

# *Proceedings*

## *RF expo*

---

## *EAST*

*September 22-24, 1992  
Tampa Convention Center  
Tampa, Florida*

*Sponsored by*  
*RF design*  
*Magazine*



*A. J. Fisher*  
*9/22/92*



# *Proceedings*

## ***RF expo*** ***EAST***

---

*September 22-24, 1992*  
*Tampa Convention Center*  
*Tampa, Florida*

**Copyright © 1992 by**  
Cardiff Publishing Company  
*All rights reserved.*

***RF design***

6300 S. Syracuse Way  
Suite 650  
Englewood, Colorado 80111

1995  
10-3  
10-11

# TABLE OF CONTENTS

## SESSIONS ON TUESDAY, SEPTEMBER 22

### Session A-1: CAD Techniques I

*Computer-Aided Design of RF Circuits by the Addition of New Components*  
Kiomars Anvari,  
Teknekron Communications Systems, Inc.....481

*Practical Applications of Non-Linear Analysis, Simulation and Measurement Techniques*  
Joel P. Dunsmore,  
Hewlett Packard Company.....1

*Practical Object Oriented CAE for RF Engineers*  
Dale A. Teets, P.E.,  
Consultant.....19

### Session A-2: Medical and Scientific Applications

*ATS-3: Celebrating 25 Years of Service in Space*  
Michael A. Cauley,  
NASA Lewis Research Center.....N/A

*Signal Synthesis for Controlled Particle Extraction from High-Energy Accelerators*  
H. Meuth, G. Heinrichs,  
H. Halling, S. Papureanu,  
Forschungszentrum Julich.....31

*Noncontacting Coal and Rock Thickness Measurement with a Vector Network Analyzer*  
Robert L. Chufo,  
U.S. Bureau of Mines.....41

### Session A-3: Wireless Communications Applications

*A Rapid Acquisition Technique of Spread Spectrum Signals Embedded in an ASIC Costas Phase Locked Loop*  
Dr. Madjid A. Belkerdid, G.G. Koller,  
University of Central Florida  
G.S. Rawlins,  
Signal Technologies, Inc.....51

*Field Strength Measurements for FCC Compliance*  
Earl McCune,  
Proxim, Inc.....485

*Cosine Transition-Shaping PSK Modulators Using Direct Digital Synthesizers*  
Roy Greeff, Bruce Williams,  
Paramax Systems Corp.....61

### Session A-4: A/D Converter Tutorial

*Understanding Data Converters' Frequency Domain Specifications*  
Robert E. Leonard, Jr.,  
Datel Inc.....69

### Session B-1: CAD Techniques II

*Reference Manual for the Smith Chart Companion RF Design Software*  
David J. Brunell.....85

*Comparing Simulation Methods for RF Designs*  
Xiaohua Xuan,  
Hewlett Packard Company.....97

*Measurement of RF Components in a Microstrip Environment*  
Chuck McGuire,  
EEsof, Inc.....N/A

### B-2: Complex Modulation

*A Network Theoretic Approach to the Reduction of Distortion in Quadrature Detectors*  
Thomas G. Xydis, Ph.D.,  
Consultant.....N/A

*The Vector Discriminator*  
Billy P. Ficklin, Marilyn F. Williams,  
SRI International.....113

*Third Generation Spread Spectrum S-Band Transponder*  
M. C. Comparini, F. Marchetti,  
Alenia Spazio s.p.a.,  
M. Siemon,  
Motorola SED.....127

### B-3: Advanced Design Methods

*Iterative Algorithms for the Design of Arbitrary Phase Finite Impulse Response Functions*  
S.M. Ritchie, C.A. Fulton, B.P. Abbott,  
University of Central Florida.....139

*Image Compression Reduces Bandwidth Requirements*  
Michael Ellis,  
U.S. Army Corps. of Engineers.....157

*High Temperature Superconductivity (HTS) and its Role in Electronics Warfare*  
Thomas N. Tuma,  
U.S.Army CECOM.....189

SESSIONS ON WEDNESDAY, SEPTEMBER 23

---

C-1: RF Systems I

---

- Choosing the Optimum Synthesizer Architecture for your Receiver Application*  
Larry Kimbrough,  
Harris GCSD.....203
- CW Rejection*  
Sherman R. Vincent,  
Raytheon Company.....N/A
- Optimizing Performance of Collocated UHF and VHF Ground to Air Radios*  
David S. Penney,  
MITRE Corporation.....N/A

---

C-2: Filter Design

---

- IF Transversal Filtering: Applications and Implementation*  
Richard D. Roberts,  
Harris Corporation.....217
- Catalog of Rhodes Filter Transfer Functions and Element Values*  
William B. Lurie,  
Consultant.....223
- Synthesis of Low-Pass Filters Containing Quads of Zeros*  
Mark Mell,  
Microsonics, Inc.....249

---

C-3: Oscillators and Synthesizers

---

- Modern Methods of Frequency Control*  
David J. Brunell.....259
- Self-Adaptive VCO*  
Dr. Wojciech Klimkiewicz,  
Theodore O. Grosh,  
SSDD Research Corporation.....271
- A CAD and Optimization Program for Microwave Oscillators*  
Axel Heitbrink, B. Roth, A. Beyer,  
Duisburg University.....281

---

D-1: RF Systems II

---

- A Simple Road Information Transmitter*  
Thomas Hack,  
Comlinear Corporation.....287
- Waveguide and Coax Components for High Power Broadcasting*  
Jim Stenberg,  
Passive Power Products.....N/A
- An Image Improvement of Microwave Diffraction Tomography for 2-Dimensional Inhomogeneous Dielectric Cylinder Based on Projection Function*  
K.W. Suh, D.Y. Lee, Y.C. Han,  
C.Y. Park, S.Y. Kim, J.W. Ra,  
SAMSUNG Electronics.....299

---

D-2: Crystal and SAW Filters

---

- Crystal Filters Having Superior Intermodulation Characteristics and the Measurement of Low Level Intermodulation in a Swept Frequency Mode*  
M. D. Howard, W. van den Akker,  
Piezo Technology, Inc.....317
- Extensions of the Holt & Gray Crystal Filter Design Technique*  
William B. Lurie,  
Consultant.....327
- Modular Implementation of the Coupling of Modes Analysis for Surface Acoustic Wave Resonators*  
S. M. Ritchie, R. P. Burke,  
University of Central Florida.....339

---

D-3: Components for New Applications

---

- A Versatile Mixed-Signal Cell-Based Array for Wireless Receiver ASICs*  
Ignatius Bezzam, Chuck Vinn,  
Raytheon-ASIC.....N/A
- Multiple Receiver Configurations Realizable With One Set of Standard Stanford Telecom ASICs*  
Herman A. Bustamante,  
Stanford Telecom.....353
- Pushing Low Quiescent Power Op-amps to Greater Than 55dBm 2-Tone Intercept, and an Automated Very Wide Dynamic Range Measurement Approach to Measure These Exceptionally Low Spurious Levels*  
Michael Steffes,  
Comlinear Corporation.....365



SESSIONS ON THURSDAY, SEPTEMBER 24

---

E-1: High Power Applications

---

*A 1KW RF Amplifier for UHF 470-860 MHz  
Television Applications*  
Mike Culling,  
Harris Allied.....385

*Design Considerations for Solid State  
High Power Amplifiers Used in Radar*  
Michael Hanczor, Mahesh Kumar,  
Paramax Systems Corporation.....393

*A Quasi-Complementary Class D Amplifier*  
Frederick Raab, Daniel Rupp,  
Green Mountain Radio Research Co.....403

---

E-2: Small-Signal Design

---

*DEVCALC — An RF Amplifier  
Design Aid*  
Donald Miller,  
MCC Panasonic.....413

*Noise Measurements on Microwave Devices  
Using a Scalar Analyzer*  
Stuart A. Fox,  
Giga-tronics, Inc.....433

*Predicting Noise Contributions in  
RF Circuits*  
Dan Pleasant,  
Hewlett Packard Company.....439

---

E-3: Test and Measurement

---

*What Ever Happened to the Q Meter?*  
Albert Helfrick,  
Embry-Riddle Aeronautical University.....445

*Selecting the Right Frequency Domain  
Measuring Instrument*  
Robert Witte,  
Hewlett Packard Company.....451

*Design and Optimization of Lumped-  
Element LCM Directional Couplers*  
Edmund Joe Tillo II, Ph.D.....467

The papers contained in these Proceedings represent the views of the the authors and/or their employers, and have been presented by them as original works. While an attempt has been made to assure accuracy, neither Cardiff Publishing Company nor *RF Design* magazine will be responsible for any copyright, trademark or patent infringement that may be contained in these papers.

# **PRACTICAL APPLICATIONS OF NON-LINEAR ANALYSIS, SIMULATION AND MEASUREMENT TECHNIQUES**

Joel Dunsmore  
Microwave Instrument Division  
1992



## **OBJECTIVE**

To examine specific, practical details in building and using vendor models in a printed circuit board amplifier design, and simulating power and harmonic distortion.



HP Design and Development

Suppliers of RF Surface Mount Amplifiers, such as Avantek, provide models and measured data for their devices. Using the linear models, the performance of the devices may be modified or extended beyond the normal usage. Careful modeling of surface mount components and Printed Circuit Board (PCB) layout can precisely predict circuit characteristics, such as gain and match. Further, with the nonlinear SPICE model for the device, power compression and harmonics distortion can be simulated, and the operating point and circuit may be changed to optimize power performance.

Vector Network Analyzers, such as the HP8753C, can be configured to make non-linear measurements. Power compression and output power, as well as second and third harmonic distortion can be measured and displayed in real time.

This paper details a broadband PCB design using the MSA-0900 amplifier. While normally limited at the low frequency to 20 MHz, this design extended the low end to 300 kHz, by adding a second feedback loop. Careful modeling and component selection ensured that the high-frequency response remained unchanged. The harmonics and power compression were modeled and measured, with remarkable agreement. Specific, practical details of building and using vendor models, setting up nonlinear simulations, and setting up and making the RF measurements with the HP 8753C Vector Network Analyzer will be stressed.

RF circuit design bridges the gap between low-frequency analog and digital circuits, and high-frequency microwave circuits. Components are often PC board mounted, like a lumped circuit, but distributed effects require microwave techniques to be applied.

## USING MANUFACTURE'S DATA

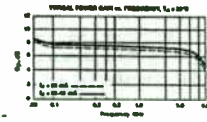
### Features

- Broadband, Minimum Ripple Cascodeable 50  $\Omega$  Gain Block
- 7.2  $\pm$  0.5 dB typical Gain Flattness from 0.1 to 5.0 GHz
- 2 dB Return Loss: 0.1 to 5.0 GHz
- 16.5 dBm typical  $P_{1dB}$  at 2.0 GHz
- Surface Mount Plastic Package
- Tape-and-Reel Packaging Option Available\*

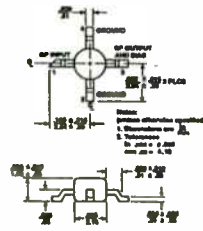
### Description

Avantek's MSA-0986 is a high performance silicon bipolar Monolithic Microwave Integrated Circuit (MMIC) housed in a low cost surface mount plastic package. This MMIC is designed for very wide bandwidth industrial and commercial applications that require flat gain and low VSWR.

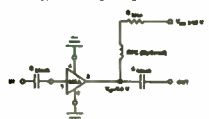
The MMIC MSA-0986 is fabricated using a 10 GHz  $f_t$  2S GHz  $f_{max}$  silicon bipolar MMIC process which utilizes an ultra self-alignment, low implantation and gold metallization to achieve excellent uniformity, performance and reliability. The use of an external bias resistor for temperature and current stability also allows bias flexibility.



### Avantek 58 Plastic Package



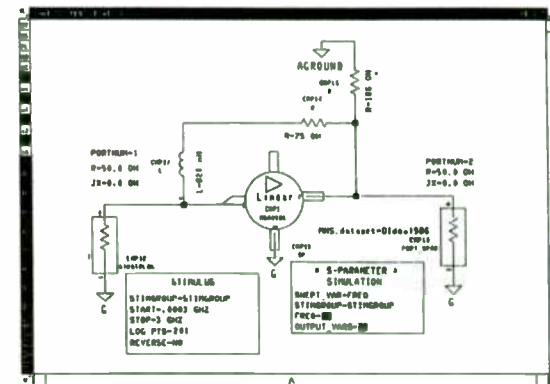
### Typical Biasing Configuration



HEWLETT  
PACKARD

1987 Document 098600010010

## DEVICE PERFORMANCE



HEWLETT  
PACKARD

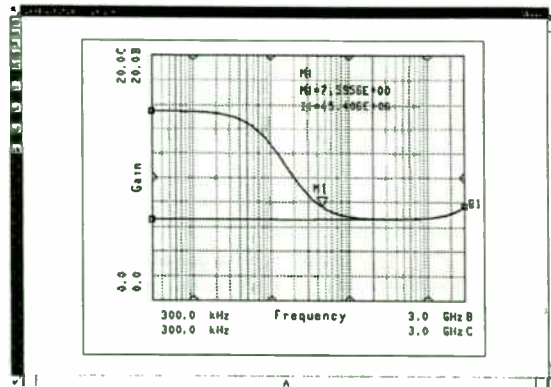
1987 Document 098600010010

The key to the amplifier circuit is the packaged silicon IC. Here is the data sheet for the Avantek MSA-0900 amplifier. It shows linear characteristics such as gain and input match, DC characteristics such as voltage versus current, and some nonlinear characteristics such as output power compression. Others, such as harmonics at some input or output power are not shown. Finally, the frequency range over which this device is to be used is much lower than the IC alone allows. The IC has an excess gain from DC to about 50 MHz, but flat gain down to 300 kHz is desired.

However, the databook does include extensive modeling information, including packaging effects, linear models and SPICE models. With the circuit simulator, designs for the low-frequency extension can be evaluated, as well as the design for harmonics versus output power. Throughout this paper the steps to use this information will be shown, as well as best practices for efficient entry and simulation. This design methodology is being used at HP, where the efforts of many engineers is leveraged over the entire lab.

A circuit using a model of the MSA-0986 amplifier, and ideal resistors and inductors was simulated as a starting point for the design. An initial simulation of the amplifier IC with only a bias resistor was performed. Then a simple feedback circuit was added to control the low-frequency gain, and the simulation was repeated. The results show that this topology holds promise for achieving the design goals for extending the low-frequency performance.

## IDEAL AMPLIFIER RESPONSE

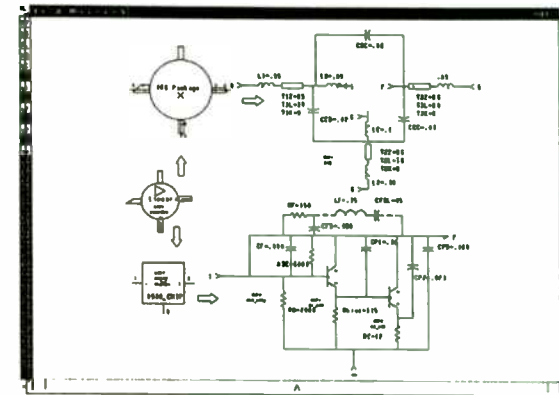


HEWLETT  
PACKARD

100 Dimensional Plot

The response of the amplifier IC shows a large gain increase below 20 MHz. From looking at the manufacturer's data book, it is clear that this is due to the limited frequency response of an internal feedback capacitor. When a low-frequency feedback path is added, with the inductor to remove the second path when the internal feedback takes over, the gain is controlled to the lower frequencies. To obtain even these simple responses of an ideal circuit, a model of the IC amplifier is needed. In the next several figures, techniques to utilize subcircuits to simply generate reusable and comprehensive models are presented. The packaged IC amplifier model will be used throughout this paper.

## COMPLETE PACKAGED IC

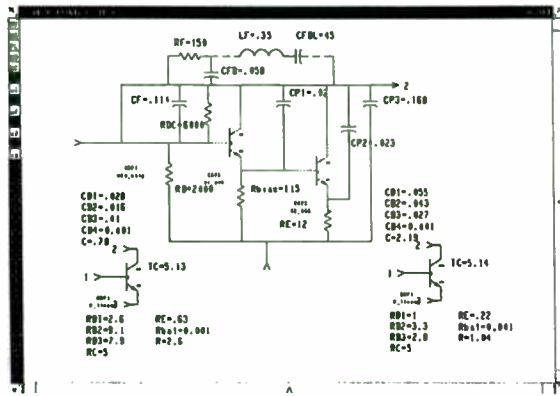


HEWLETT  
PACKARD

100 Dimensional Plot

This is the symbol created for the packaged IC. It can be used as any other simulator part, on a top-level circuit, or as part of a more complex circuit. The IC amplifier subcircuit consists of two lower level circuits. One is a general circuit for the 086 type plastic package. Any other transistor or IC device packaged in the 086 package can reuse this subcircuit. The second circuit is of the MSA-0900 IC chip. This chip can be purchased in a variety of packages, or in chip form. Making the IC subcircuit separate from the package allows it to be re-used in other applications.

## IC CHIP MODEL



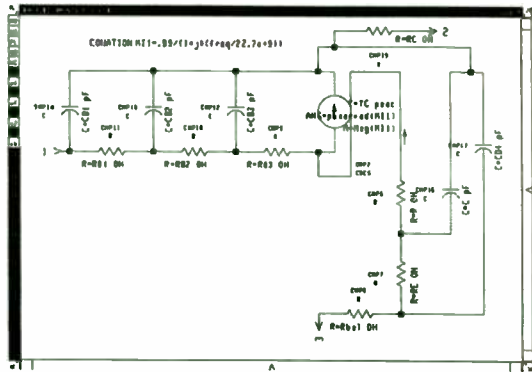
The IC consists of various thin-film resistors and some parasitic capacitors, with two transistors. One subcircuit for the IC circuit is made, connecting to the transistor subcircuits. This could be re-used for other Avantek IC types, by changing the values for resistors and capacitors, and adding the appropriate transistor model. The two transistors for the MSA-0900 amplifier will be added, making the IC complete. The transistor circuit for the linear device model are shown below the IC circuit. Q1 and Q2 use the same circuit model, with scaling factors and resistance values passed down into the model.

## MANUFACTURER'S DATABOOK

The table is titled 'Device Modeling' and contains multiple columns of data. The columns represent different parameters or models, and the rows represent different data points or configurations. The data is organized in a grid format, with numerical values for each cell. The table is used to provide detailed specifications for the device being modeled.

The Avantek databook includes all the information to produce the linear and nonlinear models.

## BUILDING THE DEVICE MODELS: LINEAR MODELS

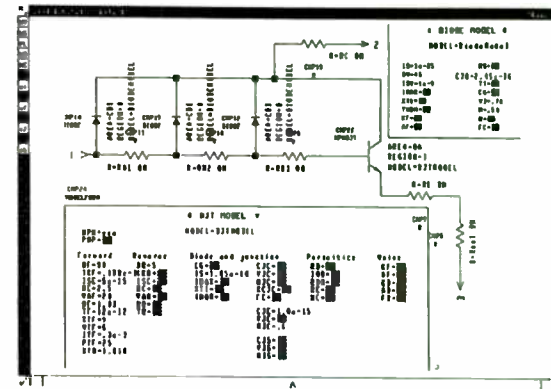


HEWLETT  
PACKARD

And Simulation through Psp

In the MSA-0900 amplifier, the linear model is shown for the transistors; there are two transistors of different sizes in the IC. Only one model is made, and the scaling factors are passed-in for the capacitances and resistance depending on the transistor. This model works well for linear simulations, but cannot be used for DC or harmonic generation simulations. For nonlinear analysis, it is only necessary to replace the linear transistor model with a nonlinear model developed next.

## NONLINEAR MODELS

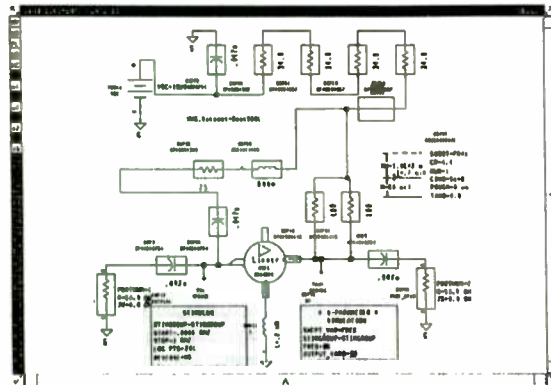


HEWLETT  
PACKARD

And Simulation through Psp

The nonlinear model for the transistor includes an ideal SPICE model transistor, surrounded by other elements such as diodes and resistors. The diodes are used to model the nonlinear capacitance effects on the device.

## COMPLETE AMPLIFIER DESIGN

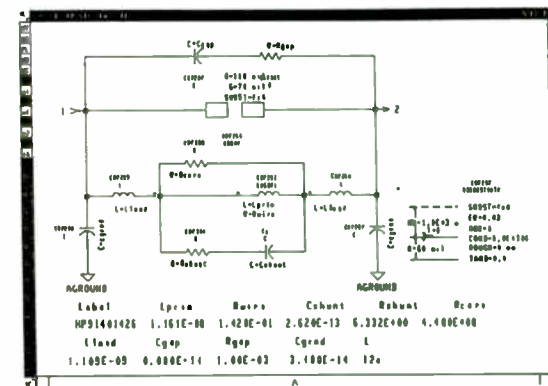


HEWLETT  
PACKARD

Just Document/Workshop Plot

A complete amplifier was designed using surface mount part models, and the IC part model. A secondary feedback path was added to control the low-frequency gain. The inductor was chosen to remove the low-frequency feedback path when the internal feedback becomes significant. The simulation used the linear model to determine the small signal response of the circuit. This design is much more complete than the first ideal circuit, including blocking capacitors on the input, output and feedback paths, and resistors chosen to provide DC bias with sufficient power handling capacity. The parasitics of each of these parts may greatly affect the circuit.

## OTHER PARTS, SURFACE MOUNT MODELS



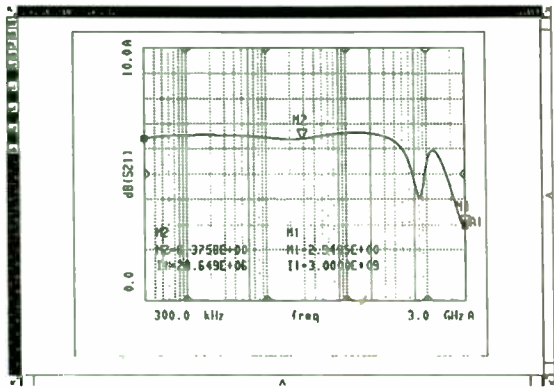
HEWLETT  
PACKARD

Just Document/Workshop Plot

The models of the surface mount parts are key to the successful simulation of the circuit. Models for a wide range of surface mount parts were developed, using network analyzer data to match the characteristics to lumped circuit models. The parts are used by specifying the part number, which is linked to the particular values needed in the lumped model. This is an example of the circuit used to model an SMT inductor.



## AMPLIFIER SIMULATION RESPONSE

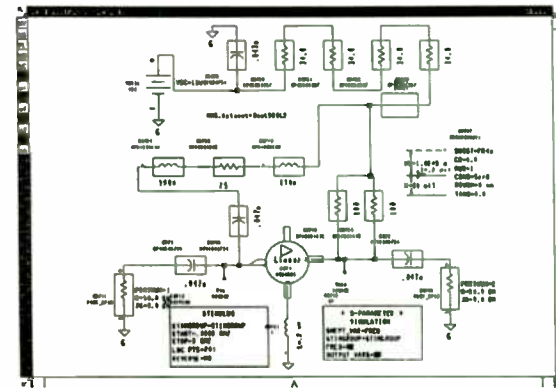


HEWLETT  
PACKARD

104 00000000000000000000

The simulator response that the external feedback path does limit the low-frequency gain. However, the high-frequency response is dramatically affected. From looking at the self-resonant frequency of the 680 nH feedback inductor, it becomes apparent that the parasitic capacitance will limit the high-frequency response, and cause a gain dip between 1 and 2 GHz.

## IMPROVE EXTERNAL FEEDBACK

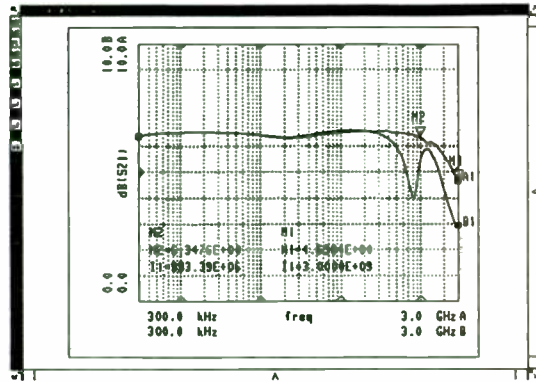


HEWLETT  
PACKARD

104 00000000000000000000

Since smaller inductors have much higher self-resonant frequencies, the 680 nH inductor was divided into two smaller inductors. Sometimes, adding inductors in series, or capacitors in parallel, will cause additional resonances. Often, any resistance between such elements will reduce this effect. The simulator should show us if any resonances might occur. The parasitics of the SMT parts are not particularly well controlled. Parasitic element values from one inductor to another can vary by 50% or more. However, it is important to include more than the simple first order model with only a shunt C, as the higher-frequency resonances will show some effect, and the circuit sensitivity to this will be discovered. An example of this is a low-pass filter. If a simple model is chosen for L and C elements, it may appear that the stop-band is well behaved. When more complex models are used, stop-band sensitivity to multiple resonances becomes clear, and steps may be taken to account for the behavior.

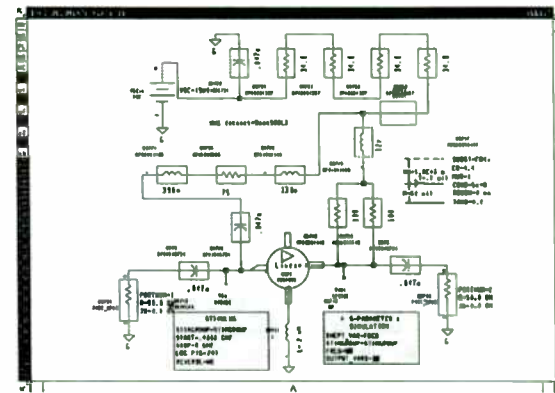
## IMPROVE FEEDBACK RESPONSE



HEWLETT  
PACKARD

Here it is seen that the low-frequency response is about the same, but the high-frequency gain dip has been removed. However, there is still a significant roll-off of the high-frequency response. Perhaps another, smaller inductor between the low-frequency feedback and the output line will reduce the gain slope.

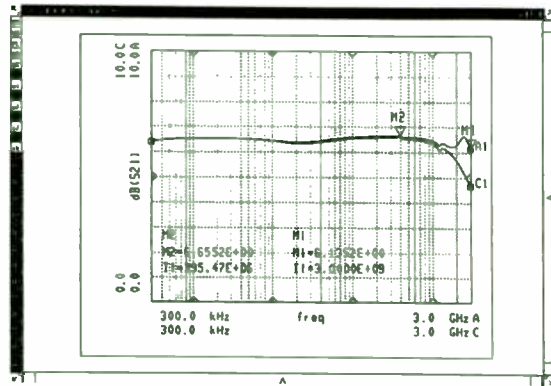
## RF CHOKE



HEWLETT  
PACKARD

Another RF choke is added to further isolate the low-frequency path. The 12 nH inductor was chosen, as it has a high impedance at frequencies above 1 GHz, where the gain slopes.

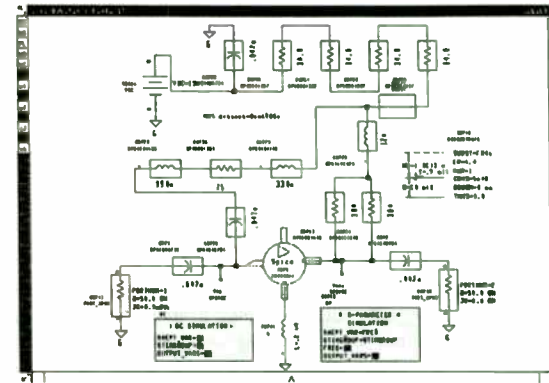
## FINAL LINEAR RESPONSE



HEWLETT  
PACKARD

Finally, it is seen that the response is as desired, with nearly flat gain over the full band. Next, it will be shown what happens when the nonlinear model is applied.

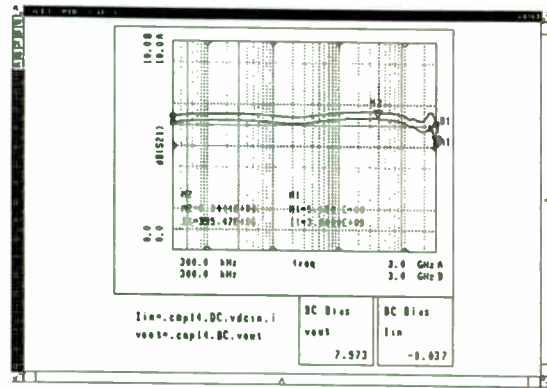
## NONLINEAR AMPLIFIER SIMULATION



HEWLETT  
PACKARD

Here, the nonlinear IC model was substituted for the linear model, and a DC as well as S-Parameter simulation was performed. Note that the simulation options block was added to set the temperature for the IC. Initial runs gave incorrect results for the bias point of the circuit. After estimating the temperature rise due to self heating in the IC, the correct bias point was obtained. The small signal response will be compared between linear and nonlinear models.

## NONLINEAR SMALL SIGNAL RESPONSE

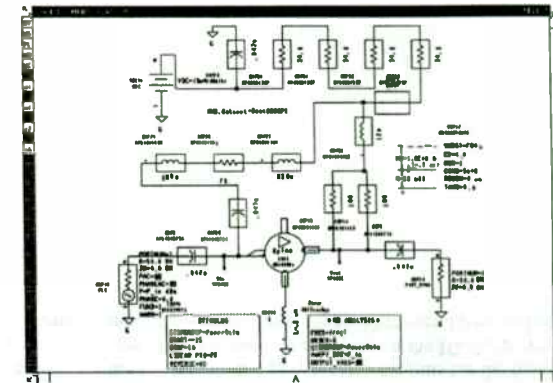


HEWLETT  
PACKARD

Just Don't Mess With My Pie

The small signal response for the nonlinear circuit differs slightly at high frequencies from the linear response. Experience shows that the linear model is often better at predicting the small signal response, as it is optimized only for the small signal case. The nonlinear model may compromise the small signal response to more accurately predict large signal performance. Here, which response is used to design to, is up to the experience of the designer to determine. The DC voltages for the operating point agree very well with the actual measured values.

## MODELING POWER AND HARMONICS

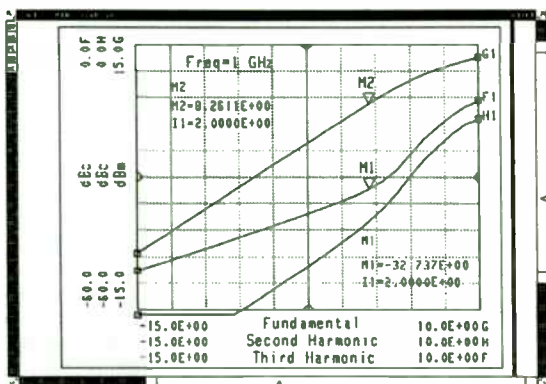


HEWLETT  
PACKARD

Just Don't Mess With My Pie

Trying to simulate the response of an amplifier to large signal power sweeps takes full advantage of the nonlinear simulator. The nonlinear analysis block has been changed to a One-Tone Harmonic Balance with a swept variable (HBSwpt). The port 1 symbol has been replaced with a Large Signal Power Port (PLS). A stimulus block has been added to control the variable, P\_in. An equation is necessary to define P\_in. With this analysis, the fundamental and 5 harmonics are analyzed for each node in the circuit, at each power level.

## SWEPT POWER RESPONSE

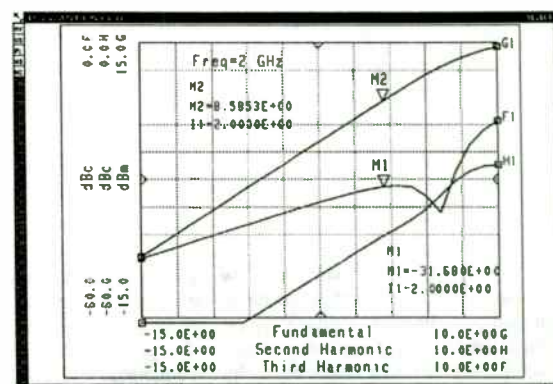


HEWLETT  
PACKARD

Just Enter-Add-Subtract Plot

This presentation shows the result of the analysis, with the output power in dBm, and harmonics in dBc, shown versus the input power. It is clear that the power compresses, and will saturate below +15 dBm. This analysis was done at 1 GHz, and it is interesting to note the behavior of the second harmonic. Normally, a distortion signal will rise with input power as a function of its order, that is, a second harmonic will increase at two times the increase in input power, and a third harmonic will increase at three times that trend until the amplifier nears compression.

## SWEPT POWER RESPONSE, 2 GHZ

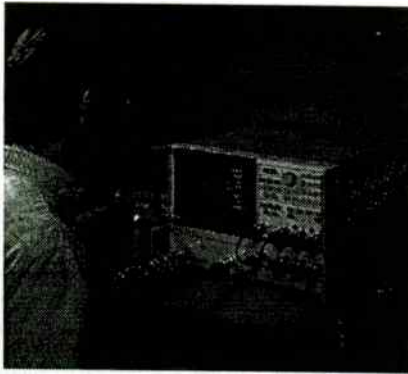


HEWLETT  
PACKARD

Just Enter-Add-Subtract Plot

The simulation was repeated at a fundamental frequency of 2 GHz. Notice now that the second harmonic now has a dip in its response. This may be due to the feedback in the amplifier, with the second harmonic cancelling itself at this frequency. Now would be a good time to try to compare the simulated results with an actual measurement.

## MEASURING NONLINEAR RESPONSES

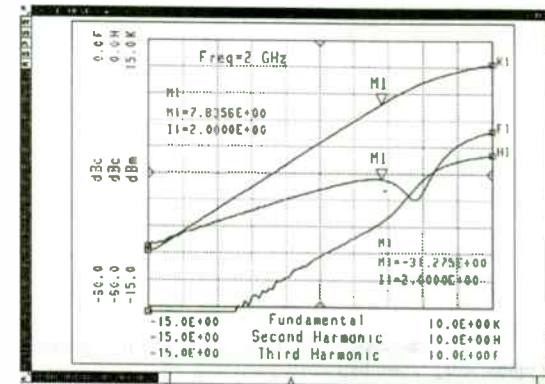


HEWLETT  
PACKARD

HP 8753C

With an RF Network Analyzer, such as the HP 8753C, the power may be swept while measuring gain, or output power. In addition, with option 002, harmonics of the output may be measured directly. The network analyzer is calibrated for power using a power meter, removing the effects of cables, fixtures and pads. The amplifier is added, and the power compression, output power, and harmonic distortion for second and third harmonics can be shown. Further, the network analyzer data may be passed directly to the MDS simulator for comparison with earlier calculated values.

## MEASURED RESULT

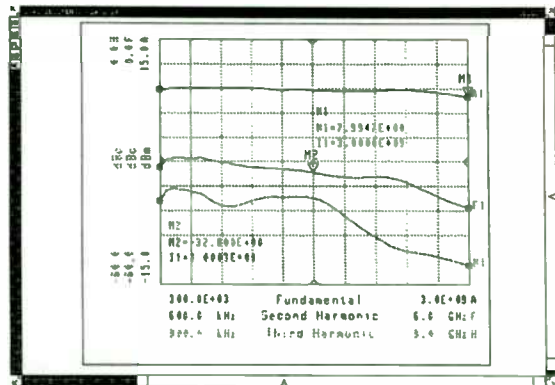


HEWLETT  
PACKARD

HP 8753C

Here is shown the results on an actual bread-board circuit using an MSA amplifier, with SMT resistors and inductors in the feedback path. The results agree remarkably well with the simulation, correctly predicting the dip in the second harmonic for a 2 GHz power sweep. The output power and compression point are within 1 dB of the simulated data. During the measurement, bias may be adjusted while sweeping harmonics real-time to determine bias effects. In fact, this method is used to tune and improve harmonic distortion of GaAs FET amplifiers used in the HP 8753 Network Analyzer.

## SWEPT-FREQUENCY HARMONICS



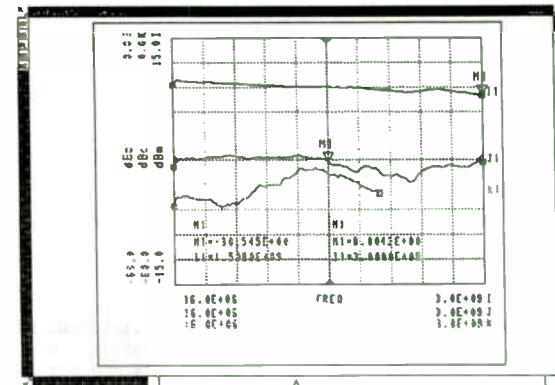
HEWLETT  
PACKARD

And Document: dshdhp010

From the power sweeps, we saw that harmonic behavior and compression varied with frequency. By replacing the input power swept stimulus with a frequency swept stimulus, and performing the harmonic balance analysis at each frequency, a swept harmonic simulation can be performed.

Here are the results of the harmonics versus frequency. Also shown is the output power, all with a 3 dBm input signal. The second and third harmonics do vary several dB over frequency, with the third harmonic getting worse at higher frequencies, up to about 1.5 GHz input. Above this frequency, the third harmonic is probably out of the gain region of the amplifier. This can now be compared to the response of a real amplifier.

## SWEPT-FREQUENCY HARMONIC MEASUREMENTS

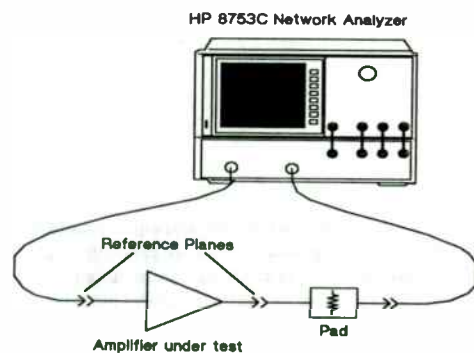


HEWLETT  
PACKARD

And Document: dshdhp010

The HP 8753C, opt 002 was used to measure the swept-frequency response of the output power and harmonics. The data was read into the HP Microwave Design System, and plotted on the same scale as the earlier simulation. The output power tracks the simulation remarkably well. The second harmonic does not show quite the roll-off of the model, but the third harmonic shows a very similar response to the simulation. The third harmonic response is limited in frequency to 6 GHz (2 GHz fundamental), due to the frequency limit of the HP 8753C.

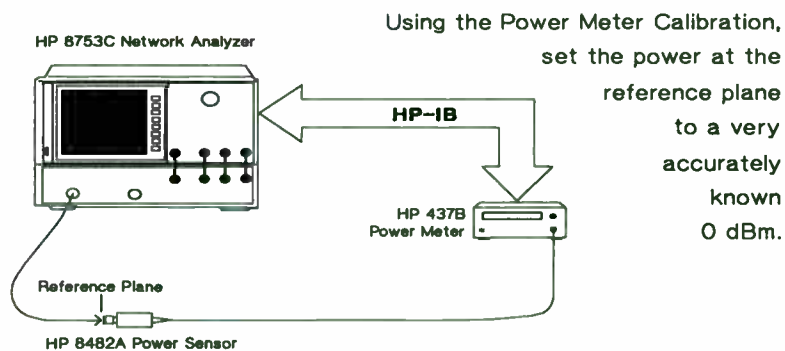
### Amplifier – Swept Power and Harmonic Measurement



The HP 8753C network analyzer may be configured to measure output power and harmonics directly, versus frequency or drive power. The harmonic measurements shown earlier were made using the equipment and methods shown next.

The calibration techniques used are an important part of the harmonic or power measurement. The receiver must be calibrated in terms of absolute power. The source will be used as a transfer standard, transferring the accuracy of a power meter to the network analyzer receiver.

### Calibrating the Source



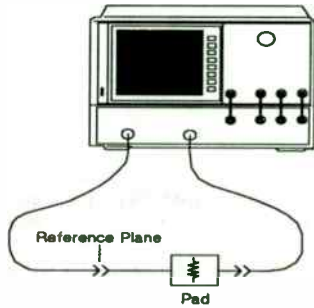
Using the Power Meter Calibration, set the power at the reference plane to a very accurately known 0 dBm.

The diagram shows the connection to perform a source power meter calibration. The power meter is controlled over HP-IB by the HP 8753. The calibration level is set to 0 dBm. The number of readings is set at two to get better accuracy. Some padding of the source can be done to improve source match, especially if lower power is normally needed. The total loss from the source output to the power meter must be less than 20 dB to get 0 dBm out. The frequency should be set to the desired frequency range and number of points. However, less points could be used during power meter cal to speed up this step. The points may be increased after power meter cal, with intermediate values interpolated from existing cal data.



### Calibrating the Receiver

HP 8753C Network Analyzer

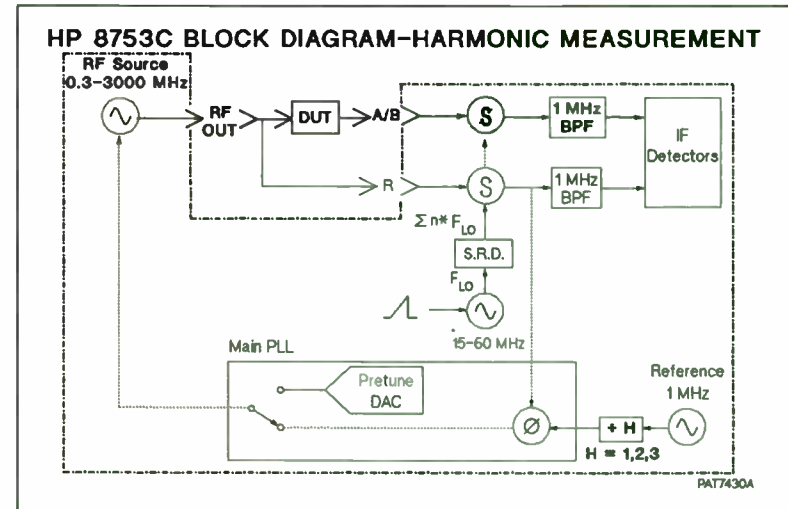


Using the calibrated 0 dBm source power, perform a response calibration on the desired receiver channel.

The receiver is now calibrated to measure absolute power (0 dBm will correspond to 0 dB on the display).

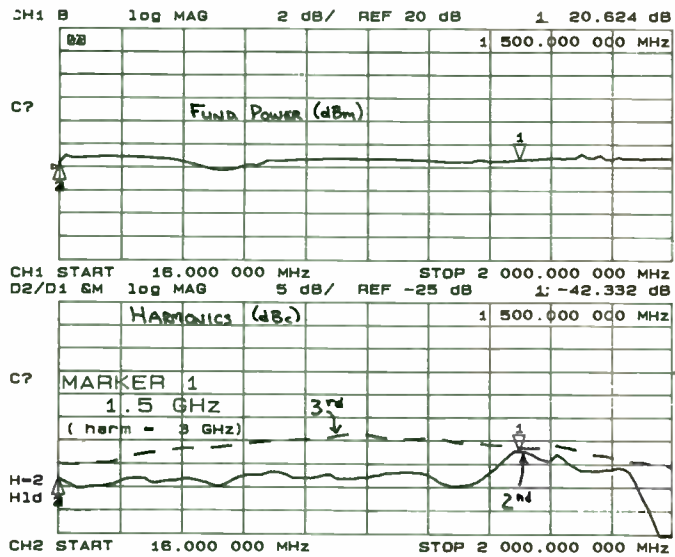
The second step is to connect the now calibrated source to the receiver input to be used to measure power. A test-set port may be used, or the input can be sent directly to the HP 8753 receiver channel. If a cable or adapters are needed they should be included now. Padding before a cable will reduce errors due to mismatch, with no loss in accuracy if added before the receiver calibration. With the source connected to the receiver, a response cal is performed. After the calibration step is complete, the receiver is calibrated in absolute power. Note that the power meter is not needed at this step. Also the number of points in this step may be different from the previous step.

### Harmonic Mode Block Diagram



The HP 8753 can make swept harmonic measurements, where the level of the second or third harmonics of the DUT is displayed versus frequency. For this measurement, the main PLL reference frequency is divided by the harmonic number; 500 kHz for second harmonic and 330 kHz for third harmonic measurements. The RF source fundamental is converted to these reference frequencies, and is rejected by the IF filter. The second or third harmonic is converted to 1 MHz, and is measured by the IF detectors.

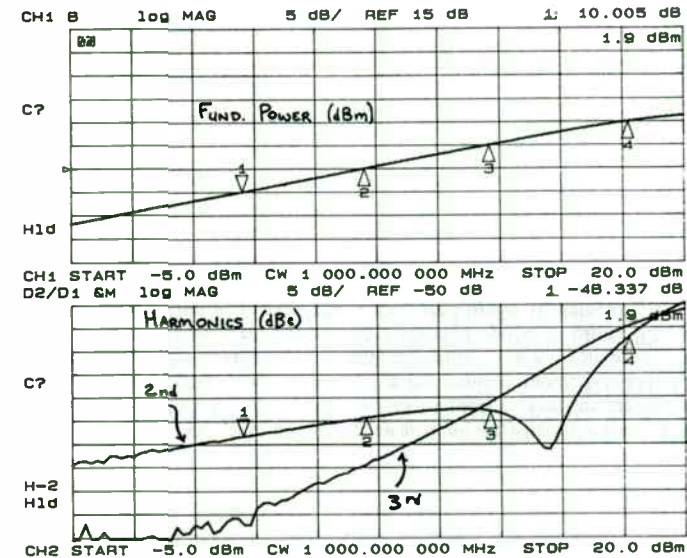
### Swept Frequency Measurements



The frequency response of harmonic distortion can be displayed directly. The dual channel display mode is used, with the channel coupling off. Channel one displays fundamental power, and channel two is in the harmonic measurement mode.

The DATA2/DATA1 display in channel two displays harmonics in dBc, relative to the fundamental. The marker displays harmonic level, reading in both fundamental and harmonic frequency.

### Swept Power Harmonic Measurements



The HP 8753 can measure power and harmonics while sweeping input power. Measurements of this type can show unusual behavior in harmonic distortion. For example, this display shows second harmonic power does not follow the uniform trend of increasing with increasing output power. One possible explanation is that the distortion is becoming symmetrical, thus reducing even order harmonics.

### Measurement Considerations

- **Keep Source Power Low**
  - +20 dBm --> 25 dBc
  - +10 dBm --> 40 dBc
  - 0 dBm --> 50 dBc
- **Keep Receiver Power Low**
  - 0 dBm --> 30 dBc
  - -20 dBm --> 50 dBc
- **Low Pass Filter Input**
- **High Pass Filter or Pad Output**
- **Use Pads to Improve Filter Match**

### Summary

- **Non-Linear Effects are Increasing Important**
- **HP 85150 CAE Tools May Be Used to Predict Non-Linear Behavior**
- **HP 8753 Network Analyzer can Directly Measure Non-Linear Behavior**

The limitations in measuring harmonic distortion come from the harmonic content of the source, and the distortion caused in the receiver. The chart shows typical values of distortion in the HP 8753.

Source harmonics in a limited band may be reduced by filtering the input signal. The receiver distortion is caused by a large fundamental signal causing harmonics in the receiver. Using a high pass filter to remove the fundamental, while passing the harmonics, can improve the receiver harmonics without affecting receiver sensitivity. Padding the receiver is a broadband way to reduce distortion. The maximum range of the receiver without filtering is about 65 dBc harmonics for second, and 75 dBc for the third.



# Practical Object Oriented CAE for RF Engineers

Dale A. Teets, P.E.  
Innovation First  
PO Box 221  
Mason, OH 45040

## Abstract

Object oriented techniques simplify the creation of powerful custom CAE in Microsoft Windows with or without C++ programming. To take advantage of this technology the engineer should understand the basics of object oriented CAE (OOCAE), and the basic components of an OOCAE framework. A fully working framework allows the engineer to concentrate on the mathematics, producing flexible custom applications in less time than is possible with BASIC or other programming techniques.

## Concept of Object Oriented Programming (OOP)

The central idea in object oriented programs is the object. An object is simply a collection of data and code that performs a particular type of task. The object template, called a class, is defined by the programmer. This class definition includes the various types of data (integers, arrays, other objects, etc.) necessary to perform the task, and the functions (methods) to define how the task is accomplished. The data and functions are called members of the class. A PLL object might have data members such as damping factor and VCO gain along with member functions to calculate loop responses. The programmer can create and use multiple instances of the objects, much

like creating and using integers or any other predefined type. Essentially the programmer is extending the C++ language with his/her own types.

But the power of OOP goes far beyond that of being able to declare and use new data types. A very powerful feature is the ability to define a hierarchy of classes where each derived class inherits characteristics of its base class. This concept of inheritance is very common in real life. Consider the analogy of the engineering profession with engineering as a base class. Mechanical, electrical, and civil engineering classes are derived from the engineering class. Each derived class has the characteristics of the base class. From the electrical class we derive those engineers who specialize in analog circuits, digital circuits, and power systems. From the analog class comes RF engineers and from there many specialties. From an OOP point of view we can have an instance of an RF engineer that inherits characteristics of the base classes analog, electrical, and engineering. A new class can be derived from RF engineering which has the knowledge of all the base classes, without rewriting the education and training from scratch. This is an important time saving feature, as entirely new classes can be created without even slightly altering the base class, and the base class is still available.

Another feature of OOP is the ability to interact with all objects derived from a given base class as though they were objects of that base class ( i.e. polymorphic behavior). From the above analogy, an engineer could be asked to perform basic mathematic analysis without knowing what type of engineer is doing it. This is especially useful in RF analysis where we might have many types of analysis but each object can be accessed as though it were the base analysis class. Thus we can tell an analysis object to analyze without knowledge of what it actually does. This feature is essential in development of a generic interface. Once developed, the interface can be the same for all custom programs written. It simply treats all objects as though they

were of the base types, handling such things as analyzing, tuning, optimizing, creating, displaying, saving, and restoring of objects. This feature reduces development time, and provides flexibility not possible with traditional programming methods.

Object oriented programming is ideal for writing custom programs due to the concept of reusability. The base classes are set up to handle the common tasks and are reused in all custom programs written. Classes are derived to handle tasks that are unique for a given application. This dramatically reduces development time and improves reliability since the base classes are not altered and remain bug-free. The object oriented approach requires a bit of forethought about how one might want to extend the classes in the future. This is a long term attitude that is more appreciated tomorrow.

#### **OOCAE Overview**

Strictly speaking, any engineering program written in an object oriented language could qualify as OOCAE. But the true power of the object oriented structure should flow through to the user; it should not just benefit the author of the program. An object oriented approach to custom engineering programs simplifies and enhances three major areas: 1) reusable event oriented interfaces, 2) manipulation of internal objects, 3) creation of a reusable framework.

Most engineers use predefined sequences in their programs that lead the user from beginning to end with options to branch. With OOP it is just as easy to respond to the user in an event oriented manner. The program responds to events from the user by executing certain routines or creating/destroying certain objects. Once the program starts it simply waits for the user to act, then passes the information to the object with the focus. An event oriented interface is easier to use and has far more potential for reusability. Since the largest portion of any program is the user interface, a large

amount of effort can be saved by reusing interface objects. This allows the engineer to concentrate on the actual math in the CAE rather than the interface.

Since objects can be accessed through a common base class without knowing what object is being manipulated, it is possible to allow the user to manipulate the internal objects via a generic interface. All engineering analysis processes have certain attributes in common. These attributes can be encapsulated in base classes with more specific objects derived from them. All the derived objects can then be controlled as though they were the base classes. Thus the user can specify which objects are to be used and their relationship. A project of objects can be dynamically created at run time, without programming. When an engineer encounters a formula in *RF Design*, all that is necessary is to type in the formulas and link them to graph objects, numeric list objects, or other objects. Using tune, analyze, and optimize features, an application is created with all the features of a professional package but appears to have been written for the specific task at hand. The entire process is much faster than writing a "quickie" BASIC program, and the result is more powerful.

For those who don't break out in hives at the thought of programming, a little C++ programming can be turned into numerous future custom programs. The basic idea is to set up (or buy) an engineering framework that can be extended in various ways in the future. Frameworks are basically collections of classes that the programmer routinely uses to accomplish basic tasks. With base classes defined in the framework, it is easy to extend those classes into more specific tasks. Taking this a step farther, it is possible to create a working executable framework for basic engineering analysis that needs no programming. When the user is allowed to manipulate the internal objects, each new class added in C++ to the framework dramatically increases the possibilities of new custom applications. As the framework grows over time, the necessity of programming decreases and the time spent creating new applications

shortens. Even if the author of the program is the only one using the program, this approach pays for itself in the long run.

### **Basic Components of an OOCAE Framework**

The construction of an OOCAE framework is essentially a task of defining classes and their relationships. It is desirable to derive from common base classes whenever possible. It is also a good idea to keep the classes as generic as possible, since it is much faster to derive than create a class from scratch. A discussion of the more important components of the Object Engineering™ framework illustrates some of the classes desired for an engineering framework.

#### **Complex Class**

The complex class defines how complex numbers are manipulated in addition, multiplication, division, trigonometric functions, and other operations. With these methods defined by a complex class, such operations as  $\tan(z)$  are as simple as though  $z$  were a real number, the compiler takes care of the details of the actual operation using the methods defined in the complex class. To use a complex object one simply declares complex  $z$ , and then uses the variable  $z$  much as one would use the built-in types integer or double floating point variables. The data members of the complex class are the obvious real and imaginary parts ( $re$  and  $im$ ). Figure 1 shows how a simple complex class could be constructed. This class is typically one of the largest classes that would be used in an engineering program. Some compilers supply a complex class ready for use.

#### **Process Classes**

The concept of an engineering process is encapsulated in a base class that we will call "Process." The Process class is generic enough that a derived class could perform most anything. It provides virtual functions for updating, executing, or preparing

to execute. Derived classes provide these functions to carry out their own specific tasks. Thus all processes can be told to execute using the process class interface without concern for how a particular process is carried out. The user could create a button to collect data from an instrument via HPIB, or calculate the optimum RC values for an active loop filter. These processes all look the same to the interface code and are accessed through the base process class. Processes also contain a ProcessData object capable of storing variables and data in a uniform manner. Each process can be linked to another process to share the ProcessData object. This essentially allows a hierarchy of processes to be formed by the user at run-time. Each process can have one parent process and many child processes. The purposes of these links is to put the object oriented power into the hands of the user. The objects may be linked in ways that the programmer could not foresee. The top process of the hierarchy actually owns the ProcessData object; the linked processes have access. Any process can of course have its own local ProcessData. The links also allow the parent to update all linked processes that want to be automatically updated. Thus if the user analyzes, tunes, or optimizes a file of formulas, all linked processes such as graphs and numeric lists can be updated without user commands. The linking, updating, and the tracking of valid objects is contained in the Process class; derived classes do not need to do anything to take advantage of these characteristics. The only user action required is to type in the names of objects, the rest is transparent to the user.

A natural extension to the Process class and perhaps the most important is the equation parser. It is the essential object in any engineering program that interprets math and commands that are typed in as text by the user. This object is most useful for performing processes that can not be predicted when the program is written. The text is parsed in a recursive decent parser (Figure 2) that assumes complex values. The parser takes advantage of the C language property of function recursion to implement

common algebraic rules and operator precedence. It supports the normal operators, relational operators, trigonometric functions, *if/for/do/while/goto*, and others, all assuming complex math. Commands can be defined to perform any task, and can be extended to provide a unique language made up by the programmer. The equation parser is also useful for giving the user access to the program's internal functions. Since the parser is derived from the process class, formulas typed in by the user can be analyzed, tuned, and optimized without a single line of programming. Thus a user can type in formulas found in an *RF Design* magazine and link to graphs, numeric lists, or other objects, and be done with the analysis faster than one could write any type of program. Yet the project is available for future use and can be shared more easily due to the Windows interface.

To enable fast efficient tuning of variables belonging to multiple processes, a *VariableTune* class is defined. It allows the user to change variable values in the *ProcessData* associated with a process and observe the results from all linked processes. Having tune and analyze features is sufficient for most purposes, but an *Optimizer* class adds a nice touch to custom CAE. A generic optimizer is less efficient than the sophisticated ones found in multi-thousand dollar CAE, but requires little extra work in OOCAE. The *Optimizer* object changes selected *ProcessData* variables until a goal is obtained. Linked processes that want to be updated are updated as the optimization proceeds. Thanks to the polymorphism of the process classes, anything can be optimized that changes variable values in the process. This also extends to formulas typed in by the user or even a test setup that alters a device under test and measures results.

#### User Interface Classes

The user interface classes eliminate the need to rewrite the user interface for each custom application. These classes take advantage of the many built-in functions

of Microsoft Windows. Windows has built-in editors, buttons, check boxes, and other controls that can be used to create powerful components of an RF analysis/synthesis program. Windows' Multiple Document Interface (MDI) is the appropriate choice for engineering applications. There are usually several windows available simultaneously, and MDI handles most of the details with very little programming. Besides the look and ease-of-use, the Windows interface has a very important advantage for custom CAE: the device independence concept that can save an engineer months of work. There is no need to learn low level details of EGA, VGA, SVGA, dot matrix printers, laser printers, HPGL plotters, or mouse drivers. It also enables other engineers with different computer setups to use the programs without concern for particular hardware requirements.

The key to the user interface is the base *Windo* class. It provides an easy way to derive unique window objects, as it handles tasks common to all windows. It handles such things as reporting that it is okay to close the window, giving the user a chance to save the contents if changed, menu changes, and handling messages from Windows that are not intercepted. An object derived from the *Windo* class can concentrate on its specific task without concern for the basic functions.

From the *Windo* class comes the *EditWindo* class with the extra capability of a text editor. It has cut, paste, and copy functions expected from an editor and serves as a base class for two of the most important windows: the *TextParseWindo* and the *ListWindo*. Each of these classes benefit from both *Windo* and *EditWindo* features. The *TextParseWindo* provides one method by which the user can enter formulas and commands that are executed by the *Parser* class. Each object gets its own parser object so it is possible to parse multiple files or execute one parser object from within another. Since the editor class is available it seems natural to use it as the display media for numeric lists. This allows the user to insert comments, cut, paste, and copy in the



listing. The formatting can be a bit tricky with proportionally spaced fonts using pixel spaced tabs, but there are fixed pitch fonts available to ease the task. ListWindo formats the variables specified by the user without special print statements usually required. In fact, no output statements are required if the list is allowed to control the process. The ListWindo knows how to get the values and format them. This means that the user does not have to memorize commands to get listings or graphs. This may sound like a lot of work for the programmer, but C++ and Windows makes it easier in the long run due to reusability.

The GraphWindo class is the base class for all types of graphs. This base class performs the functions necessary for graphing data with the exception of the actual plotting of grid and data. Each derived class knows how to plot its data and draw its grid. Thus to create an entirely new graph type, a class is derived from GraphWindo that contains the actual drawing commands. The drawing routine is made much simpler when the drawing is mapped to a common surface area. This works well for printers, plotters and screens by letting the Windows environment do the work of mapping to the actual device. The actual drawing commands need not be concerned with the output device. The graphics interface provides a generic device context to which the same commands can draw, whether it is a plotter, laser printer, SVGA screen, or whatever. To place the plot or to fill a partial page on a plotter/printer, the user can simply change the origin and extent of the mapping. The milliseconds lost in execution speed due to mapping is an excellent tradeoff for the time saved in programming.

The built-in controls in Windows can be used as they are or extended to handle specific engineering tasks. One obvious extension to the edit control is the VariableEntry class. This is a one line editor with all the features of an editor, but with the ability to set a complex variable to the value entered by the user. A related class called VariableWatch defines an object that simply displays the current value of a

variable. It uses the Windows static control and updates itself when a linked process changes the variable's value. Many other objects are possible by deriving classes from the built-in controls and the Process class. Buttons can perform user defined or programmer defined processes. The possibilities are unlimited. Listed below are some of the more useful controls provided by Windows that are easily extended or used directly.

<u>Windows Controls</u>	<u>Types/Uses</u>
Editor	Multi or single line editors with automatic scroll
List Box	Lists of strings or numbers
Buttons	Check, radio, owner draw, text
Static	Just displays something

These control objects are great for the programmer, but the real power results when the user can put them together to form unique windows. To facilitate this, a UserWindo class is defined to allow the user to place objects in a user defined window. The UserWindo can hold various controls selected and placed by the user. A separate object called a SelectBox allows the user to move and resize objects in the window. Processes for each control can be defined in text and executed by the parser object or hardcoded in C++. Naturally, it is faster to visually design a window than to create one in C++ code, an option even the programmer will enjoy when time is limited.

#### Project Manager Class

The project manager's primary function is to save and restore the objects and the configuration that is created by the user. The routine of saving and restoring the project is actually a simple matter for the project manager. This is especially true since each object knows how to save and restore itself and Windows maintains a list of windows. Each top level window is requested to save enough information to restore itself and its child windows. This information can be as brief as a name and the type of

the object, or it can be quite sophisticated. Since it is unknown what types of objects might be programmed in the future, the structure is as generic as possible. The object has the option of inserting the information directly into the project file or referencing a file name. This assumes that the object knows what it saved and how to use it in reconstruction. Again, the options are infinite because we do not know how we might extend the program, and we don't want to rewrite the project file I/O each time we write a custom program. The UserWindo also uses this file I/O concept in saving and restoring its unknown child objects. From the users point of view, the only interaction with the project manager is in the open, close, and save project commands from the menu.

#### **Putting it all Together**

There is no master object to pull everything together as one might expect. The objects are purposely designed to be as independent as possible. The program starts by creating a project manager object and the Windows MDI frame. From that point forward the user dictates what is created and how the objects are used. Thus the program consumes only enough memory to hold the created objects, rather than grabbing memory that may never be used. Figure 3 illustrates what the program might look like internally when the user creates a few objects.

Windows translates user actions such as mouse clicks and keyboard entry and passes messages about these events to the application. The interface to Windows consists of a few C style functions (as required by Windows) to receive and route these messages. Each object takes care of itself for the most part, responding to these messages from the interface. Interface messages are passed through the base Windo interface object. As each window object is created, Windows adds them to its internal window lists. Thus, to access the windows we can ask Windows for information or pass

messages. It is redundant to track Windo objects since Windows keeps all the information we need. However, we do need a way to attach our Windo objects to the window structure used by Windows. This is accomplished in a number of ways by commercially available frameworks. The most efficient way is to tell Windows to keep a pointer to the object in the window structure. Thus messages from Windows are routed to the proper object by retrieving that pointer from the window. Most of the messages are handled in the base classes or by Windows. As we move down the hierarchy, there are fewer interface tasks to handle, allowing the programming engineer to concentrate on the mathematics involved.

The processes must be tracked by the program, since Windows knows nothing of processes. Each process informs the master process list when it is created and destroyed. This list can be used to get information on any valid process in the session. This list is also presented to the user to establish a link to another process. When a process is destroyed, it informs the parent and child links to insure that an invalid process is not accessed. When a process is updated, it searches through its child process list for those processes that want to be updated, and tells them to update. Process derived classes need not do anything to take advantage of these methods.

A process object could be embedded in a Windo object, but it makes more sense to use a feature of object oriented programming called multiple inheritance. Since some objects are both Windos and Processes, it makes sense to derive them from both classes (see Figure 4). The derived class then has the characteristics of both base classes. The compiler actually does the work involved, the programmer simply benefits from the resulting inheritance. The alternative of embedding processes inside the Windo objects adds to the complexity of data and member access. It generally results in more work for the programmer and a less reliable program.

As an example of how the objects might work together, consider a simple PLL analysis. The engineer could extend the framework with a PLL class derived from the Process class. But the framework already provides everything necessary, with the exception of the loop formulas. In this case it is much faster to manipulate the internal objects at run-time, defining a project to perform the analysis. Since the built-in Parser object can handle complex math, a TextParseWindo is created to enter the formulas as shown in Figure 5. A Bode plot is created to observe the closed loop responses, and a ListWindo is created to view the phase margin. Both outputs are linked to the TextParseWindo in order to share variables and use the autoupdate option. When the TextParseWindo is told to analyze or tune it tells each of the linked output processes to update and whether the update is an analysis or tune operation. The output processes in this case use the logarithmic autosweep feature. For each frequency value the output processes tell the linked TextParseWindo object to execute and the results are collected and displayed. Figure 6 illustrates how a tune operation proceeds. Due to the small size of the formulas, constants and variables are edited directly in the TextParseWindo. Other than the typing of the names, numbers, and formulas all the actions are mouse clicks. Figure 7 shows the analysis in progress. The entire application from creation to completion of the analysis takes only a few minutes. The project and individual objects can be saved, and are easily modified or combined with other objects for future use. This example was purposely kept simple; more complicated tasks are much easier done than described.

### Getting Started

For the user of an OOCAE package, it is a simple matter to select objects from menus to perform a task then use the objects to perform the analysis. There is little to memorize when most objects present a box of options with defaults and the menus

provide a way to avoid obscure keystroke commands. It is mostly a matter of the user's imagination as to how a project is defined.

For the C++ programming engineer, the essential objects described above are developed into a framework (or purchased). Then the framework can be extended for each custom program, or the existing objects combined in a different way. Extension of the framework usually consists of deriving a class from an existing class to redefine one of the characteristics or to add a new feature. For the beginning programmer, there is also the choice of compiler to consider. From an engineering viewpoint the compiler that is easiest to use and has a built-in complex class would be the choice. Microsoft and Borland offer the most appropriate C/C++ compilers, with Borland having the edge in ease of use. (The Microsoft initial release did not include a complex class.) Often it is more efficient to purchase a book about the compiler rather than rely on the compiler documentation, since the documentation is usually reference oriented.

Purchasing a framework with source code can dramatically accelerate the learning process of both C++ and Windows. Additionally, starting with someone else's work saves time since one does not have to rewrite code that already works; simply extend it. There are libraries for numeric methods, arrays, Windows interface, imaging, graphing, and even engineering analysis that can save a tremendous amount of time and provide learning examples. A few of these that offer source code are:

Object Engineering (Innovation First, Mason OH) Engineering framework, based on Windows 3.1 MDI, works without programming.

Win++ (Blaise Computing, Berkeley CA) Windows interface class library, also offers a library for creating custom color controls.

ObjectWindows (Borland, Scotts Valley CA) Windows interface framework, optionally bundled with compiler.

Microsoft Foundation Classes (Microsoft, Redmond WA) Windows interface framework, bundled with version 7 compiler, (SDK is required), initial release does not have a complex number class.

M++ (Dyad Software, Renton WA) Comprehensive array class including eigensystem, optimization, factorization support, and complex class.

C++Views (CNS, Chanhassen MN) Windows framework.

Math.h++ (Rogue Wave, Corvallis OR) math class supporting complex numbers, FFT, linear algebra, Fourier, inverse Fourier, statistics, matrices.

The availability of these and others for C++ and Windows is an incentive to learn C++ and Windows programming rather than less popular languages and environments with limited tools available.

### Summary

The basics of an engineering framework is presented that takes advantage of object oriented techniques and the Microsoft Windows environment. The framework is designed with a reusable interface and the manipulation of internal objects to reduce the programming task, often eliminating it. Custom applications can be shaped by the user at run time, as opposed to the specific sequential operation traditionally used. This approach, powerful, yet easy to create, makes professional custom CAE practical.

### Reference

Shaw, R.H., "Anatomy of a Utility: Writing Applications with C++," *PC Magazine*, Feb 25 1992, vol. 31, No. 10.

Keuffel, W. and Hildebrand, J., "Hatch Better Programs," *Win Tech Journal*, March 1992, vol. 1, No. 2. \*\*\*

Levaro, R., "A Perfect Fit," *Win Tech Journal*, March 1992, vol. 1, No. 1. \*\*\*

Petzold, C., *Programming Windows*, Microsoft Press, Redmond, WA, 1990. \*

Norton, P. and Yao, P., *Windows 3.1 Power Programming Techniques*, Bantam, New York, NY, 1990.

Faison, T., *Borland C++ 3*, Sams, Carmel, IN, 1992.

Microsoft Corporation, *Microsoft Windows SDK: Guide to Programming*, Microsoft Corporation, Redmond, WA, 1992.

Richter, J., *Windows 3: A Developer's Guide*, M&T Books, Redwood City, CA, 1991. \*

Schildt, H., *C Power User's Guide*, McGraw-Hill, Berkeley, CA, 1988. \*\*

Press, W., Flannery, B., Teukolsky, S. and Vetterling, W., *Numerical Recipes in C*, Cambridge University Press, New York, NY, 1990.

Embree, P. and Kimble, B., *C Language Algorithms for Digital Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1991.

Smith, J., *C++ for Scientists and Engineers*, McGraw-Hill, New York, NY, 1991.

Baker, L., *C Tools for Scientists and Engineers*, McGraw-Hill, New York, NY, 1989.

\* Later version should/may be available

\*\* (Good tutorial on text parsing in the chapter entitled "Language Interpreters" with source code.)

\*\*\* (Every issue of Win Tech Journal provides good information on Windows programming, 1-800-234-0386 to subscribe)

// A sample complex class, usually much larger with more operators and functions.

