducted tests.

I should like to dedicate this paper to the memory of an old friend, now a silent key. Wes, W6YSP, was my first teacher in this art, my first contact with the wonderful world of RF, and the first one to take this gawky 15 year old kid and give him a job in the industry. Lord knows what he saw in me, but bless him for seeing it.

-30-

G. FOOTNOTES

1. Willmar K. Roberts, "A Guide To F.C.C. Equipment Authorizations", published by the author, W.K. Roberts - 4637 Van Kleek Dr. -- New Smyrna Beach FL 32069 -- (904) 427-3612.

2. Code of Federal Regulations, 47 CFR Parts 0 to 19, US Government Printing Office, available through USGPO bookstores in most major cities.

3. Office of the Chief Engineer FCC, "Receiver Certification Requirements (Bulletin OCE 24)", June 1976, Office of the Chief Engineer -- Federal Communications Commission -- POB 429 -- Columbia MD 21045.

4. Ibid, "Receiver Radiation Measurement Methods (Bulletin OCE 35)", May 1974.

5. Ibid, "Type Acceptance (Bulletin OCE 15)", May 1979.

6. Ibid, "Calibration Of A Radiation Measurement Site (Bulletin OCE 44)", undated draft copy.

7. "Spectrum Analysis ... Field Strength Measurement (Application Note 150-10)", Sept 1976, Hewlett-Packard, Palo Alto CA.

8. "Modern EMI Measurements (Application Note 63E)", ibid.

FIGURE 1 ------------------------->

A radiation test site at an airport poses certain non-standard hazards to persons conducting tests.

FIGURE 3 ------------------------->

A piece of PVC water pipe, a few plastic fittings, and some copper tape makes a remarkably inexpensive and reliable antenna.

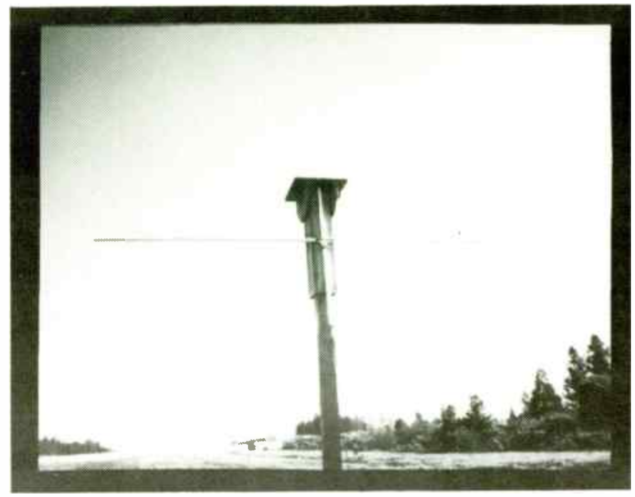


FIGURE 4 (see graph)

FIGURE 2 ------------------------->

A 4x4 mast fits into the PVC pipe sunk into concrete to make a very sturdy antenna mast structure.

FIGURE 5 ------------------------->

A loop antenna mounted on the mast is used for probing the low frequency spectrum.

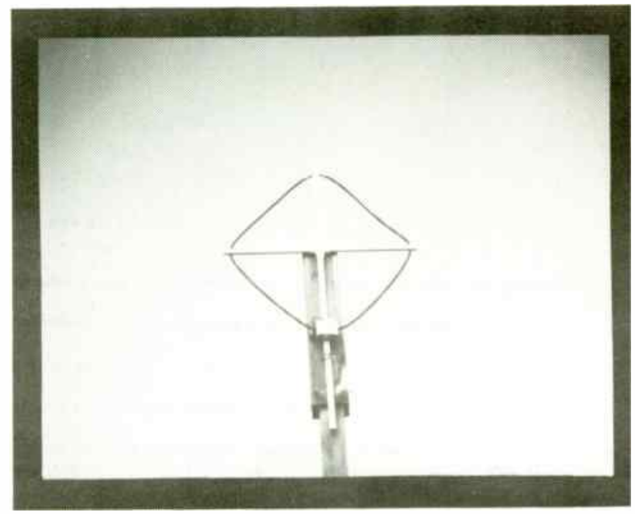


FIGURE 6 (see schematic)

FIGURE 7 (see graph)

FIGURE 8 (see drawing)

FIGURE 9 ------------------------>

The rotatable table is made from thin plywood, doweling, and strings are used to turn the table.

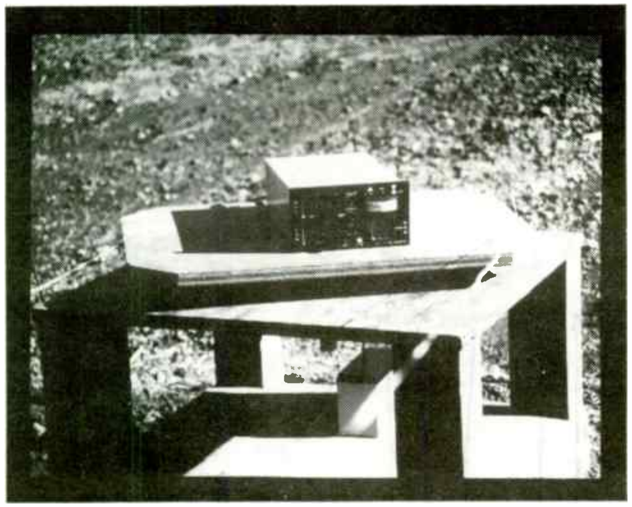


FIGURE 13 ---------------------->

The RST-572 aircraft band transceiver undergoing a type acceptance and certification for field test on the FCC approved site.

FIGURE 10 ---------------------->

An aerial view of the calibrated site. Note the dirt "emergency runoff" area at the west end of the runway.

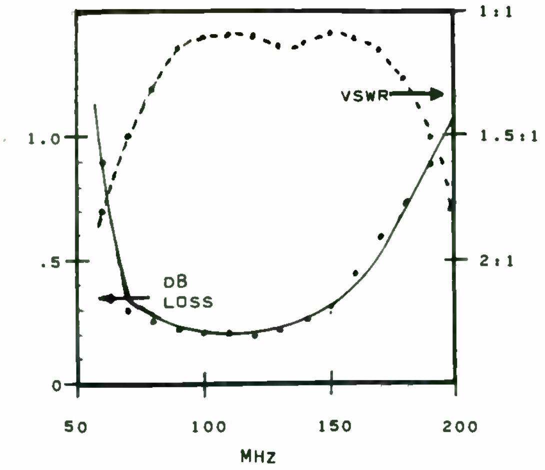
FIGURE 11 (see drawing)

FIGURE 12 (see graph)

NOTES:

1. KEEP MINIMUM SEPARATION BETWEEN COAX SHIELDS WHILE STILL KEEPING THEM FROM TOUCHING.

2. DRILL COAX AND REMOVE CENTER CONDUCTOR AT THIS POINT FROM THE RIGHT. DO NOT ALLOW BRAID TO TOUCH CENTER CONDUCTOR AT THIS POINT.



MEASURED BALUN PERFORMANCE

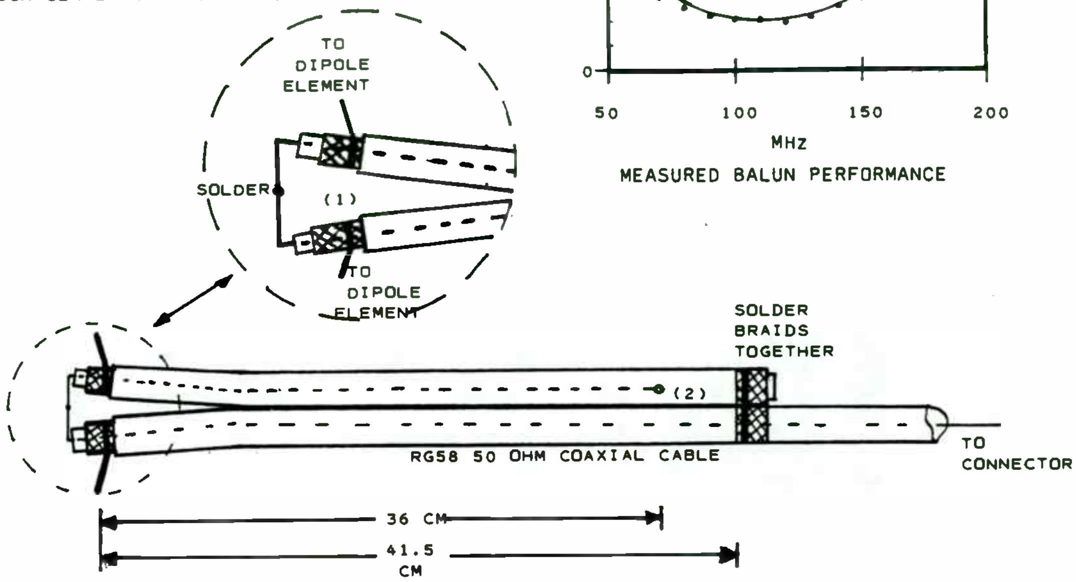
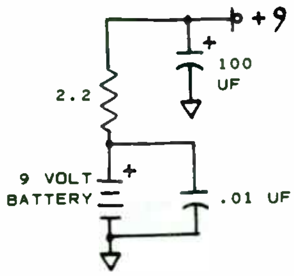


FIGURE 4. CONSTRUCTION DETAILS OF ONE OF THE BALUNS (AFTER ROBERTS).



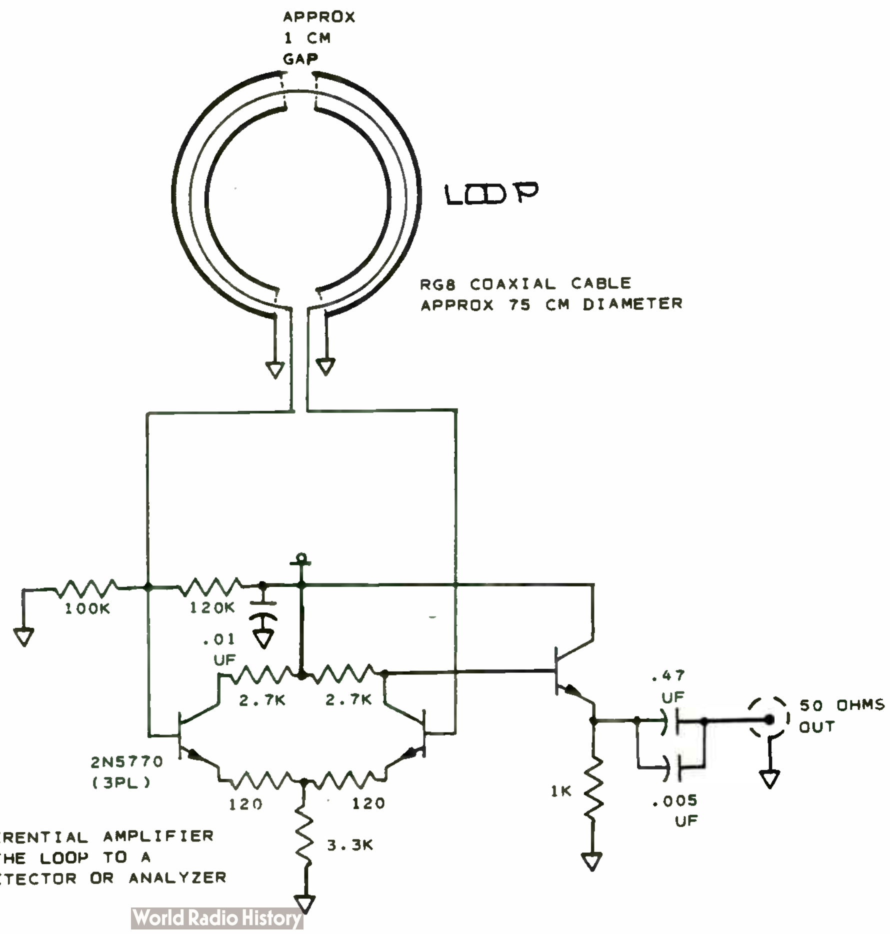


FIGURE 6 THE DIFFERENTIAL AMPLIFIER MATCHES THE LOOP TO A 50 OHM DETECTOR OR ANALYZER

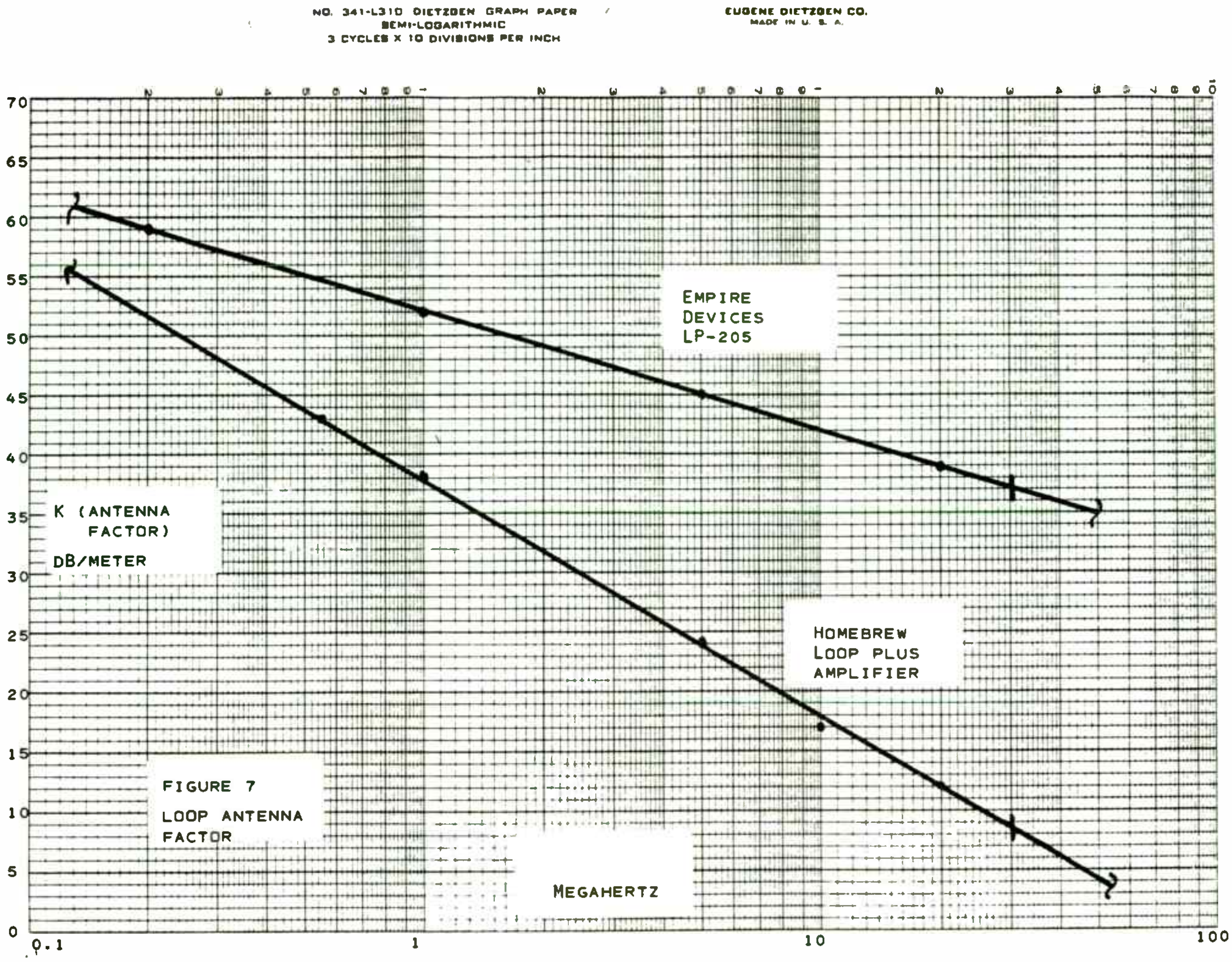


FIGURE 7
LOOP ANTENNA
FACTOR

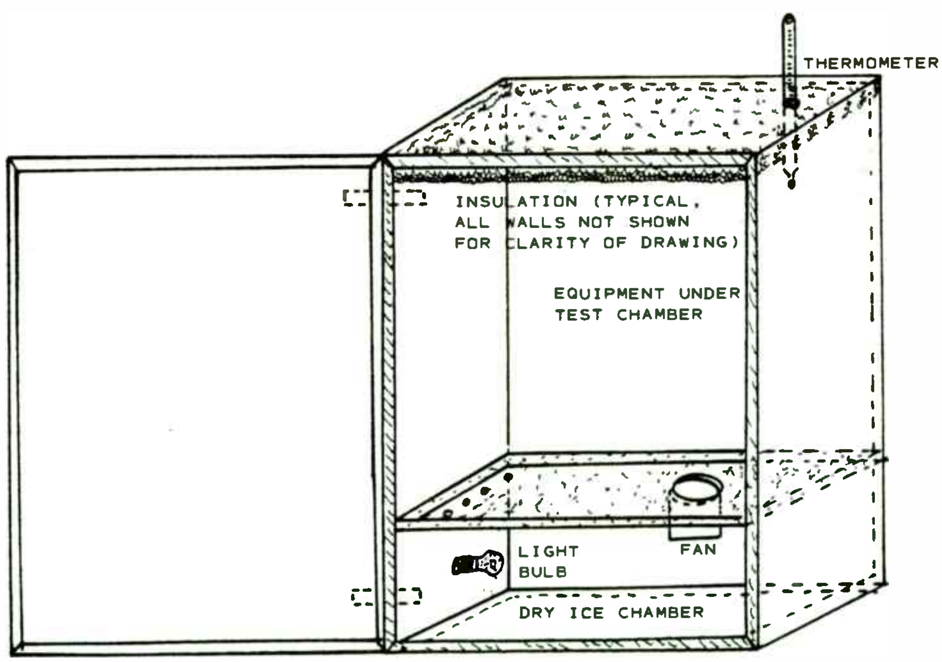


FIGURE 8 ONE POSSIBLE LAYOUT OF A THERMAL ENVIRONMENTAL CHAMBER.

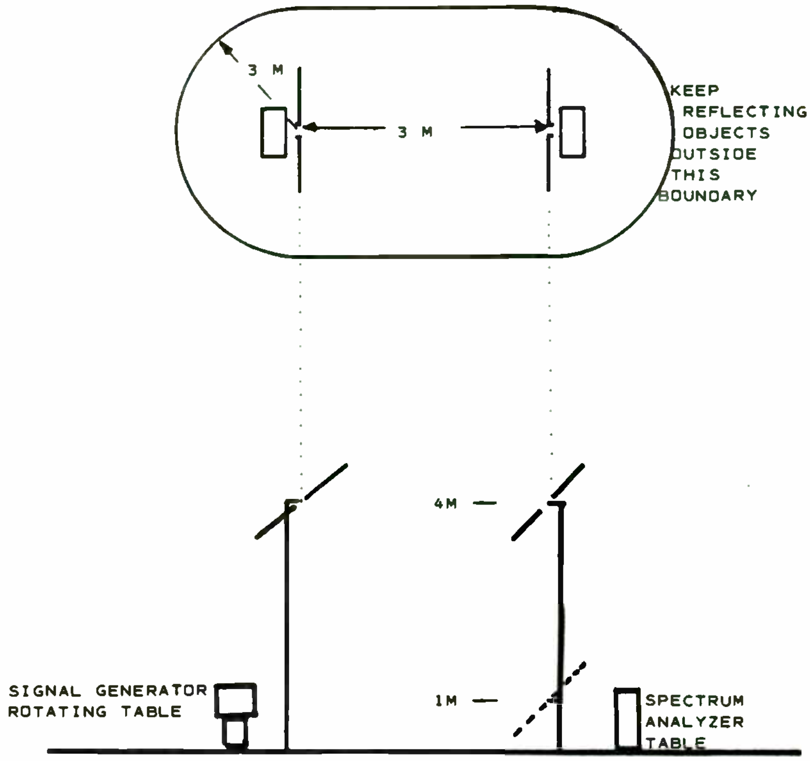


FIGURE 11 OUTLINE DRAWING OF A TYPICAL RADIATION TEST SITE

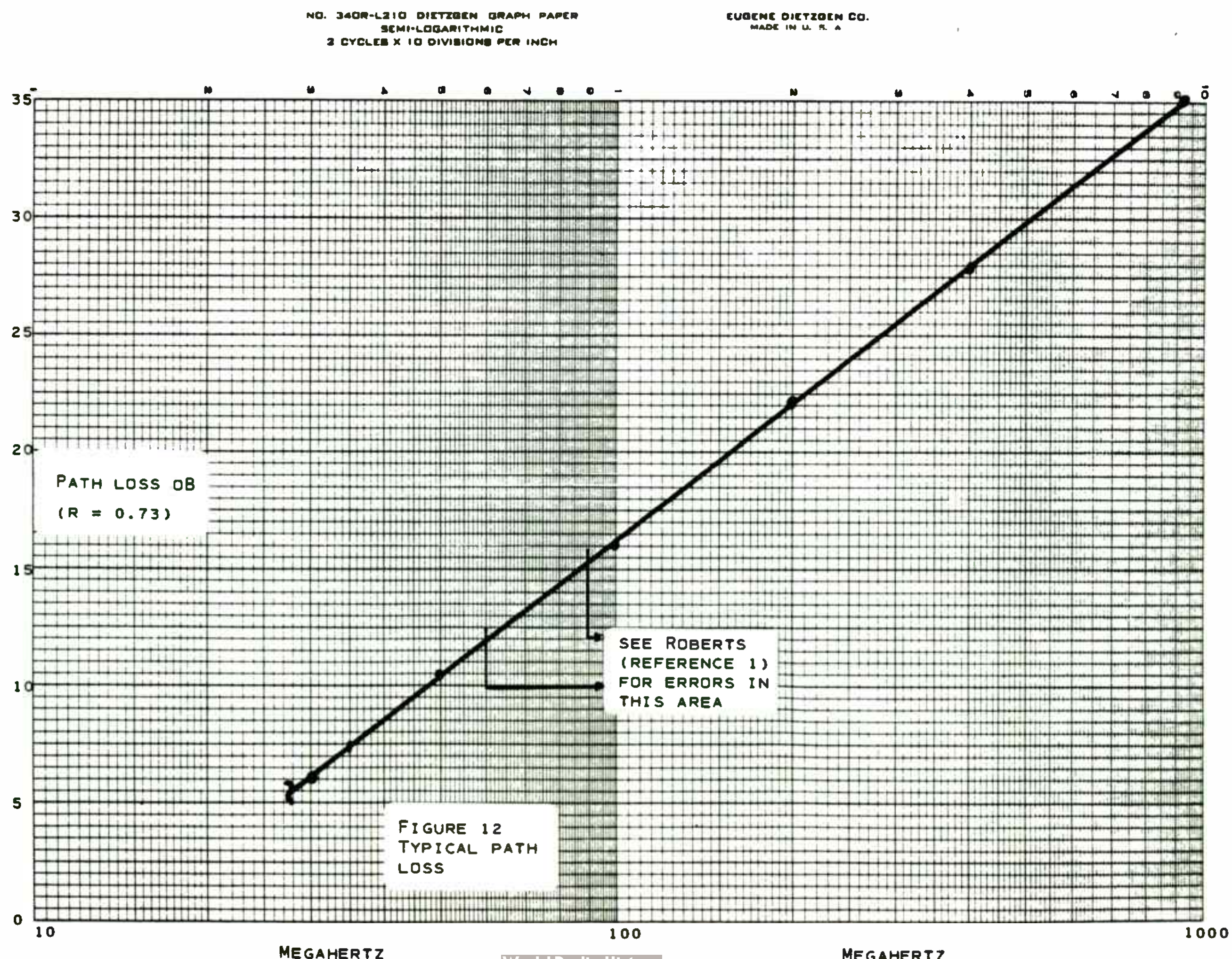


FIGURE 12
TYPICAL PATH LOSS

FEDERAL COMMUNICATIONS COMMISSION

WASHINGTON, D.C. 20554

GRANT OF EQUIPMENT AUTHORIZATION

TYPE ACCEPTANCE AND CERTIFICATION

Radio Systems Technology, Inc.
10985 Grass Valley Avenue
Grass Valley, CA 95945

Attention: Jim Weir

Date of Grant: August 31, 1981

File No.: 31010/EQU-17.9
31010/EQU 4-3-4

Application dated: June 4, 1981

NOT TRANSFERABLE

EQUIPMENT AUTHORIZATION is hereby issued to the named GRANTEE, and is VALID ONLY for the equipment identified hereon for use under the Commission's Rules and Regulations listed below

FCC IDENTIFIER BSV8YCRST571

Name of Grantee Radio Systems Technology, Inc. COPY

Manufacturer Radio Systems Technology, Inc. (USA)

Equipment Class: Communications Transceiver

Note(s)	Rule(s) Part Number(s)	Frequency Range(MHz)	Input Watts	Output Watts	Frequency Tolerance	Emission
	15,87	108-136	-	3	.002	6A3

#81-1988
PMI:sjf
G-3

FIGURE 14. THE END RESULT OF CERTIFICATION AND TYPE ACCEPTANCE TESTING

This form supersedes FCC Forms 722A & 723A.

In correspondence concerning this grant, please refer to the FCC IDENTIFIER, File No., and date of grant.

FCC Form 731A
September 1979

Line Impedance Stabilization Networks: Theory and Use

by Mark Nave
Don White Consultants, Inc.
Gainesville, Virginia 22065

Abstract

This paper analyzes conducted emissions and some of the primary tools for their measurement. Emphasis is placed on the Line Impedance Stabilization Network (LISN). The input impedance and filtering implications of LISN use for different types of conducted emissions is examined to aid in its practical application.

Key Words: LISN, Conducted Emissions

Introduction

The use of Line Impedance Stabilization Networks (LISN's) to measure the effects of filters on conducted emissions (CE) is specified in most major specs, including FCC, VDE, CISPR, MIL-STD-461 and others. While real-life conducted emissions measurements and filter performance may vary considerably from test conditions, some common reference and testing standard is necessary. LISN's allow test facilities to obtain results with greater consistency and repeatability. This paper discusses methods for analyzing common-mode and differential-mode noise and examines the effects of the LISN on CE measurements.

Conducted emissions are an important, but often misconceived electromagnetic interference (EMI) phenomenon. In addition to conducted interference, power leads can act as unintentional antennas, radiating due to CE or receiving noise from the electrical ambient. Proper filtering of the leads to and from the equipment is essential to control this phenomenon. A filter's effectiveness is dependent on an impedance mismatch to both the source (power mains) and load impedances. Figure 1 shows the statistical distribution of mains impedance with its approximate 40 dB variation. Also depicted is the variation of impedance with frequency and the $-50\ \Omega$ centroid behavior of the mean.

Noise Types

There are three basic noise types present on power buses: Differential Mode (DM) and Common Mode (CM) Types I and II. Differential-mode noise is the simplest kind of noise. It occurs between the leads of the intentional current path (phase-neutral or phase-phase), as depicted in Figure 2a. Typical sources of differential-mode noise are switching transients and motors.

Common-mode Type I noise occurs when the noise source is between safety ground and the phases, including neutral, as illustrated in Figure 2c. Common-mode Type II noise occurs when the noise source is between earth ground and all phases, including neutral, and safety ground wire. CM Type II is depicted in Figure 2d.

LISN's: Theory and Use

Measurement Variation

Figure 1 illustrates the approximate 40 dB variation of the mains impedance which can result in up to a 40 dB variation in the measured value of the CE. This effect can be understood by analyzing the interaction of the bus and the source under the assumption that the noise frequency is in the constant (50 Ω) region. Let the bus impedance (Z_b) vary by a ratio χ, $0.1<\chi<10$, so that $Z_b = \chi 50\ \Omega$. The measured voltage, V_m, is the result of voltage division across the internal impedance of the source, Z', and the bus impedance, Z_b. The expected measured voltage ($\chi=1$) is:

$$V_m = \frac{50}{50+Z'}$$

The voltage measured under varying bus impedance ($\chi \neq 1$) is:

$$V'_m = \frac{\chi 50}{\chi 50+Z'}$$

The normalized variation of the measured voltage then becomes:

$$\frac{V'_m}{V_m} = \frac{\chi 50+\chi Z'}{\chi 50+Z'}\ .$$

For a low impedance noise source, as Z' becomes small with respect to 50 Ω, the normalized variation approaches unity. This means that the varying bus impedance has little effect on the measured voltage. However, for a high impedance source, as Z' becomes large with respect to 50 Ω, the normalized variation approaches χ . The worst case variation, then, is

$$\frac{\chi_h}{\chi_l} = \frac{10}{0.1} = 100, \text{ or } 40 \text{ dB.}$$

Such a wide variation in measurements renders the data virtually useless! For this very reason, the LISN for ac mains was developed.

Schematic Concept

A LISN's purpose is to provide a stabilized impedance to conducted emissions without interfering with the normal power flow required by the Equipment Under Test (EUT). A conceptual schematic of the generator, LISN and load is shown in Figure 3a.

At the power line frequency fp, the LISN shown in Figure 3b provides a low impedance path from the power source to the load impedance Z, and a high impedance path (virtual open circuit) from the load to ground. At the noise frequency, fn(fn>>fp), the LISN provides a high impedance path from the power source to the load, and it provides an impedance approaching 50 ohms at high frequencies from the load to ground. The high impedance, low frequency impedance is provided by a capacitor to ground. The 50 Ω impedance to ground ("R" in Figure 4) is actually the input impedance of the spectrum analyzer or EMI meter used to measure the noise. All LISN output ports must be terminated in a 50 Ω impedance, either by meter input impedance or by a 50 Ω dummy load. Figure 3c shows this, whereby the LISN provides a stable impedance to the load and eliminates the effects of the varying mains impedance at noise frequencies.

Analytic Verification

The high frequency (>1 MHz) asymptote can be easily derived using the simplified schematic of Figure 4, where R_m is the mains impedance, R is the 50 Ω impedance of the EMI receiver, and L_1, C_1 and C_2 are the LISN components. Using the ladder network method of analysis, set $I_0 = 1$. Therefore,

$$V_0 = R_m. \tag{1}$$

Then,

$$I_L = I_0 + I_c = 1 + V_0\,(j\omega C_1). \tag{2}$$

Therefore,

$$V_i = V_0 + j\omega L_1(I_L) = R_m + j\omega L_1 - R_m\,\omega^2 L_1 C_1 \tag{3}$$

$$I_{in} = I_L + \frac{V_{in}}{R + \frac{1}{j\omega C_2}} = 1 + V_0(j\omega C_1) + \frac{V_{in}(j\omega C_2)}{1 + j\omega R C_2} \tag{4}$$

so

$$Z_{in} = \frac{V_{in}}{I_{in}} = \frac{V_{in}}{1 + V_0(j\omega C_1) + V_{in}\,\frac{(j\omega C_2)}{1 + i\omega R C_2}}. \tag{5}$$

Since only frequencies above 1 MHz are of interest,

$$Z_{in} = \frac{V_{in}}{1 + j\omega R_m C_1 + \frac{V_{in}}{50}}. \tag{6}$$

Substituting (3) gives

$$Z = 50\,\frac{[-\omega^2 R_m L_1 C_1 + R_m + j\omega L_1]}{[-\omega^2 R_m L_1 C_1 + R_m + 50 + j\omega L_1 + j\omega 50 R_m C_1]}. \tag{7}$$

Plugging in the values for R, C_1, C_2, L_1 gives

$$Z = 50 \cdot \frac{-\omega^2(250\times10^{-12}) + R_m + j\omega(5\times10^{-6})}{-\omega^2(250\times10^{-12}) + R_m + 50 + j\omega\ 5\times10^{-6} + j\omega(250\times10^{-6})}. \tag{8}$$

For this example, at 1 MHz and above, the ω^2 terms clearly dominate, so for f>1 MHz, $Z_{in} - 50$ Ω.

LISN's in Practice

Although the simplified diagram in Figure 3 illustrates the conceptual operation of a LISN, several details of practical LISN operation should be addressed.

Single Phase Test Set Up

In order to provide impedance stabilization for both DM and CM, the LISN is connected between phase to ground and neutral to ground. Figure 5 shows a practical single phase test set-up. Figure 5a is drawn to emphasize the effects of a LISN on CM Type I noise. At high frequencies the inductor is a virtual open circuit while the capacitor is a virtual short circuit. The high frequency equivalent circuit is shown in Figure 5b. The impedance of the two LISN's combine in parallel to present a 25 Ω impedance to the noise source.

With DM noise, the situation is altogether different. Figure 5c shows the single phase set-up redrawn to emphasize the effects of a LISN on DM noise. Under the high frequency assumptions, the equivalent circuit shown in Figure 5d results. For DM noise, the LISN's combine in series to present a 100 Ω impedance to the noise source. Use of the 50 Ω LISN has caused an unexpected impedance when used in a practical circuit, and the situation becomes worse with a three phase circuit.

LISN's on a Wye Bus

Use of the LISN on a three phase Wye bus requires four LISN's—one for each phase and neutral. Both CM and DM noise types exist on the three phase Wye bus. Figure 6a illustrates the test set-up for a Wye configuration with phase-to-phase DM noise sources, and Figure 6b illustrates the high frequency equivalent circuit with the internal impedance Z′ of the noise source. From Figure 6b it can be shown that as the internal impedance Z′ of the noise source becomes very small (Z′<<50 Ω), the LISN's impact diminishes, and the noise sources short themselves out. As Z′ becomes very large (Z′>>50 Ω), the effect of the LISN's dominates, and the impedance seen by the noise source is 100 Ω.

Figure 6c shows the Wye bus with phase-to-neutral noise sources; Figure 6d shows the high frequency equivalent circuit with the internal impedance Z′ of the noise sources. As Z′ becomes large (Z′>>50 Ω) the impedance seen by the noise source is the series combination of the phase and neutral impedances, 100 Ω. As Z′ becomes small (Z′<<50 Ω), the impedance seen by the noise source is 57 Ω. These two boundaries show an impedance stabilization effect for DM noise.

Common-mode noise types remain to be considered. Figures 6e and 6f show the Wye bus with CM Type I noise sources and the high frequency equivalent circuit, respectively. As apparent from Figure 6f, the high frequency impedance seen by the noise source is about 13 Ω. CM Type II noise is rarely a problem with conducted emissions. CM Type II problems manifest themselves as conducted susceptibility problems. This is because all five lines are bundled together in a cable above the ground plane, thus defining a loop area into which B fields are coupled. Figure 7 shows the loop area and the induced loop voltage. By reducing the loop area (lowering the cable), the induced voltage (proportional to the IAF product) will be reduced.

LISN's on a Delta Bus

Conducted emissions on a Delta power bus are virtually the same as in the case of the Wye bus for DM noise. Figure 8a shows a Delta power bus with LISN's and DM noise sources. The situation is exactly analogous to the phase-to-phase DM noise analysis for the Wye bus.

Common-mode noise analysis for the Delta bus is considerably different than for the Wye bus because the Delta bus is isolated from ground and because of the parasitic coupling from bus to ground. If the parasitic capacitance is first neglected to gain insight into the basic interaction of the LISN's, inspection reveals that the high frequency impedance seen by the noise source is about 17 Ohms. The parasitic capacitance modifies this by increasing the effective value of C1 (refer to Figure 4a). This in turn causes the effective input impedance of the LISN to approach its asymptotic value at a faster rate. This is evident by inspection of equation (7). This variation is further limited by transmission line effects at higher frequencies. The exact effect would vary as the value of the parasitic capacitance and frequency, and could only be determined on a case by case basis after considerable analysis. The result, however, would be a squaring up of the impedance versus frequency characteristic.

Measurement Techniques Using LISN'S

The most common method of measuring the value of the conducted emissions with a LISN is with an EMI meter or a spectrum analyzer. Either of these will give the sum of the CM and DM emissions. Although this is usually the method called for in specifications, it provides no information as to whether the emissions are CM or DM! Filter design topologies are quite different for CM and DM, so a method is necessary to discern the two.

Another method for measuring CE is with a current probe. The current probe makes it possible to differentiate between CM and DM. The theory is that the sum of the instantaneous currents at a point on a transmission line equals zero. Proper selection of the lines to sum (to put inside the current probe) will allow use of this principle. The application is to cancel out the DM currents to determine CM currents, and vice versa. Figure 9a shows how to use the probe so that the DM currents sum to zero and twice the CM current is measured. Figure 9b shows how the CM currents cancel and the DM currents are measured.

The LISN may also be used for susceptibility testing. If the impedance of the power mains is too low, injecting a signal of a given level may prove difficult because of the loading effect on the signal generator. This condition may be alleviated by using the LISN in the same configuration as that used for testing, except that the singal is *injected* into the LISN "output" port (now used as an input).

Summary

The effect of varying mains impedance was analyzed and found to have potentially catastrpohic results on CE measurements. In an effort to ensure repeatable measurements, a LISN was introduced to provide a stabilized mains impedance at higher frequencies. The high frequency impedance of the LISN was analytically derived to verify its value and frequency domain performance. Different noise types on different bus types with LISN's were then analyzed to determine the mains impedance that the LISN presented to the noise source.

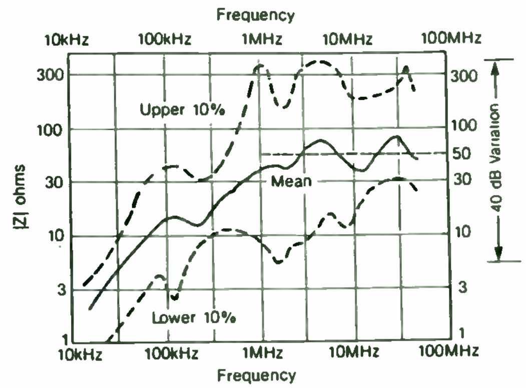
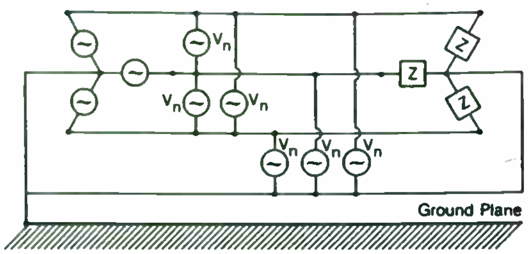
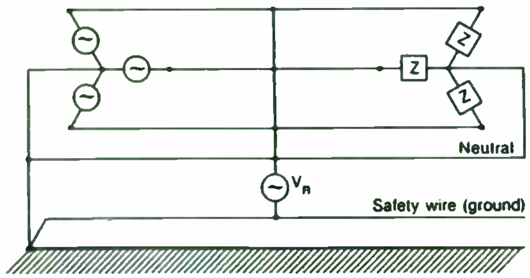


Figure 1—Absolute Mains Impedance (CM or DM) of Power Networks

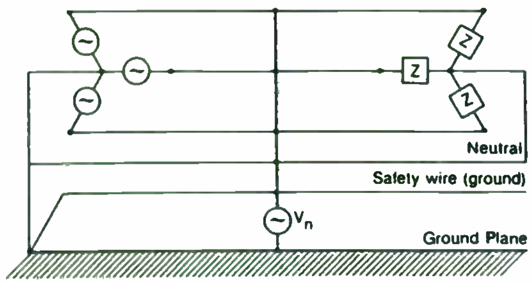


a) Three-phase differential mode noise (phase-phase and phase-neutral)

b) Single phase differential mode noise

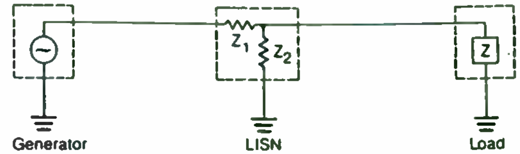


c) Common mode type I

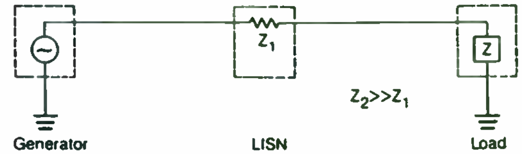


d) Common mode type II

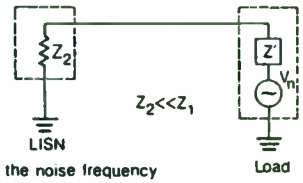
Figure 2—Noise Types



a) Conceptual schematic

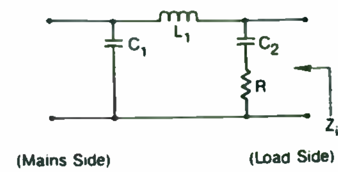


b) Simplified schematic at the power line frequency

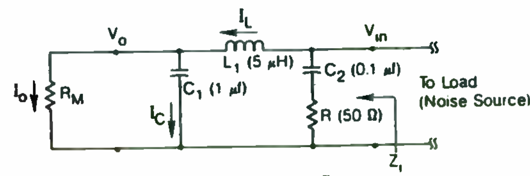


c) Simplified schematic at the noise frequency

Figure 3—Conceptual Diagrams of LISN's

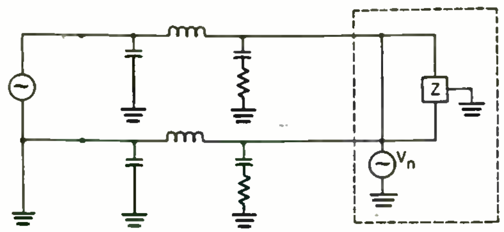


a) Simplified schematic of an LISN

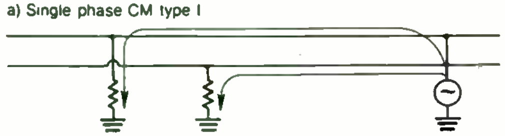


b) Simplified schematic for calculating Z_i

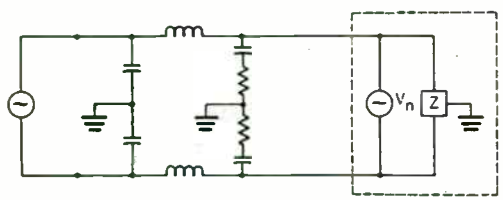
Figure 4—Functional Schematic of LISN



a) Single phase CM type I



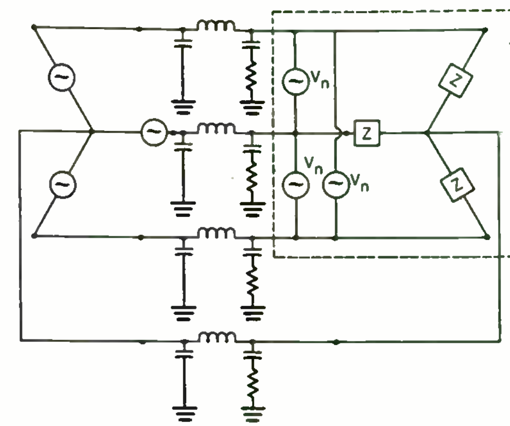
b) CM Type I high frequency equivalent circuit



c) Single phase DM noise

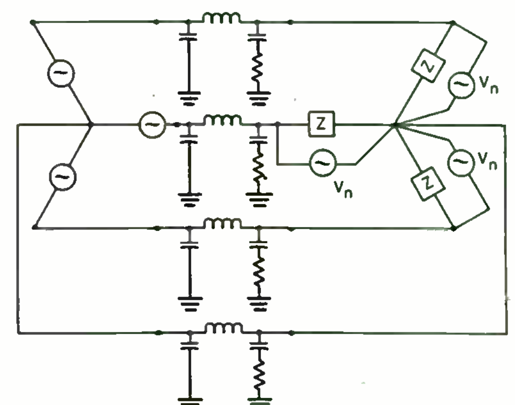
d) CM type I high frequency equivalent circuit

Figure 5—Practical Single Phase Test Set-up

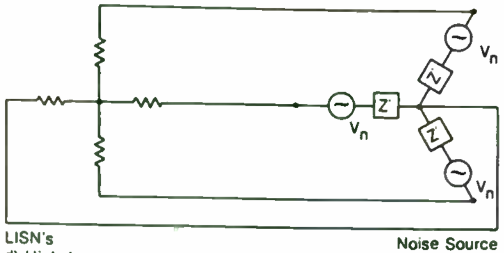


a) Wye bus LISN set-up with phase-to-phase DM noise sources

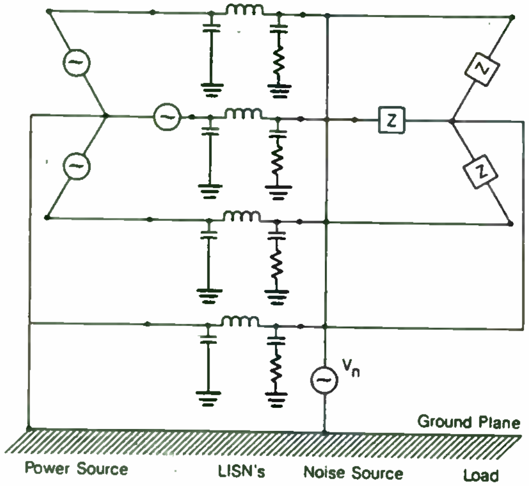
b) High frequency equivalent circuit for DM noise



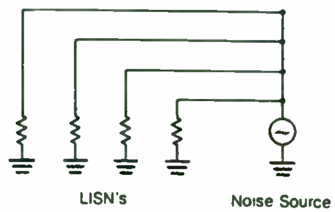
c) Wye bus LISN set-up with phase-to-neutral DM noise sources



d) High frequency equivalent circuit for phase-to-neutral DM noise sources

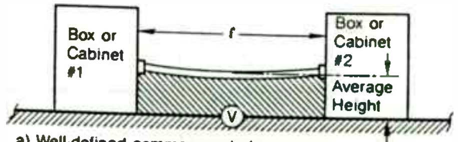


e) Wye bus LISN set-up with CM type I noise sources

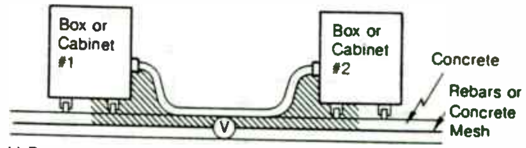


f) High frequency equivalent circuit for CM type I

Figure 6—LISN's Use on a Wye Bus

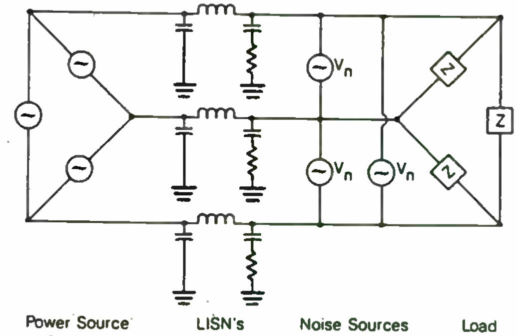


a) Well-defined common-mode loop area

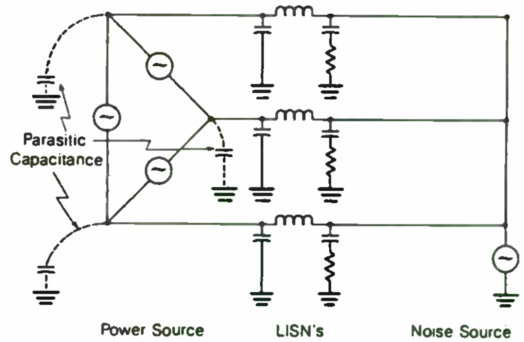


b) Poorly-defined common-mode loop area

Figure 7—Common Mode Type II Coupling

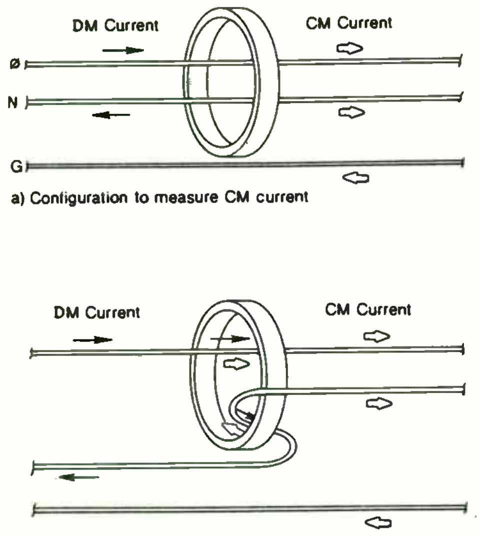


a) Delta bus with LISNs and DM noise sources



b) Simplified high frequency schematic for CM noise on a delta bus

Figure 8—LISN's Use on a Delta Bus



a) Configuration to measure CM current

b) Configuration to measure DM current

Figure 9—Differentiating Between DM and CM Currents

INTRODUCTION TO THE FIELD OF ELECTROMAGNETIC PULSE

By: Kendall Childers
Senior Scientist

Chomerics, Inc.
Woburn, Ma. 01888

INTRODUCTION

There are electromagnetic disturbances associated with the detonation of chemical explosives, therefore it was no surprise that there were electromagnetic disturbances associated with nuclear explosives. What was a surprise was the extent of the geographical coverage, the bandwidth of the energy spectrum and the amplitude of some of the disturbances. The nuclear induced disturbance is called Nuclear Electromagnetic Pulse (NEMP) or simply Electromagnetic Pulse (EMP).

EMP is caused by electrons ejected from materials by gamma-rays and X-rays emitted from the nuclear explosions. EMP goes by a variety of aliases; High altitude EMP (HEMP), Low altitude EMP (LEMP), Source Region EMP (SREMP), System Generated EMP (SGEMP), Internal EMP (IEMP), etc. The names identify the source of the EMP and are a "shorthand" used to indicate the characteristics of the EMP of interest. This paper will deal exclusively with HEMP, the most serious threat to telecommunication systems.

GENERATION OF HEMP

When the gamma-rays from an exo-atmospheric nuclear explosion descend to an altitude of about 40 km, the air density becomes sufficiently dense that there are significant interactions. By the time that the gamma-rays penetrate to an altitude of about 20 km, they are completely absorbed in the atmosphere. The primary interaction, Compton collision, results in an ejected electron and a scattered gamma-ray. The ejected electrons are turned by the earth's magnetic field (similar to the deflection [turning] of the electron beam in a TV tube by the yoke.) The process of accelerating (deflecting) charged particles generates electromagnetic radiation.

The HEMP area of coverage is determined by the area of the spherical cap enclosed by the tangent drawn from the point of the explosion to the surface of the earth. An easy set of numbers to remember is: HEMP from a nuclear explosion at an altitude of 300 miles above the surface of the earth will illuminate a 3,000 mile diameter region on the surface of the earth.

The gamma-rays from a nuclear explosion are emitted in a burst with a duration of around 10 nanoseconds. Therefore, HEMP is a pulse. The risetime of the pulse is related to the duration of the burst of gamma-rays, a few to around 10 nanoseconds. The decay of the pulse is caused by very complicated processes, the discussion of which is beyond the scope of this paper, which result in pulse lengths from about 1/10th to about one microsecond. The frequencies of interest when one is dealing with pulses are determined by taking the Fourier transform of the pulse. HEMP is typically "full energy rich" up to about 1 megahertz and has significant energy content up to at least 100 megahertz.

The degree of deflection of the ejected electrons by the earth's magnetic field depends on the direction of the electron trajectory compared to the direction of the earth's magnetic field. If the directions are perpendicular, deflection is maximized; if they are parallel, deflection is minimized. The amount of electromagnetic radiation is related to the degree of deflection, the larger the deflection the greater the radiation. The gamma-rays travel radially outward from the explosion, and the electrons are ejected "primarily in the forward direction", i.e., radially outward from the explosion. The declination of the earth's magnetic field means that (in the northern hemisphere) north of the explosion there is a region where the ejected electrons are parallel to the earth's magnetic field - in that direction there will be no HEMP. Likewise there will be a region south of the explosion where the ejected electrons are perpendicular to the earth's magnetic field - in that direction the EMP will be the maximum possible, usually characterized by a peak electric field strength of 50,000 volts/meter. In other directions HEMP will have intermediate amplitudes.

The precise characteristics of HEMP depend on the size of the nuclear explosion and the geometric relationship between the position of the explosion, the observer and the earth. Since it is not possible to specify a unique set of parameters, a composite "worst case" waveform is used. The "worst case" threat retains the nastiest characteristics of the various forms of HEMP, namely; the fastest risetime - less than 10 nanoseconds, the maximum peak electric field strength - about 50,000 volts/meter, and the longest pulse duration - about one nanosecond. The electric field strength for "worst case" HEMP is described with a double exponential (ref. 1):

$$E(t) = Eo\,[\,e^{-t/t1} - e^{-t/t2}\,]$$

where: E(t) = electric field strength as a function of time
Eo : related to peak electric field strength about 52,500 volts/meter
t = time
t1 : related to pulse width, about 250 nanoseconds
t2 : related to rise time, about 2 nanoseconds

HEMP is a plane wave with a wave impedance of 377 ohms therefore the corresponding magnetic field strength is given by:

$$H(t) = E(t)/377$$

The peak "worst case" magnetic field strength is about 133 amps/meter.

The energy density in HEMP is small, about 1 joule/square meter but the power density is large, about 7 megawatts/square meter. Whereas a significant power is incident on small (square meter) structures, significant energy is incident only on truly large (many - many square meter) structures.

One of the advantages of describing HEMP with a double exponential is that the Fourier transform is trivial. The transform is constant from zero hertz up to a frequency of $1/(2\pi\ t1)$ = 640 kilohertz where it starts to roll off at 20 dB/decade up to a frequency of $1/(2\pi\ t2)$ = 76 megahertz where it starts to roll off at 40 dB/decade.

HEMP COUPLING INTO SYSTEMS AND HEMP PROTECTION TECHNIQUES

The coupling of HEMP energy into systems is essentially an EMI/EMC problem, described by Maxwell's equations. It depends on the characteristics of HEMP and the details of the geometry of the system. We have taken care of the characteristics of HEMP by using the "worst case". The details of the geometry of the system are generally a problem, since there are only a few simple geometries for which there are analytic solutions to Maxwell's equations. With few realistic analytic solutions, complicated three dimensional computer calculations or measurements are required. Measurements are out of the question with very large systems. Therefore, as with EMI/EMC problems, simplified problems are considered to "understand" the principles of HEMP coupling, measurements are performed on small systems and detailed calculations are performed only when absolutely required. The results of estimates and calculations are that typical values of HEMP energies coupled into large systems are about 1 megajoule in dc power systems and about 10 megajoules in ac power systems such as the U.S. power grid. The susceptibility of solid-state electronic devices to HEMP damage can be estimated by subjecting them to pulsed current injection tests. The energy required to destroy a typical integrated circuit junction is about 10 microjoules and the energy required to destroy a microwave diode can be as low as 1/10th of a microjoule. These energies should be contrasted with vacuum tubes and (electro-mechanical) relays which can survive up to about one joule. This means that an enormous amount of shielding is required to protect modern circuitry.

The recommended procedure for protecting modern systems is to use Faraday cage inside of Faraday cage inside Faraday cage ... until sufficient shielding is accomplished that the system survives (Ref. 2). The Faraday cage inside Faraday cage scenario requires single point grounding between neighboring cages, separation of signal and power cables between cages, shielding of cables between cages and "terminal protection" where the cables penetrate the cages. The key feature of this scenario is that the shielding is distributed, no single shield is required to provide an inordinate amount of shielding.

HEMP shielding considerations are similar to those required for EMI/EMC. They can be divided into low frequency, high frequency, gasket and cable effects. Low frequency shielding is described by quasistatic shielding. First order high frequency shielding is described by exponential absorption (skin depth effects) but is usually dominated by leakage through seams. The gasket issue, for properly gasketed seams, is survivability. HEMP coupling into systems via cables usually dominates these other mechanisms.

The quasistatic requirement is that the wave length associated with the highest frequency of interest is much longer than the relevant dimensions of the enclosure. The quasistatic analysis determines surface charge and current distributions that the quasistatic fields induce on the enclosure, then determines the interior fields due to these charge and current distributions. It is unlikely that quasistatic electric field shielding will be a problem. It is likely that quasistatic magnetic field shielding will be poor, as it is well known that magnetic field shielding effectiveness becomes very small at low frequencies. For example, the quasistatic shielding provided by a 1.3

meter (50") diameter, 4.8 millimeter (3/16") wall thickness cylinder is nowhere less than 250 dB for the electric field but is less than 50 dB for all frequencies less than 1 kilohertz. Another shielding issue is leakage through apertures, which is included in the low frequency region because the theory has been worked out for apertures that have dimensions small compared to the wave length of the highest frequency of interest. The leakage fields fall off as the third power of the ratio of the radius of the aperture to the distance from the aperture [(a/r)**3] (Ref. 3 and 4). This means that apertures must be small and that sensitive equipment cannot be located near apertures.

High frequency shielding depends on the exponential absorption of electromagnetic energy in enclosure walls. For most enclosures the exponential absorption is so large that shielding is actually dominated by second order effects, i.e., shield imperfections at seams and seals. However, since HEMP is broadband, system resonances must be considered. Induced currents at resonant frequencies may be large enough that even with enormous shielding, sufficient energy may diffuse through the shield to cause damage.

Gasket issues are two fold, first what happens during HEMP excitation and second does the gasket survive exposure to HEMP? Shielding during HEMP exposure will probably be at least as good as it was before the HEMP exposure. The key issue is HEMP survivability; does the gasket retain its EMI/EMC sealing characteristics after HEMP? To evaluate the survivability of EMI/EMC gaskets, we have devised a HEMP simulation scenario. We calculate the HEMP-induced currents in a microwave communications relay system consisting of a tower connected to a shielded building with a wave guide and a conduit. We find that the EMI/EMC gasket must survive a damped sinewave current pulse having an amplitude of 13,500 amperes peak to peak, a frequency of 1.3 megahertz and a damping time constant of 750 nanoseconds. Our laboratory tests are current injection tests; we drive this damped sinewave current pulse through test gaskets. We find that conductive elastomer EMI/EMC gaskets containing silver coated inert particles fail the test and those that contain silver plated metallic particles, particularly when the particles are irregularly shaped, survive currents significantly larger than any realistic threat.

Though HEMP couples into short cables, usually magnetic field coupling into loops, the most serious coupling is into long cables. Current is driven in long cables by the longitudinal component of the incident electric field. In the case of shielded cables, the current is driven on the shield. The current induced on the interior cables is derived from transfer impedance coupling. HEMP can induce large transients in cables, e.g., open circuit voltages of 7,000,000 volts and short circuit currents of 14,000 amperes in overhead cables (power lines) (Ref. 1). HEMP coupling from cables into systems is controlled with terminal protection devices, similar to lightning arrestors. Risetimes associated with HEMP can be shorter than 10 nanoseconds and risetimes associated with lightning are 100 to 200 nanoseconds. HEMP terminal protection devices do not have to carry as much current as lightning arrestors, but they must act significantly faster. HEMP terminal protection devices are gas switches, metal oxide varistors (MOV's), conventional EMI/EMC filters, optical couplers and isolation transformers with back-to-back diodes.

CONCLUSION

HEMP is a severe electromagnetic disturbance caused by nuclear explosions above the earth's atmosphere. It is characterized by extensive geographical coverage, an extremely wideband energy spectrum and a very large amplitude. The amount of energy coupled into systems can be very large compared to the energy required to damage modern electronic components. HEMP protection techniques are straightforward extensions of standard EMI/EMC techniques.

REFERENCES:

1. R. Sherman et al, "EMP Engineering and Design Principles", Bell Telephone Laboratories, Inc., 1975

2. J. Miletta et al, "Defense Switched Network Design Practices for Protection Against High-Altitude Electromagnetic Pulse", HDL-SR-82-2, Harry Diamond Laboratories, Adelphi, MD, 1982

3. L. Ricketts et al, "EMP Radiation and Protective Techniques", John Wiley and Sons, 1976

4. ITT Research Institute, "EMP Design Guidelines for Naval Ship Systems", NSWC/WOL/TR 75-193, 1975

BIBLIOGRAPHY:

1. Sandia Laboratories, "Electromagnetic Pulse Handbook for Missiles and Aircraft in Flight - EMP Interaction 1-1", AFWL TR-73-68, 1973

2. K. Lee et al, "EMP Interaction: Principles, Techniques and Reference Data - EMP Interaction 2-1", AFWL-TR-80-402, 1980

FUNDAMENTALS OF RF DESIGN: PART II

A session of four papers presented January 23 and 24, 1985, at

rf technology expo 85

Disneyland Hotel Anaheim, California

INTRODUCTION

These four papers comprise the second half of a one-day course entitled "Fundamentals of RF Design," presented on two consecutive days — January 23 and 24, 1985 — at RF TECHNOLOGY EXPO 85, a technical conference sponsored by *RF Design* Magazine.

The papers are collected in this form as an aid to conference attendees actually present at the January 23 and 24 afternoon sessions. They will also be reproduced as part of the "Proceedings" of RF TECHNOLOGY EXPO 85, to be published after the event.

1. "Striplines and Microstriplines Design"
   by K.C. Gupta, Professor, University of Colorado
2. "High Power Solid State Circuit Design"
   by J.H. Johnson, President, Microwave Modules and Devices
3. "Oscillators: A Critical Systems Building Block"
   by John Morton, Engineering Manager, Microsonics Inc.
4. "And Introduction to SAW Devices: Design, Fabrication, Testing, Uses, and Future Trends"
   by Carl A. Erikson, Jr., Director of Processing Operations, Andersen Laboratories

STRIPLINES AND MICROSTRIPLINES*

by

K.C. Gupta
Department of Electrical and Computer Engineering
University of Colorado
Boulder, Colorado 80309

* Notes for short course on 'Fundamentals of RF Design" at RF Technology Expo 85 at Disneyland Hotel, Anaheim, California, January 23-25, 1985

STRIPLINES AND MICROSTRIPLINES

Stripline and microstripline are two planar transmission lines used extensively in printed microwave circuits and microwave integrated circuits. This tutorial discussion presents basic features of these lines and the design information needed for their use in RF circuits.

The basic configurations of these lines are shown in Fig. 1. The main advantage of these lines is their planar nature which results in ease of fabrication by photolithography and chemical etching processes (starting from commercially available metal clad dielectric substances).

I. STRIPLINES

Stripline (also known as triplate line) is a three conductor TEM mode transmission line with the cross-sectional configuration as shown in Fig. 1(a). It consists of a central strip with its width much greater than its thickness (W >> t) and two ground planes which extend considerably in transverse x-direction. The space between the ground planes is filled by a dielectric medium and the central strip is embedded in this dielectric. A practical embodiment of this configuration is shown in Fig. 2. Two laminates A and B (each of height b/2) are used . The lower laminate A is completely metallized on the bottom surface and has a metallic strip of the required width W on the other side. The laminate B has metallization on the top surface and is clear of any metal underneath. The two laminates are screwed (or clamped) together to yield the stripline. As thickness of the strip is very small (typically 1.4 mil) any air gap existing after pressing the two laminates together may be ignored.

The mode of propagation along a stripline is transverse electromagnetic (TEM). The ground planes are at zero potential. The electric and magnetic field lines are sketched in Fig. 3(a) and (b) respectively. Field lines are concentrated near the central strip and as one moves in the mid-plane away from the strip (along x-axis) the fields decay rapidly. This feature allows the substrate and the ground planes to be terminated in the transverse direction without affecting the transmission characteristics of the line.

The stripline geometry may be approximated by a parallel plate transmission line structure shown in Fig. 4. Boundaries denoted by MW are hypothetical open circuit walls (or magnetic walls) and E-field does not extend outside these boundaries. W_e is the effective width of the strip chosen such that energies stored in E-fields of Figs. 3(a) and 4 are equal. In terms of the geometry of this parallel plate model of the stripline, the capacitance per unit length of the stripline may be written as

$$C = 2\,\epsilon_0\,\epsilon_r\,W_e/(b/2) = 4\,\epsilon_0\,\epsilon_r\,W_e/b \qquad (1)$$

and the characteristic impedance of the line may be written as

$$Z_0 = \sqrt{L/C} = \frac{1}{v_p C} = \frac{\sqrt{\mu_0 \epsilon_0}\,\sqrt{\epsilon_r}}{C} = \sqrt{\frac{\mu_0}{\epsilon_0}}\ \frac{1}{4\sqrt{\epsilon_r}}\ \frac{b}{W_e}$$

$$= \frac{30\,\pi}{\sqrt{\epsilon_r}}\ b/W_e$$

Fringing field evaluation is needed inorder to evaluate W_e to be used in Eqs. (1) and (2). The fringing field calculations are easier for the case when the thickness of the central strip t approaches zero. For this thin strip case, Cohn* has used the conformal mapping technique to derive the total capacitance (the parallel plate capacitance plus the fringing capacitance) which leads to the following expression for the characteristic impedance of stripline:

$$Z_0\ \sqrt{\epsilon_r} = 30\ \pi\ \ K'(k)/K(k)\ \text{ohm} \qquad (3)$$

where k is a geometrical parameter given by

---

* Cohn, S.B., "Characteristic impedance of a sheilded strip transmission line," IRE Transactions Microwave Theory Tech. Vol. MTT-2, July 1954, pp 52-55.

$$k = \tanh\left(\frac{\pi W}{2b}\right) \qquad (4)$$

K represents the complete elliptical function of the first kind and K' is its complementary function. We have

$$K(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2\phi}\, d\phi\ , \quad \text{and } K'(k) = K\left(\sqrt{1 - k^2}\right) \qquad (5)$$

It is not necessary to evaluate the integral in (5) numerically since simple and accurate (8 parts per million) expressions are available for the ratio K(k)/K'(k). We have

$$\frac{K(k)}{K'(k)} = \begin{cases} \left\{\frac{1}{\pi} \ln\left(2\,\frac{1 + \sqrt{k'}}{1 - \sqrt{k'}}\right)\right\}^{-1} & \text{for } 0 \leq k \leq 1/\sqrt{2} \\ \frac{1}{\pi} \ln\left(2\,\frac{1 + \sqrt{k}}{1 - \sqrt{k}}\right) & \text{for } 1/\sqrt{2} \leq k \leq 1 \end{cases} \qquad (6)$$

where $k' = \sqrt{1 - k^2}$

Relation (3) yields the impedance of the stripline for a given geometry. However, in circuit design problems it is desirable to have formulas which yield the width of the line for a given impedance. This synthesis equation is obtained by manipulating (3), (4) and (6) and may be written as:

$$W/b = \frac{2}{\pi} \tanh^{-1} k \qquad (7)$$

where

$$k = \begin{cases} \sqrt{1 - \left\{\frac{e^{\pi a} - 2}{e^{\pi a} + 2}\right\}^4} & \text{for } a \geq 1 \\ \left\{\frac{e^{\pi/a} - 2}{e^{\pi/a} + 2}\right\}^2 & \text{for } 0 \leq a \leq 1 \end{cases}$$

with $a = Z_0 \sqrt{\epsilon_r} / (30\pi)$.

Above results are rigorously valid when strip thickness $t = 0$. Analysis and synthesis of striplines with thick strips are more complicated. Several approximate formulas are available. An expression with accuracy better than 0.5 percent, for $W/(b-t) < 20$, (given by Wheeler*) is as follows:

$$Z_0 \sqrt{\epsilon_r} = 30 \ln\left\{1 + \frac{4}{\pi}\,\frac{b - t}{W'}\left[\frac{8}{\pi}\,\frac{b - t}{W'} + \sqrt{\left(\frac{8}{\pi}\,\frac{b - t}{W'}\right)^2 + 6.27}\right]\right\} \qquad (9)$$

where

$$\frac{W'}{b - t} = \frac{W}{b - t} + \frac{\Delta W}{b - t}, \text{ and}$$

$$\frac{\Delta W}{b - t} = \frac{x}{\pi(1 - x)}\left\{1 - \frac{1}{2} \ln\left[\left(\frac{x}{2 - x}\right)^2 + \left(\frac{0.0796\, x}{W/b + 1.1x}\right)^m\right]\right\} \qquad (10)$$

in which

$$m = 2\left\{1 + \frac{2}{3}\,\frac{x}{1 - x}\right\}^{-1} \text{ and } x = t/b \qquad (11)$$

---

* Wheeler, H.A., "Transmission line properties of a stripline between parallel planes," IEEE Trans. Microwave Theory Tech., Vol. MTT-26, Nov. 1978, pp 866-876.

Equation (9) may be rearranged to yield W/b for a given Z_0 as follows:

$$W/b = \frac{W_0}{b} - \frac{\Delta W}{b} \text{ where}$$

$$\frac{W_0}{b} = \frac{8(1-x)}{\pi} \frac{\sqrt{e^A + 0.568}}{e^A - 1}, \quad A = Z_0 \sqrt{\epsilon_r}/30 \qquad (12)$$

and

$$\frac{\Delta W}{b} = \frac{x}{\pi}\left\{1 - \frac{1}{2}\ln\left[\left(\frac{x}{1-x}\right)^2 + \left(\frac{0.0796x}{W_0/b - 0.26x}\right)^m\right]\right\} \qquad (13)$$

The quantities m and x are as defined in 11.

Comparing the values of Z_0 obtained from (10) with those obtained from (3) for $t \neq 0$; we note that impedance values decrease when t is increased. If we design a line for $Z_0 = 50$ using (3) (say for b = 0.12 inch and $\epsilon_r = 2.5$), values of Z_0 for t = 0.0007 inch (1/2 oz. copper) and for t = 0.0014 inch (1 oz. copper) are 49.28 and 48.69 ohms respectively.

Variation of stripilne impedance ($Z_0 \sqrt{\epsilon_r}$) with W/b is represented graphically in Fig. 5.

Losses in Striplines

As for other types of transmission lines, the attenuation in striplines originates from conductor and dielectric losses, i.e.,

$$\alpha = \alpha_c + \alpha_d \qquad (14)$$

where α_c is the attenuation because of conductor losses and α_d is the attenuation because of dielectric losses.

Conductor Losses. At microwave frequencies, current flow in conductors is governed by the skin effect. For a semi-infinite conducting medium, the current density distribution may be expressed as

$$J = J_0 e^{-\gamma y} \qquad (15)$$

where y denotes the distance inside the conducting medium (normal to the surface) and γ is the propagation constant for plane wave in the conducting medium ($\gamma = \sqrt{j\omega\mu(\sigma + j\omega\epsilon)}$). Linear density* of the surface current J_s is obtained as:

$$J_s = \int_0^\infty J\, dy = J_0/\gamma \qquad (16)$$

Surface impedance is defined as the ratio of tangent electric field to J_s and is given by

$$Z_s = R_s + jX_s = E_{tan}/J_s = \frac{J_0/\sigma}{J_s} = \frac{\gamma}{\sigma} = \sqrt{\frac{j\omega\mu}{\sigma}}$$

$$= \sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}} \qquad (17)$$

Thus

---

* Total current per meter on the surface perpendicular to the current flow.

$$R_s = X_s = \frac{1}{\sigma\Delta} \qquad (18)$$

where Δ is the skin depth given by $\Delta = \sqrt{2/(\omega\mu\sigma)}$. Power loss per unit area of the plane conductor can now be written as $J_s^2 R_s$, where R_s is the surface resistance evaluated above and J_s is the linear current density or current per meter width (effective value) flowing in the conductor. Surface reactance X_s may be written as ωL_i where the inductance L_i is attributable to the skin effect, i.e., inductance L_i is produced by the magnetic field inside the conductor.

We note that R_s and X_s are numerically equal for a plane conductor. They are also equal for a conductor of any arbitrary shape if the radii of curvature and the thicknesses of conductors are much greater than the skin depth. This equivalence is helpful in calculating stripline losses because calculation of L_i can be carried out relatively easily by using Wheeler's incremental inductance rule.* According to this rule, L_i can be found from the external inductance L per unit length for the configuration (in our case the stripline geometry). L_i is obtained as the incremental inductance in L caused by an inward displacement of all metallic walls due to skin effect. The amount of displacement is equal to half the skin depth Δ. For example, we can apply the Wheeler's incremental inductance rule to parallel plate waveguide with magnetic walls (Fig. 6). Displacement of the two conducting walls is also shown in Fig. 6. With the walls recessed, $L + L_i = \mu_0 (h + \Delta)/W$ where L is inductances of the structure (given by $L = \mu_0 h/W$). Therefore, $L_i = \mu_0 \Delta/W$, and

---

* Wheeler, H.A., "Formulas for the skin effect," Proc. IRE Vol. 30, 1942, pp 412-424.

$$R_s = X_s = \omega L_i = \omega\mu_0 \Delta/W \text{ ohm/meter} \qquad (19)$$

Therefore the power loss in conductors of the parallel plate line is given by

$$P_c = |I|^2 R_s = |I|^2 \omega\mu_0 \Delta/W \qquad (20)$$

Now we can write α_c which is defined as ratio of the power loss in conductors to twice the power transmitted

$$\alpha_c = \frac{P_c}{2\,P(3)} = \frac{|I|^2 \omega\mu_0 \Delta/W}{2\,|I|^2 Z_o} = \frac{\omega\mu_0\Delta}{2w\sqrt{\frac{\mu_0}{\epsilon_0\epsilon_r}}\,\frac{h}{w}}$$

$$= \omega\sqrt{\mu_0\epsilon_0\epsilon_r}\,\frac{\Delta}{2h} \quad \text{neper/m} \qquad (21)$$

The above example illustrates the use of Wheeler's incremental inductance rule for calculating attenuation constant of a parallel plate line. The same procedure can be used for any other TEM mode line. A general expression for α_c of any TEM mode line may be written as

$$\alpha_c = \frac{R_s}{2\,Z_o} = \frac{\omega L_i}{2\,Z_o} = \frac{\omega}{2\,Z_o} \sum_m \frac{\partial L}{\partial n_n} \frac{\Delta_m}{2} \qquad (22)$$

where m is the number of conductor surfaces (two for parallel plate line), n is the distance in the direction of inward normal to the surface and Δ_m is the skin depth of the m-th surface (different surfaces may have different σ and hence different skin depths). The inductance L may be expressed in terms of Z_o as

$L = Z_0 / v_p$, and therefore,

$$\frac{\partial L}{\partial n} = \frac{1}{v_p} \frac{\partial Z_0}{\partial n} \qquad (23)$$

Thus (22) can be expressed in terms of Z_0 as

$$\alpha_c = \frac{\omega}{2 Z_0 v_p} \sum_m \frac{\partial Z_0}{\partial n_m} \Delta_m/2 \qquad (24)$$

Thus, if we have an explicit expression for Z, in terms of various dimensions of a line, α_c may be computed in a straightforward manner by using (24). For calculating α_c for a stripline, expression for Z_0 given by (9) for the case $t \neq 0$ can be used. Since conductor losses depend critically on the thickness of the control strip, expressions for $t = 0$ case should not be used for calculating losses. This observation is equally valid for the case of micro-stripline discussed later, as well as for other planar transmission lines not discussed here.

For calculating α_c for stripline, summation in (24) is to be carried over various conducting surfaces (total six) shown in Fig. 7.* Although $\partial Z_0/\partial b$, $\partial Z_0/\partial t$, and $\partial Z_0/\partial w$ may be obtained explicitly by differentiating (9), it is perhaps easier to find $\partial Z_0/\partial b\ \Delta_n$ etc. by computing Z_0 twice, once for nominal values of b and then for ground plane spacing $b + \Delta_g$ (as indicated in Fig. 7). In fact, $\sum_m \frac{\partial Z_0}{\partial n_m} \Delta_m/2$ itself may be obtained by computing Z_0 twice, once for nominal dimensions and second time for dimensions

* Δ_s and Δ_0 are skin depths for strip conductor and ground plane respectively.

with all the conducting surfaces recessed inwards by half the skin depth. This approach is well suited for implementation in a CAD package.

Dielectric Losses. The dielectric loss in a stripline (or for that matter in any other homogeneously filled TEM mode line) is directly proportional to the loss tangent, tan δ of the dielectric medium.

Starting from the expression for propagation constant

$$(\gamma = \sqrt{(G + j\omega C)(R + j\omega L)}),$$

one can derive α_d by putting $R = 0$ and simplifying. We get

$$\alpha_d = G\, Z_0/2 \qquad (25)$$

The transmission line conductance G is related to the capacitance C by

$$G = \omega\,(\tan \delta) C, \text{ where } \tan \delta = \omega\epsilon/\sigma \qquad (26)$$

and since

$$Z_0 = 1/(v_p C) \qquad (27)$$

we can write α_d as

$$\alpha_d = \frac{G Z_0}{2} = \frac{1}{2}\omega\,(\tan \delta) C \frac{1}{v_p C} = \pi\,(\tan \delta)\ \sqrt{\epsilon_r}/\lambda_0 \text{ nepers/m} \qquad (28)$$

where λ_0 is the free space wavelength.

Attenuation constants α_c and α_d are more often expressed in dBs which involves a multiplication by 20 $\log_{10} e$, i.e., a factor 8.68. Thus

$$\alpha_d = \frac{27.3 \sqrt{\epsilon_r}}{\lambda_0} \tan\delta \text{ dB/m} \tag{29}$$

It may be noted that because of $\frac{1}{\lambda_0}$ factor, α_d is directly proportional to frequency. On the other hand the frequency variation of the conductor loss α_c (given by 24) is contained in $\omega\Delta_m$. Since the skin depth $\Delta = \sqrt{2/(\omega\mu\sigma)}$ is inversely proportional to the square root of frequency, α_c increases only as square root of frequency. When typical numerical values are compared, the dielectric loss in general, is very small compared to the conductor loss of microwave frequencies. But at millimeter wavelengths, the dielectric loss becomes comparable to the conductor loss because of the different frequency variations of the two losses.

II. MICROSTRIP LINES

Microstrip line consists of a single dielectric substrate with ground plane metallization on one side and a strip of width W on the other surface. A cross-sectional view is shown in Fig. 8. The top ground plane of the strip-line configuration and the upper half of the dielectric laminate are not present in this case. This gives rise to the following distinguishing features of the microstriplines when compared with the striplines discussed earlier.

a) There is an easy access to the top surface of microstripline which makes it very convenient to mount discrete (active or passive) devices and to make minor adjustments after the circuit has been fabricated.

b) Because of the open nature of the structure, care has to be taken to minimize the radiation loss or interference due to the nearby conductors. In order to ensure that the microstripline fields are confined near the strip, use of higher dielectric constant substrates becomes necessary. This is advantageous from miniaturization point of view, since higher ϵ_r reduces the guided wavelength and hence the circuit dimensions.

c) Since the electromagnetic fields extend in the space (above the strip), the microstrip configuration is a mixed dielectric transmission structure. This makes the analysis and design more complicated.

Effective dielectric constant

The concept of effective dielectric is useful for transmission lines with more than one type of dielectrics filling the cross-section. The effective dielectric constant ϵ_{re} is defined such that if the transmission line cross-section is filled up uniformly with a material with $\epsilon = \epsilon_0 \epsilon_{re}$, the resulting structure will have the same phase velocity as the transmission line with composite dielectric media. Thus

$$\epsilon_{re} = (c/v_p)^2 \tag{30}$$

To illustrate the concept of ϵ_{re}, let us consider a parallel-plate line partially filled with a dielectric as shown in Fig. 9. Out of the width w, a section nW is filled with dielectric. x is a fraction less than unity and may be called filling fraction. Phase velocity for this parallel plate line may be obtained by calculating L and C and using $v_p = 1/\sqrt{LC}$. The inductance L does not depend on the dielectric filling* and is given $L = \mu_0 h/W$. The capacitance C can be considered a parallel combination of two capacitances. We get

$$C = \epsilon_0 \epsilon_r \frac{xW}{h} + \epsilon_0 \frac{(1 - x) W}{h} = \epsilon_0 \frac{W}{h} \left\{ x \epsilon_r + 1 - x \right\} \tag{31}$$

The phase velocity v_p may be derived as

$$v_p = 1/\sqrt{LC} = 1/\sqrt{\left\{\frac{\mu_0 h}{W} \cdot \epsilon_0 \frac{W}{h} \left[1 + x(\epsilon_r - 1)\right]\right\}}$$

$$= c/\sqrt{1 + x(\epsilon_r - 1)} = c/\sqrt{\epsilon_{re}} \tag{32}$$

which yields the value of effective dielectric constant ϵ_{re} as

$$\epsilon_{re} = 1 + x(\epsilon_r - 1) \tag{33}$$

We note that when $x \to 0$, $\epsilon_{re} \to 1$ and when $x \to 1$, $\epsilon_{re} = \epsilon_r$. Thus the value of ϵ_{re} varies between 1 and ϵ_r depending upon the value of the filling fraction x. For $x = 0.05$, $\epsilon_{re} = (\epsilon_r + 1)/2$, i.e., the average of the two dielectric constant values ϵ_r and 1. An alternative expression for ϵ_{re} can be written in terms of characteristic impedance Z_0 of the line. We have

$$\epsilon_{re} = (c/v_p)^2 = \left\{ (1/\sqrt{LC_0}) / (1/\sqrt{LC}) \right\}^2 = C/C_0 \tag{34}$$

* For non-magnetic dielectrics

C_0 is the capacitance for a transmission line of the same configuration but with the dielectric replaced by air. Since $Z_0 = \sqrt{L/C}$, (34) may be rewritten in terms of Z_0 as

$$\epsilon_{re} = \frac{L}{Z_0^2} \cdot Z_{oa}^2/L = (Z_{oa}/Z_0)^2 \tag{35}$$

where Z_{oa} is the characteristic impedance of the transmission line with dielectric medium replaced by air. It may be noted that, since we are dealing with non-magnetic dielectric materials the inductance L remains unchanged when the dielectric is replaced by air.

The concept of effective dielectric constant as applied to the microstripline is illustrated in Fig. 10. Because of the more complicated geometry of a microstrip line, calcuation of ϵ_{re} is not as simple as for parallel-plate waveguide discussed above. Electrostatic calculation of two capacitances C and C_0 are needed. In practice, a quasi-empirical expression suggested by Schneider* is used extensively for the design of microstrip circuits. According to this formula

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} (1 + 10\, h/W)^{-1/2} \tag{36}$$

It is instructive to look at the variations of ϵ_{re} with microstrip geometry. Typical distributions of electric and magnetic field lines in a microstrip structure are shown in Fig. 11. The electric field lines extend in the air region also, and this causes ϵ_{re} to be less than the substrate ϵ_r. The

* Schneider, M.V., "Microstrip lines for microwave integrated circuits," Bell System Tech. J., Vol. 48, 1969, pp 1421-1444.

relative amount of E-field in the air region decreases when the strip-width W is increased. In the limit when W/h becomes very large, the microstrip line converges to a parallel plate line and ϵ_{re} should tend to ϵ_r itself. In the limit $h/W \rightarrow 0$, (36) also yields $\epsilon_{re} = \epsilon_r$. The other limit of ϵ_r is reached when the line widths are very small ($W/h \rightarrow 0$). For this case, (36) yields $\epsilon_{re} = (\epsilon_r + 1)/2$. Physically, this is interpreted by observing that when $W/h \rightarrow 0$, the geometry reduces to that of a thin wire placed on a semi-infinite dielectric slab as shown in Fig. 12. This may also be viewed as a coaxial line configuration with outer conductor at an infinite distance. It can be shown that for a coaxial line half-filled with a dielectric as shown in Fig. 12, effective dielectric constant is the average value of ϵ_r and 1, which agrees with (36) in the limit $W/h \rightarrow 0$. Thus it may be concluded that ϵ_{re} for a microstrip line varies from $(\epsilon_r + 1)/2$ when W is small (i.e., for high values of Z_0) to ϵ_r when W is large (or when Z_0 has a very small value). In typical cases ϵ_{re} for a 50 ohm line on a polystyrene substrate (ϵ_r = 2.50, h = 1/16 inch) is 2.078, while that for a 50 ohm line on alumina (ϵ_r = 9.8, h = 0.025 inch), the value of ϵ_{re} is 6.606.