// Essentially extends the C++ language to include complex variables.

```
class complex {
public:
    double re;           // real part
    double im;          // imaginary part
    complex operator-(); // unary negative - operator
    complex operator+(complex&, complex&); // add complex to complex
    complex operator+(double, complex&); // add real to complex
    complex operator-(complex&, complex&); // subtract complex from complex
    complex operator*(complex&, complex&); // multiply complex by complex
    complex operator*(complex&, double); // multiply complex by real
    complex log10(complex&); // log of complex
    complex cos(complex&); // cos of complex
    complex sin(complex&); // sin of complex
    complex tan(complex&); // tan of complex
    double arg(complex&); // angle in complex plane
    complex conj(complex&); // complex conjugate
    complex polar(double, double); // create complex from polar
};
```

Figure 1. A sample complex class

// Basic operation of a recursive decent parser, parse1() is the normal entry point

// (simplified for illustration purposes, see Shildt reference for a tutorial on parsers)

```
parse1(complex* answer)
{
    parse2(answer);
    if(Equal sign) .....
}
parse2(complex* answer)
{
    parse3(answer);
    if(relational operator) ....
}
parse3(complex* answer)
{
    parse4(answer);
    if(multiply or divide) ....
}
..... Other levels
parseN(complex* answer)
{
    if(parenthesis) parse1(answer);
    else Get the value of variable or function
}
```

Typical a \* b flow:

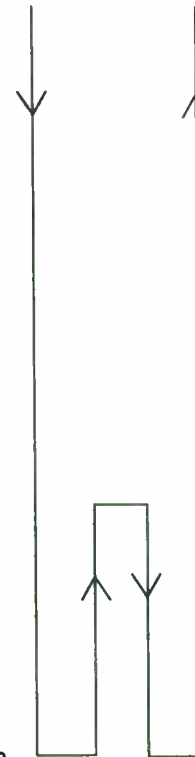


Figure 2. The Parser Implementation

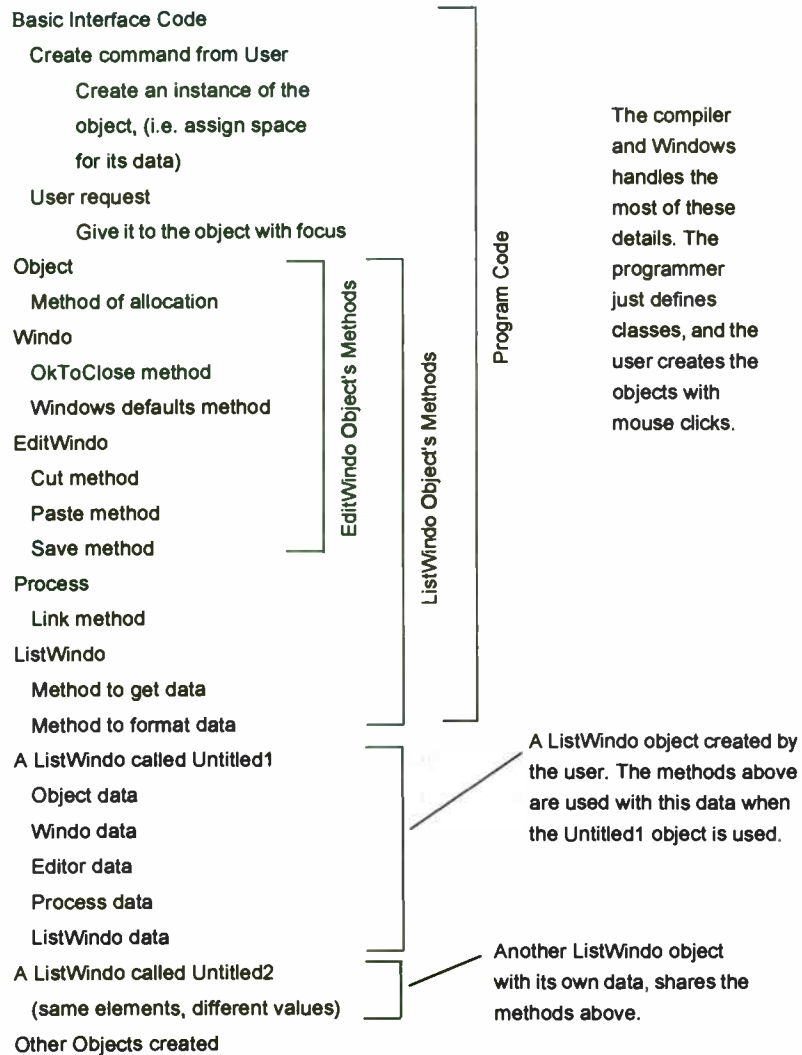


Figure 3. Internal View of Object Oriented Program

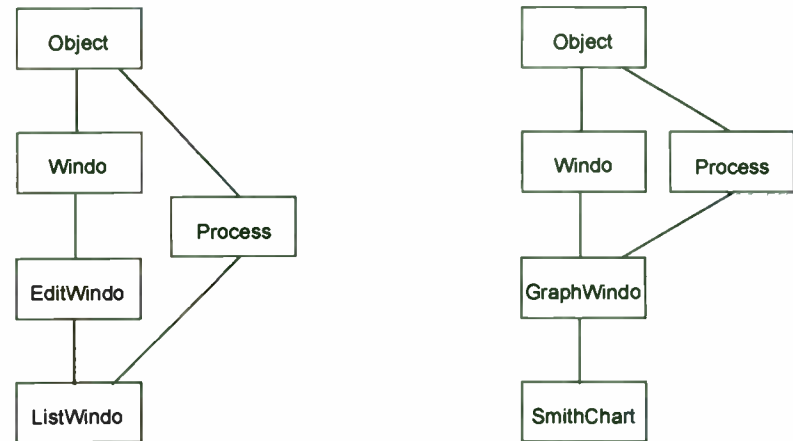


Figure 4. Typical hierarchy showing ListWindo and SmithChart

```

w = 6.2831853*Freq
R1 = 2470 // variables, in this case directly edited here
R2 = 2200
C = 0.01e-6
N = 1796
Kv = 6e6
Fs = (jw*R2*C + 1)/(jw*R1*C) // active filter
GOL = Kv/jw*4*Fs/N // open loop
GFM = 20*log(mag(1/(1 + GOL))) // closed loop (VCO)
GTR = 20*log(mag(GOL/(1 + GOL))) // closed loop (Ref)
GolMag = 20*log(mag(GOL))
GolAng = ang(GOL)
Fs = 20*log(mag(Fs))

```

Figure 5. Text typed into TextParseWindo for PLL analysis

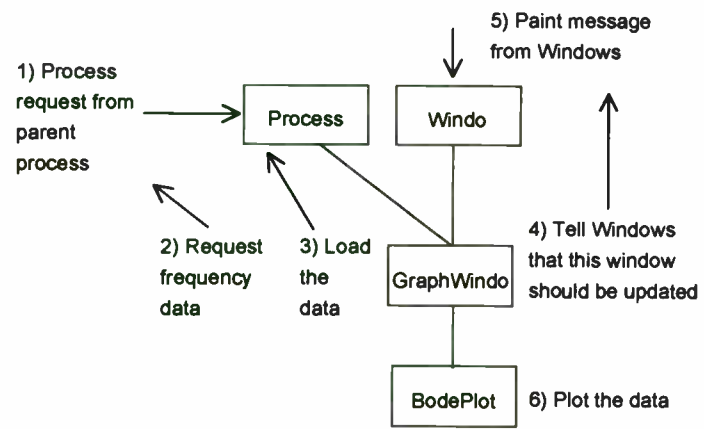


Figure 6. A typical tune operation

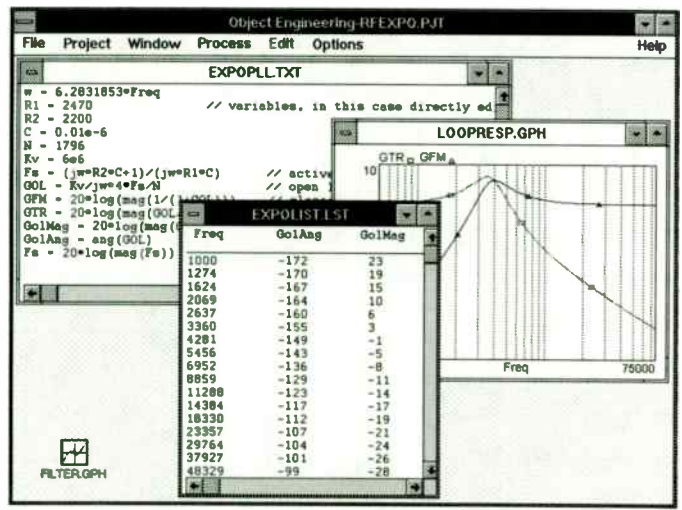


Figure 7. PLL analysis in progress

Handwritten notes, possibly a list or index, with some legible words like "List", "No.", and "Name".





## Signal Synthesis for Controlled Particle Extraction from High-Energy Accelerators

H. Meuth,<sup>(a)</sup> G. Heinrichs,<sup>(b)</sup> H. Halling<sup>(b)</sup>, and S. Papureanu<sup>(a)</sup> §  
(<sup>a</sup>) Institut für Kernphysik, (<sup>b</sup>) Zentrallabor für Elektronik  
Forschungszentrum Jülich,  
P.O. Box 1913, D-5170 Jülich 1, FRGermany

### Abstract:

Various schemes, employing magnetic and/or radio-frequency (rf) fields, may be used to extract high-energy beam particles from within an accelerator ring, to make them available to external beam lines and eventually experimental stations for medium energy and elementary particle research. Ultraslow extraction (USE) employs signals to influence the beam trajectories. This method is, due to the precision available with rf techniques, capable of a highly controllable extraction mechanism and is, therefore, preferred, when an even and constant particle flux extraction is required, especially for many hours. As opposed to acceleration and diagnosis where primarily (pure) single frequency signals are needed, the rf signals for USE require a carrier at a harmonic of the particle revolution frequency *plus* a precise tailored and sharply delimited superimposed noise band. As a novel feature, signal generation intended for COSY, a Cooler Synchrotron accelerator ring now under construction and soon to be commissioned at the research center KFA Jülich, is to employ direct digital synthesis techniques for its higher precision and flexibility. With a modular, computer-interfaced VME/VXI-frame layout, the digital system will incorporate the key features of (i) noise generation (0 to > 100 kHz), (ii) low-pass filtering functions with programmable cut-off at sharp roll-off (> 120 dB/octave) (iii) carrier frequency generation (~ 10 MHz), and, finally, (iv) double or single-sideband up-conversion of the noise band. Output to subsequent intermediate power amplifiers is rendered via digital-to-analog conversion, followed by reconstruction filtering and a DC-decoupled 50 Ω line-driving capability.

---

§ on leave from the Institute of Atomic Physics, Bukarest

### 1. Introduction:

The primary task of particle accelerators is to provide energetic particle beams with defined characteristics to the "user", to perform fundamental or applied research, or for certain manufacturing processes. Linear accelerators have a gun-like configuration, but can use any component only once for every particle. Ring accelerators, due to their recycling feature, may be more efficient, but require an extraction scheme, that, similar to a railroad switch, eject the accelerated particles out of the ring, to transport them to their final aim. Extraction schemes may simply "kick" the beam out of the ring, or slowly and continuously peel off beam particles, until no more particles are left in the ring. Thereafter, new particles must be injected, to begin a new acceleration cycle with final extraction.

To interact with particles, accelerators must employ quasi-static fields, which may guide, deflect, and focus charged energetic particles; and electromagnetic radio frequency fields for acceleration, and diagnostic beam measurements. For particle guidance at velocities of the order of the speed of light (> 0.1 c), magnetic fields are most efficient, while electric fields are used preferably at low energies. The fields are configured in certain multipoles; dipoles act as deflectors, quadrupoles act as focussing and defocussing lenses, and higher order multipoles, particularly sextupoles, are used for tailoring *transverse* (i.e., perpendicular to particle propagation) beam shape. RF fields are usually produced by high-power generators, coupled to one or more cavities; they serve to modify the longitudinal beam properties.

Characteristic radio frequencies in both linear and ring accelerators are determined by transit times in between accelerating stations. For ring accelerators, the fundamental frequency is thus  $f_{rev}$ , i.e., the particles' inverse revolution period. Since particle velocities  $v$  are to increase during acceleration, this must also be the case for the frequency. Depending on size, these frequencies may, typically, go as low as a few hundred kHz, for the fundamental in large rings, and reach some hundred MHz, with a tuning range of up to three octaves. Further, due to the (non-sinusoidal) pulse shape of the beam bunch, many higher harmonics may be present in the signal at any instant, and each harmonic contribution may be broadened due to a relative velocity or momentum spread of the particles. The advent of affordable VLSI circuits allows their use for digital rf needs, that fall into the above frequency range, see Fig. 1. As a result,

digital rf methods are receiving more and more consideration in the accelerator field: For instance, a first prototype digital system is now in place for the new cooler synchrotron COSY at the Research Center Jülich [1,2], and digital techniques are also being studied for the use of local oscillators in accelerator beam diagnostics.[3]

Quasistatic electric and magnetic fields are subject to certain limitation in rise-time and precision; field impact on particles results from the total force integrated along the particle trajectory through the multipole, giving rise to an additional uncertainty. These limitations apply especially to magnetic field components with large amounts of ferric materials involved, which must store substantial field energy, possibly close to the saturation limit. In contrast, rf techniques are, per se, capable of fast variations on the order of an rf period. Moreover, highly precise frequency reference and rf measurement techniques are well established.

Particles interact over an extended period only with those spectral components of the rf fields of frequency  $f$ , with which they are in tune, i.e.,  $f = h \times f_{rev}$ , where  $h$  is the harmonic number, and  $f_{rev}$  the revolution frequency. For a given total length of trajectory around the ring, the momentum  $p$  of particles of rest energy  $E_0$ , and the rf frequency are related in the following way

$$f(p) = \frac{h f_{\omega}}{\sqrt{\left(\frac{E_0}{cp}\right)^2 + 1}} \quad ; \quad (1)$$

thus, the higher the momentum, the higher the rf frequency. In Equ. (1),  $f_{\omega}$  is the revolution frequency of particles with speed of light,  $v = c$ .

## 2. Common Particle Extraction Schemes:

Beam extraction may be accomplished by a number of methods, depending on specific needs. To eject all particles with one single kick, high-power pulsers are used, whereby a fast discharge of a capacitor bank drives an air-coil deflection dipole magnet solenoid. Here, timing between beam pulse and discharge pulse is crucial. This method renders the entire beam bunch to the user. Due to saturation of particle detectors and the subsequent electronics, it may, however, be more desirable to control the intensity of extracted beam such, that efficiency of detectors and electronics

is optimized. For higher intensities, saturation would set in, while for lower intensities, signal-to-noise would be given away. The extraction mechanism must therefore accommodate such a requirement, in that it ejects, at any point in time, exactly only that fraction of the beam which produces the desired beam intensity, even in time.

A commonly used scheme is the "resonance extraction", also called *slow extraction*, which uses the fact, that in the operation of an accelerator, there exist points of instability, which lead to particle loss, which can be configured as a particle drift to the exterior of the ring. Such resonances may occur, when trajectories of some particles exactly close on themselves, once they have gone once, twice, or at most a few times around the ring. Since particle trajectories are essentially influenced by the multipole fields, it is possible for the particles to reach a resonance, when these fields are suitably adjusted. Now, not all particles will reach such a resonance point all at once, but depending on their momentum,  $p$ . In Fig. 2, a particle distribution is shown with a resonance point nearby. This resonance point may be moved up to the distribution by suitably cranking the fields, to eject those particles in direct touch with the resonance. It can be seen immediately, that cranking the fields evenly in time will (i) feed particles in a bell-shaped density in time, since at beginning and end of the cranking procedure, there are the few particles in the tails, while there are many during the middle phase. While this effect could be overcome by a suitable, uneven (although highly precise) field cranking speed, (ii) another problem is more serious: Fields display, due to a finite, albeit small, power supply ripple, a ripple in field strength, that translates into a jitter of the resonance position in time, resulting, locally, in a jitter in cranking. This effect leads to an ejected spiky beam intensity, and is very difficult to get rid off with regulating means alone.

## 3. Ultraslow Extraction:

A remedy for the above-mentioned shortcomings is offered by the method of *ultraslow extraction* (USE), a scheme conceived [4], and essentially pioneered [5] at the European accelerator laboratory CERN. For our considerations here, the notation *ultraslow* is not very descriptive, but stems from the fact that it is capable of a very long extraction process, over and above the *slow* extraction, permitting controlled and even particle fluxes for several hours. The USE principle uses magnetic multipole fields to only *bring* the extraction resonance *into* close *vicinity* of the beam distribution.

To actually *manipulate and extract* beam particles, it employs exclusively rf techniques. Specifically, a harmonic of the particle revolution frequency,  $h \times f_{rev}$ , is superimposed with a noise band, either static, or slowly swept, as needed. The scheme uses the fact, that beam particles slowly migrate or diffuse in momentum space,[4,6] when subjected to such an rf noise power. This may be visualized, since the spread in particle momentum  $p$  is tantamount to a spread in particle velocity,  $\beta = v/c$ , and again tantamount to a spread in frequency  $f$ , if the total length of trajectory around the ring is kept fixed. Thus, we may set

$$\frac{\Delta f}{f} = \eta(p, \gamma_{tr}) \frac{\Delta p}{p} \quad (2)$$

where the coefficient  $\eta(p, \gamma_{tr})$  is a weak function on the particle momentum  $p$ , and the machine optics, here denoted by the parameter  $\gamma_{tr}$ . The spread  $\Delta p$  is measured in reference to the beam bunch's center momentum, cf. Fig. 2. Similarly, the frequency spread is taken in reference to the center frequency. Further, what was said about the center momentum and center frequency in the context of Equ. (1), applies also for their respective spreads.

For USE, see Fig. 3, one first applies a stationary noise to broaden the particle distribution from bell-shaped to rectangular; thereafter, the noise power is turned off, not to unnecessarily heat the beam. This *shaping* procedure removes the first weakness (i) of the *slow* extraction principle, mentioned above in Sec. 2. Subsequently, a noise band is applied, that spans between the beam particle distribution, and the ring's extraction resonance, and serves as a diffusion (or *extraction*) channel, carving away particles at the edge of the particle distribution, which then diffuse toward the resonance, where they are ejected. Very much like an impedance, this diffusion channel decouples any (ii) resonance jitter from the beam feeding and ejection process. The steep roll-offs at the noise edges are essential to precisely define the cutting edge. Finally, to ensure an ejection of high efficiency, the resonance may be covered with a narrow noise band of high power density ("*chimney*"), by which particles are kept from diffusing through to the other side of the resonance region without ejection. In order to have a continued particle flux, the cutting edge of the noise band has to be moved continuously into the particle distribution, thereby slowly eating it away.

Except for the cutting edge, the noise band pattern of the extraction channel should, actually, be shaped in a subtle roll, c.f. Fig. 4, and not simply flat, as indicated in the simplified scheme of Fig. 3. This helps to gradually match the particle flux between the carving edge on the one side, and the ejection at the resonance, without which unwanted particle rejection could arise.[7]

Now, for the LEAR accelerator at CERN,[5] the coefficient  $\eta$  of Equ. (2) remains virtually constant ( $0.9 \leq \eta \leq 0.7$ ) for the entire range of operation. In contrast,  $\eta$  may vary in COSY by about two orders of magnitude for the attainable range of particle momenta; as a matter of fact, it may go through zero and become negative, i.e.,  $0.9 \leq \eta \leq -0.1$ , implying that, then, momenta above the center momentum correspond to frequencies below the center carrier. However, negative  $\eta$  may be avoided during the extraction procedure. Nevertheless, this implies, that with one and the same momentum spread, the necessary noise band width may vary by two orders of magnitude. With an expected maximum momentum spread of  $\Delta p/p \leq 5 \times 10^{-3}$ , a noise band  $5 \times 10^{-3} \leq \Delta f/f \leq 2 \times 10^{-4}$  is needed. This large window of operation, and also the interesting perspective of flexibly configuring other operation parameters led to the idea, to develop a for COSY a digital synthesis system for rf signal generation with USE. However, due to the programmability, this system need not be limited to COSY parameters.

#### 4. High-Power Components for USE:

To define all requirements on our digital signal generation system, we briefly have to touch the amplifier and beam-coupling set-up, which is mainly a high-power evolution of the somewhat simpler LEAR system.[8] To couple the noise power into the beam, propagating inside a UHV beam pipe, an inductive "narrow band" structure is used, somewhat akin to a transformer, see Fig. 5. Narrow band means here a pass-band feature with a 3-dB-breadth  $\Delta f/f \sim 0.2$ , which is much wider than the maximum noise band width needed from Equ. (2). The coupling structure, also called kicker, is built up in modules, each containing a cavity type shell, and a ferrite to have a broad resonance anywhere between 8 and 25 MHz. The actual resonance is defined by the capacitance of the coupling loop. The COSY kicker is laid out for 12 MHz center frequency. Actual final power levels of ca. 1kW total are generated by commercial rf

amplifiers, but play no role for specifying the parameters of the low level signal generation system.

### 5. Key Performance Requirements:

The condition on the carrier frequency is given by both Equ. (1), and the passband properties of the kicker. Both conditions are combined in Fig. 6, from which we may conclude that a slight tuning of the kicker structure by externally adjusting capacitance may be necessary. An entire extraction process may take anywhere from a few seconds to hours. Noise power, and achieved particle density are directly related, leading to the following conditions. The noise band width for beam shaping must be in accordance with Equ. 2, using the 4 $\sigma$ -width, while the band width for the extraction channel follows from the distance of approach to the extraction resonance, which, for our purpose, may be assumed to be of the same order. Attenuation rate or roll-off at the noise cut-off of the extraction channel is crucial to define the precision of the extracted particle population, but may be, if necessary, relaxed for the shaping channel, as long as initial and final particle flux transients are tolerable. Inversely, noise-band flatness for the shaping channel is subject to tight requirements, since it directly influences an even particle distribution, while the conditions on the extraction channel are not as severe, calling for an intentional roll. Finally, filter specifications are usually given in terms of voltage, and not in terms of power. Thus, while an *attenuation rate* of more than 120 dB/octave is conceivable, a maximum total particle number of 10<sup>12</sup> reasonably limits, in any case, the needed *voltage level ratio* to max. 120 dB.

max. carrier center frequency	< 15 MHz
max. carrier sweep speed	100 kHz/s
max. output voltage	1 V/50 $\Omega$
noise bandwidth, adjustable	1 to 100 kHz
max. stopband attenuation	100 dB
attenuation rate	> 120 dB/octave
passband flatness	< 0.1 dB

Table 1: Key parameters for signal synthesis system

All these key requirements appear to be within the reach of a fully digital system, based on technology commercially available today.

### 6. Signal Synthesis Options:

Building upon the philosophy of an all-analog USE system developed at CERN,[8] we are in the process of developing a completely new sub-system for signal generation, that relies mainly on direct digital synthesis.

#### 6.1 carrier frequency synthesis and digital-to-analog conversion:

The carrier frequencies up to 15 MHz are accessible by commercial NCO circuits.[9] To minimize possible errors, or distortions, that reconstruction filtering could cause to the precisely tailored frequency signals, a high rate of oversampling is necessary, say 100 MHz, at 12-bit signal data. This data rate is definitely feasible, as far as digital-to-analog converters (DAC) are concerned, since, recently, a number of DACs, that are capable of performing at or above 100 MHz at 12 bits [10] have become available, or are soon to be released.

#### 6.2 noise generation and filtering:

With one exception,[11] USE signal generation schemes employ to date analog white noise sources. In contrast, we are examining for our system a number of options to generate the noise bands. Unfortunately, the 100 kHz bandwidth lets commercial digital audio processors fall just short of this target. Nevertheless, one scheme examined is to use one or more digital signal processors (DSP), which provide both the noise generation, and the subsequent filtering. The data sequence could be stored intermediately in a local memory with  $\geq 10^6$  12 to 16 bit data words, to be output on a faster time scale, as needed. Direct data output from the DSP without extensive local memory is also explored; and also the use of a large memory bank, as above, without local DSP, but rather with data down-loading capability from a remote control computer. The latter option could, for instance, be accomplished with a commercial arbitrary waveform generator. Yet another option is the use of an NCO with a suitable random frequency or phase input, a method, which would permit to differentiate

between frequency, phase, and amplitude noise, allowing for the first time specifically and precisely defined beam gymnastics.[6]

These schemes could have the filter task already incorporated by outputting a suitably computed data sequence. In case of using a DSP or an NCO, a separate dedicated filter bank may be necessary. With analog techniques, the key requirements of Table I, especially the high attenuation rate, would be best fulfilled by elliptic filters, therefore suggesting suitable digital IIR filters for our purpose. Presently, filter simulations are underway to pin down the required complexity, that may involve ten or more multiplications.

#### 6.3 up-conversion:

A fully digital implementation of both double side band and single side band up-conversion are being examined. Only the latter permits generating asymmetric noise patterns of the type of Fig. 4. As a matter of fact, we could combine this way the noise spectrum of the chimney, and of the extraction channel into one single band. Digital up-conversion is possible by digital multipliers, and is used in a 16-bit, 25 MHz version [12] in our fully digital rf control system for acceleration in COSY.[2] Unfortunately, only one butterfly multiplier is available on the market today, that fits only roughly our needs, performing in 12-bit, 60 MHz.[13] As a fall-back solution, an analog up-conversion is conceivable, for which both carrier frequency signals and noise signals are first separately converted from digital to analog, and thereafter they are then applied to an analog mixer.

#### 6.4 over-all system layout:

Two probable layouts for the system are shown schematically in Figs. 7 and 8. A version using mainly NCOs is depicted in Fig. 7. The NCOs render data in quadrature. One pair generates the noise signal via real-time random phase and/or frequency input. This method has the advantage, that the generated data are already in the proper format, i.e.  $\sin [2\pi f(t) t + \phi(t)]$  and  $\cos [2\pi f(t) t + \phi(t)]$ , where  $f(t)$  and  $\phi(t)$  contain a suitable value sequence in time, needed for the noise spectrum. Subsequent digital filters will refine the spectra, if necessary. Additionally, amplitude noise can also be added. The complex multiplication perform a single sideband up-conversion. In Figure 8, the task to generate noise and to filter it, as necessary, is delegated to a

digital signal processor, clocked at 40 MHz or higher, in combination with a fast memory bank of size 1Mx16 bit with access times of < 20 ns. The memory is capable of outputting the data sequences real time to the up-conversion module. The entire system will be accommodated in a VME/VXI frame, with network (LAN) connection to a central control computer, or a workstation. The frame is supplied with a controller to manage communication and data transfer, and could contain, in addition, a second CPU to take some of the computational burden off the DSP, or manage the real-time data input/output to the NCOs.

#### 7. Conclusions:

A number of options are being explored to implement a digital signal synthesis system for use with an ultraslow extraction system with COSY. The digital scheme promises considerable flexibility in static and dynamic parameter choice. Quality performance of the synthesis system must first be checked with standard rf measurement techniques. However, only the final performance together with the intermediate power components and the particle beam will be of essence, for which beam diagnostic measurements are also being prepared.

#### Acknowledgments:

We are much indebted to G. Molinari and M. Chanel from CERN, for many fruitful discussions concerning USE; H. Rongen and A. Schnase provided invaluable advice regarding the digital circuitry and interfacing.

#### References:

- [1] P. Shirkey, H. Meuth, and H. Rongen, and A. Schnase, *A fully digital RF-signal synthesis and phase control for acceleration in the cooler synchrotron COSY*, RF-Expo East Conference 1990
- [2] A. Schnase, H. Rongen, and H. Meuth, *A Digital Synthesizer and Phase Control System for RF-Acceleration in COSY*, European Particle Accelerator Conference, Berlin 1992
- [3] D. Peterson and J. Marriner, *A Frequency Tracking Synthesizer for Beam Diagnostic Systems*, U.S. Particle Accelerator Conference, San Francisco 1991
- [4] S. van der Meer, *Stochastic Extraction, a Low-Ripple Version of Resonant Extraction*, CERN, PS/AA Note 78-6, March 1978

- [5] R. Cappel, R. Giannini, and W. Hardt, *Ultralow Extraction (Status Report)*, in *Physics at LEAR with Low-Energy Cooled Antiprotons*, Vol. 17, p. 49, U. Gastaldi and R. Klapisch, eds., Plenum, New York, 1984; R. Cappel, W.E.K. Hardt, and Ch.P. Steinbach, *Ultralow Extraction with Good Duty Factor*, in: *Proceedings of the Eleventh Intl. Conf. on High-Energy Accelerators*, p. 335, W.S. Newman, ed., Birkhäuser, Basel, 1980.
- [6] G. Dôme, *Theory of RF Acceleration and RF Noise*, in CERN Report No. 84-15, Accelerator School on Antiprotons for Colliding Beam Facilities, 1984, pp. 215-260; J.A. Ellison, B.S. Newberger, and H.-J. Shih, *Effect of RF Phase Noise on the SSC Beam*, U.S. Particle Accelerator Conference, San Francisco 1991
- [7] W. Hardt, *Moulding the Noise Spectrum for Much Better Ultralow Extraction*, CERN, PS/DL/LEAR Note 84-2, March 1984
- [8] G. Molinari, LEAR, CERN, private comm.
- [9] Stanford Telecom STEL-1175+80, 12-bit , 80 MHz
- [10] Analog Devices AD9712, 100 MHz, 12-bit DAC; Tektronix TKDA30, 250 MHz, 12-bit DAC, see also: Chris G. Martinez and John Brewer (Tektronix Microelectronics) *A Monolithic 12-Bit 100MSPS Digital to Analog Converter for Frequency Synthesis Applications*, RF expo West, San Diego, March 1992; Burr-Brown DAC650, 500 MHz, 12-bit DAC.; Triquint Gigadac TQ61XX, 1Gs/s 14-Bit GaAs DAC IC, see also Frederick G. Weiss and Tyler G. Bowman, *A 14-Bit, 1 Gs/s DAC for Direct Digital Synthesis Applications*, GaAs IC Symposium 1991, IEEE CH3059-3 91 0000-0361, p. 361.
- [11] F. Nolden, J. Pinkow, and U. Schaaf, *Experiments with Synthetic Coloured Noise at the Heavy Ion Storage Ring ESR*, European Particle Accelerator Conference, Berlin 1992
- [12] GEC Plessey PDSP16116/A, 16x16-bit Complex Multiplier
- [13] Stanford Telecom STEL-1130, 12x12-bit , 60 MHz, Quadrature Amplitude Modulator

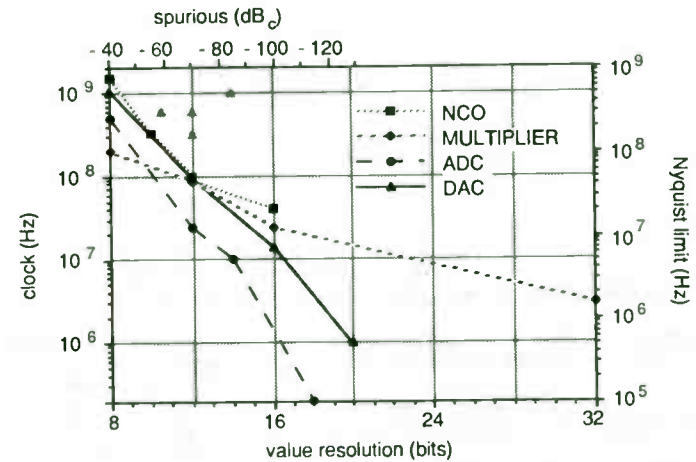


Fig. 1: Performance of present day digital rf VLSI circuits

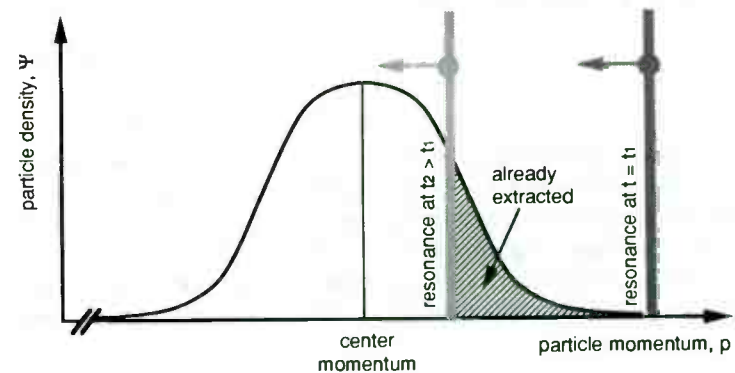


Fig. 2: Particle distribution, and resonance, vs. particle momentum for slow extraction.

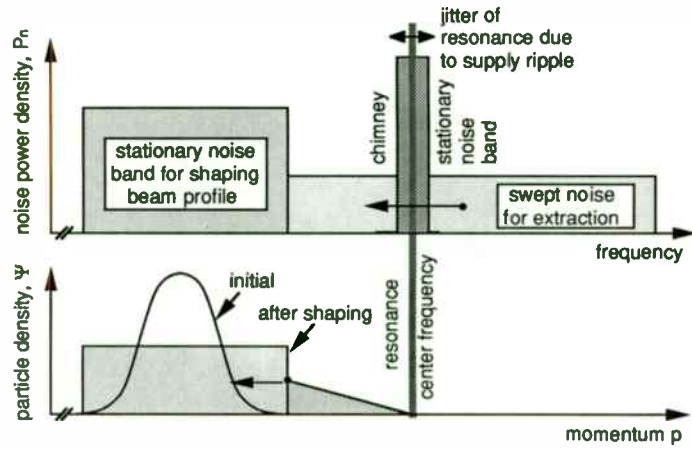


Fig. 3: Particle distribution and noise power densities for USE

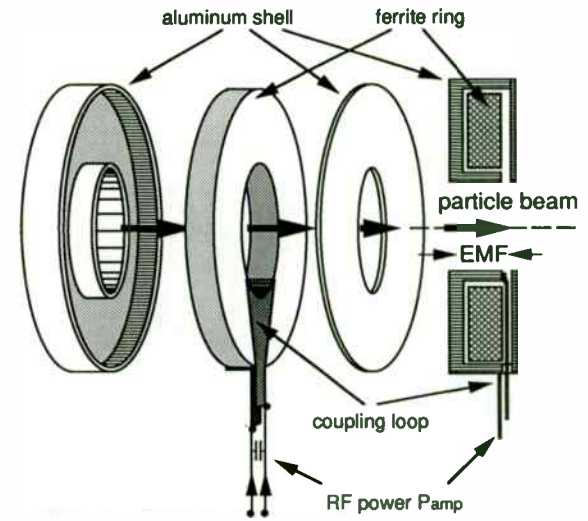


Fig. 5: kicker module for USE

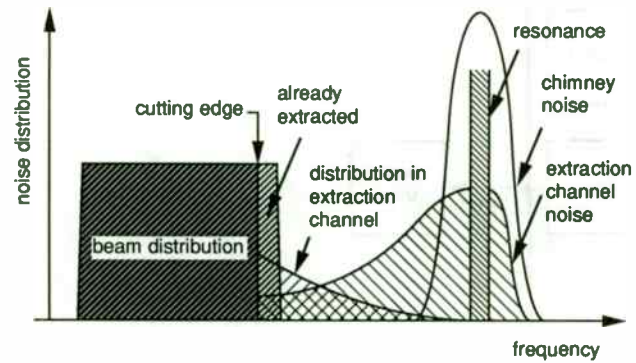


Fig. 4: Distributions with specifically moulded extraction channel noise band

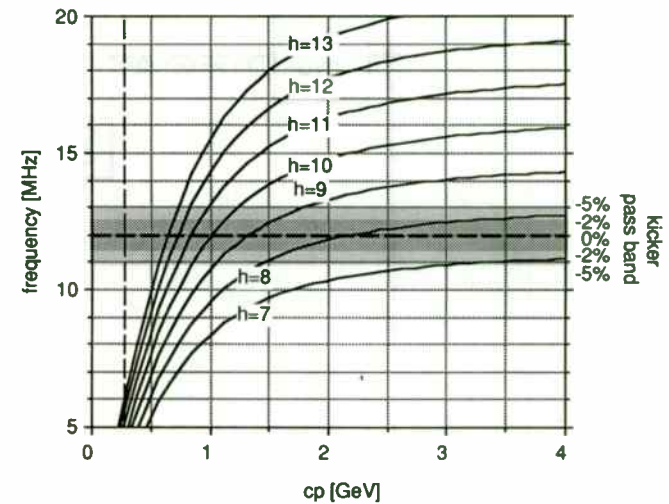


Fig. 6: Harmonics of particle revolution frequency and kicker pass band

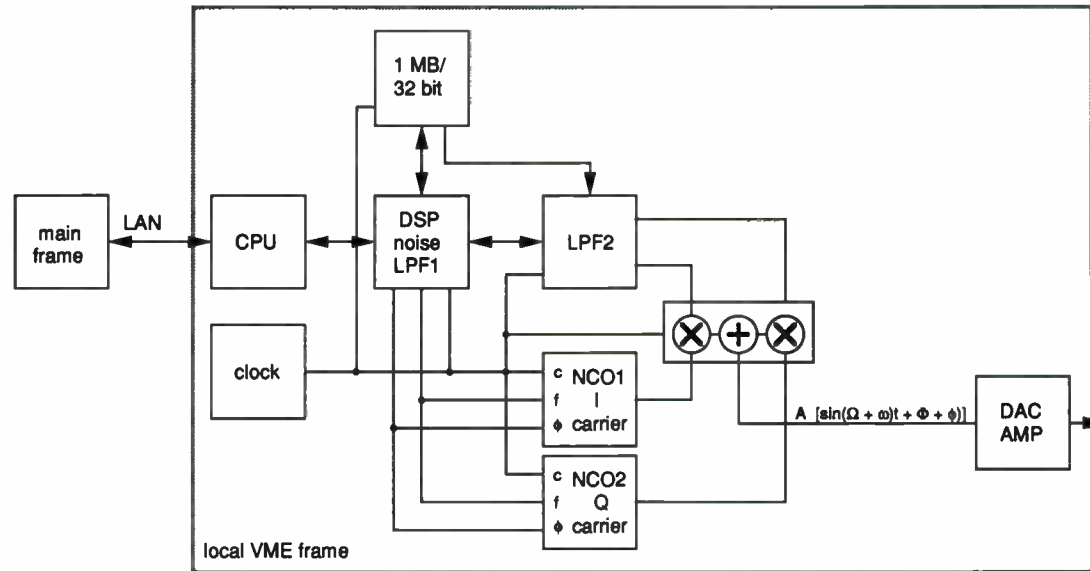


Fig. 8: Schematic layout based on DSP



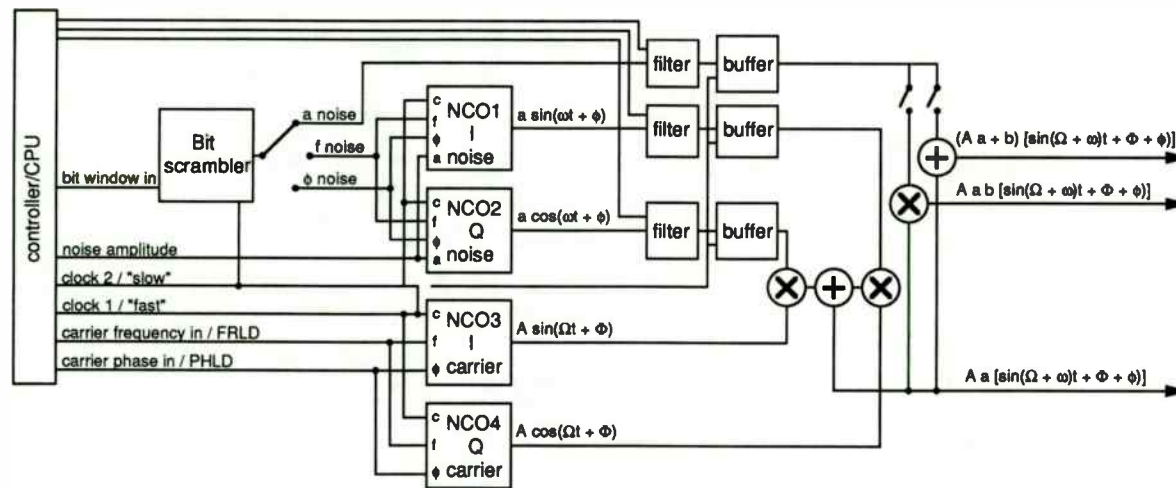


Fig. 7: Schematic layout based on NCOs



NONCONTACTING COAL AND ROCK THICKNESS MEASUREMENT WITH A  
VECTOR NETWORK ANALYZER

Robert L. Chufo  
U.S. Bureau of Mines  
Pittsburgh Research Center  
Cochrans Mill Road  
Pittsburgh, Pennsylvania 15236

Abstract

A noncontacting electromagnetic sensor has been developed by the U. S. Bureau of Mines that measures the thickness and dielectric constant of coal and rock and locates the interface in multilayer media. The technique uses a vector network analyzer, an IEEE488 bus controller, and a servo controlled L-band antenna positioner. Spatial modulation is created by moving the sensor antenna through a distance of 16 in and measuring the reflection coefficient at 401 frequencies from 0.6 to 1.4 GHz at each of 32 antenna positions. The antenna motion imparts spatial modulation to the data that, through signal processing, solves the problem of media, target, and antenna dispersion. The system along with its simple dipole antenna is calibrated by making a measurement of a reflecting metal surface. The signal processing technique avoids the use of radio frequency absorbing material to deal with multipath problems. In addition to a presentation of the system theory and architecture, field test results and application examples will be presented. The system measures coal and rock thickness of less than 3 in to over 60 in to within 1-in accuracy. The sensor has been

developed as a coal/noncoal interface detector in the computer-assisted mining machine program being conducted by the Bureau, and also has application to locating rock fractures and water filled voids.

Introduction

Ground penetrating radar (GPR) systems in general have been limited in their resolution and measurement accuracy by two problems: Penetration and focus. Electromagnetic waves of low frequency penetrate the media but do not have the ability to resolve small objects. The information content of the energy that does penetrate the media and returns to the antenna appears to be scrambled and out of focus. The problem is media dispersion. The usual GPR systems use time domain processing of electromagnetic pulse, impulse or synthetic pulse signals. Media dispersion and the added complication of antenna and target dispersion cause problems too complex to sort out in the time domain. These problems are much easier to solve in the frequency domain. The theory supporting this new research developed by the U. S. Bureau of Mines is in the frequency domain, but the signal processing architecture uses both the frequency domain and the time domain.

Previous GPR systems required knowledge of the dielectric constant of the material being measured or required that a known thickness of the unknown material be measured to calibrate the system. The present technique does not require prior knowledge of the dielectric constant because it is automatically calculated during the thickness measurement process.

The radar thickness measurement technique described in this paper uses a spatial domain concept (i.e., moving the antenna to create a modulation on the radar output that can be processed in the frequency domain). The motion occurs perpendicular to the target and is used to advantage to remove reflections that are not along the antenna motion axis, the axis of interest.

The technique uses a vector network analyzer, an IEEE488 bus controller, and a servo positioned L-band dipole antenna. The antenna motion imparts spatial modulation to the data that, through signal processing, solves the problem of media, target, and antenna dispersion. The simple dipole antenna is moved through a distance of 16 in and the reflection coefficient is measured at 401 frequencies, from 0.6 to 1.4 GHz, at each of 32 antenna positions. A nominal power level of 0 dBm or less is sufficient to penetrate to depths of over 10 ft. To date, no effort has been made to optimize the system parameters. The reader will recognize that some of the values are network analyzer default values, others have been arbitrarily chosen. The present values, whether wisely chosen or not, have produced a thickness measurement resolution of better than 1 in.

The system, including the antenna, is calibrated by measuring a reflecting metal surface. This is a hardware and environmental calibration and does not include the sample to be measured. A corrected reflection coefficient is then calculated by essentially dividing the target reflection data by the metal reflection data. The target is usually a multi-layered geological deposit as found in the roof or floor of an underground coal mine. The deposit consists of layers of coal and

slate or shale. The resulting frequency domain reflection coefficient is transformed into the time domain where reflections from each dielectric interface can be identified. The strong reflection from the air/coal interface can be separated from the weaker reflection from the coal/shale or coal/rock interface. This technology is used to determine the coal thickness remaining between the cutter head of a mining machine and the mine roof rock above the coal.

Highwall mining is done from the surface with a highwall mining machine positioned opposite a hillside. The coal vein has been previously exposed by cutting into the hillside with earth moving equipment to create what is known as a highwall. The coal is mined by cutting or augering parallel tunnels, known as entries, into the highwall. For highwall mining machine control it is necessary to keep the entry being mined parallel to the previous entries and to know the width of the coal, known as the rib, remaining between entries. Modern mining requires that the rib remaining be of sufficient width to prevent the collapse of the mine roof onto the mining machine. An excessively wide rib leaves too much coal unmined, wasting a valuable resource. A rib width between 2 ft and 4 ft is usually left in the hillside. Reflections from the air/coal and coal/air interface on both sides of the coal rib can be identified and the distance between them, rib thickness, can be determined and used to control the heading of the mining machine.

The system measures coal and rock thickness of less than 3 in to over 60 in. Similar measurements of sandstone, granite and salt have been made with even greater accuracy due to their better homogeneity

compared to coal. The technology can be used at both higher and lower frequencies. The depth of penetration will be greater at lower frequencies and the ability to resolve small objects will be greater at higher frequencies. Frequency selection can be used to advantage to avoid the detection of small detail that is of no interest or to avoid the detection of objects at depths beyond concern. Laboratory and field testing has also been done at S-band. Better resolution was obtained but with less depth of penetration. Applications of detecting clay intrusions, rock fractures, and water filled voids are also possible with this versatile technology. A by-product of this thickness measurement technique is the noncontacting measurement of the complex dielectric constant of each layer of the multilayer media being measured. The measurement of the dielectric constant is a necessary factor in the determination of the thickness of each dielectric layer.

#### Acknowledgment

The author acknowledges the contributions of Boeing Aerospace and Electronics in supplying material for this paper under U.S. Bureau of Mines contract No. H0398009, entitled Development of a Radar Coal Thickness Sensor. Walter J. Johnson is the Principal Investigator and Jerome M. Brossard is the Radar Staff Engineer on this contract. Raymond J. Helinski, Electronic Technician of the Bureau of Mines assisted in installing and modifying the software at the Pittsburgh Research Center and in the data collection during both surface and underground validation testing of the model.

#### Processing

The following discussion outlines the Radar Coal Thickness Sensor measurement technique as it currently exists in developmental form. The original research started with Kerns plane wave theory [1] and solved for each of the scattering coefficients in the matrix with a standard antenna calibration technique. However, when the antenna was within two wavelengths of the material being measured, diffraction became too strong an effect for a plane wave model. A solution was sought in simple spherical wave theory but the penalty of doing this was the necessity of finding an explicit solution for the coefficients of the model. Several approximation techniques for linear calibration were tried and it was found that a solution was possible. The coefficients did not vary when the antenna was closer than two wavelengths to a metal calibration plate. This was called the "Linear Reduction Method" [2] and it worked quite well except for some second order problems. It was assumed that these problems were nonlinear multipath effects that were not accounted for in the calibration procedure. The math model was then expanded to include the higher order effects and called the "Quick Reduction Method." Many of the higher order coefficients were lumped together to accommodate a metal plate calibration technique done with a 4-ft square metal plate placed between the sample and the antenna. This in situ calibration technique corrected for signals returned from reflectors beyond the edges of the calibration plate. In an underground mine these reflections would be from material similar to the target coal but outside the measurement area.

### Measurement

The thickness measurement process begins with a measurement of the input reflection coefficient of an antenna in close proximity to the coal surface. This measurement is taken at a wide range of frequencies and positions (e.g., 401 frequencies between 600 and 1400 MHz, and 32 positions between 4 and 20 in from the coal). This is accomplished using a vector network analyzer connected to an antenna moved in space by a linear positioner. Then the measurement plane is electrically moved from the instrument measurement plane to the plane of the antenna. Figure 1 shows the measurement test setup.

### Data

The data from this measurement are a function of both frequency and position and contain both amplitude and phase information. The data at this point represents a plot of the frequency domain history for one antenna position.

Transforming the data to the time domain at this point in the process and inspecting the time domain history for this one antenna position shows the absence of any sharp peaks around the antenna. This indicates that the information for the coal surface is corrupted by other effects such as antenna dispersion, diffraction, and multipath. These effects must be characterized and accounted for by considering the frequency domain history at each antenna position.

### Antenna Transfer Functions

To characterize the antenna, a separate test is run with a metal surface substituted for the coal surface; the same frequencies and positions are used. This provides data from a known reflection surface to obtain the antenna transfer functions. These functions are used to remove antenna dispersion from target data taken at the corresponding antenna position.

### Removing Antenna Dispersion

When the antenna transfer functions are accounted for in the data the result is the product of the antenna-to-surface-to-antenna distance, represented by the spatial delay, and the coal surface reflection coefficient. The reflection coefficient contains the information for the coal thickness. Other reflections (i.e., multipath) are also present in the resulting transfer function.

### Shifting Image Plane to Coal and Removing Diffraction

Dividing the reflection coefficient expression by the spatial delay shifts the image plane from the antenna to the surface of the coal and removes diffraction. With the antenna-to-surface-to-antenna distance a known quantity, the true reflection coefficient in the frequency domain can be determined.

### Integrating Space

By performing a spatial integration we can now decorrelate the multipath. Since spatial integration is coherent with the coal but not

coherent with any other spatial distances, the multipath will become zero sum or at least small compared with the surface reflection.

#### Transforming to the Time Domain and Range Gating

Transforming the data to the time domain and range gating removes unwanted reflections from the data by gating out all the information on either side of the desired peak. However, recent data as presented in the results section of this paper have shown that range gating is usually unnecessary as the signal return from the dielectric interfaces between the coal and rock are sufficiently distinct to make the interface easily discernible. Transforming the reflection coefficient back to the frequency domain, after gating, yields the composite reflection coefficient for just the reflections within the range gate.

#### Correlating with Theory

An earlier method [2] for determining the thickness transformed the time domain reflection coefficient back to the frequency domain so that the measured reflection coefficient could be correlated with theoretical reflection coefficients for various thicknesses and dielectric constants. The theoretical reflection coefficient that correlated best provided a statistical determination of the thickness and dielectric constant of the coal being measured. For example, the theoretical reflection coefficient that correlated best with the measured reflection coefficient was for a coal thickness of 5.3 in (for a relative dielectric 4 and loss tan of .03). The actual thickness of the rough wet coal measured in this early underground test was a nominal 6 in.

This result provided the encouragement to proceed with the development of a scheme for the direct measurement of thickness utilizing the time domain data.

#### Field Tests

The purpose of this research was to develop a coal and rock thickness sensor of sufficient accuracy to provide vertical and horizontal guidance of both room-and-pillar and highwall mining machines. In order to validate the theory developed for thickness measurement, extensive underground and surface mine testing was preformed. Over a period of two years, tests were conducted in mines with a variety of geological and environmental conditions. Test areas of both freshly mined and aged coal, from 3 to 60 in thick, were measured. The areas measured ranged from very dry to extremely wet with water dripping from the roof test area. The wet coal did not affect the thickness measurement. Coal seams with clay and metal vein intrusions of iron pyrite could be imaged and the distance from the coal surface to the intrusion was accurately measured. Surface roughness and cleating was not a problem. The average thickness of rough cleated surfaces was measured accurately. Accurate measurements were obtained even when water filled the cracks between the cleats.

#### Roof Tests

Roof thickness tests were made in production mines and in the Bruceston Research Coal Mine at the Bureau's Pittsburgh Research Center.

Figure 2 is a representative measurement of roof coal thickness. On the vertical axis, the plot shows the amplitude of the reflected signal in decibels; time in nanoseconds is shown on the horizontal axis. The large peak on the vertical axis represents the reflection from the first interface, the air/coal interface. Signals plotted to the left of the large peak represent discontinuities internal to the measurement equipment and between the antenna and the coal surface. These reflections are reduced to at least 30 dB below the air coal reflection by the calibration and spatial integration scheme. Peaks to the right of the large air/coal reflections are from discontinuities internal to the coal and shale being measured. The printout on the left shows the thickness of coal between the air/coal interface and the coal/shale interface. Measurements have identified both the thickness of the coal and the thickness of the next layer, usually shale, above the coal roof. At the L-band frequencies presently used, the depth of penetration is about 10 ft. Future roof thickness measurement research will attempt to provide a direct readout of the thickness of each layer of geological material within the penetration range of the signal. Presently, the power level of the transmitted signal is 0 dBm. This signal level, or less, is adequate to produce a good signal-to-noise ratio for the return signal measurement. The hardware will permit an increase in transmitted power of 20 dB to determine if greater penetration is best achieved through increased signal power or through the use of a lower transmitter frequency. Both the hardware and software will operate from 300 kHz to 3,000 MHz.

#### Rib Tests

Figure 3 is a plot of actual data taken at an operating highwall mine in West Virginia. Measurements were made in freshly mined entries immediately following the mining machine. The determination of rib thickness can usually be interpreted both visually, from the FFT data in the figure, and from the numerical readout from the automatic thickness measurement software. A large dielectric contrast is seen at the first air/coal interface and a somewhat smaller but still pronounced reflection can be seen as the signal exits the coal rib at the coal/air interface in the adjacent drift. The vertical bar to the right of the main peak indicates that the rib thickness is 35 in. The dielectric constant and loss in decibels-per-meter are also indicated above the rib thickness measurement printout at the left of the plot. Additional measurements were made over the range of 18 in to 50 in. The thickness in these test ribs could be determined to within 1 or 2 in.

#### Measurement of Other Geological Materials

Figure 4 shows the data obtained from a granite block. The sample measured was 36 in by 45 in by 48 in. The thickness was determined to within 3% of the actual 36-in thickness. This small error is due to a rounding error in the calculation algorithm that can be corrected if extreme accuracy is required.

Figure 5 shows data obtained from a 20 in by 44 in by 44 in Berea sandstone block. The thickness was measured to within 1 in or 5% of the 20-in dimension.



In conducting research on thickness and dielectric constant measurement it was convenient to have a calibration standard available in the surface laboratory. A satisfactory and inexpensive material is salt in the form of the cattle salt blocks found in local feed stores. The 50-lb blocks were stacked to form a wall 4 ft square by 9 in thick. Figure 6 shows the data obtained from the salt wall. The correct thickness of 9 in was measured with a dielectric constant of 4, which is similar to the dielectric of coal.

A coal sample was removed from the Bruceton mine and measured to determine the effect sample size and shape irregularities would have on the coal thickness measurement. While the single test was far from conclusive, it is felt that irregular samples with a face area of less than 4 ft square will require a higher measurement frequency for satisfactory results. The software can easily accommodate changes in any measurement parameter as it was written to be a general solution for coal and rock thickness measurement.

#### Conclusions and Recommendations

Field testing of the electromagnetic coal thickness sensor has produced results of sufficient accuracy (1 in for coal from 3 in to 60 in thick) to justify continuing with the engineering work necessary to develop a practical sensor that can be mounted on a mining machine for the determination of roof, floor, and rib thickness. Future plans are to determine the range of antenna motion necessary, the number of frequencies, and frequency range that must be used to preserve the present measurement accuracy. Minimizing the size of the data set will

reduce the software processing time and facilitate the construction of a compact machine mountable sensor.

It was found that vertical E-field polarization penetrates thicker coal ribs than horizontal E-fields. This is thought to be due to the thin horizontal ash layers having a higher loss than the coal.

The real part of the dielectric in coal varies very little from a value of 4 but the loss tangent varies a great deal. Wet, rough or heavily cleated coal had little effect on the dielectric measurement.

The ash content may be related to the loss tangent of the dielectric measurement. This would be a helpful means to identify higher quality coal.

Sub-wavelength interpretation at L-band worked down to 0.25 in when the second surface had a high contrast like clay or shale. The layering interpretation was more difficult when the second layer was sandstone with a dielectric constant approaching that of coal.

#### References

1. Kerns, D. M., Plane-Wave Scattering-Matrix Theory of Antennas and Antenna-Antenna Interactions, Nat. Bur. Stand., NBS Monogram 162, June 1981.
2. Chufo, R. L. and Walter J. Johnson, A Radar Coal Thickness Sensor, The 1991 IEEE Industry Applications Society Annual Meeting, September 28-October 4, 1991, Dearborn, MI.

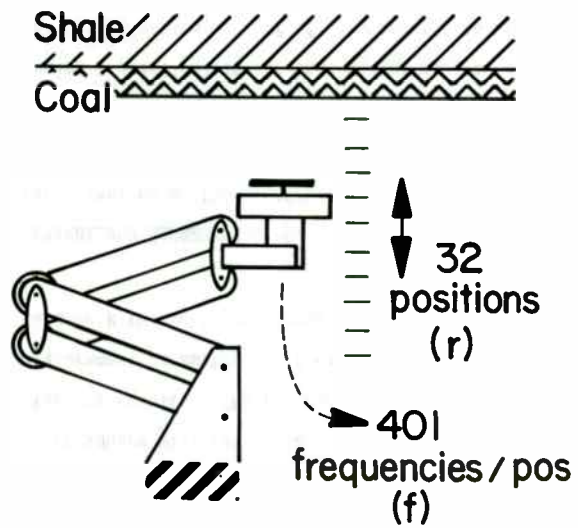


Figure 1. Measurement test setup

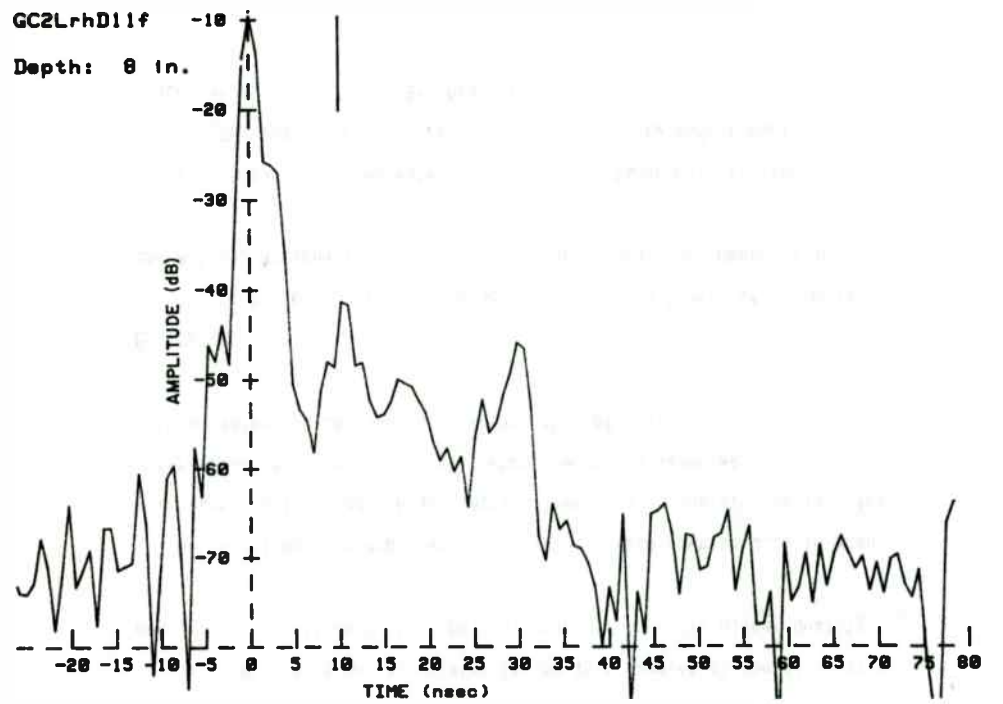


Figure 2. Roof coal thickness

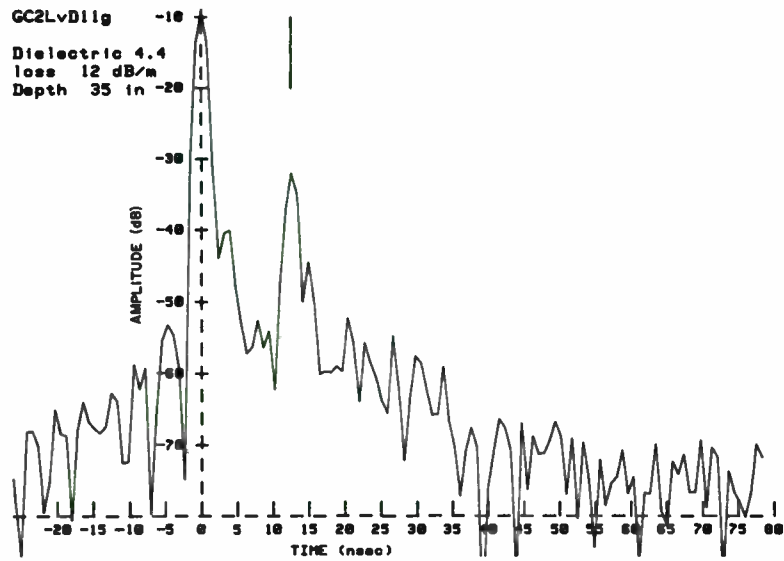


Figure 3. Rib thickness

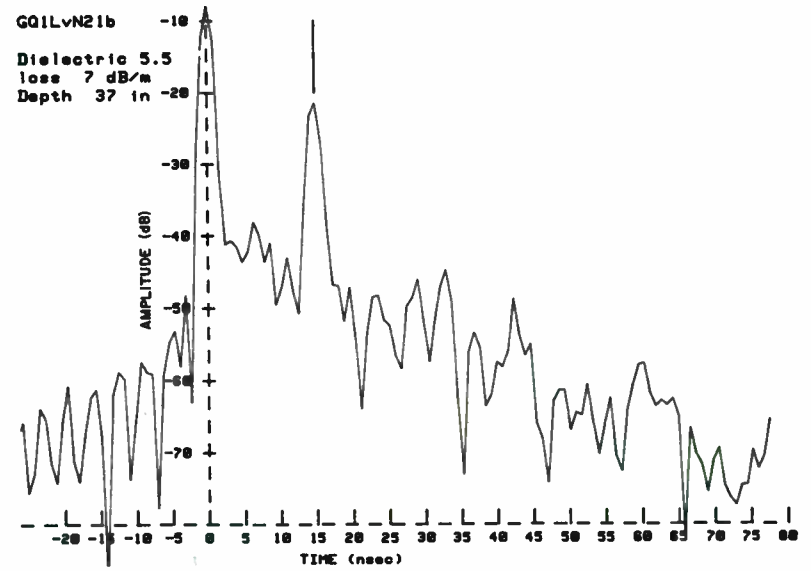


Figure 4. Granite block thickness

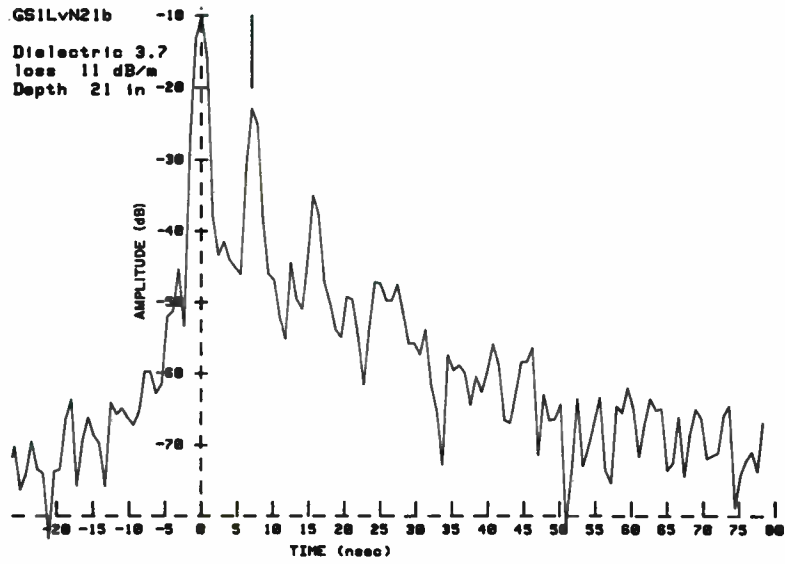


Figure 5. Sandstone block thickness

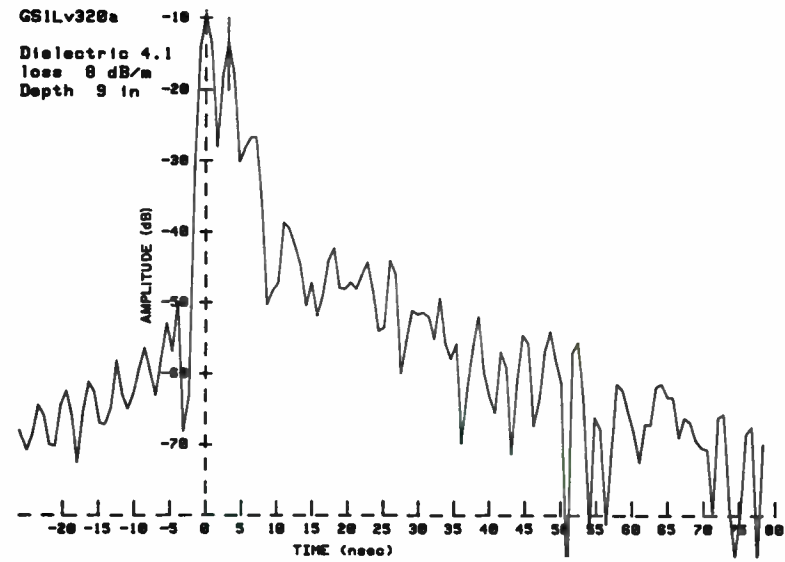


Figure 6. Salt block thickness

# A RAPID ACQUISITION TECHNIQUE OF SPREAD SPECTRUM SIGNALS EMBEDDED IN AN ASIC COSTAS PHASE LOCKED LOOP

G.G. Koller, M.A. Belkerdd, and G.S. Rawlins\*

University of Central Florida  
Electrical & Computer Engineering Department  
Orlando, FL 32816

\*Signal Technologies, Inc.  
385 Commerce Way, Suite 101  
Longwood, FL 32750

**ABSTRACT** -- A direct sequence spread spectrum (DSSS) receiver consists of three main component; a carrier loop for carrier phase acquisition, a code acquisition to line up the local and transmitted sequences within one chip, and a code tracking for fine tuning the alignment of the two sequences. This paper deals with the first two component of the DSSS receiver. A Costas phase locked loop (Costas PLL) is used as the major loop in which a novel code acquisition subsystem is nested. The Costas PLL, comprised of a number of analog components, is implemented with custom application specific integrated circuits(ASIC). Experimental results of a binary phase shifted keying (BPSK) system employing the ASIC Costas PLL are then presented for validation and performance evaluation of the ASIC chip design. The second part of this paper considers the coarse acquisition device nested within the loop. An Algebraic Despreader is proposed as an alternative to the conventional models currently used in DSSS systems. The design consists of digital components that conform readily to integrated technologies. The pseudo-noise (PN) sequence generated by the algebraic synchronizer proves to be an accurate estimate of the desired despreading sequence, to within one chip. A mathematical model is presented and shown to be consistent with the system's configuration. Curves for the probability of acquisition versus the number of clock pulses are generated through simulation and plotted for an eighth order system in various levels of additive white Gaussian noise (AWGN).

## I. INTRODUCTION

A DSSS signal is obtained when a PN sequence is used to modulate a much slower binary information signal before it is translated to a selected RF frequency by an appropriate carrier. At the receiver, similar operations are necessary to demodulate the incoming signal and retrieve the desired information. One arrangement commonly used to accomplish this is the Costas PLL. As shown in Figure 1, the Costas PLL works in concert with the embedded acquisition and tracking mechanisms to lock onto a signal and maintain synchronization. The difficulties of the system lie in the phase and frequency uncertainties between the transmitted and locally generated signals. The Costas PLL estimates the phase ambiguity of the carrier, while the code acquisition and tracking mechanisms estimate the PN sequences phase uncertainties. Since the focus of this paper is on the operations of the Algebraic Synchronizer nested in a Costas PLL, the fine tuning phase of the PN sequence, usually accomplished by a delay locked loop, is not addressed.

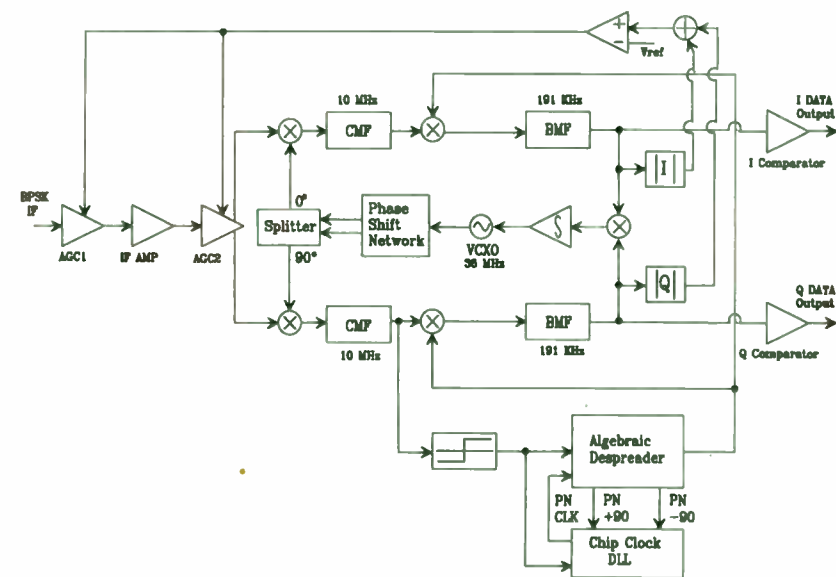


Figure 1: Costas Phase Locked Loop with Nested Acquisition and Tracking Devices.

There are several candidates for the code acquisition technique. One method uses the sliding correlator [1,2] which reduces the uncertainty region by moving the local PN sequence relative to the received signal by a fraction of a chip. This serial search method produces a high correlation value when the two sequences eventually overlap. While this approach is quite time consuming, there are variations that have improved performances. Another coarse acquisition technique for DSSS applications is the Rapid Acquisition Sequential Estimator (RASE), developed by Ward [3]. The RASE achieves synchronization by loading a portion of the transmitted sequence into shift registers of the local PN generator to induce a proper code alignment. As pointed out by Chiu and Lee in [4], however, the design does not lock onto inverted sequences (caused by data modulation) and should be replaced by the more efficient Improved Sequential Estimator (ISE).

A recent coarse acquisition technique that was first introduced in [5], and competes well with the above systems, is presented in this paper. It does not require feedback from a sync detector and operates successfully on inverted sequences. The system performs an estimate of the number of chips by which the received and locally generated PN sequences differ. This estimate is then used to synchronize the two platforms.

Section II describes the operations of the Costas PLL, suggests an ASIC implementation, and provides experimental results of a BPSK configuration. A description of the Algebraic Despreader is then given in section III, followed by a mathematical model of the design with a third order example in section IV. Section V sets up the test-beds and provides simulated results and performance curves for an eighth order system.

## II. COSTAS PHASE LOCKED LOOP

### 1. System Design

The Costas PLL of Figure 1 contains five four-quadrant multipliers. Two of the five four-quadrant multipliers are utilized for IF carrier stripping, two are employed for PN code stripping, and one is used as a phase detector to help create the loop error voltage as a control signal for an external voltage controlled oscillator (VCO). The loop also contains two operational amplifiers; one amplifier is used to process the error signal and the other is used in the power

detection loop for the automatic gain control (AGC). The dynamic response of the receiver is controlled by two independent AGCs. One AGC is based on energy contained within the spreading bandwidth while the other AGC depends on baseband energy at the Costas loop output. This separation of AGC functions permits the greatest flexibility for controlling the receiver's dynamic range, particularly for the cases of co-channel interference (CCI) and adjacent channel interference (ACI). The hybrid splitter is an active network which facilitates  $0^\circ$  and  $90^\circ$  local oscillator outputs to I and Q multipliers with a  $0^\circ$  local oscillator input. The circuit takes the single ended local oscillator input and produces differential outputs to the I and Q multipliers used for carrier stripping. Both I and Q arms of the Costas loop are equipped with absolute value circuits for power detection. These circuits are used in the AGC feedback loop to generate the energy signal for gain control. The output of the I and Q channels are followed by two comparators to facilitate the threshold decisions on the matched filtered data. Symbols from the in phase arm (I) of the Costas loop band matched filter (BMF) have the non-return to zero (NRZ) format.

The above Costas PLL has been designed and implemented with the CSL-100, an analog ASIC chip, developed and designed by Signal Technologies, Inc. The chip contains all of the above mentioned components, except for one AGC. Its circuit architecture is flexible enough to support a wide range of frequencies and circuit interconnectivity. The CSL-100 has a bias network that is used to bias the IF inputs of the Costas loop. The network uses a Darlington pair to provide the appropriate dc level at the IF inputs while providing isolation with low current demand from a bandgap reference, the dc bias source. The ASIC has a bandgap reference that is used to provide a temperature stable precision reference internally for biasing the multiplier current tails. External pin access points are provided to allow custom filtering, quadrature networks, bias networks, and other miscellaneous circuits to be connected to the ASIC.

Adding a threshold device allows for hard decision estimates from the chip matched filters (CMF) embedded in the Costas loop to be routed into an algebraic coarse synchronizer, which calculates the required PN synchronization phase.

## 2. Experimental Verification

The system of Figure 1 was designed and built and tested with the analog ASIC and supporting components to test the CSL-100 operating in the Costas PLL configuration. For simplicity, spreading and despreading was not included in the test setup. Figures 2 through 4 are oscilloscope traces of various signals for the BPSK transceiver experiment. The transmitted data at 191 Kbps is shown in Figure 2a. The PN data is then modulated by a 36 MHz RF carrier. The BPSK transmitted waveform is shown in Figure 2b.