Relation (36) is used extensively for calculating ϵ_{re}. However, Owens,* found that a modified form yields more accurate results (± 0.25 percent). This formula is

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2}(1 + 10\,h/W)^{-0.555} \tag{37}$$

* Owens, R.P., "Accurate analytical determination of quasistatic microstrip line parameters," Radio Electronic Engineer, Vol. 46, No. 7, July 1976, pp 360-364.

The ratio of free space wavelength to the wavelength along a microstripline is given by $\sqrt{\epsilon_{re}}$. This ratio is plotted in Fig. 13 as a function of W/h for different values of the substance dielectric constant.

Characteristic Impedance of Microstrip Lines

As for the case of ϵ_{re}, calculations of Z_0 for microstrip circuit design are also carried out by using quasi-empirical closed-form relations. A set of fairly accurate (± 1 percent) and simple expressions for Z_0 in terms of W, h, and ϵ_r is as follows* for narrow strips (W/h less than 3.3):

$$Z_0 = \frac{119.9}{\sqrt{2(\epsilon_r + 1)}}\left[\ln\left\{4\frac{h}{W} + \sqrt{16(h/W)^2 + 2}\right\} - \frac{1}{2}\left(\frac{\epsilon_r - 1}{\epsilon_r + 1}\right)\left(\ln\frac{\pi}{2} + \frac{2}{\epsilon_r}\ln\frac{4}{\pi}\right)\right] \tag{38}$$

For wide strips (W/h greater than 3.3):

$$Z_0 = \frac{119.9}{2\sqrt{\epsilon_r}}\left[\frac{W}{2h} + \frac{\ln 4}{\pi} + \frac{\ln(e\pi^2/16)}{2\pi}\left(\frac{\epsilon_r - 1}{\epsilon_r^2}\right) + \frac{\epsilon_r + 1}{2\epsilon_r}\left\{\ln\frac{\pi e}{2} + \ln\left(\frac{W}{2h} + 0.94\right)\right\}\right]^{-1} \tag{39}$$

* Edwards, T.C., Foundations for microstrip circuit design, New York: John Wiley & Sones, 1981, p 45.

where 'e' is the exponential base: e = 2.7182818. Formulas for synthesis (calculation of W/h for given Z_0 and ϵ_r) are as follows: For narrow strips (when $Z_0 > \{44 - 2\epsilon_r\}$ ohms):

$$\frac{W}{h} = \left\{ \frac{\exp A}{8} - \frac{1}{4 \exp A} \right\}^{-1} \tag{40}$$

where $A = \frac{Z_0 \sqrt{2(\epsilon_r + 1)}}{119.9} + \frac{1}{2} \left(\frac{\epsilon_r - 1}{\epsilon_r + 1}\right) \left(\ln \frac{\pi}{2} + \frac{1}{\epsilon_r} \ln \frac{4}{\pi}\right)$. For wide strips (when $Z_0 < \{44 - 2\epsilon_r\}$ ohms)

$$\frac{W}{h} = \frac{2}{\pi} \left\{ (B - 1) - \ln (2B - 1) \right\} + \frac{\epsilon_r - 1}{\pi \epsilon_r} \left\{ \ln (B - 1) + 0.293 - \frac{0.517}{\epsilon_r} \right\} \quad \text{where } B = 59.95\, \pi^2 / (Z_0 \sqrt{\epsilon_r}) \tag{41}$$

Recently Wheeler* has derived an expression which holds for narrow as well as wide strips. This expression has slightly poorer accuracy than those given above, but may be easily programmed on hand held calculators. These are:

$$\frac{W}{h} = 8 \frac{\left\{ \left[\exp \left(\frac{Z_0 \sqrt{\epsilon_r + 1}}{42.4}\right) - 1\right] \frac{7 + 4/\epsilon_r}{11} + (1 + 1/\epsilon_r)/0.81 \right\}^{1/2}}{\exp \left\{ \left(\frac{Z_0 \sqrt{\epsilon_r + 1}}{42.4}\right) \right\} - 1} \tag{42}$$

* Wheeler, H.A., "Transmission line properties of a strip on a dielectric sheet on a plane," IEEE Trans. Microw. Theory Tech., Vol. MTT-25, p 631-647 (1977).

For analysis, this expression may be rewritten as

$$Z_0 = \frac{42.4}{\sqrt{\epsilon_r + 1}} \ln \left\{ 1 + \left(\frac{4h}{W}\right) \left[\left(\frac{14 + 8/\epsilon_r}{11}\right) \left(\frac{4h}{W}\right) + \sqrt{\left(\frac{14 + 8/\epsilon_r}{11}\right)^2 \left(\frac{4h}{W}\right)^2 + \frac{1 + 1/\epsilon_r}{2} \pi^2} \right] \right\} \tag{43}$$

In the above results for microstrip design, the strip thickness t has been assumed to be negligible. However, when $t/h \leq 0.005$, the agreement between experimental results and calculated results for t = 0 is found to be very good. For larger values of t, formulas for finite strip thickness are needed. A set of simple and accurate formulas for finite t are as given below:*

$$Z_0 = \begin{cases} \dfrac{60}{\sqrt{\epsilon_r}} \ln \left\{ \dfrac{8h}{W_t} + 0.25 \dfrac{W_t}{h} \right\} & W/h \leq 1 \quad (44) \\[2ex] \dfrac{376.7}{\sqrt{\epsilon_{re}}} \left\{ \dfrac{W_t}{h} + 1.393 + 0.667 \ln \left(\dfrac{W_t}{h} + 1.444\right) \right\}^{-1} & W/h \geq 1 \quad (45) \end{cases}$$

* I.J. Bahl and R. Garg, "Simple and accurate formulas for microstrip with finite strip thickness," Proc. IEEE, Vol. 65, Nov. 1967, pp 1611-1612.

where $\frac{W_t}{h} = W/h + \Delta W/h$ with

$$\frac{\Delta W}{h} = \begin{cases} \frac{1.25}{\pi} \frac{t}{h} \left(1 + \ln \frac{4\pi W}{t}\right) & (W/h \leq 1/2\pi) \\ \frac{1.25}{\pi} \frac{t}{h} \left(1 + \ln \frac{2h}{t}\right) & (W/h \geq 1/2\pi) \end{cases} \tag{41}$$

Effective dielectric constant ϵ_{re} for thick microstrip is given by

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} F\,(W/h) - G\,(W/h,\ t/h) \tag{47}$$

in which the functions F and G are given by

$$F\left(\frac{W}{h}\right) = (1 + 10\ h/W)^{-1/2} \text{ and } G = \frac{\epsilon_r - 1}{4.6} \frac{t/h}{\sqrt{W/h}} \tag{48}$$

Variation of microstrip impedance with W/h is shown in Fig. 14.

Losses in Microstrips

Conductor losses in microstrip lines also may be calculated by using Wheeler's incremental inductance rule discussed earlier in connection with stripline losses. Equation (24) can be applied for microstrips also, recessions in various walls being taken as shown in Fig. 15. Symbols Δ_g and Δ_s denote skin depths for the ground plane conductor and strip conductor respectively. It may be noted that the air gap resulting from the upward recession of the bottom surface of the strip conductor is accounted while modifying h since inductance calculations do not depend upon the presence of dielectric.

Dielectric Losses

Dielectric loss in a TEM mode with uniform dielectric filling is given by (28). For a line with composite dielectric (such as a microstrip line), the concept of effect tan δ is introduced. We can write

$$(\tan \delta)_e = \frac{\sigma_e}{\omega \epsilon_{re} \epsilon_o} \tag{49}$$

where σ_e is the effective conductivity of the dielectric medium, given by

$$\sigma_e = x\,\sigma + (1 - x)\,\sigma_o \tag{50}$$

where x is the filling fraction, similar to that used for dielectric constant in (33). In (50) σ is the conductivity of the dielectric substrate that σ_o is the conductivity of the air above the substrate. Since the air is almost lossless, $\sigma_o \simeq 0$ and

$$\sigma_e = x\sigma \tag{51}$$

Also $\epsilon_{re} = x\,\epsilon_r + (1 - x)$

$$\text{or } x = (\epsilon_{re} - 1) / (\epsilon_r - 1) \tag{52}$$

Using the above definitions, attenuation caused by dielectric losses given by (28), may be modified for composite dielectric transmission lines as

$$\alpha_d = \frac{\omega}{2}\sqrt{\mu_o \epsilon_o \epsilon_{re}}\,(\tan\delta)_e = \frac{\omega}{2}\sqrt{\mu_o \epsilon_o \epsilon_{re}}\,\frac{\sigma_e}{\omega \epsilon_o \epsilon_{re}}$$

$$= \frac{\omega}{2}\sqrt{\mu_o \epsilon_o}\,\frac{1}{\sqrt{\epsilon_{re}}} \cdot x \,\frac{\sigma}{\omega \epsilon_o \epsilon_r} \cdot \epsilon_r \tag{53}$$

$$= \pi\,(\tan\delta)\,\sqrt{\epsilon_r}/\lambda_o \cdot \sqrt{\epsilon_r/\epsilon_{re}} \cdot \frac{\epsilon_{re}-1}{\epsilon_r-1}\ \text{neper/m}$$

$$= 27.3\,(\tan\delta) \cdot \frac{1}{\lambda_o}\,\frac{\epsilon_r}{\sqrt{\epsilon_{re}}} \cdot \frac{\epsilon_{re}-1}{\epsilon_r-1}\ \text{dB/m}$$

Dielectric losses become significant only at very high frequencies (e.g. millimeter waves) or when semiconductor substrates are used.

Dispersion in Microstrip Lines

All TEM mode lines are capable of propagating higher order modes at higher frequencies. However, in case of the microstrip line composite dielectric configuration causes the dominant mode itself to be slightly non-TEM. This can be shown by studying the transverse field distribution of E and H-field and applying boundary conditions at dielectric-air interface.

The distribution of the electric field lines (shown in Fig. 11) indicates that E-lines approach the dielectric-air interface obliquely, and thus both x and y directed transverse components of the electric field are present.*

---

* Coordinate system shown in Fig. 16.

Since the tangential component (x-directed) of E-field must be continuous at air-dielectric interface, the tangential component of the electrical flux density D becomes discontinuous, i.e.,

$$D_x/\text{air} \neq D_x/\text{diel} \tag{54}$$

Using Maxwell's equation for $\nabla \times H$, we can write

$$(\nabla \times H)_{x/\text{air}} \neq (\nabla \times H)_{x/\text{diel}} \tag{55}$$

If H_z is zero (i.e., if the mode is pure TEM), (3.49) yields

$$\partial H_y/\partial z_{/\text{air}} \neq \partial H_y/\partial z_{/\text{diel}} \tag{56}$$

or at the interface

$$H_y/\text{air} \neq H_{y/\text{diel}} \tag{57}$$

Inequality in (57) violates the field matching conditions for the normal component of magnetic field. Thus, it may be concluded that H_z should be a non-zero quantity for inequality in (55) and consequently the continuity of the tangential component of E-field at dielectric-air interface to be satisfied. The above argument leads to the conclusion that a pure TEM mode cannot be supported by a microstrip line. However, since the major portion of the electric field lines is concentrated below the strip, the electric flux crossing the dielectric-air boundary is small. Therefore, the deviation from TEM mode is small and may be ignored for most of the circuit design applications. The microstrip design formulas discussed earlier are based on this approximation.

The non-TEM nature of microstrip modes causes the microstrip lines to exhibit dispersion. When the frequency of a signal exciting a microstrip line is (say) doubled, the phase constant $\beta(=2\pi/\lambda_g)$ is not exactly doubled. Several transmission line structures and waveguides exhibit thisbehavior, and are known as dispersive lines or waveguides. Dispersive behavior of microstrip lines may be studied by carrying out a detailed hybrid-mode full-wave analysis* of microstrip line as a wave guiding structure. There are several methods available for such an analysis, but these become too complicated to be incorporated in a CAD program for microstrip circuits. Therefore, for CAD purposes, the dispersive nature of the microstripline is modelled approximately by considering a parallel-plate waveguide model** of the microstrip line, and considering the width and ϵ_{re} of this parallel-plate waveguide model to vary with frequency.

Measurements on microstrips*** show that the effective dielectric constant $\epsilon_e(f)$ is a function of frequency varying from ϵ_{re} in quasistatic limit to ϵ_r at very high frequencies. The following formula is suggested for $10 \leq \epsilon_r < 12$

* K.C. Gupta, "Microstrip Lines and Slot Lines," Dedham (MA) Artech House, 1979, Chap. 4.

** Kompa, G., and Mehran, R., "Planar waveguide model for calculating microstrip components," Electron. Lett., Vol. 11, 1975, pp 459-460.

*** Edwards, T.C. and Owens, R.P., "2-18 GHz dispersion measurements on 10-100 ohm microstrip lines on sapphire," IEEE Trans. MTT-24, No. 8, August 1976, pp 505-513.

$$\epsilon_e(f) = \epsilon_r - (\epsilon_r - \epsilon_{re}) / \left\{ 1 + \left(\frac{h}{Z_o}\right)^{1.33} (0.43\, f^2 - 0.009\, f^3) \right. \qquad (58)$$

where h is in millimeters and f is in gigahertz . It may be noted that dispersion effect given by (58) is significant only at x-band (and higher) frequencies where alumina substrates ($\epsilon_r \gtrsim 10$) are widely used. Corresponding measurement results and closed-form formulas are not available for other dielectric constant substrates.

In addition to variation of phase velocity with frequency, characteristic impedance of the microstrip line also varies with frequency. Using a parallel-plate waveguide model, we can write

$$Z_o(f) = \frac{\eta h}{W_e(f) \sqrt{\epsilon_e(f)}} \qquad (59)$$

where η is the wave impedance of free-space and $W_e(f)$ is the frequency dependent effective width of the microstrip line. According to Owen*

$$W_e(f) = W + \frac{W_e - W}{1 + (f/f_p)^2} \qquad (60)$$

where W_e is the quasistatic value of the effective width given by

$$W_e = \frac{\eta h}{Z_o \sqrt{\epsilon_{re}}} \qquad (61)$$

and f_p is the cut-off frequency of next higher order mode given by

* Owen, R.P., "Predicted frequency dependence of microstrip characteristic impedance using the planar waveguide model," Electron. Lett., Vol. 12, 1976, pp 269-270.

$$f_p = \frac{c}{2\, W_e \sqrt{\epsilon_{re}}} \tag{62}$$

It may be noted the frequency variation of Z_0 depends upon $W_e(f)$ as well as $\epsilon_e(f)$. However the effect of $W_e(f)$ dominates and Z_0 decreases as the frequency increases. For a 50 ohm line on alumina substrate the change is about 10 percent over the 0 to 16 GHz range.

SELECTED BIBLIOGRAPHY

1. Harlan Howe, Jr., "Stripline circuit design," Artech House, 1974.
2. K.C. Gupta et al, "Microstrip Lines and slot lines," Artech House, 1979.
3. K.C. Gupta et al, "CAD of Microwave Circuits," Artech House, 1981, Chap. 3 on "Characterization of transmission structures," pp 47-90.
4. T.C. Edwards, "Foundations for microstrip circuit design," John Wiley and Sons, 1981.

FIGURES

Fig. 1(a) A stripline
Fig. 1(b) A microstripline

Fig. 2(a) Cross-section of a stripline
Fig. 2(b) Practical embodiment of a stripline configuration

Fig. 3 Electrical and magnetic field distributions in a stripline

Fig. 4 A parallel plate model of stripline configuration

Fig. 5 Variation of $Z_0\sqrt{\epsilon_r}$ with W/b for a stripline

Fig. 6(a) Parallel plate transmission line
Fig. 6(b) Incremental displacement of walls

Fig. 7 Incremental displacement of walls for calculating conductor losses in a stripline

Fig. 8 Cross-sectional view of a microstripline

Fig. 9 A parallel plate line with mixed dielectric filling

Fig. 10 Concept of effective dielectric constant (ϵ_{re}) as applied to a microstripline

Fig. 11 Distributions of electric and magnetic fields in a microstripline

Fig. 12 Geometry of microstripline in the limit W/h → [illegible]

Fig. 13 Ratio of λ_0 to λ (wavelength along the microstripline) as a function of W/h

Fig. 14 Variation of microstrip impedance with W/h

Fig. 15 Wheeler's incremental rule applied to a microstrip configuration

Fig. 16 A typical fringing - field line in microstrip configuration

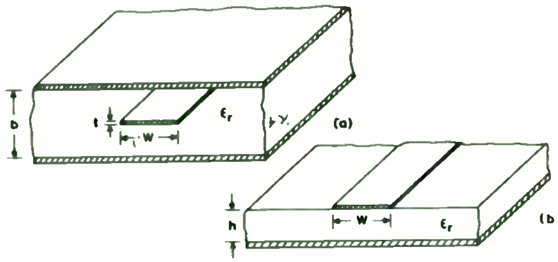
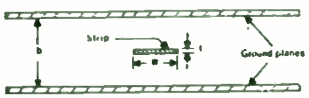


Fig. 1 (a) A stripline (b) A microstripline



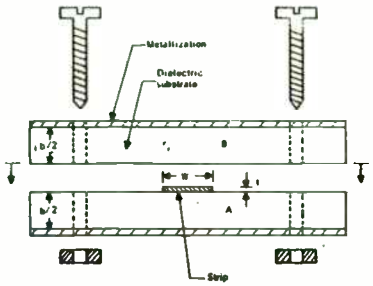


Fig. 2

(a) Cross-section of a stripline

(b) Practical embodiment of a stripline configuration

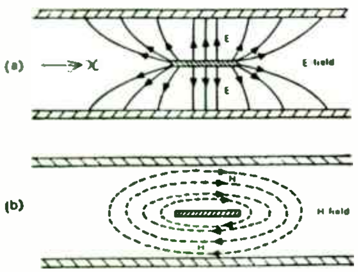


Fig. 3 Electrical and magnetic field distributions in a in a stripline

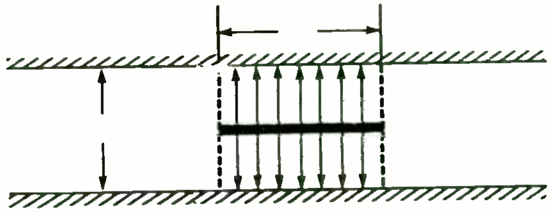


Fig. 4 A parallel plate model of stripline configuration

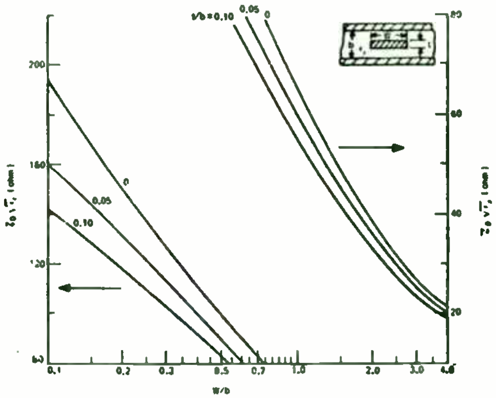
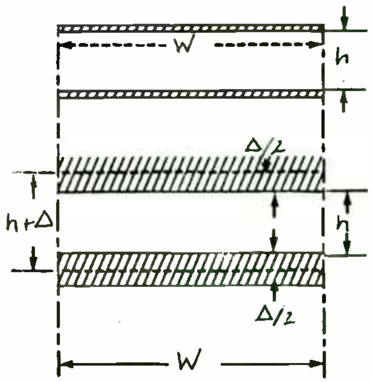


Fig. 5 Variation of $Z_o\sqrt{\epsilon_r}$ with W/b for a stripline



(a) Parallel plate transmission line

(b) Incremental displacement of walls

Fig. 6 Wheeler's incremental inductance rule

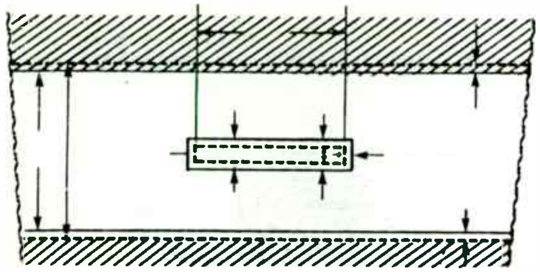


Fig. 7 Incremental displacement of walls for calculating conductor losses in a stripline

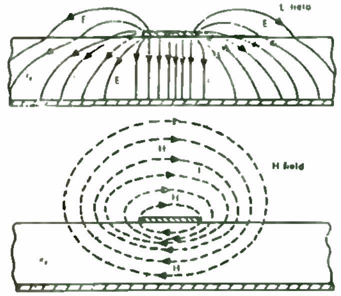
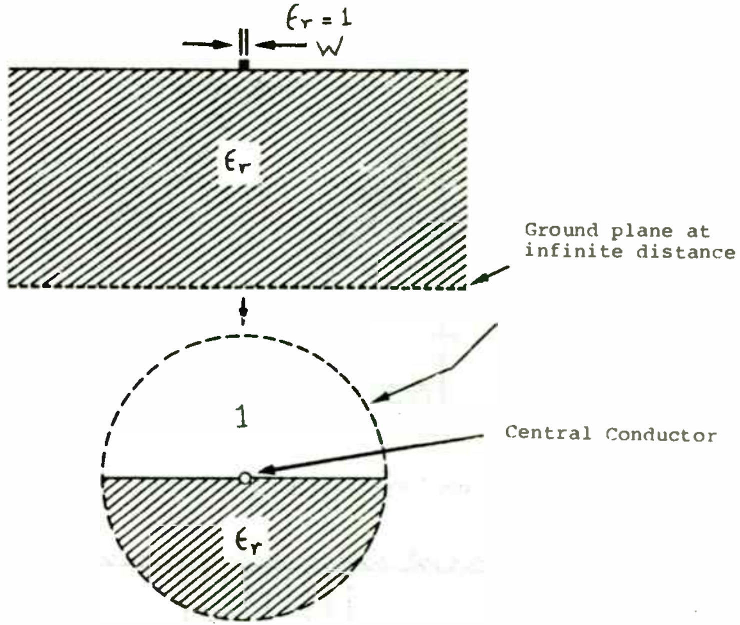


Fig. 11 Distributions of electric and magnetic fields in a microstripline

Fig. 8 Cross-sectional view of a microstripline



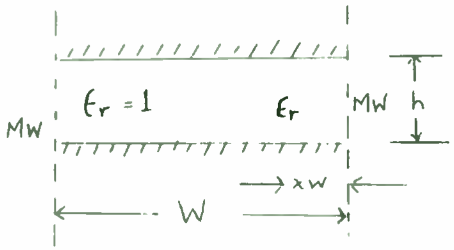


Fig. 9 A parallel plate with mixed dielectric filling

Fig. 12 Geometry of microstripline in the limit $W/h \to \infty$

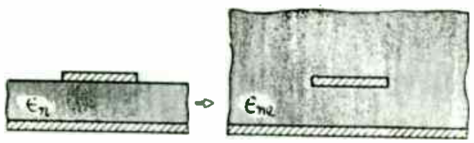


Fig. 10 Concept of effective dielectric constant (ϵ_{re}) as applied to a microstripline

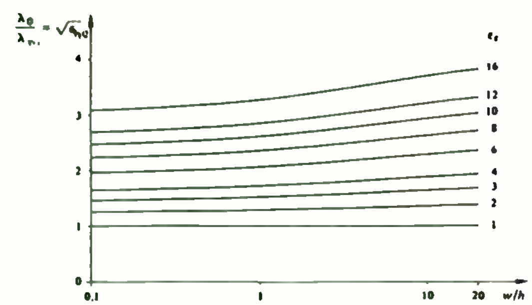


Fig. 13 Ratio of λ_o to λ_m (wavelength along the microstripline) as a function of W/h

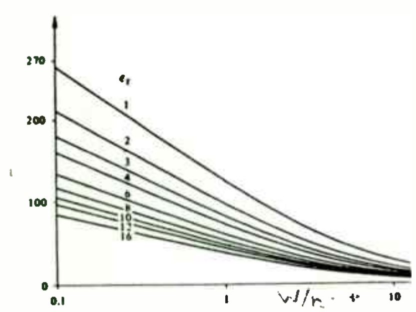


Fig. 14 Variation of microstrip impedance with W/h

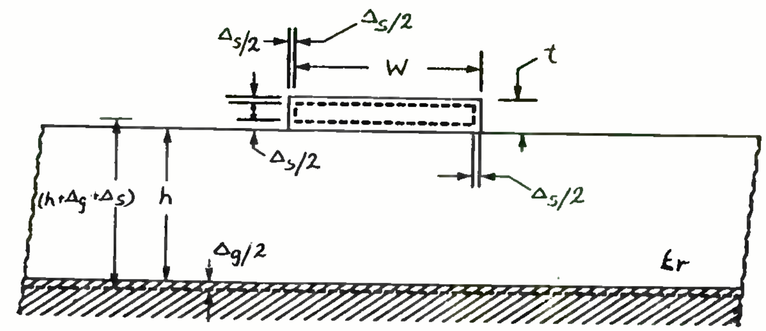


Fig. 15 Wheeler's incremental rule applied to a microstrip configuration

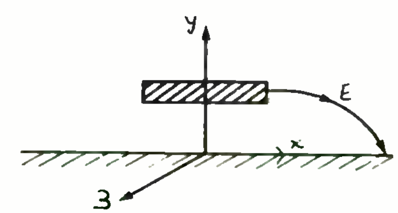


Fig. 16 A typical fringing E-field line in a microstrip configuration

HIGH POWER SOLID STATE CIRCUIT DESIGN

by:

J. H. Johnson

MICROWAVE MODULES & DEVICES, Inc.

MMD

550 ELLIS ST., MOUNTAIN VIEW, CA 94043 (415) 961-1473

HIGH POWER SOLID STATE CIRCUIT DESIGN

The intent of this session is to present a practical approach to RF circuit design for high power applications. With the instrumentation available today very accurate transistor characterization is possible. Combine this data with a properly selected transistor, good CAD and sound circuit design techniques and all the mystery that once surrounded such designs is gone. But, of course, every good engineer knows the most fundamental law, "If anything can go wrong it will". With this in mind, I would like to bring to your attention a new set of fundamental proverbs for RF Design:

- <u>If you ever see an oscillation in your design model, all production models will oscillate</u> -
  It is critically important that every possible oscillation under all possible load, drive and supply voltage conditions be identified and eliminated.

- <u>The transistor selected is never rugged enough</u> -
  In order to meet difficult performance standards, the circuit designer often sacrifices device ruggedness. Design ruggedness must be made the number one priority if you are to avoid unexplained failure, production yield problems and field failure caused by a slowly degrading load impedance.

- <u>All design margin will go away in production</u> - Design with as much margin as possible. In many cases, adding an additional stage for gain will save many times its cost in production headaches.

- <u>The transistor will always run too hot</u> - Thermal problems are the least understood of all the design problems an RF engineer faces. A direct measurement of the junction temperature under real operating conditions should be the final test.

- <u>If one of a combined group of output transistors blows they will all blow</u> - Combine transistors using techniques that provide the necessary isolation to avoid secondary destruction.

- <u>The parasitic inducance you are forced to live with is always too much</u> - Keep all parasitic inductances as small as possible so that you can maintain control over the circuit in production. The parasitics are most likely to change with time.

- <u>The transistor manufacturer will never be able to make the same device in production</u> - Use transistors that are industry standards and being manufactured in volume. Avoid moving too close to the state-of-the-art.

It is interesting and even fun to look at what can go wrong in a high power circuit; but, with good engineering design all of the common problem areas can be controlled.

TRANSISTOR SELECTION

One of the most important aspects of a good manufacturable circuit is the proper selection of the power transistors. The primary considerations are as follows:

- <u>Power Output</u> - Remember that the transistor thermal resistance, your system thermal resistance, and the permissible junction temperature determine the maximum power output. The manufacturers rated power out and even the saturated power out may be good indicators of margin, but the junction temperature sets the power output limit.

- <u>Ruggedness</u> - Will the transistor survive a badly mismatched load? Do not use a fragile transistor! If a fragile transistor is the only one you can get then reduce the supply voltage until it is adequately rugged. Reducing the voltage is much more effective than reducing drive, although reducing drive a bit helps. For amplifiers that run open loop directly into an antenna the power transistors must be able to withstand an ∞ VSWR at the lowest frequency of operation, high line voltage and maximum drive. For an amplifier with a closed loop for gain control that drives a combiner or other system component the capability to deliver rated power into a 3:1 load is a good test.

- Margin - Actual performance in a circuit is worth more than all the specifications and it is critical that the required performance is easily achieved with good margin. The easier you make it on your transistor manufacturer the easier you make it on yourself. One rule of thumb is, "Find out what the transistor manufacturer can make and find a way to use it".

- State-of-the-Art Transistors - If your performance requirements are state-of-the-art then use the latest available transistor and accept the risk involved with anything new. If your performance requirements are simple, select a transistor that has been around for a few years and is made in volume. Stay away from one-of-a-kind laboratory curiosities completely.

- Reliability - Reliability is always important regardless of the end use. Transistor metalization is important and gold metal should be used for UHF and higher frequencies. At VHF and HF aluminum is fine. Other considerations are equally important.
  1. Check a few devices for void free die attach.
  2. Look at that the overall device assembly.
  3. Make sure you have good package integrity.
  4. Good wire bonds are very important. Excessive wirebond current and the resultant bond flex can become a major failure mode in some devices.

- Balanced or Single Ended - There is nothing magic about a balanced transistor. It is simply a very useful way to combine a pair of transistors for twice the power without the decrease in impedance associated with direct paralleling. It is a particularly useful technique for wideband circuits.

- Package - The most important feature of any package is whether or not it is an industry standard. Try to select a package with minimum parasitic inductance, particularly the common lead.

- FET or BJT - There is nothing magic about either type of transistor and the choice should be made based on performance requirements. The most outstanding advantage of the FET is in its low transmitted noise. The BJT in general offers more power from a single device and is cheaper.

- The Subtle Differences - There are some subtle differences in transistors from various manufacturers that at times are very important:
  1. Ballast resistors can cause problems. Diffused resistors, while generally preferred for rugged performance, degrade noise, linearity and back IMD performance. They also perform poorly in high radiation environments.
  2. Look at both input and output impedances over a wide bandwidth and make sure there are no sharp peaks or dips.

TRANSISTOR CHARACTERIZATION

The first step is to quickly determine if the transistor you intend to use is suitable for the bandwidth that you require. First lets look at an output model assuming that the real load (R_L) needs to be:

$$R_L = \frac{(Vcc-Vsat)^2}{2\ Pout}$$

Now add the transistor and package parasitics:

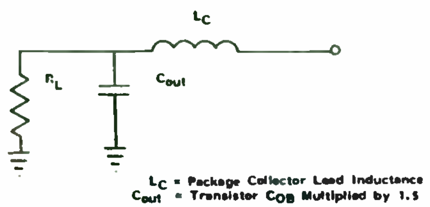


Fig. 1– Model of Transistor Output

Figure 2 shows the effect of various values for C_{OB} and L_C on the value of the real series load impedance over a common UHF bandwidth. Note that the series real load impedance is not effected by the value of L_C. L_C does, however, determine the value of the equivalent parallel resistanct at the package edge. If this value of parallel resistance (R_p) is higher than the geometric mean of the equivalent device series resistance (R_s) and the matching circuit final termination resistance (R_T), very wide bandwidths may not be possible.

Once you are satisfied that the load impedance is achievable, take a quick look at the input:

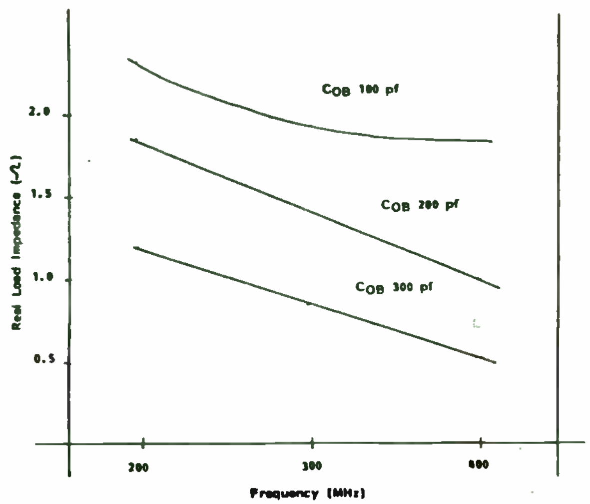


Fig. 2 - Effect of C_{out} on Real Load Impedance

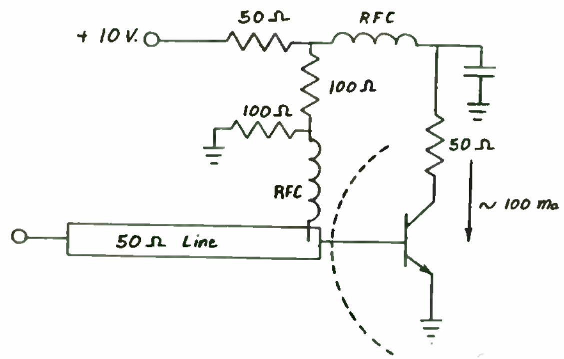


Fig. 3 - Input Z Test Circuit

The simple test fixture shown in Figure 3 turns the transistor on, loads the collector and allows the engineer to look at the input Z directly with a network analyzer. The values for Z_{in} will not be precise, but the shape of the Z_{in} versus Frequency curve will be. By looking at the shape of the Z_{in} curve, the engineer can quickly determine if it is well behaved over a specific frequency range. Other advantages of this quick measurement are:

- It is easy to compare transistors from various vendors or transistors with different internal matches.

- It is also possible to vary the collector load slightly with capacitive inductive loading to see how well the output is isolated from the input. The better the input/output isolation the easier the final design will be.

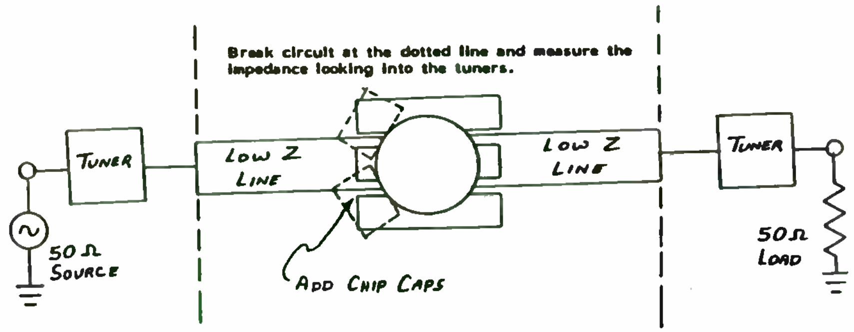


Fig 4. - Input or Output Z Measurement

Once the transistor selection is made (based on the quick Z tests) exact Z measurements must be made. If measurements are made at low impedance points,(as just discussed) the instrumentation accuracy is poor and the parasitics associated with making contact to the device are large compared to the transistor impedance. The best solution to the problem is to build a portion of the circuit in close to the transistor where the impedances are low and make the measurement at a high impedance point. Construct a circuit similar to Figure 4 with chip caps and line lengths that bring Z values up to 25 to 50 ohms. Beyond the low Z line use a tuner to complete the impedance match. Tune the circuit for desired performance. If the expected performance cannot be achieved then too much of the impedance match is being required of the tuner and accurate values of Z's cannot be measured. Once good performance is achieved, break the circuit at the dotted line and look back into the tuned section with a network analyzer. The impedance at this point should be in the 15 to 50 ohm range. Use the measured value as the termination for the transistor section of the circuit. By using the known values for the circuit components the actual transistor impedance can now be calculated (use CAD). The transistor input impedance is the conjugate of the calculated value.

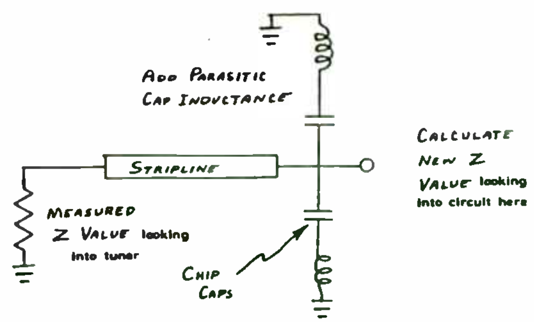


Fig. 5 - Z Calculation

BASIC DESIGN CONCEPTS

Think Low Impedance - A transistor is usually a low impedance device and in many cases an extremely low impedance device. Designing with low impedance active components requires that the engineer has the proper concept. He must think in a way that is, in many cases, opposite his previous experience:

- Current is more critical, than voltage. Components must be capable of carrying high RF currents with low loss.

- Lead inductance is important. Any lead inductance associated with the transistor or capacitors may severely degrade the amplifier performance.

- Ground paths must also be considered. In high impedance circuits the ground path is carried from component to component on a printed circuit run. This cannot be done with low impedance circuit design! A continuous groundplane on the back side of a printed circuit board is an ideal arrangement. Remember parasitic inductance in the ground path is of equal importance to the signal path.

The important thing to remember when working with low impedances is to keep all parasitic and loss terms an order of magnitude below the element being used.

Mounting the Transistor - The first step is to mount the transistor and the printed circuit board in such a way as to minimize parasitic inductances for the transistor and printed circuit board ground path. The transistor leads should be on an even plane with the printed circuit board (See Figure 6.). The space between the printed circuit board and the transistor body should be minimized.

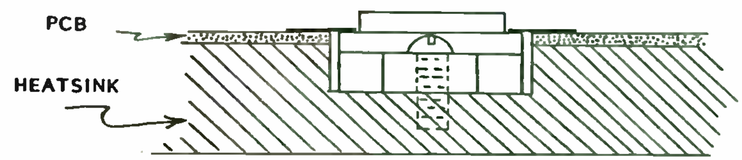


Fig. 6 - Transistor Mounting

Grounds - When designing RF power amplifiers, the technique used to ground the various components is so important that it deserves special attention. Several tips listed below will help optimize your amplifiers. REMEMBER GROUNDING BECOMES MUCH MORE CRITICAL AT EITHER HIGH POWER LEVELS OR HIGH FREQUENCIES!

- Ground the transistor common leads at the body of the transistor. Not at the ends of the leads! Not 1/8" away from the body! (See Figure 7).

- The back side of the printed circuit board should be nearly a continuous ground plane. The top side ground should be connected to the bottom side ground using straps or plated through holes under each common lead. (See Figure 8).

- The contact between the printed circuit board and the heatsink (or mounting plate) must be continuous and intimate near the body of the transistor particularly if a common flange transistor is ued. If the contact is made away from the body of the transistor, the effective common lead inductance is increased.

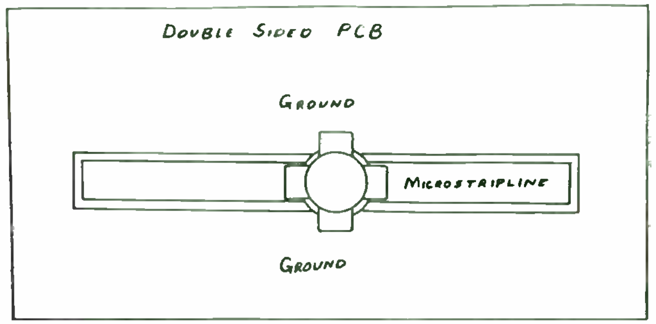


Fig. 7 - Common Lead Grounding at the Transistor Body

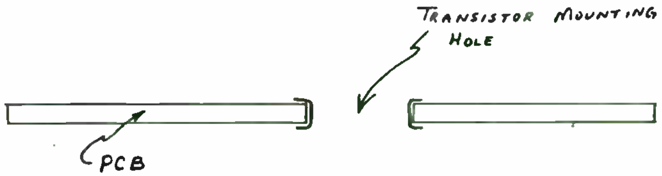


Fig. 8 - Topside Ground is Connected to the Backside Ground Using Straps or Plated Through Holes

- Components in the matching networks have critical grounds also. The ground for capacitors used near the transistor is perhaps the most critical. Remember the shunt capacitance required here is often 1 or 2 ohms and therefore the total inductive impedance in the ground return to the common lead must be extremely small. Two capacitors in parallel, one back to each common lead, will reduce the inductive impedance in the ground return and divide the high RF currents.

- Capacitors in the matching networks, at a slightly higher impedance point, still require a good ground. A direct connection to the continuous back side ground using a strap through a hole in the board or a plated through hole is the best technique.

- Single plated through holes to reach the back side ground of a printed circuit board can provide mixed results. To improve effectiveness, use many holes together.

Component Location - It is very important when using high power transistors, to make the first impedance matching step close to the transistor. By using this simple approach the overall circuit bandwidth is greater and the effective common lead inductance is reduced (See Figure 9).

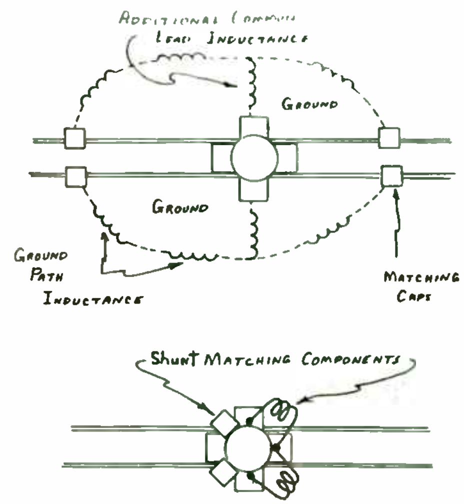


Fig. 9 Effective Common Lead Inductance May be Reduced by Proper Component Placement.

Circuit Stability - Circuit stability is one of the biggest problems an RF engineer faces. Generally a circuit/transistor combination will oscillate for three reasons:

- Excessive gain (particularly at low frequencies) combined with inphase feedback.
- Varactor type oscillations at frequencies where the circuit provides the required lead.
- Various circuit or device resonances convert a stable common emitter stage to common collector or common base at the resonant frequencies.

By carefully designing the circuit and mechanical layout to avoid the above problems, most circuits can be stabilized. An example will describe how good circuit design can eliminate the stability problem for a common emitter amplifier.