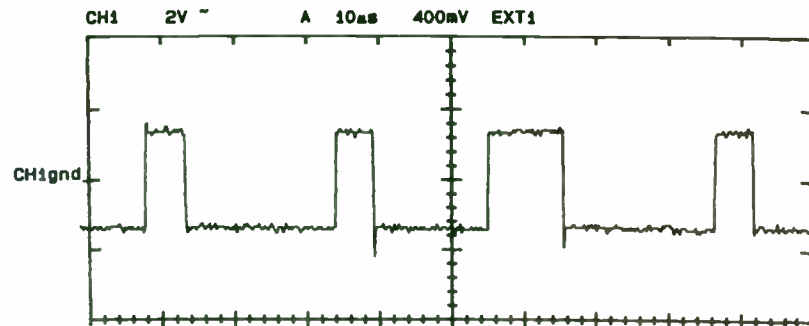


Figure 2a) Transmitted Data (191 Kbps).

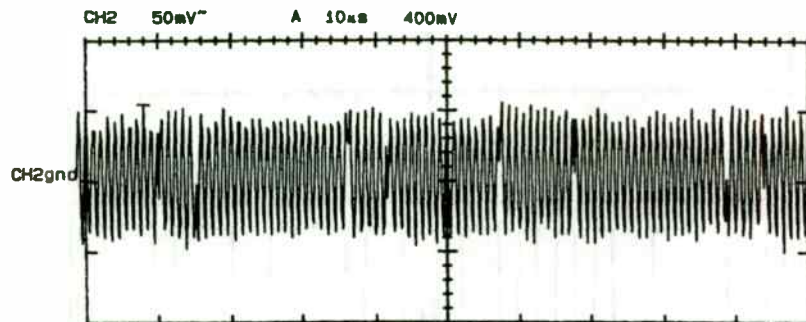


Figure 2b) BPSK Signal (36 MHz).

Additive White Gaussian Noise was then added to the received signal before entering the Costas PLL. Figure 3a is the oscilloscope trace of the transmitted data, while Figure 3b is the oscilloscope trace of the signal at the output of the BMF. Figure 4 then compares the reconstructed signal at the I comparator output to transmitted data. Except for a delay due to propagation, the signals are essentially equivalent. This experiment illustrates the capability of the CSL-100 in operating as a Costas PLL, with external support. The next set of tests for the CSL-100, would include the coarse acquisition device and the fine tuning mechanism. In the upcoming sections, focus is placed on the description, modeling, and simulation of the algebraic despreader.

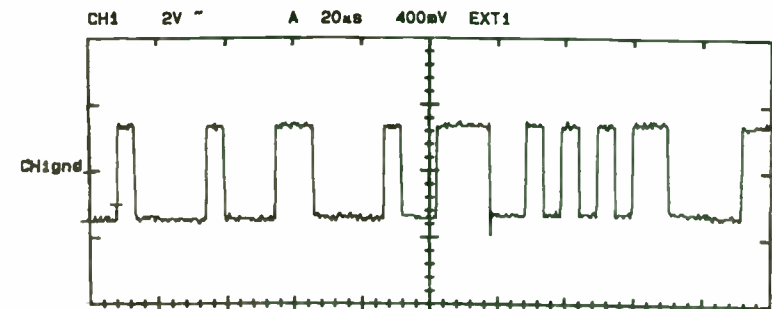


Figure 3a) Transmitted Data (191 Kbps).

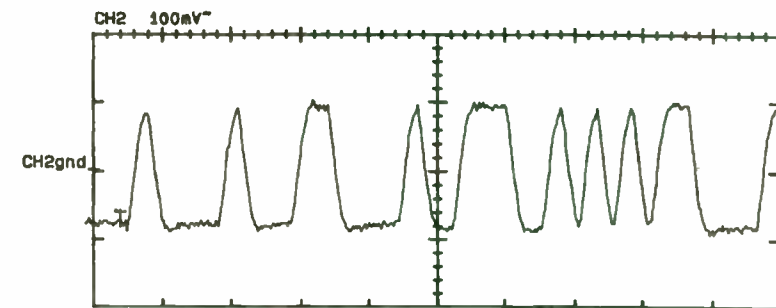


Figure 3b) Output of the Baud Matched Filter.

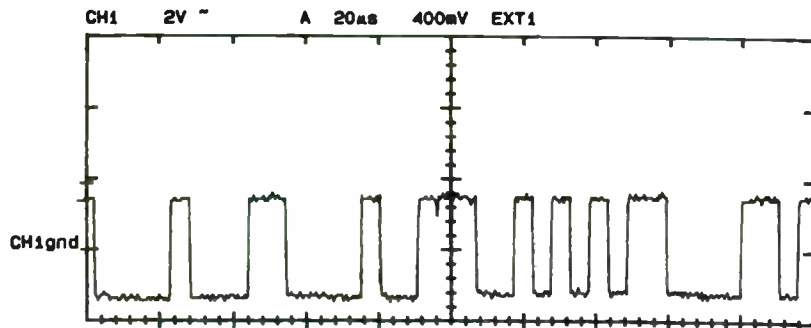


Figure 4a) Transmitted Data (191 Kbps).

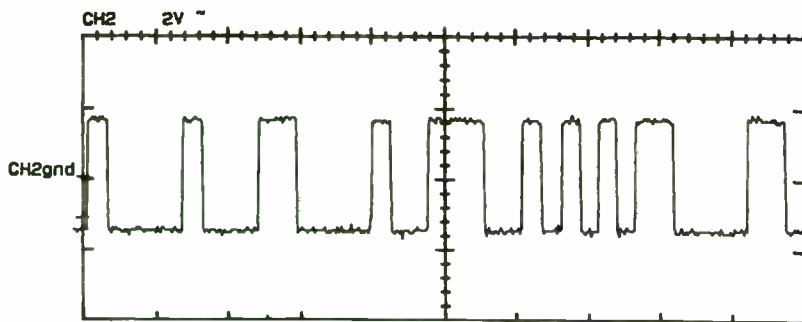


Figure 4b) Recovered Data (I Comparator Output).

### III. ALGEBRAIC DESPREADER DESCRIPTION

The algebraic despreader is a digital device that can be constructed with a pair of PN generators, two dot product calculators, and a smoothing filter. It can also be realized with a single digital ASIC and external components. The first PN generator of the despreader is used in conjunction with the first dot product operator and the incoming sequence to produce a codeword with phasing information related to the incoming sequence. This codeword is passed through the smoothing

filter and applied to the second PN generator via the second dot product operator to produce a sequence that is aligned with the incoming sequence. The smoothing filter improves the system's performance in the presence of noise.

Figure 5 is an upper level diagram of the algebraic despreader. Within each block are subordinate levels with components that perform logical operations. The overall purpose of this system is to strip off noise, data modulation, and threshold effects from the input signal to produce a resident PN sequence for the purpose of despreading. Refer back to Figure 1 for the location of this device in the Costas PLL structure.

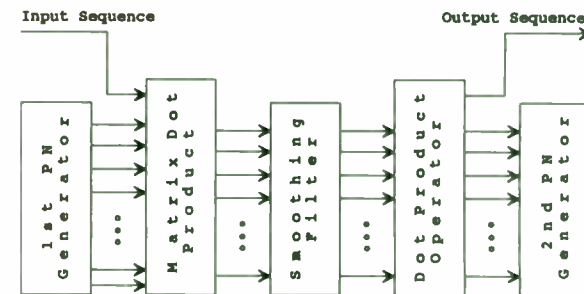


Figure 5: Algebraic Despreader.

The leftmost block of Figure 5 is an "extended", "reversed", and "modular" PN sequence generator. It is extended to operate on data-modulated sequences, reversed to assist the first dot product operator in producing constant phase vectors, and modular for efficient ASIC implementations. Figure 6 is a generalized PN generator of this type. The register contents of the generator are transported to the matrix dot product operator through parallel lines.

The next block is the dot product matrix. It performs multiple dot product calculations. Figure 7 illustrates only one of its  $n$  branches, where  $n$  is the order of the non-extended generator. The reason for  $n$  branches (instead of  $n+1$ ) is to minimize hardware while maximizing performance.



Data modulation is stripped off at this site. The other branches are identical to Figure 7, except for fewer delay elements. For example, the (i+1)th branch contains one less element than the ith branch. This shifting effect allows for uniqueness in the generation of adjacent bits of the codeword.

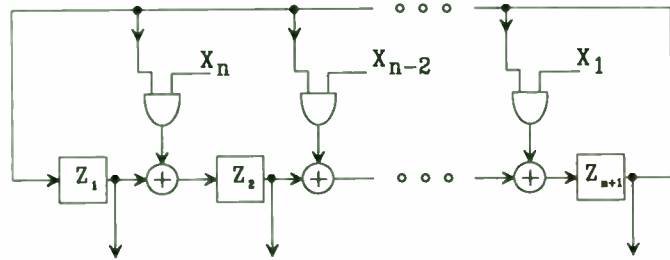


Figure 6: Extended, Reversed, and Modular PN Generator (First Generator).

As shown in Figure 7, the (n+1) register contents of the first local PN generator are ANDed with (n+1) consecutive samples of the input sequence. The AND gates are connected to an exclusive-or gate to complete the dot product operation. When all n parallel branches of the matrix dot product operator are sampled, a codeword is formed that stays constant in time, under noise-free conditions. Noise forces the codeword to temporarily change. For the nth order codeword, there are  $2^n - 1$  possible combinations, each representing a different phase.

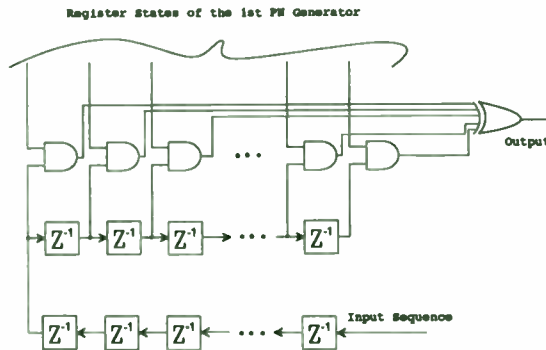


Figure 7: One Branch of the Matrix Dot Product Operator.

To stabilize the codeword, and therefore improve the performance of the system, a smoothing filter is employed between the two dot product blocks. The filter is made up of accumulators and data converters that build increasingly accurate estimates as a function of time. Figure 8 contains one branch of the n branch filter. The first block converts the binary valued data (+1 and 0) into numeric (+1 and -1) data. The running summer then adds these values as they emerge from the first block. This process is sometimes referred to as a "Random Walk", since the output of the summer walks up and down a staircase as prescribed by the input.



Figure 8: One Branch of the Smoothing Filter.

Since more correct chips than incorrect ones are expected, the summer builds a more accurate estimate as additional samples are processed. That is, the DC component at the input causes the output to eventually reach a steady state condition where all outputs are either positive or negative. The next block then converts these values into the binary format which prepares the codeword for the next stage.

Figure 9 is the dot product operator that forces the second generator to adjust its phase in relation to the input signal. The generator is of nth order to match the size of the codeword. It has a modular configuration but is not extended or reversed. It can be derived from Figure 6 if the order and feedback taps of the model are modified. All  $2^n - 1$  possible codewords are uniquely specified by this structure.

With the proper initial conditions, the output of the dot product operator is a PN sequence that can be used at the mixers (see Figure 1) to despread the incoming signal. Delay and advanced versions ( $\pm \frac{1}{2}$  chip) of the despread output can be sent to the tracking mechanism for fine tuning.

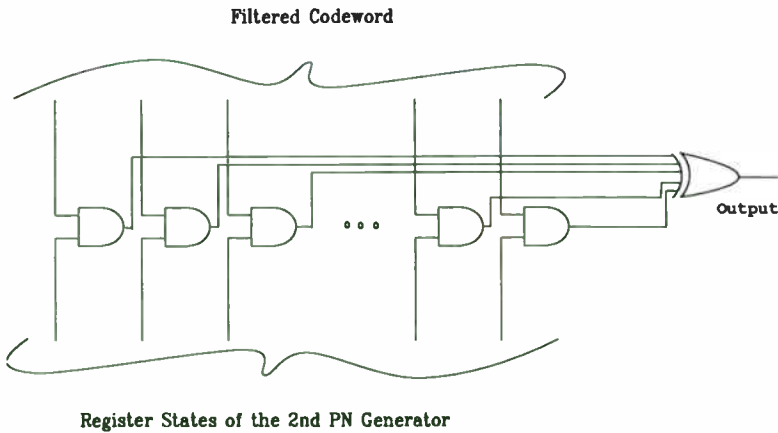


Figure 9: Dot Product Operator.

#### IV. MATHEMATICAL MODEL

As previously stated, the serial sequence entering the algebraic despreaders is potentially noisy, modulated by data, and quantized by a threshold device. Ignoring the carrier, the signal at the input to the receiver is modeled as:

$$Z(t) = s(t)a(t) + n(t) \quad (1)$$

where  $s(t)$  is the binary information,  $a(t)$  is the spreading code, and  $n(t)$  is the AWGN in the channel (see Figure 10). Sampling and quantizing this signal results in the sequence:

$$Z = Z_1 Z_2 Z_3 \dots Z_n \quad (2)$$

This sequence enters the matrix dot product operator through a set of shift registers, represented mathematically by the last row in the matrix of equation (3). A new chip enters the matrix from the lower left corner shifting the adjoining chips to the right. The other rows correspond to the other  $n$  branches of the  $(n+1)$  branch operator. The  $(i+1)$ th row differs from the  $i$ th row by only one shift. The chips exit the matrix through the upper right hand corner.

$$Z = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1(n+1)} \\ Z_{21} & Z_{22} & \dots & Z_{2(n+1)} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ Z_{(n+1)1} & Z_{(n+1)2} & \dots & Z_{(n+1)(n+1)} \end{bmatrix} \quad (3)$$

The first local PN generator of the receiver is extended, reversed, and modular [6]. Given that  $f(x)$  is the describing polynomial for the  $n$ th order generator, equation (4) shows the operation that is necessary to extend it.

$$e(x) = (1+x)f(x) \quad (4)$$

Reversing the extended equation is accomplished with the equation:

$$r(x) = x^{n+1}e(1/x) \quad (5)$$

Figure 6 is the modular arrangement, also referred to as the Galois implementation. The  $[(n+1) \times (n+1)]$  characteristic matrix,  $A_G$ , for this generator is given in equation (6) along with its state vector,  $P_G$ . The  $X$ 's represent feedback taps of the generator. The vector  $P_G$  serves as the initial conditions when the system starts up and as the present state of the generator while it is running. The multiplication of  $A_G$  with  $P_G$  produces a future state vector that can be applied to the  $Z$  matrix to produce a vector (codeword) that possesses valuable phasing information.

$$A_G P_G = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & X_{(n+1)} \\ 1 & 0 & 0 & \dots & 0 & X_n \\ 0 & 1 & 0 & \dots & 0 & X_{(n-1)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & X_1 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ \vdots \\ \vdots \\ P_{(n+1)} \end{bmatrix} \quad (6)$$

Combining the three matrices (via multiplication and the dot product operation) results in a vector codeword that stays constant for every clock pulse, provided no noise is present. If noise exists, the codeword changes with occurrence proportional to the dependent nature of the matrix dot product operator.

$$\begin{bmatrix} Z \end{bmatrix} \cdot \begin{bmatrix} A_G \end{bmatrix} \begin{bmatrix} P_G \end{bmatrix} = \begin{bmatrix} c \\ o \\ d \\ e \end{bmatrix} \quad (7)$$

The codeword is a vector of length (n+1) which reflects the extended operation that rids the incoming sequence of data modulation. At this point, the codeword can be truncated to reduce hardware overhead and to increase acquisition performance. It can be truncated to length n without loss of information. Physically, this is done by eliminating one of the (n+1) branches of the matrix dot product operator.

The codeword, assumed noisy, is passed through a smoothing filter. The filtered codeword is then applied to the dot product operator of Figure 9. It is used to adjust the phase of the second local PN generator. The [n x n] characteristic matrix of this generator,  $A_L$ , is given in equation (8) along with its present state vector,  $H_L$ .

$$A_L H_L = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & X_n \\ 1 & 0 & 0 & \dots & 0 & X_{(n-1)} \\ 0 & 1 & 0 & \dots & 0 & X_{(n-2)} \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 1 & X_1 \end{bmatrix} \begin{bmatrix} H_1 \\ H_2 \\ H_3 \\ \cdot \\ \cdot \\ \cdot \\ H_n \end{bmatrix} \quad (8)$$

The generator is operating with a random phase in relation to the input sequence, until struck by the codeword (via the dot product operator). At that time the proper phase, with respect to the incoming sequence, is forced. The final governing equation of the despreader is:

$$[c \ o \ d \ e] \cdot \begin{bmatrix} A_L \end{bmatrix} \begin{bmatrix} H_L \end{bmatrix} = [a] \quad (9)$$

where [a] represents a chip in the despreading code. Ensuing clock pulses produce subsequent chips in the PN sequence. The following example should clarify these ideas.

### Third Order Example

This example utilizes the characteristic polynomial  $f(x) = 1+x^2+x^3$ . Assuming that there is no channel noise and that the incoming sequence (after binary quantization) is 1110010 · · ·, the resulting Z matrix is:

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

Since the initial conditions of the transmitting PN generator are random, it is justifiably correct to assume that any propagation delay (in chip intervals) is absorbed by the Z matrix.

In developing the matrices for the first local generator, a couple of algebraic operations are necessary. Extending the original equation is performed as follows:

$$e(x) = (1+x)(1+x^2+x^3) = 1+x+x^2+x^4$$

The next step is to calculate the polynomial inverse:

$$r(x) = x^4 e(1/x) = 1+x^2+x^3+x^4$$

From this, the characteristic matrix,  $A_G$ , is easily obtained:

$$A_G P_G = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The initial conditions 1001 were purposely selected to contain an even number of ones. The resulting codeword is:

$$Z \cdot A_G P_G = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \text{CODEWORD}$$

The strength of this acquisition design lies within the integrity of this codeword. It must remain constant (relatively constant), throughout the entire communication process. Since noise causes it to pop in and out, a smoothing filter is used for stabilization. This example assumes that

no noise is present so there is no need for a filter. The next step is to truncate the codeword from 1110 to 111. The truncated codeword is then applied to the second local PN generator. It too is modular, but, is of order  $n=3$ . The characteristic matrix of this generator is:

$$A_L = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Strategically using the initial conditions  $H_L = 101$ , we find that:

$$[111] \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = [1]$$

In successive clock pulses,  $H_L$  is updated to future states while the characteristic matrix and codeword remain unchanged. Continued interactions in this manner produces the sequence: 1110010..., which is identical to the input sequence. It should be mentioned that improper selections of the initial conditions would have misaligned the local code relative to the incoming sequence by a constant phase shift.

### V. SIMULATION AND PERFORMANCE

Most coarse acquisition designs are evaluated by their speed in acquiring synchronization. The algebraic despreader is no exception. Before testing for speed, however, the system should be tested for validity. Figure 10 contains the model used for this test. The transmitter contains a binary data source that is spread by a higher speed PN sequence. AWGN is added in the channel. At the receiver, the incoming sequence is routed in two directions. One path leads directly to the mixer for despreading while the other to the cascaded limiter and algebraic despreader for local PN sequence generation. Figure 11 contains waveforms generated by an eighth order system. Signal 11a is the original data at the transmitter while signal 11b is the output of the local mixer. The similarities between the two signals indicate that the despreader is operating correctly.

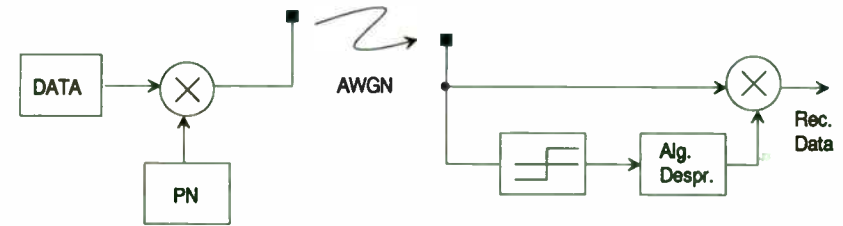


Figure 10: Test-bed to Verify Acquisition.



11a). Original Data



11b). Recovered Data

Figure 11: Example Waveforms.

To determine the speed of the system, another test-bed is used. This one simply monitors the filter outputs for the correct codeword (a stable codeword). Close inspection of the despreader design indicates that the point in time where the codeword becomes stable is the same time when synchronization is achieved. For simplicity and without loss of accuracy, data modulation was removed at the transmitter for these measurements. Here too, the channel was made up of AWGN. Figure 12 contains plots of the probability of a correct codeword versus the number of samples for various signal-to-noise ratios. As expected, the noisier environments slow down the acquisition process. With SNR = -1 dB, it takes approximately 700 samples (chips) to obtain synchronization with 90% certainty. For SNR = -3 dB and -5 dB, roughly 2200 and 7200 samples are required, respectively.

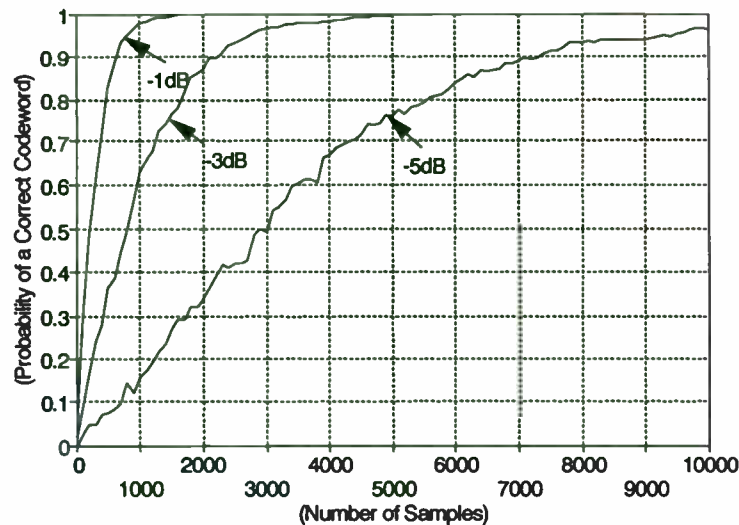


Figure 12: Acquisition Performance Curves.

## VI. CONCLUSION

The discussion started off with a review of the popular Costas PLL followed by a realization of the design using a compact and flexible analog ASIC. Experimental results yielded satisfactory performances. Focus then shifted to the coarse acquisition device. This novel new technique of algebraic despreading, embedded within the Costas PLL, was shown to be well suited for DSSS applications. A functional model was provided along with mathematical descriptions. Simulation results were then given for an eighth order example operating under different levels of input signal to noise ratios. As with the Costas PLL, a digital ASIC is available for this design.

## REFERENCES

- [1] G.G. Koller and M.A. Belkerdid, "Direct Sequence Spread Spectrum Simulation in a CDMA Environment with the Conventional Sliding Correlator," *Proceedings of the 24th Southeastern Symposium on System Theory*, Greensboro, North Carolina, March 1992.
- [2] R.L. Pickholtz, D.L. Schilling, and L.B. Milstein, "Theory of Spread-Spectrum Communications - A Tutorial," *IEEE Transactions on Communications*, Vol. COM-30, No. 5, May 1982.
- [3] R.B. Ward, "Acquisition of Pseudonoise Signals by Sequential Estimation," *IEEE Transactions on Communication Technology*, Vol. COM-13, No. 4, December 1965.
- [4] J-H.Chiu and L-S. Lee, "An Improved Sequential Estimation Scheme for PN Acquisition," *IEEE Transaction on Communications*, Vol. 36, No. 10, October 1988.
- [5] P.H. Halpern and P.E. Mallory, "Apparatus for Synchronizing to a Linear PN Sequence," US Patent 4,649,549, March 10, 1987.
- [6] J.K. Holmes, *Coherent Spread Spectrum Systems*, John Wiley & Sons, 1982.



## COSINE TRANSITION-SHAPING PSK MODULATORS

### USING DIRECT-DIGITAL SYNTHESIZERS

Roy Greeff and Bruce Williams

Paramax Systems Corporation

640 N. 2200 W. MS-F1E03

Salt Lake City, Utah 84116-2988

#### Introduction

The transmission of digital data has long been in vogue for its many advantages, e.g., ease of multiplexing, ability to error detect and correct, etc. A common major constraint placed on these (or any) transmitted signals is bandwidth. Power efficient techniques like bi-phase shift keying (BPSK) and quadrature phase shift keying (QPSK) are ideally infinite bandwidth. Therefore, systems designed around these modulation techniques, require filtering to limit spectral occupancy. Unfortunately, unless one is very careful, filter distortion effects can degrade system performance. Filter design techniques that mitigate these effects have been discussed at great length in the literature. With the advent of the direct-digital synthesizer, additional techniques have become available that greatly simplify the design of band-limited PSK modulators. This paper will discuss an easy implementation for cosine transition-shaping BPSK and QPSK modulators using direct-digital synthesizers (DDS).

#### Filter Effects

We begin our discussion of filter effects with a description of the spectrum of non-return-to-zero (NRZ) data. NRZ data has the well known "sinc-squared" normalized spectrum

$$S_{NRZ}(f) = \text{sinc}^2(fT) = \left[ \frac{\sin(\pi fT)}{\pi fT} \right]^2,$$

where  $T$  represents a bit duration, the inverse of the bit rate.

A baseband representation of the relative power spectral density is shown in figure 1.

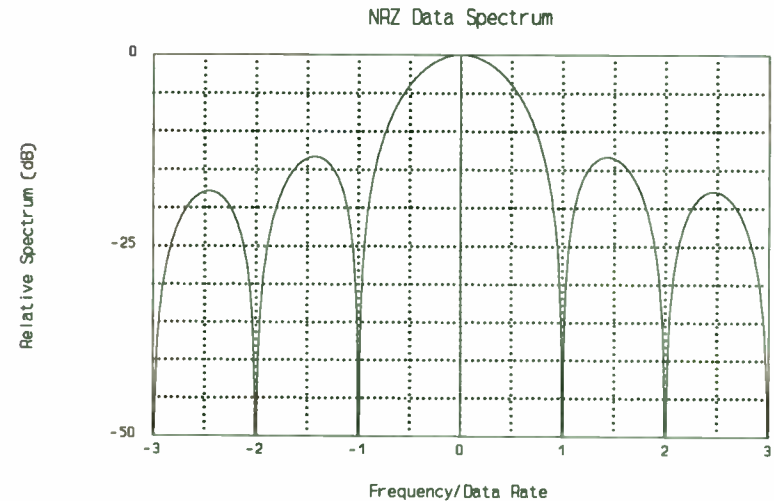


Figure 1, Normalized NRZ Data Spectrum

Non-bandlimited NRZ signals have 90.3 percent of the total energy contained in the main lobe, This is also referred to as the null-to-null band (the area between normalized frequency of -1 and 1). The center half of the null-to-null band contains 77.4 percent of

the energy. Systems, typically, employ filters to minimize the power in the sidelobes. Filters having impulse-responses of duration greater than the symbol rate (virtually all do!) cause intersymbol interference (ISI).

Additional degradations can occur when filtered PSK signals are passed through nonlinear devices. While non-band-limited PSK signals are constant envelope (magnitude of the carrier remains constant), the band-limited signals are not. When non-constant envelope signals are passed through a nonlinear devices, such as a power amplifier operating near saturation, sidelobes that had previously been suppressed can be regenerated to near pre-filtering levels. This can necessitate the addition of RF filters after power amplifiers.

#### IF filtering vs Baseband filtering

Figure 2 shows the two conventional ways that transmitted signals are often filtered. Baseband filtering has some distinct advantages over IF or RF filtering. Baseband filters allow one to change IF (or RF) frequencies without additional or new filters. Also, because these filtering schemes are done digitally at baseband, the accuracy of the filters can be significantly better than equivalent filters at IF or RF. In addition, baseband filters can also be easily changed allowing multiple data rates with only a firmware or software modification. Finally, baseband filters are easier to construct and lend themselves to digital signal processing techniques, i.e., digital filters. Unfortunately, for baseband filters to be effective one must use linear modulators (multipliers). Traditionally, nonlinear mixers are used for the multiplication function, and IF (or RF) bandpass filters are used

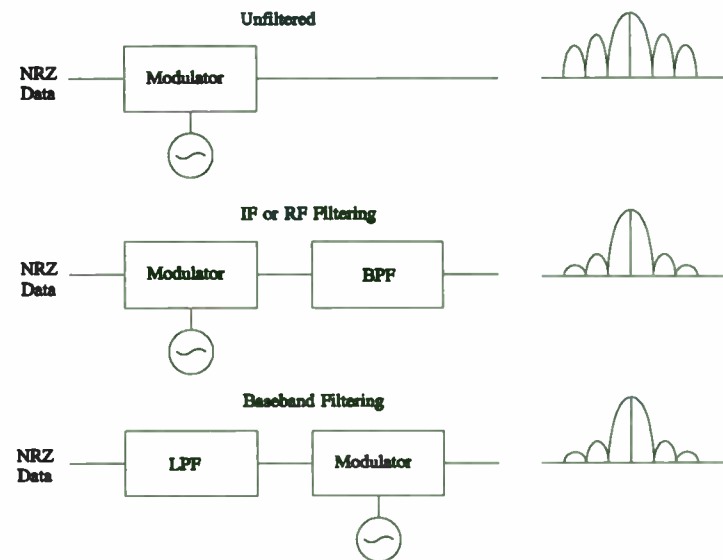


Figure 2, Filtered PSK modulation

to filter the transmitted spectrum. By utilizing direct digital synthesizers, which provide direct control over phase, frequency, and amplitude, one can create an effectively filtered PSK modulated signal. Baseband filtering can be done by controlling the output phase and amplitude of an NCO (numerically controlled oscillator). An alternative approach is to modify an NRZ data transition shape by controlling the NCO frequency; this technique creates continuous phase frequency shift keying (CP-FSK).



### Cosine Transition-Shaping PSK

By controlling the transition shape of an NRZ data stream at baseband, one can effectively filter the resultant modulated IF signal. Feher<sup>1</sup> discusses nonlinear switched filters that impose cosine transition-shaping on NRZ signals at baseband. Such filtered signals have advantages over conventionally filtered NRZ data in that these "... bandlimited signals that do not contain ISI [intersymbol interference] and are jitter-free...". Figure 3 is a block diagram of our modulator. Two CP-FSK modulators are used to create cosine transition-shaped BPSK at carrier frequency  $f_c$ . Figure 4 compares the relative spectral density envelopes of NRZ BPSK versus transition-shaped BPSK. Figure 5 displays a sample of the various baseband signals involved and a phase constellation.

In a typical BPSK system, the phase of the carrier is either  $0^\circ$  or  $180^\circ$  depending on the NRZ data input. A change in data state results in an instantaneous change in carrier phase (in an ideal, non-bandlimited system). Feher's nonlinear filter created cosine-shaped transitions by letting the phase change linearly  $180^\circ$  during 1 symbol period if a data transition occurred. Our approach is to create this phase change by frequency shift keying with an offset frequency ( $\Delta f$ ) equal to  $1/2$  of the data rate when a transition is sensed. This way a  $180^\circ$  transition occurs in each CP-FSK modulator. Two modulators are necessary because energy is mapped into a quadrature channel for a single CP-FSK modulator (see figure 5,  $Q_1$  or  $Q_2$ ). By forcing one CP-FSK modulator to a positive offset frequency and the other to a negative offset frequency, the resulting quadrature channel

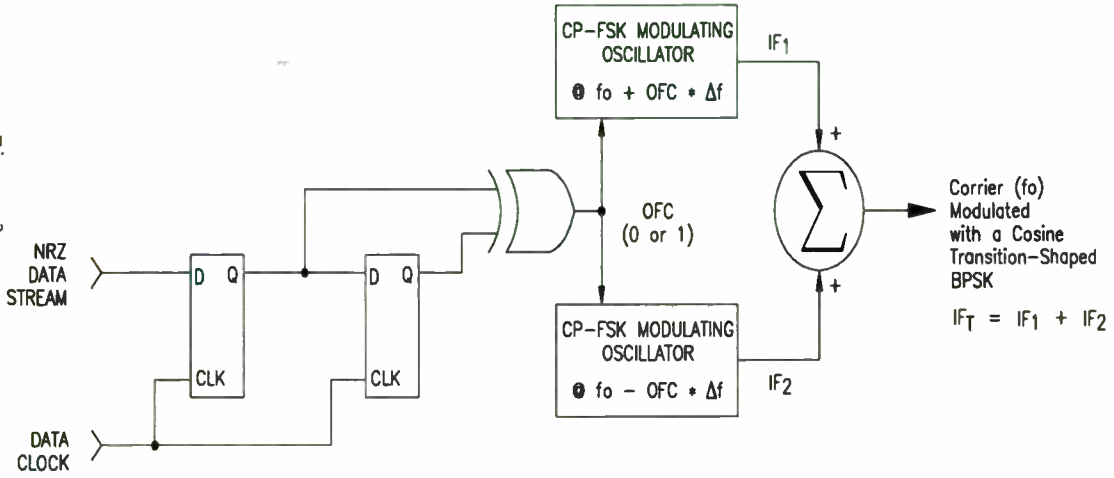
in the summation is zero ( $Q_1 + Q_2 = 0$ ). Energy is present only in the In-phase channel,  $I_c$ . When no transitions are present both CP-FSK modulators operate at the carrier frequency,  $f_c$ . The offset frequency control signal is generated by differentially encoding the NRZ data.

---

<sup>1</sup> Dr. Kamilo Feher: *DIGITAL COMMUNICATIONS Satellite/Earth Station Engineering*, Section 3.5.3, pp 118-119, Prentice-Hall 1983

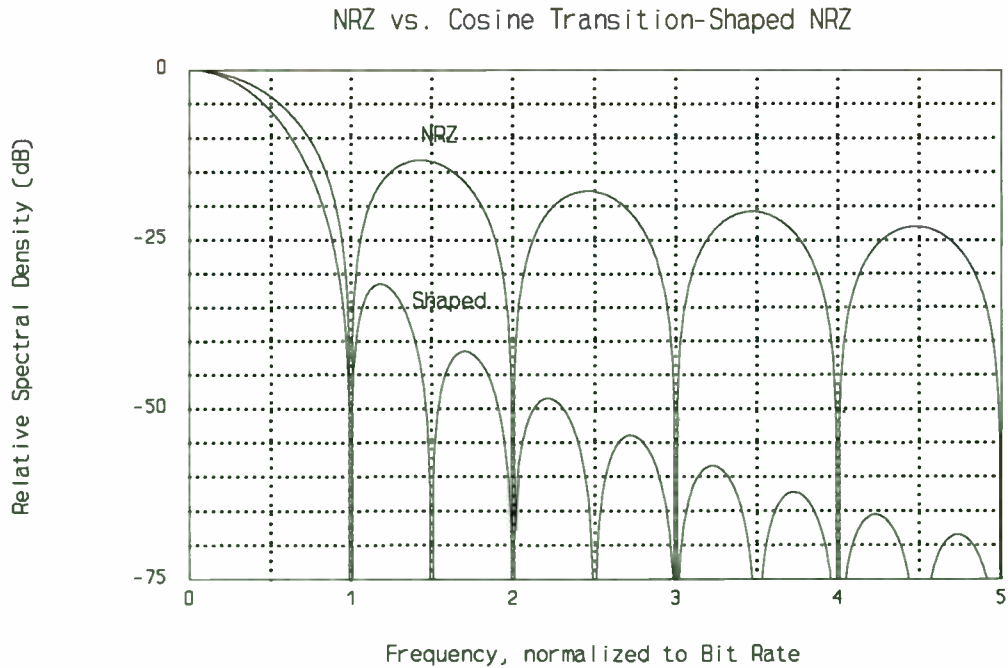
# Cosine Transition-Shaping BPSK Modulator

Figure 3



Note: \*  $\Delta f = \text{Data Rate} / 2$ , for max. filtering  
 \* OFC is the Offset Frequency Control  
 \*  $f_0$  is the carrier frequency

Figure 4



# BASEBAND SIGNAL REPRESENTATIONS

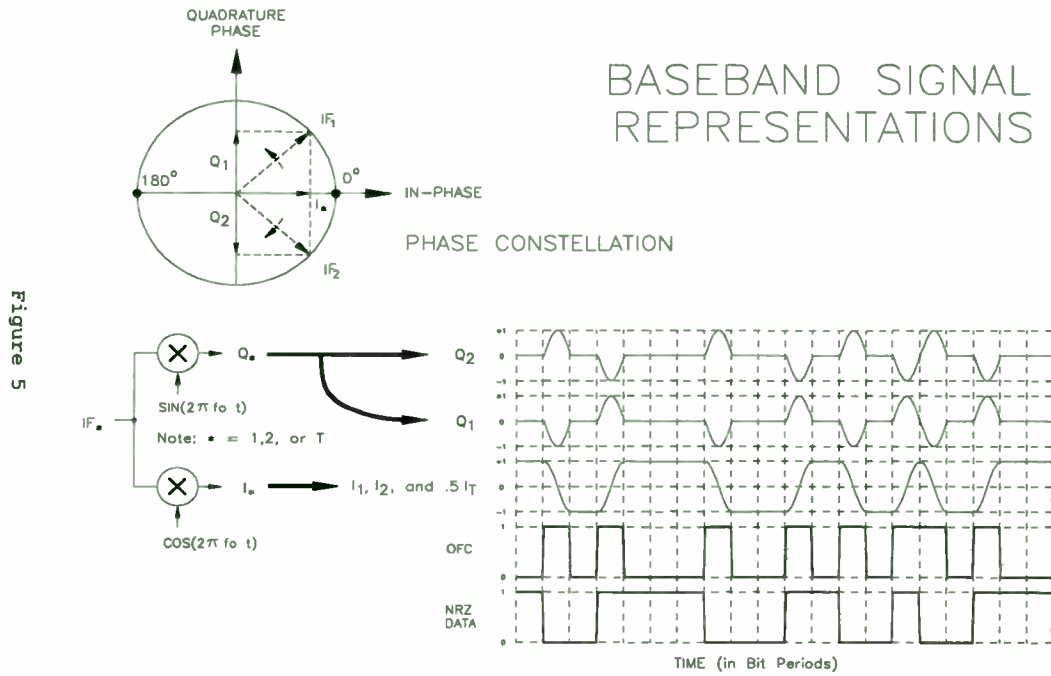
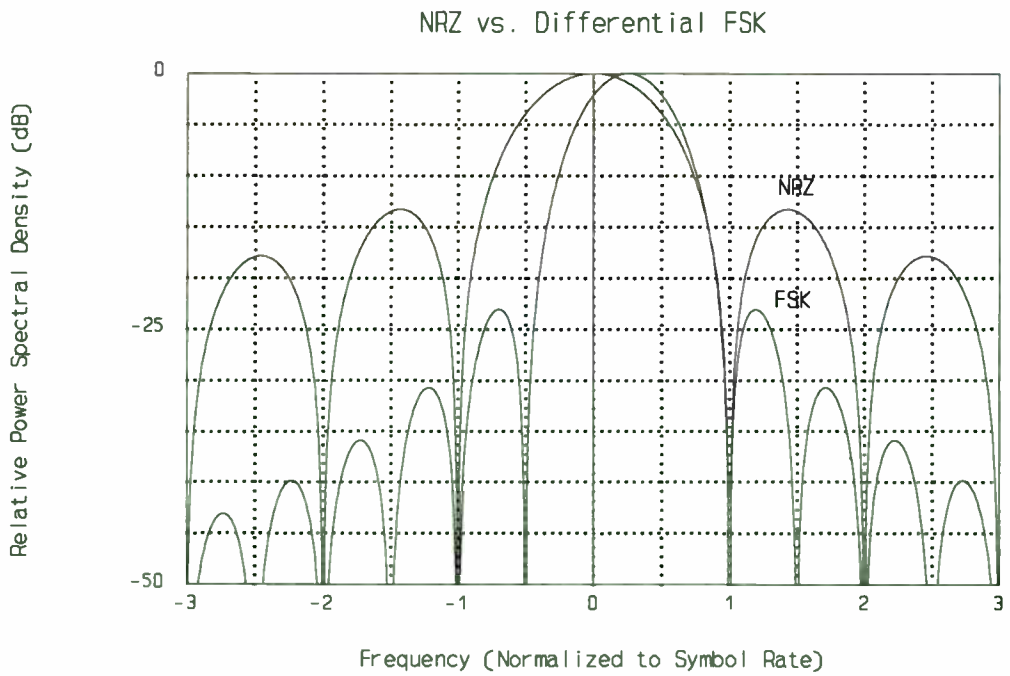


Figure 6



### Direct Digital Synthesizer's as PSK Modulators

PSK modulation using direct digital synthesizer (DDS) can be implemented by changing the phase, amplitude or frequency of the carrier. Modulation using these three methods will be examined for filtered BPSK.

For BPSK, when the data is a '0' the carrier's phase is 0 degrees. When the data is a '1' the carrier's phase is changed to 180 degrees. Using a DDS, this phase change can be easily done by changing the most significant bit (MSB) on the phase being input to the digital-to-analog converter (DAC). When the phase is changed instantaneously in this manner, the carrier's envelope is constant (see figure 7(a)). This type of modulator is superior to an analog modulator in that the phase imbalance and amplitude imbalance is virtually zero for the two phases. When this type of modulator is filtered, the phase transition time becomes finite (see figure 7(b)) and no longer has a constant envelope. If a DDS allows access to all the output phase bits, the phase can be changed in discrete steps. An example is shown on the IQ phase diagram in figure 7(c) as moving from 180 degrees to 0 degrees in a finite period of time and following a semicircle. The time required to go from 180 degrees to 0 degrees is a function of the type of filter used. More filtering occurs for slower transition times than occurs for fast transition times. This baseband filtered modulator has a constant envelope. At first glance this seems a perfect way to change the phase. The problem is that when the phase is changed in this way some energy appears in the Q vector while the phase is changing. This may make it more difficult to obtain the optimum matched filter in the associated receiver.

Amplitude modulation can also be used to create a BPSK modulator. This is done by simply multiplying the carrier by -1 for a '0' and +1 for a '1'. In a DDS NCO, this multiplication can be done digitally before feeding the DAC. Implementation of a baseband filter for this type of modulator can be accomplished by simply multiplying the carrier by discrete amplitudes varying from -1 to +1. This can be shown on the IQ phase diagram in figure 7(b) as moving from 180 degrees to 0 degrees in a finite period of time and passing through the origin (or zero carrier amplitude). The time required to go from 180 degrees to 0 degrees is a function of the type of filter used. More filtering occurs for slower transition times than occurs for fast transition times. This baseband filtered modulator does not have a constant envelope.

By coherently changing the frequency (CP-FSK), a baseband filtered BPSK modulator can be built. This is accomplished by changing the carrier frequency by a delta  $f_t$  for a period of time  $T_t$  until the phase relative to the original carrier frequency changes by 180 degrees. The frequency is then changed back to the original carrier frequency. The relationship between  $f_t$  and  $T_t$  is as follows:

$$f_t = 1/2 * T_t$$

$$f_t = \text{transition frequency}$$

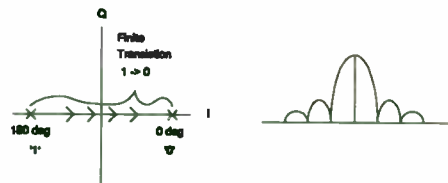
$$T_t = \text{transition time}$$

Every time the data changes from '1' to '0' the transition frequency  $f_t$  is added to the carrier for time  $T_t$ . When the data changes from '0' to '1',  $f_t$  is subtracted from the carrier for time  $T_t$ . If the transition frequency is very high, unfiltered BPSK can be simulated.

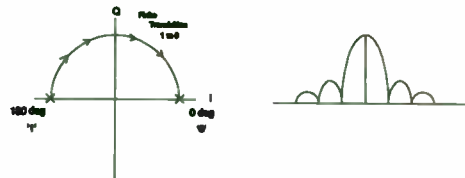
Unfiltered BPSK



Filtered BPSK



Filtered BPSK  
Constant Envelope



Filtered BPSK  
Using CFSK

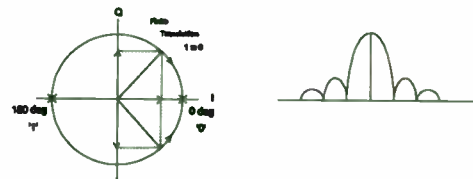


Figure 7, IQ phase diagrams

As the transition frequency becomes lower, more filtering is done. This baseband filtered modulator has a constant envelope and is similar to the phase modulated case.

Using DDS technology there are advantages to each of the above baseband filtering techniques. Direct control of the phase or amplitude of the NCO gives the most flexibility in controlling the filtered response. The cost paid is the resolution of the transition steps. Depending on the data rate, only a finite number of phase steps (dependent on how often the phase can be changed in the NCO) can be made. The same restriction can occur for the amplitude modulator. By using CP-FSK the phase can be changed with the finest resolution possible ... i.e., the phase is changed continuously. The restriction is that the shape of the transition is required to be a cosine shape. Another advantage of the CP-FSK modulator is the inherent high speed capability using only a simple NCO. At present DDS NCOs that operate with clocks in excess of 1 GHz do not allow control of all the phase bits. Using CP-FSK, only the capability to change frequencies rapidly is required to make a baseband filtered modulator.

Figure 8 shows a block diagram of a modulator implemented using CP-FSK to generate BPSK. A Stanford TeleCom STEL-1130 is used to add the outputs of two STEL-1175 running in parallel. It would also be possible to have a DAC on each NCO and then add the analog outputs. A microprocessor is used to set the output frequency and also set  $f_c$  and  $T_f$  for the filtering desired. Although this modulator has been designed for filtered BPSK, with minor modifications other modulating schemes can be done e.g., MSK, QPSK, filtered QPSK etc.

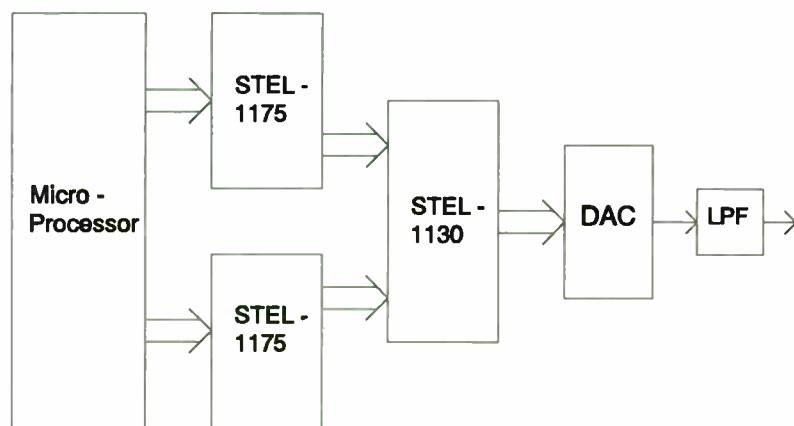
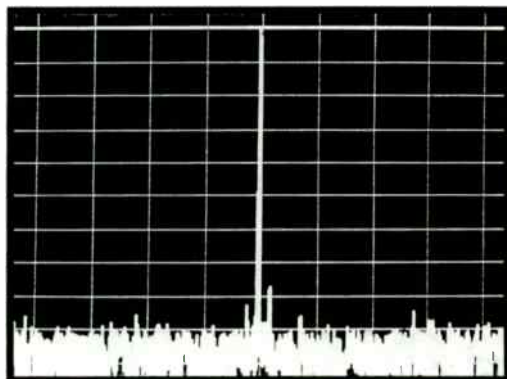


Figure 8, Block diagram of DDS Modulator

# APPLICATION NOTE AN-4

## UNDERSTANDING DATA CONVERTERS' FREQUENCY DOMAIN SPECIFICATIONS



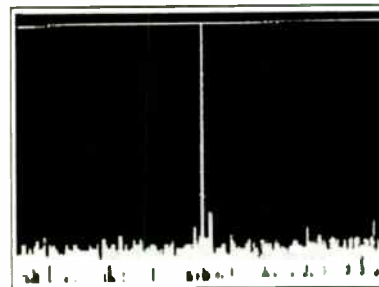
Bob Leonard  
Product Marketing Manager  
August, 1991



*Innovation and Excellence*

**Welcome to...**

UNDERSTANDING DATA CONVERTERS'  
FREQUENCY DOMAIN SPECIFICATIONS



Bob Leonard  
Product Marketing Manager



INNOVATION AND EXCELLENCE IN PRECISION DATA ACQUISITION

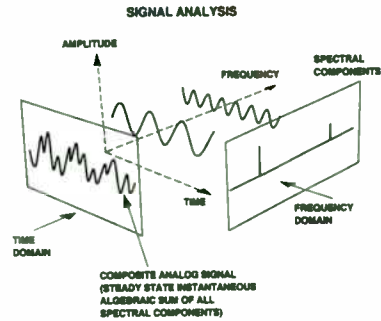
- **DATA CONVERSION COMPONENTS**
  - A/Ds
  - SAMPLING A/Ds
  - SAMPLE-HOLDS
  - D/As
  - MULTIPLEXERS
  - AMPLIFIERS
  - FILTERS
  - HDASs
- **DATA ACQUISITION BOARDS**
  - VME
  - PC/AT
  - MULTIBUS
- **PANEL PRODUCTS**
  - DIGITAL PANEL METERS
  - PROCESS MONITORS
  - PRINTERS
  - CALIBRATORS
- **POWER PRODUCTS**
  - DC-DC CONVERTERS
  - AC-DC CONVERTERS

DATEL was founded in 1970 as a producer of high performance Analog-to-Digital Converters and Data Acquisition products. Our administrative offices, engineering, modular and sub-system production facilities and hybrid production facilities qualified to MIL-STD-1772, are housed in our 180,000 square foot facility in Mansfield, Massachusetts.



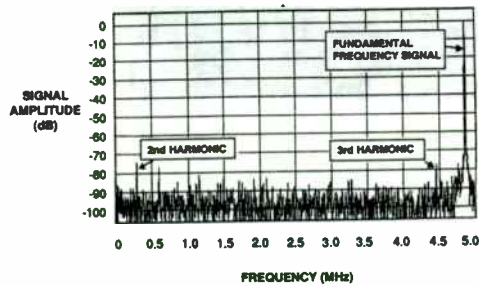
*Innovation and Excellence*

## TIME DOMAIN VERSUS FREQUENCY DOMAIN



Time domain data collected from an A/D converter is mapped into the frequency domain using the Fast Fourier Transform (FFT) algorithm.

## FFT PLOT OF ADS-130



A non-ideal FFT plot of an actual Sampling A/D illustrates the input frequency, harmonics generated and the noise floor. A single-tone frequency of 4.85 MHz was the input for this 12-bit, 10 MHz device.

## FOURIER TRANSFORM FOR A/Ds

- THE FOURIER TRANSFORM IS INTENDED TO OPERATE ON CONTINUOUS "WAVEFORM" DATA FROM  $-\infty$  TO  $+\infty$ .

$$X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft} dt \quad \text{WHERE } \omega = 2\pi f$$

- DISCRETE FOURIER TRANSFORM (DFT) IS USED ON SAMPLED A/D DATA. THE IDEAL CONTINUOUS WAVEFORM FROM  $-\infty$  TO  $+\infty$  HAS BEEN REPLACED WITH SAMPLED POINTS ON A WAVEFORM FOR A LIMITED TIME PERIOD.

$$X_D(n\Delta f) = \sum_{k=0}^{N-1} x(k\Delta t)e^{-j2\pi kn\Delta f \Delta t}$$

WHERE: N = TOTAL # OF POINTS IN THE RECORD  
nΔf = FINITE # OF FREQUENCY POINTS  
Δt = SAMPLING INTERVAL  
k = INTEGER

- THE FAST FOURIER TRANSFORM (FFT) ALGORITHM IS USED IN IMPLEMENTING THE DISCRETE FOURIER TRANSFORM DUE TO THE FFT'S MATHEMATICAL EFFICIENCY.

The Fast Fourier Transform is a mathematically efficient algorithm to supplant the Discrete Fourier Transform. Likewise the DFT is used to supplant the Continuous Fourier Transform for time sampled data. In principle, most aspects of the CFT transfer to the DFT and FFT, however there are nuances that demand attention.

N data points in the time domain produce N/2 points (amplitude and phase) in the frequency domain.

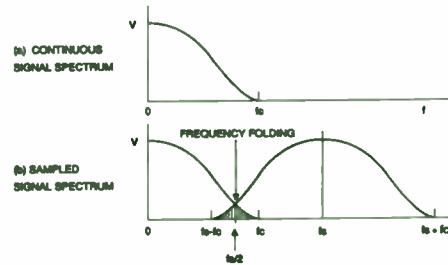
## FFT VERSUS DFT PROCESSING

# OF POINTS	FFT IMPROVEMENT
N	$N^2 / N \log_2 N$
128	x 18.2
256	x 32
512	x 56.9
1024	x 102.4
2048	x 188.2
4096	x 341.3
8192	x 630.2

The Fast Fourier Transform (FFT) requires  $N \log_2 N$  operations (multiplication & addition) while the Discrete Fourier Transform (DFT) requires  $N^2$  operations.



### FREQUENCY SPECTRA DEMONSTRATING THE SAMPLING THEOREM



The Nyquist Sampling Theorem requires that a continuous bandwidth-limited analog signal, with frequency components out to  $f_c$ , must be sampled at a rate  $f_s$  which is a minimum of  $2f_c$ .

If the sampling frequency  $f_s$  is not high enough, part of the spectrum centered about  $f_s$  will fold over into the original signal spectrum (frequency folding). Frequency folding can be eliminated in two ways: first by using a high enough sampling rate, and second by filtering the signal before sampling to limit its bandwidth to  $f_c/2$ .

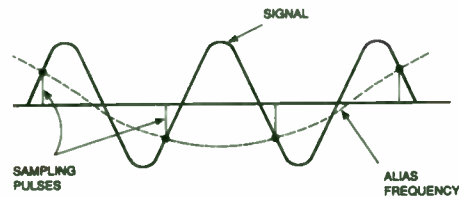
### FAST FOURIER TRANSFORM WEAKNESSES/CURES

- ALIASING / NYQUIST SAMPLING
- LEAKAGE / WINDOWING
- PICKET-FENCE EFFECT / # OF FFT POINTS

Successful application of the FFT demands an appreciation of three basic limitations; aliasing, leakage and picket-fence effect.

Observation of the Nyquist Sampling rate, utilizing windows to weight the non-infinite sequenced data and choosing an appropriate number of FFT points provide the appropriate solutions to these limitations.

### ALIAS FREQUENCY CAUSED BY INADEQUATE SAMPLING RATE



An inadequate sampling rate has the effect of producing an alias frequency in the recovered signal. Sampling at a rate less than twice per cycle results in an alias which is significantly different from the original frequency.

### FREQUENCY DOMAIN SPECIFICATIONS

- SIGNAL-TO-NOISE RATIO & DISTORTION (SINAD)
- SIGNAL-TO-NOISE RATIO WITHOUT DISTORTION
- TOTAL HARMONIC DISTORTION
- IN-BAND HARMONICS
- SPURIOUS FREE DYNAMIC RANGE (SFDR)
- TWO-TONE INTERMODULATION DISTORTION
- NOISE POWER RATIO (NPR)
- EFFECTIVE BITS

Some key Frequency Domain specifications for Sampling A/D converters are listed. Understanding how these are defined and under what conditions is as important as knowing the FFT pitfalls and cures.

## FREQUENCY FOLDING AND ALIASING

### PROCEDURE:

- ANALYZE THE INPUT FREQUENCY AS:  $F_{in} = K(F_s/2) \pm \Delta F$

### WHERE:

$F_{in}$  = INPUT FREQUENCY  
 $F_s$  = SAMPLING RATE  
 $K$  = ODD OR EVEN INTEGER (MULTIPLE OF HALF THE SAMPLING RATE)  
 (NEED TO DETERMINE  $K$  WHEN SUBSTITUTING INTO THE FORMULA)  
 $\Delta F$  = DIFFERENTIAL CHANGE IN FREQUENCY NEEDED TO EQUATE THE FORMULA  
 (NEED TO DETERMINE  $\Delta F$  WHEN SUBSTITUTING INTO THE FORMULA)

- IF  $K$  IS ODD: ALIAS FREQUENCY =  $(F_s/2) - \Delta F$
- IF  $K$  IS EVEN: ALIAS FREQUENCY =  $\Delta F$

The aliasing formulas are useful in determining where a harmonic will alias back into the signal spectrum. Conversely, a harmonic or spurious frequency can suggest possible frequencies that caused them. An example could be a system clock operating at a much higher frequency appearing as an alias in the signal spectrum.

An initial disconcertment over two unknowns and one formula is reduced once familiarity with the substitution process is practiced.

## FREQUENCY FOLDING AND ALIASING

### EXAMPLE:

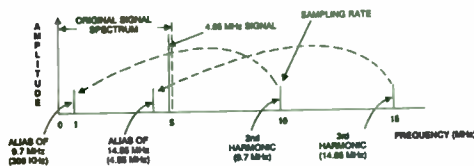
ANALYZE WHERE THE 2nd HARMONIC OF A 4.85 MHz SIGNAL WILL ALIAS WHEN DIGITIZED WITH A 10 MHz SAMPLING RATE.

$$F_s = 10 \text{ MHz} \quad F_n = 2\text{nd HARMONIC} = 9.7 \text{ MHz} \quad F_s/2 = 5 \text{ MHz}$$

SUBSTITUTING INTO THE FORMULA  $F_n = K(F_s/2) \pm \Delta F$

ENABLES THE DETERMINATION THAT:  $K = 1$  (ODD) AND  $\Delta F = 4.7 \text{ MHz}$

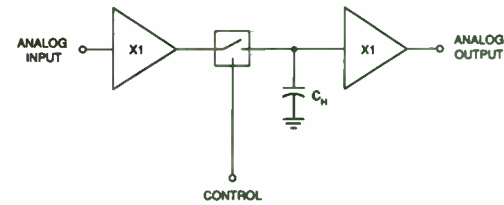
THEREFORE ALIAS FREQUENCY =  $(F_s/2) - \Delta F = 300 \text{ KHz}$



Utilizing the alias formulas, a 2nd harmonic (9.7 MHz) of a 4.85 MHz signal frequency when sampled at 10 MHz will appear as an alias frequency of 300 KHz on an FFT plot.

Similarly, a 3rd harmonic of the 4.85 MHz signal (14.55 MHz) would then yield  $K = 2$  (even) and  $\Delta F = 4.55 \text{ MHz}$ . The alias frequency would appear as  $\Delta F$ , or 4.55 MHz, on the FFT plot.

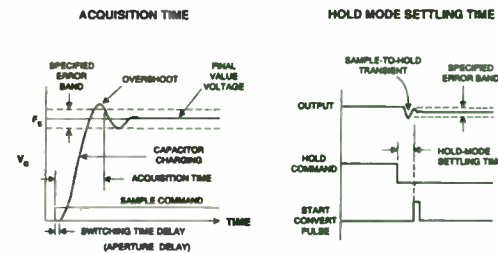
## "CLASSIC" SAMPLE-HOLD



The sampling process begins for many applications with the Sample-Hold (S/H) in front of the Analog-to-Digital Converter.

The "classic" open loop follower sample-hold architecture has a buffer in front of the switch to quicken capacitor charging and gives the S/H a high input impedance. Adding a buffer behind the hold capacitor reduces capacitor charge bleeding and output droop.

## KEY SAMPLE-HOLD SPECIFICATIONS

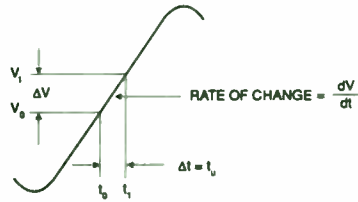


Among the key specifications for a Sample-Hold are Acquisition Time and Hold Mode Settling Time.

Acquisition Time of a S/H starts when the sample command enters and stays in the error band. Acquisition time is defined for a full-scale voltage change, measured at the hold capacitor.

When an A/D converter follows a S/H, the start conversion pulse must be delayed until the output of the S/H has had enough time to settle within the error band and stays there.

## APERTURE UNCERTAINTY



The actual voltage digitized by the data converter depends on the input signal slew rate and the "aperture time." The "aperture time" is application and architecture dependent.

The "aperture time,"  $\Delta t$ , could be the conversion time of the A/D converter used without a Sample-Hold. The "aperture time" could also be the aperture delay time specification of a Sample-Hold. Applications with repetitive sampling, such as those utilizing the FFT, can use the aperture uncertainty ( $t_a$ ) specification. The aperture delay is just phase information which is not required in identifying the frequency components.

## FREQUENCY DOMAIN SPECIFICATIONS

- SIGNAL-TO-NOISE RATIO & DISTORTION (SINAD)
- SIGNAL-TO-NOISE RATIO WITHOUT DISTORTION
- TOTAL HARMONIC DISTORTION
- IN-BAND HARMONICS
- SPURIOUS FREE DYNAMIC RANGE (SFDR)
- TWO-TONE INTERMODULATION DISTORTION
- NOISE POWER RATIO (NPR)
- EFFECTIVE BITS

Some key Frequency Domain specifications for Sampling A/D converters are listed. Understanding how these are defined and under what conditions is as important as knowing the FFT pitfalls and cures

## DETERMINING MAXIMUM FREQUENCY FOR SINUSOIDAL INPUTS

GIVEN:  $E = [dV/dt]T$

WHERE:  $E =$  VOLTAGE ERROR TOLERABLE  
 $dV/dt =$  SIGNAL SLEW RATE  
 $T =$  "APERTURE TIME"

SINUSOIDAL SIGNAL FORM:  $V_{in} = V \cdot \sin(2\pi ft)$   
 TAKING THE 1ST DERIVATIVE:  $dV/dt = (2\pi f)V \cos(2\pi ft)$   
 AT ZERO CROSSING:  $t = n/2f$  ( $n = 1,2,3\dots$ )  
 YIELDING:  $dV = (2\pi fV)dt$

MAXIMUM FREQUENCY  $f = \frac{dV}{dt(2\pi)V}$

The Sample-Hold ideally stores an instantaneous voltage (sample value) at a desired instant of time. The constraint on this time is aperture uncertainty for sinusoidal waveforms.

Noise affecting the logic threshold level of the sample-hold command creates a timing uncertainty in the switch and driver circuit (aperture uncertainty). The resulting error voltage for a given period of time increases with the rate of change of the input signal.

Determining the maximum frequency to digitize a  $\pm 10V$  sinusoid to  $\pm 1/2$  LSB accuracy to 12-bits yields (for aperture uncertainty = 50ps):

$$f = \frac{2.44 \text{ mV}}{(50\text{ps})2\pi 10V} = 776 \text{ KHz}$$

## LOGARITHMS

$$\text{dB} = 20 \log(V2/V1)$$

$$\text{FOR } V2/V1 = 10,000 \gg \gg 20 \log(10,000) = 80\text{dB}$$

$$\text{FOR } V2/V1 = 0.0001 \gg \gg 20 \log(0.0001) = -80\text{dB}$$

$$\text{FOR A 12-BIT CONVERTER, } 2^{12} = 4096$$

$$20\log(4096) = 72.2 \text{ dB} \gg \gg 72.2\text{dB}/12\text{-BITS} = 6.02\text{dB/BIT}$$

$$-0.5\text{dB} = 10^{(-0.5/20)} = 0.944$$

$$-3\text{dB} = 10^{(-3/20)} = 0.707$$

$$-6\text{dB} = 10^{(-6/20)} = 0.5$$

$$-20\text{dB} = 10^{(-20/20)} = 0.1$$

Some common algebraic manipulations of logarithms are shown as many of the frequency domain specifications will be expressed in dBs. One useful rule of thumb is 6 dB of dynamic range per bit of resolution.

The test condition of -0.5dB below Full Scale is becoming important in Frequency Domain specifications and testing. The IEEE has recommended -0.5dB as one of the test conditions for the frequency domain specifications for Waveform Instrumentation.

### SIGNAL-TO-NOISE RATIO (SNR)

FOR AN IDEAL A/D:

$$\text{SNR} = \frac{\text{RMS SIGNAL}}{\text{RMS NOISE}} = 6.02n + 1.76\text{dB}$$

WHERE n = # OF BITS

RESOLUTION	IDEAL SNR
8 BITS	49.9 dB
10 BITS	62.0 dB
12 BITS	74.0 dB
14 BITS	86.0 dB
16 BITS	98.1 dB

Signal-To-Noise Ratio is the ratio of the RMS signal to the RMS noise. The RMS noise includes all nonfundamental spectral components which in a non-ideal A/D would include the harmonics, any spurious frequencies and the noise floor of the device, but excluding dc.

Higher resolution A/Ds reduce the "quantization noise" further, enabling a better signal-to-noise ratio to be obtained.

### SIGNAL-TO-NOISE RATIO (SNR)

RMS SIGNAL

GIVEN THE SINUSOIDAL EXPRESSION:  $v(t) = A \sin \omega t$

AMPLITUDE A = 1/2 the Full Scale Range (FSR)

$$\text{RMS signal} = \frac{A}{\sqrt{2}} = \frac{\text{FSR}/2}{\sqrt{2}} = \frac{(2^n)q}{\sqrt{2}}$$

WHERE q = LSB SIZE

As the theoretical calculations for Signal-to-Noise Ratio are developed, some algebraic manipulations allow the formula to be developed as a function of the # of bits of the data converter.

### SIGNAL-TO-NOISE RATIO & DISTORTION (SINAD)

THE "CLASSIC" DEFINITION OF SNR INCLUDED THE HARMONICS

- THE SIGNAL IS THE FUNDAMENTAL FREQUENCY
- THE NOISE IS ANYTHING UNWANTED WHICH INCLUDED THE HARMONICS, ANY SPURIOUS FREQUENCIES AND NOISE FLOOR (OFFSET ERROR EXCLUDED).
- THE "CLASSIC" DEFINITION OF SNR IS NOW OFTEN STATED AS:

SIGNAL-TO-NOISE RATIO & DISTORTION (SINAD)

IF SIGNAL-TO-NOISE RATIO IS SPECIFIED WITHOUT THE DISTORTION, SINAD CAN BE DETERMINED AS FOLLOWS:

$$\text{SNR \& DISTORTION (in dB)} = -20 \text{ Log } \sqrt{(10^{-\text{SNR}} + 10^{-\text{THD}})}$$

Traditionally, when a Signal-to-Noise Ratio specification was published, the signal was the fundamental frequency and the noise was anything unwanted (harmonics, spurious frequencies, noise floor).

Today, it is important to determine if the SNR specified is with or without distortion. If the SNR w/o distortion and the Total Harmonic Distortion (THD) are specified, SINAD can be calculated.

In Data Converters, the published SINAD specification may suggest upon calculation that better SNR w/o Distortion and Harmonic Distortion specifications than published exist.

The more conservative individual guaranteed specifications are sometimes due to the subtle A/D design tradeoffs. If the harmonics are lower, the noise floor may be higher, however the guaranteed SINAD specification can still be maintained.

### SIGNAL-TO-NOISE RATIO (SNR)

RMS NOISE

QUANTIZATION ERROR =  $\pm 1/2$  LSB

RMS QUANTIZATION NOISE =  $Q_c = q/\sqrt{12}$  WHERE q = LSB SIZE

TO GET EQUAL DISTRIBUTION OF THE QUANTIZATION ERROR, A TRIANGLE WAVEFORM COULD BE ASSUMED FOR THE QUANTIZATION ERROR WITH A PERIOD T UTILIZED TO DERIVE THE RMS QUANTIZATION NOISE AFTER SOME INTEGRATION AS FOLLOWS:

$$Q_c(t)_{\text{RMS}} = \frac{1}{T} \int_{-T/2}^{T/2} \left(\frac{q}{2}\right)^2 dt = \frac{q^2}{12} \left(\frac{T}{T}\right) = \frac{q^2}{12}$$

$$Q_c(t)_{\text{RMS}} = \frac{q^2}{12} \left(\frac{T}{T}\right) = \frac{q^2}{12}$$

$Q_c(t)_{\text{RMS}} = q/\sqrt{12}$ . YIELDING:

$$\text{SNR} = \frac{V_{\text{signal}}(\text{rms})}{V_{\text{noise}}(\text{rms})} = \frac{q(2^n)/\sqrt{2}}{q/\sqrt{12}} = 2^{n-1}\sqrt{6}$$

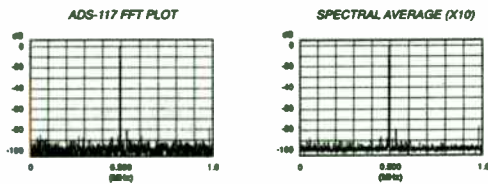
$$\text{SNR} = 20 \text{ Log } [\sqrt{6} (2^{n-1})] = 6.02n + 1.76 \text{ dB}$$

An ideal A/D is limited by the inherent quantization error of  $\pm 1/2$  LSB maximum where an LSB =  $\text{FSR}/2^n$  (n = # of bits, FSR = Full Scale Range).

The quantization error is similar to the analog input being in the presence of white noise and is often expressed as quantization noise. The root-mean-square value of the quantization error is used as the average quantization error is zero (just as likely to be +1/2 LSB as -1/2 LSB of error).

Differential nonlinearities in data converters which are not able to meet the ideal quantization errors contribute to degraded Signal-to-Noise ratio.

## SPECTRAL AVERAGING



$$\text{AVERAGE NOISE FLOOR} = 6.02n + 10 \log \left[ \frac{3N}{2 \times (\text{ENBW})} \right] \text{ dB}$$

WHERE:  $n$  = # OF BITS RESOLUTION  
 $N$  = NUMBER OF DATA POINTS  
 ENBW = EQUIVALENT NOISE BANDWIDTH OF THE WINDOW FUNCTION

PRACTICALLY:

$$\text{NOISE}_{\text{total}}(\text{rms}) = \sqrt{[\text{NOISE}_c(\text{rms})]^2 + [\text{NOISE}_j(\text{rms})]^2 + [\text{NOISE}_A(\text{rms})]^2}$$

A spectral average of the FFT allows distinguishing the harmonics and the noise floor level much easier. Now the magnitudes at each FFT point can be averaged to help overcome the randomness of the noise.

The quantization noise creates a grassy pattern at the bottom of the FFT plot known as the noise floor. An ideal 12-bit A/D with a full scale sine wave input, a 4096 point FFT and a Blackman-Harris window (ENBW = 2) yields a theoretical noise floor of 105.1 dB utilizing the formula shown.

The noise in a data converter can be attributed to the quantization noise (number of bits), the inherent converter noise (semiconductor junction noise, resistor noise) and a frequency dependent term (the noise due to aperture uncertainty).

## TOTAL HARMONIC DISTORTION (THD)

$$\text{THD}_{(\text{rms})} = 20 \text{ Log } \sqrt{[10^{(2\text{nd HARMONIC})}]^2 + [10^{(3\text{rd HARMONIC})}]^2 + \dots}$$

TOTAL HARMONIC DISTORTION INCLUDES ALL THE HARMONICS BY DEFINITION

Total Harmonic Distortion (THD) is the ratio of the rms sum of the harmonics to the rms of the fundamental signal. The harmonics appear at integer multiples of the fundamental frequency.

Integral Nonlinearities appear as harmonics in the frequency domain. The amount of Total Harmonic Distortion is related to how many codes have an integral nonlinearity which is non-ideal. Typically, a good data converter's THD improves as the input signal decreases.

Practically speaking, the first five harmonics are the major contributors to THD. Test conditions of various manufacturers may state the first 6, 9 or 40 harmonics, etc are used in the calculation.

## SIGNAL-TO-NOISE RATIO CAUTIONS

- DETERMINE IF SNR & DISTORTION (SINAD) OR SNR W/O DISTORTION IS BEING USED
- MAKE SURE FULL SCALE SIGNALS ARE BEING TESTED
  - 1) -0.5dB DOWN FROM FULL SCALE IS CONSERVATIVE TEST CONDITION
  - 2) BEWARE IF FULL SCALE NOT SPECIFIED AT ALL
- MAKE SURE THE SIGNAL FREQUENCY IS SPECIFIED
  - 1) CHARACTERIZATION TO NYQUIST IS CONSERVATIVE TEST CONDITION
- ADDITIONAL CONCERNS COULD INCLUDE:
  - 1) HOW MANY FFT POINTS WERE USED
  - 2) WHAT WINDOWING FUNCTION WAS USED

Comparison of various Data Converter Signal-to-Noise Ratio specifications requires similar test conditions.

Full Scale signals (-0.5dB down or about 94% of the Full Scale Range) characterized through Nyquist are recommended. A smaller full scale range may not exercise the data converter in a region where there are nonlinearities (harmonics in the frequency domain). Also, noise from such sources as aperture uncertainty may increase as a function of the input frequency. Typically, a good data converter's SNR degrades as the input signal decreases.

Further insight could include the frequency resolution of the FFT (how many FFT points) and the windowing functions utilized.

## SPURIOUS FREE DYNAMIC RANGE (SFDR) AND IN-BAND HARMONICS

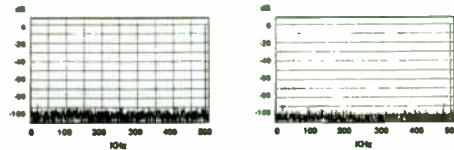
SPURIOUS FREE DYNAMIC RANGE (SFDR) AND IN-BAND HARMONICS ARE SIMILAR IN THEIR DEFINITION AS THE WORST CASE HARMONIC, SPURIOUS SIGNAL OR NOISE COMPONENT RELATIVE TO THE INPUT LEVEL. PRACTICALLY SPEAKING, THIS USUALLY ENDS UP AS BEING THE WORST HARMONIC (2nd HARMONIC).

THESE ARE DISCUSSED HERE AS SOME MANUFACTURERS USE THE SPURIOUS FREE DYNAMIC RANGE SPECIFICATION AND SOME USE THE IN-BAND HARMONICS. THESE SPECIFICATIONS DIFFER FROM TOTAL HARMONIC DISTORTION (THD) WHICH INCLUDES ALL HARMONICS.

The Spurious Free Dynamic Range (SFDR) and In-Band Harmonic specifications are intended to indicate the largest harmonic, spurious frequency or noise component. This is usually the 2nd harmonic (rarely the third), but should the harmonics be indistinguishable from the noise, then these specifications cover this situation also.

These specifications indicate the usable dynamic range to the user, the range in which frequency components other than the fundamental do not exist.

### FREQUENCY DOMAIN TEST CONDITIONS FULL SCALE INPUT CONSIDERATIONS



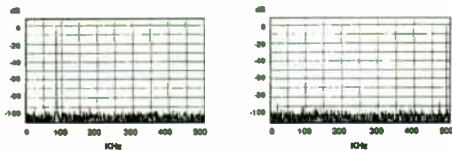
ADS-112 FFT @ -4.44 dB OF FS		ADS-112 FFT @ -0.5 dB OF FS	
2nd HARMONIC	-88.1 dB	2nd HARMONIC	-84.9 dB
SINAD	86.9 dB	SINAD	70.4 dB
EFFECTIVE BITS	11.54 BITS	EFFECTIVE BITS	11.41 BITS
DNL	±0.48 LSB	DNL	±0.55 LSB
F <sub>s</sub>	1 MHz	F <sub>s</sub>	1 MHz
F <sub>in</sub>	490 kHz	F <sub>in</sub>	490 kHz

$$\text{SNR} = 6.02n + 1.76\text{dB} - 20 \log \frac{\text{FULL SCALE AMPLITUDE}}{\text{ACTUAL INPUT}}$$

The 12-bit, 1 MHz Sampling A/D (ADS-112) is shown for different input voltages with an input frequency, F<sub>in</sub>, of 490 KHz. Note the noise floor and harmonic changes which occur at the different full scale inputs.

In comparing A/D converter specifications, it is important that comparable test conditions were employed. The conservative, -0.5 dB down from full scale, test condition is recommended.

### FREQUENCY DOMAIN TEST CONDITIONS FREQUENCY CONSIDERATIONS



ADS-112 FFT @ 90 KHz		ADS-112 FFT @ 490 KHz	
2nd HARMONIC	-88.2 dB	2nd HARMONIC	-84.9 dB
SINAD	71.1 dB	SINAD	70.4 dB
EFFECTIVE BITS	11.58 BITS	EFFECTIVE BITS	11.41 BITS
DNL	±0.51 LSB	DNL	±0.55 LSB
INPUT SIGNAL	-0.5dB of FS	INPUT SIGNAL	-0.5dB of FS
F <sub>s</sub>	1 MHz	F <sub>s</sub>	1 MHz

The FFT plot shows an input frequency spectrum up to 500 KHz for the 12-bit, 1 MHz ADS-112. It is important in comparing specifications to make sure the input frequency is defined. Note the better specifications when the frequency is closer to DC versus close to the maximum for Nyquist sampling.

The ADS-112's FFT plot shows spectral components up to half the sampling rate (F<sub>s</sub>). Make sure the sampling rate is defined for any FFT plots which do not use frequency for the X-axis. Some FFT plots have used "frequency bins" (to be discussed later) which would not allow the user to determine the sampling rate or input frequency (do not assume the maximum sampling rate, F<sub>s</sub>, was used).

### EFFECTIVE BITS

EFFECTIVE BITS USES NUMERICAL METHODS TO COMPARE HOW CLOSE A DIGITIZED SINEWAVE REPRESENTS AN IDEAL MATHEMATICAL MODEL.



D/A >>>> MONOTONICITY  
A/D >>>> NO MISSING CODES  
A/D >>>> EFFECTIVE BITS

CORRELATE EFFECTIVE BITS WITH SNR (± DISTORTION):

$$\text{EFFECTIVE BITS} = \left[ \text{SNR} - 1.76 \right] \cdot \left[ 20 \log \frac{\text{FULL SCALE AMPLITUDE}}{\text{ACTUAL INPUT AMPLITUDE}} \right] - 6.02$$

INPUT LEVEL CORRECTION FACTOR

$$\text{SNR} = 6.02n + 1.76\text{dB} + \left[ 20 \log \frac{\text{FULL SCALE AMPLITUDE}}{\text{ACTUAL INPUT AMPLITUDE}} \right]$$

Errors from Differential and Integral Nonlinearities, aperture uncertainty, noise and quantization error (noise) are involved in determining the # of effective bits.

The effective bit specification is a figure of merit on the overall A/D transfer function. As monotonicity is to Digital-to-Analog Converters and as no missing codes is to A/Ds, effective bits gives a useful insight into a Sampling A/D's transfer function as a function of frequency.

In correlating effective bits to signal-to-noise ratio, a correction factor for the formula is needed when the sinusoid input is less than full scale:

$$\text{correction factor} = \frac{\text{full scale amplitude}}{\text{actual input amplitude}} + 20\log \frac{\text{full scale amplitude}}{\text{actual input amplitude}}$$

### EFFECTIVE BITS

IDEAL SINEWAVE:  $A \sin(2\pi ft + \phi) + DC$

WHERE: A, f, φ, and DC ARE CALCULATED TO "BEST-FIT" THE A/D DATA  
A = AMPLITUDE; f = FREQUENCY; φ = PHASE; DC = DC OFFSET

THE RMS ERROR BETWEEN AN IDEAL SINEWAVE AND THE BEST-FIT SINEWAVE IS:

$$\text{ERROR}_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{m=1}^N [D_m - A \sin(2\pi f t_m + \phi) - DC]^2}$$

WHERE: E = CALCULATED rms ERROR  
N = DATA RECORD LENGTH  
D<sub>m</sub> = THE DATA

TAKING THE PARTIAL DERIVATIVE OF E<sub>rms</sub> WITH RESPECT TO EACH OF THE FOUR PARAMETERS YIELDS:

$$\text{EFFECTIVE BITS} = n - \log_2 \frac{E_{\text{rms}}}{q \sqrt{N/12}} \quad \text{WHERE } q = \text{LSB SIZE}$$

Numerical methods such as least squared minimization techniques can be used in comparing the ideal sinewave to the digitized sinewave.

As for many of the frequency domain tests, in performing the sine wave curve fit, it is important to assure that the input is not harmonically related to the sample frequency, i.e. the input frequency is not a sub-multiple of the sample frequency. Certain codes would increase in occurrence and SNR might improve as the harmonics alias back onto the fundamental.

The number of data points, D<sub>m</sub>, should also be large giving good frequency resolution. The actual frequency can then be readily determined.

## TWO-TONE INTERMODULATION DISTORTION

$$\text{TOTAL IMD} = 20 \log \left( \frac{\text{RMS OF SUM \& DIFFERENCE OF DISTORTION PRODUCTS}}{\text{RMS VALUE OF FUNDAMENTAL}} \right)$$

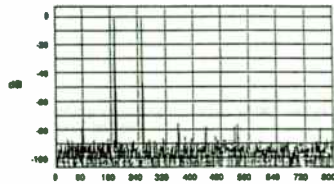
$$= 20 \log \sqrt{\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} [(IMD(mn, \pm na_1))]^2}$$

$$\text{TOTAL IMD} = 20 \log \sqrt{I_{IMD}(F_1 + F_2/20)^2 + I_{IMD}(F_1 - F_2/20)^2 + I_{IMD}(2F_1 + F_2/20)^2 + \dots + I_{IMD}(2F_1 - F_2/20)^2 + I_{IMD}(F_1 + 2F_2/20)^2 + I_{IMD}(F_1 - 2F_2/20)^2 + \dots}$$

Typically used in communication applications such as when multiple frequencies are multiplexed onto a single carrier, two-tone intermodulation distortion is the output distortion resulting from one sinusoidal input signal's interaction with another signal at a different frequency (tone).

As in harmonic distortion, nonlinearities in the A/D transfer in the frequency domain as intermodulation distortion, occurring for two-tone inputs at the sum and difference frequencies.

## TWO-TONE INTERMODULATION DISTORTION



SHOWN ARE TWO EQUAL AMPLITUDE SINUSOIDS,  $\sin(2\pi F_1 t)$  AND  $\sin(2\pi F_2 t)$  OF DIFFERENT FREQUENCIES  $F_1, F_2$ .

- HARMONICS STILL APPEAR AT  $nF_1$  AND  $nF_2$ , WHERE  $n$  IS AN INTEGER
- NOW ALSO SEE COMPONENTS AT  $mF_1 \pm nF_2$  AND  $mF_2 \pm nF_1$  WHERE  $m, n$  ARE ANY INTEGERS.

FOR THE FFT SHOWN:

- 2nd ORDER TERMS:  $(F_1 + F_2)$ ,  $(178 \text{ KHz} + 258 \text{ KHz} = 432 \text{ KHz})$   
 $(F_1 - F_2)$ ,  $(178 \text{ KHz} - 258 \text{ KHz} = -80 \text{ KHz})$   
 $(2F_1 + F_2)$ ,  $(2 \times 178 \text{ KHz} + 258 \text{ KHz} = 608 \text{ KHz})$   
 $(2F_1 - F_2)$ ,  $(2 \times 178 \text{ KHz} - 258 \text{ KHz} = 98 \text{ KHz})$   
 $(F_1 + 2F_2)$ ,  $(178 \text{ KHz} + 2 \times 258 \text{ KHz} = 698 \text{ KHz})$   
 $(F_1 - 2F_2)$ ,  $(178 \text{ KHz} - 2 \times 258 \text{ KHz} = -338 \text{ KHz})$

An FFT is shown with two input frequencies (tones) applied. The input amplitudes should combine to produce a resultant signal amplitude of -0.5 dB down from full scale. Utilizing individual signals that are less than -6 dB down from full scale will prevent clipping of the input signals. Also, these input signals should not be closer than the frequency resolution of the FFT.

Normally, two input signals close to the upper end of the bandwidth (1/2 the sampling rate) are used as test signals (lower frequency tones show here to easily demonstrate the IMD). This results in the 2nd order terms being spaced far apart from the original signals while there are third order terms which are spaced close to the input signals. The third IMD terms which are close to the input signals are difficult to filter.

## NOISE POWER RATIO (NPR)

FOR AN IDEAL A/D:

$$\text{NPR} = \frac{\text{FULL SCALE OF THE A/D}}{\text{K(QUANTIZATION NOISE)}} = \frac{Q^2}{KQ^2/12} = \frac{2^n}{K/12}$$

WHERE:

Q = QUANTIZATION LEVEL ( $Q = \sqrt{12}$  IS QUANTIZATION NOISE)  
 $n$  = # OF BITS

K = LOADING FACTOR  $\frac{\text{(FULL SCALE)}}{\text{RMS NOISE LEVEL}}$

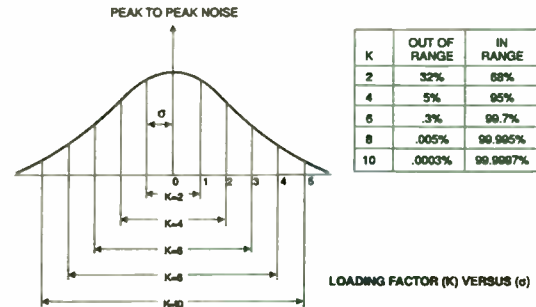
$$\text{NPR}_{dB} = 6.02n + 20 \log(\sqrt{3}/K) \text{ dB}$$

Noise Power Ratio (NPR) is a specification developed for Frequency Division Multiplexed Communication Equipment. In a multi-channel environment, the signals in the off channels can create thermal and intermodulation effects on the channel of interest.

Practical NPR testing simulates the random noise which would occur in the multi-channel environment through the use of a noise generator. A narrow-band notch filter is switched in and out and a noise receiver utilized in determining the ratios. FFT based NPR testing usually demands excessive data records to obtain the relevant frequency resolution for the band of interest.

Testing is performed at low, medium and high frequencies as various errors from nonlinearity, intermodulation or crossover distortion, and phase distortion may be more prevalent at certain frequencies.

## NOISE POWER RATIO (NPR)



The Noise Power Ratio (NPR) calculation requires a loading factor, K, be used to prevent clipping and its associated distortion. Here, the peak-to-peak noise is graphed as a function of sigma ( $\sigma$  = rms noise level).

For a 12-bit accurate system, choosing a value of 8 for K will assure that only a small amount of the noise (1 LSB = 0.024% FSR) is outside the full scale range of the converter.

## INPUT BANDWIDTH



**LARGE SIGNAL BANDWIDTH:**  
FREQUENCY WHERE THE MAXIMUM SINUSOIDAL INPUT SIGNAL HAS DECREASED BY 3dB AS DERIVED FROM THE DIGITAL OUTPUT DATA.

**FULL POWER BANDWIDTH:**  
MAXIMUM FREQUENCY FULL SCALE SINE WAVE DIGITIZED WITHOUT SPURIOUS OR MISSING CODES.

**SMALL SIGNAL BANDWIDTH:**  
BANDWIDTH WHERE THE AMPLITUDE IS 1/10 THE VALUE (-20dB) OF THE MAXIMUM INPUT AMPLITUDE.

A/D converters should have wide input bandwidths for handling transients or pulse type analog inputs. Sinusoids digitized at Nyquist undergo full scale steps from one conversion to the next. However the step function is not able to track the input signal between conversions and presents a more stringent test.

A step function, such as created when a multiplexer switches between multiple channels or perhaps full scale steps between pixels (white/dark edges) in an imaging application, may also undergo full-scale steps from one conversion to the next.

Wide input bandwidths also prevent phase shifts from occurring in such applications as I & Q channels in radar. Additionally, fast recovery times to input overvoltage conditions which are not clamped are achieved when wide input bandwidths are present.

## TRANSIENT RESPONSE & OVERLOAD RECOVERY TIME

- TRANSIENT RESPONSE >>>> ACQUISITION TIME
- OVERLOAD RECOVERY TIME:

TIME AFTER THE INPUT RETURNS TO THE A/D CONVERTER'S NORMAL OPERATING INPUT RANGE UNTIL THE TIME THAT THE A/D CAN MAKE A PROPER CONVERSION.

THE OVERVOLTAGE CONDITION NEEDS TO BE SPECIFIED. +/- 20% ? +/- 2X INPUT?



Acquisition time, the time needed for a sample-and-hold to acquire a full scale step signal is sometimes referred to as transient response. Full scale steps, pulses or transients can be a more difficult test for an A/D than a sinusoidal type input. A wide input bandwidth and fast acquisition time (transient response) assures the A/D can handle these transients.

## ADS-130 12-BIT, 10 MHz SAMPLING A/D CONVERTER

### FUNCTIONAL SPECIFICATIONS

Apply over the operating temperature range:  
±15V and ±5V power supplies unless otherwise specified

FUNCTIONAL SPECIFICATIONS	MIN	TYP	MAX	UNITS
Resolution	12	12	12	bits
Conversion Rate	100,000	100,000	100,000	conversions/sec
Aperture Uncertainty	±0.5	±0.5	±0.5	ns
Alien Rejection	-	-	-	dB
Linearity	-	-	-	LSB
Gain Error	-	-	-	LSB
Offset Error	-	-	-	LSB
Temperature Coefficient	-	-	-	LSB/°C
Power Consumption	-	-	-	mW
Supply Current	-	-	-	mA
Operating Voltage	-	-	-	V
Package	-	-	-	Pin

DYNAMIC PERFORMANCE	MIN	TYP	MAX	UNITS
Conversion Rate	100,000	100,000	100,000	conversions/sec
Aperture Uncertainty	±0.5	±0.5	±0.5	ns
Alien Rejection	-	-	-	dB
Linearity	-	-	-	LSB
Gain Error	-	-	-	LSB
Offset Error	-	-	-	LSB
Temperature Coefficient	-	-	-	LSB/°C
Power Consumption	-	-	-	mW
Supply Current	-	-	-	mA
Operating Voltage	-	-	-	V
Package	-	-	-	Pin

© The same specifications apply to Inboard Conversion.

Key specifications for determining how an A/D converter will handle sinusoidal or pulse/step applications are shown. Note the minimum-maximum specifications over temperature and the frequency domain specifications as a function of the full scale input (-0.5 dB) and frequency.

## FREQUENCY DOMAIN TEST CONDITIONS FFT SIZE/ACCURACY CONSIDERATIONS

$$\text{FFT ACCURACY} = \frac{-2.4}{n \sqrt{N}} \text{ dB}$$

WHERE:  $n$  = # OF BITS RESOLUTION  
 $N$  = NUMBER OF DATA POINTS

### ACCURACY OF FFT

NUMBER OF DATA POINTS:	256	4096
RESOLUTION		
12-BIT	0.021dB 0.24%	0.005dB 0.06%
14-BIT	0.018dB 0.21%	0.0045dB 0.051%
16-BIT	0.016dB 0.18%	0.0039dB 0.045%

The size of the FFT (number of data points) utilized determines the frequency resolution and the accuracy of the FFT. The table shows that taking more samples improves the accuracy of the FFT, however note that the square root relationship prevents a dramatic accuracy gain. The 4096 point FFT yields a 0.06% accuracy which is more suitable for 12-bit A/Ds.



## COHERENT & NONCOHERENT SAMPLING

### COHERENT SAMPLING:

AN INTEGER NUMBER OF SINEWAVE CYCLES ARE USED IN THE DATA RECORD FOR THE FFT.

- FREQUENCY CONTENT OF THE INPUT SIGNAL HAS TO BE KNOWN
- PROVIDES THE LARGEST DYNAMIC RANGE
- ELIMINATES "LEAKAGE" AND THE NEED FOR "WINDOWING"

### NONCOHERENT SAMPLING

A NON-INTEGER MULTIPLE OF THE SINEWAVE CYCLE IS IN THE DATA RECORD.

- NEEDED WHEN THE FREQUENCY CONTENT OF THE INPUT SIGNAL IS UNKNOWN
- A "WINDOW" FUNCTION TO TIME-WEIGHT THE DATA IS REQUIRED.

The Fourier Transform expects to work on continuous data from  $-\infty$  to  $+\infty$ . In performing an FFT, the infinite amount of data has been truncated to contain a finite number of samples.

At the heart of this mathematical exercise in transforming time domain data to the frequency domain is the goal that complete cycles of sinusoidal waveforms are being analyzed. Coherent sampling assures that complete cycles are used.

Noncoherent sampling utilizes windowing functions to approximate this ideal. This is accomplished by having the data at the ends (partial cycle) go to zero. Careful selection of the window function is required to minimize the fundamental frequency leaking into other frequency "bins."

## NON-COHERENT SAMPLING - FFT WINDOW ASPECTS

- FREQUENCY RANGE/RESOLUTION
- SPECTRAL LEAKAGE
- 3 dB BANDWIDTH
- PICKET-FENCE EFFECT (SCALLOP LOSS)
- EQUIVALENT NOISE BANDWIDTH
- WORST CASE PROCESSING LOSS

Non-coherent sampling is required when the input frequency is unknown and is therefore more representative of an actual A/D application. The windows utilized in non-coherent sampling attenuate the sidelobe errors while spreading out the main-lobe response.

A number of characteristics of the window chosen and FFT become important depending on the application.

## COHERENT SAMPLING

THE ANALOG INPUT SINEWAVE FREQUENCY AND THE A/D SAMPLING RATE NEED TO OBSERVE THE FOLLOWING RELATIONSHIP FOR COHERENT SAMPLING:

$$f_{in} = \frac{nF_s}{N}$$

WHERE:

- $f_{in}$  = INPUT SINEWAVE FREQUENCY
- $F_s$  = SAMPLING RATE
- $n$  = PRIME NUMBER (1,3,5,7,11, ETC)  
(PRIME #s ASSURE EACH DATA POINT IS UNIQUE)
- $N$  = NUMBER OF SAMPLES TAKEN

NOTE: THE INPUT FREQUENCY AND THE SAMPLING RATE NEED TO BE FREQUENCY LOCKED.

Coherent sampling requires the input frequency to be frequency locked to the A/D sampling rate. Assuring complete cycles eliminates the need for windowing functions to force the data at the boundaries to zero.

Coherent sampling becomes important with higher resolution A/Ds as the window functions reach their limit. The 4-term Blackman-Harris window's sidelobe is -92 dB down and begins to be a factor in 14 and 16-bit A/D testing.

## FREQUENCY RANGE/RESOLUTION OF THE FFT

$$\text{FREQUENCY RANGE OF THE FFT} = F_r = \frac{\text{SAMPLING RATE}}{2}$$

$$\text{MINIMUM FREQUENCY RESOLUTION OF THE FFT (}\# \text{ OF FREQUENCY BINS)} = \Delta f = \frac{\text{SAMPLING RATE (}F_s\text{)}}{\text{RECORD LENGTH OF FFT (}N\text{)}}$$

$$\text{FREQUENCY RESOLUTION} = \Delta f = \frac{\text{(ENBW) SAMPLING RATE}}{\text{RECORD LENGTH}}$$

WHERE ENBW = EQUIVALENT NOISE BANDWIDTH

Per Nyquist, the frequency range,  $F_r$ , is the Sampling rate divided by two. The frequency resolution,  $F_s/N$ , is the minimum FFT bin width. Depending on the particular window utilized, the bin width increases by the Equivalent Noise Bandwidth (ENBW).

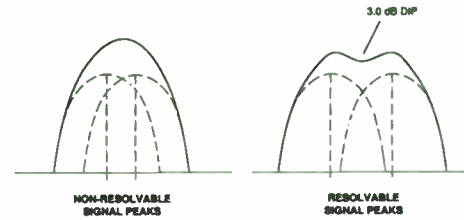
### SPECTRAL LEAKAGE



- SPECTRAL LEAKAGE IS CAUSED BY FINITE DATA RECORDS NOT REPRESENTING COMPLETE CYCLES. DISCONTINUITIES AT THE BOUNDARIES OCCUR WHEN COMPLETE CYCLES ARE NOT REPRESENTED.
- WINDOW FUNCTIONS WEIGHT THE DATA HEAVILY IN THE CENTER OF THE DATA RECORD AND BRINGS THE SIGNAL TO ZERO AT THE BOUNDARIES.

Spectral leakage occurs when energy from one frequency spreads into adjacent ones. This occurs when a fraction of a cycle exists in the waveform that is subjected to the FFT. Window functions are used to reduce the spectral leakage.

### 3.0 dB BANDWIDTH (IN BINS)



The 3.0 dB bandwidth (in frequency bins) is defined as the minimum bandwidth where two closely spaced signals can still be resolved. Two signals separated by less than their 3.0 dB bandwidth are erroneously resolved into a single spectrum component.

The ideal 3.0 dB bandwidth is equal to one FFT bin width, requiring a windowing function with characteristics of an ideal bandpass filter (0 dB in the passband, steep attenuation outside).

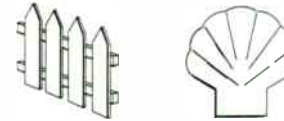
### SPECTRAL LEAKAGE



- THE WINDOWING OF THE DATA IS SIMILAR TO MULTIPLYING THE SIGNAL BY A RECTANGULAR WINDOW. MULTIPLICATION IN THE TIME DOMAIN IS ANALOGOUS TO CONVOLUTION IN THE FREQUENCY DOMAIN.
- THE FOURIER TRANSFORM OF A RECTANGULAR WINDOW IS THE POPULAR (SIN X)/X OR SINC FUNCTION. CONVOLVING THE SPECTRUM OF DATA,  $x(n\Delta t)$ , WITH THIS FUNCTION SMEARS THE CALCULATED SPECTRUM OF  $x(t)$ . THE SPECTRAL AMPLITUDES OF  $x(t)$  WILL LEAK THROUGH THE SIDELOBES OF THE SINC FUNCTION, SPREADING ENERGY FROM ONE FREQUENCY INTO ADJACENT FREQUENCIES.

The various window functions modify the main lobe and side lobes of the basic sinc function shown for a rectangular window.

### PICKET FENCE EFFECT (SCALLOP LOSS)



- LOSS IN GAIN FOR A FREQUENCY MIDWAY BETWEEN TWO BIN FREQUENCIES
- SCALLOP LOSS VARIES DEPENDING ON WINDOW CHOSEN
- AMPLITUDE OF THE KTH SPECTRAL LINE =  $\sqrt{A_k^2 + B_k^2}$

$$\text{WHERE: } A_k = \frac{1}{N} \sum_{m=0}^{N-1} W_m D_m \cos\left(\frac{2\pi k(m-1)}{N}\right)$$

$$B_k = \frac{1}{N} \sum_{m=0}^{N-1} W_m D_m \sin\left(\frac{2\pi k(m-1)}{N}\right)$$

$A_k$  = MAGNITUDE OF SINE  
 $B_k$  = MAGNITUDE OF COSINE  
 $W_m$  = WEIGHTING FUNCTION  
 $D_m$  = AMPLITUDE OF THE DATA POINT  
 $N$  = NUMBER OF DATA POINTS

The picket-fence effect or scallop loss (dB) can be significant for multi-tone signals. The FFT output points (frequency bins) occur at the sample frequency divided by the number of FFT points. The deviation in gain from 0 dB within  $\pm 1/2$  frequency bin is the scallop loss.

This scallop loss can vary from 3.92 for the rectangular window to 0.83 for the 4-term Blackman-Harris window.

### EQUIVALENT NOISE BANDWIDTH (ENBW)



- DETERMINING THE SPECTRAL AMPLITUDE OF A SIGNAL IS AFFECTED BY THE NOISE IN THE BANDWIDTH OF THE WINDOW
- NARROW BANDWIDTH FILTERS ARE USED TO MINIMIZE THIS NOISE
- ENBW IS THE WIDTH OF A RECTANGULAR FILTER WITH THE SAME PEAK POWER GAIN WHICH WOULD ACQUIRE THE SAME NOISE POWER AS THE WINDOWING FUNCTION UTILIZED.

A window function acts like a filter, weighting the inputs for the final calculation over its entire bandwidth. To detect low-level signals, windows with narrow bandwidths are used to minimize the noise.

The equivalent noise bandwidth (in frequency bins) can vary from 1.0 for the rectangular window to 2.0 for the 4 term Blackman-Harris window.

### COMMON FFT WINDOWING FUNCTIONS

#### FOUR TERM BLACKMAN-HARRIS

- FLATNESS OF MAIN LOBE - ACCURATE PEAK AND POWER MEASUREMENTS

#### HANNING

- IMPROVED FREQUENCY RESOLUTION OVER RECTANGULAR OR TRIANGULAR WINDOWS.

#### HAMMING

- OPTIMIZED TO LOWER THE 1ST SIDE LOBE.
- DETECTION OF CLOSE FREQUENCY LINES IMPROVED

#### RECTANGULAR

- TRANSIENT SIGNALS
- HIGHEST SIDELOBES LIMIT ABILITY TO:
  - (1) RESOLVE TWO CLOSE SPECTRAL COMPONENTS
  - (2) DETECT LOW AMPLITUDE FREQUENCIES (MASKED BY LEAKAGE)

#### TRIANGULAR

- CANNOT DETECT A WEAK SPECTRAL LINE IN THE PRESENCE OF STRONG SIGNALS

A variety of window functions exist with five of the more common listed here. Window selection may depend on ease of processing, single or multi-tone testing, presence of a small signal in the near vicinity of a large signal, etc..

### WORST CASE PROCESSING LOSS

- COMBINES SCALLOP LOSS WITH THE PROCESSING LOSS WHICH OCCURS AT THE DATA RECORD BOUNDARIES.
- THE WORST CASE PROCESSING LOSS REDUCES THE SIGNAL-TO-NOISE RATIO DUE TO WINDOWING AND THE WORST CASE FREQUENCY LOCATION.
- MAINLY AN ISSUE FOR MULTI-TONE TESTING

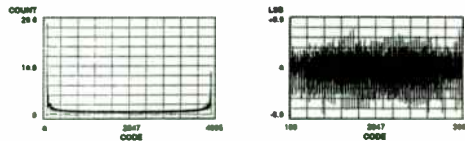
The worst case processing loss varies from 3.0 to 4.3 dB for various windowing functions. This loss becomes important when detecting low level signals in the presence of many other frequencies where contribution to a spectral component may be shared by different frequency bins.

### FFT WINDOW CHARACTERISTICS

	HIGHEST SIDELOBE (dB)	3.0 dB BANDWIDTH (BINS)	SCALLOP LOSS (dB)	ENBW (BINS)	WORST CASE PROCESS LOSS (dB)
4-TERM BLACKMAN-HARRIS	-92	1.90	0.83	2.0	3.85
HANNING (COS <sup>2</sup> X)	-32	1.44	1.42	1.5	3.18
HAMMING	-43	1.3	1.78	1.36	3.10
RECTANGULAR	-13	0.89	3.92	1.00	3.92
TRIANGULAR	-27	1.28	1.82	1.33	3.07

Some of the major characteristics are shown for some of the common FFT windows. Note the -92 dB side lobe for the 4-term Blackman Harris window used in testing precision A/Ds

### HISTOGRAMS (CODE DENSITY TEST)



ADS-131 HISTOGRAM  
@ 2.45 MHz Fin

ADS-131 DIFFERENTIAL  
NONLINEARITY @ 2.45 MHz Fin

- HISTOGRAM USED IN TESTING DIFFERENTIAL NONLINEARITY
- SINEWAVE INPUTS VERSUS DC RAMPERS ALLOW TESTING AS A FUNCTION OF FREQUENCY

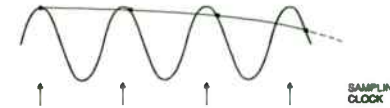
• PROBABILITY DENSITY FUNCTION FOR A SINEWAVE -  $P(V) = \frac{1}{\pi \sqrt{A^2 - V^2}}$

WHERE: A = PEAK AMPLITUDE OF SINEWAVE  
P(V) = PROBABILITY OF OCCURENCE AT VOLTAGE V

The histogram plot of the 12-bit, 5 MHz ADS-131 sampling A/D shows the possible output codes along the X axis and the frequency of occurrences along the Y axis for a sinewave. This is then compared to the probability density function of an ideal sinusoid in determining the differential nonlinearity of the A/D.

Sinusoids allow testing differential nonlinearity as a function of the input frequency (non-coherent sampling). More occurrences are seen at the peak of the sinewave versus at the zero crossing where maximum slew rate of the sinusoid occurs. Ramp testing as an input for a histogram is usually used for slow changing signals only. Discontinuities at the peak would require infinite bandwidth for no distortion.

### BEAT FREQUENCY TEST



- BEAT FREQUENCY TESTING IS A FIGURE OF MERIT TEST GIVING QUICK INSIGHT INTO AN A/D'S GROSS DYNAMIC ERRORS.
- DIFFERENTIAL NONLINEARITIES AND MISSING CODES APPEAR AS DISTORTIONS OR DISCONTINUITIES IN THE LOW FREQUENCY SINEWAVE.
- THE BEAT FREQUENCY IS SET TO CHANGE BY ONE LSB FROM ONE CONVERSION TO THE NEXT:

$\Delta f = \frac{F_s}{2^n}$  WHERE:  $\Delta f$  = BEAT FREQUENCY  
 $F_s$  = SAMPLING RATE  
 $n$  = # OF BITS

Setting the full scale input sinewave frequency equal to the sampling rate of the A/D would ideally give the same output code for each conversion. Offsetting the input sinewave frequency such that from one conversion to the next a maximum of one LSB change occurs results in a beat frequency,  $\Delta f$ .

This slow beat frequency gives a quick, visual demonstration of the A/D's performance utilizing computer graphics or reconstruction D/A's. The D/A's update rate should be frequency locked with the A/D sampling rate.

### HISTOGRAM ACCURACY

NUMBER OF SAMPLES (N) REQUIRED FOR A PARTICULAR PRECISION IN LSBs ( $\beta$ ) AND CONFIDENCE (C) FOR AN n-BIT CONVERTER IS:

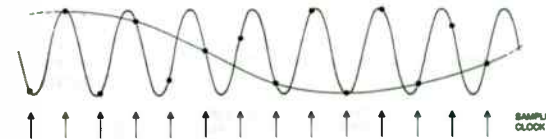
$N = \frac{(Z_\alpha)^2 \cdot 2^n}{\beta^2}$  WHERE:  $Z_\alpha$  = # OF STANDARD DEVIATIONS FOR ANY CHOSEN  $\alpha$   
 $\alpha = \frac{100 - C}{200}$

C%	99	98	97	96	95	94	93	92	91	90
$\alpha$	0.005	0.01	0.015	0.02	0.025	0.03	0.035	0.04	0.045	0.05
$Z_\alpha$	-2.58	-2.33	-2.17	-2.06	-1.96	-1.88	-1.82	-1.75	-1.7	-1.65

The differential nonlinearity for a 12-bit converter with a 99% confidence to 0.1 LSB accuracy would require 4.3 million samples. Dividing the 4.3 million samples by 4096 codes ( $2^{12}$  where  $n=12$ ) is about 1000 records. Using 100 records yields 1/4 LSB accuracy.

Histogram testing is not popular for integral nonlinearity. Integral nonlinearity would require more samples and associated longer testing time. Drifts in the input signal's amplitude or offset or the A/D's offset or gain would produce erroneous results. An alternative method would be to use an FFT where nonlinearity appears as harmonics.

### ENVELOPE TEST



- SIMILAR TO BEAT FREQUENCY TEST WITH THE OFFSET FREQUENCY BEING AT TWO SAMPLES PER CYCLE.
- THE A/D IS REQUIRED TO SLEW EFFECTIVELY FROM PEAK TO PEAK ON THE SINEWAVE FROM ONE CONVERSION TO THE NEXT.
- A DIGITAL DIVIDER IS USED TO SEND EVERY OTHER SAMPLE TO THE D/A, ALLOWING A LOW FREQUENCY BEAT FREQUENCY TO BE OBSERVED.

The envelope test is more demanding than a beat frequency test as it requires the A/D to slew from one extreme of the full scale range to another on successive conversions.

## LOW-PASS FILTERS FOR SIGNAL GENERATORS

### HARMONICS OBSERVED ON THE FFT PLOT MAY BE DUE TO:

- SIGNAL GENERATOR
  - RESOLUTION/FREQUENCY DEPENDENT
- FILTER USED TO REMOVE THE SIGNAL GENERATOR HARMONICS
- DRIVE/IMPEDANCE MATCHING WHEN PASSIVE FILTER USED
- TEST FIXTURING
- THE A/D CONVERTER

Careful analysis of the sources of the harmonics appearing on an FFT plot of an A/D would go beyond the A/D itself. The signal generator, anti-alias filtering, and test fixturing are all potential causes of harmonics.

## COMMON ANTI-ALIAS FILTER TYPES

**BUTTERWORTH**  
FLATTEST RESPONSE NEAR DC, MODERATELY FAST ROLL-OFF, ATTENUATION RATE - 6dB/OCTAVE, STABLE PHASE SHIFT,  $F_c$  OCCURS AT -3dB, CONSTANT AMPLITUDE EMPHASIS VERSUS TIME DELAY OR PHASE RESPONSE. OVERSHOOTS ON STEP RESPONSE.

**CHEBYSHEV**  
RAPID ATTENUATION ABOVE THE CUTOFF FREQUENCY WITH SOME PASSBAND RIPPLE.  $F_c$  OCCURS WHEN THE ATTENUATION EXCEEDS THE SPECIFIED RIPPLE. HAS A SQUARE AMPLITUDE RESPONSE THAN THE BUTTERWORTH BUT LESS DESIRABLE PHASE AND TIME DELAY.

**CAUER (ELLIPTICAL)**  
SURPASSES OTHER DESIGNS FOR CRITICAL AMPLITUDE APPLICATIONS, VERY SHARP ROLLOFF RATE WITH SOME RIPPLE, SQUAREST POSSIBLE AMPLITUDE RESPONSE WITH POOR PHASE AND TRANSIENT RESPONSE.

**BESSEL FILTERS**  
OPTIMIZED FLAT PHASE RESPONSE OVER WIDE INPUT FREQUENCY, MODERATE ATTENUATION RATE,  $F_c$  OCCURS WHERE PHASE SHIFT IS 1/2 THE MAX PHASE SHIFT, AMPLITUDE NOT AS FLAT AS BUTTERWORTH, ROLLOFF IS SLOW, GAIN ROLLOFF COULD MODIFY AMPLITUDE, IMPORTANT IN PULSE TRANSMISSION AS AVOIDS OVERSHOOT/UNDERSHOOT.

Some of the common anti-alias filter types are listed. For frequency analysis of high resolution, high-speed A/D converters, the Cauer (Elliptical) filter finds broad usage.

## PASSIVE VERSUS ACTIVE FILTERS

### PASSIVE FILTERS

- UTILIZE RESISTORS, CAPACITORS AND INDUCTORS
- LOWER NOISE THAN ACTIVE TYPE ESPECIALLY AT HIGHER FREQUENCIES
- LESS HARMONIC DISTORTION
- NO POWER SUPPLY REQUIRED
- IMPEDANCE MATCHING ISSUES
- LARGE SIZE
- STABLE VERSUS FREQUENCY

### ACTIVE FILTERS

- UTILIZE OP AMPS, RESISTORS AND CAPACITORS
- BULKY AND EXPENSIVE INDUCTORS ELIMINATED
- EASY TO CONTROL GAIN
- TUNABLE
- SMALL SIZE
- LOW WEIGHT
- MINIMAL SHIELDING PROBLEMS
- HIGH INPUT IMPEDANCE
- LOW OUTPUT IMPEDANCE

A comparison of the virtues of passive versus active filters are shown. As the resolution and frequency increase, the passive filters become the filters of choice.

## FINAL SUMMARY

### SEMINAR OBJECTIVES:

UPON COMPLETION OF THIS SEMINAR YOU WILL BE ABLE TO UNDERSTAND THE VARIOUS A/D CONVERTER FREQUENCY DOMAIN SPECIFICATIONS.

DATEL THANKS YOU FOR YOUR PARTICIPATION

**Instructor Biography:**

Bob Leonard is the Product Marketing Manager for Component Products at DATEL. Bob received his BSEE degree from Northeastern University and has been with DATEL since 1974. In his spare time, Bob enjoys reading and various sports including golf, tennis, basketball, softball, soccer and running.

**References:**

1. Robert E. Leonard Jr., "Data Converters: Getting to Know Dynamic Specs," *Electronic Design*, November 8, 1990.
2. Robert E. Leonard Jr., "Picking the Right Sample-and Hold Amp for Various Data-Acquisition Needs," *Electronic Engineering Times*, August 17, 1987.
3. Robert E. Leonard Jr., "High-speed A/D converter Designs: Layout and Interfacing Pitfalls," *I & CS*, January 1987.
4. Ray Ushani [DATEL, Inc.], "Subranging Analog-to-Digital Converters, Architecture, Design Considerations, Parameters and Applications," *EDN*, three-part series beginning April 11, 1991.
5. Frederick 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.
6. "Dynamic Performance Testing of A to D Converters," Hewlett Packard Product Note 5180A-2.
7. Y.C. Jenq, PH.D., "Asynchronous dynamic testing of A/D converters," *Tektronix Handshake*, Summer 1988.

Copyright © March, 1991 by DATEL, Inc.  
Second Printing August, 1991

**Reference Manual for the  
Smith Chart Companion  
RF Design Software**

David J. Brunell  
Department of Electrical Engineering  
Florida Atlantic University  
Boca Raton, FL 33431

**Abstract**

The Smith chart is a very useful tool for the analysis of transmission line problems, as well as the design of high-frequency amplifiers. Its utility lies in the fact that relationships between complex impedances and reflection coefficients are represented in an intuitive, graphical way. The price paid for this graphical technique is poor accuracy and interpolation headaches reminiscent of slide rule days. The Smith Chart Companion software has been designed to off-load the computational tasks to a programmable calculator, while maintaining the graphical ease of use inherent in the chart itself.

**Installation**

To run the Smith Chart Companion you need a Hewlett-Packard HP 48 calculator and some method for loading the software, which is provided on IBM-compatible diskettes or in listing format. The program occupies approximately 8k of memory and

does not require any RAM or library plug-in cards. The program should be loaded according to the instructions given in the READ.ME file contained on the distribution diskette. Contact the author for a copy of the software, which is available for \$5 to cover distribution costs.

**Overview**

The Smith Chart Companion program is divided into three main directories: LINE, AMP, and STRIP, corresponding to transmission line, high-frequency amplifier, and microstrip solutions. Within each of these directories is a number of *objects* which can be one of the following: variable, program, solver-type program, or equation. A variable stores numerical data and it may be necessary to modify or inspect variables to provide program input or output. A program is executed by pressing the menu key which is immediately below its name. Data for the program is either contained in pre-existing variables or prompted for during execution. A solver-type program is also executed by pressing the corresponding menu key, but in this case a variable menu is presented and operation proceeds as it would when using the built-in solver application. An equation object is used with the built-in solver by making it the current equation, as described in the HP 48 Owner's Manual. Note that we do not classify objects in strictly the same sense as in the Hewlett-Packard literature, but as to the way the object is used to solve the problem.

There are two excellent books available which cover the use of Smith charts and the design of high-frequency amplifiers [1], [2]. These texts are recommended for reference while familiarizing oneself with the operation of this software. In addition, a

complete amplifier design problem with a key-by-key solution will be presented at the conference and is available separately from the author.

### Program Descriptions

#### Section 1 - Objects in the LINE directory

#### XLINE

Object type: Solver-type program

Variables: ZIN = The input impedance looking into the line  
ZL = The load impedance at the far end of the line  
Z0 = The characteristic impedance of the line  
D/λ = The length of the line in wavelengths

Relevant equation:

$$ZIN = Z0 \frac{ZL + j Z0 \tan(D/\lambda)}{Z0 + j ZL \tan(D/\lambda)}$$

Description: This program allows one to compute the basic parameters associated with impedances on a lossless, terminated transmission line. Upon entering any three of the variables, the fourth one can be solved for with the exception of Z0.

#### REFL

Object type: Solver-type program

Variables: GAMMA = The complex reflection coefficient  
ZIN = The input impedance looking into the line

Z0 = The characteristic impedance of the line

Relevant equation:

$$GAMMA = \frac{ZIN - Z0}{ZIN + Z0}$$

Description: For a given characteristic line impedance, the reflection coefficient at a certain point can be calculated if the input impedance at that point is known. To find the load reflection coefficient, simply let ZIN=ZL. Any unknown variable can be solved for if the other two are known. This process is equivalent to determining the reflection coefficient of a point plotted on the Smith Chart using the angular and radial scales.

#### SWR

Object type: Program

Variables: GAMMA = The complex reflection coefficient of the line  
SWR = The voltage standing wave ratio

Relevant Equation:

$$SWR = \frac{1 + |GAMMA|}{1 - |GAMMA|}$$

Description: Once a value of GAMMA has been found using the REFL program or stored manually, executing the SWR object will display the corresponding vswr value. This is equivalent to reading the vswr radial scale on the Smith chart.

#### OPEN

Object type: Equation



**Variables:** BOC = The susceptance of the open-circuited line  
Y0 = The characteristic admittance of the line (1/Z0)  
D = The length of the line in wavelengths

**Relevant Equation:**  $BOC = Y0 \tan(2\pi D)$

**Description:** This equation relates the parameters associated with admittances on an open-circuited transmission line. Calculator must be in radian mode. Setting Y0=1 allows use of normalized susceptance. Remember that capacitive susceptances are positive.

## SHORT

**Object type:** Equation

**Variables:** BSC = The susceptance of the short-circuited line  
Y0 = The characteristic admittance of the line (1/Z0)  
D = The length of the line in wavelengths

**Relevant Equation:**  $BSC = -Y0 \cot(2\pi D)$

**Description:** This equation relates the parameters associated with admittances on an short-circuited transmission line. Calculator must be in radian mode. Remember that inductive susceptances are negative.

## UNIT

**Object type:** Program

**Variables:** None

**Relevant Equation:** None

**Description:** For use in conjunction with the Smith chart in designing matching networks. Given a normalized, complex impedance or admittance on the stack, calculates the distance in that coordinate system, along a constant resistance or conductance circle, to the opposite system's unit circle, following the two possible paths.

**Example:** Design a matching network to transform an impedance of 25+j10 to 50 ohms. We first normalize the impedance to 0.5+j0.2 and plot this point on the Smith chart. In order to reach the center of the chart (50 ohms), we must first use a series element to move us to the unit admittance circle. To determine the two distances to the unit admittance circle, enter the following keystrokes: (.5 SPC .2 ENTER UNIT. The two paths are (0, 0.3) and (0, -0.7), corresponding to a series inductor and a series capacitor, respectively. If we choose the series inductor (which has a reactance of 0.3\*50=15 ohms) by pressing DROP +, then convert it to an admittance by taking the reciprocal, the resulting value of (1, -1) shows us that we are indeed on the unit admittance circle and that a parallel capacitor of 1/50=0.02S susceptance will complete the match.

## PARL

**Object type:** Program

**Variables:** None

**Relevant equation:**

$$Z_{||} = \frac{1}{\frac{1}{Z_1} + \frac{1}{Z_2}}$$

**Description:** Given two complex impedances on stack levels 1 and 2, this program finds the impedance of the parallel combination. Alternatively, this program can combine two admittances in series. It is useful for performing calculations to verify Smith chart solutions to matching problems.

## XL

**Object type:** Equation

**Variables:** XL = inductive reactance

F = frequency

L = inductance

**Relevant Equation:**

$$XL = 2\pi FL$$

**Description:** This is an equation for use with the solver, and any variable can be found if the other two are known.

## XC

**Object type:** Equation

**Variables:** XC = capacitive reactance (always negative)

F = frequency

C = capacitance

**Relevant Equation:**

$$XC = -\frac{1}{2\pi FC}$$

**Description:** This is an equation for use with the solver, and any variable can be found if the other two are known.

## Section 2 - Objects in the AMP directory

Note that programs in this section which reference the S, Z, Z0, FMIN, GAM0, and RN variables require that appropriate values be stored in these variables before program execution is begun.

## NCIR

**Object type:** Program

**Variables:** Fi = Noise figure in dB of the desired noise circle

FMIN = minimum noise figure of device at chosen freq. and bias point

RN = equivalent noise resistance of transistor

GAM0 = input reflection coefficient corresponding to FMIN

Ni = intermediate parameter used in calculations

CFi = the center of the noise figure circle corresponding to Fi

RFi = radius of noise circle

**Relevant equations:**

$$Ni = \frac{Fi - FMIN}{4RN} |1 + GAM0|^2$$

$$CFi = \frac{GAM0}{1+Ni}$$

$$RFi = \frac{1}{1+Ni} \sqrt{Ni^2 + Ni(1 - |GAM0|^2)}$$

Description: Before this program can be run, the noise parameters of the transistor must be stored in variables FMIN, GAM0, and RN. These parameters are dependent on bias conditions and frequency of operation and are either given by the manufacturer or determined experimentally. The program will prompt for the desired noise figure, which must be greater than or equal to the minimum noise figure. The radius and center of the noise circle is then calculated and displayed. Once this circle is plotted on the Smith chart, any value of source impedance which is within this circle will provide a noise figure better than Ni.

### SCM

Object type: Program

Variables: S = the S-matrix of the device in question

GAM\_Ms = source reflection coefficient necessary for simultaneous conjugate match

GAM\_ML = load reflection coefficient necessary for simultaneous conjugate match

Relevant equations:

$$GAM_{Ms} = \frac{B_1 \pm \sqrt{B_1^2 - 4|C_1|^2}}{2C_1}$$

$$GAM_{ML} = \frac{B_2 \pm \sqrt{B_2^2 - 4|C_2|^2}}{2C_2}$$

$$B_1 = 1 + |S_{11}|^2 - |S_{22}|^2 - |\Delta|^2$$

$$B_2 = 1 + |S_{22}|^2 - |S_{11}|^2 - |\Delta|^2$$

$$C_1 = S_{11} - \Delta S_{22}^*$$

$$C_2 = S_{22} - \Delta S_{11}^*$$

Description: This program computes the source and load reflection coefficients necessary for a simultaneous conjugate match at the input and output of a bilateral transistor amplifier stage. Before running this program, the S-matrix must be set up with the appropriate values.

### GLCIR

Object type: Program

Variables: GL = gain associated with output matching network

S = S-matrix of transistor

gL = normalized gain factor

RL = radius of gain circle

dL = distance from center of chart to center of gain circle

Relevant equations:

$$gL = GL(1 - |S_{22}|^2)$$

$$RL = \frac{\sqrt{1 - gL}(1 - |S_{22}|^2)}{1 - |S_{22}|^2(1 - gL)}$$

$$dL = \frac{gL |S_{22}|}{1 - |S_{22}|^2(1 - gL)}$$

Description: This program computes the location and radius of the constant-gain circle corresponding to a gain of GL. The center of the circle is a distance of dL from the center of the Smith chart and lies along a line drawn from the center of the chart to the point  $S_{22}$ .

### GSCIR

Object type: Program

Variables: Gs = gain related to input matching

S = S-matrix of transistor

gs = normalized gain factor

Rs = radius of gain circle

ds = distance from center of chart to center of gain circle

Relevant equations:

$$gs = Gs(1 - |S_{11}|^2)$$

$$Rs = \frac{\sqrt{1 - gs}(1 - |S_{11}|^2)}{1 - |S_{11}|^2(1 - gs)}$$

$$ds = \frac{gs |S_{11}|}{1 - |S_{11}|^2(1 - gs)}$$

Description: This program computes the location and radius of the constant-gain circle corresponding to a gain of Gs. The center of the circle is a distance of ds from the center of the Smith chart and lies along a line drawn from the center of the chart to the point  $S_{11}$ .

### GAMIN

Object type: Program

Variables: GAML = the load reflection coefficient

GAMIN = the input reflection coefficient

S = the S-matrix of the two-port.

Relevant equation:

$$GAMIN = S_{11} + \frac{S_{12}S_{21}GAML}{1 - S_{22}GAML}$$

Description: Given the load reflection coefficient corresponding to a prescribed output termination, this program calculates the reflection coefficient at the input to the transistor or two-port.

## GAMOUT

Object type: Program

Variables:  $GAMs$  = the source reflection coefficient  
 $GAMOUT$  = the input reflection coefficient  
 $S$  = the s-matrix of the two-port.

Relevant equation:

$$GAMOUT = S_{22} + \frac{S_{12}S_{21}GAMs}{1 - S_{11}GAMs}$$

Description: Given the source reflection coefficient corresponding to a prescribed input termination, this program calculates the reflection coefficient at the output of the transistor or two-port.

## GTMAX

Object type: Program

Variables:  $K$  = stability factor of two-port  
 $GTmax$  = maximum transducer power gain of two-port  
 $S$  = the S-matrix of the two-port

Relevant equation:

$$GTmax = \frac{|S_{21}|}{|S_{12}|} (K - \sqrt{K^2 - 1})$$

Description:  $GTmax$  is the maximum transducer power gain available from a given two-port. It is calculated under the assumption that the amplifier is unconditionally stable. Before running this program, run the STAB program to determine stability conditions. The value of  $K$  is automatically calculated and need not be entered.

## GPCIR

Object type: Program

Variables:  $Gp$  = power gain for which circle is to be drawn  
 $S$  = S-matrix of the two-port  
 $gp$  = normalized power gain  
 $Rp$  = radius of power gain circle  
 $Cp$  = center of power gain circle

Relevant equations:

$$gp = \frac{Gp}{|S_{21}|^2}$$

$$Rp = \frac{[1 - 2K|S_{12}S_{21}|gp + |S_{12}S_{21}|^2gp^2]^{1/2}}{|1 + gp(|S_{22}|^2 - |\Delta|^2)|}$$

$$Cp = \frac{gpC_2^*}{1 + gp(|S_{22}|^2 - |\Delta|^2)}$$

Description: For cases where the two-port cannot be considered unilateral, the power gain circles are plotted on the load reflection coefficient chart. The circles are independent of the input matching network and will equal a conjugate match at the input.

## UFM

Object type: program

as:  $S$  = S-matrix of the two-port

$U$  = unilateral figure of merit

$G_T$  = transducer power gain

$G_{TU}$  = unilateral transducer power gain

Key equations:

$$\frac{1}{(1+U)^2} < \frac{G_T}{G_{TU}} < \frac{1}{(1-U)^2}$$

$$U = \frac{|S_{11}| |S_{12}| |S_{21}| |S_{22}|}{(1-|S_{11}|^2)(1-|S_{22}|^2)}$$

Conclusion: When a transistor amplifier is not unilateral, it is often desirable to consider what uncertainty in gain can be expected by analyzing the amplifier as if it were unilateral. High and low bounds can be placed on the ratio of bilateral to unilateral transducer power gain as shown in the equation above. This program will compute the value of  $U$  (lower values signify lower uncertainty), and display the minimum and maximum values for  $G_T/G_{TU}$ .

Program

as:  $S$  = S-matrix of the two-port

$K$  = stability factor of the two-port

$\Delta$  = determinant of the S-matrix

$r_s$  = radius of the input stability circle

$C_s$  = center of input stability circle

$r_L$  = radius of the output stability circle

$CL$  = center of output stability circle

Relevant equations:

$$K = \frac{1-|S_{11}|^2-|S_{22}|^2+|\Delta|^2}{2|S_{12}S_{21}|}$$

$$r_L = \frac{|S_{12}S_{21}|}{|S_{22}|^2-|\Delta|^2}$$

$$CL = \frac{(S_{22}-\Delta S_{11}^*)^*}{|S_{22}|^2-|\Delta|^2}$$

$$r_s = \frac{|S_{12}S_{21}|}{|S_{11}|^2-|\Delta|^2}$$

$$C_s = \frac{(S_{11}-\Delta S_{22}^*)^*}{|S_{11}|^2-|\Delta|^2}$$

Description: This program should be run first before any amplifier design is attempted.

It will determine under which conditions the amplifier will be stable. It can be shown [2] that a necessary and sufficient condition for unconditional stability is that  $K > 1$  and  $|\Delta| < 1$ . The program first calculates the values of  $K$  and  $|\Delta|$ , then determines whether or not the transistor is unconditionally stable. If it is, then a message indicating that is displayed

and execution stops. If the transistor is not unconditionally stable, then the input and output stability circles are found and displayed, and in each case (input and output) indicates whether or not the center of the Smith represents a stable point. The stability circles represent the boundary between stable and unstable operation.

### S->Z

Object type: Program

Variables: S = S-matrix of two-port

Z = Z-matrix of two-port

Z0 = Reference impedance matrix

Description: This program converts an S-matrix to its equivalent Z-matrix representation. Z0 is the reference impedance matrix, where Z0<sub>11</sub> is the input reference impedance, Z0<sub>22</sub> is the output reference impedance, and the other elements are zero. Normally, Z0 is left as  $\begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}$ .

### Z->S

Object type: Program

Variables: S = S-matrix of two-port

Z = Z-matrix of two-port

Z0 = Reference impedance matrix

Description: This program converts a Z-matrix to its equivalent S-matrix representation.

Z0 is the reference impedance matrix, where Z0<sub>11</sub> is the input reference

impedance, Z0<sub>22</sub> is the output reference impedance, and the other elements are zero. Normally, Z0 is left as  $\begin{bmatrix} 50 & 0 \\ 0 & 50 \end{bmatrix}$ .

### Section 3 - Objects in the STRIP directory

#### IMP

Object type: Equation

Variables: Z0 = characteristic impedance of microstrip transmission line

εR = relative permittivity of dielectric substrate material

H = height of microstrip above ground plane

W = width of microstrip

εEFF = effective relative permittivity

F1, A, B = intermediate variables

Relevant equations [3]:

$$Z_0 = \frac{60}{\sqrt{\epsilon_{eff}}} \ln \left[ \frac{f_1 h}{w} + \sqrt{1 + \left( \frac{2h}{w} \right)^2} \right]$$

$$f_1 = 6 + (2\pi - 6) \exp \left\{ - \left[ \left( \frac{30.666h}{w} \right)^{0.753} \right] \right\}$$

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left( 1 + \frac{10h}{w} \right)^{-0.5}$$

$$a = 1 + \frac{1}{49} \ln \left[ \frac{\left(\frac{w}{h}\right)^4 + \left(\frac{w}{52h}\right)^2}{\left(\frac{w}{h}\right)^4 + 0.432} \right] + \frac{1}{18.7} \ln \left[ 1 + \left(\frac{w}{18.1h}\right)^3 \right]$$

$$b = 0.564 \left[ \frac{\epsilon_r - 0.9}{\epsilon_r + 3} \right]^{0.53}$$

Description: The IMP equation allows use of the solver in determining unknown parameters of microstrip transmission lines. IMP has already been made the current equation in the STRIP directory, so to access it press <right shift> SOLVER and the menu of variables will appear. After solving for the desired parameter, hit the VAR key to return to the variable menu. At this point, pressing εEFF then EVAL will compute the effective relative permittivity. The guided wavelength is now found using the formula  $\lambda_g = \lambda_0 / \sqrt{\epsilon_{EFF}}$ .

### COUP

Object type: Directory

Description: COUP is a subdirectory of the STRIP directory and contains further objects relating to the solution of coupled microstrip lines. Description of these other objects follows.

### Z0E

Object type: Equation

Variables: Z0E = Odd mode characteristic impedance

Z0 = impedance of equivalent single microstrip line

C = desired coupling factor in -dB (eg., -10dB)

Relevant equation:

$$Z0E = Z0 \sqrt{\frac{1 + 10^{\frac{C}{20}}}{1 - 10^{\frac{C}{20}}}}$$

Description: This equation be used in conjunction with the solver to determine the value of Z0E necessary for the desired degree of coupling C. Z0O (odd mode impedance) is then found from  $Z0O = Z0^2 / Z0E$ .

### COUP

Object type: Program

Variables: u\_se = W/H for a single microstrip with Z0=Z0E/2

u\_so = W/H for a single microstrip with Z0=Z0O/2

εr = relative permittivity of substrate material

u = W/H for coupled area of microstrip

g = S/H where S=spacing of coupled lines

Relevant equations: see [4]

Description: This program determines u and g, the normalized width and spacing of the coupled microstrip lines. Before running this routine, the values for u\_se and u\_so should be first determined with the IMP equation described in



the previous section.

### $\epsilon_{RE}$

Object type: Expression

Variables:  $\epsilon_{RE}$  = the relative permittivity for the even coupled mode

Relevant equation: see [3]

Description: Pressing the  $\epsilon_{RE}$  menu key followed by <EVAL> will calculate and display the even coupled mode relative permittivity

### $\epsilon_{RO}$

Object type: Expression

Variables:  $\epsilon_{RO}$  = the relative permittivity for the odd coupled mode

Relevant equation: see [3]

Description: Pressing the  $\epsilon_{RO}$  menu key followed by <EVAL> will calculate and display the odd coupled mode relative permittivity

### References

- [1] P. Smith, *Electronic Applications of the Smith Chart*, New York: McGraw-Hill, 1969.
- [2] G. Gonzalez, *Microwave Transistor Amplifiers*, Englewood Cliffs: Prentice Hall, 1984.
- [3] E. Hammerstad and O. Jensen, "Accurate models for microstrip computer-aided design," *IEEE MTT-S Digest*, 1980.
- [4] S. Akhtarzad, T. R. Rowbotham, and P. Johns, "The design of coupled microstrip lines," *IEEE Trans. Microwave Theory and Techniques*, vol. 23, no. 6, June 1975.

Faint, illegible text, possibly bleed-through from the reverse side of the page.

# Comparing Simulation Methods for RF Designs

Xiaohua Xuan  
Hewlett-Packard  
1400 Fountaingrove Parkway  
Santa Rosa, CA 95403

## ABSTRACT

HP recently introduced a new circuit simulator called *HP Impulse* that overcomes some of the limitations of SPICE for the high-frequency designer. This survey of high-frequency circuit simulation techniques discusses how HP's new simulator differs from harmonic balance and SPICE. These differences are illustrated in several common applications.

## TRANSIENT CIRCUIT SIMULATION

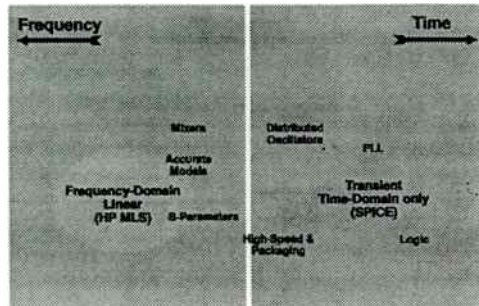
- HP has introduced a new simulator "HP Impulse"
- How is HP Impulse different from SPICE or HP MNS?
- Which simulator is appropriate in certain applications?



HP Impulse is a new kind of circuit simulator for the high-frequency designer. It overcomes some of the weaknesses of existing simulators and has its own strengths and weaknesses. This paper reviews the theory underlying HP Impulse and other simulators. Based on this information, the student is led through a series of common high-frequency applications and is asked to identify the simulation techniques that are (and are not) appropriate for the task.

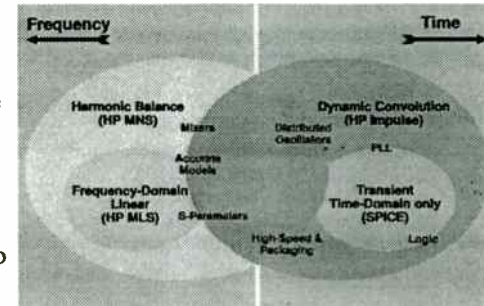
## YESTERDAY'S HIGH FREQUENCY CIRCUIT SIMULATION TOOLS

- Tools excluded large classes of useful circuits
- Drastic assumptions needed to simplify problems so that tools could be used



## TODAY'S HIGH-FREQUENCY CIRCUIT SIMULATION TOOLS

- New tools supercede the old tools
  - Provide more sophisticated information
  - Overlap in some areas
- Complement each other, do not compete



In the recent past, high-frequency designers had two primary simulation tools: frequency-domain linear simulators and SPICE.

In the late 1980s, harmonic balance was added. It is a frequency-domain tool that extends the older, linear simulators to nonlinear circuits.

With HP Impulse, there is now a high-frequency adaptation of SPICE. HP Impulse is a superset of SPICE that adds important capabilities for the high-frequency designer (notably the ability to use frequency-domain models and data in the time domain). It is based on the new dynamic convolution technique. HP Impulse and HP MNS together are therefore modernized evolutions of the previous generation of circuit simulation tools.

## ABOUT HP IMPULSE

- HP Impulse complements HP MNS and is a high-frequency replacement for SPICE
- HP Impulse has 2 key contributions vs. SPICE:
  - Use S-parameter data directly
  - Accurate models for microstrip, etc.
- HP Impulse is fully integrated into HP MDS



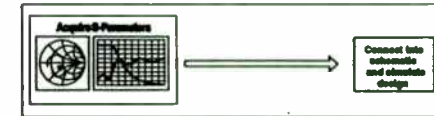
www.hp.com/go/impulse

## HP IMPULSE VS. SPICE

EXAMPLE: USING S-PARAMETER FREQUENCY RESPONSE

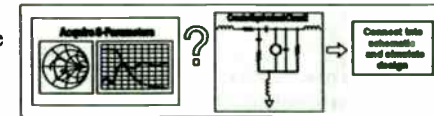
### HP Impulse

Linear procedure  
Excellent accuracy  
Requires no knowledge of internal circuit



### SPICE

Uncertain procedure  
Requires complete knowledge of equivalent circuit

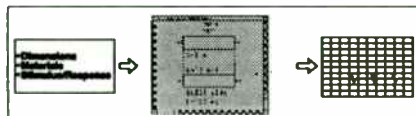


www.hp.com/go/impulse

## HP IMPULSE VS. SPICE EXAMPLE: COUPLED PC BOARD (MICROSTRIP) TRACES

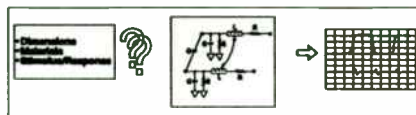
### HP Impulse

Linear procedure  
Excellent accuracy  
Saves modeling time



### SPICE

Uncertain procedure  
Questionable accuracy



## MAJOR SIMULATOR CATEGORIES

- **SPICE** - Simulates all components exclusively in the time domain.
- **HARMONIC BALANCE** - Simulates most components in the frequency domain. Nonlinear components are simulated in the time domain and transformed to the frequency spectra.
- **DYNAMIC CONVOLUTION** - Simulates most components in the time domain. Frequency-domain components are transformed to impulse responses.



The following slides will demonstrate that harmonic balance and dynamic convolution are symmetric in many ways with respect to their use of frequency and time domain. Harmonic balance solves its admittance matrix in the frequency domain. It enforces KVL and KCL at each frequency (spectral line).

The SPICE and dynamic convolution techniques solve their admittance matrices in the time domain. They enforce KVL and KCL at each time point. When viewed at this level of complexity, all simulators are quite similar.

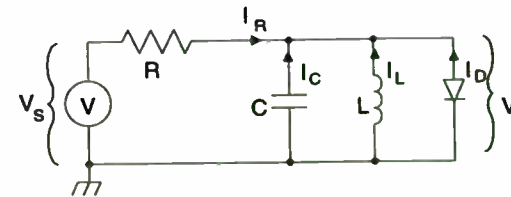
## DEFINITION OF TERMS

- **KIRCHHOFF'S CURRENT LAW (KCL):**  
The net sum of the currents flowing into a node is zero.
- **KIRCHHOFF'S VOLTAGE LAW (KVL):**  
The sum of the voltages around any closed loop is zero.



www.hewlett-packard.com

## TIME-DOMAIN SIMULATION (SPICE)



Kirchhoff's Current Law (evaluated at each time point)

$$I_R + I_C + I_L + I_D = 0$$

$$\frac{(V_L - V_s)}{R} + C \frac{dV_L}{dt} + \frac{1}{L} \int V_L dt + I_{SAT} (e^{V_L/V_T} - 1) = 0$$

Iterative method  
required

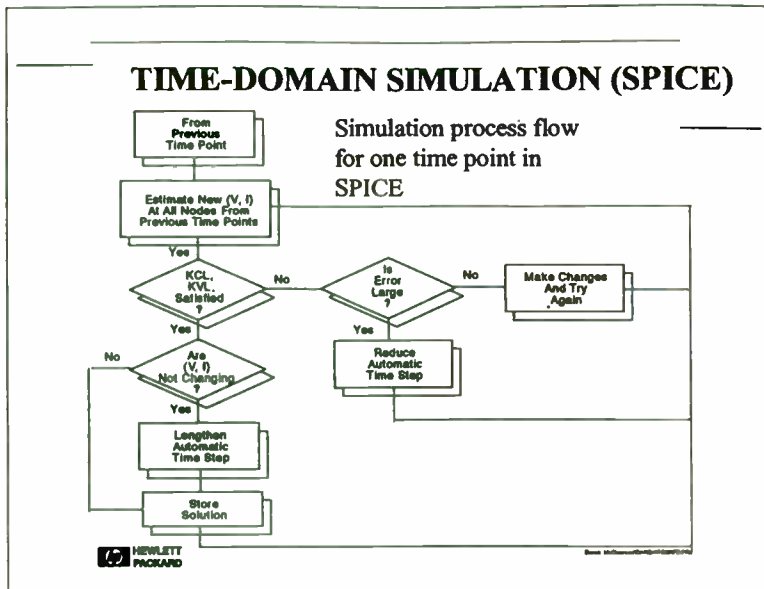


www.hewlett-packard.com

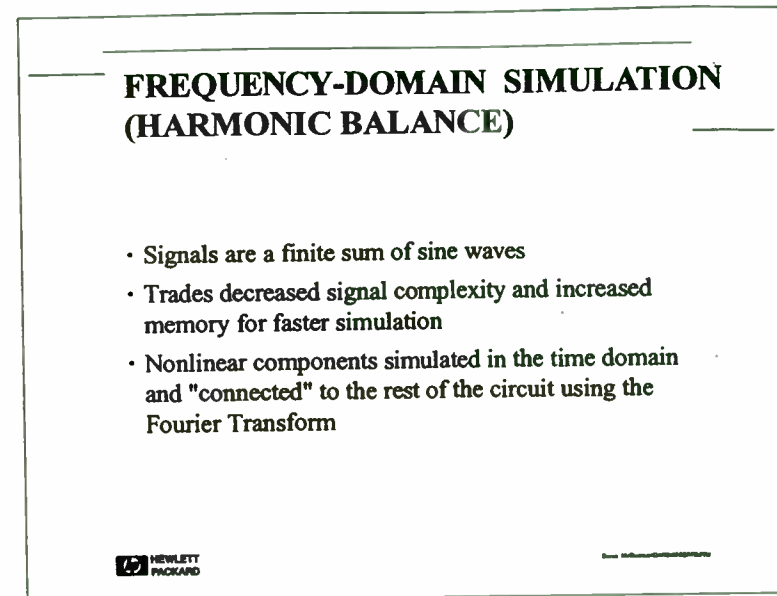
Most circuit simulators factor an admittance matrix to solve the matrix equation  $[I] = [Y][V]$ . By using an admittance matrix, the software designer is inherently satisfying two fundamental laws of circuit theory, Kirchhoff's Current Law (KCL) and Kirchhoff's Voltage Law (KVL).

KCL states that the net sum of the currents into a node is zero. KCL implies that no charge is created or destroyed anywhere in the circuit. The magnitude of residual error currents, summed over all the nodes of a circuit, is a common measure of the accuracy of a simulation. KVL is another measure of simulation accuracy. KVL states that the sum of the voltages around any closed loop of circuit nodes is zero.

This circuit shows that the sum of the currents entering the "output" node (at each point in time) is zero. If there were no nonlinear elements in the circuit, the instantaneous currents through the resistor, capacitor, and inductor could be calculated directly using the present voltage and current conditions, along with conditions in the recent past. However, the current through the diode depends on the other currents because of the voltage drop across the resistor. Therefore, the complete set of voltages and currents is determined using an iterative process.