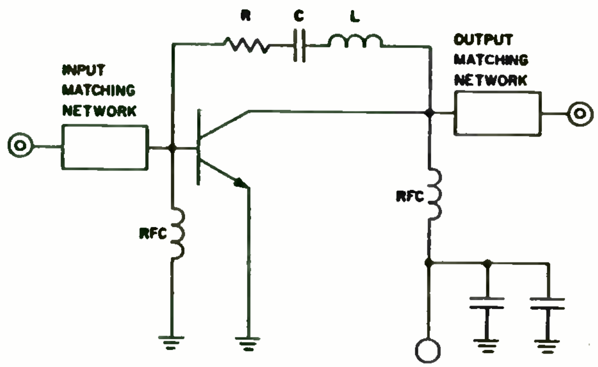


Fig. 11 - Collector to Base Feedback Reduces the Transistor Low Frequency Gain

- Any resistance or inductance in the transistor emitter provides negative feedback which decreases gain and makes the transistor more stable. A transistor with large emitter resistors is easier to stabilize. A transistor with very low emitter lead inductance (the best stripline packages) is slightly more difficult to stabilize. (But it has more gain.)

- A transistor built on higher resistivity material will have a higher collector resistance and seems to be more stable.

Oscillations due to the varactor mode or resonant grounding are more difficult to eliminate. In most cases the transistor chip sees a zero or very low impedance at the frequency of oscillation. By carefully looking at the circuit near the device with a vector impedance meter or network analyzer, zeros can be spotted and eliminated. In many cases simply moving a component slightly or using two capacitors in parallel instead of one will solve the problem. One other check is often useful before beginning circuit design. Look at the transistor performance in a high quality test fixture to insure that there are no unusual ripples or negative resistance regions in the P_{in} versus P_{out} curve. The output power should also change smoothly with variations in V_{cc}.

Low frequency spurious signal generation in common emitter RF power amplifiers is due primarily to the extremely high (30-40dB) low frequency gain. Generally the transistor is merely oscillating in one of the classical modes rather than the more commonly discussed parametric modes. The input and output impedance required to sustain oscillation is provided by the dc feed networks, not the matching networks. There are two techniques to prevent these unwanted oscillations:

1. Present the device a source and load impedance which will not sustain any oscillation.
2. Lower the low frequency gain of the transistor.

The best source and load impedance to prevent low frequency spurious is a low pure resistance. The following circuit provides this termination. (See Figure 10).

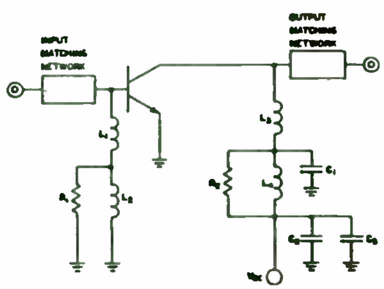


Fig. 10 - Resistive Loading of the Collector and Base Reduces the Tendency of the Transistor to Generate Spurious Signals.

- L_1 and L_3 need to be small RF chokes at the carrier frequency. L_2 and L_4 need to be as large as possible and still be able to handle the required current. (10µh is a good value). C_1 is a small bypass at the carrier frequency and should be as small as possible without cutting into the output power. C_2 and C_3 must provide a solid bypass at all frequencies including the very low ones. (.22µf and 10µf are good choices). At low frequencies the base will see R_1 and the collector will see R_2. R_1 and R_2 should be some low value like 10 ohms to 50 ohms. In less stubborn cases, the base feed network may be used with the simplified collector network shown in Figure 11 with good results.

The second method involves using negative collector base feedback to lower the below some selected frequency. (See Figure 11.) The values of the feedback network are selected as follows:

- L-make large enough so that the feedback network has no effect at the carrier frequency. The lead inductances of R and C are often enough without any additional L.

- C - make large enough for good coupling down to the lowest frequency of interest.

- R - a small value of about 10 ohms to 100 ohms is usually selected.

Both base and collector RFC's must be small to get the maximum benefit from the feedback. If either of these techniques is used, most transistors can be stabilized. However, a transistor can have several features that make it easier to stabilize.

- The low frequency gain of the transistor needs to be as low as possible. This is controlled mainly by h_{FE}. A value of 20 to 30 is usually more stable than a value of 80 to 100.

IMPEDANCE MATCHING NETWORKS

Most circuit designs today utilize a combination of discrete component impedance matching and a transformer or balun. Figure 12 shows a typical block diagram.

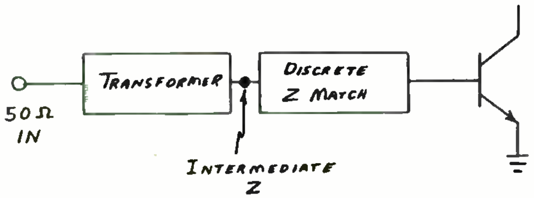


Fig. 12 - Typical Impedance Matching Approach

Generally the collector match will be similar to the input.

Discrete matching network - The discrete portion of the matching network can be designed using any one of several basic networks. Figure 13 shows a few commonly used examples. (See Figure 13).

As an example consider matching a transistor to an intermediate Z of 12.5 ohms. The transistor input is shown, but the output Z can be matched in a similar way.

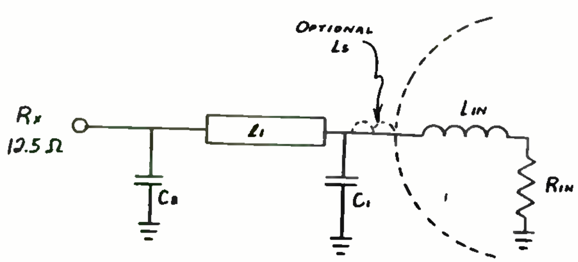
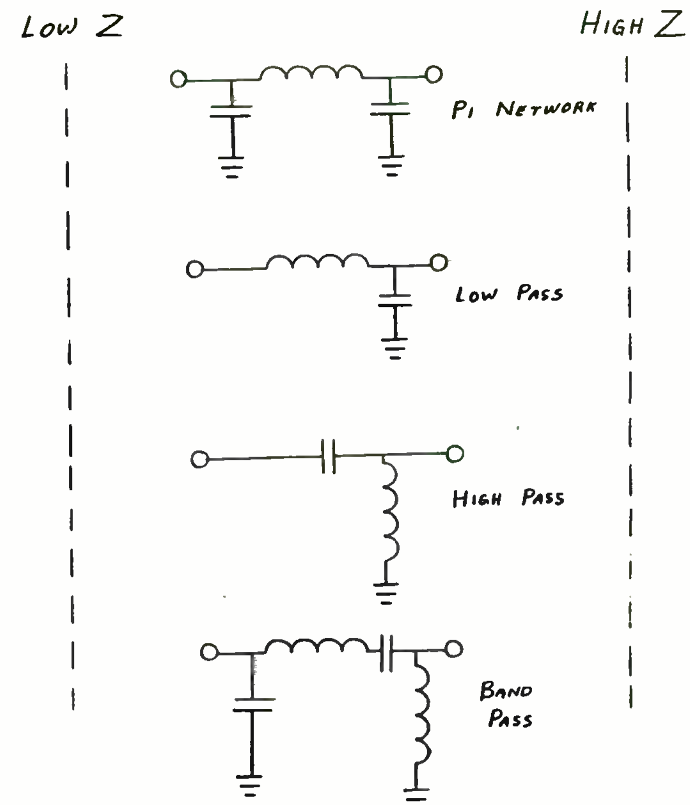


Fig 14. - Input Matching Network

Fig. 13 - Basic Matching Networks



Selected microstrip $_1$ to have a Z_o that is equal to R_x. By making this Z_o choice (even though it is not perfect from a bandwidth standpoint) the shunt capacitor may be moved along $_1$ in the final tracking of the circuit. If $_1$ turns out to be too long the extra length will have a Z_o equivalent to the circuit impedance at that point and therefore have no effect.

The next step is to convert the transistor series input impedance to parallel equivalents. Ideally the resulting R_p should be the geometric mean between R_{IN} and T_X.

$$R_p = \sqrt{R_{IN} \cdot R_X}$$

If R_p is less than $\sqrt{R_{IN} \cdot R_X}$ a small series inductance (L_s)* should be added to L_{IN} to increase R_p. If R_p is higher than optimum, live with it. Actual values for C_1, C_2 and the length of ℓ can be obtained from a Smith chart or just estimated from experience. Once the preliminary values are selected, enter them into a suitable CAD program that optimizes the values over the appropriate bandwidth. The Z_o and ℓ_1 could also be entered as a variable if the bandwidth is extremely critical. If a transformer (a 4:1 as example) is to be used, any output parasitics can also be added to the circuit.

If the bandwidth cannot be achieved with the simple pi network, a multiple step match should be tried.

*L_s is usually required on the transistor output. Generally C_1 needs to be as close to the transistor body as possible on the input.

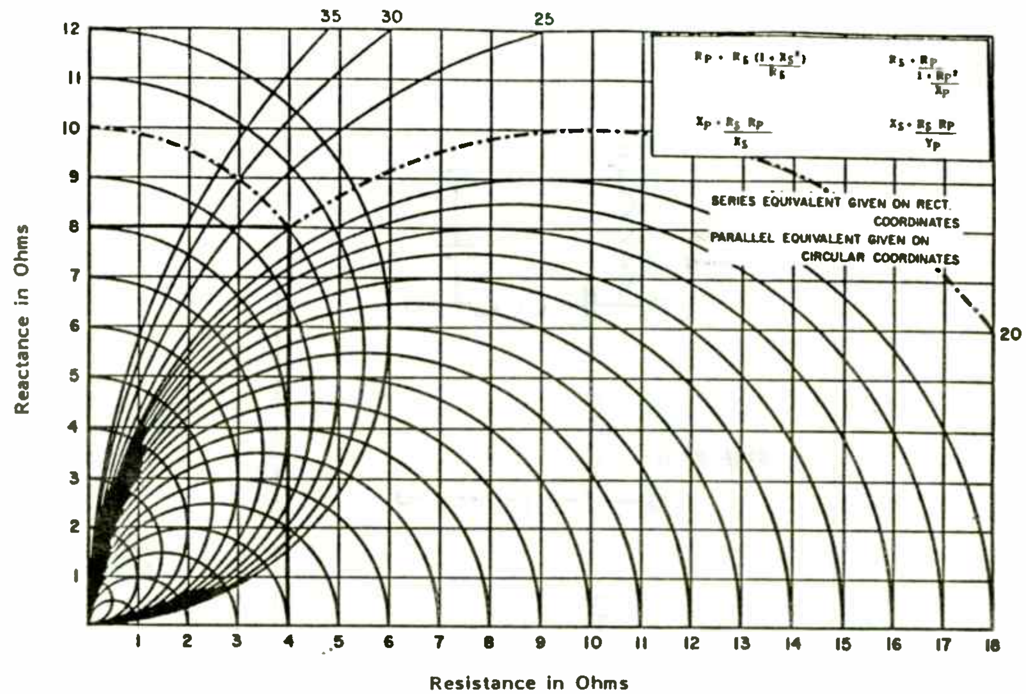
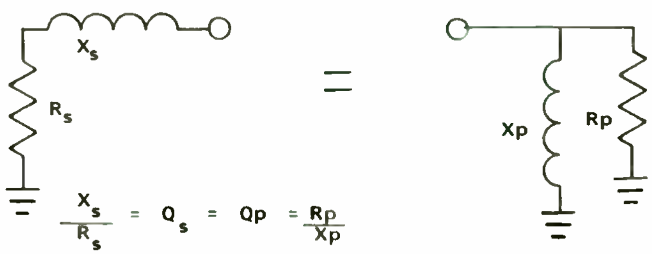


Fig. 15 - Series to Parallel and Parallel to Series Conversion - Example; if Zs = 4 + j8 then Zp = 20 + j10



TRANSFORMERS AND BALUNS

Transformers are extremely useful circuit elements because they can generally provide discrete predictable impedance transformation over very wide bandwidths. The 4:1 transformer is perhaps the most common of all. Several forms are shown in Figure 16.

(See Figure 16)

A quick look at currents and voltages in the 4:1 transformer explains how the impedance transformation occurs. Assume a resistive load R_L is added to the transformer as shown in Figure 17. Assume a current I is flowing through the load causing a voltage V across R_L.

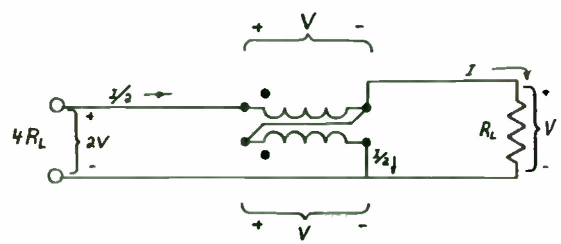


Fig. 17 - How Does the 4:1 Transformer Work

If V is present across R_L, the same voltage will be impressed across the secondary of the transformer. Since the turns of the secondary and primary are the same, the voltage V will be impressed across the transformer primary also. The polarity of the voltages is shown by the placement of the small dot on the transformer diagram. The voltage at the transformer input is the sum of the voltage across the transformer primary and R_L, or 2V. If a total current I is to flow through R_L, 1/2 must be provided from each transformer winding. Again since both primary and secondary have the same

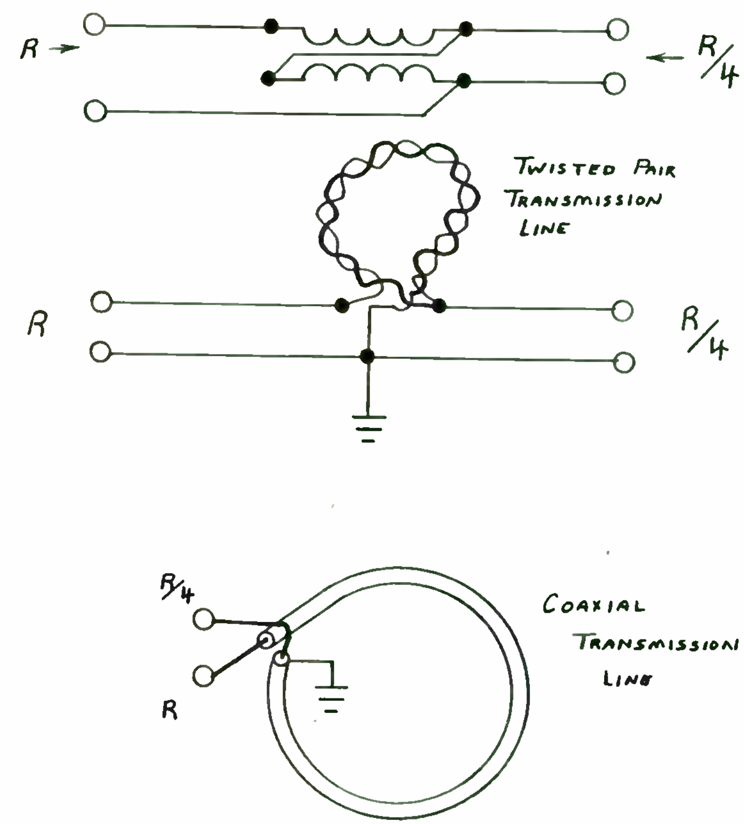


Fig. 16 - 4:1 Impedance Matching Transformers Can Be Constructed Using Any Suitable Transmission Line

number of turns (or the same voltage across them) the currents through each will be the same. Therefore, the input impedance will be $2V/I/2=4R_L$. More complex transformers provide other transformation ratios as shown in Figure 18.

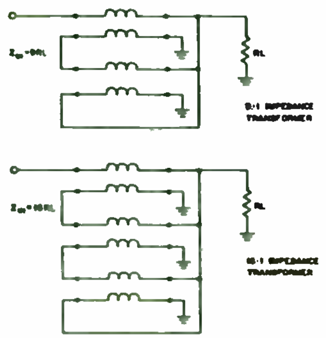


Fig. 18 - Typical Configuration of 9:1 and 16:1 Transformers

Transmission line transformers of this type can be used over bandwidths of more than 2 decades. The ideal transformation ratios are, of course, for low frequencies, but with a few precautions the upper frequency limit can be quite high (UHF):

- Transformer interconnects must be as short and precise as possible to avoid parasitics.

- Select a transmission line impedance (Z_o) that is the mean between input and output.

$$Z_o = \sqrt{R_{in} R_{out}}$$

- If the desired transmission line impedance is near 50 ohms a twisted pair of wires may be used. Multiple wires may be paralleled for a twisted pair impedance that is lower than 50 ohms. A coaxial cable is generally preferred at impedances much below 50 ohms. Semi-rigid coax is excellent.

- The length of the transmission line is also important. If it is too long ($\lambda/8$ is approximate maximum) the loss and phase shift will be excessive. If it is too short the impedance to ground (common or rejection mode impedance) will load the input signal.

- Core materials of various permeabilities may be used to surround the transmission lines to improve the performance of the transformer.

  a) Increases the series inductance of the transmission line conductors and thus extend the low frequency performance of the transformer.

  b) Does not alter the characteristic impedance of the transmission line.

  c) Allows the electrical length of the transmission line to remain relatively short.

  d) Is not the medium used to couple power from input to output; thus small cross section ferrite can accommodate large power levels and still remain unsaturated.

A simple balun transformer also has many applications. Figure 19 shows a 1:1 balun. The simple balun is usually used to connect a balanced load to an unbalanced system. It is also useful as a phase shift network.

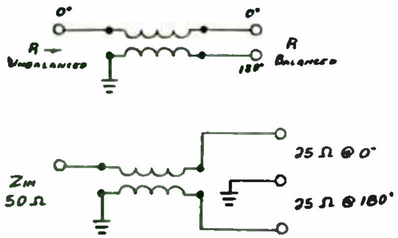


Fig. 19 - Simple Balun Transformer

You can also think of the balun as providing a pair of unbalanced outputs that are out of phase by 180°.

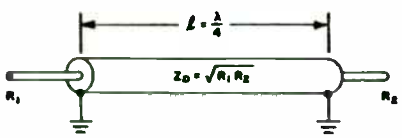


Fig. 20 – λ/4 Transformer

Quarter wave transmission lines also make useful impedance transformers A real impedance R_1 can be transformed to a value R_2 at a frequency f by using a transmission line such that the characteristic impedance is equal to $Z_0 = \sqrt{R_1 \; R_2}$ and the line is electrically a quarter wave at the frequency f (see Figure 20).

The smaller the transformation ratio, the larger the achievable bandwidth. A very common use of this technique is to match a 50 ohm system down to approximately 12.5 ohms over an octave bandwidth by using a 25 ohm transmission line that is a quarter wave long at the center frequency of the band.

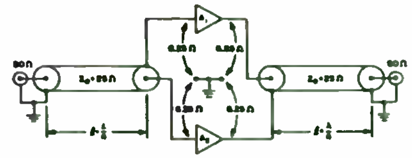


Fig. 21
Matching Networks for Push-Pull

The quarter wave transformer can also be used as a balun. By using two such baluns a very simple push-pull amplifier can be constructed. See Figure 21.

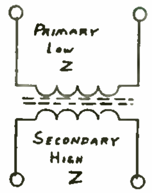
The input quarter wave section provides:

- 180° phase shift between input signals to amplifiers A_1 and A_2.
- 12.5 ohms between A_1 and A_2, or 6.25 ohms for each amplifier and ground.

The output quarter wave section provides:

- In phase combining of signals through the entire amplifier system (180° phase shift between signals of A_1 and A_2).
- 12.5 ohms between A_1 and A_2 or 6.25 ohms from each amplifier and ground.

Transformers that operate in the more classical transformer mode are quite useful at the lower frequencies (HF & VHF). One common type is constructed using ferrite beads, hollow brass tubes (some silver plated) and a few turns of insulated wire. Two parallel sections are constructed where the brass sleeves fit snuggly through the center of a length of ferrite beads. A copper strap is used to interconnect the two brass sleeves at one end of the transformer. This serves as a primary winding of the transformer. The secondary winding is formed by winding the appropriate number of turns of insulated wire through the center of the brass tubes. Use the largest possible wire size to improve performance. This construction places the current carrying surfaces very close to each other and also very close to the ferrite materials, thus minimizing the amount of leakage flux which tends to limit upper frequency performance of a classic transformer. The impedance transformation ratio is equal to the square of the turns ratio. This technique works very well up through 30 MHz and has actually been used successfully in wideband amplifiers up to 90 MHz. A further advantage of this technique is that it provides a means to construct push-pull amplifier modules that can be very easily reproduced. One very common use is for a 28 volt 100 watt, 2-30 MHz, push-pull, two transistor linear amplifier module.



Z Ratio	Primary Turns	Secondary Turns
4:1	1	2
9:1	1	3
16:1	1	4

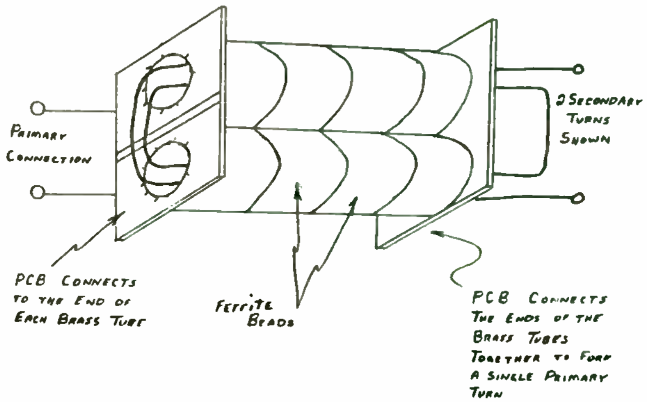


Fig. 22- Classical Impedance Matching Transformer

The 50 ohms transmission is correctly transmitted at all frequencies down to where the common mode impedance of the 25 ohm line is approaching 25 ohms. The operation of the network is as follows:

- The 25 ohm line is terminated at one end with 25 ohms.
- Then at the other end of the 25 ohm line, we have a "floating" 25 ohm impedance.
- Consequently, if we connect a 25 ohm load from one conductor (it can be either) to ground, we now have a total impedance of 50 ohms from the other conductor to ground.

If points 2 and 3 are connected together, we now have the classic 4:1 impedance ratio transmission line transformer. From our description of how Figure 23 works, it should be immediately apparent that the 4:1 transformer will only work well at frequencies where the phase shift between the voltages are 2 and 3 is small. This is true only at frequencies whose wave length is large compared to the length of the 25 ohms line or frequencies where the length of the line is exactly one or more wave length.

Consider now the network shown in Figure 24.

When the widest bandwidth is important, another class of transformer may be the best choice. This type of transformer was first described in the CTC "Solid Circuits" handbook as an "Equal Delay Transformer". A brief theoretical discussion will clarify the advantages of the "Equal Delay Transformer". The basic principle of operation of transmission line transformers is that a terminated transmission line has a well-defined impedance to any differential signal applied to it and a much higher impedance to any common mode applied to it (which may be enhanced by coiling the transmission line or using a ferrite core). Consequently, one end of the line can be considered as "floating" or "elevated" relative to the other end.

To illustrate what we are discussing, consider Figure 23.

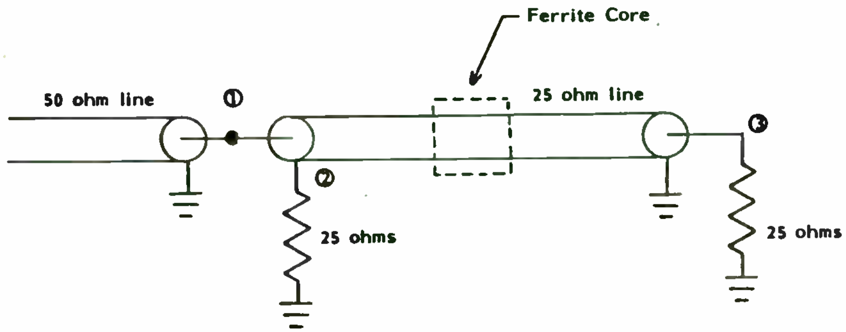


Fig. 23 - Understanding Transformer Operation

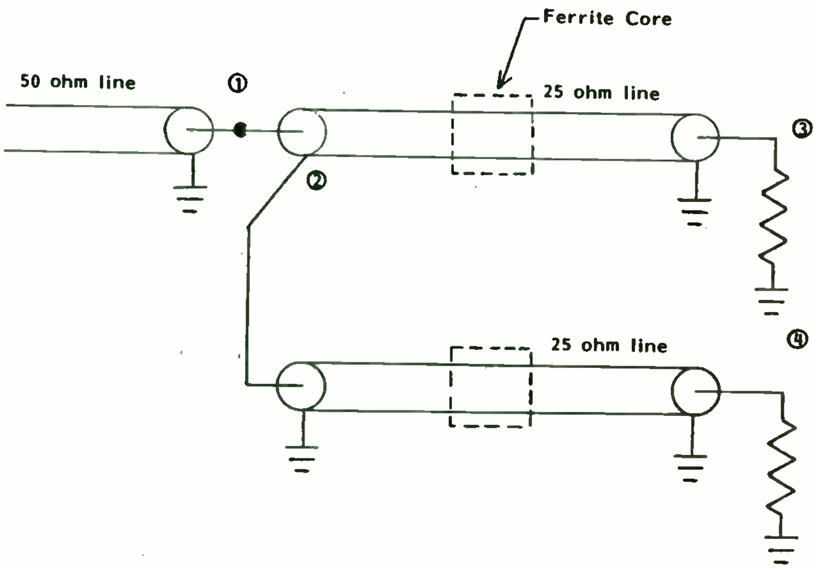


Fig. 24 - Equal Delay Transformer

The two 25 ohms lines are exactly the same length. The circuit in Figure 24 behaves exactly the same as that in Figure 23 . However, there is no phase difference between the voltages across both 25 ohm loads at high frequencies. Consequently, if we connect points 3 and 4 together, we have a 4:1 impedance ratio unbalanced to unbalanced transformer whose high frequency cut off is essentially unrelated to the electrical length of the 25 ohm.

The transformer, as represented by Figure 24 with points 3 and 4 interconnected, represents the very simplest of this class of transformer. With suitable ingenuity and more lines and cores, transformers can be configured which will have different impedance ratios, and/or balanced output, or act as hybrids with similar high frequency properties. (See Figure 25). This class of transformer should prove most useful when it is necessary to achieve extremely wide bandwidths.

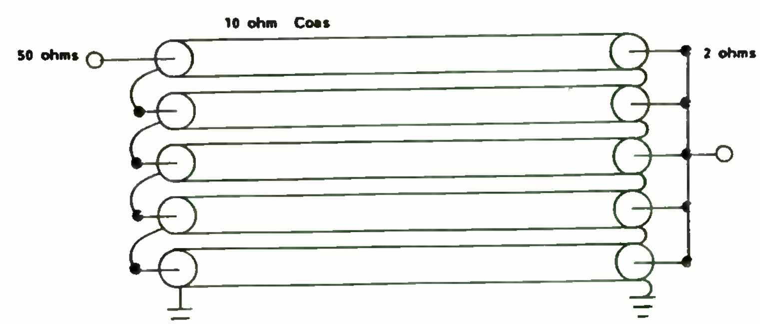


Fig. 25 - An Example of an Equal Delay Transformer

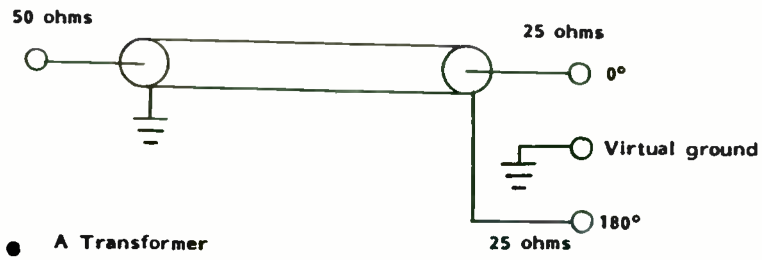
Balanced Transistors

Balanced transistors have become very popular throughout the industry since they were introduced by Lee Max at CTC in 1976 . A balanced transistor is simply a way of combining a pair of transistor chips that results in some very significant advantages. The two chips operate 180° out of phase with the other, with the midpoint being at RF ground. If good balance is maintained there will always be a zero RF potential at the midpoint (a virtual ground). The primary advantages are:

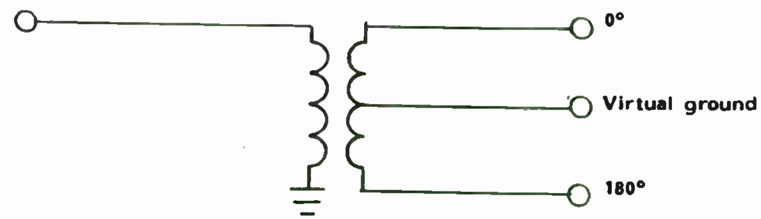
- Input Z is four times higher than a parallel combination of similar chips.
- Wider bandwidth is possible.
- A greater variety of impedance matching networks (internal to the transistor or external) is possible.
- Reduced even order harmonics.
- Better efficiency.
- The virtual ground that exists inside the transistor package reduces the effective common lead inductance.
- The internal structure of the transistor is simpler and less critical

Matching a balanced transistor is often somewhat frightening the first time. It is, however, quite simple. The 180° phase shift network can be achieved in several ways as shown in Figure 26:

- A Balun



- A Transformer
(Z determined by turns ratio)



- A microstrip phase shift

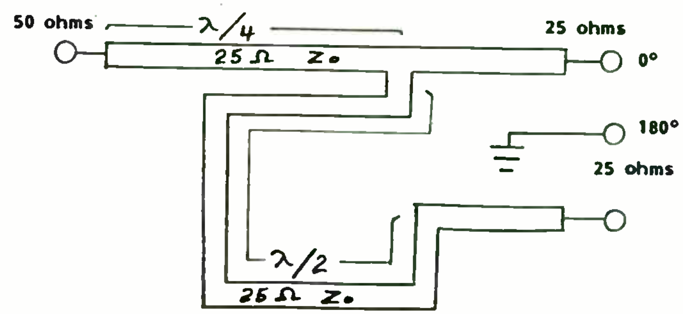


Fig. 26. - 108° Phase Shift Networks

Other parts of the matching network are usually calculated using measured values from one half of a balanced transistor as if it were a single ended device. (See Figure 27)

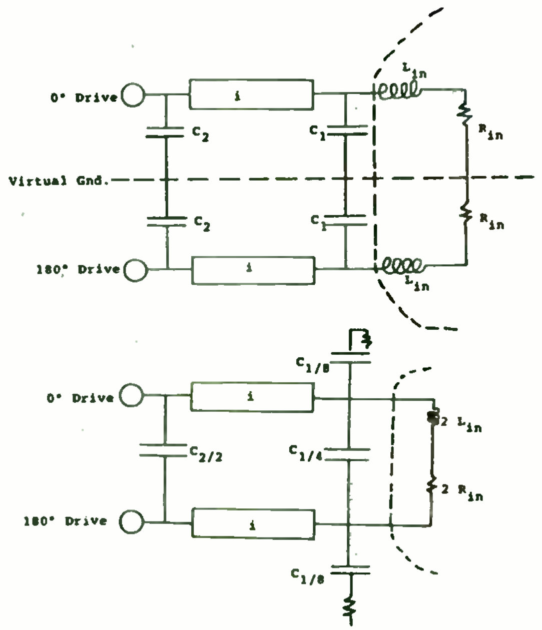


Fig. 27 - Converting Single Ended Impedances to Balanced Impedances

In a real circuit C_1 is spread among 3 capacitors to reduce parasitic inductance and divide the first RF currents. It is important to notice that the RF current through a capacitor in a balanced circuit such as C_2 is twice the value through C_2 in the single ended model.

In general a balanced transistor should be used in applications requiring very high power or wide bandwidth. From a systems standpoint, an ideal building block for very high power amplifiers is a hybrid combined pair of balanced transistors. This combination provides several important advantages:

- Low even harmonics
- Good 50 ohms input and output match
- Good delivered power into mismatched load

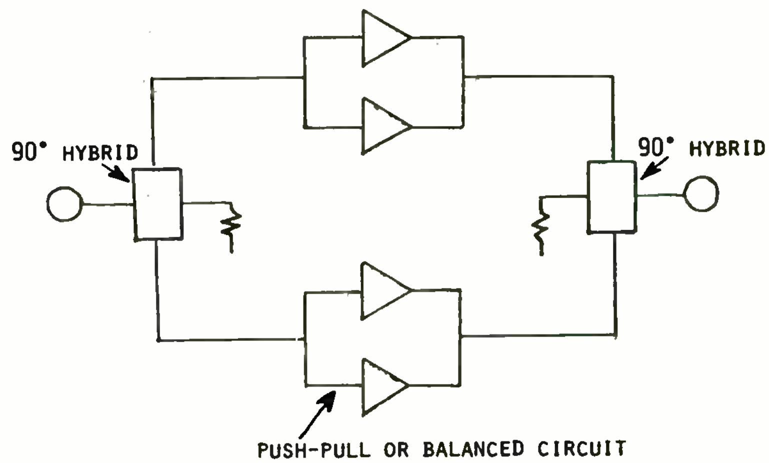


Fig. 28 - Ideal System Building Block

WHICH IS THE BEST CLASS OF OPERATION?

A, B or C. Most designers think of the Class C, zero bias, as the "normal" mode of operation. Class C is generally used in FM transmitters, CW power and a variety of pulse applications (radar). Class C is the easiest to design and is very efficient. Disadvantages are lower gain and poor linearity. Linearity can be greatly improved with a small bias or Class B operation. Usually the quiescent current of 50-100 ma is allowed to flow. An additional plus for Class B is about 2 dB improvement in gain. In many cases designers use Class B just to improve gain. If a bipolar transistor is used, some special precautions must be taken to achieve the appropriate bias over the temperature range. The bias voltage must track the change in the transistor VBE with temperature (approximately 2MV/°C). Figure 29 shows a good bias circuit that can provide peak base currents of several amps depending on the exact output transistor used. The diode used to sense the temperature should be closely coupled thermally to the power transistor. A common stud rectifier mounted to the heatsink adjacent to the transistor works nicely.

If greater linearity is required, Class A is required. Class A improves gain an additional dB over Class B, but the efficiency is reduced to 20-50%. Class A may be biased in a way similar to Class B or by using emitter degeneration or collector base dc feedback. Figure 30 shows a general example. If the degeneration or feedback is adequate, no other adjustments are required.

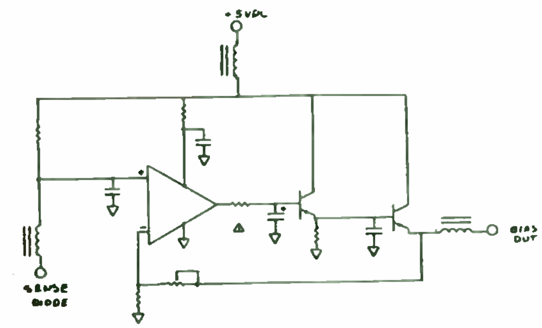


Fig. 29 - Typical Bias Circuit for a Class B (or AB) Amplifier

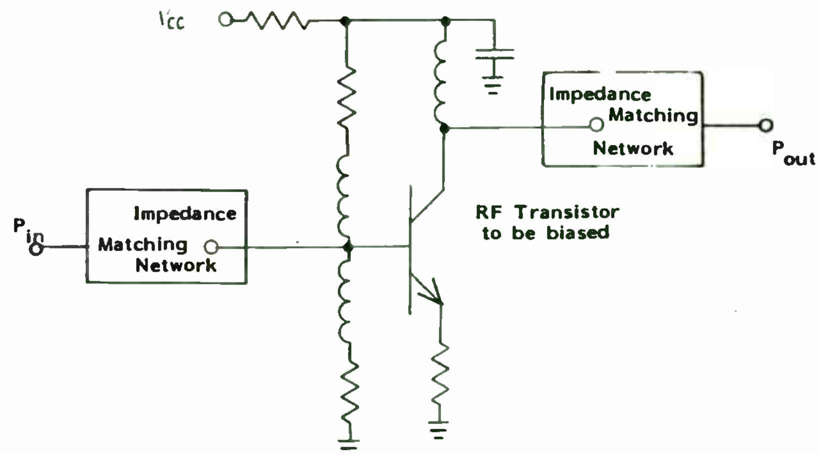


Fig. 30 - Typical Class A Bias Technique

Thermal Considerations

The one controlling factor in any high power RF design is the resulting transistor junction temperature under worst case conditions. Maximum permissible temperature will:

- Limit the power output
- Limit the allowed load mismatch
- Control the size of the heatsink

Generally, today's reliability considerations limit operating junction temperature to 150° to 160°C with excursions to 200°C under worst case load mismatch. This will vary, of course, depending on the transistor MTF and exact requirements.

The most important recommendation is that the design engineer measure the actual junction temperature and not depend on vendor supplied transistor data or calculations. Several important factors lead to this recommendation:

- Transistor manufacturers tend to publish thermal data that is measured under ideal conditions such as:
  1) Using lower than maximum rated power.
  2) Using dc power rather than RF power.
  3) Heatsink held to lower than realistic temperature.
- Many thermal resistance values of materials change with temperature.
- The thermal resistance value of the interfaces between materials is difficult to calculate.

If you have no means to make actual measurements and must rely on calculations, make sure that every heat barrier is considered. Each thermal resistance may be combined with the other values as with a ordinary resistance network.

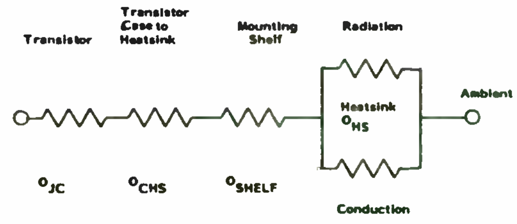


Fig. 31 - Levels of Thermal Resistance in Typical Transistor to Ambient Condition

The junction temperature may then be calculated as follows:

$$O_T \times P_D = T$$

O_T Combined total thermal resistance

P_D Total power dissipated

T Change in junction temperature

The transistor junction temperature changes rather drastically with temperature as shown in Figure 32.

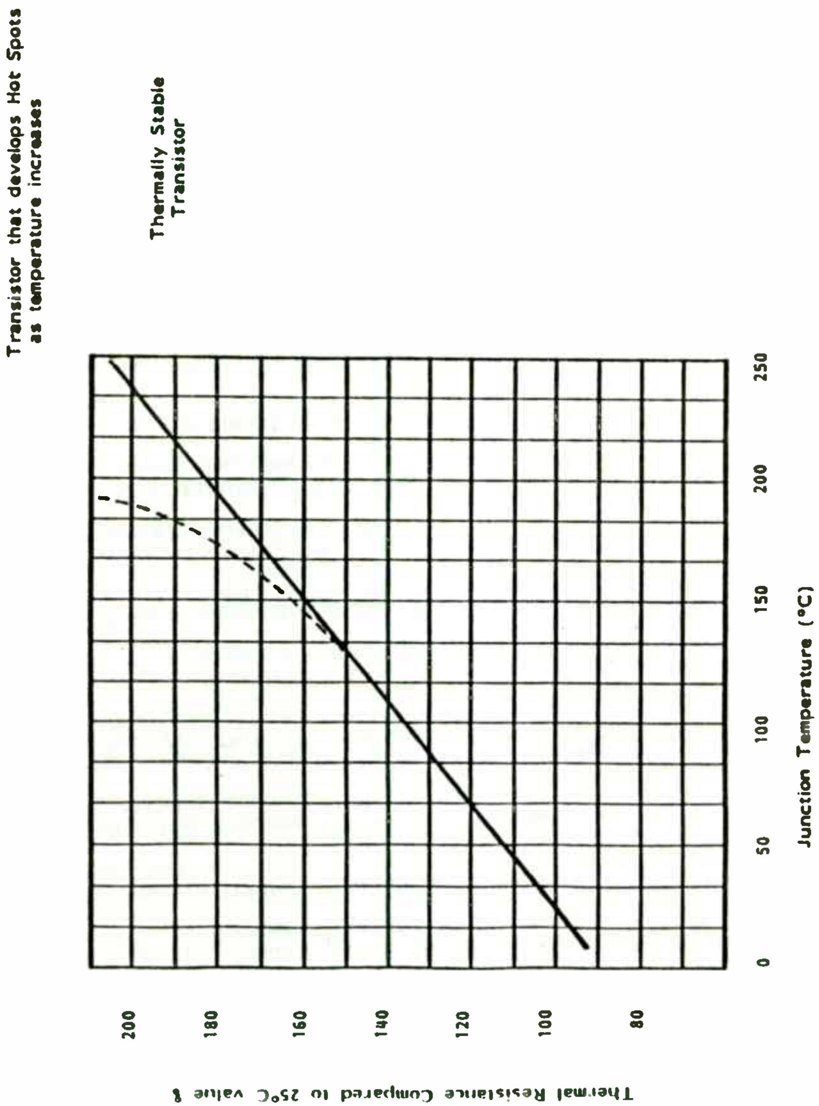


Fig. 32 - Change in Thermal Resistance with Temp

Time also plays an important part in thermal design as shown in Figure 33 and Figure 34.