A transient simulation is a history of the voltages and currents over an interval of "time," evaluated at a sequence of discrete time steps. The flow chart indicates what happens for one time point. Note the introduction of an adaptive time step algorithm. Over the course of a simulation, the software recognizes periods of inactivity and performs fewer analyses. When the voltages and currents are changing rapidly, perhaps during the rising edge of a clock signal, the simulator performs many more analyses. Optimized especially for logic circuits, the adaptive time step algorithm can increase simulation speed significantly. There is less benefit for circuits driven by sine waves.



Harmonic balance is able to calculate the steady-state solution directly, bypassing any transient settling or start-up intervals. Harmonic balance is a frequency-domain approach. Its primary assumption is that all signals are a summation of a finite number of sine waves, which implies that all signals are periodic. This is a reasonable assumption for components one might characterize using a network analyzer or spectrum analyzer.



## FREQUENCY-DOMAIN SIMULATION (HARMONIC BALANCE)

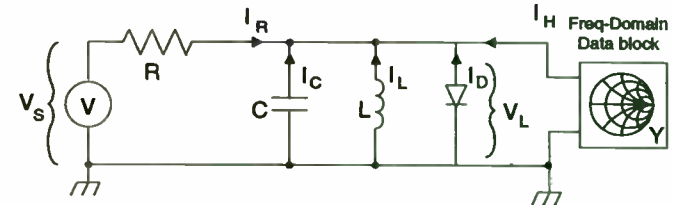
- The DC operating point of the circuit is calculated.
- A linear AC analysis is used to estimate the high-frequency voltage and current spectrums.
- KVL and KCL are applied to the estimates for each frequency of the spectrum.
- If KVL and KCL are not satisfied, the voltage and current spectrums are adjusted and the process repeats.
- The solution is stored.



© 1987 Hewlett-Packard Company

Harmonic balance calculates voltage and currents using iterative, matrix-driven techniques much like time-domain simulators. One important difference is that harmonic balance must solve for the magnitudes and phases of all spectral lines simultaneously, leading to high memory usage. This is equivalent to solving for the entire waveform at once. If the signals are simple in the frequency domain (such as an amplifier driven by a sine wave), then harmonic balance is extremely efficient.

## FREQUENCY-DOMAIN SIMULATION (HARMONIC BALANCE)



Kirchhoff's Current Law (evaluated at each frequency):

$$\frac{(V_L - V_s)}{R} + j\omega V_L C + \frac{V_L}{j\omega L} + \mathcal{S}\{i_D(t)\} + V_L Y_{11} = 0$$

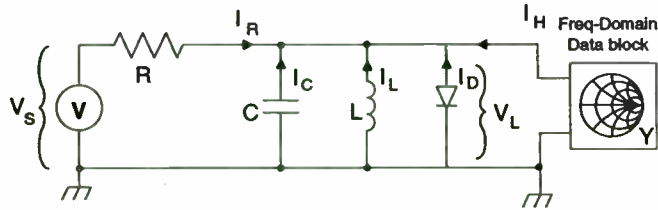
Iterative solution  
required



© 1987 Hewlett-Packard Company

This circuit adds a 1-port data device to the circuit of an earlier slide illustrating the operation of SPICE. It shows that the sum of the currents entering the "output" node (at each spectral line in the frequency domain) is zero. The currents through the resistor, capacitor, inductor, and data block are calculated in closed form with simple algebra, as opposed to the numeric differentiation and integration needed in the time domain. In harmonic balance, the iterating is done in the frequency domain using voltage and current spectrums instead of instantaneous time values of voltage and current.

## FREQUENCY-DOMAIN SIMULATION (HARMONIC BALANCE)



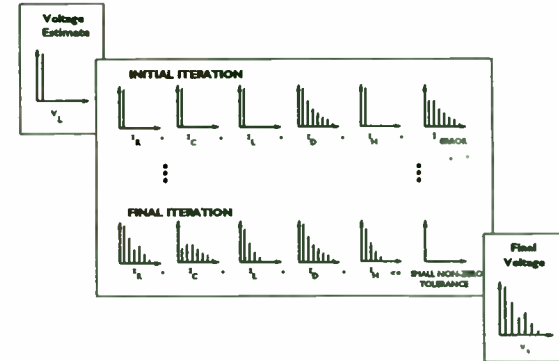
Kirchhoff's Current law. (solved iteratively for all frequencies)

$$\begin{aligned}
 1 \text{ GHz: } & i_k + i_c + i_l + i_b + i_h = 0 \\
 2 \text{ GHz: } & i_k + i_c + i_l + i_b + i_h = 0 \\
 3 \text{ GHz: } & i_k + i_c + i_l + i_b + i_h = 0 \\
 \vdots & \\
 \vdots &
 \end{aligned}$$



Doc 54890A01-00/000100

## CONVERGING TOWARD A SOLUTION (HARMONIC BALANCE)



Doc 54890A01-00/000100

Continuing from the previous slide, the instantaneous current through the diode depends on the instantaneous voltage drop across the resistor. The node voltage waveform is transformed to the time domain and applied to the diode equation to obtain a current waveform. The diode current waveform is then transformed back to the frequency domain, where it is combined with the other currents and tested against KCL and KVL.

Since the net current into each node is driven toward zero to satisfy KCL, the frequency spectrum of these error currents must also go to zero. Therefore, KCL is satisfied implicitly at each spectral frequency at every node.

It was noted in a previous slide that an iterative process was needed to calculate the voltages and currents throughout the circuit. In harmonic balance, the iterating is done in the frequency domain using voltage and current spectra. If one ignores the presence of the Fourier Transforms, a voltage spectrum is input to the nonlinear diode, and a current spectrum results. Therefore, it can be stated that the simulator progressively refines its estimate of the voltage spectrum until KCL and KVL are satisfied. When this happens, all of the harmonics are "balanced," hence the name of the technique.

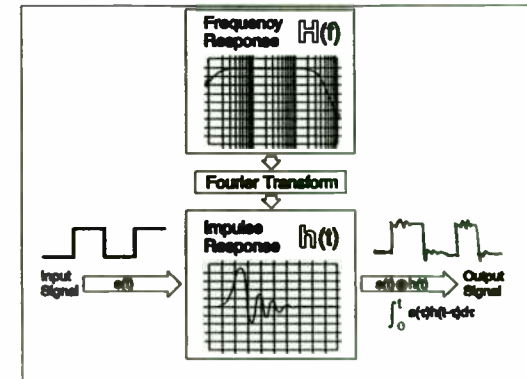
## TIME-DOMAIN SIMULATION (HP IMPULSE)

- Extends functionality of SPICE using "dynamic convolution" technique.
- The frequency responses of all S-parameter blocks and dispersive components are converted to time-domain finite impulse responses.
- Each response is convolved with the signal passing through it to obtain the time-domain response.
- Active devices are evaluated in the time-domain.



From HP Impulse User's Guide

## WHAT IS CONVOLUTION? (TIME DOMAIN)



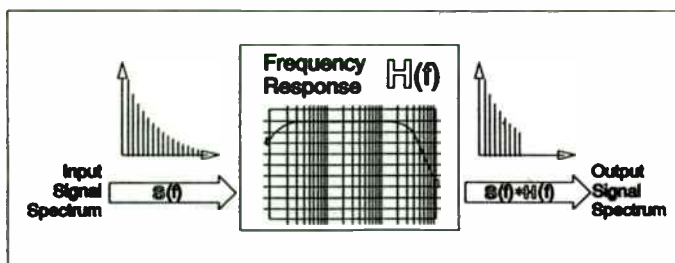
From HP Impulse User's Guide

HP Impulse is a new type of circuit simulator that overcomes some of the limitations of SPICE for the high-frequency designer without giving up any of the advantages of SPICE. It uses the "dynamic convolution" simulation technique, developed commercially by HP's High-Frequency Design Software operation (Santa Rosa, California).

Like SPICE, HP Impulse is a time domain simulator. However, it also allows frequency-domain components (such as microstrip transmission lines and measured S-parameters) to be used in the time domain. The integration of HP Impulse into the HP Microwave Design System creates additional benefits. In the absence of frequency-domain components, one would use HP Impulse in much the same manner as SPICE.

Convolution is a mathematical process that maps the characteristics of one waveform onto another. One property of the Fourier Transform is that convolution in either the time or frequency domain is analogous to algebraic multiplication in the other domain. Thus convolution of a time-domain signal with the impulse response of a network is equivalent to multiplying the frequency spectrum of the signal with the frequency response of the network.

## WHAT IS CONVOLUTION? (FREQUENCY DOMAIN)



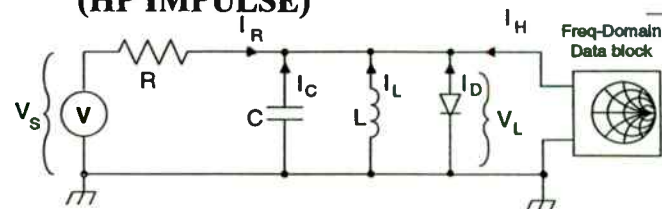
HEWLETT  
PACKARD

Doc 11499-9001-0000000000

Ideally, multiplication in the frequency domain is equivalent to convolution in the time domain. But there are subtle differences when the discrete-time Fourier summations are not infinite. Since transient signals are only meaningful in the time domain, HP Impulse does not try to transform them to the frequency domain. Rather, frequency-domain circuit elements are transformed to time domain equivalents and convolved with the signals passing through them.

This limits some of the frequency-domain components that can be used in HP Impulse, but allows the time-domain signals to be completely arbitrary. In some ways, this is symmetric with the treatment of time-domain (nonlinear) elements in harmonic-balance.

## TIME-DOMAIN SIMULATION (HP IMPULSE)



Kirchhoff's Current Law (evaluated at each time point):

$$I_R + I_C + I_L + I_D + I_H = 0$$

$$\frac{(V_L - V_s)}{R} + C \frac{dV_L}{dt} + \frac{1}{L} \int V_L dt + I_{SAT} (e^{V_L/V_T} - 1) + \int_0^t V_L(\tau) Y_{11}(\tau - t) d\tau = 0$$

Iterative solution required      Convolution

HEWLETT  
PACKARD

Doc 11499-9001-0000000000

This circuit is a time-domain interpretation of an earlier slide illustrating the operation of harmonic balance. It shows that the sum of the currents entering the "output" node (at each point in time) is zero. The data block simply adds another term to the equation.

Dynamic convolution is used to calculate the instantaneous current "IH" of the data block. First, the magnitude and phase of  $Y_{11}$  vs. frequency are converted to an impulse response. Then the history of the node voltage is convolved with the time domain version of  $Y_{11}$  to calculate the instantaneous current. KCL is applied and the simulator iterates toward a self-consistent solution as described in earlier slides. In most other respects, the internal operation of HP Impulse resembles SPICE.

## ANALYSIS STRENGTHS

HP Impulse (Time-Domain)	HP MNS (Frequency-Domain)
Arbitrary signals	Arbitrary frequency response
Transient response	Steady-state response
Boolean/logic states	Behavioral system models
Switched/chopped circuits	High-Q/reverberant circuits



HP Impulse and HP MNS are trademarks of Hewlett-Packard Company.

## ANALYSIS STRENGTHS

- HP Impulse overcomes important limitations of SPICE without giving up any of SPICE's advantages.
  - Use S-parameter data directly
  - Accurate models for microstrip, etc.
- For some applications, harmonic balance (frequency domain) is still the best choice.
- For all other high-frequency circuits, HP Impulse is a direct replacement for SPICE.



HP Impulse and HP MNS are trademarks of Hewlett-Packard Company.

Signal complexity and desired output parameters determine which simulator should be used. Stimulus/response and swept-parameter simulations usually indicate that harmonic balance should be used. Harmonic balance is well-suited to tests that are done with one or two sine wave stimuli, such conversion loss vs. drive level, distortion vs. frequency, and gain compression.

If the signals become more complex or involve boolean logic states, time-domain simulation (HP Impulse) should be considered. HP Impulse is excellent for testing non-repeating transient behavior, indicated by the use of an arbitrary waveform generator or a wide-bandwidth oscilloscope. HP Impulse is also useful for verifying the dynamic performance of a design after it has been optimized and tested using harmonic balance.

If you could only use one circuit simulation tool for all the high-frequency work you did, you might choose the time-domain tool, HP Impulse (dynamic convolution). SPICE became the industry standard circuit simulator because of its versatility, and HP Impulse only extends the capabilities of SPICE. Conversely, HP MNS, is extremely efficient for many frequency-oriented circuits, but an equally bad choice for other circuits. In situations where HP MNS is efficient, HP MNS is usually chosen instead of HP Impulse. An average high-frequency designer would typically use both HP Impulse and HP MNS. Other electronic disciplines might favor HP Impulse more heavily.

## APPLICATION SUMMARY

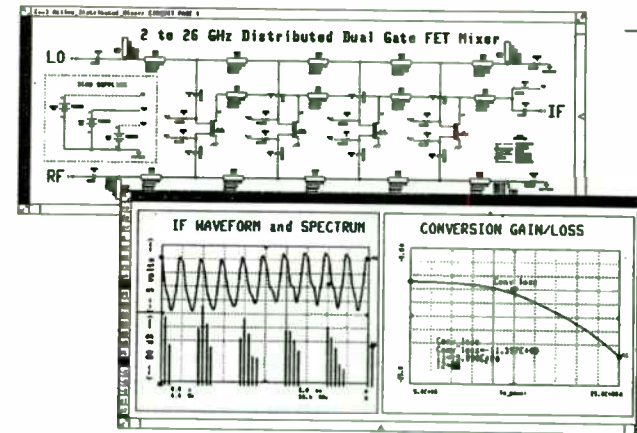
- Distributed GaAs FET mixer
- MMIC traveling wave amplifier
- High-Q SAW oscillator
- Low-Q oscillator
- 4-Quadrant analog multiplier IC
- General purpose analog circuits
- TTL/CMOS phase-locked loop
- Logic signal integrity



Draw: M01000000000000000000

The second half of this High-Frequency Design Forum paper offers several application examples for discussion. HP Impulse, HP MNS, or both simulators can be applied to each circuit, depending on the application and the nature of the performance tests. The examples sample a continuum of high-frequency, analog, and digital applications. They amplify the differences between HP Impulse and HP MNS and are not intended to represent either product completely.

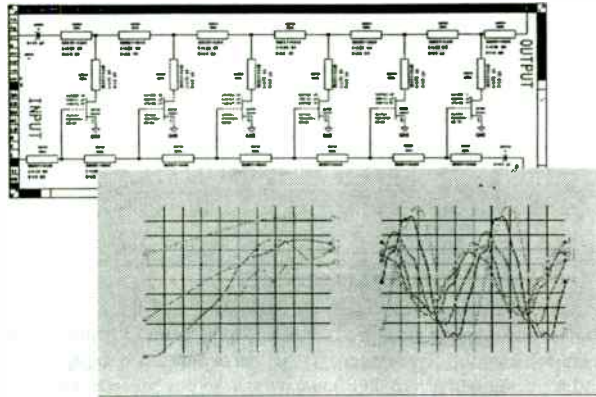
## DISTRIBUTED GaAs FET MIXER



Draw: M01000000000000000000

Because of the need to simulate both the IF and second order frequencies, time-domain simulators require many time points and are usually slower than harmonic balance. In the figure, a plot of conversion efficiency vs. LO power level is simple for harmonic balance but time-consuming for HP Impulse. HP Impulse would be useful for situations where the baseband signal is not a sine wave, or if there are multiple heterodyned frequency translations elsewhere in the circuit. In these circumstances, the memory usage and analysis times of harmonic balance can increase, making the time-domain approach more attractive.

## MMIC TRAVELING WAVE AMPLIFIER

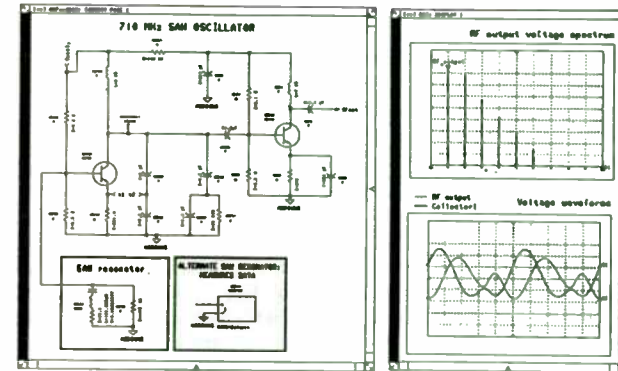


HEWLETT  
PACKARD

Source: Multisim@hp.com/2000/02/04/20000204

Similar in character to the distributed active mixer, the TWA has been used heavily to promote the advantages of harmonic balance, precisely because of the inadequacies of SPICE. The plot of harmonics vs. power level is best done with HP MNS. If the amplifier is pulsed, stimulated with an arbitrary waveform, dynamically unstable, or used in an automatic gain control circuit, HP Impulse becomes the superior tool.

## HIGH-Q SAW OSCILLATOR



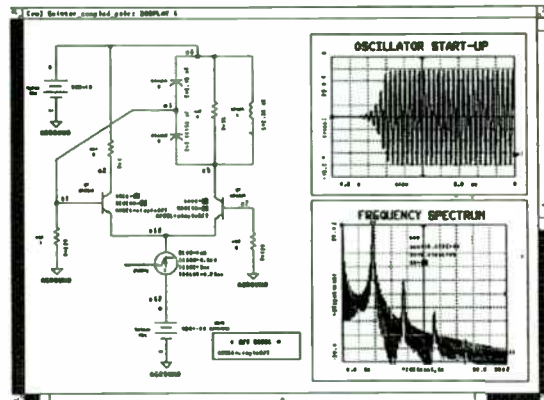
HEWLETT  
PACKARD

Source: Multisim@hp.com/2000/02/04/20000204

Oscillators may require on the order of "Q" cycles to settle to steady state conditions of frequency stability and power level. For this reason, designers typically lower the Q to make the transient simulation reasonable, or simply use frequency-domain tools and techniques. High-Q oscillators are often designed using linear Smith Chart principles without any nonlinear simulation.

HP Impulse can be used to verify that the circuit oscillates at the desired frequency, starting from power-on transients at time=0. Spurious resonances can exist in external bias circuitry and undesired overtones of some crystals, if modeled properly in the design environment.

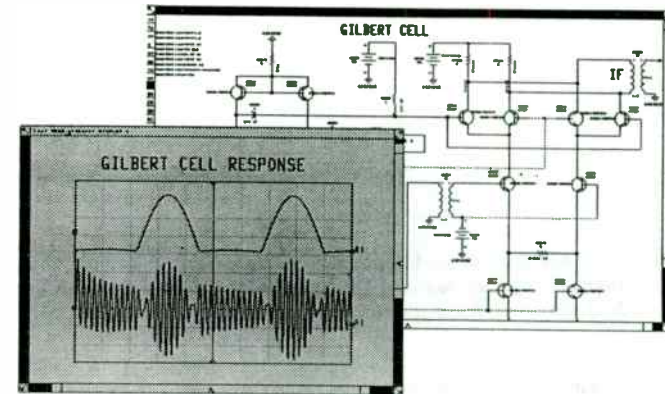
## LOW-Q OSCILLATOR



HEWLETT  
PACKARD

Simulation Results

## 4-QUADRANT ANALOG MULTIPLIER IC



HEWLETT  
PACKARD

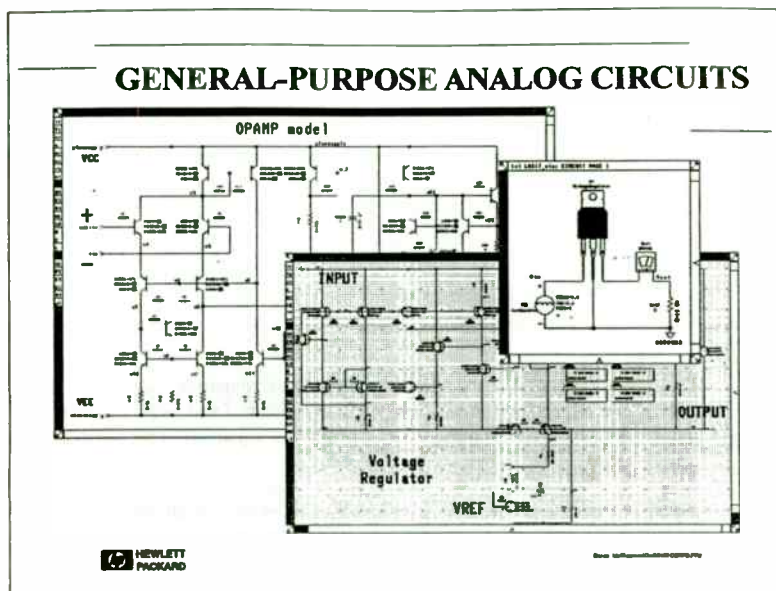
Simulation Results

HP Impulse is an excellent choice for designing low-Q oscillators because it can account for all feedback paths and sources of instability. HP MNS will only find an oscillation if the user asks for the oscillator simulation mode and places a special probe in the circuit. HP Impulse can easily verify tuning, frequency pulling due to dynamic conditions, subharmonics, spurious modes, start-up transients, and even inadvertent amplifier instability, all aspects of oscillator design that HP MNS is not able to model. HP Impulse also provides accurate models for designing distributed-element, high-frequency oscillators in microstrip, waveguide, etc.

This analog multiplier circuit is used as a mixer or as a gain control element. In its target application, frequency-selective filters and distributed PC board traces surround this IC. HP Impulse can simulate these circuit elements, allowing the designer to use the IC in its target application. This might be difficult in SPICE.

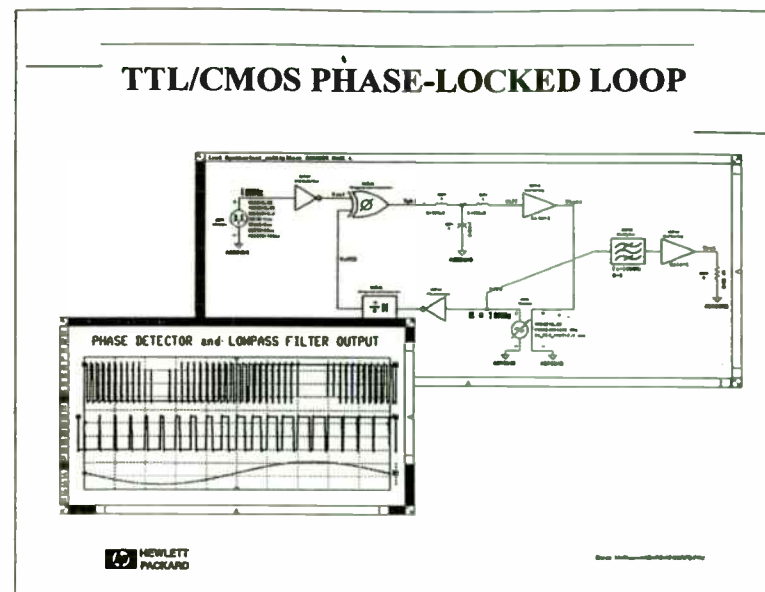
With sine wave inputs, the circuit can also be simulated using HP MNS. If the circuit is part of an AGC loop, has arbitrary waveform inputs, or is part of a larger circuit with several mixers and independent frequencies, HP MNS is less useful. Because HP Impulse can simulate the circuit over the widest range of conditions and applications, it is an excellent verification tool to use on a mature design.





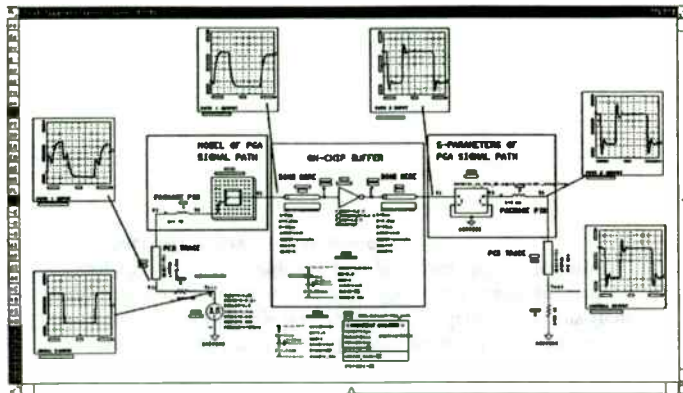
Shown here are two representative analog circuits, an operational amplifier and a voltage regulator. HP Impulse is clearly the preferred simulator, even if HP MNS is useful in some way. These circuits do not use any of the special features of HP Impulse, so the SPICE results are quite comparable to the HP Impulse results.

There are two reasons why one might use HP Impulse instead of SPICE for general-purpose analog design. HP Impulse allows you to consolidate both high- and low-frequency design in one environment so you can combine elements of both in the same simulation. Also, the HP Microwave Design System has a dedicated post-processor, links to test equipment, modeling tools, and other features that may be of assistance.



This phase-locked loop is configured as a frequency multiplier. It includes a mixture of several contradictory features that prevent it from being simulated by any other general-purpose circuit simulator, including HP MNS or SPICE. It combines logic and analog functions, frequency-domain and time-domain behavioral components, high-frequency and low-frequency circuits, lumped elements, and possibly distributed elements. HP Impulse could be used to simulate the transient response and capture range, and even document typical waveforms for benchtop tuning.

## LOGIC SIGNAL INTEGRITY



This circuit represents a signal path through a printed circuit board and a ceramic pin grid array (PGA). One of the signal paths is modeled using microstrip and stripline components. A different signal path has been characterized with S-parameter data. This simulation tests signal integrity directly without the use of equivalent RLGC networks for the transmission lines. It also includes nonlinear circuits that react to the voltage waveforms inside the IC to change the loading and fan-out dynamically.

HP MDS can also capture the waveform from an oscilloscope for use as a stimulus in a simulation. This allows you apply an actual waveform to a package that is still being developed to test the risetime, crosstalk, and other parameters.

## CONCLUSION

- HP Impulse complements HP MNS and is a high-frequency replacement for SPICE
- HP Impulse has 2 key contributions vs. SPICE:
  - Use S-parameter data directly
  - Accurate models for microstrip, etc.
- HP Impulse is fully integrated into HP MDS

HP has introduced a new transient circuit simulator that fills a gap between SPICE and high-frequency measurements and techniques. This simulator is called HP Impulse. It is the product of original research at the High-Frequency Design Software laboratory at the HP Santa Rosa Systems Division and is fully integrated into the HP Microwave Design System.

## THE VECTOR DISCRIMINATOR

Marilyn F. Williams

Billy P. Ficklin

SRI International

333 Ravenswood Avenue

Menlo Park, CA 94025-3493

### 1 Introduction

In a typical digital communications system, information can be transmitted efficiently via a propagating medium by modulating a sine-wave carrier with baseband signals representing the information, and then transmitting that band of signals around the carrier via the medium or over a propagation channel to a receiver. Carrier and data clock acquisition, recovery, synchronization, and tracking are necessary prerequisites to data reception. The signal, recovered from unwanted additions and distortions, is demodulated, and an estimate of the original information is presented. The methods of modulation and demodulation play vital roles in carrier and data clock recovery and information transmission.

The fundamental modulation methods are amplitude (AM), frequency (FM), and phase (PM), or combinations of the three. We have developed a vector discriminator technique for demodulating FM signals that allows accurate detection of small frequency changes without long integration times, dc drifts, or fast Fourier transform (FFT) methods. This technique has been used to detect instantaneous Doppler shifts of a few hertz to accuracies of better than 0.1 Hz on carrier frequencies of 50.0 MHz. The simplicity of the processing circuits in the digital realm and the advent of high-speed analog-to-digital converters makes the technique applicable to a large class of problems. It is expected that applications will be found in communications, sonar, and radar systems implementations.

Multistatic radars and sonars with almost any waveform modulation can use the technique to detect the presence, location, and velocity of targets. Multiple frequency shift keying (M-ary FSK) communications signals can be discriminated to multiple dc levels for integration, smoothing, and detection, and can be further demodulated to recover the originally transmitted symbols.

### 2 Modulation Techniques

At the most basic level, a modulated signal can be represented at any instant in time by a function,  $S(t)^{1,2}$  where:

$$S(t) = A(t) \cdot \sin[\omega_c t + \phi(t) - \phi(0)] , \quad \omega_c = 2\pi f_c \quad (1)$$

where  $A(t)$  is the instantaneous amplitude variation in volts,  $\phi(t)$  is the instantaneous phase variation in radians or degrees,  $\phi(0)$  is the initial phase difference, and the instantaneous frequency of the signal in radians/s is:

$$\omega(t) = \omega_c + d \phi(t)/dt \quad (2)$$

The rate of change of  $A(t)$  and  $\phi(t)$  are much less than the carrier frequency,  $\omega_c$  radians/s.

Individually,  $A(t)$ ,  $\omega(t)$ , and  $\phi(t)$  can be determined within fixed constants using the formulas derived in Appendix A relating the baseband components or the in-phase,  $I(t)$ , and quadrature,  $Q(t)$ , outputs of the vector demodulator shown in Figure A-1. The instantaneous amplitude,  $A(t)$ , can be calculated and recovered by the relationship

$$\text{constant} \cdot A^2(t) = B^2(t) = I^2(t) + Q^2(t) \quad (3)$$

and the instantaneous phase by

$$\Delta\omega t + \phi(t) - \phi(0) = \arctan [Q(t)/I(t)] \quad (4)$$

where  $\Delta\omega = \omega_c - \omega_o$ , the difference between the carrier frequency,  $\omega_c$ , and the local oscillator frequency,  $\omega_o$ , used in the vector demodulator.

The instantaneous frequency difference,  $\Delta\omega$ , is given by differentiating the instantaneous phase<sup>3</sup> given in Equation (4).

$$\frac{d[\Delta\omega t + \phi(t) - \phi(0)]}{dt} = \frac{d \Delta\omega}{dt} \cdot t + \Delta\omega + \frac{d \phi(t)}{dt} - \frac{d \phi(0)}{dt} \quad (5)$$

\* References are listed at the end of this paper.

$$\frac{d[\arctan Q(t)/I(t)]}{dt} = \frac{I(t) \cdot \frac{dQ(t)}{dt} - Q(t) \cdot \frac{dI(t)}{dt}}{I^2(t) + Q^2(t)} \quad (6)$$

Since  $d\phi(0)/dt = 0$  and if the rates of change in  $\omega$  and  $\phi(t)$  are small during the time period of interest,  $\Delta\omega$  is a good approximation for the instantaneous frequency difference between  $\omega_c$  and  $\omega_o$ .

Amplitude and phase demodulation are not discussed in detail here, since they are adequately covered in many textbooks. Digital amplitude modulation (on-off keying) can be recovered by the use of Equation (3) combined with some filtering to smooth out the fluctuation due to noise. The presence or absence of a signal is determined by a threshold set to accept signals larger than a certain magnitude as signal and reject those below as noise. The amplitude measured using Equation (3) can also be used as part of an automatic gain control (AGC) circuit. Amplitude modulation is the one most susceptible to noise caused by amplitude variations and must have the largest signal-to-noise ratio (SNR) for a specified error rate.

Angular modulation techniques (PM and FM) can use power-efficient fixed-output transmitters and limiters in the receivers to decrease amplitude variations due to noise, fading, or changes in path loss. In addition, the modulation scheme may be coherent or noncoherent depending on whether or not the local oscillator of the receiver is phase-locked to the received signal. Binary phase shift keying (BPSK), a coherent modulation scheme, can be recovered by the use of a phase detector and one or more phase-locked loops. The phase of the receiver local oscillator initially must synchronize and lock to the received signal carrier to start data transfer, and then must continue to track the received signal to recover the data clock and maintain information transfer. Differential binary phase shift keying (DBPSK) demodulation does not require synchronization to the carrier. DBPSK data can be demodulated with some small loss in quality over BPSK, but without the requirement of phase-locking the local oscillator to the incoming signal.

### 3 FM Demodulation

Based on Equation (5), there are several processes whereby FM modulated digital signals can be detected. They usually provide some means of creating an output voltage that is linearly related to the frequency of the received signal and are called frequency discriminators. A limiter is usually inserted before the discriminator circuit to reduce amplitude modulation. An FM detector can be constructed with a differentiating circuit followed by an envelope detector. Other techniques

with either a phase- or frequency-locked loop coherently detect the modulation. In addition to coherent discriminators, differential delay line and zero crossing detectors can be constructed that produce outputs that are linearly related to the frequency of deviation<sup>4</sup> of the carrier. Some techniques work best at radio or intermediate frequencies, while others work best at baseband.

The vector discriminator described here is a baseband technique. It is based on a calculation of the tangent of the angle between two consecutive vectors and could be viewed as a complex (vs. real) differential type detector. It does not require phase-locking and therefore is a noncoherent technique. In this respect, it is similar to the differential phase detector.

### 4 Vector Discriminator

Baseband in-phase,  $I(t)$ , and quadrature,  $Q(t)$ , components can be produced with a vector demodulator that are related to the original signals by relationships derived in Appendix A. If a signal is frequency modulated, the location in I-Q space of a complex signal at two consecutive instants in time can be represented by two vectors. Given two vectors, as shown in Figure 1, the tangent of the angle,  $\theta$ , between the two is equal to the vector cross product divided by the vector dot product as shown:

$$\tan \theta = \frac{\vec{V}_1 \times \vec{V}_2}{\vec{V}_1 \cdot \vec{V}_2} = \frac{I(t_1) \cdot Q(t_2) - I(t_2) \cdot Q(t_1)}{I(t_1) \cdot I(t_2) + Q(t_1) \cdot Q(t_2)} \quad (7)$$

where  $V_1 = I(t_1) + jQ(t_1)$ ,  $V_2 = I(t_2) + jQ(t_2)$ .

Another approach starting from Equation (6) yields nearly the same result. If two discrete measurements of  $I(t)$  and  $Q(t)$  are made at  $t_1$  and  $t_2$ , so that  $t_2 = t_1 + \Delta t$ , and  $\Delta t$  is very small, then

$$\frac{dI(t)}{dt} = \frac{I(t_2) - I(t_1)}{\Delta t} \quad (8)$$

$$\frac{dQ(t)}{dt} = \frac{Q(t_2) - Q(t_1)}{\Delta t} \quad (9)$$

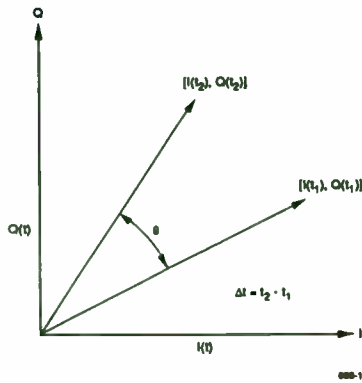


FIGURE 1 I&Q PLANE REPRESENTATION OF VECTOR DEMODULATOR OUTPUTS

Substituting Equations (8) and (9) into Equation (6) yields an approximation to the derivative of the instantaneous phase,

$$\frac{d[\arctan Q(t)/I(t)]}{dt} = \frac{I(t_1) \cdot Q(t_2) - I(t_2) \cdot Q(t_1)}{I^2(t_1) + Q^2(t_1)}, \quad (10)$$

which is almost the same as Equation (7). Equation (7) is the more accurate, while Equation (10) is the more convenient of the two to use.

If there is a constant, but small, frequency difference,  $\Delta\omega$ , between the carrier,  $\omega_c$ , and the local oscillator,  $\omega_o$ , the angle,  $\theta$ , between successive samples of  $I(t)$  and  $Q(t)$  will be a constant, which is related to the difference frequency:

$$\Delta\omega = \frac{1}{\Delta t} \arctan \theta \quad (11)$$

If  $\tan \theta < 0.1$ , then it is approximately equal to  $\theta$ , and if  $\Delta t$  is small, then the instantaneous frequency is approximately equal to  $\theta$  divided by  $\Delta t$ , or:

$$\Delta\omega = \frac{1}{\Delta t} \cdot \frac{I(t_1) \cdot Q(t_2) - I(t_2) \cdot Q(t_1)}{I^2(t_1) + Q^2(t_1)} \quad (12)$$

Equation (12) relates the instantaneous difference frequency,  $\Delta\omega$ , to two sequential complex baseband measurements. The critical conditions that must be met before Equation (12) is usable are (1)  $\tan \theta \approx \theta$  and (2)  $\Delta t$  is small. Subroutines that calculate the instantaneous difference frequency are given in Appendix B.

By letting  $A(t)$  in Equation (1) be a constant,  $A$ , a constant amplitude frequency modulated signal, can be created by modulating a carrier with a single-frequency signal,  $(\Delta\phi/\rho) \cos pt$ , so that the signal  $S_{fm}(t)$  results:

$$S_{fm}(t) = A \cdot \cos(\omega_c t + \Delta\phi \sin pt) \quad (13)$$

where  $\Delta\phi$  is the phase excursion,  $\Delta\phi/\rho$  is the modulation index, and  $\omega_c$  is the carrier. The instantaneous output of the vector discriminator is

$$\frac{\theta}{\Delta t} = \Delta\omega + \frac{\Delta\phi}{\rho} \cos pt \quad (14)$$

If the local oscillator frequency,  $\omega_o$ , of the vector demodulator is close to but not coherent with  $\omega_c$ , the output can be low-pass filtered to remove  $(\Delta\phi/\rho) \cos pt$ , and  $\Delta\omega$  fed back to control a voltage or digitally controlled local oscillator to function similarly to a Costas loop for achieving frequency-coherent detection. Frequency-modulated signals can be demodulated with this technique by high-pass filtering (the output of the instantaneous vector discriminator) to eliminate  $\Delta\omega$  and recover  $(\Delta\phi/\rho) \cos pt$ . The original modulation is recovered without additional differentiating or processing of the output.

The process shown in Figure 2 demodulates an FM IF signal by first limiting the signal to remove amplitude variations. In-phase and quadrature (I&Q) baseband signals are generated by mixing the clipped IF signal and quadrature outputs of the local oscillator. After low-pass filtering the I&Q signals to remove the high-frequency (sum) components and to prevent aliasing, the I&Q signals are digitized with analog-to-digital (A/D) converters. The digitized baseband waveforms can then be further filtered (low-pass, high-pass, band-pass, or matched filtered as the case may be) before calculation of  $\tan \theta$  by using consecutive samples of I&Q and Equation (12).

When the sampling frequency,  $f_s$ , is much higher than the highest baseband frequency passed by the first low-pass filter,  $\tan \theta$  will always be less than 0.1 and approximately equal to  $\theta$ .

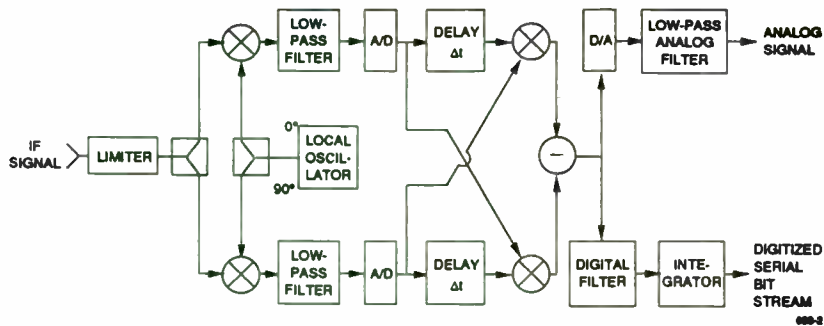


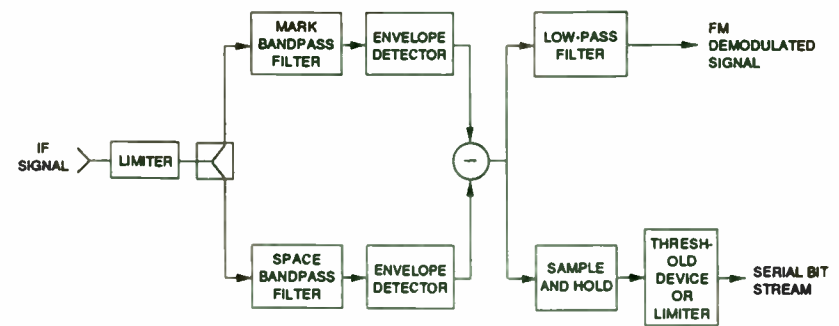
FIGURE 2 FUNCTIONAL BLOCK DIAGRAM OF FM DISCRIMINATOR

The sample period,  $\Delta t$ , is equal to  $1/f_s$ . The calculated digital instantaneous frequency difference between the local oscillator and the IF signal is then  $\theta/\Delta t$  or  $\theta \cdot f_s$ . The digitally demodulated signal may be clipped, smoothed, and integrated to recover a serial bit stream if the original FM modulation was binary frequency shift keying (BFSK). The digital signal,  $\theta \cdot f_s$ , may be converted to an analog signal with a digital-to-analog (D/A) converter, then filtered and amplified and used to drive speakers for audio communication or entertainment. The analog or digital signal may be used to control a voltage or digitally controlled, agile, frequency-hopping oscillator that can be switched from one frequency to another in less than 100 times the sampling period.

### 5 Frequency Discrimination By Other Techniques

The conventional method of detecting noncoherent FSK signals and analog frequency modulated signals is shown in Figure 3. The signal at some intermediate frequency is split and fed to two filters: one tuned to  $f_1$  (the upper or mark frequency with a bandwidth  $B_q$ ) and the other tuned to  $f_2$  (the lower or space frequency with an identical bandwidth,  $B_q$ ). The envelope-detected outputs are subtracted, then smoothed or integrated and, in the case of BFSK, fed to a threshold device or hard limiter. The output of the limiter is then a serial bit stream.

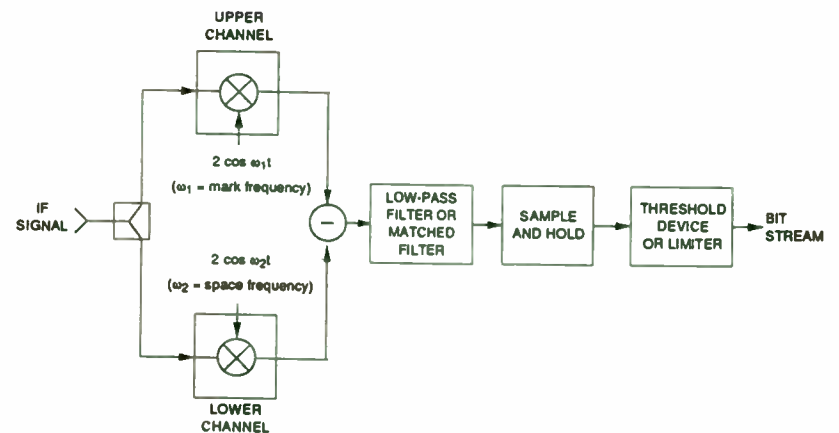
Coherent detection of FSK signals is usually accomplished by the process shown in Figure 4. The input signal is split into two parts that are then coherently detected by the two phase detectors (mixers)—one for the mark frequency and one for the space frequency. The mixer outputs are differenced and then low-pass or matched filtered before a limiter generates the detected serial bit stream. The oscillators that generate the mark and space frequencies must be coherent or synchronized with the originally transmitted signals. The local oscillator driving the I&Q mixers of



Source: Couch, Leon W., II, "Digital and Analog Communications Systems" University of Florida, Electrical Engineering, Gainesville, Florida, Macmillan Publishing Co., New York (1983).

688-3

FIGURE 3 NONCOHERENT FSK SIGNAL PROCESSOR OR DEMODULATOR



Source: Couch, Leon W., II, "Digital and Analog Communications Systems" University of Florida, Electrical Engineering, Gainesville, Florida, Macmillan Publishing Co., New York, (1983).

688-4

FIGURE 4 COHERENT FSK DEMODULATOR

a quadrature detector is controlled by a phase-locked loop (PLL). The phase of the incoming carrier signal is recovered by stripping off the modulation, low-pass filtering, and using the dc levels created by the slow phase fluctuation to control the PLL.

FSK demodulators constructed with the discriminator shown in Figure 2 achieve nearly the same bit error rate (BER) performance as the coherent FSK demodulator shown in Figure 4, but without the requirement of coherency. In Figure 2, the signal necessary to phase-lock the local oscillator to the carrier frequency is available in the output of the instantaneous discriminator to be used or not, as required. As long as the separation between the mark ( $f_1$ ) and space ( $f_2$ ) frequency is much greater than the difference between transmit and receive oscillators,  $\Delta f_{CO}$ , the information can be readily recovered:

$$\Delta f_{CO} \ll f_1 - f_2 .$$

Delay line discriminators<sup>5</sup> are constructed using a delay element, a phase shifter, a multiplier or mixer, and a low-pass filter as shown in Figure 5. The signal is split into two channels, one applied directly to an input port of a double balanced mixer, and the other delayed by a time element, then applied to the other input port of the double balanced mixer. The mixer operates as a phase detector and converts changes in frequency into changes in phase. The filter removes the high-frequency products of mixing and smooths the data. Delay line discriminators are used to construct high-frequency discriminators for instantaneous frequency measurement receivers.

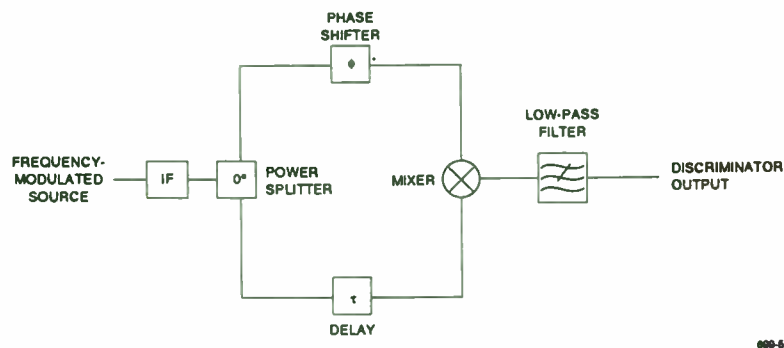


FIGURE 5 DELAY LINE DISCRIMINATOR

The delay line discriminator has a transfer function given by

$$\Delta V_{fm} = K_o \cdot 2\pi T_d \Delta f_{fm} \frac{\sin(\pi f_m T_d)}{(\pi f_m T_d)} , \tag{15}$$

where  $\Delta V_{fm}$  is the voltage out of the discriminator,  $K_o$  is a gain constant,  $T_d$  is the time delay, and  $\Delta f_{fm}$  is the short-term frequency deviation.

The output follows a  $\sin x/x$  response and is periodic in  $\omega$  equal to  $2\pi f_m$ , with the first null occurring at  $1/T_d$ . The short-term frequency deviation should remain less than  $1/(2\pi T_d)$ , where the response is linear and at a maximum. This requirement limits the practical bandwidth of the delay line discriminator.

Zero-crossing techniques have been successfully used to construct discriminators for laboratory frequency-measuring instruments. The time difference between zero crossings, the signal period, is accurately measured and its reciprocal used as an indication of the frequency.

The frequency of real signals can be found by using the vector discriminator algorithms on the output of a signal that has been processed by a discrete Hilbert transform (DHT), and the real signal delayed by the DHT delay in place of  $I(t)$ , as shown in Figure 6.

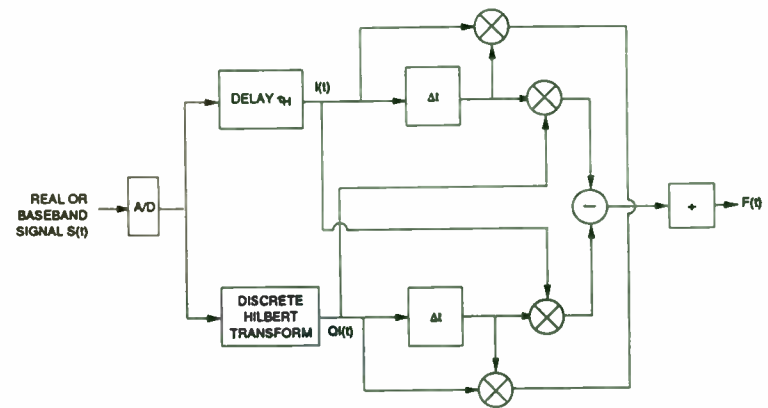


FIGURE 6 INSTANTANEOUS FREQUENCY FROM A REAL SIGNAL

## 6 Examples of Vector Discriminator Outputs

An example of detection of an FSK signal embedded in noise is given in Figures 7 through 10. Figures 7 and 8 show examples of digitized I&Q signals versus time that were low-pass filtered with a cut-off frequency equal to one-half of the frequency shift. The unsmoothed discriminator output is shown in Figure 9. Figure 10 shows the serial bit stream that results after bandpass filtering the I&Q signals around the mark and space frequency, before discriminating. Additional smoothing or low-pass filtering and clipping was performed on the discriminator output.

Figure 11 compares the simulated or calculated bit error rate (BER) performance of the vector discriminator to theoretical BER performance of other demodulation or discriminator techniques<sup>1,2</sup>. The BER performance versus signal-to-noise ( $E_b/N_0$ ) given Gaussian noise and BFSK signaling are shown to be much better than noncoherent FSK. The BER performance of the vector discriminator (labeled DFSK) is not as good as coherent FSK, DPSK, or BPSK, but the requirement to obtain carrier synchronization is not present. Its resistance to signal fading or pulsed noise should be much greater than the coherent schemes. The ability to resist capture by a continuous-wave (CW) interference source is close to the theoretical value of a limiter discriminator, as is also shown in Figure 11. If the CW interference source is greater than the signal, then a high error rate occurs. If the signal is greater than the CW interference, then a low error rate occurs. The performance of noncoherent FSK discriminators degrades when CW interference is 3 dB less than the signal, and is somewhat dependent upon the modulation index.

An example of the discriminator output, when the original signal was a linear chirp, is shown in Figures 12 and 13. Figure 12 is the magnitude versus time of the received and processed signal. Figure 13 is the frequency versus time calculated using the same data.

Figure 14 shows relative amplitude versus time for a received meteor burst signal that is BPSK modulated. Figure 15 shows the frequency shift caused by motion of the decaying ionized meteor cloud that was calculated by first stripping off the modulation and then determining the difference frequency by the instantaneous discriminator algorithm. Figure 16 is the histogram of Doppler shift observed for the same meteor. Figures 15 and 16 show the time variation of frequency and the spectral distribution and resolution better than can be determined by using discrete FFTs for a fixed amount of data.

The discriminator is useful in instantaneous frequency measurement (IFM) receivers to measure the frequency of hopping signals. A sequence of sine waves stepping from 4.0 to 6.0

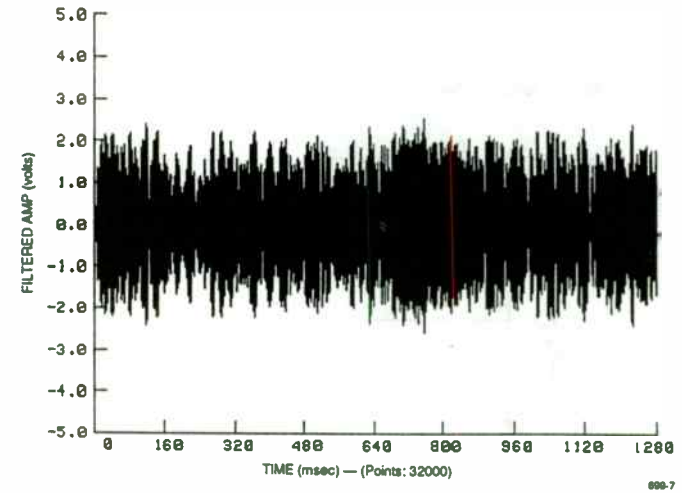


FIGURE 7 DIGITIZED IN-PHASE SIGNAL WITH NOISE VERSUS TIME

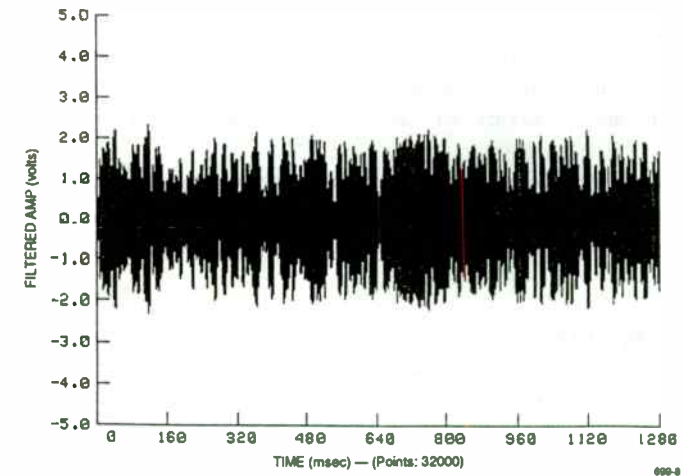


FIGURE 8 DIGITIZED QUADRATURE SIGNAL WITH NOISE VERSUS TIME



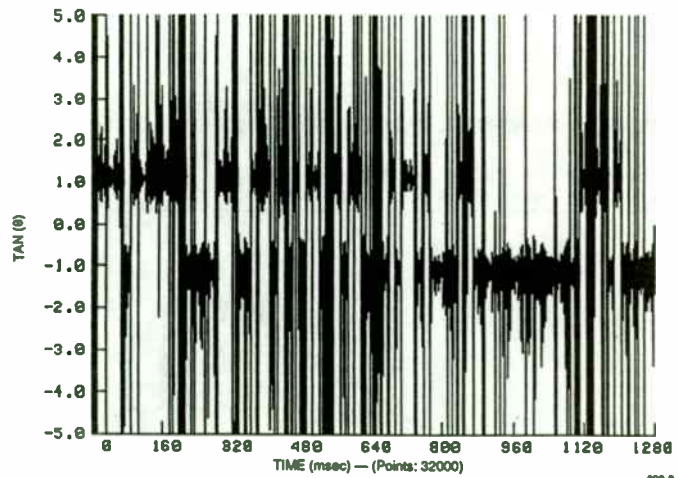


FIGURE 9 UNSMOOTHED DISCRIMINATOR OUTPUT FROM SIGNALS SHOWN IN FIGURES 8 AND 9

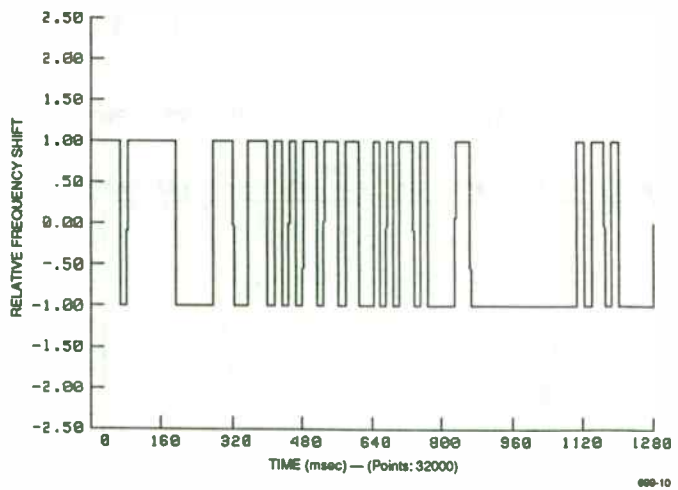
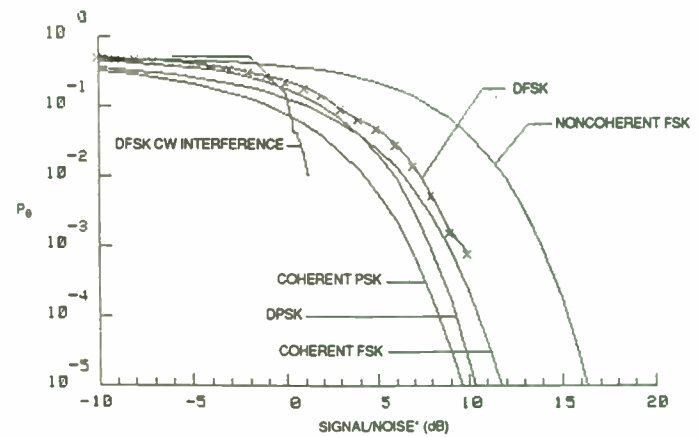


FIGURE 10 RESULTANT SERIAL BIT STREAM AFTER BANDPASS FILTERING



\*GAUSSIAN NOISE UNLESS OTHERWISE SPECIFIED

FIGURE 11 BER PERFORMANCE COMPARISON OF CFSK, NCFsk, DfSK, DPsK, AND BPsK VERSUS GAUSSIAN NOISE

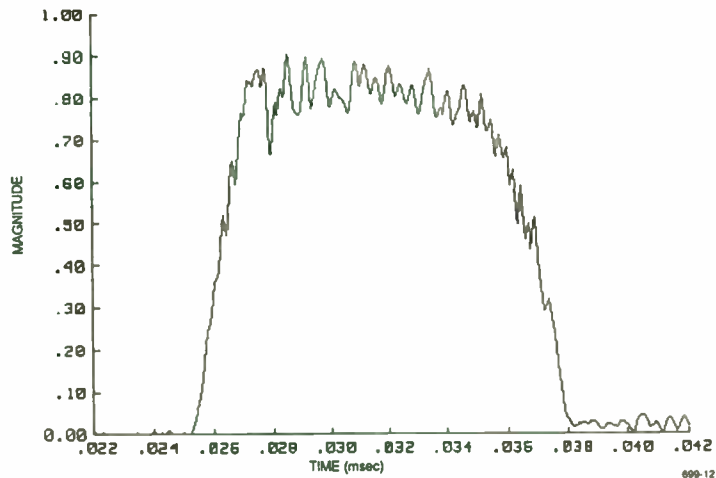


FIGURE 12 MAGNITUDE VERSUS TIME OF A LINEAR CHIRP SIGNAL

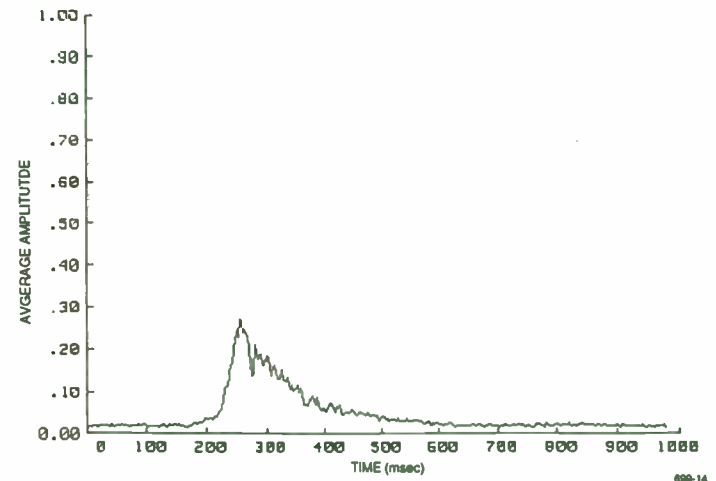


FIGURE 14 RELATIVE AMPLITUDE VERSUS TIME OF RECEIVED METEOR BURST SIGNAL

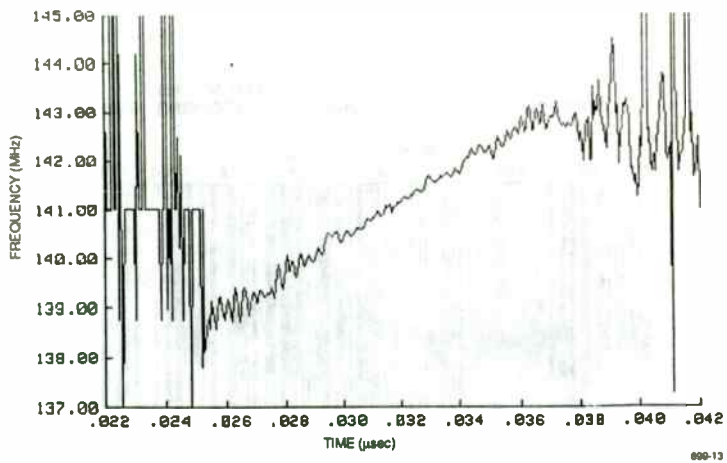


FIGURE 13 FREQUENCY VERSUS TIME OF THE CHIRP SIGNAL SHOWN IN FIGURE 12

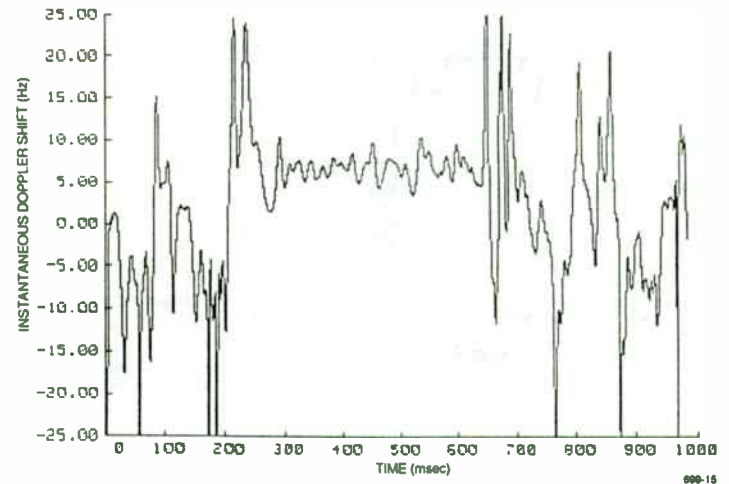


FIGURE 15 INSTANTANEOUS FREQUENCY DIFFERENCE BETWEEN RECEIVED SIGNAL AND STABLE LOCAL OSCILLATOR

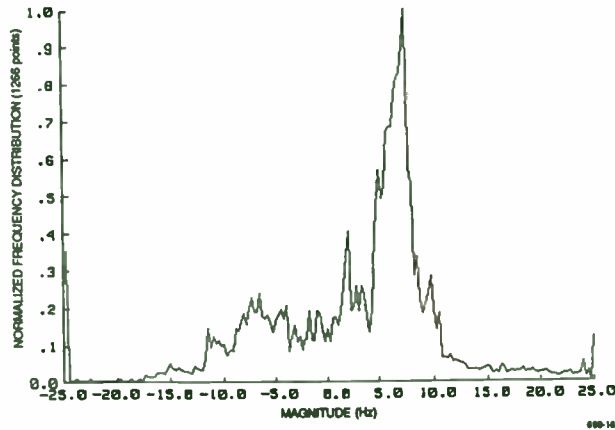


FIGURE 16 HISTOGRAM OF DOPPLER SHIFT CAUSED BY MOTION OF IONIZED PARTICLES

MHz was generated and demodulated with a vector demodulator and a 5.0 MHz local oscillator to create signals that varied from -1.0 to 1.0 MHz. As shown in Figure 17, the vector discriminator output vs. time is the instantaneous frequency of the original signals offset by the local oscillator frequency.

### 7 Applications of the Discriminator

If a serial bit stream of information or symbols is transmitted by frequency shift keying between two frequencies (BFSK),  $\omega_c + \omega_s$  and  $\omega_c - \omega_s$ , and  $\omega_0 = \omega_c$ , the instantaneous discriminator output is equal to  $+\omega_s + \Delta\omega$ , or  $-\omega_s + \Delta\omega$ . The discriminator output may be clipped, integrated, filtered, correlated, or further processed to reduce noise and recover the transmitted information in the serial bit stream if  $\omega_s > \Delta\omega$ . The serial bit stream can be used for synchronizing and recovering data clocks, data, and protocol information, decoding encoded signals, and decrypting encrypted information.

Multiple frequency shift keying (M-ary FSK) signals can be discriminated to multiple dc levels for integration, smoothing, detection, and further demodulation to recover the originally transmitted symbols. This discriminator is useful in communication, location, and time distribution systems that use spread-spectrum techniques (frequency-hopping, MSK, direct sequence, chirp) either singly or in combinations.

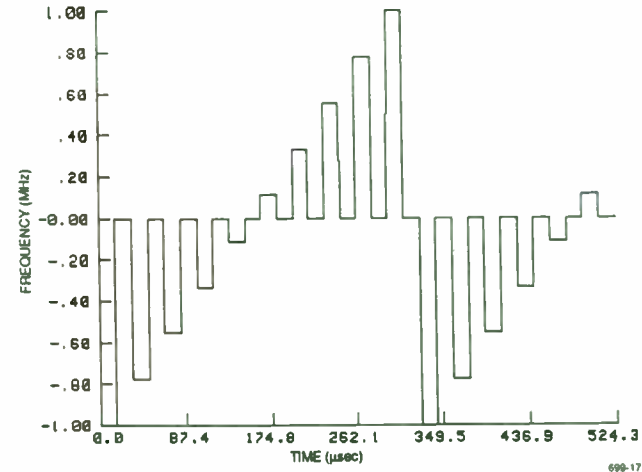


FIGURE 17 INSTANTANEOUS FREQUENCY VERSUS TIME CALCULATED FROM A SEQUENCE OF DATA GENERATED BY STEPPING SINE WAVES IN FREQUENCY

In monostatic radar<sup>6</sup> and sonar applications, the Doppler shift of the returned signal and transmitted signal frequency is related to the target radial velocity and speed-of-signal propagation by:

$$\frac{\omega_d}{\omega_t} = \frac{2v}{c-v} = \frac{2v}{c}, \quad (16)$$

where

$\omega_d$  = Doppler frequencies, rad/s

$\omega_t$  = transmitted frequency, rad/s

$v$  = target radial velocity, m/s

$c$  = speed of propagation, m/s.

Continuous wave, hopped, pulsed, chirped, or coded signals may be used with this discriminator to detect accurately and quickly the velocity of targets. The location of the target can be determined in all of the above cases except CW unless several CW signals are used. The ground clutter can be measured accurately and subtracted from signals to improve moving target detection in the presence of clutter.

Bistatic or multistatic radars with almost any waveform modulation can use the discriminator to detect the presence, location, and velocity of targets. These radars can operate out of line of sight (LOS) of each other because it is not necessary that the receivers be phase-locked to the transmitter. The oscillators must be stable so that their frequencies do not change rapidly during data sampling periods. The total path length can be determined by synchronizing the transmitters and receivers to a common timing signal (such as WWVB, WWVH, WWV, or GPS) that propagates beyond LOS. The Doppler frequency shift is proportional to the rate of change of the total path length and the transmitted frequency. It is possible to determine the trajectory of the target from Doppler measurements only. The vector discriminator can be used to detect and track (using the simple algorithm with unprecedented accuracy) moving ionized clouds in the ionosphere or the neutral wind if meteor trail or auroral-scattered signals are observed.

## Appendix A VECTOR DEMODULATOR

The vector demodulator shown in Figure A-1 is a circuit commonly used to translate modulated signals to baseband from intermediate or radio frequencies. It is specially useful when digital representation of radio frequency signals are desired. If properly done no information in the modulated signal is lost.

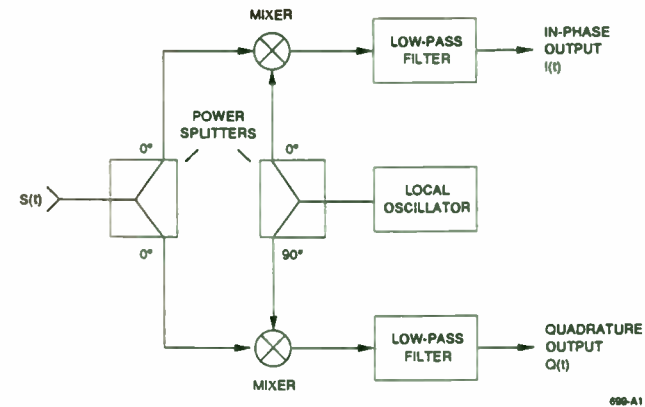


FIGURE A-1 VECTOR DEMODULATOR FUNCTIONAL BLOCK DIAGRAM

The signal  $S(t)$  from Equation (1) is split into two equal signals by a  $0^\circ$  phase-shift power divider; one half is fed to the in-phase mixer and the other half to the quadrature mixer. The in-phase mixer is driven by a signal,  $C \cdot \sin(\omega_0 t + \phi(0))$ , in phase with the receiver stable local oscillator (LO) that is operating with a constant amplitude,  $C$ , at a frequency of  $\omega_0$  and at an initial phase shift,  $\phi(0)$ , relative to the phase of  $S(t)$ . The quadrature mixer is driven by a signal,  $C \cdot \cos[\omega_0 t + \phi(0)]$ , from the same LO with a  $90^\circ$  phase difference from the signal driving the in-phase mixer. The outputs of the in-phase mixer are the sum and difference components generated by the products of sinusoids:

$$\begin{aligned} \frac{1}{2} \cdot S(t) \cdot C \cdot \sin[\omega_0 t + \phi(0)] &= \frac{1}{2} \cdot C \cdot A(t) \cdot \sin[\omega_c t + \phi(t)] \cdot \sin[\omega_0 t + \phi(0)] \\ &= \frac{C}{2\sqrt{2}} \cdot A(t) \cdot \{ \cos[(\omega_c - \omega_0)t + \phi(t) - \phi(0)] - \cos[(\omega_c + \omega_0)t + \phi(t) + \phi(0)] \} \quad (A-1) \end{aligned}$$

The output of the quadrature mixer is:

$$\begin{aligned} \frac{1}{2} \cdot S(t) \cdot C \cdot \cos[\omega_0 t + \phi(0)] &= \frac{1}{2} \cdot C \cdot A(t) \cdot \sin[\omega_c t + \phi(t)] \cdot \cos[\omega_0 t + \phi(0)] \\ &= \frac{C}{2\sqrt{2}} \cdot A(t) \cdot \{ \sin[(\omega_c - \omega_0)t + \phi(t) - \phi(0)] + \sin[(\omega_c + \omega_0)t + \phi(t) + \phi(0)] \} \end{aligned} \quad (A-2)$$

If  $\omega_0$  is near  $\omega_c$  but not coherent with it, then a new low-frequency variable near zero is in order,  $\Delta\omega = \omega_c - \omega_0$ . Both mixer outputs are low-pass filtered to retain the functions of  $\Delta\omega t$  and to eliminate the high-frequency components that contain functions of  $(\omega_c + \omega_0)t$  in order that baseband signals remain, but with some loss or attenuation, D. The amplitude constants,  $C \cdot D/(2\sqrt{2})$ , can be consolidated to create a new amplitude function, B(t), that differs from A(t) by a constant. The in-phase, I(t), and quadrature, Q(t), functions of time shown in Figure A-2 are then:

$$I(t) = B(t) \cdot \cos[\Delta\omega t + \phi(t) - \phi(0)] \quad (A-3)$$

$$Q(t) = B(t) \cdot \sin[\Delta\omega t + \phi(t) - \phi(0)] \quad (A-4)$$

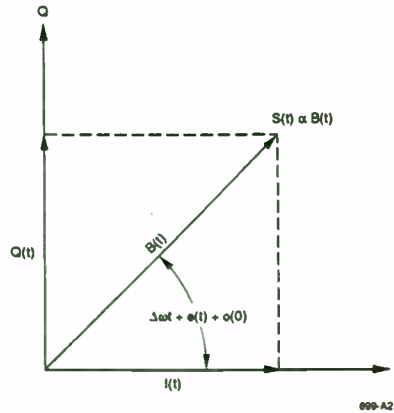


FIGURE A-2 B(t) REPRESENTATION IN I&Q PLANE

## Appendix B

### VECTOR DISCRIMINATOR SUBROUTINES

Figure B-1 is a BASIC language subroutine that calculates the vector discriminator output given in-phase and quadrature inputs according to Equation (12). Figure B-2 calculates the discriminator output with the same inputs but uses a half-angle formula. Figure B-3 is a FORTRAN language subroutine that calculates the vector discriminator output using Equation (12). Figure B-4 is the subroutine that strips off the BPSK modulation and calculates twice the frequency shift of the carrier which, when divided by two, gives the instantaneous Doppler shift or frequency variation of the signal source.

```

10 !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
20 ! SUB VECTOR_DISCRIM (VECTOR DISCRIMINATOR) !
30 ! SAMPLE_RATE -- SAMPLING RATE IN HZ !
40 ! I_IN(+), Q_IN(+)-- IN PHASE & QUADRATURE INPUT ARRAYS !
50 ! NPTS -- NO OF SAMPLE POINTS IN ARRAYS !
60 ! DISCRIM_OUT(+)-- OUTPUT OF THE VECTOR DISCRIMINATOR !
70 ! IN RADIANS/SEC !
80 !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
90 !
100 SUB Vector_discrim(Sample_rate,I_in(+),Q_in(+),Npts,Discrim_out(+))
110 OPTION BASE 1
120 !
130 Vector_discrim:
140 FOR I=1 TO Npts-1
150 I1=I_in(I)
160 I2=I_in(I+1)
170 Q1=Q_in(I)
180 Q2=Q_in(I+1)
190 V_cross_v=I1*Q2-I2*Q1
200 V_dot_v=I1*I2+Q1*Q2
210 IF (V_dot_v=0) THEN V_dot_v=MINREAL
220 Discrim_out(I)=(V_cross_v/V_dot_v)*Sample_rate
230 NEXT I
240 Discrim_out(Npts)=Discrim_out(Npts-1)
250 SUBEND!
    
```

000-B1

FIGURE B-1 BASIC LANGUAGE SUBROUTINE CALCULATING THE VECTOR DISCRIMINATOR USING EQUATION (12)

```

10 | !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!! |
20 | SUB HALF_ANG_DISCRM (HALF ANGLE DISCRIMINATOR) |
30 | SAMPLE_RATE -- SAMPLING RATE IN HZ |
40 | I_IN(*), Q_IN(*) -- IN PHASE & QUADRATURE INPUT ARRAYS |
50 | NPTS -- NO OF SAMPLE POINTS IN ARRAYS |
60 | DISCRIM_OUT(*) -- OUTPUT OF THE VECTOR DISCRIMINATOR |
70 | IN RADIANS/SEC |
80 | !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!! |
90 | |
100 SUB Half_ang_discrm(Sample_rate,I_in(*),Q_in(*),Npts,Discrim_out(*))
110 OPTION BASE I
120 |
130 Half_angles:
140 FOR I=1 TO Npts-1
150 I1=I_in(I)
160 I2=I_in(I+1)
170 Q1=Q_in(I)
180 Q2=Q_in(I+1)
190 V_cross_v=I1*Q2-I2*Q1
200 V_dot_v=I1*I2+Q1*Q2
210 Denom=V_dot_v+I1*I1+Q1*Q1
220 IF (Denom=0) THEN Denom=MINREAL
230 Discrim_out(I)=(2.*V_cross_v/Denom)*Sample_rate
240 NEXT I
250 Discrim_out(Npts)=Discrim_out(Npts-1)
260 SUBEND!

```

000-82

FIGURE B-2 BASIC LANGUAGE SUBROUTINE CALCULATING THE VECTOR DISCRIMINATOR USING A HALF-ANGLE FORMULA

```

SUBROUTINE VECTOR(SRATE,DATAI,DATAQ,NPTS,VCTDIS)
CALCULATE VECTOR DISCRIMINATE
C SRATE -- SAMPLING RATE IN HZ
C DATAI,DATAQ -- IN PHASE AND QUADRATURE INPUT ARRAYS
C NPTS -- NO OF SAMPLE POINTS IN ARRAYS
C VCTDIS -- OUTPUT OF VECTOR DISCRIMINATOR (RAD/SEC)
DIMENSION DATAI(1),DATAQ(1),VCTDIS(1)
DATA REALMN/1.0E-38/
DO 100 I=1,NPTS-1
DI1=DATAI(I)
DI2=DATAI(I+1)
DQ1=DATAQ(I)
DQ2=DATAQ(I+1)
VCV=I1*Q2-I2*Q1
VDV=I1*I2+Q1*Q2
IF (VDV) 5,5,10
5 VDV=REALMN
10 CONTINUE
VCTDIS(I)=(VCV/VDV)*SRATE
100 CONTINUE
VCTDIS(NPTS)=VCTDIS(NPTS-1)
RETURN
END

```

000-83

FIGURE B-3 FORTRAN LANGUAGE SUBROUTINE CALCULATING THE VECTOR DISCRIMINATOR USING EQUATION (12)

```

10  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
20  !      SUB BPSK
30  !  SAMPLE_RATE      -- SAMPLING RATE IN HZ
40  !  I_IN(+), Q_IN(+)-- IN PHASE & QUADRATURE INPUT ARRAYS
50  !  NPTS            -- NO OF SAMPLE POINTS IN ARRAYS
60  !  DISCRIM_OUT(+)-- OUTPUT OF THE VECTOR DISCRIMINATOR IN Hz
70  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
80  !
90  SUB Bpsk(Sample_rate,I_in(+),Q_in(+),Npts,Discrim_out(+))
100 OPTION BASE 1
110 Bpsk:
120 ALLOCATE I_times_q(Npts)
130 Clip_iq(I_in(+),Q_in(+),Npts,I_times_q(+)) !CLIP I & Q AND RETURN I+Q
140 Low_pass_filter(I_times_q(+),Npts) !LOW PASS FILTER I+Q
150 !
160 ALLOCATE I2mq2(Npts)
170 Clip_i2mq2(I_in(+),Q_in(+),Npts,I2mq2(+)) !CLIP I+I - Q+Q AND RETURN
180 Low_pass_filter(I2mq2(+),Npts) !LOW PASS FILTER I+I-Q+Q
190 !
200 FOR I=1 TO Npts-1
210   I1=I_times_q(I)
220   I2=I_times_q(I+1)
230   Q1=I2mq2(I)
240   Q2=I2mq2(I+1)
250   V_cross_u=I1+Q2-I2+Q1
260   V_dot_u=I1+I2+Q1+Q2
270   Denom=V_dot_u+I1+I1+Q1+Q1
280   IF (Denom=0) THEN Denom=MIHREAL
290   Discrim_out(I)=V_cross_u/Denom
300 NEXT I
310 Discrim_out(Npts)=Discrim_out(Npts-1)
320 !
330 DEALLOCATE I_times_q(+),I2mq2(+))
340 Frctan_r(Discrim_out(+)) !CONVERT TO ANGLE
350 Fac=(2.*PI)*2.*Sample_rate
360 MAT Discrim_out=Discrim_out/(Fac) !Hz
370 SUBEND!
380 !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
390 SUB Clip_iq(I_in(+),Q_in(+),Npts,I_times_q(+))
400 OPTION BASE 1
410 C: p_iq: !
420 FOR I=1 TO Npts
430   Sign_i=SGN(I_in(I))
440   IF Sign_i=0 THEN Sign_i=1
450   Sign_q=SGN(Q_in(I))
460   IF Sign_q=0 THEN Sign_q=1
470   I_times_q(I)=Sign_i*STgn_q
480 --EXT I
490 SUBEND!
500 !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
510 SUB Clip_i2mq2(I_in(+),Q_in(+),Npts,I2mq2(+))
520 OPTION BASE 1
530 C: p_i2mq2: !
540 FOR I=1 TO Npts
550   I1=I_in(I)
560   Q1=Q_in(I)
570   Sign_i2mq2=SGN(I1+I1-Q1+Q1)
580   IF Sign_i2mq2=0 THEN Sign_i2mq2=1
590   I2mq2(I)=Sign_i2mq2
600 --EXT I
610 SUBEND!

```

## REFERENCES

1. Shanmugam, K. Sam, "Digital and Analog Communication Systems," University of Kansas, John Wiley & Sons, New York (1979).
2. Couch, Leon W., II, "Digital and Analog Communication Systems," University of Florida, Department of Electrical Engineering, Gainesville, Florida, Macmillan Publishing Co., New York (1983).
3. Papoulis, Athanasios, "Probability, Random Variables, and Stochastic Processes," Second Edition, McGraw-Hill Book Co., New York (1984).
4. Ziemer, Rodger E. and Roger L. Peterson, "Digital Communications and Spread Spectrum Systems," University of Colorado, Department of Electrical Engineering, Colorado Springs, Colorado, and Motorola, Inc., Government Electronics Group, Scottsdale, Arizona, Macmillan Publishing Company, New York (1985).
5. Hewlett-Packard, "Phase Noise Product Note 11729C-2, Characterization of Microwave Oscillators, Frequency Discriminator Method" (1985).
6. Skolnik, Merrill I., "Introduction to Radar Systems," Second Edition, McGraw-Hill Book Co., New York, (1980).

FIGURE B-4 BASIC LANGUAGE SUBROUTINE FOR BPSK TO DETERMINE INSTANTANEOUS DOPPLER SHIFT





### **A THIRD GENERATION SPREAD SPECTRUM S-BAND TRANSPONDER**

Massimo C. Comparini, F. Marchetti  
Alenia Spazio S.p.A.  
Via Saccomuro 24, 00131, Rome, Italy

Marty Siemon  
Motorola Inc., Strategic Electronics Division  
2501 S. Price Road, Chandler, AZ, USA

#### **ABSTRACT**

A new generation TT&C S-band transponder has been developed to support the spread spectrum modulation schemes required for operation with the future European Data Relay Satellite System (DRSS) and the existing NASA Tracking and Data Relay Satellite System (TDRSS).

The data relay systems provide the communications services of tracking, telemetry and command for spacecraft in low earth orbits (LEO users). Both the European and U.S. systems consist of two or more geostationary satellites, a single ground facility for the management of services (Darmstadt and White Sands respectively) and a network of payload operations centers to complete the user communications architecture.

In addition to spread spectrum operation, the transponder has the capability to communicate directly to a ground station in a residual carrier mode compatible with ESA and NASA ground networks.

The transponder was codeveloped by Alenia Spazio of Rome, Italy and Motorola's Strategic Electronics Division of Chandler, Arizona. Motorola has developed two previous generations of space qualified spread spectrum transponders for NASA and Alenia Spazio is the leading company in Europe for TT&C systems. The goal of the codevelopment was to combine the talents of the two companies to develop a transponder, using state of the art technologies, that would be compatible with both the ESA and the NASA data relay system satellites.

Starting with a baseline digital signal processing architecture that Motorola had developed, experts from the two companies formed a team to produce a design that would incorporate both system's requirements. An advanced development model (ADM) of the new design has been built and successfully tested. Motorola contributed the digital hardware, synthesizer components and software to the ADM while Alenia developed the RF and analog circuits using Microwave Integrated Circuit and Hybrid technology. A

joint team performed the design of a CMOS signal processing ASIC with over 51,000 gates of complexity.

#### **THE DATA RELAY SATELLITE SYSTEM CONCEPT**

The DRS/TDRS system configurations each provide two operational satellites in geostationary orbit, visible almost continuously to satellites in low earth orbits of up to 1000 Km altitude. Figure 1 shows the Data Relay System concept.

The orbital positions of the two satellites have been optimized to maximize the percentage of an orbit during which a user spacecraft can relay data. For the U.S. TDRS the locations of the satellites are at 41° West Longitude and 171° West Longitude, the European DRS foresees satellites at 44°W and 59°E. Compatibility between the NASA and ESA systems will insure 100 % orbital coverage, closing the zone of exclusion that exists with the current NASA system for some users.

Three basic communication functions of tracking, telemetry and command are provided:

\*Multiple Access

\*S-Band Single Access

\*Ku-band Single Access (NASA TDRS)  
Ka-band Single Access (ESA DRS)

The multiple access service is intended for low telemetry data rate users (up to 150 kb/sec) who require real time dedicated return links. The S-band single access provides the service at higher data rates (up to 6 Mbit/s) for higher EIRP users. The Ku-band TDRSS service is similar to single access S-band service but allows data rates to 300 Mbit/sec.

In the European DRS, to avoid interference with other systems a Ka-band frequency has been chosen. A maximum data rate of 150 Mbit/sec is foreseen for the wideband channel. The alternative frequency band for the high-data rate users is at optical wavelengths. A laser-optical communications experiment will fly on the European system pre-operative satellite, ARTEMIS, to verify the characteristics of optical interorbital communications. It is planned to continue using optical IOLs on the operational DRS satellite.

In order to define the TDRSS/DRSS signal structure a number of constraints were analyzed and considered. The regulation of spectral power flux density on the Earth's surface is a significant limitation and has a direct effect on the signal structure and modulation formats. Suppressed carrier modulation was selected for ease of

implementing direct sequence spreading with a pseudo-noise (PN) code sequence. The need for simultaneous command and range capability is met by using two different spreading codes modulated in quadrature on the forward link. An unbalanced QPSK scheme is used to minimize command signal loss while maintaining sufficient power for the ranging channel.

#### ***TRANSPONDER FUNCTIONAL REQUIREMENTS***

The transponder provides the user spacecraft portion of an S-band telecommunications link between the earth and an earth-orbiting spacecraft in one of two modes. In the spread spectrum mode, direct sequence spread forward and return link signals are relayed to a ground station through a geosynchronous relay satellite. The relay satellite systems allow the user to have global communications coverage from a single ground station. In the residual carrier mode, the signals are transmitted directly to a ground station which requires multiple ground stations for full orbit coverage.

A functional block diagram of the transponder is shown in figure 2. The transponder consists of both transmit and receive functions and communications can be receive only, transmit only or full duplex. The forward link (receive) frequency is modulated with command and ranging data and is selectable from 2025 to 2120 MHz using 180 channels. The corresponding return link transmitter is modulated with spacecraft telemetry data, science data and turnaround ranging information. The transmit frequency can be exactly coherent with the receive frequency in the ratio of 240/221 for fine range Doppler tracking and is between 2200 and 2300 MHz.

#### ***Spread Spectrum Functional Requirements***

The codevelopment design works with either the Single Access or Multiple Access S-band services. The forward link inphase component power is 10 dB greater than the quadrature component and is modulated with the command data modulo two added with a short (1023 chip) PN spreading sequence at approximately 3 M chips per second. Command data rates can be up to 300,000 bits per second. The quadrature component of the signal is modulated with a longer PN code sequence (256 times the short code length) for turnaround ranging. The normal dynamic range for the spread spectrum receiver mode is approximately 40 dB from -140 dBm to -100 dBm.

In addition to reducing the spectral power flux density, the short code sequences provide for rapid acquisition time and good cross correlation properties permitting up to 85 users to operate on a single frequency through code division multiplexing. The first function in the signal acquisition process is for the receiver to acquire and despread the short PN code. Once the PN code tracking loop is closed non-coherently, the receiver must acquire and track the suppressed carrier and demodulate the non-return-to-zero phase shift

key (NRZ-PSK) command data. Hard or soft (3-bit) command data decisions and reconstructed bit timing are provided to the spacecraft command system.

For the return link, two independent telemetry data channels are provided via the in-phase and quadrature components. Data are modulo-2 added to PN spreading codes and modulated onto the two carrier components. The carrier frequency can be coherent with the forward link signal or noncoherent. The code sequences can be a long ranging code synchronized to the forward link code for coherent operation or short sequences for one-way return link operation. A staggered QPSK scheme is used to have the best performance in terms of out-of-band power reduction. Convolutional encoding is required and can be provided by the transponder.

#### ***Residual Carrier Mode Functional Requirements***

In the residual carrier mode, the receiver must detect, acquire, track and demodulate a residual carrier forward link signal. The normal dynamic range for this signal type is approximately -100 dBm to -50 dBm. The signal format is linear phase modulation of the carrier by a 16 KHz or 8 KHz subcarrier which is BPSK modulated with command data. CW ranging tones can be simultaneously phase modulated on the forward link at a specified modulation index. These signals are demodulated to baseband and provided to the transmitter section for phase modulation of the return link carrier.

The return link signal can be coherent with the forward link or noncoherent and is phase modulated with a linear telemetry input signal. If desired, a 1.024 MHz subcarrier generated internally in the transponder can be BPSK modulated with an external digital telemetry data stream and linearly phase modulated on the return link carrier at a specified modulation index.

#### ***TRANSPONDER ARCHITECTURE***

A detailed block diagram of the transponder is shown in figure 3. The receiver uses two RF downconversions to a second IF frequency of about 9 MHz. The final downconversion to baseband signals for demodulation and carrier tracking is performed in the digital domain after the second IF signal is A to D converted at approximately a 32 MHz sampling rate.

The receiver first LO signal is tunable through the use of a Gallium Arsenide Numerically Controlled Oscillator and D to A converter. These GaAs components are also used in the transmitter chain for the frequency control and modulation functions. The first LO is varied as required to provide a constant first IF frequency regardless of channel assignment. This allows narrowband filters to be used in the IF chain for good RFI rejection.

An analog wideband AGC function is implemented to compensate the IF gain for the wide dynamic range seen by the receiver. This AGC is active above approximately -100 dBm (the residual carrier mode range).

A CMOS signal processing ASIC device combined with a high performance reduced instruction set (RISC) microprocessor is used to implement all of the receiver carrier tracking and demodulation functions. A signal processing ASIC in the transmitter is used to control the GaAs NCO's frequency and phase and also implements premodulation filtering of the data in the digital domain for shaping of the transmitted output spectrum.

The modulated IF signal from the GaAs NCO/Modulator is upconverted and applied to the input of the 5 watt power amplifier. The power amplifier is designed for a flat amplitude vs frequency characteristic over the entire 100 MHz tuning range.

### *Receiver Processing*

The receiver RF/analog section performs the following functions:

- Low Noise amplification and downconversion of the received S-band signal to an IF frequency capable of being sampled by an ADC and processed by a CMOS ASIC chip.
- Wideband AGC function for controlling the receiver gain in a strong signal environment and setting the operating point for the A/D converter operation.
- Removal of the carrier Doppler (phase locked loop tracking of the forward link carrier) via a Gallium Arsenide NCO under control of the microprocessor (Long Loop Option).
- RF filtering of the signal to reject spurious interfering signals.

The architecture chosen for the Receiver RF section is based upon a double conversion super-heterodyne configuration. This approach was judged to be the best compromise between hardware complexity and the ease of implementing the filtering functions. This configuration also distributes the overall gain among different frequencies along the RF/IF chain, to prevent undesirable oscillations that may occur when a large amount of gain is concentrated at one frequency.

A frequency synthesizer provides the RF receiving section with the needed Local Oscillator signals for conversion. The RF receiver analog section makes extensive use of MIC technology on alumina substrates.

An S-Band Low Noise Amplifier (LNA) is the first stage of the receiver RF analog chain. It consists of two MESFET stages, preceded by an isolator. The amplifier achieves a .9 dB max. noise figure (1.6 dB including isolator) over the full 100 MHz bandwidth.

The main design task for the LNA was to minimize the thermal noise from the signal source impedance, this was accomplished by minimizing the transfer of noise power in the two-port network by mismatching the input network. Doing so, however, the probability of potential instability was increased because of the high input VSWR. We arrived at the final solution by using a suspended line in the input matching network, which provides the LNA with optimum noise figure and unconditional stability. To create a suspended line an etching of alumina substrate ground plane in an area that lies underneath the line was done.

A fairly steep shape-factor SAW filter is employed to provide minimum attenuation of the desired signal components while maintaining good rejection of interfering signals that may be present at the receiver input along with the desired signal.

A noncoherent 2nd order AGC loop keeps the output signal plus noise power constant. A wide band detector is used to detect the amplitude variations of the incoming signal. The resulting signal is integrated, DC amplified and routed to both the 1st IF amplifiers and a 2nd IF PIN diode attenuator. For typical spread spectrum operation, the signal to noise ratio in the AGC bandwidth is negative and the AGC is essentially based on the noise power. The AGC time constant is selected to prevent a pulsed RFI source from degrading the receiver significantly.

The synthesizer frequency reference is implemented using a TCXO circuit. The TCXO circuit is basically a voltage controlled crystal oscillator, whose temperature behavior depends on an analog compensation network. This network consists of four NTC thermistors arranged in such a way to compensate for the crystal's temperature variations.

The TCXO is based upon a grounded-base oscillator configuration, which allows the stray capacitance to be tuned out, making it suitable in high frequency applications. A careful design and a suitable choice of the components have led to very good phase noise performance.

Another circuit which plays a key role in the frequency synthesizer is a times six frequency multiplier. The circuit is basically a quasi class-B biased amplifier, a slight bias is allowed to lower the input power level. The output waveform is rich in even harmonics; an idler circuit placed in the output mesh, tuned at the 4th harmonic increases the overall efficiency of the circuit.

The IF signal is quantized and samples provided to the CMOS ASIC for the signal processing. The ASIC performs the high speed receiver signal processing functions including:

- Downconversion of IF samples to baseband in-phase (I) and quadrature (Q) signals.
- Despreading of the forward link signal by correlation with local code replicas of the spreading code (spread spectrum mode).
- Generation of error signals for the acquisition, tracking and lock estimation functions of the carrier, subcarrier (residual carrier mode), PN code and bit synchronizer phase locked loops.
- Matched filter detection of the forward link command data and soft decision quantization to three bits.
- Coherent AGC for the full receiver dynamic range.
- Implementation of code NCO, receiver PN code generators, carrier NCO for the short loop Doppler removal, and subcarrier and bit synchronization tracking loops under control of the external microprocessor.
- Demodulation of the forward link ranging tones to baseband for driving the transmitter in residual carrier mode.

The receiver portion of the ASIC also contains the circuitry for intermodule connection and external spacecraft interfaces for control and telemetry functions.

The microprocessor performs the low rate baseband processing tasks most likely to be changed or customized by various users. It also implements signal processing software functions and management functions.

The microprocessor utilizes the baseband data samples to implement the following signal processing software functions:

- Simultaneous search for spread spectrum and residual carrier signal.
- Sequential detection algorithms for searching and synchronizing to the receiver short and long codes

- Sidelobe search and mainlobe verification algorithm to avoid tracking a sidelobe of the PN code autocorrelation function.
- Noncoherent code tracking loop filter (early-late loop) and frequency estimation to the carrier tracking loop to aid the carrier acquisition.
- Carrier tracking loop filter function using a modified Costas loop error function and coherent drive to the transmitter GaAs NCO.
- Provide programmable loop constants for the bit synchronizer and a bit lock indicator and squelch function by processing the bit loop error signal.
- Second order subcarrier tracking loop filter in residual carrier mode.

In addition to the signal processing tasks the microprocessor performs tasks related to control and telemetry including:

- At power-up, read the configuration data stored in the external programming connector.
- Monitor the receipt of 16 bit serial control words from the spacecraft and interpret them to configure the transponder operating states.
- Provide serial telemetry words and update the status of the bilevel telemetries.

#### *Transmitter Function*

The transmitter RF section provides the filtering and the up conversion of the synthesized carrier generated by the GaAs NCO. A 300 MHz chain performs the signal filtering and amplification and a second upconversion provides the S-band signal to the power amplifier.

The combination of the receiver and transmitter GaAs NCO frequency control provides the 100 MHz tuning range required for frequency flexibility and coherent turnaround capability.

The power amplifier section of the transmitter has been designed with a modular approach and can be easily changed to provide an output from 100 mW to 5 Watts.

The digital section of the transmitter generates the modulated signal and includes the CMOS ASIC to perform the signal processing tasks needed by transmitter operation.

The transmitter ASIC with the GaAs NCO provides frequency synthesis and phase modulation for the following modes of operation:

- QPSK modulation of the carrier with programmable Q:I power ratios. The power ratio can be set with finer than 0.1 dB resolution to virtually any value.
- BPSK modulation of the carrier (or PSK of any phase step size) with a single data channel.
- Convolutional encoding of the data selectable by serial command.
- Digital filtering of the modulating data streams to limit the output spectral bandwidth of the transmitter.
- Sampled linear modulation in the residual carrier mode using an external ADC.
- BPSK modulation by a digital bit stream of an internally generated 1.024 MHz subcarrier which is then linearly phase modulated onto the RF carrier.

Modulation of the return link carrier can be accomplished from several sources.

One source is in the pre-modulation filter output. The return link pre-modulation filter shapes the modulated transmitter output spectrum to meet NASA and ESA spectral mask requirements.

A second source of modulation is from the nine bit linear modulation input to the ASIC. This signal can be a BPSK modulated sine wave subcarrier, an analog ranging tone from the receiver, or both.

The microprocessor is not involved in real time signal processing in the transmitter, but is used only to control transmitter modes and characteristics. The microprocessor performs the following functions related to the transmitter ASIC:

- When the carrier loop has calculated the GaAs NCO frequency control word, the microprocessor loads it into the receiver ASIC parallel to serial shift register. This data is then automatically shifted into the transmitter ASIC.
- If the modulation filter is to be used, the microprocessor uses the serial interface to load the modulation filter with the delta phase words that control the slew rate characteristics of the filter.

- A control word is provided through the serial interface that selects programmable or selectable modes in the ASIC, such as selection of linear or QPSK modes, and turning the convolutional encoders on or off.

### ***DUAL STANDARD TRANSPONDER TECHNOLOGY***

The dual standard transponder has been designed using a high level of integration providing attractive performance in terms of mass and dimensions. To obtain these results different technologies have been used for the analog and digital design.

#### ***Radiation Hardened CMOS ASICs***

The transponder includes two CMOS ASICs which perform the complex signal processing tasks required for the receiver and transmitter operation. The chips are identical and consist of a receiver portion (about 40,000 gates) and a transmitter portion (about 12,000 gates). Some of the input/output pins are multiplexed on the chip so a programming pin is used to configure the chip in its target application.

The CMOS ASIC is manufactured using the LSI Logic LRH 10,000 family. The LRH 10,000 series is manufactured using an advanced radiation hardened 1.5-micron, planarized, silicon-gate fabrication process.

The transponder ASIC is implemented on the LRH10129 array which offers a gate complexity of about 129000 gates (over 51,000 are used). In the receiver application the ASIC provides the complex down conversion circuitry, digital correlators, bit synchronizer and demodulation circuits, PN code generators, coherent AGC and spacecraft interfacing circuitry. For the transmitter, a full digital premodulation filter, a subcarrier generator/modulator and convolutional encoding circuitry are implemented.

The ASICs are reconfigured, by means of numerous control signals, according to the detected signal type (spread spectrum or residual carrier) and the operating state of the transponder.

The chip is the key element in obtaining state-of-art performance in conjunction with attractive mechanical design.

#### ***Gallium Arsenide Numerically Controlled Oscillators and Digital to Analog Converter***

One of the key elements of the dual standard transponder is the Gallium Arsenide NCO which allows simple digital tracking loop implementations, complete frequency synthesis and phase modulation capability.

The chip has been developed using TriQuint Semiconductor's QLSI standard cell family. This logic family is implemented using Source Coupled FET Logic based on fully differential, current steered logic similar to silicon ECL. The QLSI cells are fabricated using the QED enhancement/depletion Gallium Arsenide process.

The NCO offers a 32 bit frequency control word and an independent 11 bit phase control word which allows the simultaneous direct digital frequency synthesis and modulation capability required for the transponder's operation. A clock rate of 260 MHz is used in the dual standard transponder design.

The GaAs NCOs, under the control of the microprocessor, are the key to providing operation over the complete 100 MHz frequency band (180 in-flight commandable channels). In addition they provide the possibility for operation in a long loop configuration for the Doppler removal and carrier tracking and a fully digital implementation of all the modulation formats (QPSK, BPSK, UQPSK, linear phase) needed for the communication to the Data Relay Satellite or directly to ground.

A 12-bit amplitude word is provided to the GaAs DAC for the generation of the analog signal. This component is a monolithic 14-bit IC which contains input protection, a current source generator, bit current switches, and clock and control circuitry. The dual standard transponder uses a version designed for 12 bits of linearity. The component has been specifically designed for direct digital synthesis applications.

#### *Special Features*

The Dual standard transponder has been designed to meet ESA and NASA requirements specifications and to provide a transponder easily adaptable to future user needs. The following features, some of which are not required by the system specifications, have been incorporated into the design to provide maximum flexibility in transponder use:

#### *Sampling Process*

The spread spectrum or residual carrier signal is quantized at an intermediate frequency (IF) and a complex full digital quadrature demodulation is performed inside the receiver ASIC. The ASIC is able to accept either 4 bit or 1 bit (hard decision) quantized samples of the IF signal for processing.

The four bit architecture provides the best performance in terms of overall bit error rate performance and this architecture was employed in the ADM transponder.

The hard quantization mode permits simplifications in the digitizing hardware at the expense of a small increase in implementation loss. These losses are minimized by a proper selection of IF bandwidth and sampling rate.

#### *Long Loop/Short Loop Configuration*

The receiver architecture allows the Doppler removal and the tracking of the carrier in two different ways, either long loop or short loop. In the long loop configuration the Doppler removal is performed by the receiver Gallium Arsenide NCO. The NCO is controlled by the microprocessor which updates the phase at a rate from 1-4 KHz depending on the forward link data rate and dynamics. In this mode the sampling process is coherent and implementation loss is minimized.

In the case of the short loop configuration a CMOS NCO internal to the ASIC signal processor is used for carrier tracking. The elimination of the GaAs NCO and DAC in the receiver can provide about 5 watts of power savings for a fixed frequency user.

#### *Command Words*

Serial control words set operational modes of the transponder. These words are input to the transponder as 16-bit serial words through one of two serial command input ports which consist of data, clock, and enable lines.

Notable control features include:

- Command data rate selection
- PN code selection
- Frequency channel selection
- Fine frequency control

The data rate selection command word allows the spread spectrum data rate to be selected in orbit from 125 to 300,000 bps without limitation. Initial code and carrier acquisition is performed at 125 bps. If the received Eb/No is above the command squelch threshold for the selected data rate, the receiver and the bit synchronizer are configured for the programmed rate.

An external programming connector includes a data rate change disable pin. If set the only data rate available is the default rate set by the programming connector.

The PN code selection command word provides selection of any one of the 85 codes in the code library.

The frequency selection command word controls the center frequency of the transponder, selecting one of the 180 available S-band channels. This selectability can be disabled by a pin on the programming connector.

In addition to the coarse frequency channel selection, a special feature, fine frequency offset, has been provided. In this mode fine frequency control words can be used to offset the frequencies of the receiver or transmitter (independently or together) to provide on-board Doppler compensation and oscillator aging frequency compensation for long term frequency management. The frequency word contains a 13 bit field in which a number in the range +4095 to -4096 is specified. The frequency will be offset by this word according to the equation:

$$F_r = F_c + (F_w \times F_{res})$$

where  $F_r$  is the actual frequency,  $F_c$  is the selected center frequency,  $F_w$  is the frequency control word and  $F_{res}$  is the frequency resolution (about 19.7 Hz). This yields a tuning range of approximately +/- 80.6 KHz.

#### ADM TESTING

The ADM was extensively tested in spread spectrum and residual carrier modes during 1991 and 1992. The transponder also successfully demonstrated operation with the NASA TDRSS system in July of 1992. The units overall performance was excellent. This third generation transponder contains the flexibility, advanced capabilities, and performance required to meet the rigorous requirements of the ESA DRSS and NASA TDRSS user transponder environment in the years ahead.

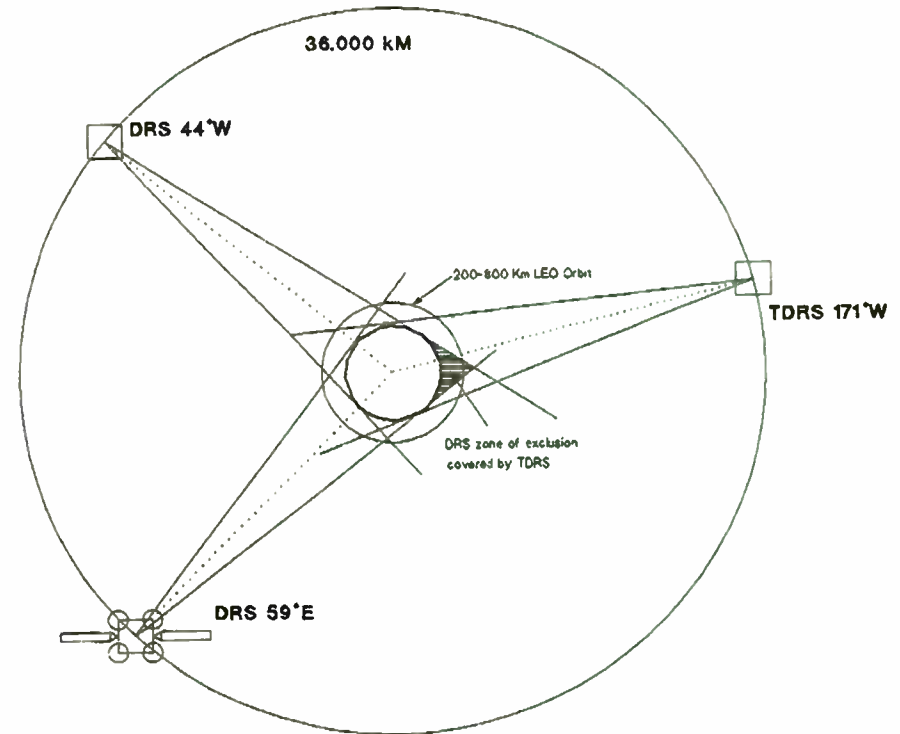
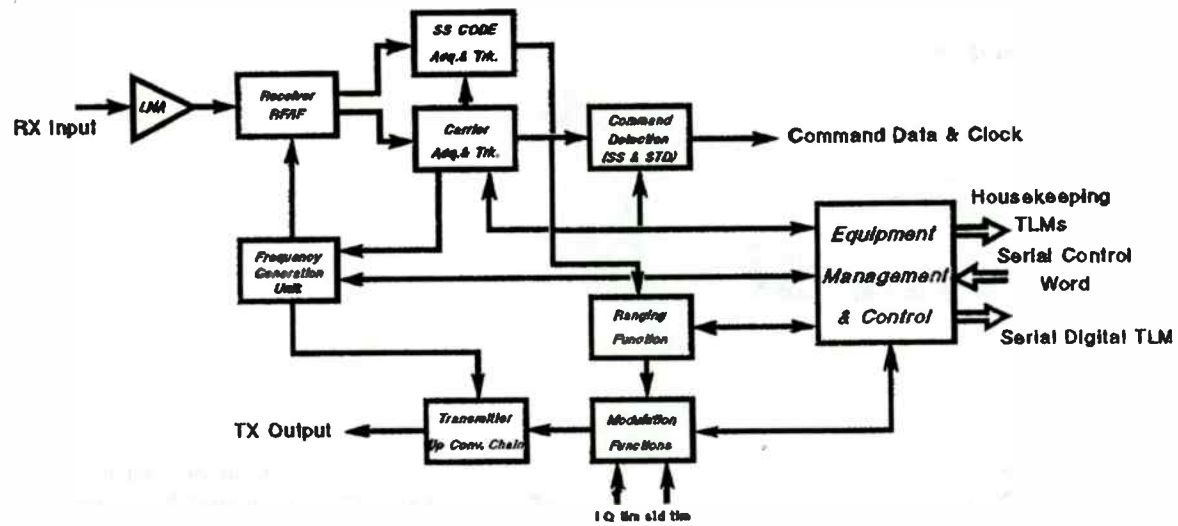
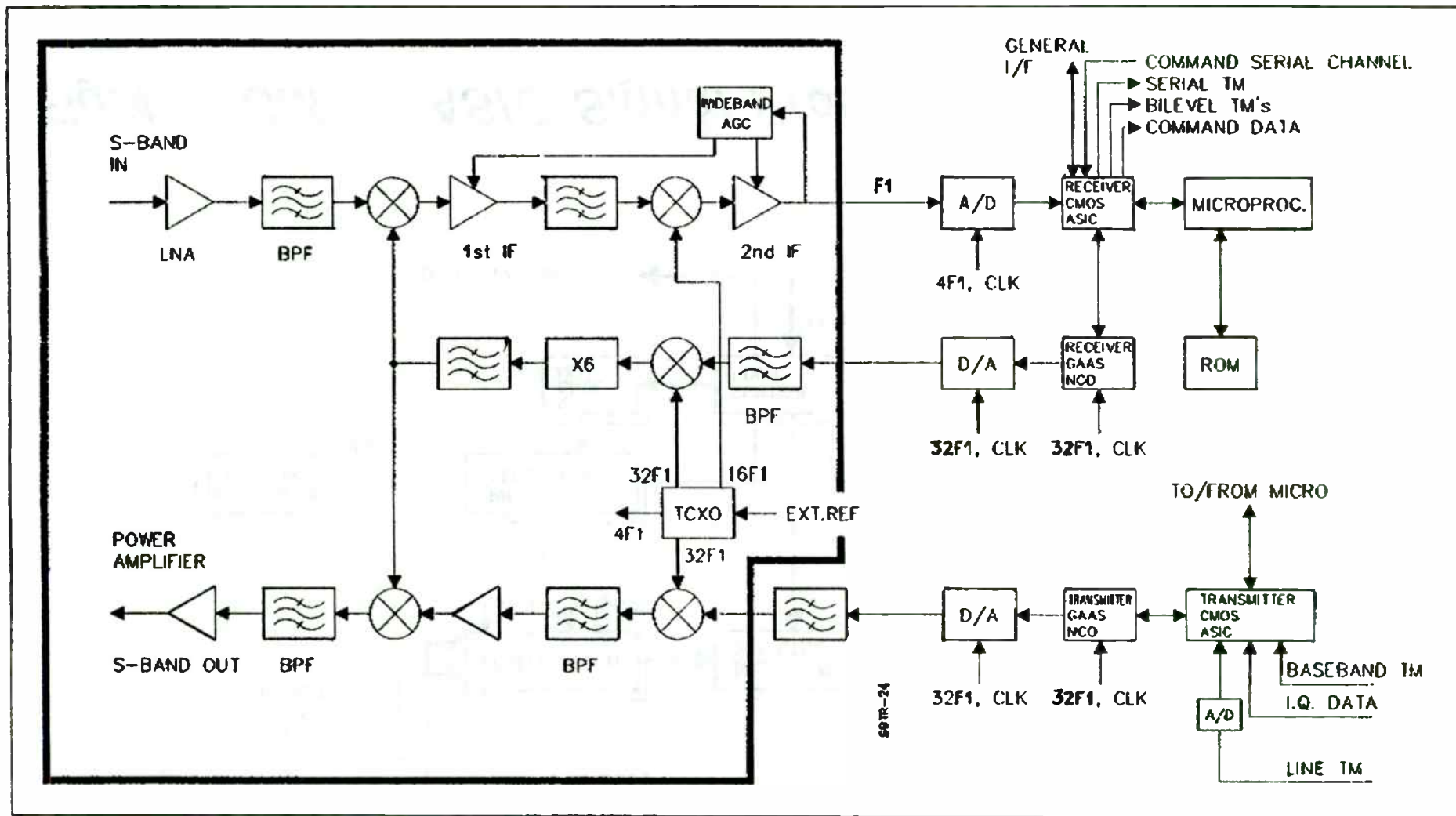


Fig.1 Data Relay Satellite System Concept

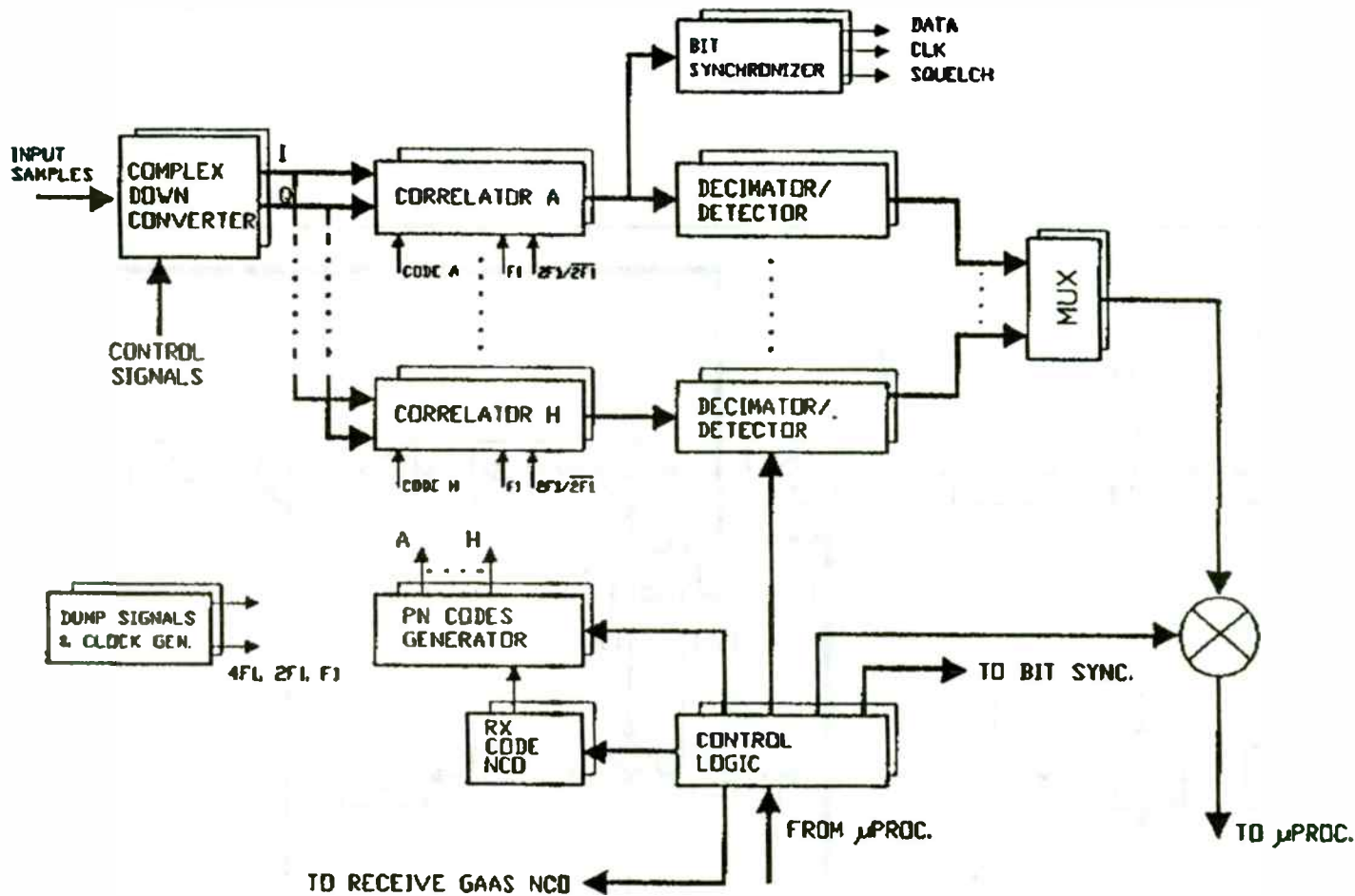


*Fig.2 S-Band Transponder Functional Block Diagram*





**Fig.3 S-Band Transponder Detailed Block Diagram**



**Fig.4 CMOS ASIC Signal Processing Block Diagram**

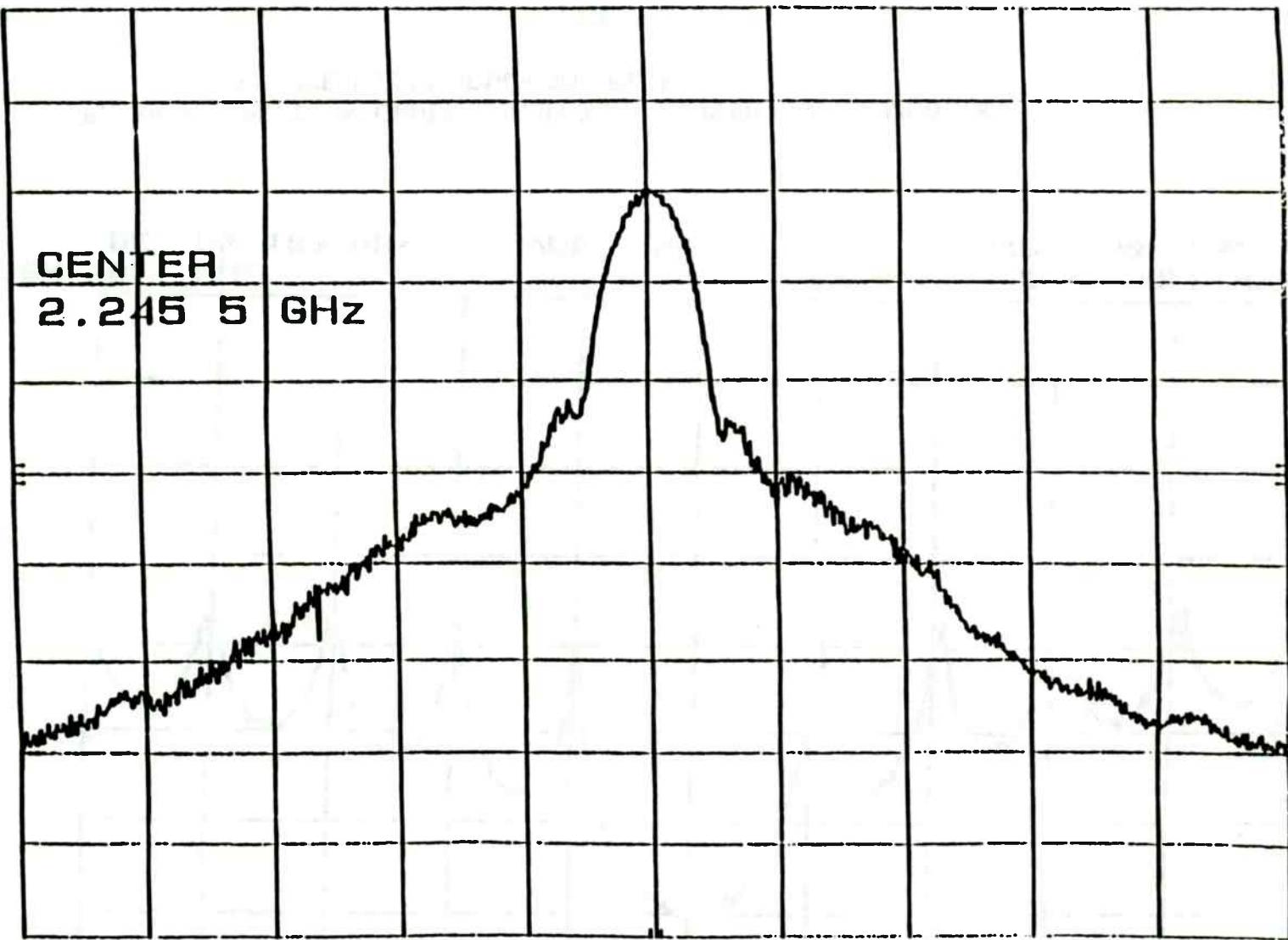
h<sub>2</sub>

REF -8.0 dBm

ATTEN 10 dB

MHz 2.245 50 GHz  
-28.00 dBm

10 dB/



CENTER 2.245 5 GHz

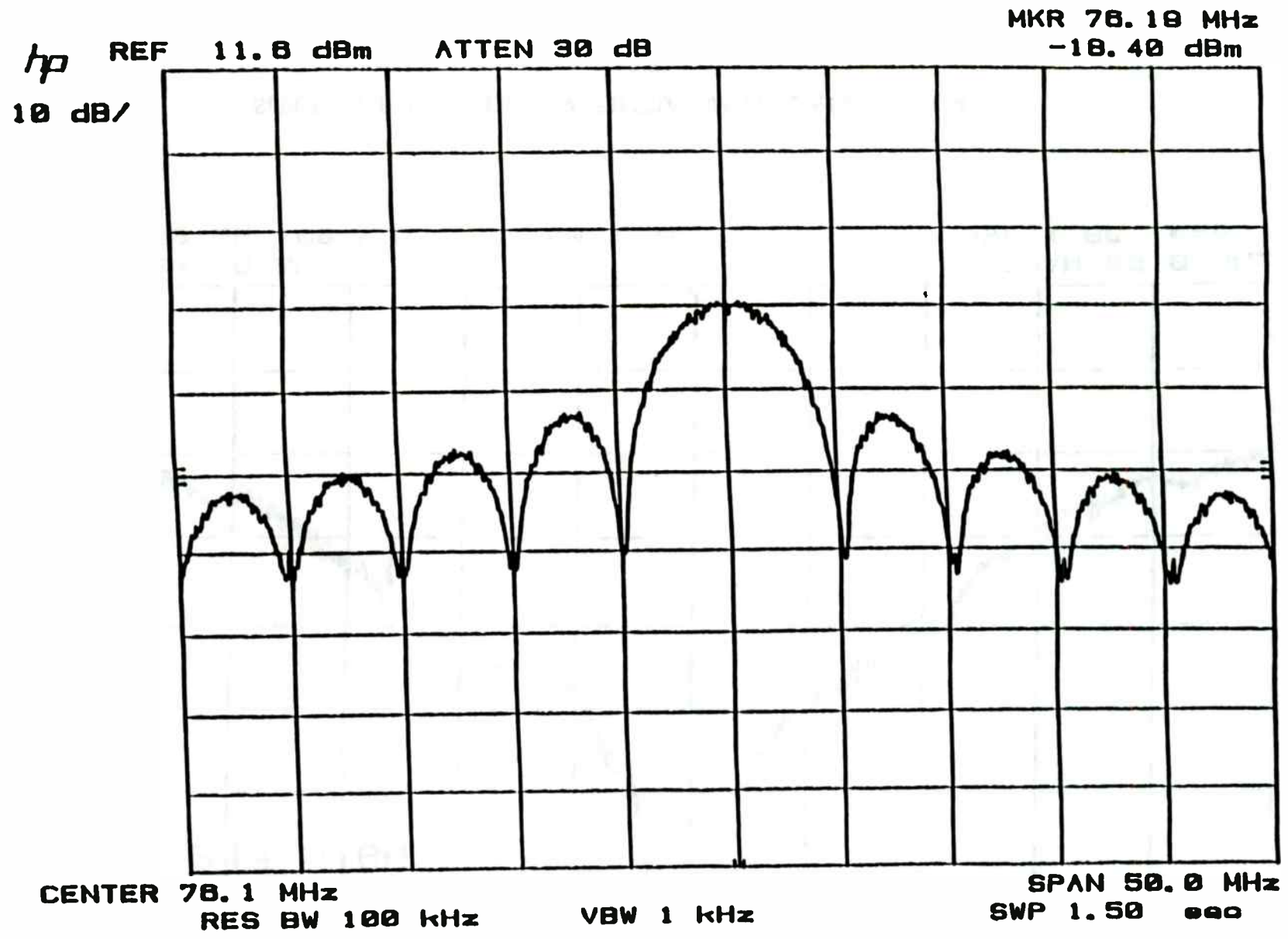
RES BW 100 kHz

VBW 1 kHz

SPAN 60.0 MHz

SWP 1.00 sec

SPECTRAL PLOT OF THE SHAPED TRANSMITTER PN SPECTRUM



BASEBAND SPECTRUM GENERATED BY TRANSMITTER DIGITAL MODULE  
WITHOUT THE PREMODULATION FILTER

**Iterative Algorithms for the Design of Arbitrary Phase  
Finite Impulse Response Functions**

C. A. Fulton, S. M. Richie, and B. P. Abbott\*

University of Central Florida  
Dept. of Electrical and Computer Engineering  
Orlando, FL 32816-2450

\*RF Monolithics  
4441 Sigma Road  
Dallas, TX 75244

**ABSTRACT**

Several algorithms have been developed which solve the complex Chebyshev approximation problem for FIR filter design [1]-[4]. In this paper the adaptive approach by Chit and Mason [1] is combined with the weight adjustment scheme developed by Lim, Lee, Chen, and Yang [5]. This hybrid algorithm retains the superior versatility of the original Chit and Mason routine while requiring only 1% to 5% of the execution time. The large decrease in number of iterations results in reduced numerical noise, thus larger order filters are possible. Several examples are presented which illustrate the power (and speed) of the modified algorithm; these include the design of a 30<sup>th</sup> order nearly linear phase filter, the design of a 60<sup>th</sup> order chirp filter, the design of a 127<sup>th</sup> order filter with shaping cost function, and the design of an 800<sup>th</sup> order sine delay filter. No parameter adjustment is required between examples.

**1.0 INTRODUCTION**

The design of filters having a specified frequency response optimized in the complex Chebyshev sense (maximum magnitude of the complex error minimized) is quite useful in a variety of applications; delay equalization in surface acoustic wave transducers and the design of nearly linear phase filters with a nearly constant group delay are but two examples. Indeed, the design of filters having other than linear or minimum phase requires that the optimization be done in the complex domain. Several authors [1,2,3,4] have developed algorithms to solve this complex Remez problem. This paper will investigate a modification to the Chit and Mason approach [1] in which the weight adjustment scheme developed by Lim, Lee, Chen, and Yang [5] is incorporated. The resulting program exhibits superior versatility, robustness, and speed.

Problem Statement

The objective is to find the FIR filter coefficients that, for a given filter order  $N$ , will minimize the maximum magnitude of the complex (weighted) error in the frequency domain. Since the design is to be optimized in the Chebyshev or minimax sense, the magnitude of the deviation (error) will exhibit an equiripple behavior. There will be at least  $N+2$  but no more than  $2N+3$  extremals, each having magnitude  $\delta$  [1-5].

Given a desired complex frequency response  $D(w)$  which is to be approximated by the designed frequency response  $H(w)$ , the complex error function  $E(w)$  is defined as

$$E(w) = [D(w) - H(w)] W_d(w)$$

where  $W_d(w)$  is a prescribed weighting (cost) function controlling relative values of maximum error at different frequencies. The larger the cost at a given frequency, the smaller will be the deviation at that frequency compared to its value had the cost been unity. Since the designed frequency response of an  $N^{\text{th}}$  order filter is the discrete Fourier Transform of its coefficients  $h(k)$ , then

$$E(w) = [D(w) - \sum_{k=0}^N h(k) e^{-jkw}] W_d(w)$$

The minimax design requires finding those  $N+1$  coefficients such that  $\delta = \min\{\max\{|E(w)|\}\}$  [1-5].

Alternate Approaches

McClellan and Parks [6,7] developed a fast, flexible algorithm for designing FIR linear phase filters by unifying the expression for their frequency response. The response of a general filter can be written in terms of its real-valued gain function  $G(w)$  (which can have negative values) and its phase  $\Theta(w)$ ,

$$D(w) = G(w) e^{j\Theta(w)}$$

Linear phase filters have a frequency response which can be written as

$$H(w) = G(w) e^{j[A\omega - N\omega/2]}, \quad A = 0 \text{ or } 1.$$

There are thus four possible cases for these filters: odd order with  $A$  zero, odd order with  $A$  unity, even order with  $A$  zero, and even order with  $A$  unity. Note that the linear phase term,  $e^{j\omega N/2}$ , corresponds to a time shift in the impulse response and can thus be ignored. The frequency response is then either purely real or purely imaginary; this greatly simplifies the problem since the constant phase term can be ignored when dealing with the magnitude of the error. By writing

the gain as  $G(w) = Q(w)/P(w)$  where  $P(w)$  is a linear combination of cosine functions and  $Q(w)$  is an appropriate trigonometric expression, McClellan and Parks were able to unify the expressions for the frequency response. This allowed a single algorithm, the real Remez exchange, to solve all classes of linear phase filters [6,7]. The real Remez exchange iteratively finds the location of the extremal frequencies in an equiripple linear phase filter; it proceeds according to the flow chart in Figure 1 [8].

The linear restriction on phase greatly simplifies the Chebyshev problem, reducing it from the complex to the real domain and thus allowing use of the powerful alternation theorem [7]. In order to design filters other than linear or minimum phase, a complex Chebyshev routine is required. Preuss generalized the real Remez exchange to solve the complex problem; it proceeds according to the flow chart in Figure 2 [2,3].

Schulist [3] modified the Preuss algorithm, producing a marked decrease in the execution time. He proposed two alterations: that all extremals be utilized in the calculations, and that Gaussian relaxation replace Newton interpolation for non-linear phase filters. His procedure is summarized in Figure 3 [3].

Chen and Parks [4] developed an algorithm for finding the real FIR filter coefficients that approximate a complex Chebyshev design. The linear complex problem is converted into a nearly equivalent non-linear real problem; this real function is then transformed so that linear programming can be used to reach a solution. Their method requires a large amount of memory and is slow to converge; execution time increases exponentially with increasing frequency grid density [3,4].

Chit and Mason [1] use an adaptive approach to solve the complex Chebyshev problem for real FIR filter coefficients; their algorithm is unusual in that its power lies in the time domain rather than the frequency domain. They use the fact reported by several authors that a Chebyshev design can be produced by a least mean square fit if an appropriate weighting function is chosen; these costs which distort the function to produce the equiripple behavior are found iteratively [6]. The Chit and Mason method is exceptionally versatile but quite slow; literally hundreds of iterations are often required for convergence. The remainder of this paper will investigate this algorithm; modifications are made which will drastically improve the speed.

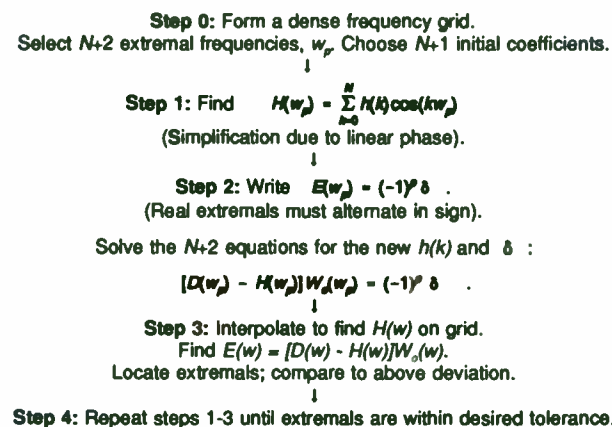


Figure 1- Real Remez Exchange

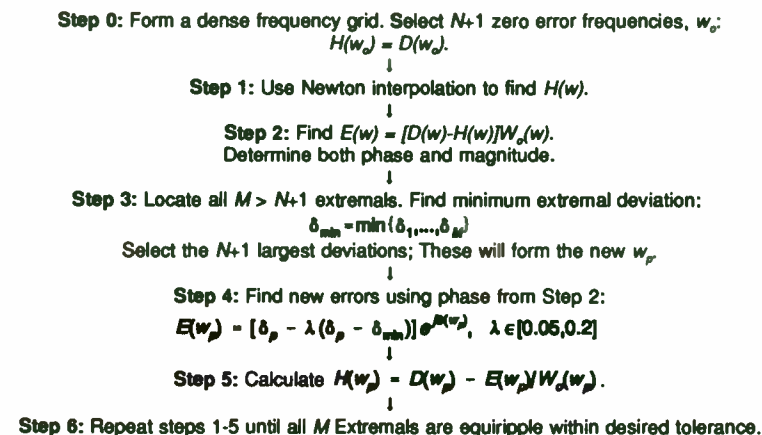


Figure 2 - Preuss Algorithm

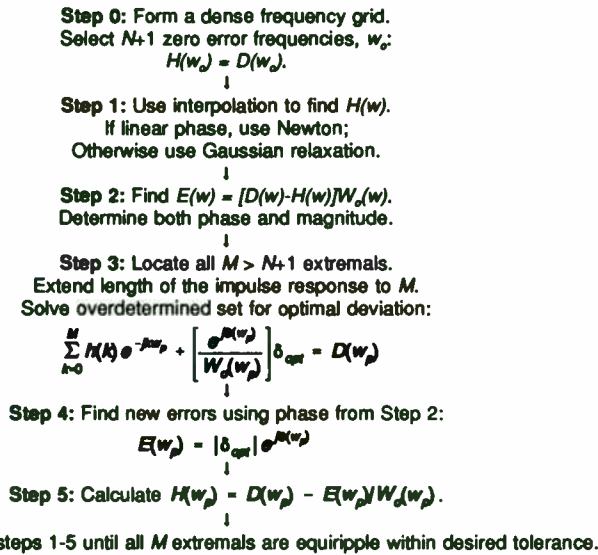


Figure 3 - Schulist Algorithm

## 2.0 CHIT AND MASON ALGORITHM

This chapter presents an implementation of the Chit and Mason algorithm as described in [1]; examples are included for comparison. With a single exception, the results are in good agreement with those reported in [1]; the chirp filter, however, is designed with less than ten percent of the deviation quoted by Chit and Mason. The reason for this implementation's superior performance is currently unknown.

### Algorithm

The desired filter order  $N$ , gain function  $G(w)$ , phase function  $\theta(w)$ , and weighting function  $W_d(w)$  must be specified over the frequency interval  $[0, \pi]$ . This interval is halved because the filter coefficients are assumed real; therefore, the gain is an even function and the phase is an odd function. The algorithm proceeds according to the flow chart in Figure 4 [1]:

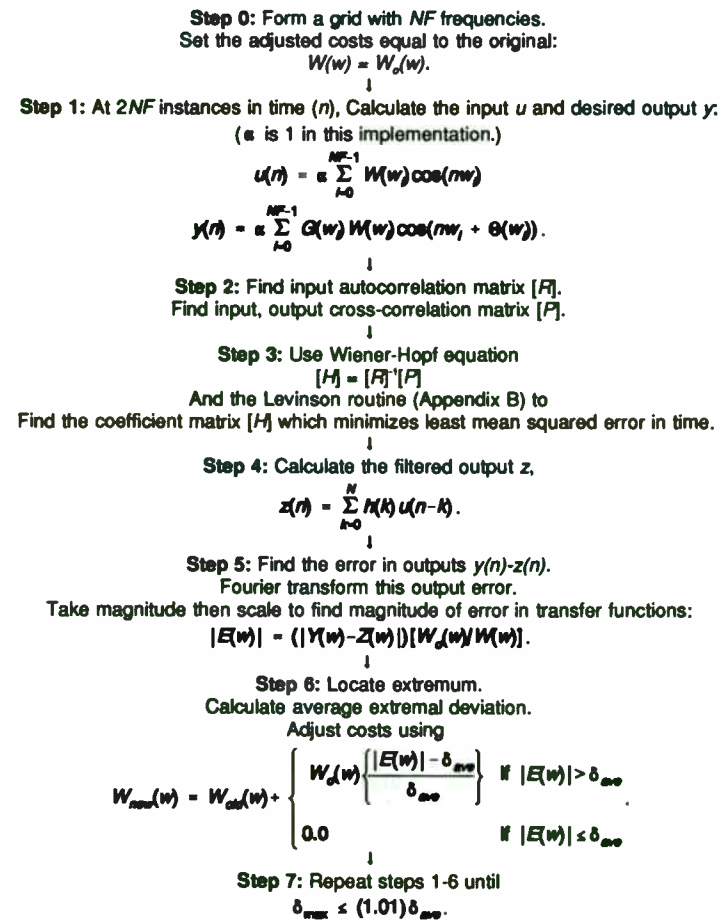


Figure 4: Chit and Mason Algorithm

Chit and Mason recommend  $NF$  be equal to five times the filter order. From Table I in the next section, the extremal deviation is seen to be sensitive to grid density in an unpredictable fashion; therefore, accuracy cannot be optimized by choosing a specific  $NF$ . This implementation thus chooses to minimize speed by setting  $NF$  equal to a power of two, thereby allowing use of the fast Fourier transform. The largest power of two that is smaller than five filter orders is found to be effective for all but the 1200<sup>th</sup> order filters. Memory requirements for the latter force the use of a discrete (slow) Fourier transform with a less dense grid; 3600 frequencies is found to be sufficient (perhaps not necessary) for these to converge.

#### Examples

The program is run on a 486/33 Intel based computer; stated execution times do not include much of the final print to the output file. The criteria for convergence is that the maximum deviation be within one percent of the average extremal deviation. The frequency  $f$  is normalized to  $2\pi$  thus parameters are specified from 0.0 to 0.5. The filter specifications are as follows:

##### Example 1) Twin Delay All Pass

$G(f)=1$ ,  $W_d(f)=1$   
 Group Delay 1: 34,  $f < 0.25$   
 Group Delay 2: 26,  $f \geq 0.25$

##### Example 2) Sine Delay All Pass

$G(f)=1$ ,  $W_d(f)=1$   
 $\theta(f) = -60\pi f + 2\pi(1 - \cos(2\pi f))$

##### Example 3) Chirp All Pass

$G(f)=1$ ,  $W_d(f)=1$   
 $\theta(f) = -60\pi f - 8\pi(2f-5)^2$

##### Example 4) NLP Band Pass, Group Delay 46

Passband:  $0.12 \leq f \leq 0.19$ ,  $G(f)=1$   
 Stopband:  $f \leq 0.1$  and  $f \geq 0.15$ ,  $G(f)=0$

$$W_d(f) = \begin{cases} \frac{10}{1-9f} & , 0.00 \leq f \leq 0.10 \\ 1 & , 0.12 \leq f \leq 0.19 \\ \frac{10}{9f-1.25} & , 0.15 \leq f \leq 0.25 \\ 10 & , 0.25 \leq f \end{cases}$$

##### Example 5) NLP Low Pass, Group Delay 12

Passband:  $f \in [0.0, 0.08]$ ,  $W_d(f)=1$ ,  $G(f)=1$   
 Stopband:  $f \in [0.12, 0.5]$ ,  $W_d(f)=10$ ,  $G(f)=0$

As was previously discussed, the selection of grid density affects the accuracy of the solution in an unpredictable fashion; this is clearly illustrated in Table I. This implementation thus chooses to optimize speed. The frequency grid contains 128, 256, 512, and 2048 frequencies, respectively, for the filters of order 30, 60, 127, and 800 (the largest power of two smaller than five filter orders).

The first four examples are taken from [1] in order to verify the validity of the current implementation as well as to collect data on executions times and iteration numbers, parameters which were not specified. The first three designs are all-pass filters with unity costs; the fourth is a bandpass filter with shaping costs. Note that the typographical errors in [1] have been corrected. The final example is from Chen and Parks [5], Preuss [3] and Schullist [4]; a direct comparison will be made between this Chit and Mason with the alternate algorithms.

The first design is an all-pass twin delay filter with unity weighting. The group delay is 34 for frequencies less than 0.25 and 26 for those greater or equal to 0.25. At 60<sup>th</sup> order this filter is designed by the current implementation in 5.42 minutes, 212 iterations, with a maximum deviation of 0.0384 and 31 extremals. These results are in good agreement with Chit and Mason [1], who also obtained a peak error of 0.0384 with 31 extremals.

The second example is an all pass filter with unity costs and a sinusoidal phase,

$$\theta(f) = -60\pi f + 2\pi(1 - \cos(2\pi f)).$$

At 60<sup>th</sup> order this filter is designed by the current implementation in 3.02 minutes, 181 iterations, with 32 extremals and a maximum deviation of 0.000961. These results are also in close agreement with [1] and thus shall not be illustrated here. The third filter is an all pass chirp with



unity costs and phase given by

$$\theta(f) = -80\pi f + 2\pi(1 - \cos(2\pi f)).$$

Table I: Dependence on Grid Density

Filter	Maximum Deviation & Execution Time			
	$NF=2^2$	$NF = 5N$	$NF = 5N+5$	$NF=2^{11}$
Twin, All $N=60$	0.0384 5.42 min	0.0388 10.47 min	0.0323 9.37 min	0.0389 24.97 min
Sine, All $N=60$	9.61E-4 3.02 min	9.72E-4 7.18 min	9.73E-4 7.78 min	9.66E-4 15.15 min
Chirp, All $N=60$	1.06E-3 3.13 min	1.05E-3 7.12 min	1.05E-3 7.10 min	1.06E-3 16.52 min
NLP, Band $N=127$	0.0230 19.9 min	0.0244 45.05 min	0.0261 33.00 min	0.0233 76.9 min
NLP, Low $N=30^a$	0.0381 3.32 min	0.0437 6.27 min	0.0348 5.70 min	0.0383 16.88 min

The current implementation requires 3.13 minutes, 123 iterations, for convergence with a maximum deviation of 0.00106 and 31 extremals. The Chit and Mason algorithm in [1] had a much larger peak error of 0.02036; the cause of this large discrepancy is unknown. The superior design is illustrated in Figures 5-9.

The fourth example is a nearly linear phase band pass filter with a group delay of 46 and shaping costs. The passband extends over the interval  $0.12 \leq f \leq 0.13$  and the two stopbands have frequencies  $f \leq 0.1$  or  $f \geq 0.15$ . The cost function is given by

$$W_d(f) = \begin{cases} \frac{10}{1-9f} & , 0.00 \leq f \leq 0.10 \\ 1 & , 0.12 \leq f \leq 0.13 \\ \frac{10}{9f-1.25} & , 0.15 \leq f \leq 0.25 \\ 10 & , 0.25 \leq f \end{cases}$$

At 127<sup>th</sup> order, the current implementation completes this design in 19.9 minutes, 314 iterations, with a maximum deviation of 0.0230 and 66 extremals. This is an improvement to the 0.0278 and 67 extremals reported in [1], the cause of which could be related to choice of grid density (see Table I). Figures 10-14 reveal the results. Table II summarizes the comparison of extremal deviations between the current implementation of the Chit and Mason algorithm and that from [1].

Table II: Accuracy of Chit and Mason Implementations

	Error, From [1]	Error, This Paper
Twin, All $N=60$	0.0384	0.0384
Sine, All $N=60$	close to 0.00093	0.000961
Chirp, All $N=60$	0.02036	0.00106
NLP, Band $N=127$	0.0278	0.0230

The final example is a nearly linear phase low pass filter with a group delay of 12. The passband extends to  $f=0.06$ , inclusive, with unity costs; the stopband begins at  $f=0.12$  with  $W_d(f)=10$ . This choice of costs results in an unweighted error ripple one order of magnitude greater in the passband than in the stopband. At 30<sup>th</sup> order this filter is designed by the current

implementation of Chit and Mason in 3.32 minutes, 760 iterations, with a maximum deviation of 0.0381 and 18 extremals. The results are shown in Figures 15-19.

This same 30<sup>th</sup> order NLP low pass filter was designed on a VAX 11/750 by Chen and Parks [5], Preuss [3], and Schulist [4]. The Chen and Parks algorithm required 18 minutes, 206 iterations, with a peak error of 0.0436; the Preuss algorithm took about 3.2 minutes, 113 iterations, with a maximum error of 0.0426; the Schulist algorithm needed only about 15 seconds, 8 iterations, with a maximum deviation of 0.0425. While the Schulist algorithm has an impressive speed advantage, the Chit and Mason routine does yield the most accurate design. Table III summarizes these results; Table IV compares the filter coefficients found by this implementation with those by Preuss.

The Chit and Mason algorithm is exceptionally versatile, accurately handling filters as diverse as nearly linear phase to chirps without needing adjustment between designs. Accuracy is good but convergence is slow, especially for higher orders. The number of iterations is excessive; improvements in this area are investigated in the next chapter.

Table III: Comparison for NLP, 30<sup>th</sup> Order

	Chen/Parks	Preuss	Schulist	Chit/Mason
Peak Error	0.0436	0.0426	0.0425	0.0381
Time (Machine)	18 min (VAX)	3.2 min (VAX)	15 sec (VAX)	3.3 min (486/33)
Iterations	206	113	8	760

Table IV: Filter Coefficients for NLP, 30<sup>th</sup> Order

x.01	Preuss	Chit/Mason	x.01	Preuss	Chit/Mason
h(0)	-0.2710	-0.2979	h(16)	+7.1853	+7.1136
h(1)	-0.9775	-0.9446	h(17)	+3.1904	+3.1575
h(2)	-1.3469	-1.3540	h(18)	-0.1505	-0.1334
h(3)	-1.9356	-1.9112	h(19)	-2.3986	-2.3316
h(4)	-1.9985	-1.9946	h(20)	-3.3919	-3.2925
h(5)	-1.4779	-1.4582	h(21)	-3.2641	-3.1617
h(6)	-0.0213	-0.0001	h(22)	-2.3642	-2.2796
h(7)	+2.3986	+2.4272	h(23)	-1.1068	-1.0790
h(8)	+5.6292	+5.6696	h(24)	+0.0720	+0.0485
h(9)	+9.2845	+9.3244	h(25)	+0.9267	+0.8510
h(10)	+12.775	+12.807	h(26)	+1.3155	+1.1968
h(11)	+15.459	+15.474	h(27)	+1.3255	+1.2105
h(12)	+16.794	+16.778	h(28)	+0.9617	+0.6702
h(13)	+16.460	+16.411	h(29)	+0.7337	+0.6547
h(14)	+14.175	+14.399	h(30)	+0.24936	+0.1936
h(15)	+11.184	+11.097			

### 3.0 MODIFICATIONS

This chapter discusses the different modifications made to the Chit and Mason algorithm [1] in an effort to increase its speed. Incorporation of the Lim, Lee, Chen, and Yang frequency weighting routine [5] proves most beneficial, resulting in execution times of 1% to 5% the original. Specific examples will be given in Chapter IV.