Fig. 33 - Range of Temp. Rise for Transistor Chips when rated power is applied at time = O

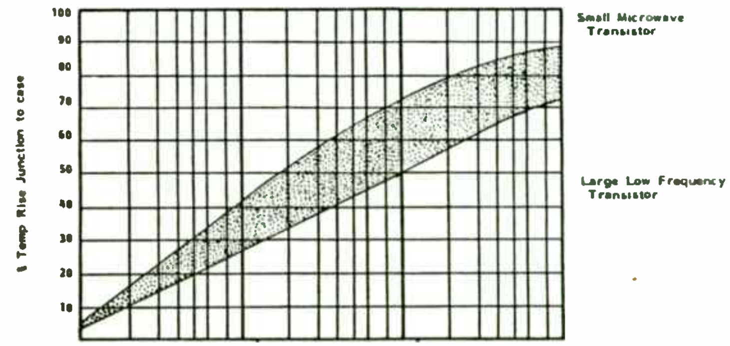
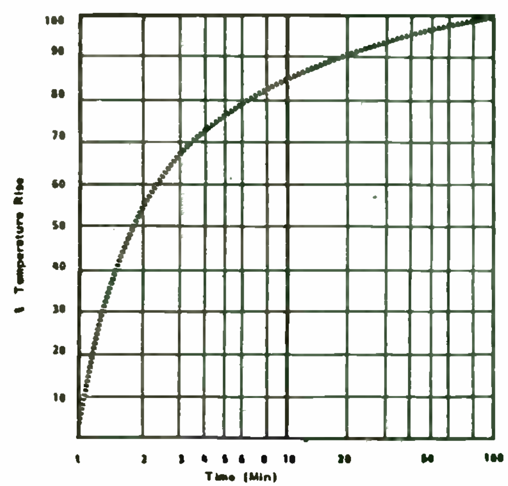


Fig. 34 - Typical Temperature rise with time for a finned heatsink (θ_{HA} = -.8°C/W) in Still Air (30°C) after a continuous 200W of power begins to dissipate



REFERENCES

1. Johnson, J., "Balanced Transistors: A Look Inside Those Integrated Two-Chip Amplifiers," MicroWaves, Vol. 19, No.2, Feb. 1980,pp 54-59

2. Max, L.B., "Balanced Transistors: Apply Wideband Techniques to Balanced Amplifiers," MicroWaves, Vol. 19, No.4, April, 1980, pp.83-88

3. Bailey, R.L., "Large-Signal Nonlinear Analysis of a High-Power High-Frequency Junction Transistor," IEEE Trans. on Electron Devices, Vol. ED-17, No.2

4. Johnson, J.H.,(Editor), Solid Circuits: Circuit Techniques, Applications Book 2.2.8.0A, Communications Transistor Corporation, March, 1973

5. Wisherd, D., "900 MHz 12 Volt Applications", Applications Note 2.1.8.8A CTC, October, 1974

6. C.L. Ruthroff, "Some BroadBand Transformers", Proceedings of the IEEE, Vol. 47, pp. 1337-1342 (August 1959)

7. Octavius Pitzalis and T.P. Course, "Broadband Transformer Design for RF Transistor Power Amplifiers," 1968 Electronic Components Conference Proceedings (May 1968).

8. L.B. Max, "Balanced Transistors: A New Option for RF Design," Micro-Waves, Vol. 16, No. 6, pp. 42-46 (June 1977).

9. F.H. Raab, "Output Circuit Analysis of Class A and B Push-Pull RF Power Amplifiers," Paper prepared for Communications Transistor Corp. (March 1979)

10. M.V. Joyce and K.K. Clarke, Transistor Circuit Analysis, Addison-Wesley Publishing Company, Reading, MA (1961)

11. H.L. Krauss, C.W. Bostian and F.H. Raab, Solid-State Radio Engineering, John Wiley and Sons, Inc. New York, NY, to be published

12. B.E. Rose, "Notes on Class D Transistor Amplifiers," IEEE Journal of Solid State Circuits

13. R.L. Bailey, "Large-Signal Nonlinear Analysis of A High-Power High-Frequency Junction Transistor," IEEE Transactions on Electron Devices, Vol. ED-17, No.2, pp. 108-119 (February 1970)

14. F.H. Raab, "The Class BD High-Efficiency RF Power Amplifier:, IEEE Journal of Solid State Circuits, Vol. SC-12, No. 3, pp. 291-298 (June 1977)

15. F.H. Raab, "Effects of Circuit Variations On The Class-E Tuned Power Amplifier", IEEE Journal of Solid State Circuits, Vol. SC-12, No. 2 (April 1978)

16. K.K. Clarke and D.T. Hess, Communications Circuits: Analysis and Design, Addison-Wesley Publishing Company, Reading, MA(1971)

17. C. Bowick, RF Circuit Design: Howard W. Sams & Co, Inc, Indianapolis, Indiana, (1982)

MICROSONICS INCORPORATED

60 WINTER ST. WEYMOUTH, MA 02188-3336
(617) 337-4200 TWX 710-338-6833

OSCILLATORS

CRITICAL SYSTEM

BUILDING BLOCKS

John R. Morton
Engineering Manager
Microsonics Incorporated
60 Winter Street
Weymouth, MA 02188

A SUBSIDIARY OF SIGNAL TECHNOLOGY CORPORATION

MICROSONICS INCORPORATED

OSCILLATORS

CRITICAL SYSTEM - BUILDING BLOCKS

Index

MICROSONICS INCORPORATED

INTRODUCTION

Oscillator circuits are fundamental building blocks used in almost every electronic system developed; analog, R.F. or digital. Systems that we use everyday could not function without them. For example there is at least one oscillator in your digital watch or clock, television, radio, VCR, microwave oven, automatic garage door opener, and personal computer.

Oscillators may be thought of as the heart of electronic systems since they provide the timing and synchronising for digital systems and they are used as frequency standards for RF and analog systems.

Since oscillators are critical system building blocks and the number of applications are increasing as systems become more complex, many engineers are finding themselves confronted with the challenge to design, specify and buy oscillators that provide optimum performance which is tailored to their specific system requirements. For the new engineer especially this can be a difficult task. It is indeed rare that oscillator circuits or devices reported in the literature can be used without modifications since detail, not described fully in the description, often are significant for optimum system performance.

It is the purpose of this paper to provide some insight into:

1) What an oscillator is and its applications.
2) Fundamental design conditions.
3) Commonly classified circuit examples.
4) Oscillator technolog comparisons.
5) Crystal oscillators.
6) Oscillator specifications and definitions of terms.
7) Performance trade-offs.

MICROSONICS INCORPORATED

BACKGROUND

Electromagnetic wave theory shows that a signal can be radiated in space effectively only if the radiating antenna is on the order of one-tenth (1/10) or more of the wave length corresponding to the frequencies of the signals to be radiated. The first problem of wireless communications was that of transmitting low frequency information (i.e.,voice) using practically small antennas. As an example: to radiate a maximum voice frequency of 10kHz an antenna of approximately 30,000 meters in length would be required. Additionally multiple transmissions on a single channel create severe interference and signal distortion.

To solve this problem the process of frequency translation or modulation was developed. The process of shifting a low frequency information spectrum to a higher frequency, that can be radiated with a practical antenna, is accomplished by multiplying the low frequency information signal by a high frequency sinusoidal (carrier) signal at the transmitter. Recovery of the information signal at the receiver is then accomplished by the use of appropriate demodulation techniques that retranslate the information hack to its original position. This technique also allows multiple transmissions by shifting each one to a different assigned frequency or channel.

The techniques of modulation and demodulation demand very stable oscillators for generating the transmitter carrier signal and the local oscillator signals for demodulation at the receiver.

Several block diagrams of simplified communications systems are shown in Figures 2, 3 and 4.

Figure 2 SIMPLIFIED A.M. TRANSMITTER BLOCK DIAGRAM

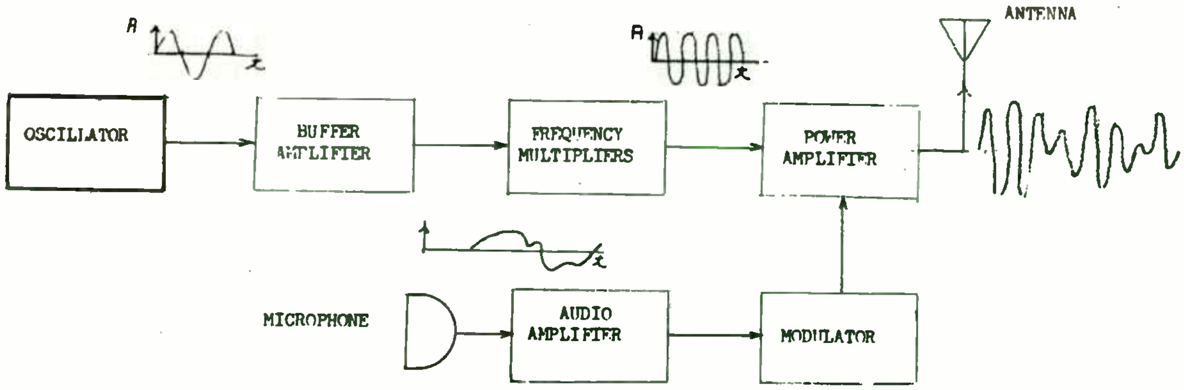
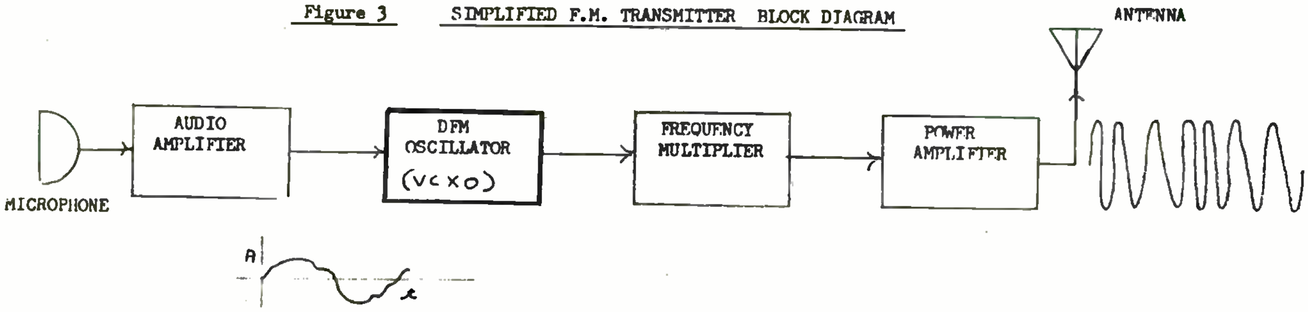
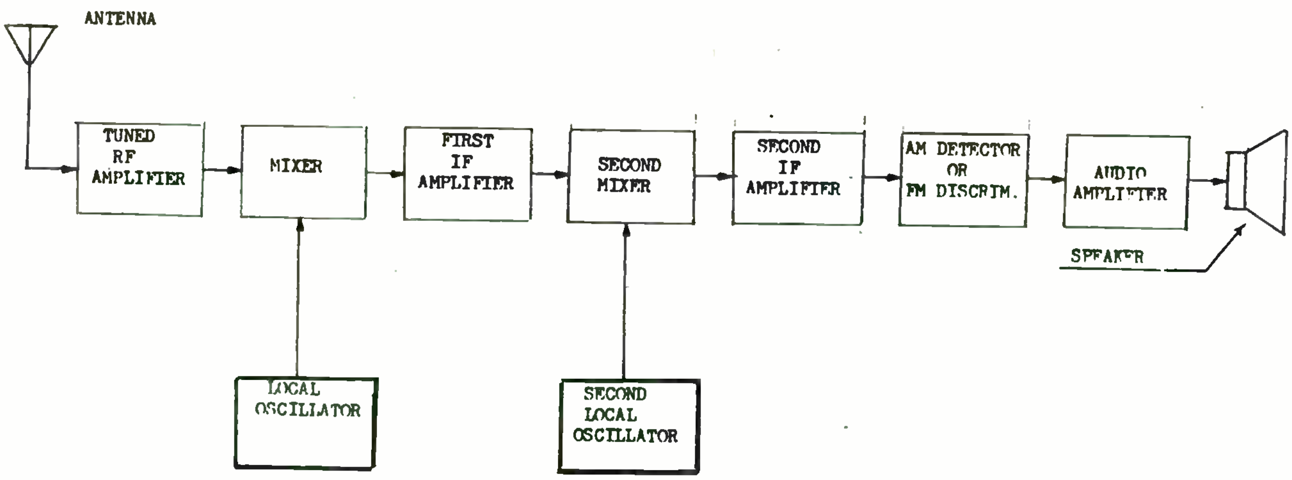


Figure 3 SIMPLIFIED F.M. TRANSMITTER BLOCK DIAGRAM



DUAL CONVERSION

Figure 4 SUPERHETRODYNE RECEIVER BLOCK DIAGRAM



MICROSONICS
INCORPORATED

Today, with an increasing number of users of a limited frequency spectrum, channel spacings are becoming narrower. This is in turn is increasing the need for improved oscillator stabilities to keep transmitted signals within their assigned channels. Also as the number of channels for each system has expanded; the need for seperate stable oscillator pairs, for each channel receiver and transmitter has multipled. This in turn has created a demand for frequency synthesizers to reduce the total number of oscillators per system. While frequency synthesizers provide a practical solution they have created even tougher requirements for oscillators including; tighter stabilities for the reference oscillator, improved phase noise, tighter control of oscillator output waveforms and power, and wide swing voltage controlled oscillators whose output frequency varies linearily with input control voltages.

A simplified block diagram of a frequency synthesizer is shown in Figure 5.

So far we have discussed a variety of Frequency Division Multiplexed system requriements for oscillators. Here the signals are mixed in the time domain but maintain their identity in the frequency domain. Each channel occupies a different frequency band.

There is also an increasing use of Time Division Multiplexed systems where the frequency spectra of the sampled signals occupy the same frequency region and are mixed beyond recognition. However each signal occupies a distinct time interval and the information is received with the use of synchronous switching or gating circuits. In these systems highly stable clock oscillators are required for the synchronous timing circuits.

Figure 6 shows a block diagram for a simplified Time Division Multipled system.

The proliferation of digital computers and digital circuits has also increased the need for stable ocsillators for use in timing and control. It is difficult to envision any digital system without at least one clock oscillator being used as a timing reference for controlling system functions. Most of us see an application of clock oscillators every day since they are also used as the timing reference for many wristwatches and real time clocks, both digital and analog.

Oscillators are commonly used in test equipment as frequency and time reference sources for frequency generators and signal analyzers. Other applications include time interval meters, temperature and pressure gauges, navigation equipment, missile guidance, and electronic warfare systems.

Several applications for oscillators are shown in Figures 7, 8, 9 and 10.

FIGURE 5

FREQUENCY SYNTHESIZER

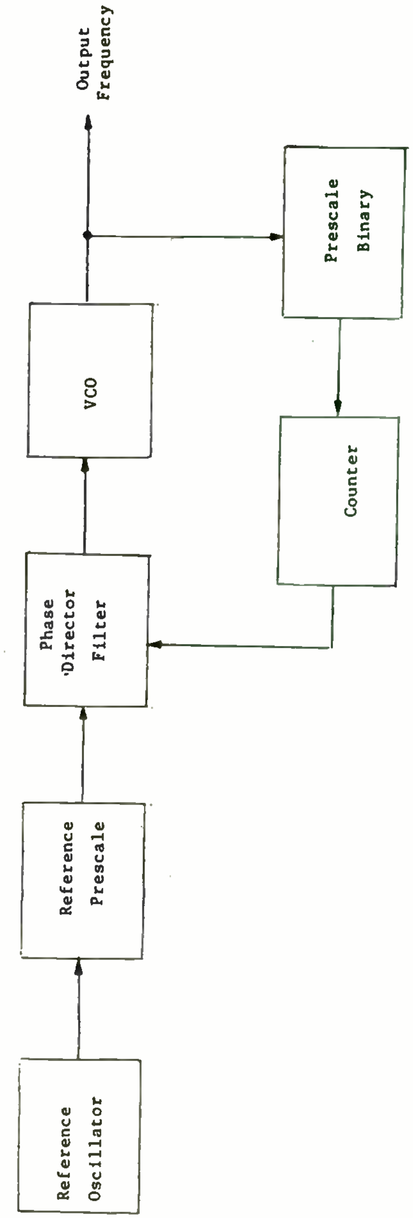
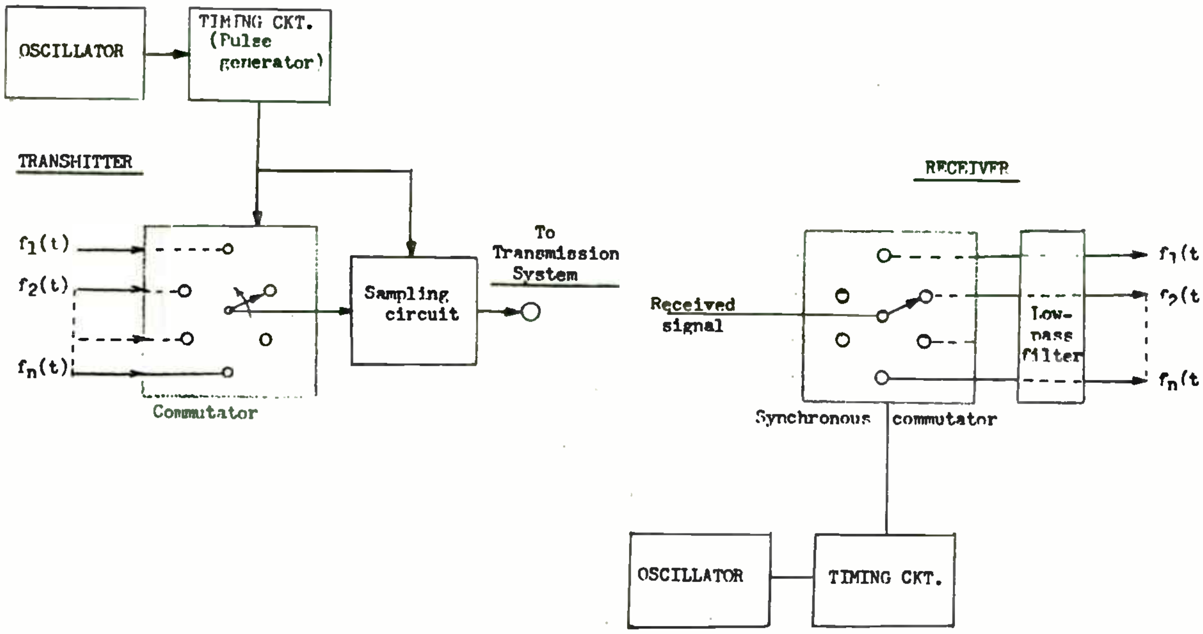


Figure 6

TIME DIVISION MULTIPLEXING



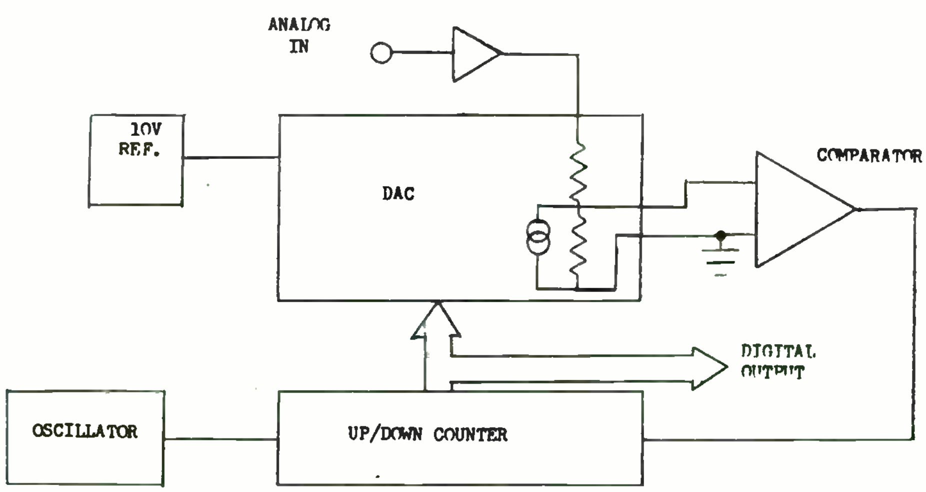


FIGURE 7

TRACKING A/D

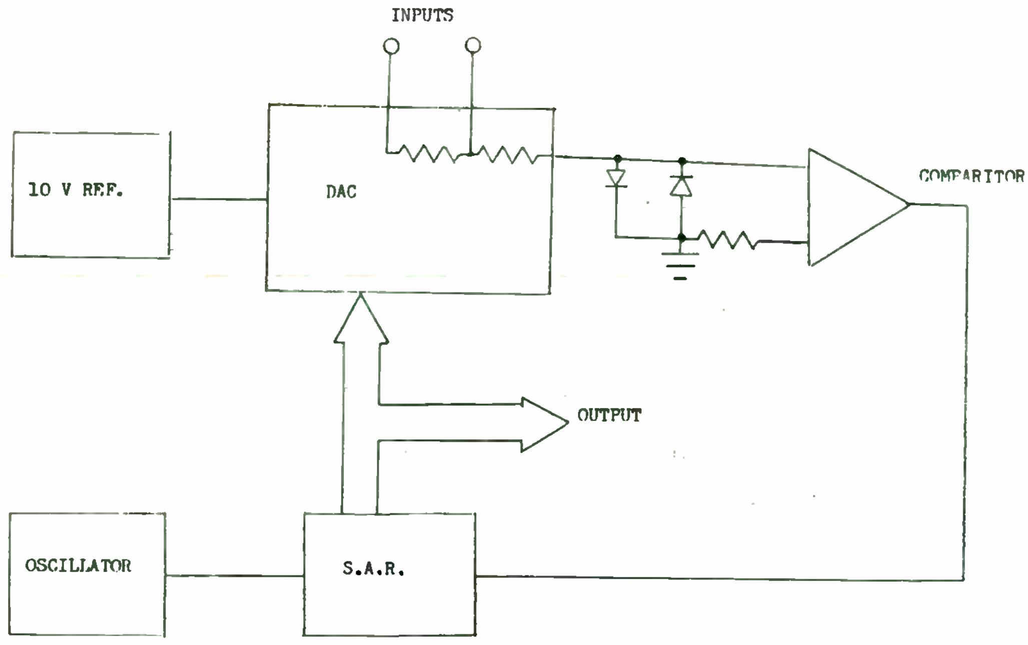


Figure 8

HIGH SPEED 12-BIT ADC

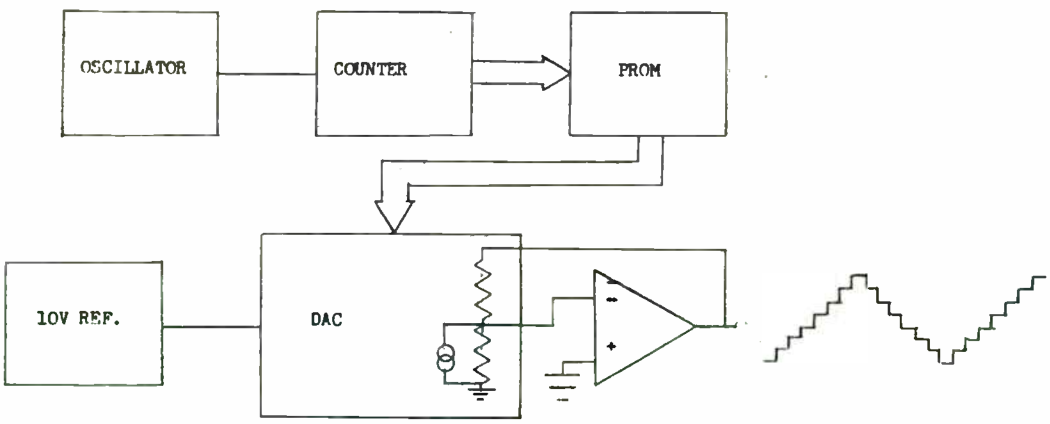


Figure 9

WAVEFORM GENERATOR

MICROSONICS
INCORPORATED

OSCILLATOR - TECHNICAL

Definitions:

1. Webster's New Collegiate Dictionary

   "OS-CIL-LA-TOR

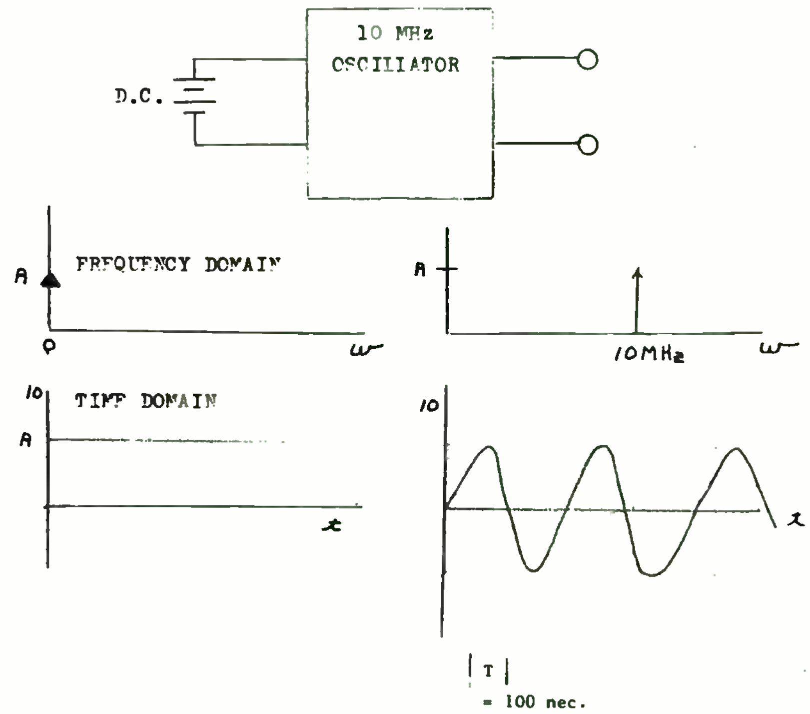
   a) One that oscillates
   b) A device for providing alternating current
      esp: A radio frequency or audio frequency generator."

2. J. Groszkowski   Frequency of Self-Oscillations

   "Self oscillations - electric oscillators in which electrical energy, usually d.c., is converted into electrical oscillations of a desired frequency."

OSCILLATOR CIRCUIT BLOCK

FIGURE 11



MICROSONICS INCORPORATED

OSCILLATOR TECHNOLOGIES

As many experienced R.F. amplifier designers will tell you the possible combinations of circuit elements that maybe used to create an oscillator are endless; even when you do not want one. Several types of oscillator circuits will be discussed in this section including:

- Relaxation oscillators
- Resonant circuit and phase shift oscillators
- A variety of resonantor oscillators
- Laser/Maser oscillators
- Atomic Frequency Standards

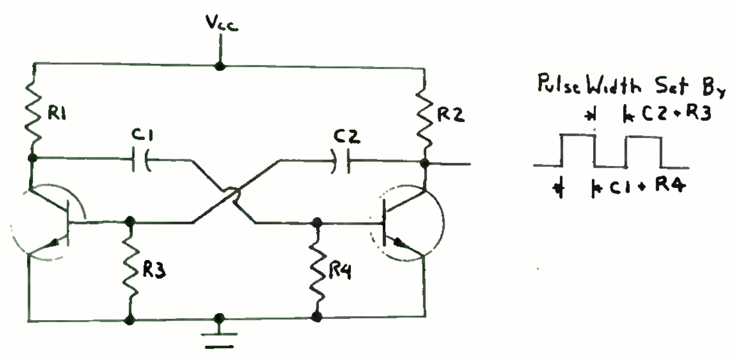
MICROSONICS INCORPORATED

RELAXATION OSCILLATORS

Relaxation oscillators generate a continuous output of pulses at a specific pulse repetition rate (PRR) without any external signal other than a D.C. supply voltage. In relaxation oscillators the output frequency or PRR is determined by capacitor/resistor or inductor/resistor combinations rather than by a conventional inductor-capacitor tuned circuit. Energy builds up in the capacitor until the capacitor is charged to a certain voltage level. The circuit then "relaxes" and the capacitor discharges; hence the term relaxation oscillator.

Figure 18 shows a collector coupled multivibrator circuit. The collector coupled multivibrator is a 2-stage R.C. coupled amplifier with the outputs of each collector being coupled to the opposite transistor base through the capacitors C1 and C2. The output frequency is controlled by the R.C. time constants of R_3, C_2 and R_4, C_1.

FIGURE 18



Collector Coupled Multivibrator

MICROSONICS
INCORPORATED

OSCILLATOR CRITERION

An oscillator is a circuit that provides an output, usually with a specific frequency and waveform, without receiving any external signals; i.e., it is a self-generating circuit. However an external power source, usually d.c. is required.

Figure 12 shows a feedback control block diagram model of an oscillator consisting of an amplifier, a feedback network and a summing junction. From the loop analysis we can see that:

$$S_i + S_f = (AB)\ S_i$$
$$\text{or } AB = 1$$

This is one of the first conditions of self-generation in an oscillator. That is: the loop gain (AB) must be equal to unity or

$$|AB| = 1$$

Also it is easy to see that if S_f were not in phase with Si then S_f would cancel some of Si and the output So would eventually stop. However when Si and S_f are identically in phase Si is increased and the output So will continue. This represents another condition of sustained oscillations that is: The total phase shift as the signal proceeds from the input through the amplifier and feedback network back to the input must be zero or stated another way when n is an integer:

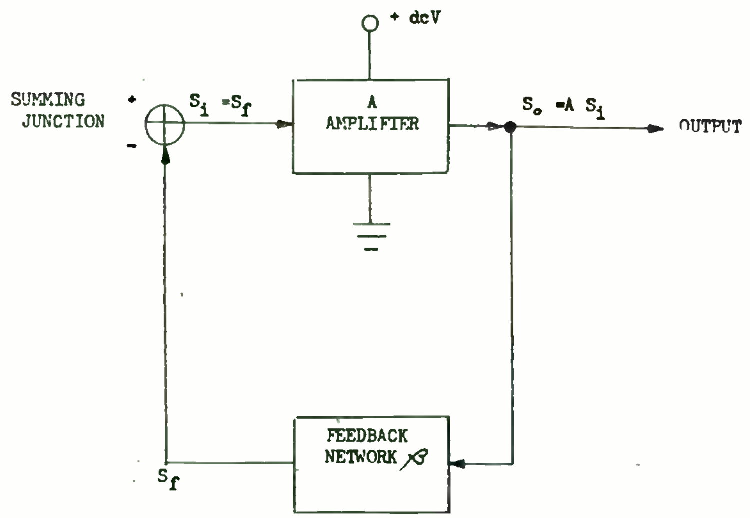
$$O_f = 2\ n \quad = n \text{ X } 360°$$

These two conditions are called the Barkhausen criterion [3]. "Oscillations will be sustained if at the oscillator frequency, the magnitude of the product of the open loop gain transfer function is unity. This implies that $|AB| = 1$ and that the phase OAB = 0. This type of feedback is often referred to as regenative or positive feedback. In practice the feedback network, previously represented as B, usually contains a resonant structure made up of an equivalent inductance and capacitance that resonates at the desired operating frequency.

In practice it is ncessary to have the loop gain $|AB|$ slightly larger than unity since the amplitude is generally limited by gain device nonlinearities (i.e., the amplifier self-limits to provide $|AB| = 12$. Also the feedback networks (B) phase shift usually requires adjustment to compensate for gain device parasitic reactances and circuit "stray" impedances.

Figure 12

OSCILLATOR FEEDBACK CONTROL BLOCK DIAGRAM



S_f = Feedback Signal = $\beta S_o = A\beta S_i$

$S_i = S_f$ so that $S_i = A\beta S_i$

MICROSONICS
INCORPORATED

RESONANT CIRCUIT OSCILLATORS

Figure 13 shows a resonant circuit oscillator using a single NPN transistor and Transformer Feedback. The operating frequency is determined by W^2 1/LC with the transformer feedback being -180°C and the base to collector phase shift being +180°C such tht the total feedback phase difference is zero. Note that the amplifier is based in its active region initially and dynamic self bias results from R_3C_3. This action results in Class C operation and unity gain at the operating frequency.

There are several common types of resonant LC circuit feedback oscillators, two of these are shown, without DC components in Figures 14 and 15. The general operating frequency equations are also shown.

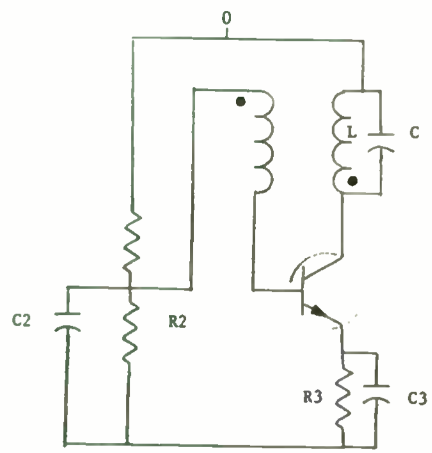
The Hartley oscillator is shown of Figure 14. It is similar to the transformer coupled oscillator except that, instead of two seperate coils, it uses a single coil that is tapped to provide the signal feedback by means of the mutual inductive coupling between the coils L1 and L2 formed by the tap.

The Colpitts oscillator of Figure 15 is similar to the Hartley oscillator except that the pair of capacitors C1 and C2 are used in place of the tapped inductor. Feedback from the collector to the base now occurs capacitively by means of the voltage divider effect of C1 and C2.

Notice that the resonant circuits provide an 180 degree phase shift that is opposite to that of the base to collector to provide a net phase shift of 0 degrees at the operating frequency.

FIGURE 13

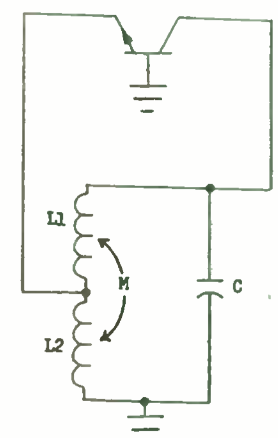
TRANSFORMER COUPLED FEEDBACK OSCILLATOR



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

Figure 14

HARTLEY OSCILLATOR

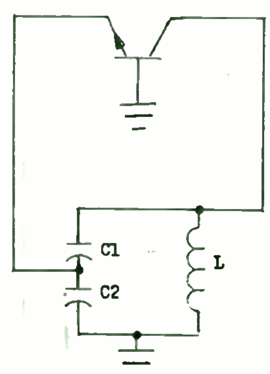


$$\omega_o^2 = \frac{1}{C(L1 + L2 + 2M) - (L1 \cdot L2 - M^2)\frac{h_{ob}}{h_{ib}}}$$

MICROSONICS INCORPORATED

Figure 15

COLPITTS OSCILLATOR

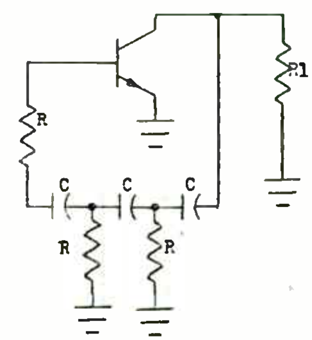


$$\omega_o^2 = \frac{C1 + C2}{LC1 \cdot C2} + \frac{h_{ob}}{C1 \cdot C2 \cdot h_{ib}}$$

A phase shift oscillator is shown in Figure 16. In this oscillator the RC networks provide the necessary feedback phase shift to sustain oscillators.

FIGURE 16

PHASE SHIFT OSCILLATOR



$$\omega_o^2 = \frac{1}{C^2(4R \cdot R1 + 6R^2)}$$

$$h_{fe} = 23 + \frac{29R}{R1} + \frac{4R1}{R}$$

MICROSONICS
INCORPORATED

RESONATOR OSCILLATORS

It is easy to see that resonant LC circuits can be replaced with other electronic components that have equivalent circuits. Examples of such devices are discussed in the succeeding paragraphs annd they include:

- Mechanical — Or electromechanical resonators
- Crystals — Quartz or ceramic
- S.A.W. Devices — Surface Acoustic Wave Devices
- Microwave Resonators
  - Transmission lines
  - Metallic cavity
  - Dielectric resonators

OSCILLATOR TECHNOLOGIES

Oscillator technologies are often classified by the resonant structure used in the feedback network. These include:

Mechanical Resonators

One type of mechanical resonator uses a metal tuning fork with its resonant frequency dependent upon the length width, thickness and material used for the tuning fork tines. Energy is coupled in and out of the device with Piezo Electric ceramic transducers that convert electrical signals to a mechanical vibration at the input and the tuning fork mechanical vibration back to an electrical signal at the output. These devices are limited by their mechanical properties to frequencies to approximately 10kHz. A mechanical tuning fork example is shown in Figure 19.

Crystal Resonators

Crystals are piezoelectric devices which mechanically vibrate when excited with an alternating voltage. The device resonant frequencies and Q are dependent upon the crystal dimensions, how the crystal is orientated with respect with its axes, how the electrodes are applied, and how it is mounted in its holder.

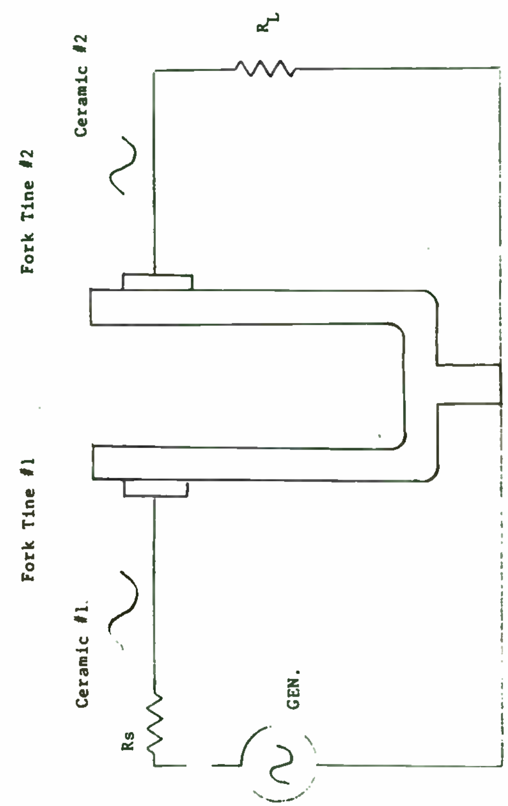
An equivalent circuit crystal is shown in Figure 17. Crystal oscillators may be designed using piezoelectric crystals made from special ceramics or quartz.

Piezoelectric crystal oscillators are inherently very stable due to the high resonator circuit Q's and their excellent time and temperature stability make them one of the better choices when tight stability performance is required

Quartz crystals may be used for frequencies from 10kHz to over 250MHz with Q's of 10000 to over 1 million. Ceramic resonators may be used for frequencies from 100kHz over 1MHz with circuit Q's of 3000.

FIGURE 19

MECHANICAL TUNING FORK RESONATOR



MICROSONICS
INCORPORATED

Figure 17

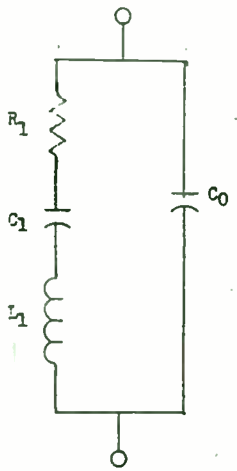
CRYSTAL QUARTZ RESONATOR CIRCUIT
and
REACTANCE CURVES

R_1 = Motional (Quartz) Resistnace (Loss)

C_1 = Motional (Quartz) Capacitance

L_1 = Motional (Quartz) Inductance

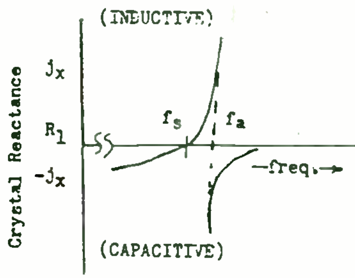
C_0 = Holder (Electrode) Capacitance



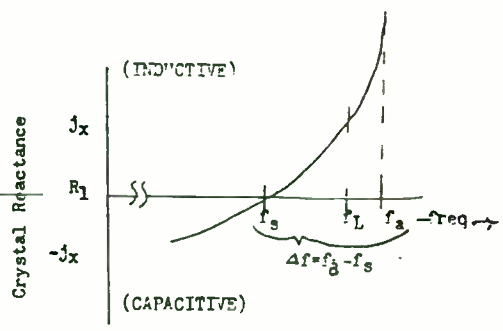
$$F_s = \frac{1}{2\pi\sqrt{L_1 C_1}}$$

$$F_a = F_s\left[1+\frac{C_1}{C_0}\right]$$

EQUIVALENT CIRCUIT



REACTANCE Vs FREQUENCY



EXPANDED REACTANCE Vs FREQUENCY

MICROSONICS
INCORPORATED

Surface Acoustic Wave (SAW) Devices

While crystal oscillators use the bulk properties of piezoelectric quartz and ceramic crystals, surface acoustic wave oscillators utilize the surface acoustic wave propogation properties of piezoelectric materials (quartz or lithium niobate) to convert electromagnetic wave energy into surface waves with lower wave velocities which allows small rugged, high frequency devices to be formed which are relatively insensitive to mounting procedures. SAW devices usually require at least 2 sets of electrode patterms of carefully controlled interdigitated fingers to convert the electronic signal to a surface wave and back again. A block diagram is shown in Figure 21.

SAW oscillators provide size and cost advantage about 100MHz to quartz crytal oscillators with frequency multipliers while providing improved frequency stabilities over LC circuit oscillators that operate in that frequency range.

Microwave Resonators

A wide variety of microwave oscillators are currently being used. Some of these include the use of conventional circuit topologies with special high frequency semiconductor devices and unique resonant feedback network structures such as:

- Transmission line or microstrip resonant structures as shown in Figure 19. Frequencies range from 300MHz to over 2GHz.

- Metal cavity oscillators as shown in Figure 20. The cavities correspond to tuned circuits with relativerly high Q's. Frequencies range from 300MHz to over 6GHz.

- Dielectric resonator oscillators use resonant structures that are made from dielectric materials of the barium titanate variety with alumina and fused quartz. These devices have been used to frequencies over 16GHz. An example is shown in Figure 41.