### Initial Cost Dependence

The Chit and Mason cost adjustment scheme [1] controls the number of required iterations for convergence; this routine needs revamping in order to obtain a significant improvement in speed. Analysis reveals an important factor responsible for the sluggish response: the dependence on the original weighting function. The Chit and Mason cost adjustment routine can be rewritten in the form

$$W_{new}(w) = W_{old}(w) \begin{cases} \left(1 + \beta \left[\frac{W_d(w)}{W(w)}\right]\right), & \text{if } |E(w)| > \delta_{ave} \\ 0 & , \text{if } |E(w)| \leq \delta_{ave} \end{cases}$$

where  $\beta = (|E(w)|/\delta_{ave}) - 1$ . The adjustment then depends on two parameters,  $\beta$  and the cost ratio  $W_d(w)/W(w)$ , both of which decrease as the program progresses. The first dependence is desirable; smaller adjustments need be made as convergence approaches lest the alteration be too strong and divergence result.

The second dependence, however, is deleterious. As the number of iterations increases, the cost ratio decreases for all errors larger than the average extremal deviation. The ratio can drop under one percent (as seen in Chapter II), resulting in an excessively small adjustment per iteration; thus the convergence becomes overdamped.

The first modification attempted eliminates this undesirable dependence on the original cost function. The adjustment is simply made proportional to the current weighting:

$$W_{new}(w) = W_{old}(w) \begin{cases} (1 + \beta), & \text{if } |E(w)| > \delta_{ave} \\ 0 & , \text{if } |E(w)| \leq \delta_{ave} \end{cases}$$

Although this alteration is minor, its impact is not; execution time decreases by a factor of 5 to 15. There is frequently a slight (less than 1%) increase in the maximum deviation; Table V summarizes the effects of eliminating the dependence on initial costs.

Table V: Effect of Dependence on Original Costs

Filter	Original (Dependent)			Altered (Independent)		
	Time	# Iter	$\delta \times 10^{-4}$	Time	# Iter	$\delta \times 10^{-4}$
Twin, All N=60	5.42 m	212	384	0.68 m	27	384
Sine, All N=60	3.02 m	118	9.61	0.53 m	21	9.64
Chirp, All N=60	3.13 m	123	10.6	0.55 m	22	10.6
NLP, Band N=127	19.9 m	314	230	2.82 m	45	231
NLP, Low N=30	3.32 m	760	381	0.22 m	51	382

An attempt was then made to increase speed by removing the selectivity of the adjustment. Instead of simply forcing the largest errors down, the small errors were pulled up simultaneously; all errors received an adjustment each iteration. This method did not yield any significant advantage; the more sophisticated approach of Lim, Lee, Chen and Yang was required for further optimization.

### Lim Hybrid

Lim, Lee, Chen, and Yang [5] developed an algorithm for finding the least mean squares weighting function that produces an optimal Chebyshev design. Their routine finds the  $r(w)$  that

yield an equiripple error by minimizing  $\sum_{k=0}^{M-1} r(w) E^2(w)$ .

The  $r(w)$  correspond to the  $W^2(w)$  in the Chit and Mason approach.

The Lim, Lee, Chen, and Yang method uses an envelope  $B(w)$  to find the adjustment in weights. The envelope is just a series of line segments connecting the extremals in a frequency band; for this calculation the band edges are considered extremals. The deviations are adjusted in the following manner to prevent the envelope from misleadingly approaching zero. The extremal deviation is compared to those of the immediately adjacent peaks in the same frequency band; the smaller error of these two (or one) neighboring peaks is found. If the value of the current deviation is less than 10% of this minimum value, then the current peak error is redefined

as 10% of this minimum.

Once the envelope is defined, the weights are adjusted according to

$$r_{\text{new}}(w) = r_{\text{old}}(w) \left[ \left( \frac{E(w)}{W_d(w)} \right) \left( \frac{W_d(w)}{E(w)} \right)_{\text{opt}} \right]^{\gamma}$$

where the optimal power was found to be 1.5 [5]. This equation is slightly modified for incorporation into the Chit and Mason algorithm. The actual cost update becomes

$$W_{\text{new}}(w) = W_{\text{old}}(w) \left[ \left( \frac{E(w)}{B_{\text{opt}}} \right) \right]^{\gamma}$$

The new optimal  $\gamma$  was found to be 0.8, in good agreement with the observation that  $r(w)$  corresponds to  $W^2(w)$ .

Incorporation of this routine has a dramatic impact on performance. The execution time drops to as little as one percent of the original, taking with it the iteration number and thus numerical noise. The design of very large order filters now becomes feasible as well as practical.

#### 4.0 EXAMPLES

The filters specified in Chapter II will be designed by the hybrid algorithm for direct comparison with the original Chit and Mason. Filters of large order (800 and 1200) will be designed, demonstrating some of the hybrid's potential. The convergence process will be investigated.

##### Moderate Order

The hybrid requires just 9 seconds, 6 iterations, to design either the twin delay, sine delay, or chirp filter of 60<sup>th</sup> order; the respective peak deviations are 0.0385, 0.000961, and 0.00107 and are in good agreement with the original results. The NLP band pass filter of order 127 requires only 22 seconds, 6 iterations, with a maximum error of 0.0243. This peak error is about 6% larger than that determined by the unmodified Chit and Mason routine; however, the execution time for the hybrid is 98% less. The NLP low pass filter of order 30 is designed in a scant 2 seconds, 11 iterations, with maximum deviation of 0.0380; this agrees well with the previous results. Figures 20-40 compare the designs of the new with the old.

##### Large Order

With the exception of the twin delay all pass specification, the original Chit and Mason

algorithm cannot successfully design filters at 800<sup>th</sup> order; the numerical noise is too excessive. A lengthy 12.70 hours, 396 iterations, is required for convergence in the one case in which it does succeed. The performance of the hybrid routine is far superior. Although it also fails in designing the nearly linear phase specifications, it has no difficulty whatsoever with the others.

At order 800 the hybrid requires a mere 15.4 minutes, 8 iterations, to complete the twin delay filter; the maximum deviation is 0.0183. The 800<sup>th</sup> order sine delay filter and the chirp each require 13.5 minutes, 7 iterations, with respective deviations of 0.000511 and 0.000767. Table VI summarizes the effect of incorporating the Lim, Lee, Chen, and Yang cost adjustment scheme into the Chit and Mason algorithm.

These three filters can also be constructed at order 1200, though much more slowly since memory requirements prevent the use of a fast Fourier transform. A grid containing 3600 frequencies is used. The procedure requires between 45 and 71 minutes, depending on the filter; this is still much faster than the time required by the original routine to produce a filter of only 800<sup>th</sup> order (if it succeeds).

The twin delay filter of order 1200 requires 70.0 minutes, 9 iterations, with peak deviation of 0.0181. The sine delay is relatively speedy, requiring only 46.7 minutes, 6 iterations, with a maximum error of 0.000505; the chirp filter can be designed in 54.6 minutes, 7 iterations, with a maximum deviation of 0.000760. Results for the sine delay are illustrated in Figures 41-42.

Table VI: The Hybrid Impact

Filter	Order		Original	Hybrid
Twin Delay, All Pass	60	Time	5.42 min	9 sec
		# Iter.	212	6
		Error	0.0384	0.0385
	800	Time	12.7 hrs	15.4 min
		# Iter.	396	8
		Error	0.0183	0.0183
Sine Delay, All Pass	60	Time	3.02 min	9 sec
		# Iter.	118	6
		Error	0.000961	0.000961
	800	Time	fails	13.5 min
		# Iter.	fails	7
		Error	fails	0.000511
Chirp, All Pass	60	Time	3.13 min	9 sec
		# Iter.	123	6
		Error	0.00106	0.00107
	800	Time	fails	13.5 min
		# Iter.	fails	7
		Error	fails	0.000767
NLP, Band Pass	127	Time	19.87 min	22 sec
		# Iter.	314	6
		Error	0.0230	0.0243
NLP, Low Pass	30	Time	3.32 min	2 sec
		# Iter.	760	11
		Error	0.0381	0.0380

Convergence

Figures 43-47 demonstrate the hybrid's vastly superior method of convergence. Note that curves for the original routine have been truncated for clarity; they continue for a hundred or more iterations, ever so gradually decreasing towards the optimal deviation. The large jump in the hybrid curve in Figure 47 corresponds to a change in the number of located extrema's.

**5.0 CONCLUSION**

The Chit and Mason approach coupled with the Lim, Lee, Chen, and Yang weight adjustment scheme proves to be a superior method for finding real FIR filter coefficients that are optimized in the Chebyshev sense. This hybrid algorithm demonstrates the exceptional versatility of the original Chit and Mason yet requires 1% to 5% of the time. It also needs no adjustment, no human intervention or intuition, in order to proceed with different designs.

Additionally, the hybrid shows great promise for constructing filters of very large order, 800 and more, which is of particular value to surface acoustic wave device design. The reduced number of iterations make this possible by decreasing the numerical noise; the increase in speed makes large order design more practical. Further investigation is required to eliminate the difficulty with large order nearly linear phase filters.

**LIST OF REFERENCES**

- [1] Chit, Nassim N. and Mason, John S. "Complex Chebyshev Approximation for FIR Digital Filters." *IEEE Trans. on Signal Processing*, vol. 39, no. 1, (January 1991): 49-54.
- [2] Preuss, Klaus "On the Design of FIR Filters by Complex Chebyshev Approximation." *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, (May 1989): 702-712.
- [3] Schulist, Matthias "Improvements of a Complex FIR Filter Design Algorithm." *Signal Processing*.
- [4] Chen, Xiangkun and Parks, Thomas W. "Design of FIR Filters in the Complex Domain" *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 2, (February 1987): 144-152.
- [5] Lim, Yong C.; Lee, Ju-Hong; Chen, C.K.; and Yang, Rong-Huan "A Weighted Least Squares Algorithm for Quasi-Equiripple FIR and IIR Digital Filter Design." *IEEE Trans. on Signal Processing*, vol. 40, no. 3, (March 1992): 551-558.
- [6] McClellan, James H. and Parks, Thomas W. "A Unified Approach to the Design of Optimum FIR Linear-Phase Digital Filters." *IEEE Trans. on Circuit Theory*, vol. CT-20, no.

6, (November 1973): 697-701.

- [7] McClellan, James H.; Parks, Thomas W.; and Rabiner, Lawrence R. "A Computer Program for Designing Optimum FIR Linear Phase Digital Filters." *IEEE Trans. on Audio and Electroacoustics*, vol. AU-21, no. 6, (December 1973): 506-526.
- [8] Bellanger, Maurice Digital Processing of Signals. New York: John Wiley & Sons, Inc., (1984): 126-128, 291-293.
- [9] Press, William H.; Flannery, Brian P.; Teukolosky, Saul A.; and Vetterling, William T. Numerical Recipes in C. New York: Cambridge University Press (1991): 398-470.
- [10] Blahut, Richard E. Fast Algorithms for Digital Signal Processing. Massachusetts: Addison-Wesley Publishing Company (1985): 352-359.

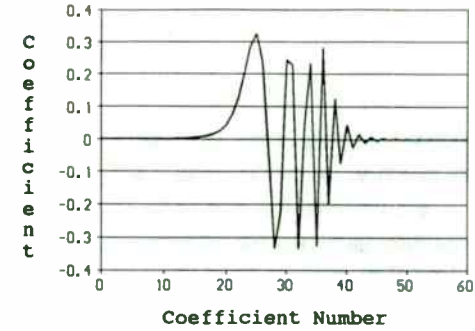


Figure 5: Chirp, N=60 (Chit & Mason) Filter Coefficients

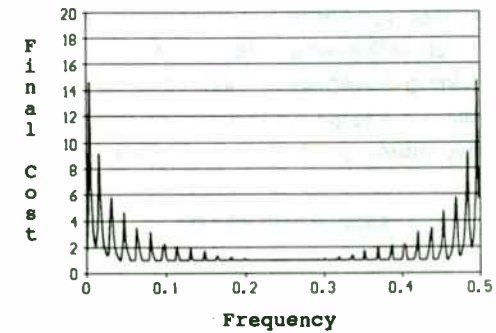


Figure 6: Chirp, N=60 (Chit & Mason) Final Adjusted Costs vs Frequency

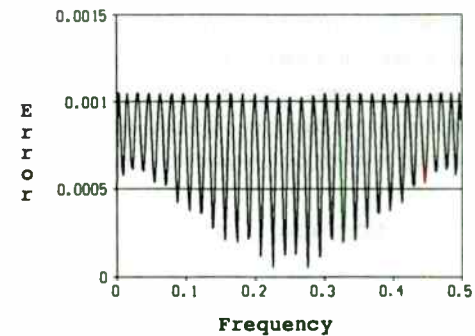


Figure 7: Chirp, N=60 (Chit & Mason) Magnitude of Complex Error vs Frequency

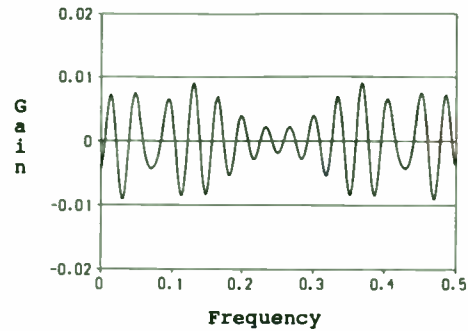


Figure 8: Chirp, N=60 (Chit & Mason)  
Gain (dB) vs Frequency

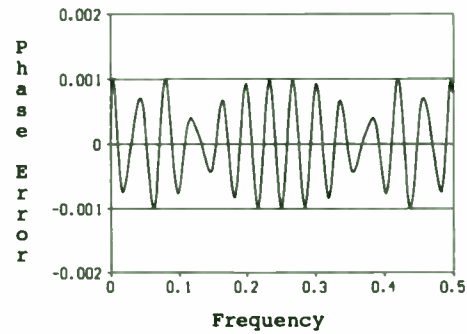


Figure 9: Chirp, N=60 (Chit & Mason)  
Phase Error (rad) vs Frequency

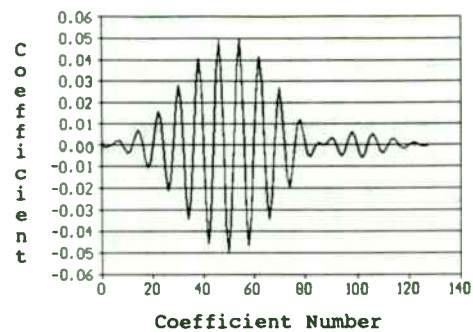


Figure 10: NLP Band Pass, N=127 (C & M)  
Filter Coefficients

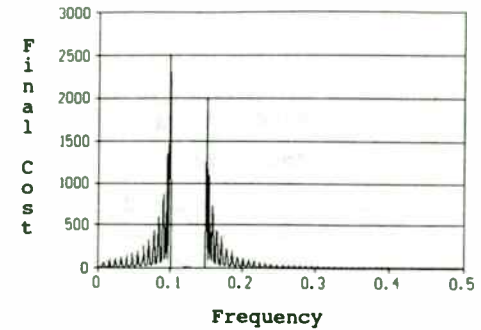


Figure 11: NLP Band Pass, N=127 (C & M)  
Final Adjusted Costs vs Frequency

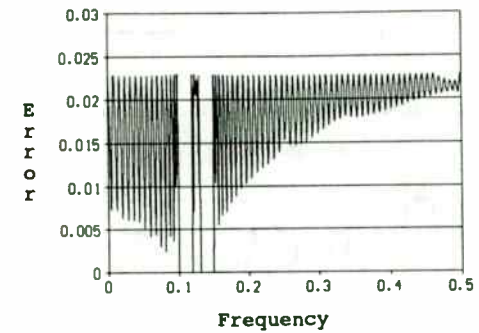


Figure 12: NLP Band Pass, N=127 (C & M)  
Magnitude of Complex Error vs Frequency

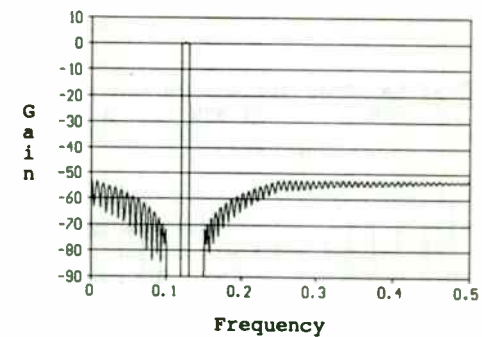


Figure 13: NLP Band Pass, N=127 (C & M)  
Gain (dB) vs Frequency

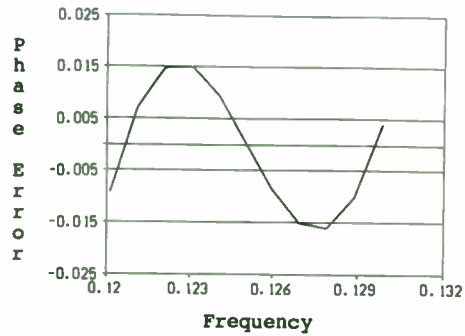


Figure 14: NLP Band Pass, N=127 (C & M)  
Phase Error (rad) vs Frequency

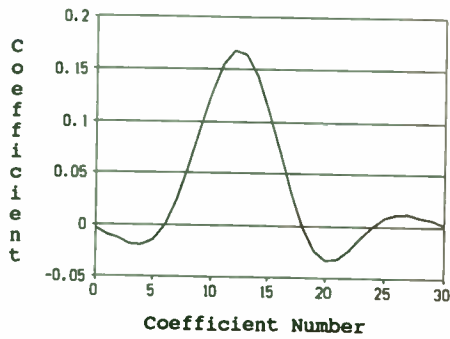


Figure 15: NLP Low Pass, N=30 (C & M)  
Filter Coefficients

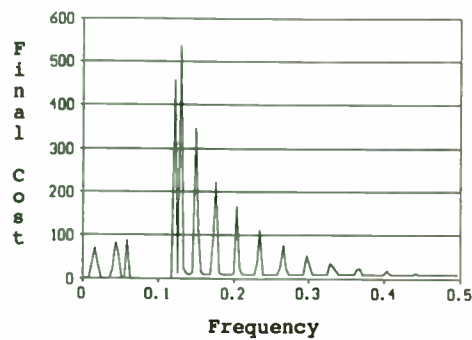


Figure 16: NLP Low Pass, N=30 (C & M)  
Final Adjusted Costs vs Frequency

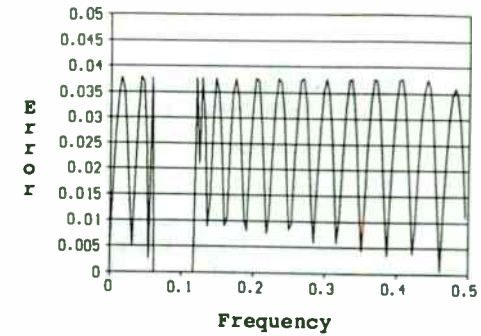


Figure 17: NLP Low Pass, N=30 (C & M)  
Magnitude of Complex Error vs Frequency

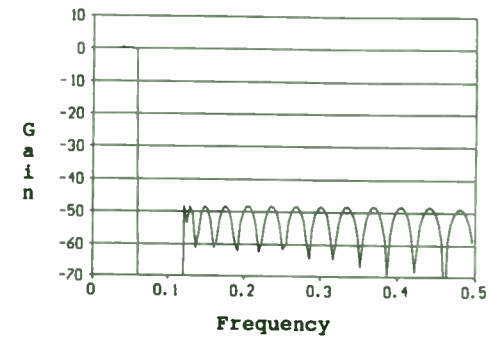


Figure 18: NLP Low Pass, N=30 (C & M)  
Gain (dB) vs Frequency

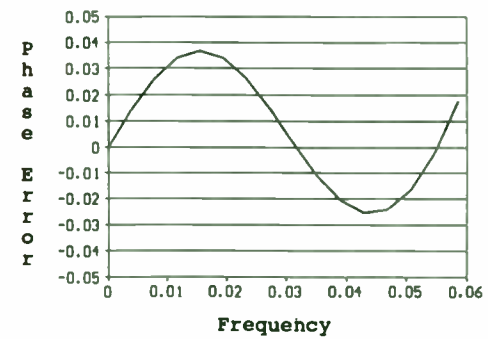


Figure 19: NLP Low Pass, N=30 (C & M)  
Phase Error (rad) vs Frequency



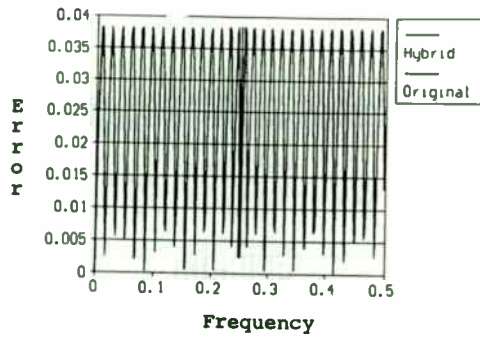


Figure 20: Twin Delay, N=60 (Comparison)  
Magnitude of Complex Error vs Frequency

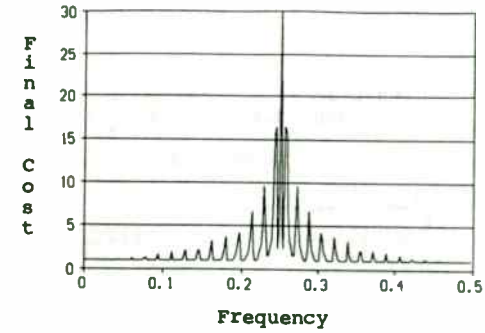


Figure 23: Twin Delay, N=60  
Final Adjusted Costs: Original

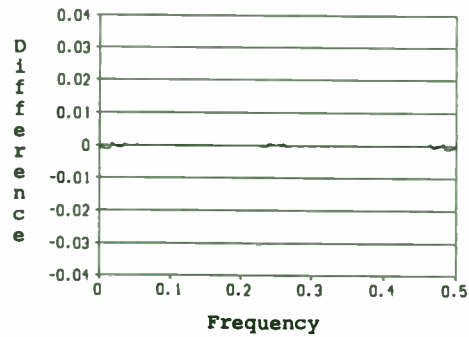


Figure 21: Twin Delay, N=60 (Comparison)  
Original Error Minus Hybrid Error

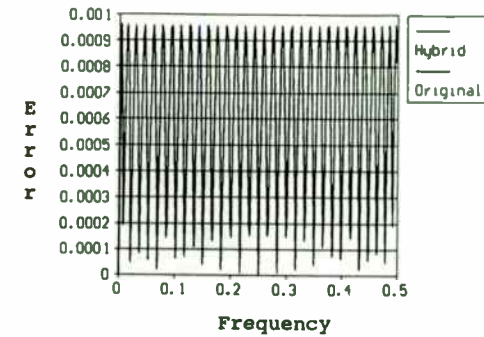


Figure 24: Sine Delay, N=60 (Comparison)  
Magnitude of Complex Error vs Frequency

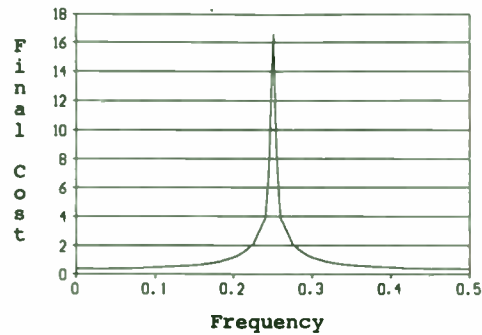


Figure 22: Twin Delay, N=60  
Final Adjusted Costs: Hybrid

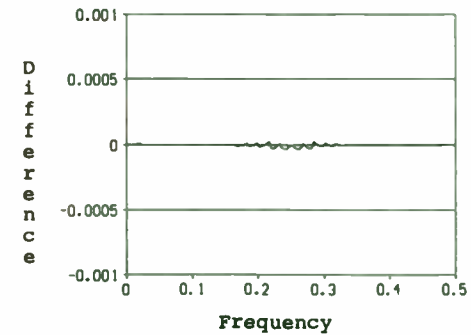


Figure 25: Sine Delay, N=60 (Comparison)  
Original Error Minus Hybrid Error

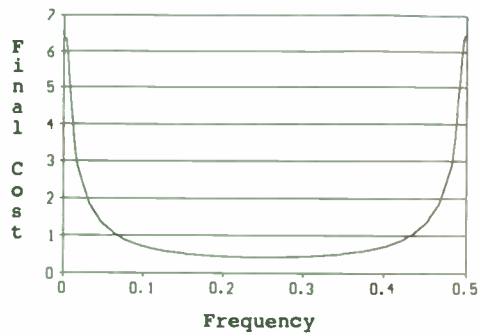


Figure 26: Sine Delay, N=60  
Final Adjusted Costs: Hybrid

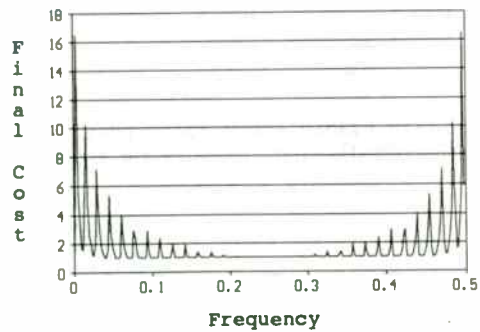


Figure 27: Sine Delay, N=60  
Final Adjusted Costs: Original

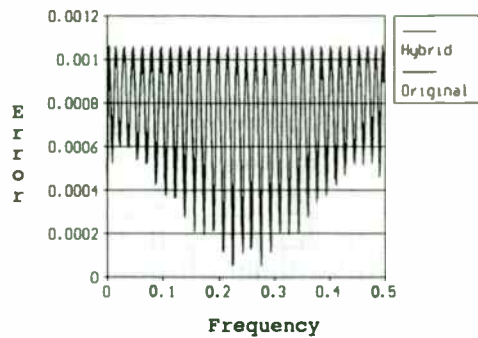


Figure 28: Chirp, N=60 (Comparison)  
Magnitude of Complex Error vs Frequency

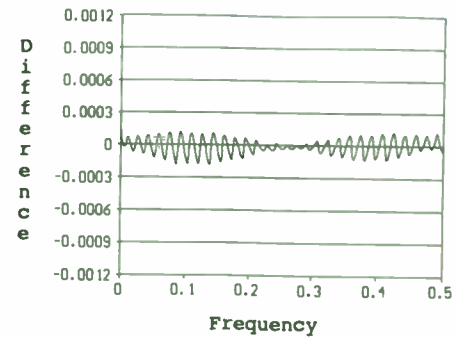


Figure 29: Chirp, N=60 (Comparison)  
Original Error Minus Hybrid Error

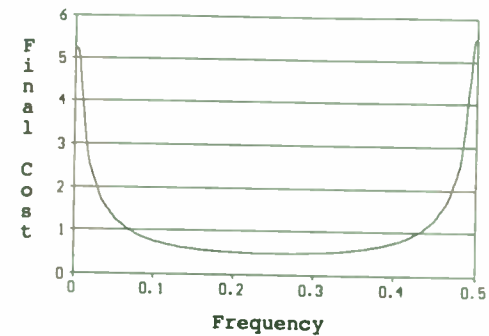


Figure 30: Chirp, N=60  
Final Adjusted Costs: Hybrid

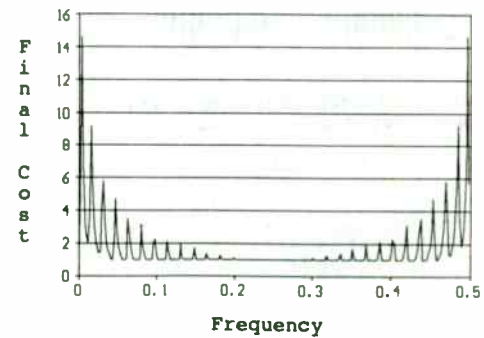


Figure 31: Chirp, N=60  
Final Adjusted Costs: Original

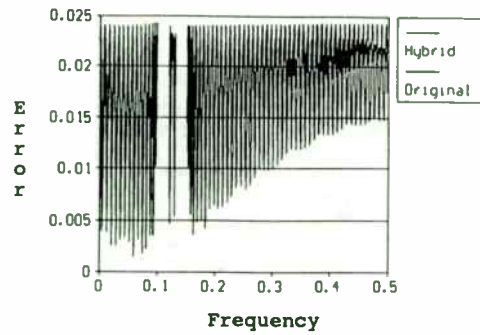


Figure 32: NLP Band Pass, N=127  
(Comparison)  
Magnitude of Complex Error vs Frequency

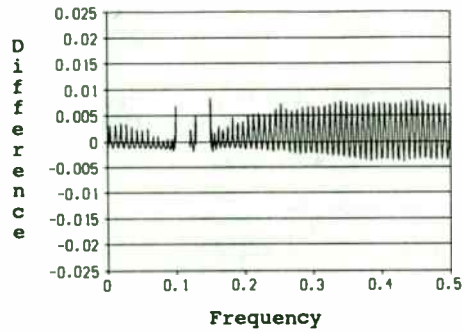


Figure 33: NLP Band Pass, N=127  
(Comparison)  
Original Error Minus Hybrid Error

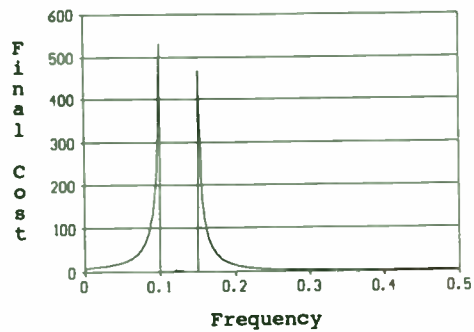


Figure 34: NLP Band Pass, N=127  
Final Adjusted Costs: Hybrid

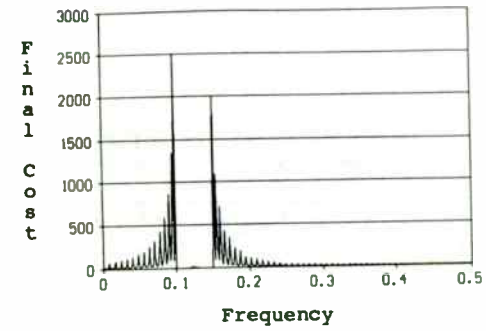


Figure 35: NLP Band Pass, N=127  
Final Adjusted Costs: Original

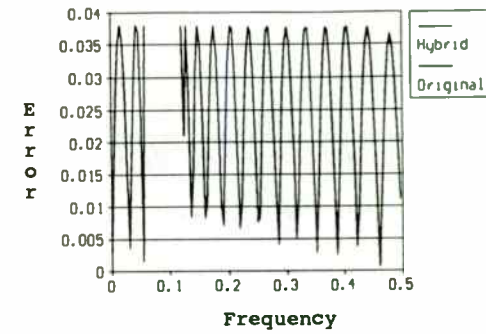


Figure 36: NLP Low Pass, N=30  
(Comparison)  
Magnitude of Complex Error vs Frequency

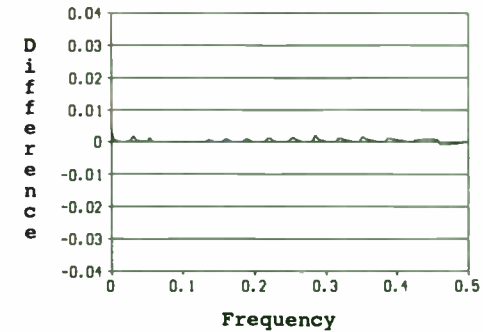


Figure 37: NLP Low Pass, N=30  
(Comparison)  
Original Error Minus Hybrid Error

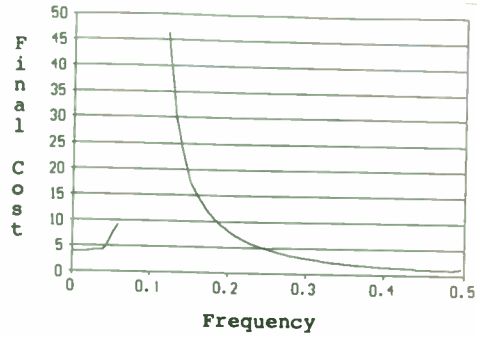


Figure 38: NLP Low Pass, N=30  
Final Adjusted Costs: Hybrid

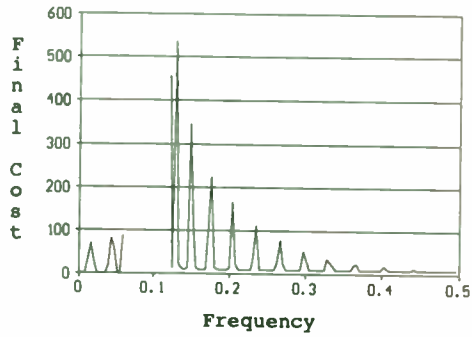


Figure 39: NLP Low Pass, N=30  
Final Adjusted Costs: Original

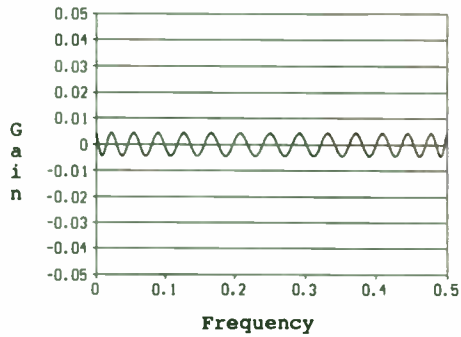


Figure 41: Sine Delay, N=1200 (Hybrid)  
Gain (dB) vs Frequency

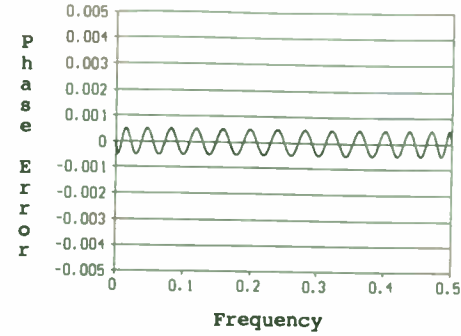


Figure 42: Sine Delay, N=1200 (Hybrid)  
Phase Error (rad) vs Frequency

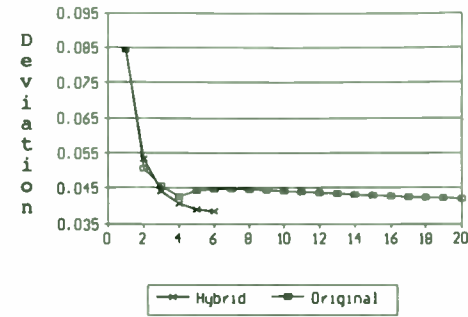


Figure 43: Twin Delay, N=60  
Peak Deviation vs Iteration Number

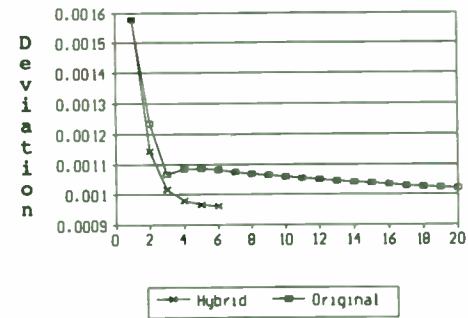


Figure 44: Sine Delay, N=60  
Peak Deviation vs Iteration Number

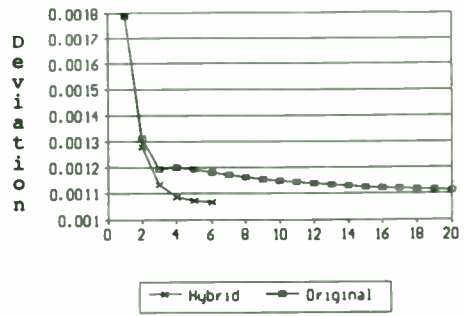


Figure 45: Chirp, N=60  
Peak Deviation vs Iteration Number

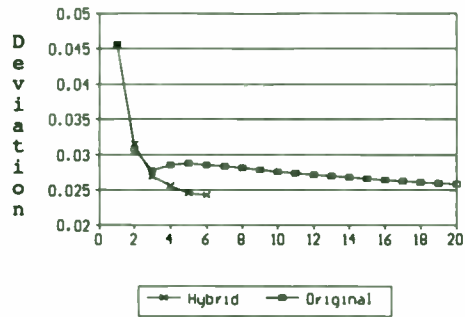


Figure 46: NLP Band Pass, N=127  
Peak Deviation vs Iteration Number

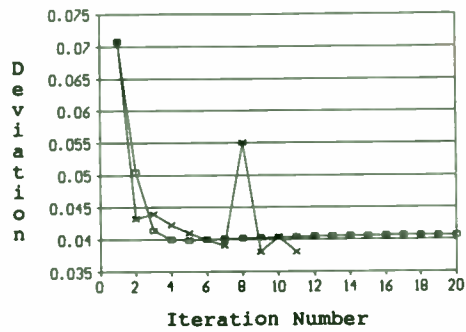


Figure 47: NLP Low Pass, N=30  
Peak Deviation vs Iteration Number



## Image Compression Reduces Bandwidth Requirements

Michael G. Ellis  
Research Electrical Engineer  
U.S. Army Corps. of Engineers  
Waterways Experiment Station (WES)  
3909 Halls Ferry Rd.  
Vicksburg, MS 39180-6199

**Abstract:** This article reviews the standard lossless and lossy techniques that comprise image compression algorithms. The term "lossless" means that the compressed file can be reconstructed exactly while "lossy" means that there is some loss of information in the reconstructed image. The most important components of the Joint Photographic Experts Group (JPEG) algorithm are targeted including Huffman encoding and the Cosine Transform. LZW, Singular Value Decomposition, fractal compression, and other techniques are also discussed. The feasibility of using these methods for bandwidth reduction is analyzed relative to several High Definition TV (HDTV) standards.

### Section I. Lossless Coding Techniques

#### 1. Entropy Coding

The term "entropy" is borrowed from thermodynamics and has a similar meaning.

The higher the entropy of a message, the more information it contains. The entropy of a data file is defined as the information per byte for that particular file. If the 8 bit ASCII character set that makes up a data or image file is not uniformly distributed, then the entropy will be less than 8 bits/byte. In entropy coding, the goal is to encode a block of M pixels containing M\*B bits, where B represents the number of bits per pixel in the raw data field. The entropy for any general file is given as

$$H = - \sum_{i=1}^M [ p_i \log_2(p_i) ] \quad (1)$$

In an 8 bit character set, there are 256 values that can be represented by a single byte. The probability,  $p_i$ , is the total occurrences of the  $i$ th byte divided by the total number of bytes in the data file.

The program, ENTROPY.BAS, uses equation (1) to estimate the amount of compression achievable by entropy coding. ENTROPY.BAS works only under QuickBasic 4.5 because of the "open binary" statement in line 70. A C language version, ENTROPY.C, is provided to illustrate the syntax differences between BASIC and C. Most compression code is written in C since the more optimized C compiler produces executables which are ten times faster than compiled BASIC.

#### 2. Huffman Compression

This classic compression technique, introduced in 1952 by David A. Huffman, is perhaps the most well known technique for entropy compression. It achieves the

minimum amount of redundancy possible in a fixed set of variable-length codes.

This does not mean that Huffman coding is an optimal coding method. It means that it provides the best approximation for coding symbols when using fixed-width codes.

The entropy of a data file is given by equation (1). In general,  $\log_2(p_i)$  will not be an integer so that the achieved data rate exceeds H bits/pixel; however the data rate can be made to asymptotically approach H with increasing block size.

Huffman coding operates similar to the Morse code in which the most frequently used symbols are given the shortest code. However, it is not obvious how to optimally assign a set of varying length codes to a set of symbols to be transmitted. Huffman coding optimizes by assigning a varying length code to a set of symbols based on the probability of occurrence. This does not exclude the possibility of other lossless techniques producing a higher compression rate.

Suppose our total set of characters is represented by eight ASCII values with probabilities in Table 1. Eight ASCII values can be represented by 3 bits/pixel; however, for the given probabilities of occurrence, the entropy is only 2.6984 bits/pixel. Therefore, Huffman coding can provide a compression scheme with a bit rate approaching 2.6984 bits/pixel.

The Huffman code is a variable length code in which no code can be a prefix for any other Huffman code. This allows unambiguous decoding of the transmitted data stream.

Table 1

Sample 3-bit character set with Probabilities of Occurrence

Character #	Probability	Entropy (H)
1	.30	.521089 = $-0.3 \log_2(0.3)$
2	.23	.487667 = $-0.23 \log_2(0.23)$
3	.15	.410544 = $-0.15 \log_2(0.15)$
4	.08	.291508 = $-0.08 \log_2(0.08)$
5	.06	.243533 = $-0.06 \log_2(0.06)$
6	.06	.243533 = $-0.06 \log_2(0.06)$
7	.06	.243533 = $-0.06 \log_2(0.06)$
8	.06	.243533 = $-0.06 \log_2(0.06)$
	1.0	2.6984 bits/pixel = total entropy

The Huffman codebook that is generated by the tree algorithm must be transmitted prior to the actual data in order to set up the decoder. Programs for Huffman coding require two passes over the data. Probabilities are generated in the first pass and the actual Huffman compressed file is generated during the second pass.

#### The Huffman Coding Algorithm

1. Arrange the symbol probabilities  $P_i$  in a decreasing order, and consider them as leaf nodes of a tree.
2. While there is more than one node:
  - Merge the two nodes with smallest probability to form a new node whose probability is the sum of the two merged nodes.
  - Arbitrarily assign 1 and 0 to each pair of branches merging into a node.
3. Read sequentially from the root node to the leaf node where the symbol is located.



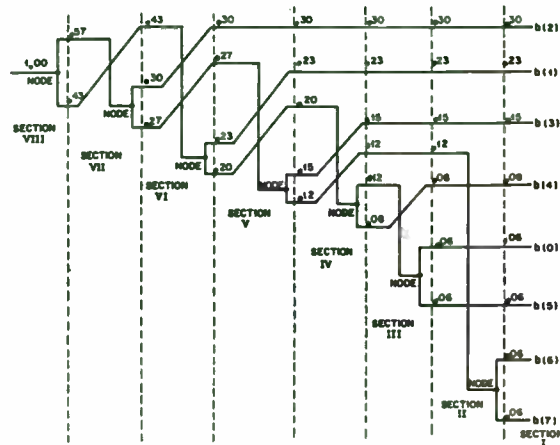


Figure 1. Construction of a binary Huffman code proceeds from right to left. At each section the two bottom-most branches are combined to form a node and followed by a reordering of probabilities into descending order. These probabilities are then used to start the next section.

The actual achieved compressed data rate is 2.71 bits/pixel and can be determined as follows.

0.30	*	2 bits =	0.60
0.23	*	2 bits =	0.46
0.15	*	3 bits =	0.45
0.08	*	3 bits =	0.24
0.06	*	4 bits =	0.24
0.06	*	4 bits =	0.24
0.06	*	4 bits =	0.24
0.06	*	4 bits =	0.24

Total = 2.71 bits/symbol

The coding sequence from Figure 2 is

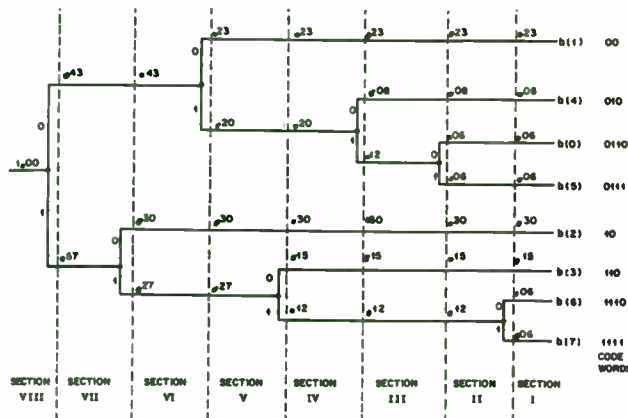


Figure 2. After the code construction, the tree is rearranged to eliminate crossovers, and coding proceeds from left to right. At each node a step-up produces a zero and a step-down a one. Resulting code words are shown at the right for each of the eight levels. Note that no code word is a prefix of any other code word.

Char #	Probability	Huffman Codeword
1	0.30	10
2	0.23	00
3	0.15	110
4	0.08	010
5	0.06	0110
6	0.06	0111
7	0.06	1110
8	0.06	1111

In high performance data compression, Huffman coding faces a significant problem. The program has to pass a complete copy of the coding statistics to the expansion program. For a Huffman Order-0 encoder, there are 8 bits per byte and

the probability table that is passed to the decoder may occupy as little as 256 bytes. However, a Huffman Order-0 encoder sometimes cannot achieve compression rates close to the entropy given in equation (1). A Huffman Order-1 encoder groups the symbols into 16 bit words, but boosts the statistics table from 256 to 65,536 bytes. Though compression ratios will undoubtedly improve when moving to order-1, the overhead of passing the statistics table will probably nullify any gains.

Adaptive Huffman coding solves this situation without paying any penalty for added statistics. It does so by adjusting the Huffman tree based on data previously seen, with no knowledge about future statistics. When using an adaptive model, data does not have to be scanned once in order to generate statistics. Instead, the statistics are continually modified as new characters are read in and coded. A more detailed discussion of Adaptive Huffman coding can be obtained in [1].

### 3. Arithmetic Coding

Arithmetic Coding is a lossless technique that can compress below the entropy level. It generally provides 10% to 20% better compression than Huffman. The output from an arithmetic process is a single number less than 1 and greater than, or equal to, 0. This single number can be uniquely decoded to exactly recreate the original byte stream.

To construct the output number, the symbols are assigned a set of probabilities. The message "KING HENRY" would have a probability distribution as follows:

<u>Character</u>	<u>Probability</u>
Space	1/10
E	1/10
G	1/10
H	1/10
I	1/10
K	1/10
N	2/10
R	1/10
Y	1/10

Individual characters can now be assigned along a probability line as follows.

<u>Character</u>	<u>Probability</u>	<u>Range</u>
Space	1/10	$0.00 \geq r > 0.10$
E	1/10	$0.10 \geq r > 0.20$
G	1/10	$0.20 \geq r > 0.30$
H	1/10	$0.30 \geq r > 0.40$
I	1/10	$0.40 \geq r > 0.50$
K	1/10	$0.50 \geq r > 0.60$
N	2/10	$0.60 \geq r > 0.80$
R	1/10	$0.80 \geq r > 0.90$
Y	1/10	$0.90 \geq r \geq 1.00$

The most significant portion of an arithmetic coded message belongs to the first symbol, or "K", in the message "King Henry". To decode the first character properly, the final coded message has to be a number greater than or equal to 0.5, and less than 0.6. Therefore the low end for this range is 0.5, and the high end is 0.6.

During the rest of the encoding process, each new symbol will further restrict the possible range of the output number. The next character to be encoded, the letter "I", has a the range from 0.4 to 0.5 in the subrange 0.5 to 0.6. Applying this logic will further restrict the number to 0.54 to 0.55. Continuing this process results in the final low value of 0.5464063556, which will uniquely decode into "KING HENRY".

<u>New Character</u>	<u>Low Value</u>	<u>High Value</u>
K	0.5	0.6
I	0.54	0.55
N	0.546	0.548
G	0.5464	0.5466
Space	0.54640	0.54642
H	0.546406	0.546408
E	0.5464062	0.5464064
N	0.54640632	0.54640636
R	0.546406352	0.546406356
Y	0.5464063556	0.546406356

Arithmetic coding is best accomplished using 32 bit binary integral math. The decimal point is implied at the left side of the word. The initial value of HIGH is \$FFFFFFFF, and LOW is \$00000000. Decoding is accomplished by first transmitting a table of statistics to the decoder, and then by transmitting a series of integers to be decoded. Typically, arithmetic coding is not used in shareware since IBM holds the patent for this technique.

#### 4. Lempel, Ziv, Welch (LZW) Compression

LZW is related to two compression techniques known as LZ77 and LZ78. LZ77 is a "sliding window" process in which the dictionary consists of a set of fixed-length phrases found in a "window" in the previously processed text. The size of the window is generally somewhere between 2K and 16K bytes, with the maximum phrase length ranging from perhaps 16 to 64 bytes. LZ78 takes a completely different approach to building a dictionary. Instead of using fixed-length phrases from a window in the text, LZ78 builds phrases up one symbol at a time, adding a new symbol to an existing phrase when a match occurs.

LZ77 has a major performance bottleneck. When encoding, it has to perform string comparisons against the look-ahead buffer for every position in the text window. As LZ77 tries to improve compression performance by increasing the size of the window, (and thus the dictionary), this performance bottleneck only gets worse. LZSS seeks to avoid some of the performance problems in the LZ77

algorithm. Under LZ77, the phrases in the text window were stored as a single contiguous block of text, with no other organization on top of it. LZSS still stores text in contiguous windows, but it creates an additional data structure that improves on the organization of the phrases. As each phrase passes out of the look-ahead buffer and into the encoded portion of the test windows, LZSS adds the phrase to a tree structure. The savings created by using the tree not only makes the compression side of the algorithm much more efficient, it also encourages experimentation with longer window sizes. Doubling the size of the text window now might only cause a small increase in the compression time; whereas, before it would have doubled it.

LZ78 is similar to LZ77 but abandons the concept of a text window. Under LZ77 the dictionary of phrases was defined by a fixed window of previously seen text. With LZ78, the dictionary is a potentially unlimited list of previously seen phrases.

LZW is an improved LZ78 algorithm in which the phrase dictionary is preloaded with single-symbol phrases equal to the number of symbols in the alphabet. LZW never outputs single characters, only phrases.

LZW compression is shown using " BET BE BEE BED BEG" as the input stream. The quotation marks are used only to indicate that the character stream begins with a "space". The first 256 codes (0 to 255) are not shown since the following table is generated by LZW as an appendix to the first 256 characters.

Step;	Character In;	Code Output;	New Code
1	Space	none	[already in table]
2	B	ASCII for space	space B at (256)
3	E	ASCII for B	BE at (257)
4	T	ASCII for E	ET at (258)
5	Space	ASCII for T	T space at (259)
6	B	none	[space B already in table]
7	E	256	space BE at (260)
8	Space	ASCII for E	E space at (261)
9	B	none	[space B already in table]
10	E	none	[space BE already in table]
11	E	260	BEE space (262)
12	Space	none	[E space already in table]
13	B	261	E space B at (263)
14	E	none	[BE already in table]
15	D	257	BED at (264)
16	space	D	D space at (265)
17	B	none	[space B already in table]
18	E	none	[space BE already in table]
19	T	260	BET at (266)
20	<EOF>	ASCII for T	

The first seven steps are explained as follows.

Step 1: A "space" is read. The ASCII value for "space" is already contained in the table. No action is taken.

Step 2: A "B" is read. There is no "space B" combination in the table. The ASCII value for "space" is output and "space B" is assigned code 256.

Step 3: An "E" is read. There is no "BE" combination in the table, so "BE" is assigned code 257 and the ASCII value for "B" is output.

Step 4: A "T" is read. There is no "ET" combination in the table, so "ET" is assigned code 258 and the ASCII value for "E" is output.

Step 5: A "space" is read. There is no "T space" combination in the table, so "T space" is assigned code 259 and the ASCII value for "T" is output.

Step 6: A "B" is read. The combination "space B" already exists in the table so no action is required.

Step 7: An "E" is read. The combination "space BE" is assigned code 260, and code 256 for "space B"

The LZW decompressor takes the stream of codes from the compression algorithm and uses them to recreate the exact input stream. One reason for the efficiency of the LZW algorithm is that it does not need to pass the dictionary to the decompressor since the table can be built exactly as it was during compression, using the input stream as data. This is possible because the compression algorithm always outputs the phrase and character components of a code before it uses it in the output stream, so the compressed data is not burdened with carrying a large dictionary. Typically 12-bit codes words will be used in an LZW algorithm for 4096 possible phrases to accommodate the standard 256 8-bit character set plus the additional phrases that are constructed as the input stream is processed.

The patent for LZW is assigned to Unisys, which has made public its intention to protect its intellectual property rights. LZW compression is defined as part of the CCITT V.42bis specification, and Unisys has defined specific terms under which it

will license the algorithm to modem manufacturers. It has not stated that it will apply the same terms to any parties manufacturing other types of products.

#### Lossless Compression Tests Results

Tests were run on six lossless compression routines at WES. The results of the test are given below. The compression time and performance is given for each file and package. The first 3 files are executable programs. The next two are text files. JBARB2Y.COL and JGOLDY.COL are 8-bit grayscale images. The last four files are 8-bit RGB color images. The source and compiled code for the following programs is available from Reference [6].

ARC, ZIP, and COMPRESS implement variations of the LZW algorithm. The highest performance package (in terms of compression) is LHARC by Haruyasu Yoshizaki which is built on an LZ77 algorithm using a dictionary based method on a sliding window that moves through the text. All of these packages are compared against standard entropy coding (Table 2) which consistently provided the worst compression.

## Section II. Lossy Compression Techniques

### 1. The Fourier Transform

Lossy compression methods are useful for image compression and are often based on the Fourier transform. In general, a two dimensional Fourier transform of an

Table 2

FILE	ARC	LHARC	ZOO	ZIP	COMPRESS	ENTROPY
WP.EXE	2.55s	6.35s	5.11s	3.53s	23.55s	
244736	222,275	175,828	223,621	176,564	230,025	222,383
100%	90.8%	71.8%	91.4%	72.1%	94.0%	90.8%
WS.EXE	1.78s	2.40s	2.20s	1.34s	5.84s	
78336	69,158	60,704	75,191	61,982	74,007	69,214
100%	88.3%	77.5%	96.0%	79.1%	94.5%	88.4%
PROCOMM.EXE	1.85s	4.60s	2.78s	2.90s	11.30s	
165,296	104,312	82,072	103,510	80,782	105,895	129,729
100%	63.1%	49.7%	62.6%	48.9%	64.1%	78.5%
SILVERD.DOC	7.18s	35.17s	11.29s	24.71s	62.44s	
1,096,064	403,771	365,080	412,111	359,696	363,231	586,292
100%	36.8%	33.3%	37.6%	32.8%	33.1%	53.5%
REGISTER.TXT	0.66s	0.66s	0.69s	0.55s	1.00s	
6801	3478	2903	3645	2946	3476	4004
100%	51.1%	42.7%	53.6%	43.2%	51.1%	58.9%
JBARB2Y.COL	6.40s	11.71s	10.59s	5.92s	38.78s	
403200	377,351	352,790	412,625	373,014	381,029	377,173
100%	93.6%	87.5%	102%	92.5%	94.5%	93.5%
JGOLDY.COL	4.40s	11.03s	7.53s	6.40s	36.9s	
403200	373,323	339,233	372,283	350,122	350,123	380,240
100%	92.5%	84.1%	92.3%	86.8%	86.8%	94.3%
BARB.RGB	3.21s	13.15s	4.28s	9.54s	25.6s	
403,200	150,193	152,410	152,041	165,471	146,998	248,399
100%	37.3%	37.8%	37.7%	41.0%	36.5%	61.6%
GOLD.RGB	2.73s	14.78s	3.84s	13.73s	20.72s	
403,200	122,381	124,775	119,857	135,986	119,843	235,154
100%	30.3%	30.9%	29.7%	33.7%	29.7%	58.3%
BOATS.RGB	2.65s	19.79s	3.53s	14.87s	17.92s	
403,200	94,844	98,351	95,231	106,958	96,405	204,358
100%	23.5%	24.4%	23.6%	26.5%	23.9%	50.6%
BOARD.RGB	2.67s	15.93s	3.59s	14.29s	19.09s	
403,200	107,201	108,918	107,458	118,593	105,775	218,287
100%	26.6%	27.0%	26.7%	29.4%	26.2%	54.1%

Compression time is shown in seconds for a IBM compatible 486/33MHz PC. The next two rows show the file size after compression and the size of the compressed file as a percentage of the original.

image will produce a large number of coefficients close to zero. For example, if 95% of the transform coefficients can be approximated by zero, then the image may be adequately represented using only 5% of the Fourier coefficients. The sparse matrix that results is usually Huffman encoded. Although LZW and Arithmetic codes generally provide more compression than Huffman, the latter tends to work better on sparse data files that contain a large number of common values.

A two dimensional Fourier Transform is required for image files since images are two dimensional. The 2-D transform is equivalent to taking the one dimensional transform of each row, and then taking the one dimensional transform of each column, for an image matrix. The problem with the Fourier Transform is that it leaves us with a real component, and an imaginary component, for each pixel in the image. By symmetrically extending the image before taking the 2-D Fourier transform, the imaginary, or sine, terms can be forced to zero.

Assume that an image is represented by the 4x4 matrix, U, such that

$$[U] = \begin{bmatrix} 127 & 123 & 119 & 110 \\ 115 & 111 & 103 & 98 \\ 99 & 87 & 85 & 83 \\ 94 & 82 & 81 & 79 \end{bmatrix}$$

then the two-dimensional Fourier Transform will produce 32 terms, or a real term (cosine) and an imaginary term (sine) for each of the 16 pixel values in the matrix. By symmetrically extending U, either the cosine terms, or the sine terms, can be forced to be identically zero.

A symmetrical extension of U given by

0	0	0	0	0	0	0	0	0
0	127	123	119	110	119	123	127	0
0	115	111	103	98	103	111	115	0
0	99	87	85	83	85	87	99	0
0	94	82	81	79	81	82	94	0
0	99	87	85	83	85	87	99	0
0	115	111	103	98	103	111	115	0
0	127	123	119	110	119	123	127	0

will produce only non-zero cosine coefficients under a 2-D Fourier Transform.

This modification of the Fourier Transform is called a Cosine Transform and allows the matrix U to be fully represented by 16 unique coefficients, most of them close to zero. If a 2-D Fast Fourier Transform is performed on the symmetrically extended matrix, the results are equivalent to a Fast Cosine Transform. It is important to note that the coefficients produced by a Cosine Transform are not the same as the real part of the Fourier Transform applied to the original matrix, U.

An antisymmetric extension of U given by

0	0	0	0	0	0	0	0	0
0	127	123	119	110	-127	-123	-119	-110
0	115	111	103	98	-115	-111	-103	-98
0	99	87	85	83	-99	-87	-85	-83
0	94	82	81	79	-94	-82	-81	-79
0	-127	-123	-119	-110	127	123	119	110
0	-115	-111	-103	-98	115	111	103	98
0	-99	-87	-85	-83	99	87	85	83
0	-94	-82	-81	-79	94	82	81	79

will produce only non-zero sine coefficients under a 2-D Fourier Transform. This modification of the Fourier Transform is called a Sine Transform.

## 2. Representation of 8-Bit Images

The images used in this article are raw 8-bit grayscale images. The word "raw" is used to enforce the idea that there is no header on the file. The test image used for lossy compression, DBSJ.4, contains 448 columns and 280 rows for a total of 125,540 bytes.

Each byte controls the grayscale intensity of a pixel on the screen with 0 representing black, and 255 representing white for a total of 256 shades of gray.

A separate palette file controls whether the image is grayscale or color. The palette files used with the program "Image" consist of 768 bytes, or 256 bytes for the color red, 256 bytes for green, and 256 bytes for blue. If a grayscale image is desired, then the palette "gray.pal" is used and includes identical 256 byte segments for red, green, and blue. The actual computation of the color of each pixel in an RGB image can be seen easily in line 380. For color images, bits 5, 6, and 7 determine the intensity of the red component, bits 2, 3, and 4 define the green intensity, and bits 0 and 1 provide four levels of blue. The palette file defines the mapping of the red, green, and blue bits to various intensity levels. VGA mode 13 is set up in line 220 for 320x200 resolution with 256 colors, or shades of gray, depending on the palette that is selected. The program, IMAGE, is useful for displaying an image prior to compression, and then showing the degradation caused by a lossy compression/decompression routine.

While the program "Image" can let the user visually determine the quality of a reconstructed image by comparison to the original, the program "MSE" computes the mean-square-error between the original image and the reconstructed image to provide an analytical measure of quality. The mean-square-error is defined as

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - x_i')^2 \quad (2)$$

where  $X_i$  is the  $i$ th pixel in the original image,  $X_i'$  is the  $i$ th pixel in the reconstructed image, and  $N$  is the number of pixels. In general, a MSE of 5 or less indicates very little loss of detail.

#### 1. The Discrete Cosine Transform (DCT)

The DCT is simply a separate mathematical method for generating the Cosine Transform without explicitly using the Fourier Transform. It is based on defining a matrix [C] such that

$$C(k,n) = 1/\sqrt{N}, \quad \text{for } k,n = 0$$

$$C(k,n) = \sqrt{2/N} * \cos(3.14159*(2*k+1)*n/(2*N)), \text{ otherwise.}$$

If an image is divided into blocks of size 8x8, then the matrix, C, must also be 8x8 and becomes

$$C = \begin{bmatrix} .35355 & .35355 & .35355 & .35355 & .35355 & .35355 & .35355 & .35355 \\ .49039 & .41573 & .27778 & .0975 & -.0975 & -.27778 & -.41573 & -.49039 \\ .46193 & .19134 & -.19134 & -.46193 & -.46194 & -.19134 & .19133 & .46193 \\ .41573 & -.09754 & -.49039 & -.27778 & .27778 & .49039 & .09754 & -.41573 \\ .35355 & -.35355 & -.35355 & .35355 & .35355 & -.35355 & -.35355 & .35355 \\ .27778 & -.49039 & .09754 & .41573 & -.41573 & -.09754 & .49039 & -.27778 \\ .19134 & -.46194 & .46193 & -.19133 & -.19134 & .46194 & -.46193 & .19133 \\ .09754 & -.27778 & .41573 & -.49039 & .49039 & -.41573 & .27777 & -.09754 \end{bmatrix}$$



If a matrix  $U$  represents the pixels in an 8x8 block of an image, then the DCT is defined by the equation

$$[V] = [C] [U] [C^T] \quad (3)$$

using matrix multiplication and  $[C^T]$  as the transpose of matrix  $[C]$ .  $[V]$  becomes a sparse matrix with few large coefficients. For example,

$$\text{let } [U] = \begin{bmatrix} 48 & 48 & 51 & 53 & 54 & 54 & 57 & 62 \\ 50 & 51 & 52 & 52 & 57 & 59 & 59 & 63 \\ 51 & 52 & 55 & 56 & 59 & 61 & 62 & 67 \\ 53 & 55 & 57 & 57 & 62 & 64 & 63 & 66 \\ 56 & 58 & 57 & 60 & 65 & 67 & 68 & 68 \\ 56 & 59 & 61 & 63 & 66 & 66 & 70 & 72 \\ 60 & 61 & 64 & 65 & 67 & 70 & 72 & 72 \\ 62 & 62 & 65 & 70 & 69 & 70 & 73 & 76 \end{bmatrix}$$

which represents the first 8x8 subblock in the grayscale test file DBSJ.4. The actual transform file,  $V$ , is rounded to integer values and becomes

$$[V] = \begin{bmatrix} 61 & -5 & 0 & 0 & 0 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

which contains a large number of zeroes, but retains almost all of the information in the original subimage  $[U]$ . Entropy analysis of the original matrix  $[U]$  reveals that a compression of 1.77 to 1 is possible using Huffman encoding (as shown by running ENTROPY.EXE). However, a compression of 25.6 to 1 is possible using Huffman encoding of the matrix  $[V]$ . This is the basis of the Joint Photographic Expert Group (JPEG) algorithm in which images are divided into 8x8 subblocks; each subblock is processed independently by a DCT algorithm, and then the resulting file is Huffman encoded.

The inverse DCT can be used to regenerate an approximation to the matrix  $[U]$  using the equation

$$[U'] = [C^T] [V] [C] \quad (4)$$

where  $[U']$  represents the reconstructed image. Using the inverse DCT, the matrix  $[U']$  becomes

$$[U'] = \begin{bmatrix} 47 & 48 & 50 & 53 & 55 & 58 & 60 & 61 \\ 48 & 49 & 51 & 54 & 57 & 59 & 61 & 62 \\ 50 & 51 & 53 & 56 & 58 & 61 & 63 & 64 \\ 53 & 54 & 56 & 58 & 61 & 64 & 65 & 67 \\ 55 & 57 & 58 & 61 & 64 & 66 & 68 & 69 \\ 58 & 59 & 61 & 64 & 66 & 69 & 71 & 72 \\ 60 & 61 & 63 & 65 & 68 & 71 & 73 & 74 \\ 61 & 62 & 64 & 67 & 69 & 72 & 74 & 75 \end{bmatrix}$$

which is nearly identical to the original image [U] and differs by a mean-square-error of 1.85 indicating that visual differences will not be perceptible. A compression scheme using the DCT can be summarized in the following steps.

1. Divide the image into 8x8 blocks.
2. Perform the DCT on each 8x8 block.
3. Replace values close to zero with zero.
4. Huffman encode the resulting file.

Decompression involves the following.

1. Huffman decode the compressed file.
2. Perform the inverse DCT on each 8x8 block.

The entire DBSJ.4 image is shown in Figure 3 before and after compression, at a rate of 25:1. The programs DCT.BAS and INVDCT.BAS were used for the transformations. The final compression of 25:1 was achieved by truncating any DCT coefficients, less than 5, to zero and then entropy encoding of the remaining DCT coefficients.

The JPEG algorithm is basically an 8x8 DCT with some enhancements. JPEG is generally considered feasible for compression of still images in the 10:1 to 25:1

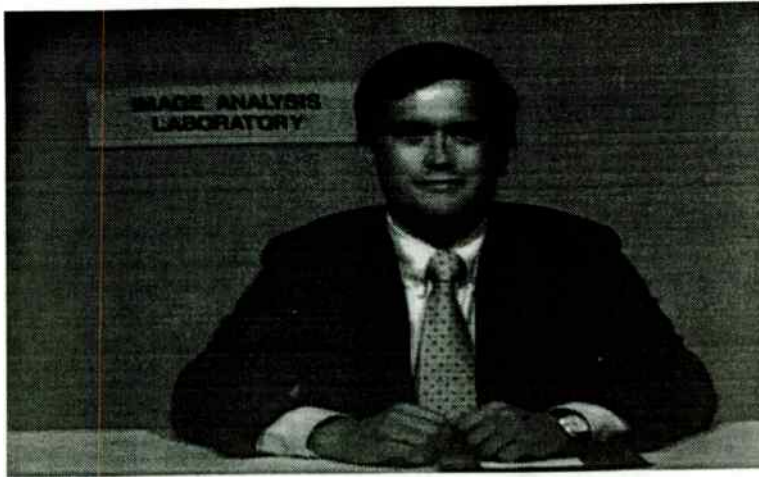
range. Another standard, MPEG, is used for motion video and exploits the images differences frame-by-frame. MPEG is considered feasible for the compression of real time video images at the rate of 50:1.

The program, DCT.BAS, performs the DCT on any grayscale image file. X and Y dimensions are required as input parameters. The output file is given the extension .DCT.

The .DCT file is a binary file. The user can perform Huffman encoding directly on this file to achieve a high level of compression. Even higher levels of compression can be obtained by defining a threshold level in the DCT file such that coefficients below the threshold are replaced by zero before Huffman encoding.

An inverse DCT can be obtained using INVDCT.BAS. This program is very similar to DCT.BAS. The major change is that the indices of matrix [C] are swapped in lines 630 and 730. Since matrix [C] is orthogonal, the transpose of [C] is identical to the inverse of [C]. The quality of the restoration depends on the threshold level set previously for the truncation of the DCT coefficients.

These two programs will also handle other types of transforms including Hadamard and Sine transforms simply by modifying the subroutine that defines the transform matrix [C]. The Hadamard and Sine transforms do not provide as much compression as the Cosine transform, so their discussion will be deferred to [9].



## JPEG and MPEG

While the JPEG algorithm was defined for still images and is based on the DCT, MPEG is meant primarily for motion video and obtains additional compression by exploiting interframe redundancy. Basically, if an  $N \times N$  pixel block changes very little in successive frames, then the DCT coefficients saved from a previous frame can be reused to represent the  $N^2$  pixel values. MPEG is the basis for many proposed HDTV systems.

### 2. Singular Value Decomposition (SVD)

Singular Value Decomposition (SVD) is useful mainly as a gauge against which to measure the performance of other transform techniques and in image restoration. SVD concentrates the maximum amount of energy in the fewest eigenvalues and is optimal in the least square sense.

Let  $[U]$  be an image matrix. The matrices  $[U][U^T]$  and  $[U^T][U]$  are nonnegative, symmetric, and have the identical eigenvalues,  $\{\lambda_m\}$ . Assuming that  $U$  is an  $N \times N$  matrix of  $N^2$  pixels, there are at most  $r \leq N$  nonzero eigenvalues. It is possible to find  $r$  orthogonal,  $N \times 1$  eigenvectors  $\{\Phi_m\}$  of  $[U^T][U]$  and  $r$  orthogonal  $N \times 1$

Figure 3. (a) Original Image (top), (b) Restored (bottom) after 25:1 compression.  
The MSE is 25.39 using a threshold level of 5 in the DCT coefficient matrix.

eigenvectors  $\{\Psi_m\}$  of  $[U][U^T]$ , that is

$$[U^T][U] \Phi_m = \lambda_m \Phi_m, \quad m = 1, \dots, r \quad (6)$$

$$[U][U^T] \Psi_m = \lambda_m \Psi_m, \quad m = 1, \dots, r \quad (7)$$

The matrix  $U$  has the representation

$$[U] = \Psi \Lambda^{0.5} \Phi^T = \sum_{m=1}^r \Psi_m \Phi_m^T \sqrt{\lambda_m} \quad (8)$$

where  $\Psi$  and  $\Phi$  are  $N \times r$  matrices whose  $m$ th column are the vectors  $\Psi_m$  and  $\Phi_m$ ,

respectively, and  $\Lambda^{0.5}$  is an  $r \times r$  diagonal matrix, defined as

$$\Lambda^{0.5} = \begin{bmatrix} \sqrt{\Gamma_1} & & \\ & \ddots & \\ & & \sqrt{\Gamma_m} \end{bmatrix} \quad (9)$$

Equation 8 is called the spectral representation, the outer product expansion, or the singular value decomposition (SVD) of  $[U]$ . The nonzero eigenvalues of  $[U^T][U]$ ,  $\Gamma_m$ , are also called the singular values of  $[U]$ . If  $r \ll N$ , then the image containing

$N^2$  samples can be represented by  $(N + N)r$  samples of the vectors

$$\{\lambda_m^{1/4} \Psi_m, \lambda_m^{1/4} \Phi_m; m = 1, \dots, r\}.$$

Since  $\Psi$  and  $\Phi$  have orthogonal columns, the SVD transform of the image  $U$  is defined as

$$[U] = \Psi \Lambda^{0.5} \Phi^T \quad (10)$$

which is a separable transform that diagonalizes the given image.