Other unique microwave device oscillators include:

- Magnetron oscillators
- Klystron oscillators
- Traveling wave oscillators
- Masers

FIGURE 21
S.A.W. DELAY LINE
OSCILLATOR

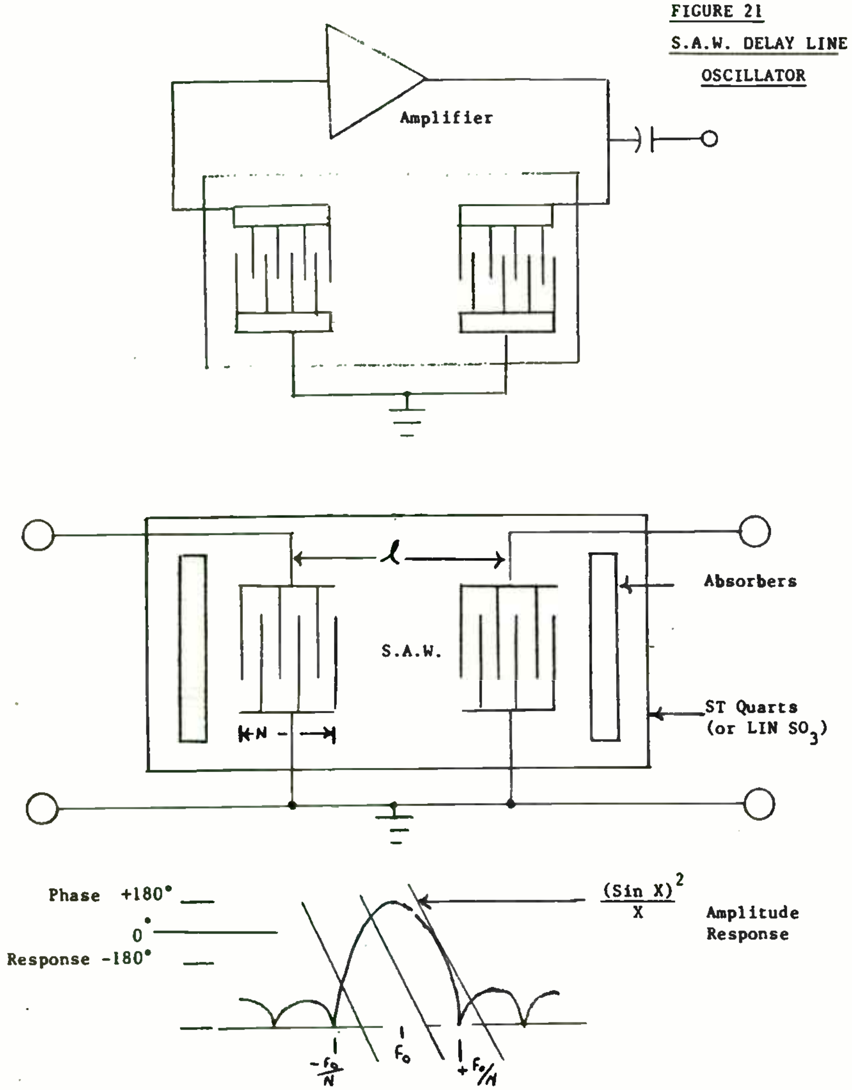


Figure 19
MICROSTRIP OSCILLATOR

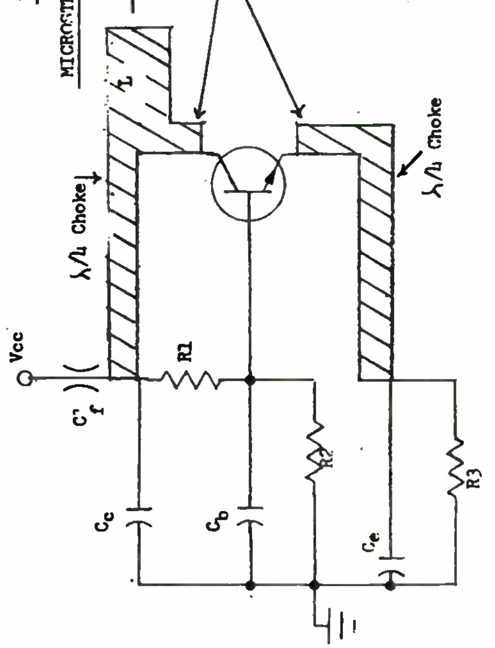
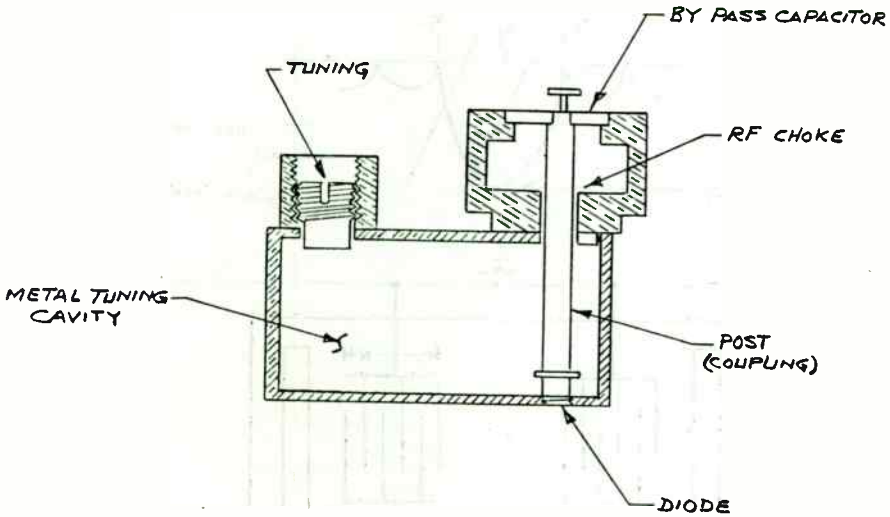


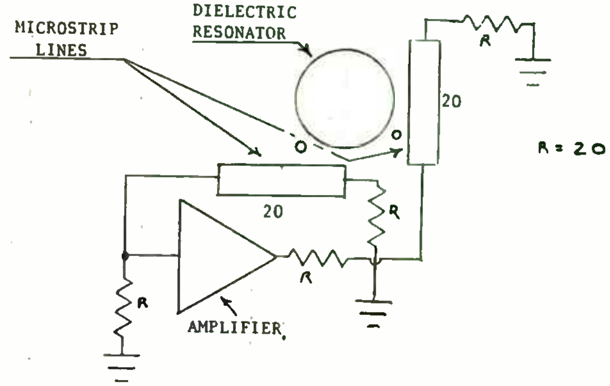
FIGURE 20



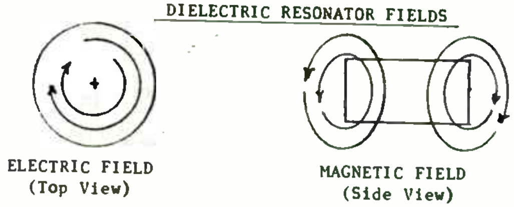
GUNN DIODE / CAVITY OSCILLATOR

Figure 41

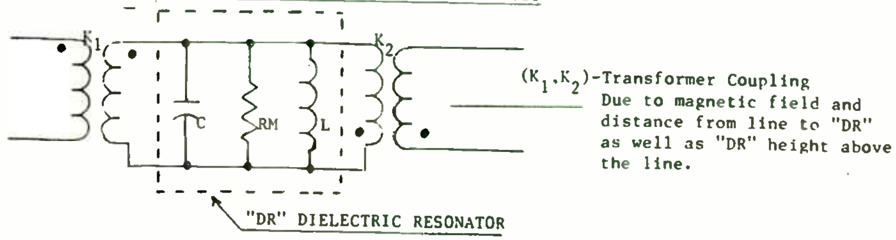
DIELECTRIC RESONATOR OSCILLATOR



DIELECTRIC RESONATOR FIELDS



DIELECTRIC RESONATOR EQUIVALANT CIRCUIT



MICROSONICS
INCORPORATED

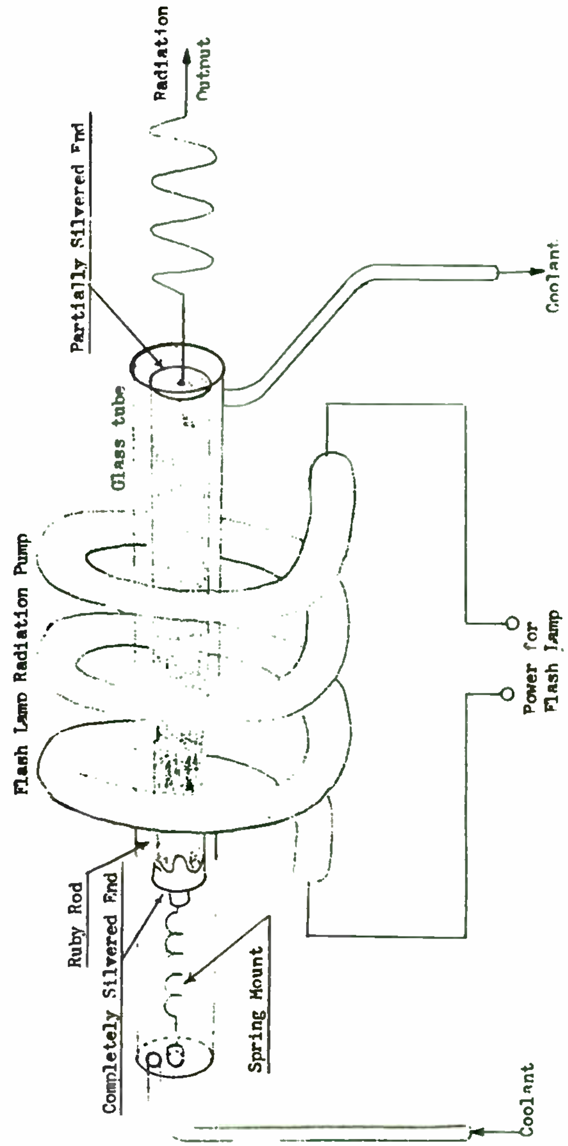
LASER/MASER - LIGHT/MICROWAVE GENERATING OSCILLATOR

Figure 22 shows a simplified construction of a laser system. The term Laser is an abbrevation for Light Amplification by Stimulated Emission of Radiation. The laser emitted radiation is at light frequencies usually the infrared or visible red spectrum and the radiation is at microwave frequences for Masers. The stimulation is produced by applying an input from a radiation pump. Here the Laser is made from a ruby crystal for infrared output that utilizes a green light for the radiation pump to produce an output with a frequency of 300 X 10^{-12}Hz or a wavelength of 1 X 10^{-4}cm.

The basis of the Laser/Maser action is the fact that the emitted radiation is a result of the electrons in the Laser/Maser atoms changing energy levels by first raising to a higher unstable level and then dropping back to the lower level to release the electromagnetic radiation. The output frequency or wavelength is determined by the difference in the two energy levels for the electrons.

In the Laser of Figure 22 the radiation pump is the flash lamp which is similar to a neon bulb. This radiation supplies energy to the solid ruby rod to stimulate its emission. The output is radiated through the partially silvered end ot the right end of the ruby crystal.

The other end is completely silvered so that radiation can build up its intensity by repeated reflection between the ends. For a Maser the output may be taken through a waveguide and the ruby rod maybe replaced with a gas mixture of helium and neon.



CONSTRUCTION OF A RUBY LASER

Figure 22

STABILITY

Several factors must be carefully considered when specifying or selecting an oscillator function. Two of the most important factors are: the desired operating frequency and the overall frequency stability required. We have already discussed examples for determining operating frequency but not frequency stability.

Generally the most stable oscillators are those whose resonant feedback networks have the highest Q making them less sensitive to variations of of other circuit components. Another way to compare oscillator circuit stabilities is to determine the phase slope of the frequency determining networks. The circuit with the smallest df/dO characteristic will be the most stable. This is one reason that quartz crystals with Q's over 100,000 are much more stable than LC networks with Q's of 10 to 50. Figure 34 shows "Bode" plots of gain ahd phase characteristics for a LCR network with different Q's. Notice that the frequency vs. phase slope decreases as the circuit Q increases.

The frequency stability (frequency error) of an oscillator is a measure of its ability to maintain as nearly a fixed frequency over as long a time interval as possible while being exposed to system environmental conditions. Usually when frequency stability is considered it is one of three (3) types: a) environmental - temperature, supply and load variations and mechanical shock and vibraton; b) long-term frequency drift referred to as aging; c) short-term frequency stability or phase stability.

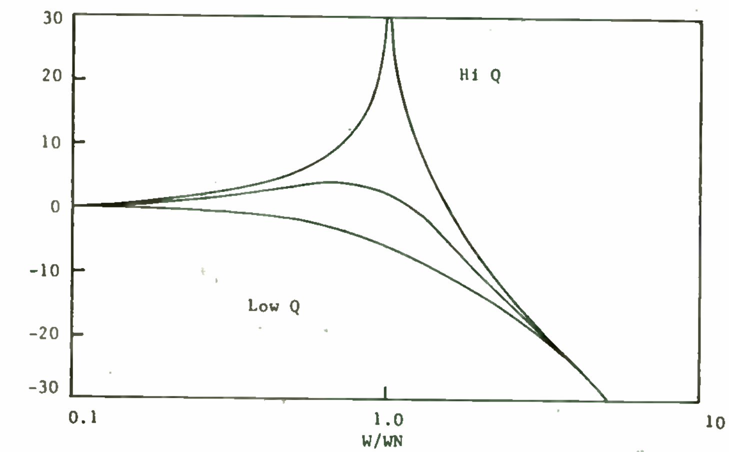
Often oscillator circuits are designed with temperature compensation networks that cancel the temperature drift characteristics of those circuit components to which the oscillator frequency is most sensitive. Another approach is to encase the oscillator in an oven that is controlled to maintain the oscillator at a constant temperature regardless of the outside ambient temperature.

Long term aging drifts are usually dependent upon the mechancical, thermal and chemical properties of the frequency determining circuit elements used. Also components manufacturing process and process controls can play a critical role in their aging characteristics. As an example high temperature burn-in and temperature cycling can stress relieve and stabilize inductors to provide an order of magnitude improvement in their long-term aging characteristics and those effects upon oscillator drift.

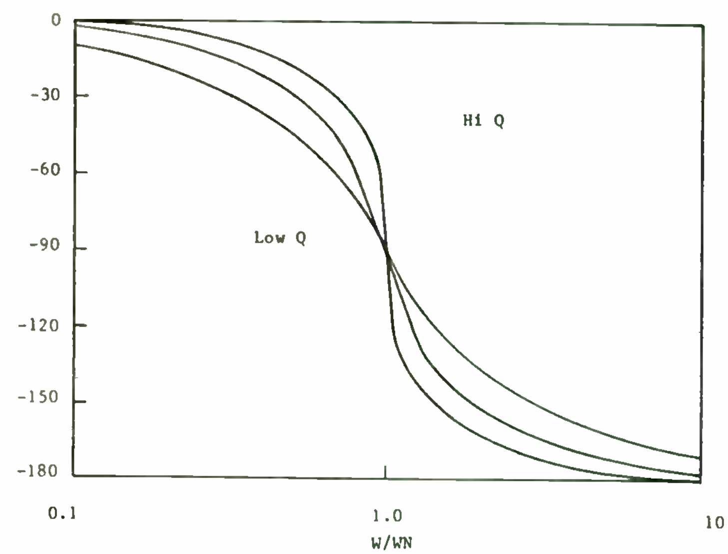
As shown in Figure 5 frequency synthesizer circuit techniques may be employed to frequency lock low stability sources to tight stability reference oscillators such as quartz crystal or atomic standards to provide improved system frequency stabilities. These approaches are especially useful for multiple frequency and microwave applications.

Oscillator stability will be discussed further in the Crystal Oscillator Section.

FIGURE 34



BODE AMPLITUDE PLOT



BODE PHASE PLOT

MICROSONICS INCORPORATED

ATOMIC STANDARDS

As the demand has increased for tight stability oscillators so has the demand for ultra stable standards to which these oscilllators can be compared and calibrated against. Generally the frequency standards stability must be at least one (1) order of magnitude better than the oscillator that is being calibrated.

Standards are one of two types: primary or secondary. The distinction between them is that the standard does not require any other reference for calibaration; whereas the secondary standard requires calibrations both during manufacturing and at intervals in the field.

The three most common types of frequency standards include:

- Quartz crystal (usually ovenized) - Secondary Standard
- Rubidium gas cell controlled oscillator - Secondary Standard
- Cesium atomic beam controlled oscillator - Primary Standard

The Cesium beam standards are used where ultra high accuracy is required. In fact the NBS frequency standard is the Cesium beam type. The Cesium beam standard is a quantum electronic device that provides access to one of nature's invariant frequencies, that of the Cesium atom. The resonance frequency will not drift or age and thdrefore can besguaranteed to provide of specified accuracy.

Ruhidium standards are similar to Cesium beam standards in that an atomic resonant element prevents drift of a quartz oscillator through a frequency lock loop. Yet the rubidium gas cell is dependent upon the gas mixture and presssure within the cell. Therefore it must be calibrated both during manufacturing and in the field.

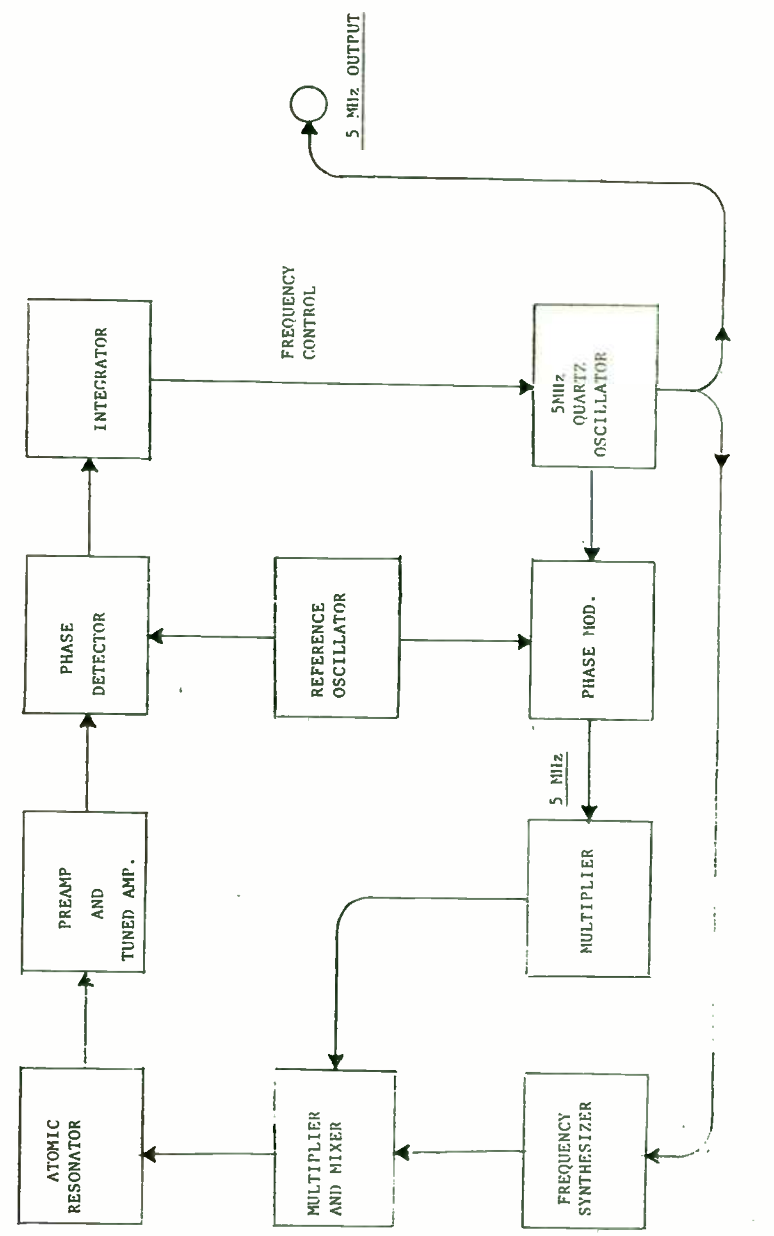
Quartz crystal oscillators are used in virtually all frequency standards. However quartz crystal oscillators exhibit an aging rate (frequeny drift over time) that must be periodically readjusted in the field. In atomic standards the oscillator drift is constantly being corrected by the rubidium or cesium beam devices through the use of an electronic frequency lock system as shown in Figure 30. The crystal oscillator is multiplied and synthesized to the atomic resonance frequency (6834 +MHz for rubidium and 9192 +MHz for cesium). The signal is frequency modulated to sweep through the atomic resonance frequency causing the beam intensity in cesium tube or tranmitted light through the rubidium cell to vary. The output signal is amplified and through a phase detector controls the frequency of a low noise crystal oscillator. Frequency dividers may also be applied to produce other outputs at lower frequencies. The table below compares the relative long term drift of the three standards:

	Drift per year
Quartz crystal (ovenized)	1 X 10^{-7}
Rubidium	1 X 10^{-10}
Cesium	3 X 10^{-12}

FIGURE 30

ATOMIC FREQUENCY STANDARD

BLOCK DIAGRAM



MICROSONICS
INCORPORATED

OSCILLATOR TECHNOLOGY COMPARISONS

Oscillator Techonologies vs. Frequency Range

The chart in Figure 23 outlines the discussed oscillator technologies as a function of the frequency range for which they may be used. It should be noted that several oscillator techniques may be employed to provide the same output frequency. This highlights the question as to which oscillator technology is best for your application. This question might be better answered by evaluating your system stability, size, and input power requirements and using the charts in Figures 23 and 24.

Oscillator Techonogies vs. Stability (temperature and aging) are charted in Figure 24. It is easy to notice that as tighter stabilities are required the number of oscillator techonology options are significantly reduced. When stabilities better than 1 X 10^{-5} are desired the choices must include the use of quartz crystal controlled oscillators either directly, as a referenced oscillator or phase locked to an atomic standard depending upon the actual system stability that is required.

FIGURE 23

OSCILLATOR TECHNOLOGIES Vs. FREQUENCY RANGE

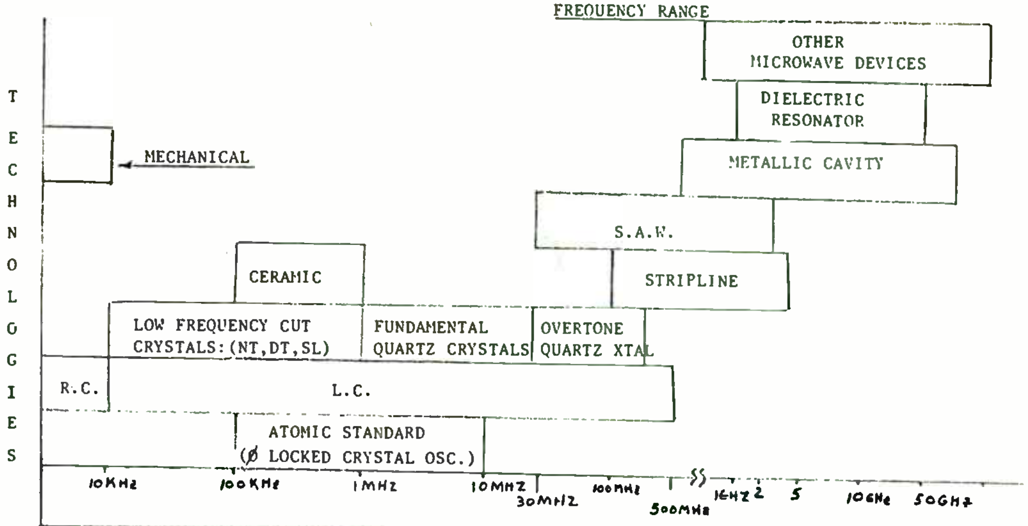
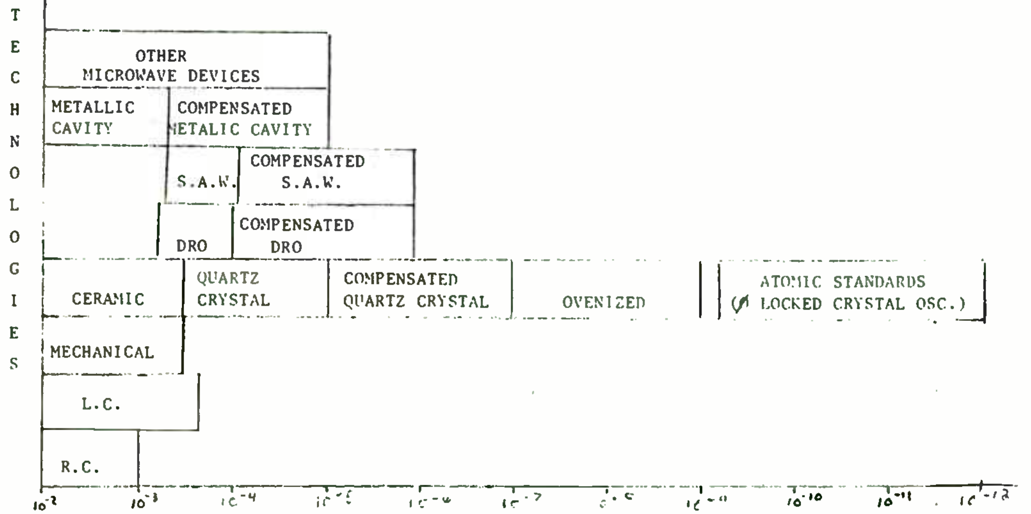


FIGURE 24

OSCILLATOR TECHNOLOGIES Vs. STABILITY



MICROSONICS
INCORPORATED

QUARTZ CRYSTAL OSCILLATORS - THE LOGICAL CHOICE FOR MOST REASONS

The advantages of extremely high Q, small size and excellent temperature stability have made quartz crystals the natural selection for the majority of R.F. oscillator applications over the past 50 years.

As stated earlier quartz crystals are piezoelectric devices that vibrate as an alternating voltage is applied to the electrodes. If the frequency of this voltage is very close to the natural frequency of the quartz crystal the amplitude of vibration will become very large. The strain of these vibrations causes the quartz to produce a sinusoidal field which controls the crystal equivalent circuit impedance. The resonant frequency and Q are dependent upon the crystal dimensions, how the surfaces are oriented with respect to its axes and how the electrodes are applied. An example of how a specific type of quartz crystal (AT) is cut from a quartz bar is shown in Figure 25. It is the orientation of the quartz crystal with respect to the atomic lattice Z axis that also governs the resulting temperature characteristic.

A diagram of a typical quartz crystal is shown in Figure 26. Figure 27 shows the circuit symbol, an equivalent circuit schematic and impedance diagrams for a quartz crystal. Figure 28 shows frequency vs. temperature characteristics of typical AT-cut crystals which are most common for R.F. oscillator applications. Figures 35 and 36 show Frequency vs. Temperature characteristics for other crystal cuts, the ST and SC cut.

AT-cut crystals may be operated on their fundamental mode as discussed earlier as well as on mechanical overtone mdoes which are odd multiples (3rd, 5th, 7th, etc.) of the fundamental mode. Since these modes have higher Q values they are more difficult to tune to exact frequency with external circuit reactances. The circuits shown previously can be made to operate with a higher frequency crystal mode by adding a tuned circuit to "trap" out the undesired lower frequency modes. An inductor may also be added to tune out controls of the crystal hold to increase the circuit tuning.

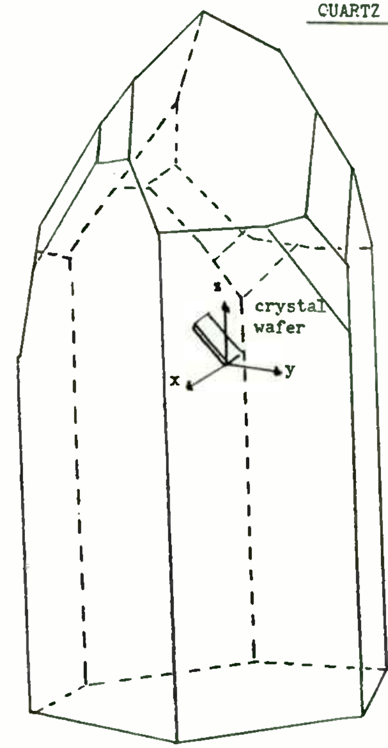
There are a number of crystal oscillator configurations. Three of these are shown in Figure 29. The gate oscillator is used primarily in low stability logic applications. The modified Colpitts circuit is often used because of its simplicity and the relative isolation between the base and collector circuits. The Pierce circuit is used for high frequency applications using quartz crystals operated on their overtone modes.

The circuit in Figure 37 is a Colpitts oscillator that will operate on the quartz crytal 3rd overtone mode due to the addition of L1 and C1 that is tuned to "trap" out the fundamental frequency mode.

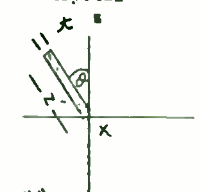
Block diagrams for typical TCXO and ovenized voltage controlled oscillator circuits are shown in Figures 38 and 40.

Figure 25

QUARTZ CRYSTAL BAR



AT Cut Crystal



"Y" Cut Plate Viewed Along "X" Axis

$\measuredangle\theta$ = Angle Reference to "Z" Axis That Governs AT Cut Temp. Charactor

Thickness t Determines the Operating Frequency $fr \propto \frac{1}{t}$

QUARTZ CRYSTAL ADVANTAGES:

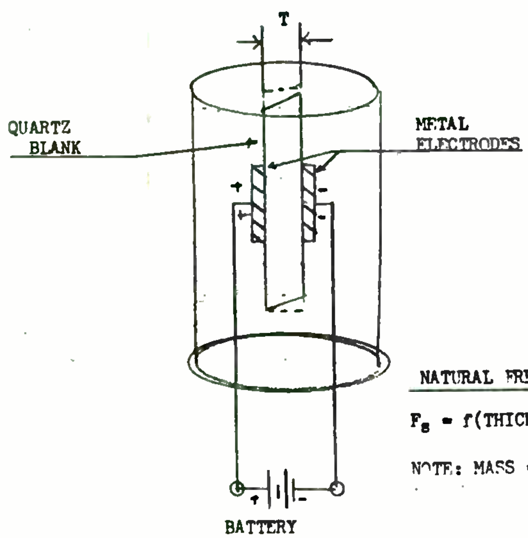
Very High Q = Low Loss = Requires large circuit impedance change to create small frequency error.

Small Size .530"H x .402"W x .185"T

Good Temperature Stability = $\Delta f/f = \pm 10 \times 10^{-6}/25°C \pm 55°C$ (UNCOMPENSATED)

Figure 26

AT
THICKNESS SHEAR
MODE
QUARTZ CRYSTAL



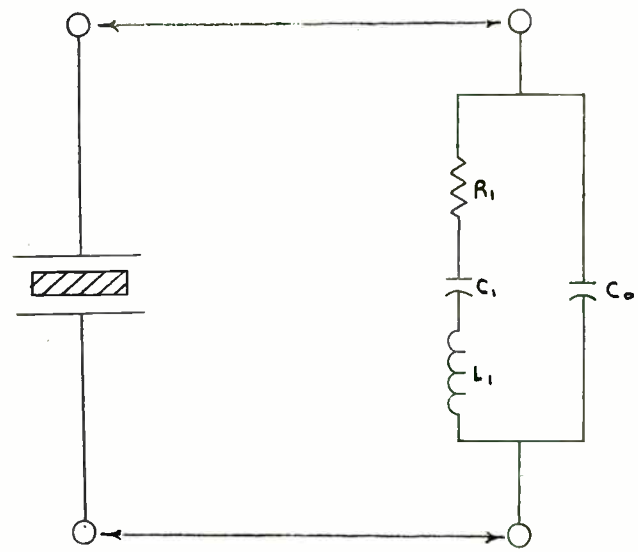
NATURAL FREQUENCY

F_s = f(THICKNESS, MASS*)

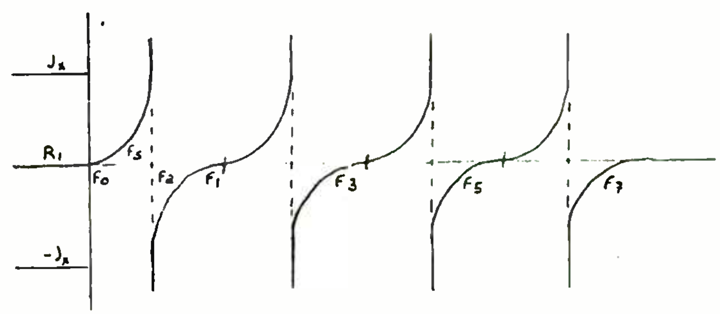
NOTE: MASS = Electrode Mass

FIGURE 27

CIRCUIT EQUIVALENT



OVERTONE MODE IMPEDANCE DIAGRAMS



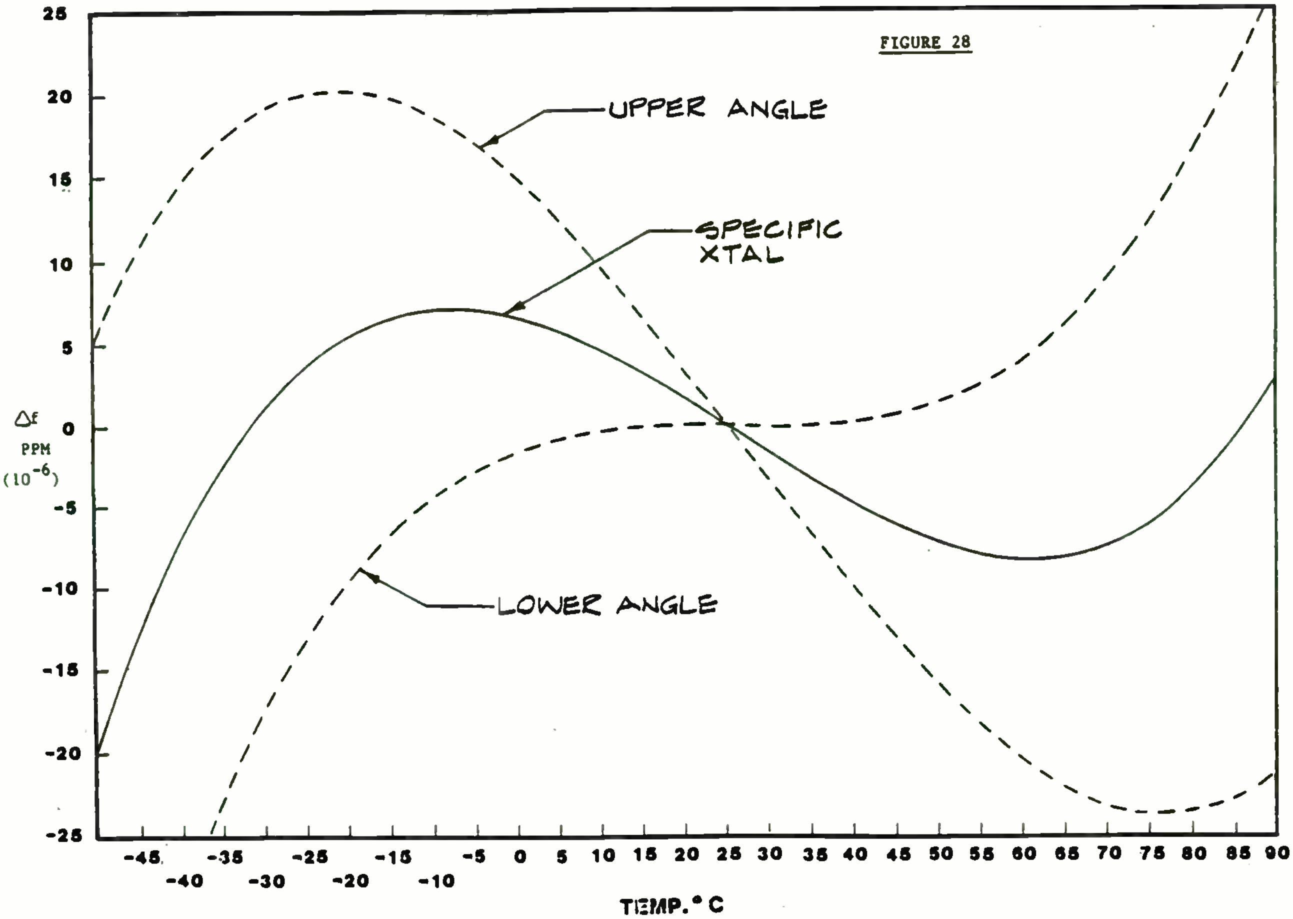


FIGURE 28

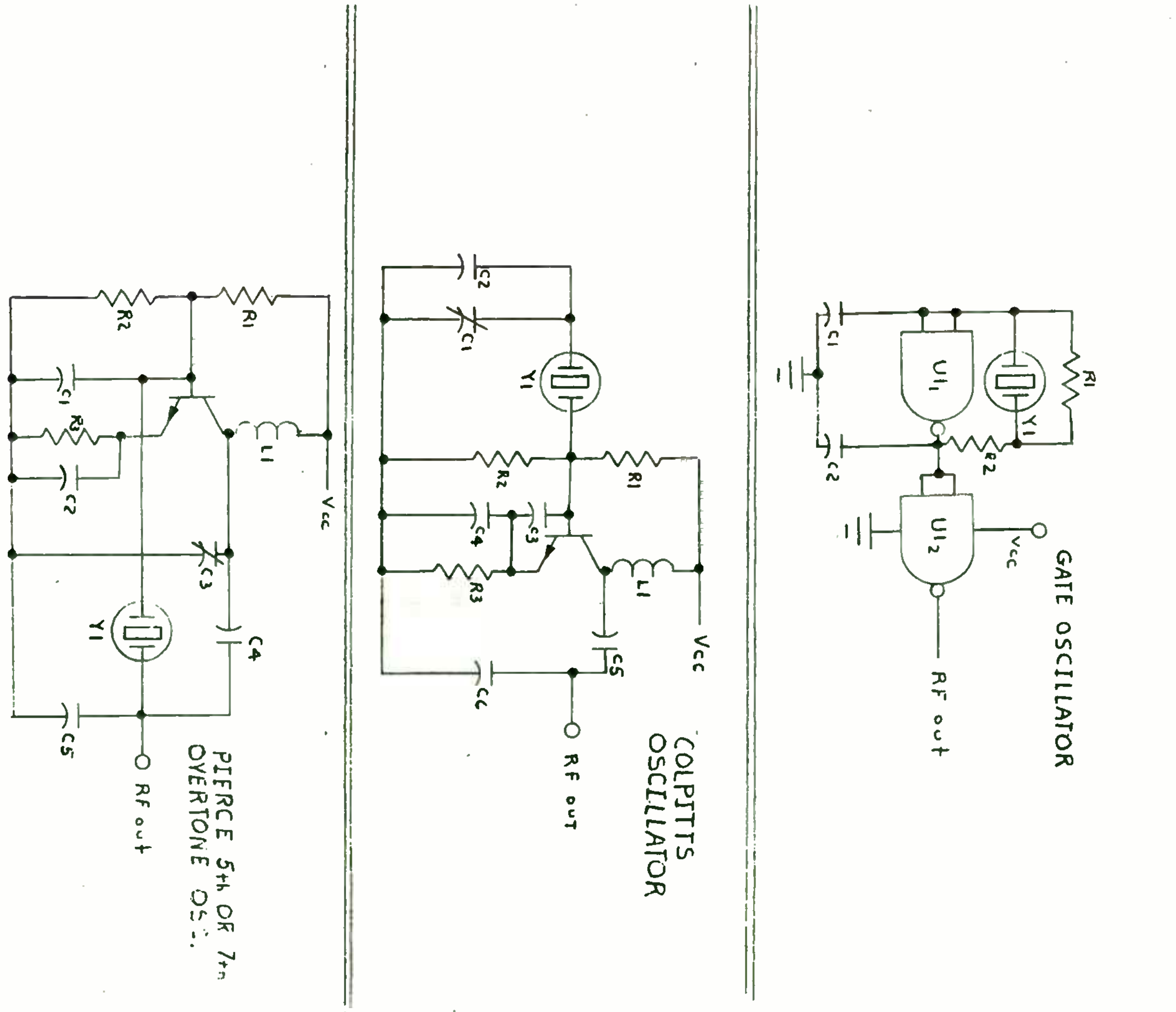


FIGURE 29

FIGURE 35

TYPICAL ST CUT TEMPERATURE/FREQUENCY CHARACTERISTICS

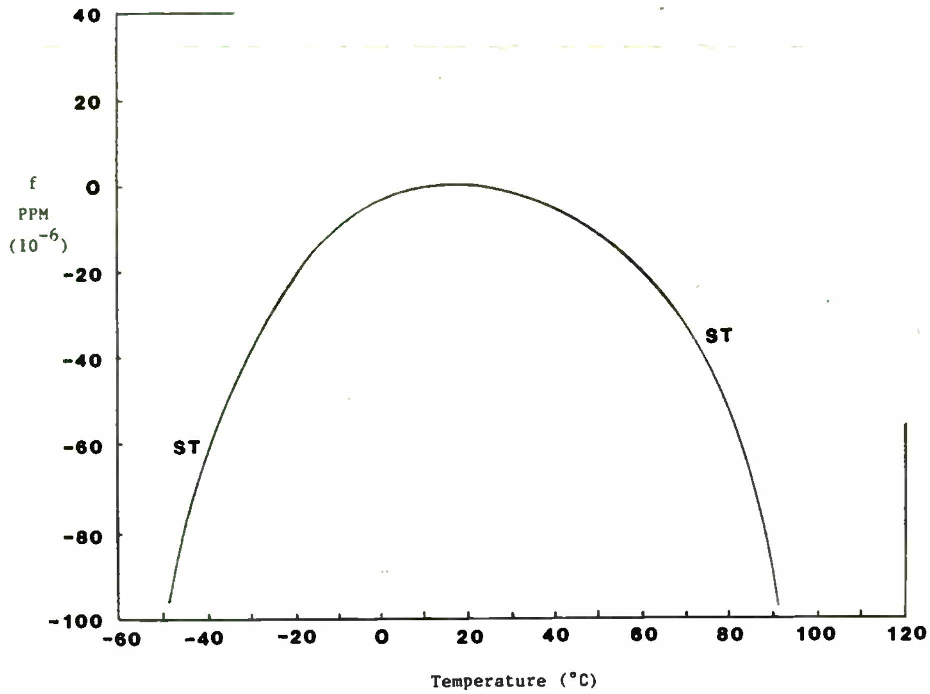
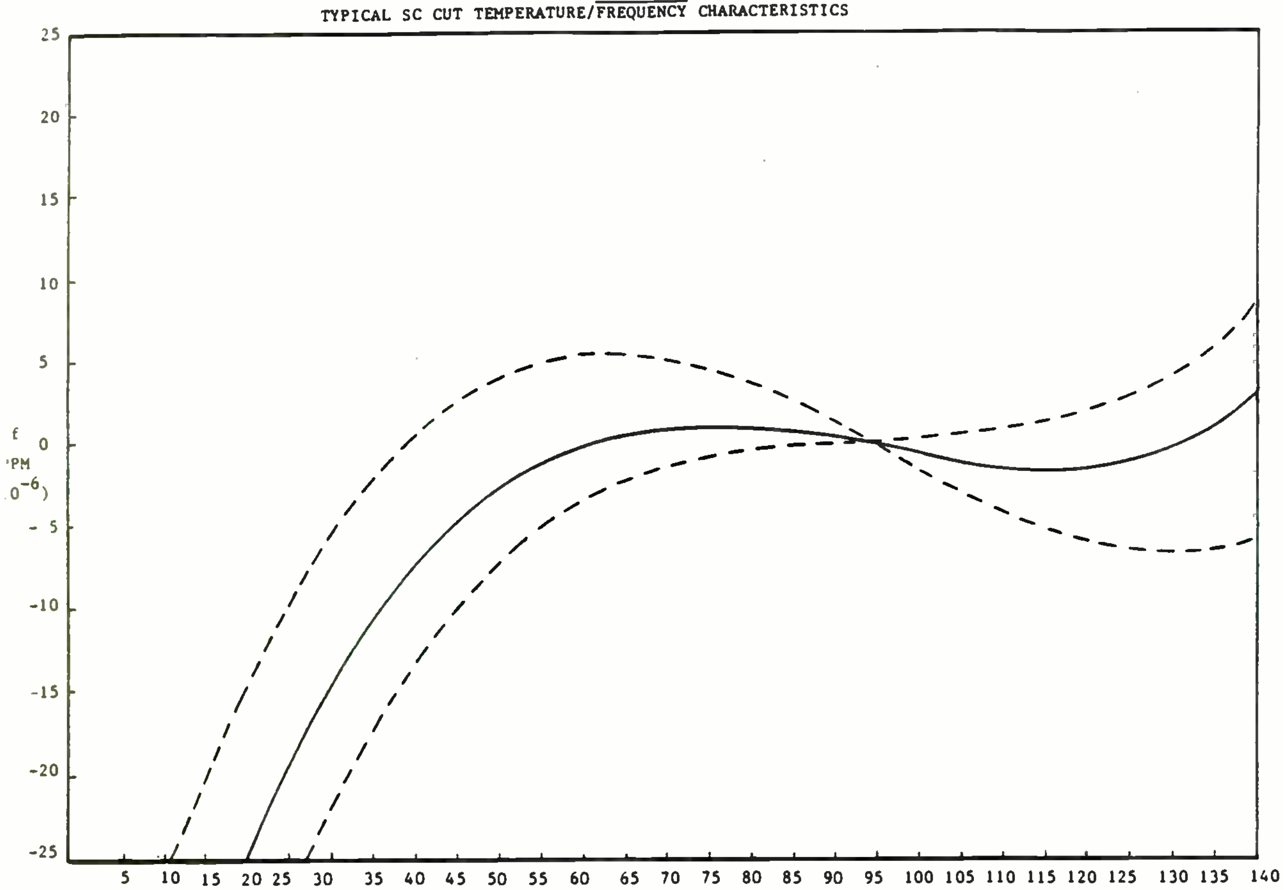
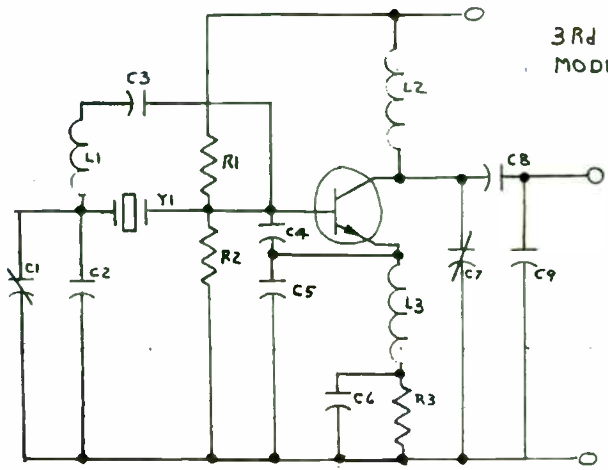


FIGURE 36

TYPICAL SC CUT TEMPERATURE/FREQUENCY CHARACTERISTICS



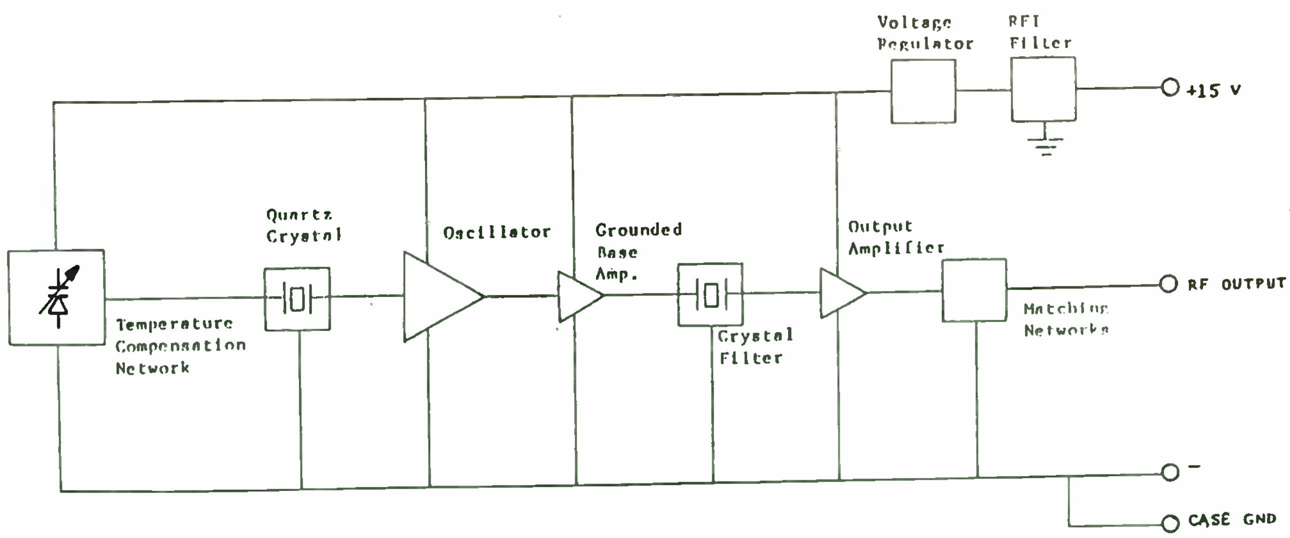


3Rd OVERTONE
MODIFIED COLPITTS.

REVISIONS

LTR	DESCRIPTION	DATE	APPROVED

Figure 38: TCXO CIRCUIT
BLOCK DIAGRAM



UNLESS OTHERWISE SPECIFIED DIMENSIONS ARE IN INCHES TOLERANCES ON
FRACTIONS ± 1/64 DECIMALS ±.01 ±.005 ANGLES ± ½°
ALL SURFACES √

REMOVE ALL SHARP EDGES AND BURRS 1/64 MAX

CONTR	
DRAWN	
CHK	
PROJ ENG.	
MFG. ENG.	
QUAL MGR.	

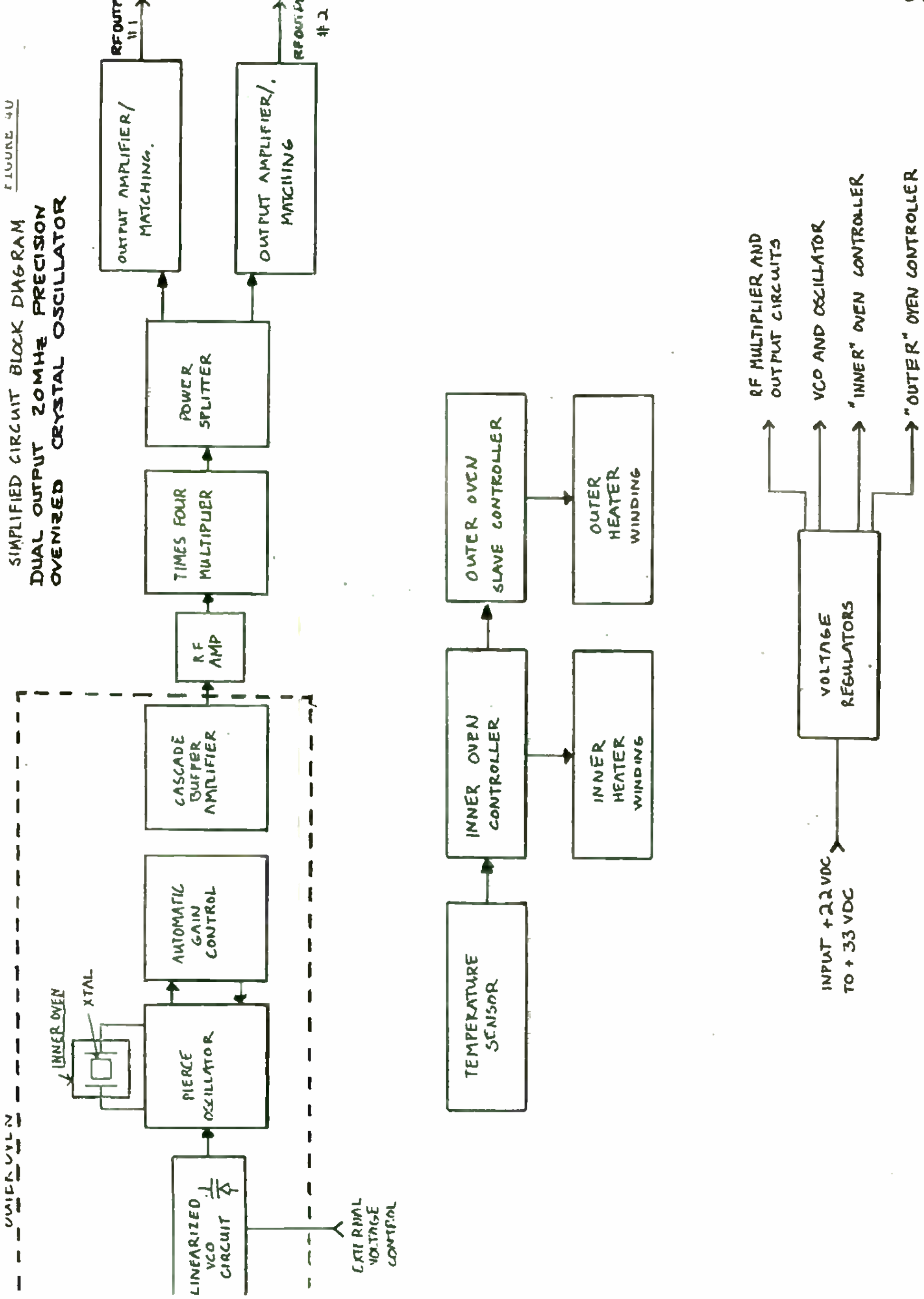
NEXT ASSY | USED ON

APPLICATION

MICROSONICS
INCORPORATED

SIZE	CODE IDENT NO
A	11869

SCALE | SHEET



MICROSONICS INCORPORATED

Temperature compensation of quartz crystal oscillators can be accomplish by the addition of temperature compensation networks. These provide reactance that is in series with the quartz crystal which ideally varies t oscillator frequency in a manner that is equal in magnitude but opposite sign to that of the crystal so that the net frequency changes is zero ov the operating temperature range. This is shown in Figure 31. While variety of compensation network techniques are being used today, most of th use a varactor (voltage variable capacitor) to which a temperature compe sating voltage is applied. As the DC temperature compensating volta varies over temperature the varactor equivalent capacitance is varied effect a frequency change that cancels that of the quartz crystal by itsel Depending upon the frequency vs. temperature stability desired each tempe ature compensating network must be individually adjusted to provide t proper voltage change for each crystal oscillator since no two crystal oscil ators are identical. Crystal oscillators that employ temperature compensati networks are called TCXO's (Temperatue Compensated Xtal Oscillator). TCXO' with frequency vs. temperature stabilities of 1-2 parts per million (1×10^{-6} over wide temperature ranges may employ very sophisticated analog and/ c digital compensation networks. Usually these temperature compensation techniques require the use of automatic test equipment and network synthesi computer programs to accurately compensate TCXO's on a mass production basis

Typical quartz crystal oscillator frequency vs. time (long term aging) characteristics are shown in Figure 32. Note that the characteristics ar primarily due to the quartz crystal and the aging can be positive or negative Quartz crystals are available that age at rates better than 1×10^{-11}/day some of the mechanisms that cause aging include:

1) Mass transfer due to contaninations within the quartz crystal enclosure. This can be minimized with proper process controls of cleaning and the environmental control including the crystal package sealing.
2) Stresses caused by the quartz crystal mounting and electrode deposition.
3) Oscillator circuit effects such as crystal drive level, oven control or temperature compensation cicuitry.
4) Materials outgassing and impurity diffusion.

Vocabulary

Over the years a special vocabulary of quartz crystal controlled oscillato terms has evolved. Some of these terms are listed on the next page fo your convenience.

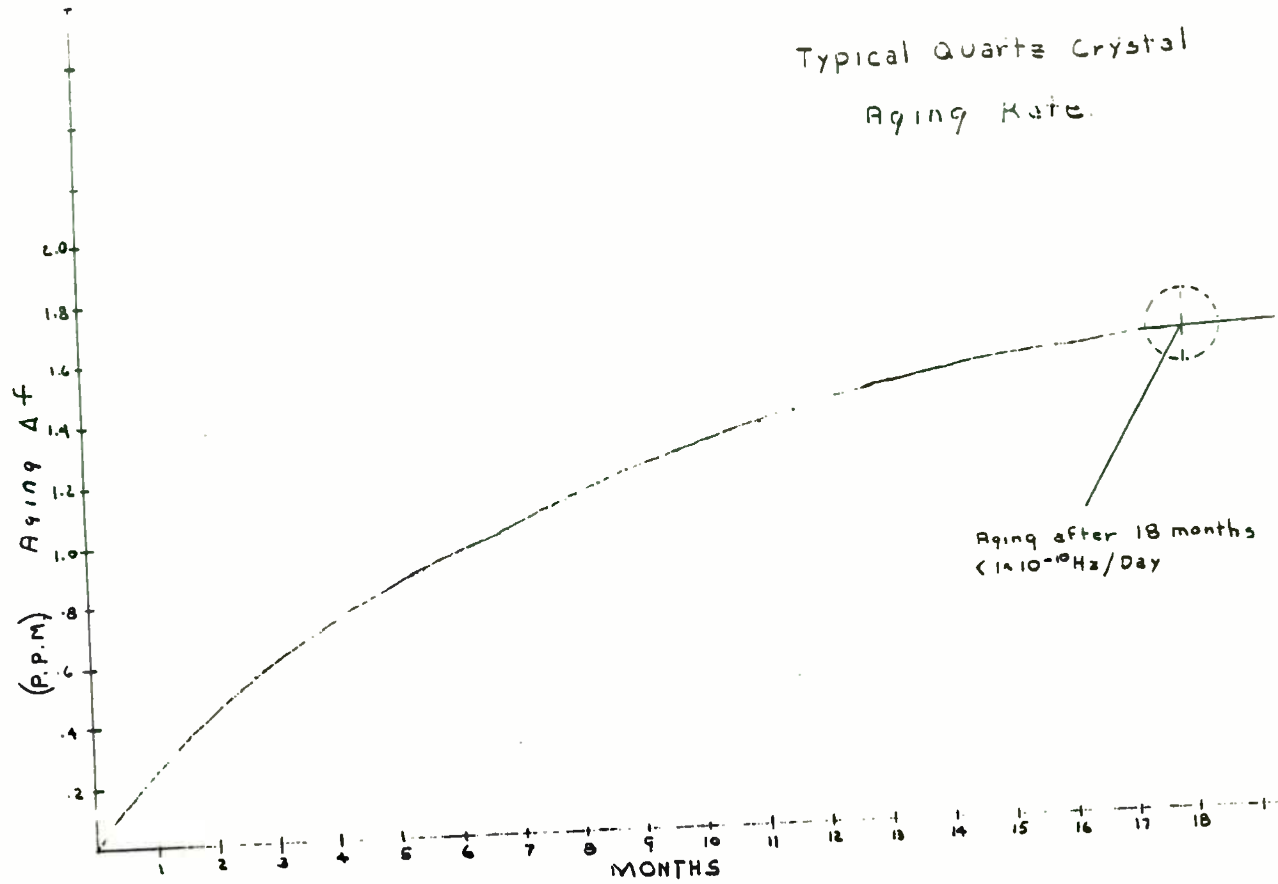
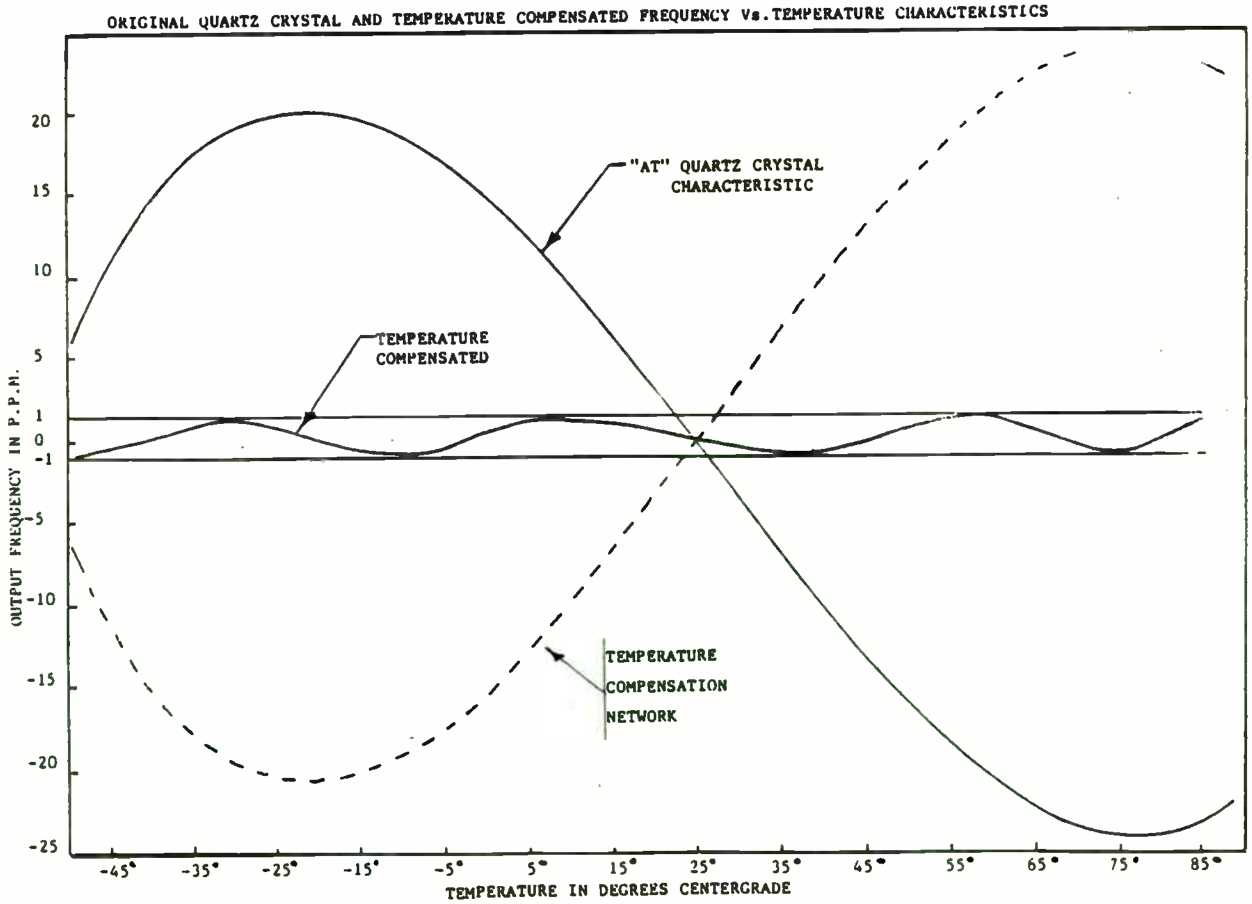


FIGURE 31

ORIGINAL QUARTZ CRYSTAL AND TEMPERATURE COMPENSATED FREQUENCY Vs. TEMPERATURE CHARACTERISTICS



MICROSONICS
INCORPORATED

CRYSTAL OSCILLATOR STABILITY VOCABULARY

Quartz crystal oscillators are often classified into one of four (4) types: crystal clock oscillaor (XO), ovenized crystal oscillators (OXO), voltage controlled crystal oscillators (VCXO), and temperature compensated crystal oscillators (TCXO).

Clock oscillators (XO) are usually a loose stability ($\gtrless$ $\pm$100PPM -55 to +85°C) crystal oscollators dependent primarly upon the specific quartz resonator for its frequency vs. temperature performance.

Ovenized quartz crystal oscillators provide much improved stabilities over clock oscillators (OXO) with stabilities of 1×10^{-10} being achieved when the crystal and/or oscillator temperature is controlled with the use of specially designed heaters and proportional controllers. Often multiple ovens and controllers may be designed with the total oscillator encased in a well insulated package using low-loss insulation or evacuated Dewar flasks.

Voltage controlled crystal oscillators (VCXO) are cyrstal oscillators that employ a varactor (voltage variable capacitor) network to provide a means of adjusting the frequency with an external control voltage. These oscillators are often used in automatic frequency control ciccuits or phase locked loops similar to that of Figure 15 as well as for direct F.M. applications as shown previously in Figure 4. The stability is dependent upon the specific crystal used and the external control voltage.

Temperature compensated crystal oscillators (XCXO) are similar to VCXO's in that a varactor is used to vary the frequency. However for TCXO's the varactor control voltage is generated internally by a temperature compensation network that uses themistors or diodes to provide a voltage that is proportional to the oscillator temperature. This voltage is then "shaped" through sophisticated analog or digital circuits to provide a temperature compensation voltage to the varactor. The resulting vibration nearly cancels the natural quartz crystal frequency vs. temperature characteristic to prove a compensated frequency change that often is several orders of magnitude better than the crystal without the compensation network. Frequency stabilities of 5×10^{-8} over wide operating temperatures have been achieved using highly specialized varactor compensation circuits.

Voltage controlled temperature compensated crystal oscillators (VCTCO's) combine the performance attributes of an external voltage control of the VCXO with the frequency vs. temperature stability (at a reference control voltage setting) of the TCXO.

MICROSONICS
INCORPORATED

While we have discussed oscillator frequency and stability in some detail there are some other specifications that should be considered.

These, along with frequency and stability definitions are listed as follows:

OSCILLATOR SPECIFICATIONS

Frequency Accuracy: The frequency setting tolerance at room temperature at time of shipment. Often oscillators are designed with tuning elements to allow periodic readjustment of the output frequency.

Stability: Usually specified as a minimum/maximum change from the nominal frequency as a +/- percentage, +/- parts per million (PPM) or decimal factors 10^{-6}, 10^{-7}, 10^{-10}, etc.

Aging: Expressed as ppm or decimal portions per time from days to years. Example: $\leq$1.0 x 10-6/year for TCXO's and $\leq 1 \times 10^{-9}$/day for OCXO's

Short Term Stability: Short term frequency stability usually is expressed as an Allen Variance (pp-10 per 10 ms) or in terms of SSB phase noise in the frequency domain.

Frequency Stabilities vs Supply Voltage and Load Variation: Frequency change from nominal as the supply voltage or load are varied. These stabilities are met by the use of voltage regulation circuits and buffer stages. The tighter the stability requirements, the more sophisticated the design.

Frequency vs. Temperature Stability: Usually expressed as +/-PPM (X 10^{-6}) or decimal factors (1^{-7}, 10^{-8}, 10^{-10}) over an operating temperature reference to the nominal frequency desired. Expl: +/- 1PPM -55°C to +85°C.

WARM UP TIME: The length of time that elapses, after power is applied to the oscillator circuit, for the output frequency and power to stabilize within specification limits. This is an important parameter for ovenized oscillators since the oven requires a finite time to reach its operating temperature.

TURN ON POWER: The amount of power that an oscillator consumes initially when power is first applied. This is an important ovenized oscillator parameter since peak power is required at turn-on from a cold start until the oscillator oven reaches its operating temperature.

MICROSONICS
INCORPORATED

Harmonic Signals: Signals which are coherently related to the output frequency.

Subharmonics: Harmonics of the actual oscillator (usually in oscillators that includes frequency multipliers).

Spurious Signals: Signals, other than the nominal output, which are not coherently related to the output frequency.

Phase Noise: The short-term frequency variations in the output frequency which appear as energy at frequencies other than the carrier. It is usually expressed in terms of dBc or as a RMS frequency deviation in a specified bandwidth at a specified frequency removed from the carrier. Phase noise is often specified at specific offset frequencies from the nominal output signal.

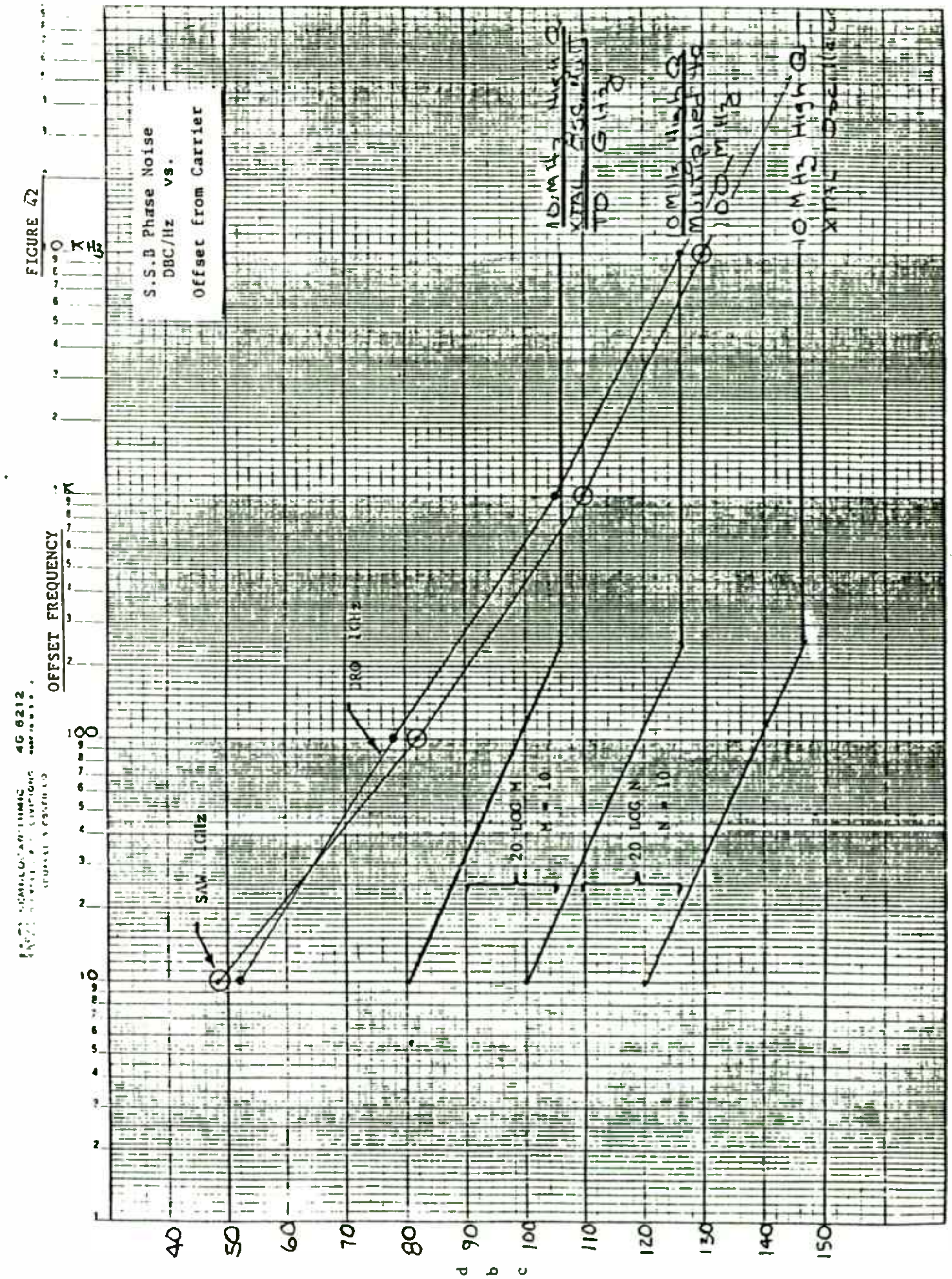
Output Power: The output power is usually defined in dbm into a specified load for a sine wave output.

Output Wave Forms: Output waveforms for digital logic output usually define the required logic levels "0", "1", duty cycle or symmetry, and number of digital devices to be connected in parallel as the oscillator load.

D.C. Voltage Inputs: Oscillators may require several different dc voltage supplies and currents depending upon the type of oscillator, logic families to be driven, and number of output signals. Ovenized oscillators usually require a high peak current supply seperate from that of the oscillator.

Size and Mechanical: This item defines the volume, dimensions and form factor as well as the type of packaging and input/output terminal locations.

Enviornmental and Screening: Environmental operating specifications for oscillators are of course dependent upon the system applications. However these items are often omitted or not properly considered in the early system development phases resulting sometimes in unpleasant surprises in the field.



MICROSONICS
INCORPORATED

OSCILLATOR PHASE NOISE

Oscillator phase noise performance is becoming increasingly important as the frequency spectrum becomes more crowded and low level signal detection techniques are becoming more sophistcated.

The chart in Figure 42 shows:

- 10MHz quartz crystal controlled oscillator phase noise characteristic
- 100MHz phase noise characteristic obtained by 10X multiplication of the 10MHz oscillator.
- 1GHz phase noise characteristic for the 10MHz oscillator multiplied by 100X.
- Phase noise characteristic of a typical bipolar dielectric resonator and oscillator SAW oscillator.

As the chart shows; quartz crystal oscillators provide the optimum phase noise performance, even when multiplied by 1GHz, for "close-in" offset frequencies of less than 1kHz. This is do to the high quartz crystal Q. However the other two (2) technologies, DRO and SAW, provide optimum noise floor (greater than 5kHz) performance because they do not require frequency multiplication to provide the 1GHz output; whereas the quartz crystal oscillator phase noise performance is degraded by 20 log of the multiplication factor or 40 db in this example.

In applications where an optimum phase noise floor performance is required at high output frequencies DRO and SAW oscillator technologies may provide the best choice. However if this requirement is combined with the need for tight stability an option may be to use the DRO or SAW oscillator as a VCO in a frequency control phase locked loop that is controlled by a stable quartz crystal controlled oscillator. The properly designed phase locked loop can also yield this benefit of the excellent close-in phase noise performance of the crystal oscillator and the noise floor of the DRO and SAW.

MICROSONICS
INCORPORATED

VCXO SPECIFICATIONS

VCXO:	A quartz crystal controlled oscillator which allows the frequency to be changed by applying an external control voltage.
VCTCXO:	A temperature compensated VCXO.
Frequency Deviation:	The amount that the center frequency will change as a function of control voltage; usually specified in +/-%. As the deviation is made larger, the other stabilities, temperature, aging will degrade. Example: +/-1% f/ OVDC ±5VDC
Linearity:	The allowable error from the best straight line for the f/ V characteristic.
Response Slope:	The slope of the voltage versus the control voltage, ± f/ V
Modulation Frequency:	The maximum or minimum control input frequency usually less than 10kHz.

MICROSONICS
INCORPORATED

Since most military specifications are comprehensive and organized they are often used vs. environmental testing standards even for industrial and consumer applications.
Some of the most common are:

Item	Specification
General Crystal Oscillator Specifiation	MIL-0-55310
Vibration	MIL-STD-202
Thermal Shock	MIL-STD-202
Moisture	MIL-STD-202
Microelectronics	MIL-STD-883
Radiation/Hardening	Various Classified

MICROSONICS
INCORPORATED

60 WINTER ST., WEYMOUTH, MA 02188-3336
(617) 337-4200 TWX 710-338-6833

CRYSTAL OSCILLATOR SPECIFICATION SHEET

FREQUENCY AND STABILITY:

Frequency ______

Stability, Specify Frequency Versus

Temperature ______

Time ______

Supply Variation ______ Load ______

Oper. Temp Range ______ °C to ______ °C

Freq. Adjustment ______

Power Supply:

Osc ______ ± ______ AT ______ ma

Oven ______ ± ______ AT ______ ma

Additional Comments: ______

OUTPUT WAVE FORM:

Sine ☐

Output Level ______

Harmonic Dist. ______ Subharmonics ______

Load ______

Square:

☐ TTL ☐ CMOS ☐ ECL

Rise Time ______ Fall Time ______

Measured From ______ to ______

Duty Cycle ______ at ______ level

Levels: Logic "0" ______ Logic "1" ______

Load: ______

Environmental:

Storage Temp Range ______ °C to ______ °C

Vibration ______

Shock ______

Moisture ______

Others ______

Voltage Control: (if applicable)

Freq. Deviation ______

Control Voltage ______

Linearity ______

Input Impedance ______

Modulation Freq. ______

ΔF/ΔV Slope Pos ☐ ______ Neg ☐ ______

Special ______

Mechanical:

Size ______

Termination ______

Finish ______

Mounting ______

Additions ______

Special Notes: ______

ST A SUBSIDIARY OF SIGNAL TECHNOLOGY CORPORATION

ANOTHER OSCILLATOR COMPARISON CHART

Technology	Frequency Range	Aging	Temperature Stability -40 to +85°C	Relative Size	Input Power	Approx. Cost
Mechanical R/C Osillator	DC to 100kHz	±20%	2%	Small	50mw	<$50
LC Oscillator	kHz to 500MHz	±20%	2%	Small	50mw	$20 to $150
DRO (Dielectric Resonator Oscillator)	3GHz to 16GHz	?	±400PPM	Small	100mw	$200 - $300
SAW - Surface Acoustic Wave Oscillator	25MHz to 200MHz	5PPM	±100PPM	Small	100mw	$75 - $200
Quartz Xtal Clock XO	800MHz to 200MHz	1 to 5PPM/ year	±5PPM		25 -50mw	$5 to $50
Quartz Temperature Compensated TCXO	10MHz to 60MHz	.2 to 1PPM/ year	±2PPM	Small	50mw	$40 - $200
Quartz Ovenized OCXO or PCXO	4MHz to 100MHz	.05 to 1PPM/ year	±.03PPM	Moderate to large	1 watt - 6 watts	$200 to $1000
Cesium	5MHz	.000010PPM 1×10^{-11}/year	$\pm 1 \times 10^{-11}$ (0-50°C)	Very large	25 watts	>$35K

TUTORIAL ON

SURFACE ACOUSTIC WAVE TECHNOLOGY

PRESENTED AT

RF TECHNOLOGY EXPO '85

Disneyland Hotel
Anaheim, CA

by

Carl A. Erikson, Jr.
Director of Processing Operations
Andersen Laboratories
1280 Blue Hills Avenue
Bloomfield, CT 06002

OUTLINE

WHAT IS A S.A.W. DEVICE?

- Parts of a SAW device
- Terminology
- Reasons for use
- Types of devices
- Applications

DESIGN CONSIDERATIONS

- Substrate
- Mask
- Package

FABRICATION SCHEME

TESTING

FUTURE TRENDS

THREE PARTS OF A SURFACE ACOUSTIC WAVE DEVICE

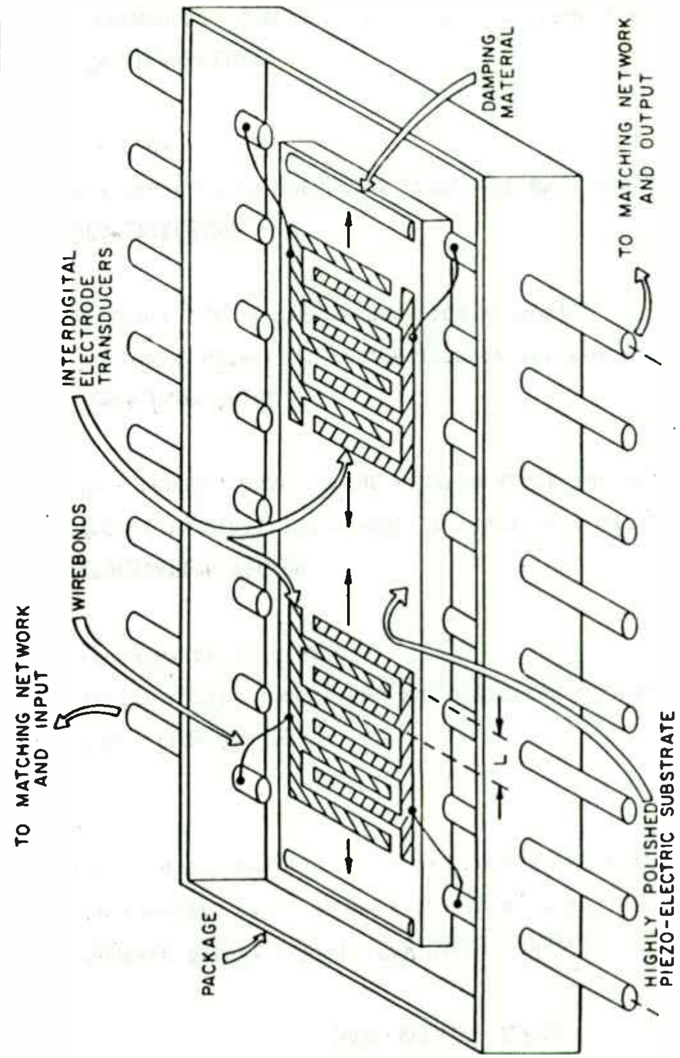
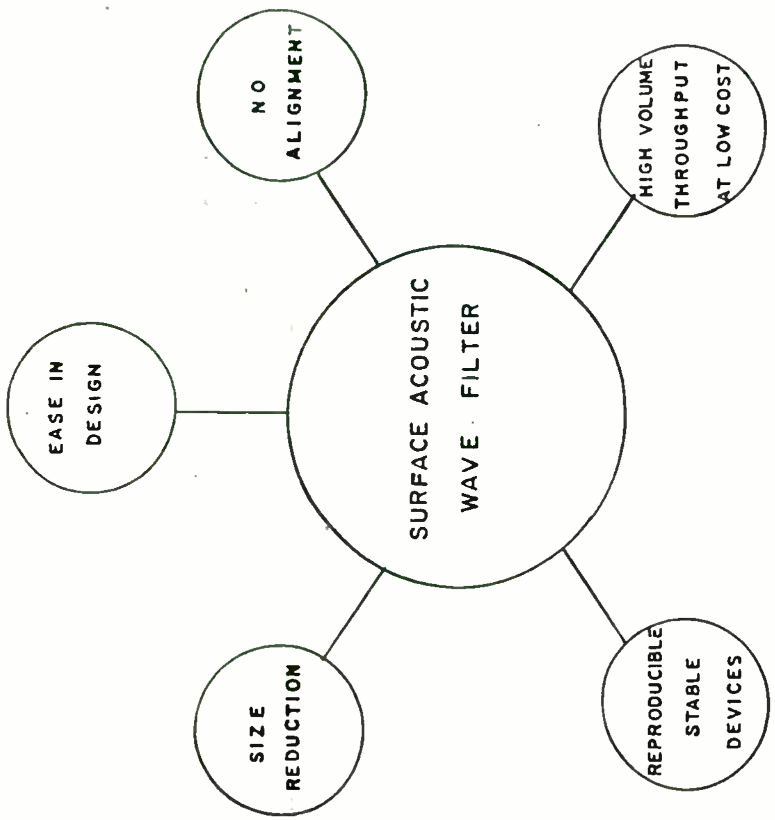


FIGURE 1



WAVE-RELATED TERMS

1. **Surface Acoustic Wave (SAW Raleigh Wave)**

   An acoustic wave, propagating along a surface of an elastic substrate, whose amplitude decays exponentially with substrate depth.

2. **Power Flow Vector**

   Vector giving magnitude and direction of power per unit area propagating in a wave.

3. **Propagation Vector**

   Vector in direction normal to lines of constant phase with magnitude proportional to the reciprocal of the wavelength.

4. **Power Flow Angle**

   The angle between the direction of the power flow vector and the direction of the propagation vector.

5. **Beam Steering**

   SAW propagation phenomena described by a non-zero angle of power flow.

6. **SAW Diffraction**

   A phenomenon analogous to optical diffraction due to the finite aperture of the source causing SAW beam spreading and wave-front distortion.

7. SAW Coupling Coefficient (K_s)

SAW electromechanical coupling coefficient is defined as follows:

$$K_s^2 = 2\left|\Delta v/v\right|$$

where $\Delta v/v$ is the fractional phase velocity change produced by short-circuiting the surface potential.

8. Acoustic Regeneration (for SAW)

The generation of a secondary acoustic wave by the potential variations of an electrode caused by primary SAW passing under it.

9. Acoustic Propagation Loss

The ratio of the power transmitted in a SAW beam to the power received in a cross-section of the same width, expressed in dB. Propagation loss includes power dissipation or scattering due to material damping, diffraction, defects and radiation into the air above the substrate.

10. Mass Loading

The change in phase velocity of a SAW produced by a thin layer on the substrate of higher density than that of the substrate. (Alternate definition) perturbations in reflections, velocity, dispersion, etc., due to loading effects of thin films on the substrate surface.

11. Beamwidth

The spatial distance in units of wavelength which contain 50% of the acoustic energy.

12. Surface Skimming (Shallow-Bulk) Acoustic Wave

A horizontally polarized bulk shear wave radiated almost parallel to the substrate surface. These waves are not a single mode as with the SAW but have many of the properties of a SAW.

TRANSDUCER-RELATED TERMS

1. Interdigital Transducer (IDT)

A comb structure applied to the surface of a substrate consisting of interleaved metal electrodes whose function is to transform electrical energy into acoustic energy or vice versa by means of the piezoelectric effect.

2. Unidirectional Interdigital Transducer (UDT)

Transducer capable of radiating and receiving surface acoustic waves in a single direction.

3. Multiphase Transducer

Interdigital transducer having more than two inputs which are driven in different phases. Usually used for unidirectional transducers.

4. Phase Coded Transducer

An IDT in which the phase of the signal from an individual tap is determined simply by the polarity of the connections to the bus bars.

5. Weighted-Response Transducer

A transducer intended to produce a surface wave with spatial distribution corresponding to a weighted impulse respone.

TRANSDUCER-RELATED TERMS (contd)

6. Focussing IDT

An IDT with curved electrodes to focus the launched acoustic wave to a narrower beam width.

7. Finger

An element of the IDT comb electrode.

8. Dummy Finger

A passive finger which may be included in an IDT in order to suppress wavefront distortion.

9. Bus Bar

A common electrode which both connects individual fingers of an IDT together and the transducer to an external circuit.

10. Aperture

Maximum IDT finger overlap length.

11. Finger Overlap (acoustic aperture)

The length of a finger pair between which electromechanical interaction is generated.

12. Apodization

Response weighting due to the change of finger overlap along the length of the IDT.

TRANSDUCER-RELATED TERMS (CONTD)

13. WITHDRAWAL WEIGHTING

RESPONSE WEIGHTING BY REMOVAL OF A FINGER OR REPLACEMENT OF AN ACTIVE FINGER WITH A DUMMY FINGER IN THE IDT.

14. CAPACITIVE WEIGHTING

RESPONSE WEIGHTING BY INCLUDING CAPACITANCE IN SERIES WITH THE CONNECTION OF EACH ELECTRODE AS PART OF THE IDT: THE CAPACITANCE VALUE IS VARIED FROM ELECTRODE TO ELECTRODE.

15. SERIES WEIGHTING

RESPONSE WEIGHTING BY FINGER SEPARATION INTO INDIVIDUAL ELEMENTS BY SEPARATION FROM BUS BAR.

16. PHASE WEIGHTING

RESPONSE WEIGHTING BY CHANGE IN PERIOD OR PHASE OF FINGER ARRANGEMENT INSIDE THE IDT.

17. STRIP-TO-GAP-RATIO

THE RATIO OF THE METALLIZED SURFACE TO THE FREE SURFACE WITHIN THE IDT.

18. DOUBLE (SPLIT) ELECTRODE

QUARTER WAVELENGTH SPACED (CENTER-TO-CENTER) FINGERS, TYPICALLY ONE-EIGHTH WAVELENGTH WIDE, USED TO REDUCE REFLECTIONS FROM TRANSDUCERS.

DEVICE-RELATED TERMS

1. SURFACE ACOUSTIC WAVE FILTER

A FILTER UTILIZING SURFACE ACOUSTIC WAVES WHICH ARE USUALLY GENERATED BY AN INTERDIGITAL TRANSDUCER AND PROPAGATE ALONG A SUBSTRATE SURFACE TO A RECEIVING TRANSDUCER.

2. SAW RESONATOR FILTER

A TYPE OF SAW FILTER OFFERING A HIGH Q DUE TO EFFICIENT REFLECTORS IN THE FABRY-PEROT RESONATOR CAVITY STRUCTURE.

3. MULTISTRIP COUPLER (MSC)

AN ARRAY OF METAL STRIPS DEPOSITED ON A PIEZOELECTRIC SUBSTRATE IN A DIRECTION TRANSVERSE TO THE PROPAGATION DIRECTION WHICH CAN TRANSFER ACOUSTIC POWER FROM ONE ACOUSTIC TRACK TO AN ADJACENT TRACK.

4. REFLECTOR

A SAW REFLECTING COMPONENT WHICH NORMALLY MAKES USE OF THE PERIODIC DISCONTINUITY PROVIDED BY AN ARRAY OF METAL STRIPS, DOTS OR GROOVES.

5. SHIELDING ELECTRODE

ELECTRODE INTENDED FOR THE REDUCTION OF ELECTROMAGNETIC INTERFERENCE SIGNALS.

DEVICE-RELATED TERMS (CONTD)

6. SUPPRESSION CORRUGATION

GROOVES IN THE NON-ACTIVE SIDE OF THE SUBSTRATE FOR SUPPRESSING BULK WAVE SIGNALS.

7. ACOUSTIC ABSORBER

MATERIAL WITH HIGH ACOUSTIC LOSS AT OPERATING FREQUENCY PLACED ON ANY PART OF SUBSTRATE FOR ACOUSTIC ABSORPTION PURPOSES.

8. ACOUSTIC WAVEGUIDE

A PERTURBATION ALONG THE DIRECTION OF PROPAGATION OF A SAW TO PRODUCE A DECREASED PHASE VELOCITY AND HENCE TRANSVERSE CONCENTRATION AND GUIDING OF THE SAW.

9. BEAM COMPRESSORS

STRUCTURES ON THE SURFACE OF A SUSBSTRATE TO INCREASE THE POWER DENSITY IN A SAW BY DECREASING ITS LATERAL EXTENT:

(A) HORN: TAPERED STRUCTURE OF REDUCED VELOCITY TO PRODUCE GRADUAL REDUCTION OF TRANSVERSE WIDTH OF BEAM.

(B) MULTISTRIP BEAM COMPRESSOR: A MULTISTRIP COUPLER WITH SPACING OF THE STRIPS CHOSEN SO THAT ONE TRACK IS APPRECIABLE WIDER THAN THE OTHER.

(C) LENSES: REGIONS OF DECREASED PHASE VELOCITY SO SHAPED AS TO PRODUCE FOCUSSING OF AN INCIDENT SAW BEAM.

DEVICE-RELATED TERMS (CONTD)

10. CONVOLVER

A THREE PORT DEVICE WHOSE OUTPUT SIGNAL IS THE CONVOLUTION OF TWO SIGNALS APPLIED SIMULTANEOUSLY TO THE INPUT PORTS.

11. CHIRP FILTER:

A FILTER WHOSE GROUP DELAY IS A NON-CONSTANT FUNCTION OF THE INSTANTANEOUS FREQUENCY OF THE INPUT SIGNAL.

12. LINEAR FM CHIRP FILTER

A CHIRP FILTER WHICH MANIFESTS A LINEAR DELAY VARIATION WITH FREQUENCY.

13. REFLECTIVE ARRAY COMPRESSOR (RAC)

A TYPE OF DEVICE WHICH USES REFLECTIONS OF THE SURFACE ACOUSTIC WAVE FROM OBLIQUE GROOVES OR STRIPES TO ACHIEVE THE DESIRED DISPERSIVE DELAY FUNCTION.

14. REFLECTIVE DOT ARRAY (RDA)

A TYPE OF DEVICE WHICH USES REFLECTIONS OF THE SURFACE ACOUSTIC WAVE FROM OBLIQUE ROWS OF METALLIC DOTS.

15. SAW OSCILLATOR

AN OSCILLATOR THAT USES A SAW DEVICE (RESONATOR OR DELAY LINE) AS THE MAIN FREQUENCY CONTROLLING ELEMENT.

DEVICE-RELATED TERMS (CONTD)

16. **Oscillator Mode**

Frequency or frequencies for which the total phase shift around the oscillator loop is an integer multiple of 2π.

17. **Excess Gain**

The value of the positive gain (in decibels) at any specified frequency for the open oscillator loop measured under small signal conditions (no limiting action). The source and load impedance must be specified.

18. **Single-Mode SAW Oscillator**

A SAW oscillator in which there is only one frequency which satisfies the oscillation conditions of having positive excess gain and total phase shift of $N \cdot 2\pi$ (where N is a positive integer).

19. **Multimode SAW Oscillator**

A SAW oscillator in which more than one frequency satisfies the oscillation conditions.

20. **Delay Line**

A device which operates over some defined range of electrical and environmental conditions as a linear passive circuit element. The transfer characteristic has a modulus and argument (phase) which can be constant or a function of frequency.

DEVICE-RELATED TERMS (CONTD)

21. **Delay Line, Nondispersive**

A delay line which nominally has constant delay over a specified frequency band. The argument (phase) of the transfer function is a linear function of frequency.

22. **Delay Line, Dispersive**

A delay line which has a transfer characteristic with a constant modulus and an argument (phase) which is a nonlinear function of frequency. The phase characteristic of devices of common interest is a quadratic function of frequency, but in general may be represented by higher order polynomials or other nonlinear functions.

23. **Delay Time: (Nondispersive Delay Line)**

The transit time of the envelope of an RF tone Burst.

24. **Phase Shift: (Dispersive and Nondispersive Delay Lines)**

The total number of degrees or radians between the phase of the CW input signal and the CW output signal as the delay device is operated at a given frequency within the band of operation. The phase shift is nominally a linear-function of frequency within the frequency band of operation for a nondispersive delay device.

25. **Phase Delay: (Dispersive and Nondispersive Delay Lines)**

The ratio of total radian phase shift, ϕ, to the specified radian frequency, ω. Phase delay is nominally constant over the frequency band of operation for non-dispersive delay devices.

DEVICE-RELATED TERMS (CONTD)

26. GROUP DELAY: (NONDISPERSIVE DELAY LINE)

THE DERIVATIVE OF RADIAN PHASE WITH RESPECT TO RADIAN FREQUENCY, $\partial\phi/\partial\omega$. IT IS EQUAL TO THE PHASE DELAY FOR AN IDEAL NONDISPERSIVE DELAY DEVICE, BUT MAY DIFFER IN ACTUAL DEVICES WHERE THERE IS RIPPLE IN THE PHASE VS. FREQUENCY CHARACTERISTIC.

27. DELAY DISPERSION: (DISPERSIVE DELAY LINE)

THE CHANGE IN PHASE DELAY OVER A SPECIFIED OPERATING FREQUENCY RANGE.

28. DISPERSIVE BANDWIDTH: (DISPERSIVE DELAY LINE).

THE OPERATING FREQUENCY RANGE OVER WHICH THE DELAY DISPERSION IS DEFINED.

29. DELAY SLOPE: (DISPERSIVE DELAY LINE)

THE RATIO OF THE DELAY DISPERSION TO THE DISPERSIVE BANDWIDTH.

30. CENTER FREQUENCY DELAY: (DISPERSIVE DELAY LINE)

THE PHASE DELAY OF THE DEVICE AT THE CENTER FREQUENCY, F_0, GENERALLY EXPRESSED IN MICROSECONDS.

31. BULK WAVE SIGNALS

UNWANTED SIGNALS CAUSED BY BULK WAVE EXCITATION EXISTING AT THE FILTER OUTPUT.

32 SPURIOUS REFLECTIONS

UNWANTED SIGNALS CAUSED BY REFLECTION OF SAW OR BULK WAVES FROM SUBSTRATE EDGES OR ELECTRODES.

DEVICE-RELATED TERMS (CONTD)

33. TRIPLE TRANSIT ECHO (TTE)

UNWANTED SIGNALS IN A SAW WHICH HAVE TRANSVERSED 3 TIMES THE PROPAGATION PATH BETWEEN INPUT AND OUTPUT IDT'S CAUSED BY ACOUSTIC REFLECTIONS AT THE IDT'S.

34. FEEDTHROUGH SIGNAL

THE UNDELAYED SIGNAL RESULTING FROM DIRECT COUPLING BETWEEN THE INPUT AND THE OUTPUT OF THE DEVICE.

35. MULTIPLE TRANSIT SIGNALS: (DISPERSIVE AND NONDISPERSIVE DELAY LINES).

SPURIOUS SIGNALS HAVING DELAY TIME RELATED TO THE MAIN SIGNAL DELAY BY SMALL ODD INTEGERS. SPECIFIC MULTIPLE TRANSIT SIGNALS MAY BE LABELED THE THIRD TRANSIT (TRIPLE TRANSIT), FIFTH TRANSIT, ETC. THERE IS OFTEN A TRADEOFF AVAILABLE BETWEEN MULTIPLE TRANSIT SIGNAL LEVELS AND BANDWIDTH, DELAY TIME, INSERTION LOSS, AND VSWR.

36. NON-MULTIPLE TRANSIT SPURIOUS SIGNALS: (DISPERSIVE AND NON-DISPERSIVE DELAY LINES

SIGNALS NOT RELATED TO THE MAIN SIGNAL DELAY BY A SIMPLE INTEGER MAY BE LABELED BY THE DELAY TIME OF THAT SIGNAL.

37. BANDWIDTH: (DISPERSIVE AND NONDISPERSIVE DELAY LINES).

A SPECIFIED FREQUENCY RANGE OVER WHICH THE AMPLITUDE RESPONSE DOES NOT VARY MORE THAN A DEFINED AMOUNT. NOTE: TYPICALLY, AMPLITUDE VARIATIONS TO SPECIFY BANDWIDTH ARE 1 dB AND 3 dB.

38. TIME-BANDWIDTH PRODUCT

THE PRODUCT OF THE DEVICE TIME DURATION AND THE CHIRP BANDWIDTH.

39. COMPRESSION GAIN

10 LOG OF THE RATIO OF THE MAGNITUDE OF THE PEAK POWER OF A COMPRESSED PULSE TO THE RMS NOISE POWER MEASURED. FOR AN UNWEIGHTED CHIRP PULSE COMPRESSION SYSTEM, THE VALUE IS 10 LOG (TB) WHERE TB IS THE TIME BANDWIDTH PRODUCT.

40. INSERTION LOSS:(1) (PULSE DELAY LINE)

THE RATIO OF THE INPUT PULSE POWER TO THE OUTPUT POWER OF THE MAIN PULSE EXPRESSED IN DECIBELS.

(2) (CW DELAY LINE)

THE RATIO OF INPUT POWER TO TOTAL OUTPUT POWER, NORMALLY EXPRESSED IN DECIBELS.

(NOTE: BOTH THE SOURCE IMPEDANCE AND LOAD IMPEDANCE MUST BE SPECIFIED.)

INSERTION LOSS CONSISTS OF THE FOLLOWING COMPONENTS:

(1) PROPAGATION LOSS DUE TO SAW ATTENUATION
  (A) INTERACTION WITH THERMALLY EXCITED ELASTIC WAVES (TEMPERATURE DEPENDENT)
  (B) SCATTERING BY CRYSTALLINE DEFECTS, IMPURITIES, PITS AND SCRATCHES IN THE OPTICAL POLISH (TEMPERATURE INDEPENDENT)

DEVICE-RELATED TERMS (CONTD)

  (C) AIR LOADING LOSS (PRESSURE DEPENDENT)
  (D) HARMONIC CONVERSION LOSS (POWER LEVEL DEPENDENT)

(2) LOSSES DUE TO TRANSDUCER GEOMETRY
  (A) BEAM STEERING
  (B) DIFFRACTION

(3) TRANSDUCER LOSSES
  (A) CONDUCTION LOSSES
  (B) BULK WAVE EXCITATION
  (C) IMPEDANCE MISMATCH

41. INSERTIONLOSS RIPPLE:

THE PEAK-TO-PEAK VARIATION OF THE INSERTION LOSS, I.E., THE DIFFERENCE BETWEEN THE MAXIMUM AND MINIMUM INSERTION LOSS, OVER A SPECIFIED FREQUENCY RANGE OF THE DEVICE.

Some Applications of SAW Devices

DELAY LINES

Fusing, MTI Radar, communciations path length equalizer, altimetry, time ordering, target simulation

WIDEBAND DELAY LINES

Recirculating digital storage, nuclear experiments

DIFFERENTIAL DELAY LINES

Data communciations

BANDPASS FILTERS AND RESONATORS

Color TV, radar, CATV, repeaters, transposers, ECM, frequency synthesis

OSCILLATORS

Stable source VHF to microwave, local oscillators for communciations and coherent radar

TAPPED DELAY LINES

Fourier transformation, acoustic image scanning, clutter reference radar, ECM deception

DISPERSIVE DELAY LINES (CHIRP)

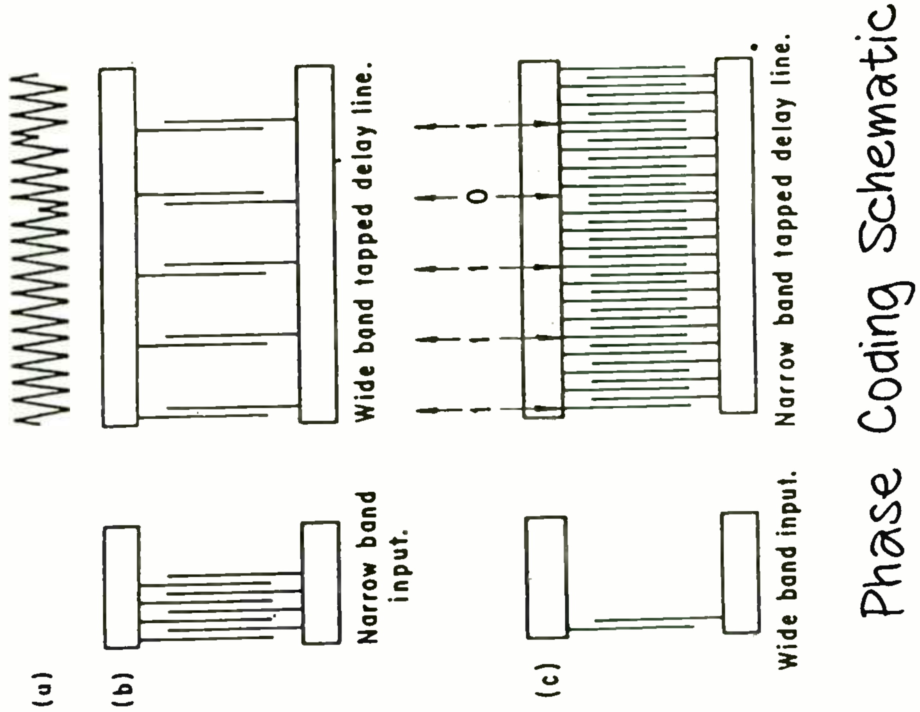
Radar pulse compfession/ expansion, variable delay for target simulation, fourier trnasforms(spectral analysis), compressive receiver, group delay equalization

PSK FILTERS

Spread spectrum communciations, radar military ATC

CONVOLVERS

Synchronizer for spread spectrum communciators, fourier transformation



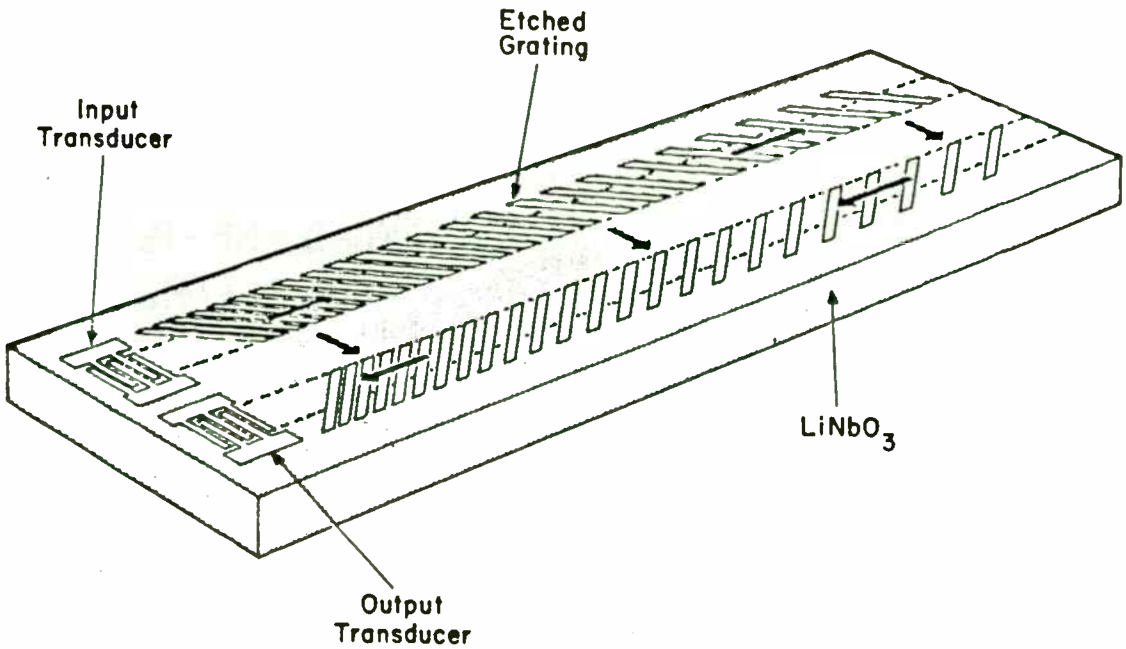
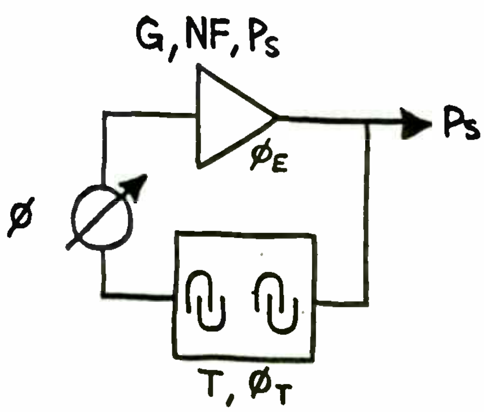
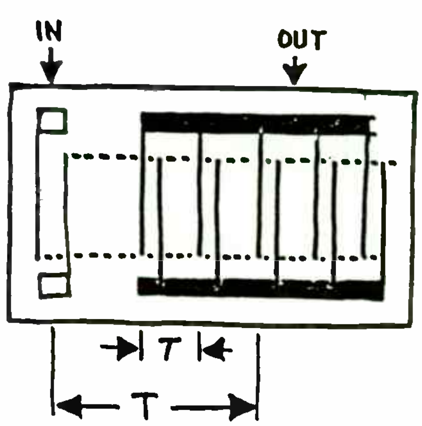


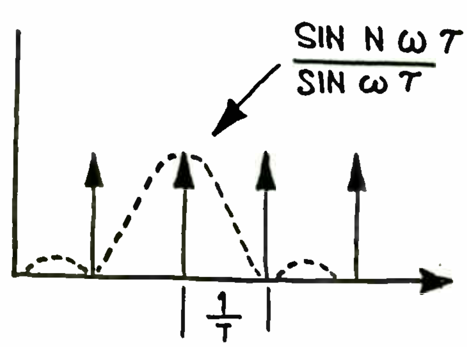
Figure 1

SAW OSCILLATORS



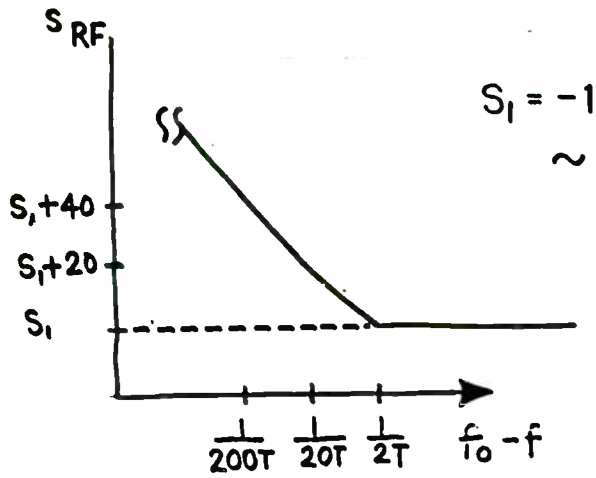
$$\phi_T + \phi_E + \phi = 2n\pi$$
$$\phi_T = 2\pi f_o T >> \phi_E$$
$$\Delta f_o \leq 1/2T$$
$$Q = \pi f_o T$$





OSCILLATORS

SAW Oscillators 2



$S_1 = -176 + G + NF - P_s$ dbc/Hz
~ -145 dbc/Hz

NOISE POWER SPECTRUM

Short term stability $= \frac{\sqrt{S_1}}{Q} \sim \frac{6 \times 10^{-8}}{Q}$ per sec.

Medium term stability $\sim .03$ ppm/C°^{2} ref. 25°C

Long term stability ~ 10 ppm/year

f_0 100 to 1000 MHz Q to 30 f_0

T to 10 μS $\Delta f = \frac{1}{4T}$

594

Comparison of Properties of Various Oscillators

Oscillator	Approx. Frequency Range	Effective Loaded Q	Max. Freq. Deviation ppm	Temp. Coeff in ppm/°C (−30°C to +70°C)
Conventional Quartz XTAL	$<10^8$ Hz	$5000-2.10^6$	≈500	<1 ppm/°C
LC (including cavity oscillators)	10^3-10^{11} Hz	Typically $10-10^4$	as large as required	Typically 10 ppm/°C
SAW	$10^7-2 10^9$ Hz	$200-10^4$ by choice of 1	10^2-10^4 by choice of 1	average value ≈1 ppm/°C

Comparison

SAW SENSORS

Some Physical Parameters Which Can Be Measured:

- Pressure
- Force
- Displacement
- Acceleration
- Vibration/Shock
- Temperature

Measurement is accomplished by noting the change in the transducer characteristics due to mechanically or thermally induced stress.

For SAW resonators and/or delay line oscillators, usually it is the frequency shift.

ADVANTAGES

- Fast Response
- High Sensitivity
- Small Size
- Good Output Signal

DISADVANTAGES

- Packaging (Environmental)
- Sensitive to Shock/Vibration (mounting)

TYPICAL SAW MATERIALS

HUGHES ⑥

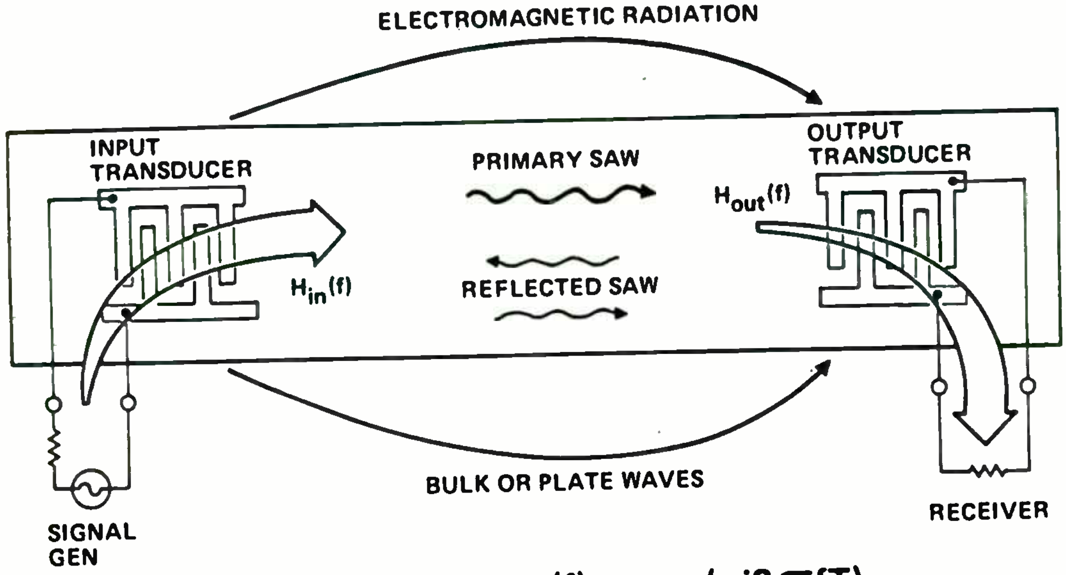
	YZ $LiNbO_3$	"MDC" $LiTaO_3$	ST-X QUARTZ	(001) (110) $Bi_{12}GeO_{20}$
ELECTRO-MECHANICAL COUPLING CONSTANT k^2	0.044	0.0067	0.0015	0.014
RELATIVE DIELECTRIC CONSTANT	50	48	4.4	38
VELOCITY M/SEC	3488	3378	3156	1681
LINEAR TEMPERATURE COEFFICIENT OF DELAY $1/\tau\ d\tau/dT$	$94 \times 10^{-6}/^{\circ}C$	$64 \times 10^{-6}/^{\circ}C$	0	–
TYPICAL FRACTIONAL BANDWIDTH APPLICABILITY RANGE	> 15%	5 – 15%	≤ 5%	5 – 20%

95227-23 (5-1-79)

Courtesy of W.R. Smith, Jr.

FILTER TRANSFER FUNCTION AS PRODUCT OF TRANSDUCER TRANSFER FUNCTIONS

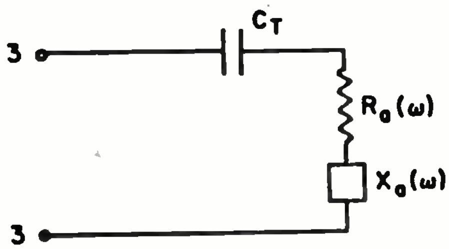
HUGHES (19)



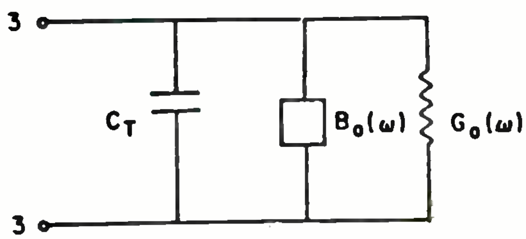
$$H(f) \cong H_{in}(f) \cdot H_{out}(f) \cdot \exp(-j2\pi fT)$$

95227-19 (5-1-79)

Courtesy of W.R. Smith, Jr.



(a)

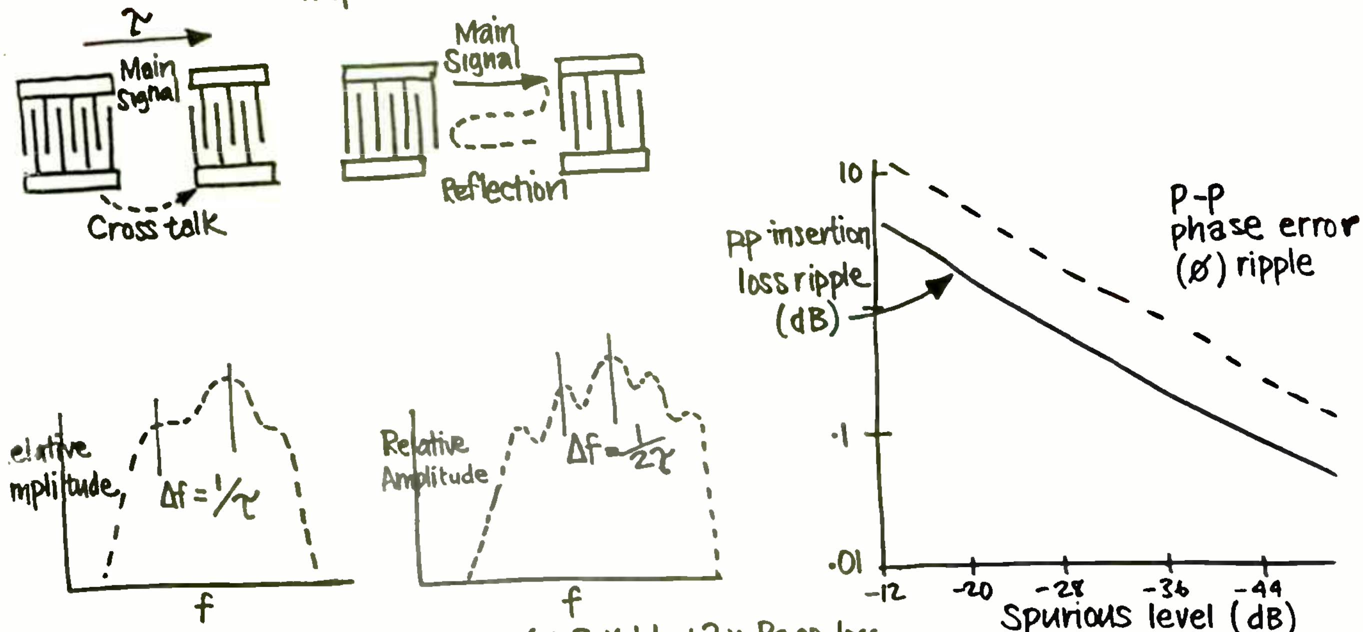


(b)

EQUIVALENT CIRCUITS FOR IDT

SECONDARY EFFECTS

Triple transit echo and feedthrough.



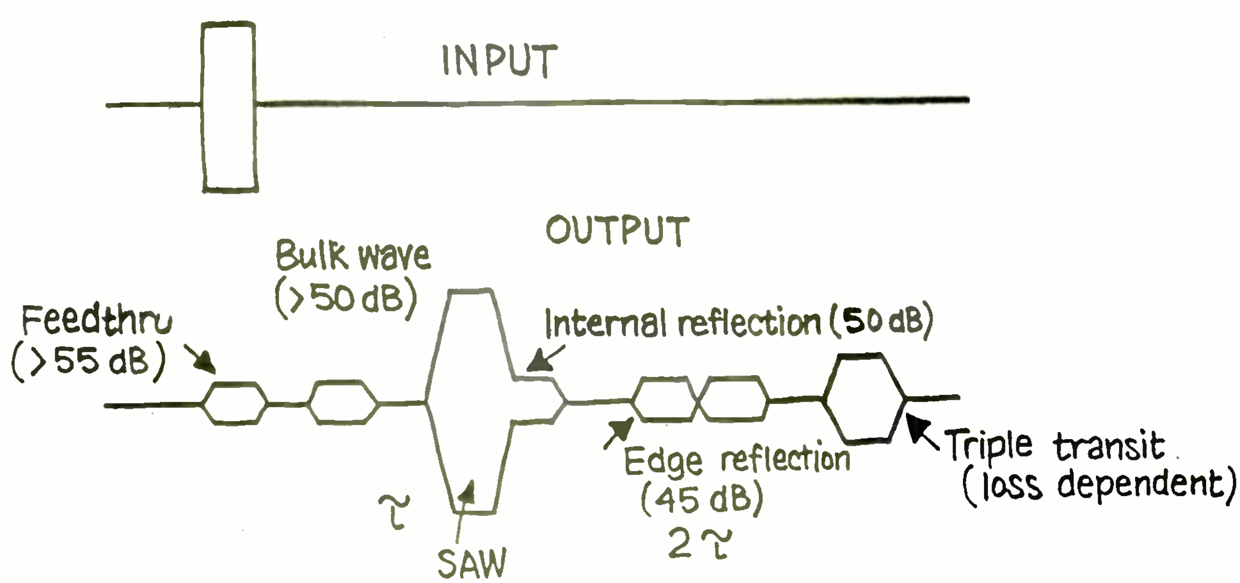
TTE = 6 + 2 x I.L. + 2 x Pr. op. loss

597

SECOND ORDER EFFECTS

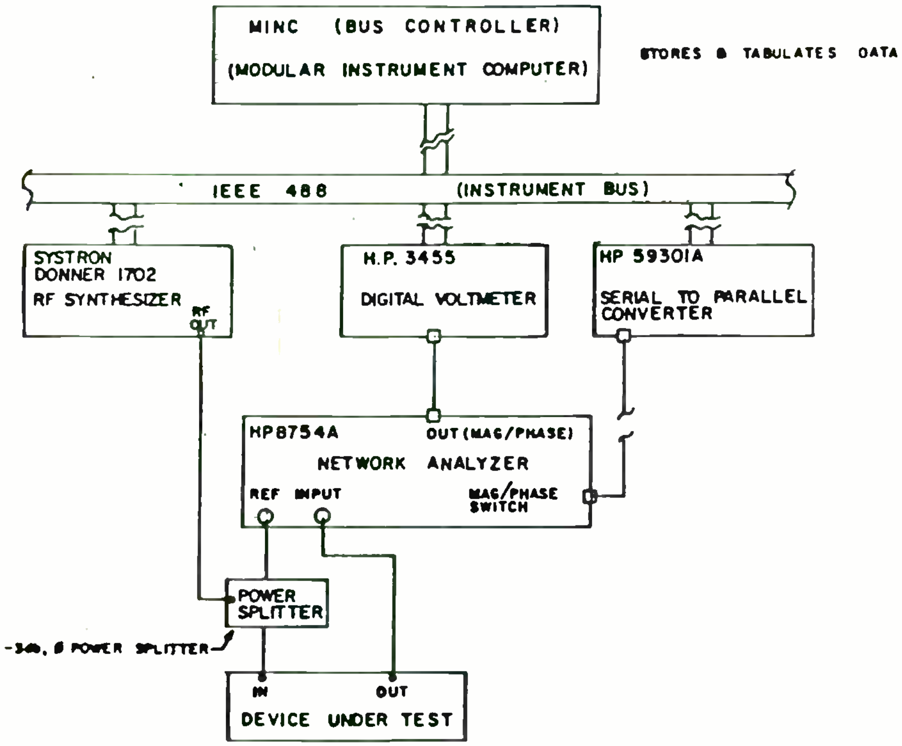
Time domain spurious

Edge reflections – angle, absorber
Bulk waves – roughen
Internal reflections – metal – adjust velocity
– mass – clever design



TYPICAL TEST STATION FOP DISPERSIVE
DELAY DEVICES

FIG I



FUTURE TRENDS

High Frequency SAW Devices (>400 MHz)

- Surface Quality of Substrate Material
- Metallization Techniques - "liftoff, plasma etching, ion beam etching"
- Particulate Contamination
- Resolution of Structures - exposure source, multi-layer resist techniques, direct write vs. contact or projection printing.
- Handling - better assembly techniques
- Inspecting - SEM

SAW/Hybrid Integration

- Passive Tuning Networks
- Active Amplification Circuits
- High Frequency Devices
- Packaging and Size Reduction
- Programmability

SAW Materials

- Lithium Tantalate
- Zinc Oxide
- Lithium Tetraborate
- Berlinite
- Gallium Arsenide

FUTURE TRENDS (CONTINUED)

- High Volume Applications
  - Mature Technology
  - Low Cost
- Devices and Systems
  - Grooved Devices
    - reflective array compressors (RAC's)
    - resonators
    - buried IDT
  - Convolvers
  - Channelized Receivers
  - Compressive Receivers
- Automation
  - Automated Equipment
    - inspection of substrates
    - inspection of etched wafers
    - damping
    - wire bonding
    - testing
  - Automated SAW Device Factory
    - integration of CAD/CAM/CAT

INTEGRATED APPROACH TO FABRICATION OF HIGH FREQUENCY SAW DEVICES

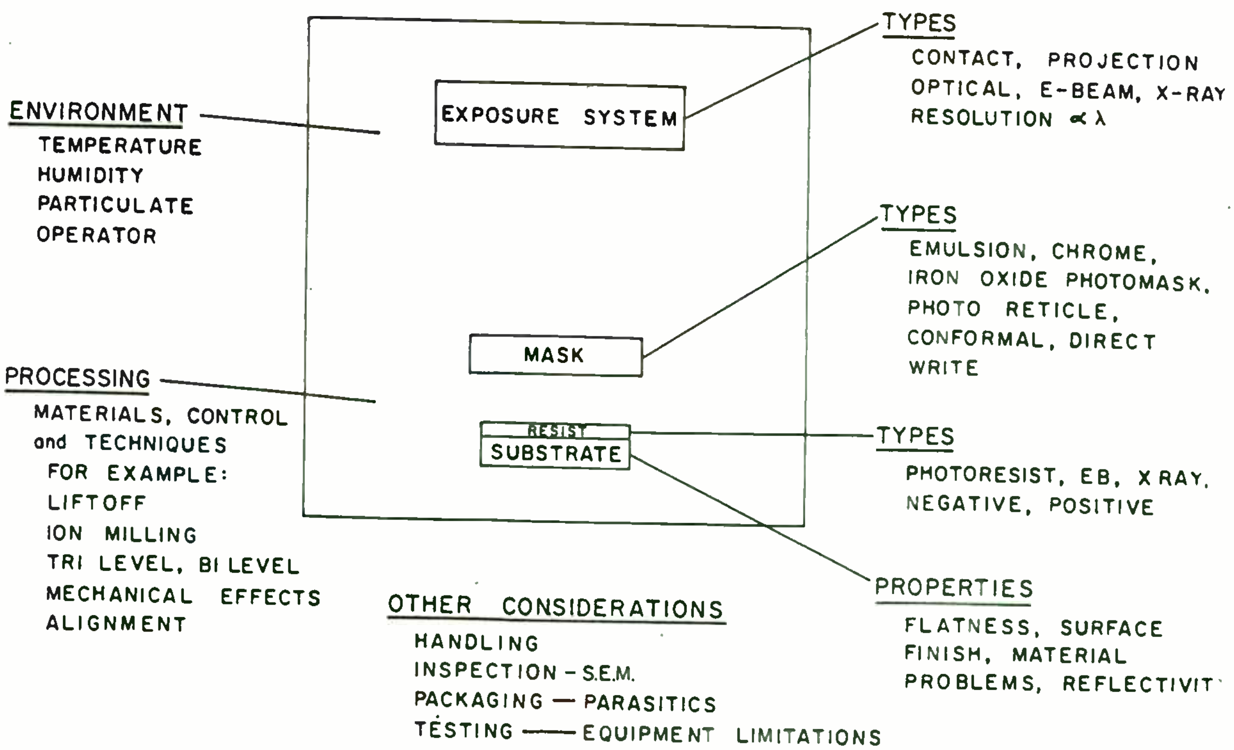


FIGURE 2