The image  $[U_k]$  generated by the partial sum

$$[U_k] = \sum_{m=1}^k \sqrt{\lambda_m} \Psi_m \Phi_m^T, \quad k \leq r \quad (11)$$

is the best least squares rank- $k$  approximation of  $[U]$  if  $\lambda_m$  are in decreasing order

of magnitude. For any  $k \leq r$ , the least squares error

$$\epsilon_k^2 = \sum_{m=1}^M \sum_{n=1}^N |u(m,n) - u_k(m,n)|^2, \quad k = 1, 2, \dots, r \quad (12)$$

reduces to

$$\epsilon_k^2 = \sum_{m=k+1}^r \Gamma_m \quad (13)$$

This means the energy concentrated in the transform coefficients is maximized by the SVD transform for the given image.

Example:

$$\text{Let } [U] = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}$$

The eigenvalues of  $[U^T][U]$  are found to be  $\lambda_1 = 18.06$ ,  $\lambda_2 = 1.94$ ,

which give  $r = 2$ , and the SVD transform of  $[U]$  is

$$\Lambda^{0.5} = \begin{bmatrix} 4.25 & 0 \\ 0 & 1.39 \end{bmatrix}$$

The eigenvectors are found to be

$$\Phi_1 = \begin{bmatrix} 0.5019 \\ 0.8649 \end{bmatrix} \quad \Phi_2 = \begin{bmatrix} 0.8649 \\ -0.5019 \end{bmatrix}$$

From above,  $\Psi_1$  is obtained via

$$\Psi_m = \frac{1}{\sqrt{\lambda_m}} [U] \Phi_m$$

to yield

$$[U_1] = \sqrt{\lambda_1} \Psi_1 \Phi_1^T = \begin{bmatrix} 1.120 & 1.94 \\ 0.953 & 1.62 \\ 1.549 & 2.70 \end{bmatrix}$$

as the best least squares rank-1 approximation of  $[U]$ . The energy concentrated in the  $K$  samples of SVD is greater than the energy concentrated in any  $K$  samples of the other transform methods.

Example: Let  $[U]$  be an  $N \times N$  images matrix. It is desired to use SVD to achieve a  $N:r$  compression using only  $r$  of the largest eigenvalues. The  $[U]$  matrix can then be approximated as

$$[U] = \begin{bmatrix} \psi_{11} & \dots & \psi_{1r} & 0 & \dots & 0 \\ \vdots & & \vdots & & & \vdots \\ \psi_{n1} & \dots & \psi_{nr} & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_r} \\ & & & & 0 \\ 0 & \dots & & & & 0 \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} \\ \vdots & \vdots & & \vdots \\ \phi_{r1} & \phi_{r2} & \dots & \phi_{rn} \\ 0 & \dots & & 0 \\ \vdots & & & \vdots \\ 0 & \dots & & 0 \end{bmatrix}$$

The total number of coefficients that must be encoded and transmitted are  $rxN$  coefficients from matrix  $\Psi$  plus  $rxN$  coefficients from matrix  $\Phi$  or  $2rxN$  coefficients total, which only achieves half of the desired compression. A small improvement can be obtained by recognizing that the columns in  $\Psi$  and the rows in  $\Phi$  are orthogonal. If  $\Psi$  has  $r$  orthogonal columns such that

$$\Psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \dots & \psi_{1r} & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & & \vdots \\ \psi_{n1} & \psi_{n2} & \dots & \psi_{nr} & 0 & \dots & 0 \end{bmatrix}$$

with  $r \times N$  non-zero elements, then the  $rN$  elements of  $\Psi$  can be reconstructed using the conditions for orthogonality from the matrix  $\Psi'$  with

$$\Psi' = \begin{bmatrix} 0 & \dots & \dots & \dots & 0 \\ \Psi_{11} & 0 & \dots & \dots & 0 \\ \dots & \Psi_{12} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \Psi_{n1} & \Psi_{n2} & \dots & \Psi_{nr-1} & 0 \end{bmatrix}$$

where  $\Psi'$  has  $rN - 1/2 r^2$  non-zero elements. The number of elements that are required to be transmitted are  $rN - 1/2 r^2$  for  $\Psi$  and  $rN - 1/2 r^2$  for  $\Phi$  or  $r(2N - r)$  coefficients total, as opposed to  $r \times N$ . Therefore a degradation factor,  $D$ , can be computed as

$$D = 2 - r', \text{ where } r' = r/N$$

and  $r'$  represents the percent compression. Therefore, if we try to achieve 90% compression with SVD on a data file of size  $N^2$  by using only 1/10th of the eigenvalues, the actual size of the data file that has to be transmitted is  $0.1 \times (2 - 0.1) \times N^2$ , or 19% of the original size so that a compression of only 81% is achieved.

Image compression using SVD is of limited use for three reasons including

- (1) the vast amount of computations required,
- (2) the degradation factor limits the actual compression to approximately 1/2 of the theoretical limit, and
- (3) each coefficient in the  $\Psi$  and  $\Phi$  matrices will require more than a single byte of accuracy.

Example:

If the SVD transform of an image matrix,  $[U]$ , produces eigenvalues that are zero, then by not transmitting these zero eigenvalues, and their corresponding zero rows and columns in  $\Psi$  and  $\Phi^T$ , respectively, lossless image compression can be achieved. If  $[U]$  is an  $256 \times 256$  matrix given by

$$[U] = \begin{bmatrix} 0, 1, 2, 3, 4, 5, \dots, 255 \\ 0, 1, 2, 3, 4, 5, \dots, 255 \\ 0, 1, 2, 3, 4, 5, \dots, 255 \\ \vdots \\ \vdots \\ \vdots \\ 0, 1, 2, 3, 4, 5, \dots, 255 \end{bmatrix}$$

then the SVD transform produces 255 eigenvalues that are zero and one eigenvalue that is non-zero. The (unnormalized)  $\Psi$  and  $\Phi^T$  matrices become

$$\Psi^T = [1, 1, 1, 1, \dots, 1]$$

and

$$\Phi = [0, 1, 2, 3, 4, 5, \dots, 255].$$

The original  $U$  matrix can be reconstructed by simple matrix multiplication of  $\Psi$  and  $\Phi^T$ .

All of the energy is compacted into a single eigenvalue for a theoretical possible compression of  $1/256$ , but notice that a total of 512 bytes must be transmitted in order to send the 256 element  $\Psi$  matrix and the 256 element  $\Phi$  matrix. Therefore, only a compression of  $1/128$  can be realized which is consistent with the degradation factor

$$D = 1/256 \times (2 - 1/256) \approx 2$$

with all the energy compressed into  $1/256$  of the eigenvalues. Since the other 255 eigenvalues are exactly 0, the entire 65536 element  $U$  matrix can be reconstructed exactly from 512 transmitted elements. Huffman coding would not be able to achieve any compression on matrix  $[U]$  since all of the possible ASCII values occur with equal probability.

Example:

DBSJ.4 represents a  $448 \times 280$  8-bit grayscale image. The SVD of DBSJ.4 was taken using MATLAB. Figure 4 represents the image quality for a compression ratio of 100:1.

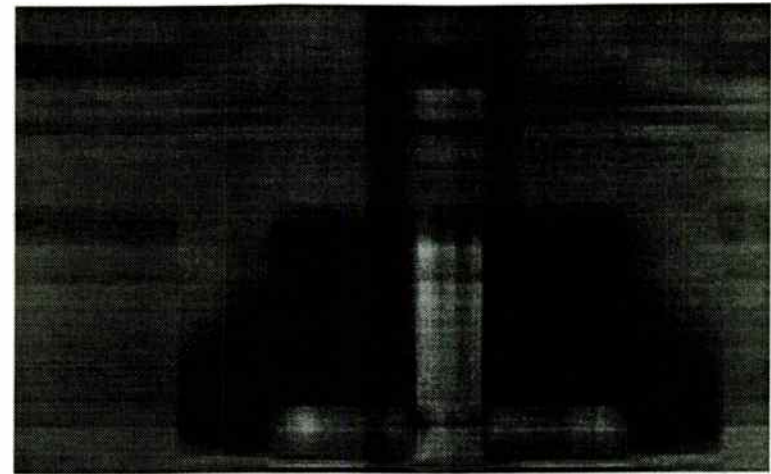


Figure 4. DBSJ.4 images compressed with SVD at 100:1 with a mean-square error of 553.

### 3. Fractal Compression

Fractal compression is a new technique that promises compression ratios of 500:1 without noticeable loss in image quality. It is based on taking a portion of the image, and reproducing the rest of the image with a set of affine transformations that consist of translations, rotations, and scaling. A set of iterated function system (IFS) codes define the image. Once the IFS codes for an image have been determined, then a simple iterated procedure is used to regenerate the image.

The iterated function is defined by the IFS codes. Essentially,  $X$  and  $Y$  coordinates of the new pixel are computed from  $X$  and  $Y$  coordinates of the previous

pixel. The IFS codes given below will illustrate the procedure. The program, IFS.BAS, will generate a solid rectangle from these IFS codes in a random and rather spectacular way.

IFS Codes for a Square

A	B	C	D	E	F	P
.5	0	0	.5	1	1	.25
.5	0	0	.5	50	1	.25
.5	0	0	.5	1	50	.25
.5	0	0	.5	50	50	.25

From the IFS codes, new values of X and Y are computed from one of four sets of equations as follows.

Equation Set 1

$$\begin{aligned} \text{New X} &= A * X + B * Y + 1 \\ \text{New Y} &= C * X + D * Y + 1 \end{aligned}$$

Equation Set 2

$$\begin{aligned} \text{New X} &= A * X + B * Y + 50 \\ \text{New Y} &= C * X + D * Y + 1 \end{aligned}$$

Equation Set 3

$$\begin{aligned} \text{New X} &= A * X + B * Y + 1 \\ \text{New Y} &= C * X + D * Y + 50 \end{aligned}$$

Equation Set 4

$$\begin{aligned} \text{New X} &= A * X + B * Y + 50 \\ \text{New Y} &= C * X + D * Y + 50 \end{aligned}$$

The variable P indicates that the probability that any one set is used to compute the new X and Y coordinates is 0.25. After each iteration of one set of equations, the X and Y coordinates are plotted, and another set of equations is randomly chosen to compute the next point.

The procedure is illustrated graphically in Figure 5. The X and Y coordinates are scaled by 0.5 and either a 1 or 50 is added to the X and/or Y values. This procedure is repeated until the entire rectangle is painted.

Any image can be reduced to sets of IFS codes. Reference [8] provides other IFS codes that can be used with IFS.BAS for more exotic images.

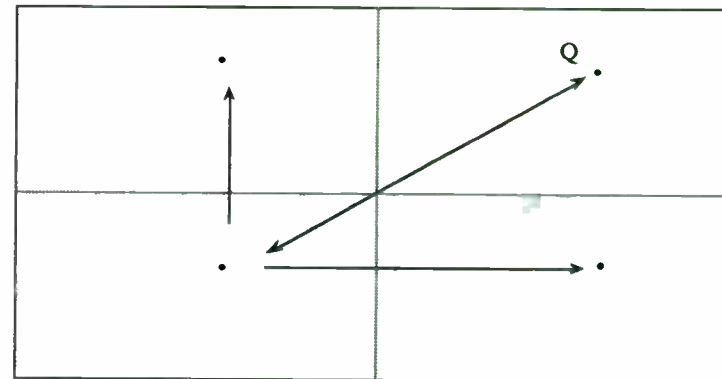


Figure 5. Random Mapping of Point Q for a Fractal Process



### Section III. Image Compression and Bandwidth Reduction

An analog NTSC cable TV system requires 6 MHz per channel. It might seem that a compression of 50:1 would reduce the bandwidth by a factor of 50 (to 120,000 Hz), but it does not. Using an estimated resolution of 512x480 for TV with 8-bit color and a frame rate of 30 times per second, the required bit rate for digital transmission is 58,982,000 bits/sec, or approximately 60 Mbps. Since compression techniques are inherently digital, the 6 MHz analog signal was converted to a 60,000,000 bit/sec digital signal and now requires an absolute minimum bandwidth of 30 MHz (based on one bit per symbol). Compression by 50:1 would result in a total transmission bandwidth of at least 600,000 KHz.

Shannon's law dictates the maximum transmission speed of digital data. This speed limit is determined by only two factors which include the bandwidth of the transmission channel, and the signal to noise ratio of the channel. In a digital system, the maximum data rate, called the channel capacity  $C$  is bounded by Shannon's law which is given by

$$C \text{ bits/second} = BW \log_2(1 + S/N)$$

where  $BW$  is the bandwidth of the channel in Hertz. From this equation, it can be seen that if there is no noise in the channel ( $S/N = \text{infinity}$ ), then the channel capacity becomes infinite regardless of the bandwidth.

As an example, consider an analog phone line with a bandwidth of 4000Hz and

assuming that it is digitized at 8000 times per second using 8 bits per byte. If no other noise is present on the line, the signal to noise ratio will be 48 dB since each bit of resolution in the A/D convertor decreases the quantization noise by 6dB.

To use Shannon's law, the  $S/N$  ratio must be expressed in linear terms, or  $S/N = 10^{(48/10)}$ . The channel capacity then becomes

$$C = 64,000 \text{ bits/sec} = 4000 \log_2(1 + 10^{(4.8)})$$

and will not be affected by the number of bits encoded in a symbol. The theoretical limit essentially states that it is possible to build dial-up modems that transmit without error at rates up to 64,000 bits/sec. Shannon's law does not tell us how to design hardware to achieve the theoretical maximum bit rate. The common quadrature phase shift keying (QPSK) and quadrature amplitude modulation (QAM) techniques used for modems have inherent deficiencies that will not allow them to obtain the maximum theoretical bit rate; however, Shannon's law can be used to determine the theoretical limit on the channel capacity.

The distinction between bit rate and symbol rate is vital in determining bandwidth requirements. The QPSK modulation technique encodes two bits in a symbol since there are 4 possible phases (0 degrees, 90 degrees, 180 degrees, and 270 degrees) of the carrier. A change of phase of the carrier represents that transmission of a single symbol (baud) which conveys two bits of information according to the following table.

Table 3

Differential Phase	Dibits
0 degrees	00
90 degrees	01
180 degrees	11
270 degrees	10

Although it is common to speak of a 1200 baud modem, the bit rate is 1200 bits/sec and the baud rate (with QPSK modulation) is really 600 baud.

QAM varies both the amplitude and phase of the carrier in order to encode bits into symbols. 16-QAM means the carrier can assume one of 16 possible states in order to encode  $\log_2(16)$ , or 4, bits per symbol. The constellation for a 16-QAM carrier is shown below where the distance from the origin represents the amplitude of the carrier, and the angle of the point represents the phase of the carrier.

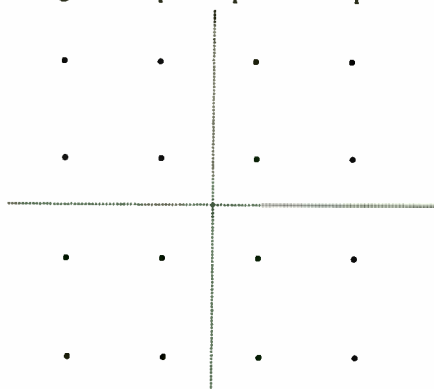


Figure 5. 16-QAM Constellation

The difficulty with using 64-QAM, 256-QAM, or higher is that the susceptibility to noise increases exponentially as the number of bits per symbol is increased.

The bandwidth required for digital transmission of data is given by

$$BW = \frac{\text{Data Rate (symbols/sec)} \times (1 + r)}{2}$$

where  $r$  represents the Nyquist rolloff factor of the raised cosine filter and  $r$  falls between 0 and 1. The basis for this equation will not be derived but is given in [12].

In the previous example, if the 64,000 bit per second data rate is to be transmitted using frequency shift keying (FSK) such that only one bit is encoded per symbol, then the symbol rate becomes 64,000 symbols per second. The required transmission bandwidth is usually taken as 64,000 Hz ( $r = 1$ ) in most texts, although it can approach as low as 32,000 Hz by using a sharp cutoff Nyquist filter ( $r = 0$ ).

Throughout the remainder of this article, it will be assumed that the required minimum transmission bandwidth is 1/2 of the symbol rate.

Digital HDTV has a proposed resolution of 1440x960 pixels/frame x 12 bits/pixel x 30 frames/sec (12 bits/pixel = 8bits/pixel for luminance + 4 bits/pixel for chrominance at half-resolution). The data rate is about 500 Mbits/sec. The aim of most HDTV systems is to broadcast at about 15 Mbits/sec. This requires approximately 33:1 compression. Encoding with one bit/symbol would require a transmission bandwidth of 7.5 MHz. Encoding with 4 bits/symbol would increase

the susceptibility of the transmitted signal to noise but would only require a bandwidth of 1.875 MHz, as a minimum.

The Common Intermediate Format (CIF) is a common format for transmitting video images at 30 frames/sec over a T1 line (capacity 1.544 Mbits/sec) or a Quarter-CIF (QCIF) picture at 10 frames/sec over Integrated Services Digital Network (ISDN) telephone line (capacity 64 Kbits/sec).

A CIF image is composed of luminance channel with a resolution of 288 lines/frame  $\times$  352 pixels/line and 8 bits/pixel, and two chrominance channels ( $C_b$  and  $C_r$ ) with half resolution, i.e., 144 lines/frame  $\times$  176 pixels/line and 8 bits/pixel. At 30 frame/sec, the data rate is 36.5 Mbits/sec. Transmission over a T1 line for videoconferencing applications requires a compression ratio of about 24:1. A QCIF image is 1/4th the size of a CIF image and at 10 frames/sec requires about 3 Mbits/sec. Thus, video phone over ISDN network requires about 48:1 compression.

#### Advanced Digital Television (ADTV)

ATDV was developed by the Advanced Television Research Consortium and uses MPEG compression to provide a data rate of 24 million bits per second. It has two trellis-coded 32-QAM data carriers to provide a wide bandwidth standard priority channel and a narrower bandwidth high-priority channel, all within a 6 MHz transmission bandwidth. The high priority channel provides the viewable picture,

and the additional standard-priority channel provides the full HDTV quality. The ADTV spectrum is shown in Figure 6.

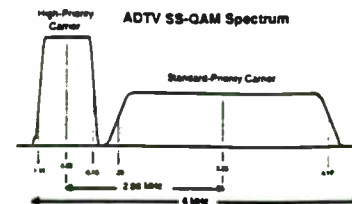


Figure 6. ATDV 32-QAM channel spectrum.

ADTV's 59.94 field rate is identical to NTSC, thus eliminating temporal artifacts and the need for frame synchronization in mixed ADTV-NTSC environments. Its 1440 $\times$ 960, 1050-line scanning format is cost effective in the production studio. The 2:1 vertical ratio with 525 line NTSC video and 2:1 horizontal ratio with the CCIR Rec. 601 sampling standard used in the 525 line D1 tape recorders offers economical transcoding in mixed ADTV-NTSC production environments. The 16 Mbit DRAM frame memories in an ADTV receiver are predicted to cost about \$13 each by 1996.

#### Digital Spectrum Compatible (DSC)

DSC was developed by Zenith and AT&T to provide all digital definition

television simulcast by compressing the wide bandwidth digital signal into a 6 MHz channel. It uses a unique four-level vestigial sideband (4-VSB, or 2 bits per symbol) modulation technique to assure noise free and interference free reception.

The 4-VSB coding is complemented by a two-level digital data system (2-VSB, or one bit per symbol). The resulting bi-rate coding system identifies and selects the most important picture information on a scene-by-scene basis and automatically transmits that data in a two-level (one bit/symbol) binary mode. The two-level digital coding makes the system far more tolerant to noise and other interference at greater distances from the transmitter.

DSC uses a 17 Mbps data rate and a 787.5 line progressive scanning format to eliminate artifacts due to interlacing video.

#### American Television Alliance Digicipher (ATVA)

ATVA was developed by General Instrument Corporation and provides another alternative for squeezing a HDTV signal into a 6 MHz bandwidth. Compression is based on the DCT transform. With ATVA, there are two distinct transmission modes, 32-QAM and 16-QAM. The broadcaster can select the mode, and the receivers can auto-configure to the mode being transmitted. The 32-QAM mode requires 16.5 dB signal-to-noise ratio for error free transmission, whereas the 16-QAM mode requires only a 12.5dB signal-to-noise ratio.

Figure 7 shows a block diagram of the encoder. The digital video encoder accepts YUV inputs with 16:9 aspect ratio and 1050 line interlace (1050/2:1) at a 59.94 field rate. The YUV signals are obtained from analog RGB input by low pass filtering, A/D conversion, and an RGB-to-YUV matrix. The sampling frequency is 53.65 MHz for R, G, and B. The digital video encoder implements the compression algorithm and generates a video data stream.

The multiplexer combines the various data streams into one data stream at 18.22 Mbps. The forward error correction (FEC) encoder adds error correction overhead bits and provides 24.39 Mbps of data to the 32-QAM modulation. The symbol rate of the 32-QAM signal is 4.88 MHz.

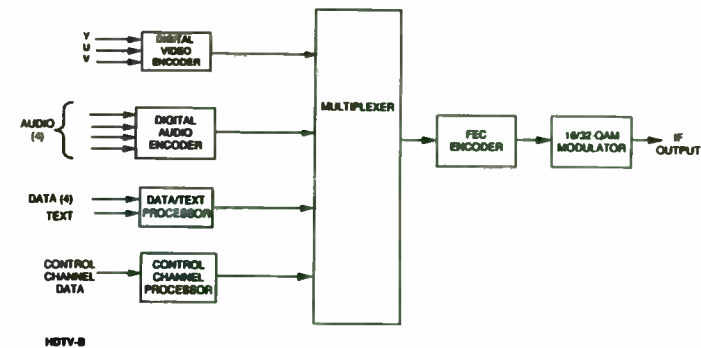


Figure 7. ATVA encoder block diagram

## Summary

This document is a review of image compression techniques for application to the digital transmission of motion video. Lossless compression techniques are covered since they represent a component in a lossy technique. Emphasis is given to the discrete cosine transform since it is the basis for the JPEG still image compression algorithm and the MPEG motion picture algorithm. Basic modulation theory provides a lower bound on the limits of bandwidth and data rate. Finally, three competitive HDTV standards are discussed.

### ENTROPY.BAS

```
10 DIM J(257), K1(257)
20 CLASS
30 SUM = 0
40 ICOUNT = 0
50 INPUT "Type name for input file or <RET> "; F$
60 PRINT
70 OPEN F$ FOR BINARY AS #1
80 A$ = INPUT$(1, 1)
90 IF EOF(1) THEN GOTO 130
100 ICOUNT = ICOUNT + 1
110 J(ASC(A$)) = J(ASC(A$)) + 1
120 GOTO 80
130 FOR I = 0 TO 255
```

```
140 IF J(I) = 0 THEN GOTO 160
150 SUM = SUM - (J(I) / ICOUNT) / LOG(2) * LOG(J(I) / ICOUNT)
160 NEXT I
170 PRINT "The entropy = "; SUM; " bits per symbol."
180 PRINT "A lossless compression of "; 8 / SUM; " to 1 is possible with entropy coding."
190 CLOSE #1
200 PRINT
210 PRINT "Type any key to end "
220 A$ = INKEY$
230 IF A$ = "" THEN GOTO 220
240 END
```

### MSE.BAS

```
10 REM Program Mean-Square-Error
20 REM this program computes the mean square error between two
25 REM files
30 n = 1
40 CLASS
50 INPUT "Type filename #1 "; F$
60 INPUT "Type filename #2 "; G$
70 OPEN F$ FOR BINARY AS #1
80 OPEN G$ FOR BINARY AS #2
90 A$ = INPUT$(1, 1)
100 B$ = INPUT$(1, 2)
110 IF EOF(1) THEN GOTO 1000
120 SUM = SUM + (ASC(A$) - ASC(B$)) ^ 2
```

```

130 n = n + 1
140 GOTO 90
1000 SUM = SUM / n
1010 PRINT "The mean square error = "; SUM

```

#### ENTROPY.C

```

#include <math.h>
#include <stdio.h>
/*Huffman Estimator*/
/* Mike Ellis */
main ()
{
char string[80];
unsigned char p;
int i,k;
float j[256],pixel;
float entropy;
FILE *input_file;
FILE *infile;
    entropy = 0;    /*Set all Variable to 0 */
    for ( i = 0; i<=255; i++)
    {
        j[i]=0;
    }
    pixel = 0;
    printf ("\nType name of file "); /*Open the input file*/
    scanf ("%s",string );

```

```

input_file=fopen(string,"r+b"); /*For Read + Binary */
while (!feof(input_file)) /*Read to end-of-file*/
{
    fscanf (input_file,"%c",&p); /*input as unsigned char*/
    pixel = pixel + 1; /*count characters read*/
    k=p; /*convert char to integer*/
    j[k]=j[k]+1; /*number of times that */
} /*this char was read */
for (i = 0; i <=255; i++) /*compute entropy*/
{
    if (j[i] !=0)
    {
        entropy = entropy + j[i]*log(j[i]/pixel);
    }
}
entropy = -entropy/(pixel*log(2));
printf ("\nEntropy = %f",entropy); /*print the entropy*/
printf("\nEntropy encoding can achieve");
printf(" a %f to 1 compression.\n",8/entropy);
}

```

#### IMAGE.BAS

```

10 REM
20 REM PROGRAM "IMAGE" FOR DISPLAY GRAYSCALE OR COLOR IMAGES
30 REM USE UNDER QUICKBASIC VERSION 4.5
40 REM

```

```

50 CLASS
60 INPUT "Type Image Filename "; C$
70 DEFINT A-Z
80 INPUT "Type X Dimension "; XDIM
90 INPUT "Type Y Dimension "; YDIM
100 GOSUB 210          'set up 256 color palette
110 OPEN C$ FOR BINARY AS #1
120 WINDOW SCREEN (0,0)-(XDIM, YDIM)
130 FOR I = 1 TO YDIM
140 A$ = INPUT$(XDIM, I)
150 FOR J = 1 TO XDIM
160 B = ASC(MID$(A$, J, 1))
170 PSET (J, I),      'display the pixel
180 NEXT
190 NEXT
200 GOTO 200
210 INPUT "Type Palette Filename "; P$
220 SCREEN 13
230 OPEN P$ FOR BINARY AS #1
240 DIM RED(256), GREEN(256), BLUE(256)
250 FOR K = 0 TO 255
260 A$ = INPUT$(1, 1)
270 RED(K) = ASC(A$)
280 NEXT K
290 FOR K = 0 TO 255
300 A$ = INPUT$(1, 1)
310 GREEN(K) = ASC(A$)

```

```

320 NEXT K
330 FOR K = 0 TO 255
340 A$ = INPUT$(1, 1)
350 BLUE(K) = ASC(A$).
360 NEXT K
370 FOR I = 0 TO 255
380 PALETTE I, (INT(BLUE(I) / 4)) * 65536! + 256 * (INT(GREEN(I) / 4)) + INT(RED(I)) / 4
390 NEXT I
400 CLOSE #1
410 RETURN

```

#### DCT.BAS

```

10 COMMON DATA 1%()          'More than 64K block
20 CLEAR
30 DIM C(8, 8), DATA1%(90, 1, 8, 8), TEMP(8, 8)
40 DIM TEMP2(8, 8)
50 MAX(1, 1) = -100000!
60 MIN(1, 1) = 100000!
70 REM
80 REM Collect Input Data 8 X 8 Byte Blocks
90 REM
100 INPUT "Type Name of Input File "; F$
110 IF F$ = "" THEN FILES: GOTO 100
120 K = INSTR(1, F$, ".")
130 IF K > 0 THEN A$ = LEFT$(F$, K - 1) ELSE A$ = F$
140 INPUT "Type X Range "; XINDEX
150 INPUT "Type Y Range "; YINDEX

```

```

160 XINDEX = (XINDEX / 8)
170 YINDEX = (YINDEX / 8)
180 K1 = XINDEX * YINDEX
190 G$ = A$ + ".dct"
200 OPEN F$ FOR BINARY AS #1
210 OPEN G$ FOR OUTPUT AS #2
260 GOSUB 870      'Define cosine transform matrix
290 REM
300 FOR J = 1 TO YINDEX
310 PRINT J * 8
320 FOR Y = 1 TO 8
330 I1 = 0
340 A1$ = INPUT$(8 * XINDEX, 1)
350 FOR I = 1 TO XINDEX
360 FOR X = 1 TO 8
370 I1 = I1 + 1
380 A$ = MID$(A1$, I1, 1)
390 DATA1%(I, 1, Y, X) = ASC(A$) - 128
400 NEXT X
410 NEXT I
420 NEXT Y
430 GOSUB 590      'Perform the cosine transform
440 FOR X = 1 TO 8
450 FOR I = 1 TO XINDEX
460 FOR Y = 1 TO 8
470 K = (DATA1%(I, 1, X, Y))
480 S$ = S$ + CHR$(K + 128)

```

```

490 NEXT
500 NEXT
510 PRINT #2, S$;
520 S$ = ""
530 NEXT
540 REM
550 NEXT
560 CLOSE #1
570 CLOSE #2
580 END
590 REM
600 REM
610 FOR I = 1 TO XINDEX
620 REM FOR k = 1 TO 8
630 FOR L = 1 TO 8
640 TEMP1 = 0
650 FOR M = 1 TO 8
660 FOR N = 1 TO 8
670 TEMP1 = TEMP1 + DATA1%(I, 1, M, N) * C(L, N)
680 NEXT
690 TEMP(M, L) = TEMP1
700 TEMP1 = 0
710 NEXT
720 NEXT
730 REM NEXT
740 FOR L = 1 TO 8
750 TEMP1 = 0

```



```

760 FOR M = 1 TO 8
770 FOR N = 1 TO 8
780 TEMP1 = TEMP1 + C(M, N) * TEMP(N, L)
790 NEXT N
800 DATA 1%(I, 1, M, L) = TEMP1 / 8
810 TEMP1 = 0
820 NEXT
830 NEXT
840 REM
850 NEXT I
860 RETURN
870 REM
880 REM Generate Cosine Transform Matrix
890 FOR K = 0 TO 7
900 FOR N = 0 TO 7
910 IF K = 0 THEN C(K + 1, N + 1) = 1 / SQR(8)
920 IF K <> 0 THEN C(K + 1, N + 1) = SQR(.25) * COS(3.14159 * (2 * N + 1) * K / 16)
930 NEXT N
940 NEXT K
950 RETURN

```

#### INVDCT.BAS

```

10 COMMON DATA 1()      'More than 64K block
20 CLEAR
30 DIM C(8, 8), DATA1(90, 1, 8, 8), TEMP(8, 8)
40 MAX(1, 1) = -100000!

```

```

50 MIN(1, 1) = 100000!
60 REM
70 REM Collect Input Data 8 X 8 Byte Blocks
80 REM
90 INPUT "Type Name of Input File "; F$
100 IF F$ = "" THEN FILES: GOTO 90
110 K = INSTR(1, F$, ".")
120 IF K <> 0 THEN A$ = LEFT$(F$, K - 1) ELSE A$ = F$
130 INPUT "Type X Range "; XINDEX
140 INPUT "Type Y Range "; YINDEX
150 XINDEX = (XINDEX / 8)
160 YINDEX = (YINDEX / 8)
170 KI = XINDEX * YINDEX
180 H$ = A$ + ".rec"
190 OPEN F$ FOR BINARY AS #1
200 OPEN H$ FOR OUTPUT AS #2
250 GOSUB 820          'Set up C matrix
280 REM
290 FOR J = 1 TO YINDEX
300 PRINT J * 8
310 FOR Y = 1 TO 8
320 FOR I = 1 TO XINDEX
330 FOR X = 1 TO 8
340 A$ = INPUT$(1, 1)
350 DATA 1(I, 1, Y, X) = ASC(A$) - 128
360 NEXT X
370 NEXT I

```

```

380 NEXT Y
390 GOSUB 550           'Perform Inverse Cosine
400 REM               'Transform
410 FOR X = 1 TO 8
420 FOR I = 1 TO XINDEX
430 FOR Y = 1 TO 8
440 K = CINT(DATA1(I, 1, X, Y))
450 IF K > 127 THEN K = 127
460 IF K < -128 THEN K = -128
470 PRINT #2, CHR$(K + 128);
480 NEXT
490 NEXT
500 NEXT
510 REM
520 NEXT
530 CLOSE #1: CLOSE #2
540 END
550 REM
560 REM
570 REM
580 FOR I = 1 TO XINDEX
590 FOR L = 1 TO 8
600 TEMP1 = 0
610 FOR M = 1 TO 8
620 FOR N = 1 TO 8
630 TEMP1 = TEMP1 + DATA1(I, 1, M, N) * C(N, L)
640 NEXT

```

```

650 TEMP(M, L) = TEMP1
660 TEMP1 = 0
670 NEXT
680 NEXT
690 FOR L = 1 TO 8
700 TEMP1 = 0
710 FOR M = 1 TO 8
720 FOR N = 1 TO 8
730 TEMP1 = TEMP1 + C(N, M) * TEMP(N, L)
740 NEXT N
750 DATA1(I, 1, M, L) = TEMP1 * 8
760 TEMP1 = 0
770 NEXT
780 NEXT
790 REM
800 NEXT
810 RETURN
820 REM
830 REM Generate Cosine Transform Matrix
840 FOR K = 0 TO 7
850 FOR N = 0 TO 7
860 IF K = 0 THEN C(K + 1, N + 1) = 1 / SQR(8)
870 IF K <> 0 THEN C(K + 1, N + 1) = SQR(.25) * COS(3.14159 * (2 * N + 1) * K / 16)
880 NEXT N
890 NEXT K
900 RETURN

```

IFS.BAS

```

10 KEY OFF
20 INPUT "Type Data File for IFS codes "; IFSS$
30 OPEN IFSS$ FOR INPUT AS #1
40 I = 1
50 INPUT #1, A(I), B(I), C(I), D(I), E(I), F(I), P(I)
60 P(I) = P(I) + P(I - 1)
70 IF EOF(1) THEN GOTO 100
80 I = I + 1
90 GOTO 50
100 SCREEN 9: CLASS
110 CLOSE #1
120 MINX = 10000: MAXX = -10000: MINY = 10000: MAXY = -10000
130 COLOR 10
140 REM WINDOW (-3, -3)-(5, 15)
150 X = 0: Y = 0: NUMITS = 1000!
160 FOR N = 1 TO NUMITS
170 K1 = INT(RND * 100)
180 FOR J = 1 TO I
190 IF (K1 >= P(J - 1) * 100) AND (K1 < P(J) * 100) THEN K = J
200 NEXT J
210 NEWX = A(K) * X + B(K) * Y + E(K)
220 NEWY = C(K) * X + D(K) * Y + F(K)
230 X = NEWX
240 Y = NEWY
250 OLDK = K
260 IF (N > 10) AND (NUMITS < 10000) THEN GOSUB 330

```

```

270 IF (N > 10) AND (NUMITS > 10000) THEN PSET (X, Y)
280 NEXT
290 IF NUMITS > 10000 THEN GOTO 380
300 X = 0: Y = 0: NUMITS = 10000000#
310 WINDOW (MINX - .5 * ABS(MINX), MINY - .5 * ABS(MINY))-(MAXX +
.5 * ABS(MAXX), MAXY + .5 * ABS(MAXY)): CLASS : GOTO 160
320 END
330 IF X > MAXX THEN MAXX = X
340 IF X < MINX THEN MINX = X
350 IF Y > MAXY THEN MAXY = Y
360 IF Y < MINY THEN MINY = Y
370 RETURN
380 A$ = INKEY$
390 IF A$ = "" THEN GOTO 380
400 END

```

IFS Codes for a Square

A	B	C	D	E	F	P
.5	0	0	.5	1	1	.25
.5	0	0	.5	50	1	.25
.5	0	0	.5	1	50	.25
.5	0	0	.5	50	50	.25

## References

1. "The Data Compression Book", Mark Nelson, M&T Books, 1991
2. "A Universal Algorithm for Sequential Data Compression", Ziv, Jacob, and Lempel, Abraham, IEEE Transactions on Information Theory, 1977
3. "A Technique for High-Performance Data Compression", Welch, Terry, IEEEComputer, June 1984
4. "Compression of Individual Sequences via Variable-Rate Coding", Ziv, Jacob, and Lempel, Abraham, IEEE Transactionsof Information Theory, Sept. 1978
5. "Programmer's Guide to the EGA and VGA Cards", Second Edition, Ferraro, Richard F., Addison-Wesley, 1990
6. "DR. Dobb's Journal November 1991 Software Listings", M&T Books, 501 Galveston Drive, Redwood City, CA 94063-4728, 1-800-533-4372
7. "386-MATLAB User's Guide", The Mathworks, Inc., 508-653-1415
8. "Fractals Everywhere", Barnsley, Michael, Academic Press, Inc., 1988
9. "Fundamentals of Digital Image Processing" Jain, Anil K., Prentice-Hall, 1989
10. NCSA Image 3.1, National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign
11. "Methods and Apparatus for Image Compression by Iterated Function System",U.S. Patent 4,941,193, Barnsley, Michael, July 10, 1990
12. "Digital and Analog Communication Systems", Couch, Macmillan Publishing Co., 1983

### Test Questions

1. Give them a Huffman coded file and have them decode it.
2. Have them do a simple Huffman tree
3. A matrix multiply question
4. A syntax difference question of Basic versus C
5. Symmetrical extension for DCT using FFT.
6. A question about 64-QAM

□



HIGH TEMPERATURE SUPERCONDUCTIVITY (HTS)  
AND ITS ROLE IN ELECTRONICS WARFARE

Thomas N. Tuma

U.S. Army CECOM Signals Warfare Directorate  
Research and Technology Division  
Vint Hill Farms Station  
Warrenton, VA 22186

ABSTRACT

DURING THE PAST EIGHT YEARS, THE QUEST FOR MATERIALS HAVING THE PROPERTY OF SUPERCONDUCTIVITY AT "HIGH" TEMPERATURES (I.E. FAR ABOVE ABSOLUTE ZERO) HAS YIELDED ASTONISHING RESULTS. FIRST, VARIETIES OF YTTRIUM-BARIUM-COPPER-OXIDES, COOLED WITH LIQUID NITROGEN TO 77 DEGREES KELVIN WERE PRODUCED AND MORE RECENTLY OTHER CERAMICS CAPABLE OF BEING SUPERCONDUCTIVE AT 130 DEGREES K ARE OPENING THE WAY TO THE ROUTINE DESIGN OF A GROWING FAMILY OF ELECTRONIC COMMUNICATION COMPONENTS, TO COMPLEMENT THE PIONEER HTS ACHIEVEMENTS IN ULTRA POWERFUL MOTORS, ELECTROMAGNETS, ETC.

THE DEBUT OF HTS COMPONENTS IN THE RF FIELD WAS IN THE DESIGN OF A VARIETY OF MICROWAVE COMPONENTS, WHOSE FREQUENCY-DICTATED SMALL SIZES MAKES THEM EASY TO PACK IN A CRYOGENICS ENVIRONMENT. WITH THE EXPERIMENTAL EVIDENCE THAT AC RESISTIVITY DECREASES STEADILY AS ONE DECREASES THE VALUE OF RF FREQUENCY USED, IT IS EVIDENT THAT THIS BENEFICIAL PHENOMENON HAS TO BE EXPLOITED IN THE HF AND VHF BANDS. RECEIVERS WITH SUPERCONDUCTING COMPONENTS WILL HAVE IMPRESSIVE SENSITIVITY; HTS TRANSMITTER COMPONENTS ALLOW HIGHER ERP'S AND ADDITIONALLY SAVE VOLUME AND WEIGHT. MODERN ARMY JAMMING SYSTEMS WILL DEFINITELY SUFFER A METAMORPHOSIS AS A RESULT OF THE RESEARCH IN THIS FIELD, PIONEERED AT CECOM SWD IN WARRENTON VIRGINIA.

1.0 HISTORICAL INTRODUCTION

On July 10 1908, Dr. Heike K. Onnes, professor of experimental Physics at the University of Leyden in The Netherlands succeeded to liquefy helium for the first time. Helium liquefies at 4.2 degrees K under standard atmospheric pressure. This rare product was used by Onnes to investigate resistivities of various metals in this temperature regime. Earlier experiments had shown that by lowering the temperature of conductors, their DC resistivities decreased, and by extrapolation they would vanish at the (unreachable) absolute zero. In Onnes' experiments an unexpected result was observed: the resistivity of mercury suddenly became zero when it got cooled by helium at 4.2 degrees K. After ascertaining that there were no short circuits and that the measurements were accurate, he tested other metals; among them Indium, Tin and Lead. Their critical temperatures (transition to superconductors) were respectively 3.4, 3.7 and 7.2 degrees K (Ref 2).

Dr. Onnes received the Nobel award in Physics in 1913. Large scale commercial applications of superconductivity were soon envisioned as in a George Orwell creation. At that time it was not known that external magnetic fields exceeding a critical value  $H_c$ , large throughput current or power values and high frequencies, as well as temperatures above a critical value  $T_c$  would quench the effect of superconductivity. This quenching extended to some of these preliminary envisions of applications, making it necessary to re-examine the potential applications of low temperature superconductivity.

In 1933, Walter Meissner and Robert Ochsenfeld, in Germany, discovered that an external magnetic flux of moderate amplitude was not only excluded from entering a superconductor (diamagnetism), but it was also dynamically expelled from its interior (at some expense of energy). This experience, the "Meissner Effect" demonstrated that superconductivity is a magnetic phenomenon and is used as a test on whether a material is superconductive. In 1952, Alexei Abrisokov, in what was called the Soviet Union, produced a new family of materials: Type II superconductors. These types can conduct large currents and expel even larger values of magnetic fields, therefore they can be used to make powerful electromagnets which do not quench to the normal state under normal use conditions.

HTS, or superconductors with high critical temperature ( $T_c$ ) were first reported in April 1986 by the Zurich IBM group led by Georg Bednorz, who produced a Lanthanum-Barium-Copper-Oxide with a  $T_c$  of 30 degrees K. The modern types of HTS materials were arrived at in Huntsville, Alabama in January, 1987 by Wu, Torng and Ashburn, who were seeking a substitute for Barium in La-Ba-Cu-O. This quest led to the discovery of Yttrium-Barium Copper Oxide or  $Y_1 Ba_2 Cu_3 O_7$ . This easily produced ceramic has a  $T_c$  of 90° K (Fig 1), so liquid nitrogen is more than sufficient to cool it to the superconductive state. On March 18 1987 the American Physical society held a meeting in New York City where the new ceramics were introduced to the public and scientific communities. Presently HTS materials make several families, each composed of various proportions of the elements Y-Ba-Cu-O, (LaSr)-Cu-O, Bi-Sr-Ca-Cu-O and Tl-Ba-Ca-Cu-O.

## 2.0 IMPORTANT SUPERCONDUCTIVITY PHENOMENA

The superconductive state's most notable macroscopic features are zero DC (and very low AC) resistivity and near perfect diamagnetism under moderate magnetic fields. The last property, the Meissner Effect culminates with the expulsion of magnetic fields from the interior of the material, except for a thin region at the surface, where a screening current is established to cancel the magnetic field in the interior.

In the superconductive state electrons associate in loose pairs (Cooper Pairs) separated by a "coherence length" interval and interact with the lattice through phonon interactions (lattice vibrations). In the DC case and with a perfect lattice, the Cooper pairs flow coherently along the crystals in the same plane without interruption or loss. In the AC case, the fact that electrons have nonzero momentum and kinetic energy, prevents the electrons from reacting instantaneously to the AC electric field, so this lag creates a secondary normal electric field along the surface of the conductor which in turn will drive normal (single) electrons producing ohmic loss. The theory explaining this well understood phenomenon is referred to as the two fluid model (Ref 1).

At frequencies higher than the energy gap value divided by Planck's constant, the extra energy absorbed by the material excites the electrons so that the Cooper pairs dissociate and the electrons become normal, thus quenching superconductivity.

The superconductive state is also destroyed by high current densities exceeding a critical value. This critical current is a



function of the temperature, the crystal geometry and the state of the material, and depends too on whether the magnetic field produced by the surface current exceeds the material's  $H_c$ .

If an applied magnetic field exceeds the  $H_c$  value, the screening current that would prevent H from penetrating the material, would require a large amount of energy to be established. If the energy required to produce the screening current is larger than the energy required to transition to its normal state, the system will follow the least energy demanding path and the material will become normal.

The two known types of superconductors (Types I and II) have respectively one and two  $H_c$  values. In type I superconductors, if  $H > H_c$ , then H penetrates the material completely, quenching it. In type II materials, which have an  $H_{c,low}$  and an  $H_{c,high}$ , two phenomena can occur: if  $H < H_{c,low}$  or if  $H > H_{c,high}$  there is respectively no flux penetration at all, or there will be quenching as in type I. For H values between both  $H_c$ 's of type II, the H field penetrates in a partial and organized way, forming a "mixed state" of pinning points or vortices whose radii are comparable to the coherence length. These vortices are distributed in regular intervals within the otherwise superconductive material. As H approaches  $H_{c,high}$ , these vortices group together more densely until the material runs out of superconducting regions and becomes normal. The HTS material discovered in 1987 by Wu et. al. belongs to the Type II class.

The mixed state of Type II superconductors allows for much higher critical current densities and magnetic field strengths. This difference can be as high as three orders of magnitude.

### 3.0 CHARACTERISTICS OF HIGH TEMPERATURE SUPERCONDUCTOR MATERIALS

Modern HTS materials are ceramics composed of copper oxides and non-metallic chemical elements, as mentioned in Section 1.0. All HTS material belong to the Type II class. Indeed, the most notable difference between Types I and II is the lack in composition of common metallic conductors as in Type II.

The crystal structure of an HTS material is similar to a "perovskite" unit cell (Fig 2) where a metal anion lies at the center of a cube of cations, surrounded by an octahedral configuration of oxygen atoms. In these unit cells, the Cu-O atoms are strongly linked and all lay in two planes in each cell. These planes correspond to the a-b plane of the cell and it is only in these planes that the conduction phenomenon occurs. Additionally, the zero DC or low AC resistivities occur at the surface of the sample, decaying exponentially as one penetrates into the material. Based on these facts surface resistivity ( $R_s$ ) is a significant parameter of HTS materials. These anisotropies of HTS materials have to be taken into consideration in the design of superconductive electronics components.

The three structural types of HTS materials in production today are bulk, thin films and thick films. Bulk HTS materials are free standing poly or single crystals made by sintering, i.e., extruding a composite of an HTS powder and a termoplastic binder which is heated and shaped into a continuous monofilament called green fiber. Other desired shapes can be made, such as a straight line with a certain cross section or a helix. After sintering, the wire is melt

textured, in other words, it is zone-wise molten under a constant magnetic field and made to recrystallize into either one very long crystal or a collection of parallel crystals with undiscernible grain boundaries. Melt textured bulk materials can withstand current densities of 20,000 Amp/cm as well as magnetic fields of 8 Teslas (80,000 Gauss) without quenching.

Thin films are made by deposition onto substrates by processes like laser ablation, off-axis RF magnetron sputtering or electrophoretic deposition. They can be deposited into single or poly-crystals, where the single crystal or epitaxial format offers the best values of surface resistivity and critical temperature. For several years thin films had the lowest available values of surface resistivity, especially at VHF and UHF frequencies. Today, bulk and thin films are in even competition. Thick films are the newest and hitherto less tested materials. They have properties common to both bulk and thin films (Fig 3, Table 1).

At the present there is a project of RF testing of HTS materials of the three kinds as produced by numerous private industries and universities. This is being done at the Naval Research Laboratory in Washington, DC (Ref 10). The tests are designed to qualify the various samples in terms of the possibility of production of unwanted power dependence effects and non linear mechanisms of signal degradation.

#### 4.0 RF AND EW APPLICATIONS OF HTS MATERIALS

Since the onset of high temperature superconductivity, the world

engineering community in general has been eagerly looking for new applications provided by high temperature superconductivity to improve the performance of electrical and electronic equipment.

The dawn of RF use of HTS materials was in thin film technology, which was applied into microwave components. Two important examples of RF applications of HTS in the microwave frequency world are resonators and distributed element bandpass filters. These devices can greatly reduce system noise temperature and improve the stability of a signal. An inherent asset of HTS filters is their very high Q values, so they can satisfactorily replace cavity filters. At this time, thin film technology seems best suited for the production of these devices of planar geometry.

The most important application envisioned for HTS devices is for the front end of HF, VHF and UHF transmitting and receiving equipment. To this date, the front end devices that can be supercooled are antenna elements, impedance matching networks or couplers, tuners, delay lines, filters, resonators and Butler matrices (Figs 7, 8, 9; Ref 6).

Consider for example superconducting delay lines. These devices are used routinely military ESM applications such as in signal processing, like calculating fast Fourier transforms, correlation functions and other analogical calculations where the value of a signal is needed at a shifted time. The HTS delay line's advantage over normal delay lines is that a normal temperature delay line has unbearable losses, while their tiny HTS counterparts can do the same job without appreciable losses.

Next, consider the application of HTS technology to advance the state of the art of High Frequency (HF) antennas. The already demonstrated HTSC characteristics such as the 4 to 6 order of magnitude reduction in surface resistance of HTSC bulk materials was used to theoretically show that the size of HF antennas systems could be reduced by at least two orders of magnitude\*. The more recently demonstrated high critical current (10,000A/cm<sup>2</sup> at 1 Tesla) melt textured HTS material shows that transmit applications are also attractive. HTS material have applications for instance in small antennas, beam formers and RF memories for jammers. The advent of high reliability small size closed cycle cryogenic coolers such as the Oxford Systems cooler employed in the NRL Superconductor Space Experiment with a three-year lifetime has removed many of the logistic and reliability objections to devices requiring cryogenic support (Ref 7).

HTS helical transmission lines are a particularly attractive component for reducing the size of HF antenna systems. HTSC materials allow the dimensions of helical transmission lines and resonators to be reduced by a factor of over 100. For instance 1/2 wavelength lines at 15 MHz are only 10 cm (4 in.) long. Dielectric loading of HTSC delay lines will provide an other order of magnitude reduction in delay line length. These lines can be used to make lossless multistub matching networks for small antennas, lossless channelized filter banks frequencies, and a number of other components which will greatly reduce the size of and increase the performance of HF antenna systems.

#### 4.1 Role of HTS in Electrically Small Antennas

The reduction of antenna size in the HF frequency range is important to improve mobility and decrease observability of HF systems. Superconductivity can contribute to this goal in at least three distinct ways:

- A. Lossless filters can be inserted between the antenna and the high impedance amplifier in electrically small antenna applications to reduce the dynamic range requirements and improve system performance
- B. Optimal bandwidth antenna elements a factor of 10 more bandwidth than an equivalent size dipole can be realized with superconductors
- C. Slow wave superconducting helical delay lines make miniature lossless stub matching networks for tuned electrically small antenna elements.

The simplest way to obtain a large bandwidth with a small receiving antenna is to operate in an essentially unmatched condition with a high impedance amplifier directly mounted on the antenna. Channelizing the receiver with sharp cutoff filters would solve a long standing dynamic range problem for this receiver configuration. However, existing filters have too much insertion loss in an already compromised signal to noise ratio situation. However, superconducting filters have the combination of high Q and

essentially low insertion loss that is required to solve the HF dynamic range problem.

#### 4.2 Use of HTS Electrically Small Antennas with Matched Loads

Superconductors can also be employed to make electrically small antennas for both transmit and receive applications. In addition, applications such as DF are of interest which can utilize more sensitivity than HF radar or communications receiver because of the differential mode of DF operation. Antennas operated with matched loads are required for these applications. Two basic problems arise with electrically small matched antennas. First, the bandwidth becomes impractically small as the electrical size of the antenna is reduced. Second, the antenna and especially the matching network dissipates much of the incident power as the electrical size of the antenna is reduced. Superconductivity can solve both of these problems.

Many recent evaluations of antenna applications of HTS have come to the conclusion that there is no benefit in making the antenna itself out of superconductors. This is certainly true for dipoles if only the radiation properties are considered. However, compared to the dipole antenna a factor of somewhat more than 10 in bandwidth can be achieved by using a more complex antenna. Bandwidth is the most serious penalty which must be overcome to achieve a practical electrically small antenna. These optimal bandwidth antennas have ohmic losses that rival the radiation resistance resulting in as much power dissipated ohmically in the antenna as is radiated. In

addition, the heat loss from the cryogenic section to the outside world via an exposed antenna is not tolerable from a systems point of view. In transmitter applications, the dissipation of even a small fraction of the total power in the antenna has a great effect on cryogenic requirements. Therefore, from a practical point of view the antenna element itself must be superconducting.

The major source of inefficiency in small antennas is dissipation in the matching networks. The large size of the required networks when they are made with normal metal conductors is also a problem. Superconducting slow wave delay lines solve both of these deficiencies. In addition, the superconducting circuits (an array of stagger tuned antennas) needed to achieve a usable bandwidth can be engineered without incurring intolerable losses.

In a contract with Foster Miller, Inc. of Waltham MA., the most problematic part of the research proposed in that program was the construction of the HTSC coils for the slow wave transmission line needed for the miniature matching stubs. This was a major technical accomplishment and forms a first step in the application of HTS to reduction of antenna size in the HF frequency regime. Cera Nova as a subcontractor demonstrated melt textured wires with a critical current of 50,000 Amp/cm<sup>2</sup>. Although the simple sintered wire used for that work is not suitable for transmitter applications, the CPS superconductor melt textured wire has more than sufficient current carrying capacity for jamming applications. Melt texturing of YBaCuO HTSC wire is also an innovation which opens the way for transmitter applications.

#### 4.3 Size Reduction Goals

HTS materials provide an opportunity to reduce the length of a matched dipole antenna element by a factor of about 100 without losing efficiency. The arithmetic illustrating this point is shown in Table 1 where the length of a quarter wave antenna at the ends of the HF band are compared with two reduced size superconducting antennas. The matching network for the electrically small antennas adds an additional length which is coincidentally about the same length as the antenna itself. Recently, Foster Miller concluded analytically that delay lines of 800 n-sec delay per 0.3m can be built.

Ultimately, small HF antennas will be used for a number of applications such as mobile DF will follow from the insertion of HTSC components into antenna systems. We foresee individual antenna elements including signal conditioning elements weighing less than 35 Kg. These antennas could be deployed (by two soldiers) from the back of a small vehicle (for large base line arrays) and connected with fiberoptic cable to a central processing unit in the vehicle. Such a concept would greatly reduce the logistic complexity of deploying sensitive long base line arrays.

#### 4.4 HTS Bulk Wire Versus Thin Film Technology

One objective of the mentioned Foster Miller program was to apply the basic HTSC technology in the development of a two element DF array. The demonstration involved the use of HTS (bulk wire) technology which is especially suitable for low HF. The key element

in application of HTSC to HF antenna systems is slow wave structures which permit size reduction. Helical delay lines, for instance, can be substantially shortened. Most of the effort in HTS small antennas has focused on traditional HTS thin film microwave transmission lines such as microstrip lines with propagation velocities nearly equal to that of light. The theme of the thin film studies was to compare the loss of standard copper transmission line structures and demonstrate greater efficiency with HTS.

In the HF regime much of the reduction in resistivity resulting from HTS can be used to achieve size reduction. Nonetheless, there is still an potential for improvement in efficiency.

The HTS bulk technology is most effective at lower frequencies (at least HF and perhaps lower frequencies). The materials already developed have sufficient advantage compared to normal metals in the HF region to support applications. The bulk technology is suited for the fabrication of large structures (order of 0.1m) needed after size reduction (in the HF regime). Since the bulk wire structures are stand alone mechanical structures, they can be combined with dielectric material optimized for the specific application. The wire technology becomes less useful as the frequency of interest increases beyond 300 MHz where the coil fabrication is difficult due to the small size of the circuits and the reduction in resistance of HTSC relative to that of copper is no longer as great.

#### 5.0 OTHER ROLES OF HTS IN ELECTRONICS WARFARE.

The Electronics Countermeasures Laboratory of CECOM-SWD's Research and Technology Division is engaged in the development of effective, state of the art communications electronic warfare SIGINT equipment. This effort includes the conceptualization, development and testing of complete jamming systems designed to perform target signal reception and analysis and subsequent jamming at frequencies ranking from HF to UHF.

HTS technology has application in many and diverse components of EW jamming hardware. The receiver and the transmitter of a jammer contain analogue and digital signal processing and power handling electronics which will benefit significantly from HTS and can be easily packaged in a cryogenics environment. With smaller and lighter electronic components of nullified resistivity one expects immediate benefits on effective jamming power, speed, bandwidth, resolution and dynamic range.

Many components of a receiving/transmitting antenna of a jammer will also dramatically benefit with superconductivity. It has been proven practically that the coupler of a transmitting antenna is able to handle previously forbidden power levels when its components are made superconductive. Similarly, the delay lines used in a powerful pattern-steering array jamming antenna as well as in a precise direction finding antenna are significantly improved with HTS. In the case of jamming the advantage is more pronounced since ordinary, non superconductive delay lines are impedingly large, heavy and highly inefficient for they dissipate

too much ohmic power.

Future milestones in this planned technology are integrating components into systems, studying the feasibility of constructing a complex jamming subsystem, interfacing such components to existing hardware and producing signal processing algorithms and related software.

Further research by Foster Miller demonstrated that lossless filters can be inserted between the antenna and the high impedance amplifier in electrically small antennas to reduce dynamic range requirements and improve system performance. Additionally, with the use of lossless filters bandwidth can be improved by a factor of 10 over an equivalent size ordinary antenna. Finally, it was proven that HTS helical delay lines make excellent miniature lossless stub matching networks for electrically small tuned antenna elements.

BAA programs, such as an ongoing contract with TRW Incorporated of San Diego, Ca., are currently being planned to demonstrate enhanced radiation efficiency HTS antennas for ground based EW applications.

As TRW states, impedance matching networks are made so that the sum of the electrical lengths of the matching inductor and the antenna's radiator totals one quarter wavelength at the chosen frequency, however the tuner's ohmic losses  $R_t$  and ground return losses  $R_g$  reduce the potential value of the radiation efficiency  $N$ .

Efficiency  $N$  of a system is defined as the ratio of power output

to power input. It follows then that

$$N = \frac{R_r}{(R_g + R_n + R_r)}$$

Where  $R_g$  is the ground resistance,  $R_n$  is the impedance matching network's resistance and  $R_r$  is the antenna's radiation resistance.

The antenna efficiency can be increased from 25% to 80%, depending on the ground return losses. Although most ohmic losses will occur in the inductor, it is inexpensive to pack the entire matching network in a cryogenic environment for even better results.

The above arguments leads us to conclude that since a high Q matching circuit limits instantaneous bandwidth, a rapid antenna tuning device is required for matching impedances at a hopper's rate in a multi-threat ECM environment. Fortunately, such a device is now a reality with AEL's Agile HF antenna coupler. It makes use of rapid, high power switches to select the correct values of L and C. A 10 meter antenna operating at one (1.0) MHz can thus radiate at 63% efficiency.

Hypres Incorporated has been working in low and high temperature superconductivity basic research and in cooperation with Harris Corporation is applying its technology to communication and other systems. One of them is the production of low noise amplifiers and electrically small, low noise (non HF) antennas, together with the integration of HTS components into monolithic substrates with significant cost reduction.

Another Hypres Inc. project is an HTS multi-port mode forming network for an array antenna in the shape of a matrix has been developed. This is referred to as a Shelton-Butler matrix, it consists of hybrids and fixed phase filters. A 2-port Butler matrix shall be demonstrated in a future contract. Together with Foster Miller, a tunable circular 4-element array antenna for DF uses has been conceptualized.

The above descriptions of US Army Sigint EW applications and related contracts is an active, open file which extends as stated at the time of the publication of this article. Changes and additions will follow as this emerging technology gains further momentum.

## 6.0 RECOMMENDED BIBLIOGRAPHY

This article discusses only selected issues of HTS applications in the RF field. The reader is invited to research in some of the literature below, which itself refers to further material.

1. High Temperature Superconductors and Their Applications in Passive Antenna Systems. Jeffrey T. Williams and Stuart A. Long. IEEE Antennas and Propagation, Aug. 1990.
2. Engineer's Guide to High Temperature Superconductivity. James D. Doss. Wiley-Interscience, 1989.
3. Fabrication and Evaluation of Superconducting Devices. Babbit, Lenzing and Rachlin (US Army LABCOM). Microwave journal. April 1991.
4. Towards Explaining Superconductivity. John Maddox. Nature, Vol. 344. April 5, 1990.
5. Superconductivity, a Future of Technological Excellence. Internal Report by the Du Pont Superconductivity Experimental Station, Wilmington, DE.
6. High Temperature Superconducting Small Antennal, Technology Briefing. Foster Miller, Inc., Waltham MA.
7. The High Temperature Superconductivity Space Experiment. Dennis C. Webb and M. Nisenoff. (Naval Research Lab). Microwave Journal, September 1991.
8. High Temperature Superconductors in High Frequency Fields. Fundamentals and Applications. Piel, Chaloupka and Muller, University of Wuppertal, D-56, Germany.
9. A Survey of Possible Passive Antenna Applications of HTS. Robert Dinger, Don Bowling and Anna Martin. IEEE Transactions on Microwaves. Vol. 39, No. 9. September 1991.
10. Measurements of Power Dependence Effects in High Temperature Superconductor Materials at Radio Frequencies. NRL Report. Apr, 91.
11. A Half Loop 5000 MHz Antenna Made From HTS YBCO and TBCCO. Don Bowling and Anna Martin; Robert Dinger. 1991 DARPA HTS Workshop, 30 Sept 1991, Seattle, Washington.
- Superconductive Antennas. R.C. Hansen. IEEE Transactions on Aerospace and Electronics Systems. Vol. 26, No. 2. March 1990.
12. The Effective Microwave Surface Impedance of High Tc thin Films. Klein, Chaloupka, Muller, Orbach and Piel. Bergische Universitat Wuppertal, West Germany.
13. Superconductors and MMICs in Microwave Systems, a Special Report. Kumar Agrarwal. Applied Microwaves. Spring 1992.
14. Passive Microwave Device Applications of High Tc Superconducting Thin Films. W. G. Lyons and R. S. Withers (MIT) Microwave Journal. November 1990.



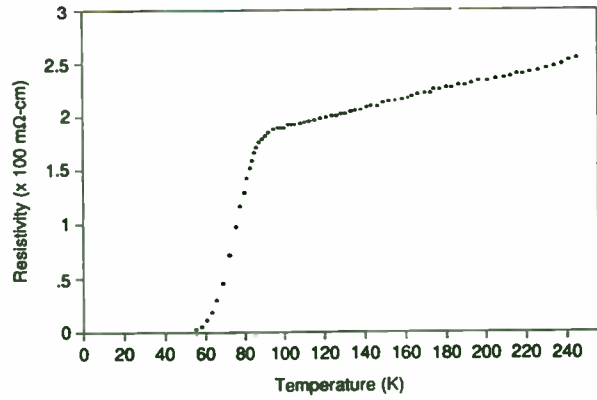


Figure 1. Surface Resistivity -vs- Temperature for the world's first sample for YBCO superconductors. (Courtesy, University of Alabama in Huntsville).

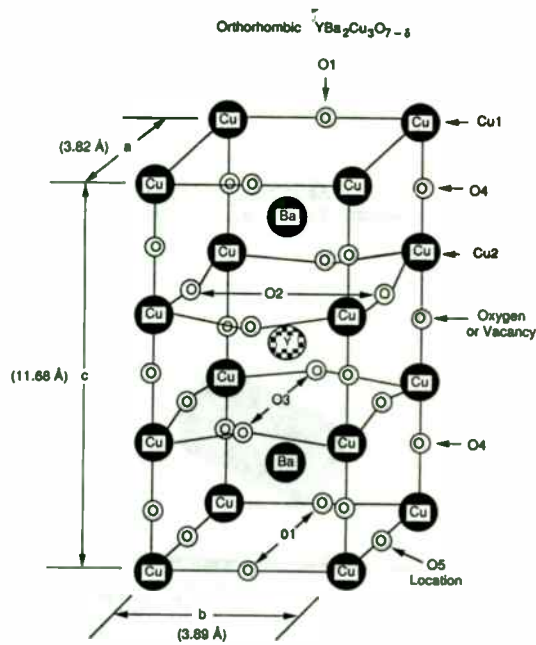


FIGURE 2. The molecular structure of the oxygen-deficient perovskite unit crystal  $YBa_2Cu_3O_{7-x}$ . (Courtesy James D. Doss, Ref. 2)

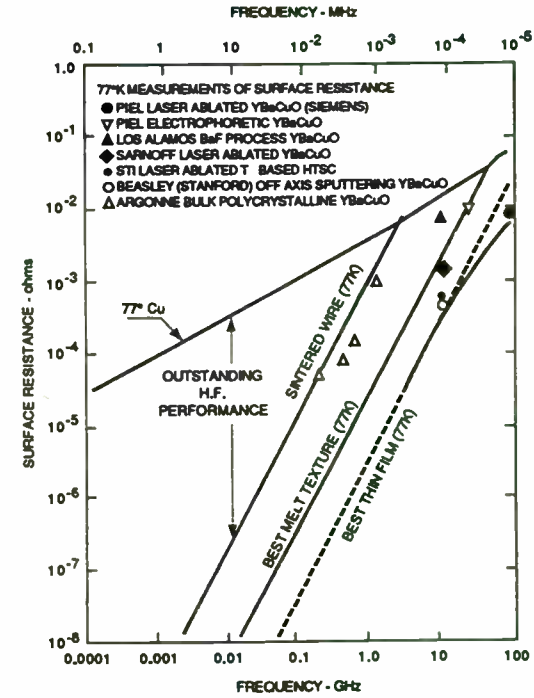


Figure 3. Experimental Data showing the dependence of  $R_s$  on Frequency. Note that best results occur in HF.

TABLE 1. Approximate combinations of current density ( $J$ ) and magnetic field ( $H$ ) required for typical applications of superconductors.

Application	$H$ (T)	$J$ ( $A/cm^2$ )
Interconnects	0.1	$5 \times 10^6$
AC Transmission Lines	0.2	$10^5 - 10^6$
Power Transformers	$0.3^a - 3^b$	$10^5$
DC Transmission lines	0.2	$2 \times 10^4$
SQUIDS	0.1	$2 \times 10^7$
SQMEs (Energy Storage)	$2.5^c - 5^d$	$5 \times 10^5$
Motors, Generators	$2.5^c - 5^d$	$4 \times 10^4 - 10^{5b}$
Magnetic Separation	2-5	$3 \times 10^6$
MAGLEV	5-6	$4 \times 10^4$
Fusion	10-15	$10^5$
Fault Current Limiters	> 5	$> 10^5$

Figure 4.

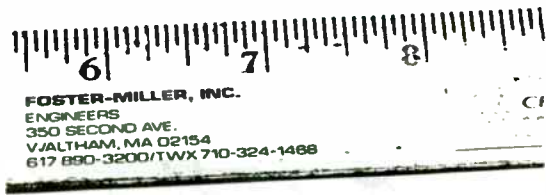


Figure 5. HTS coil used in the development of a delay line and high Q coupler. (Courtesy, Foster Miller).



Figure 6. Superconductive Archimedes Spiral used as an inductor in an impedance matching network. (courtesy, TRW)

## Tunable Circular Array for DF Applications

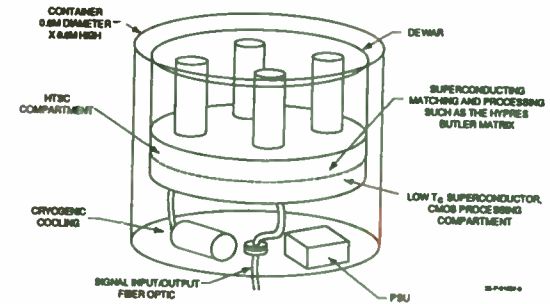


Figure 7. System concept for a phased array of antennas and Butler Matrix using HTS technology. (Courtesy, Foster Miller).

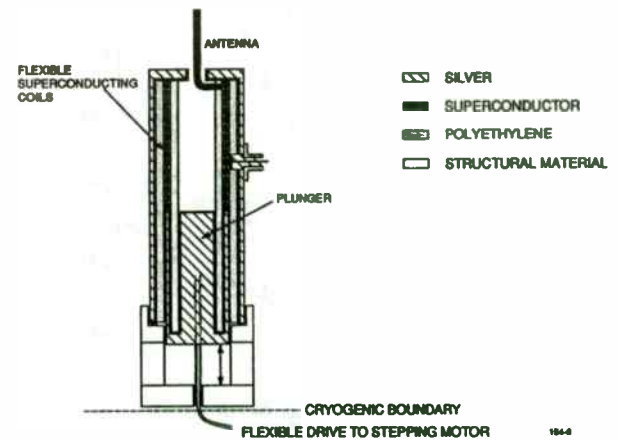


Figure 8. Design by Foster Miller Inc. for a super-conducting turnable antenna matching networks.

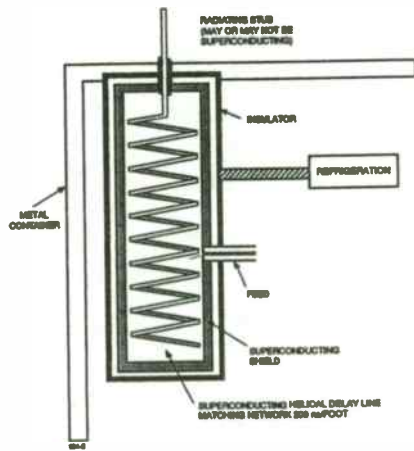


Figure 9. Design by Foster Miller Inc. for a tapped delay line and High Q Tanks made with HTS materials.

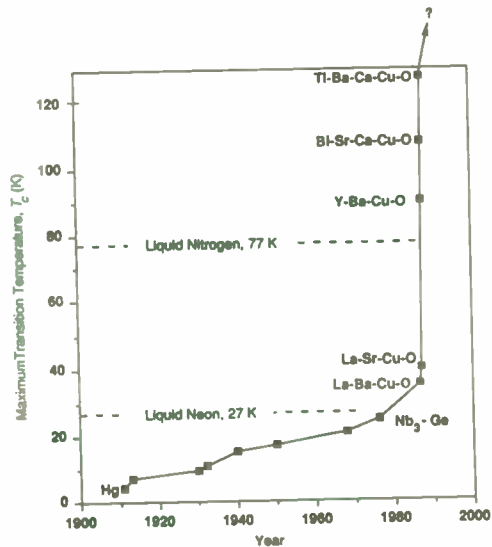
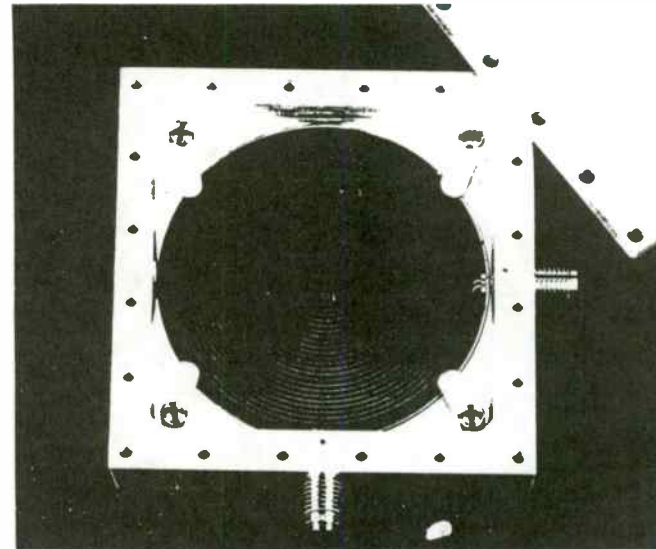


Figure 10. Progress made in HTS materials. (Courtesy, James D. Doss, Ref. 2)

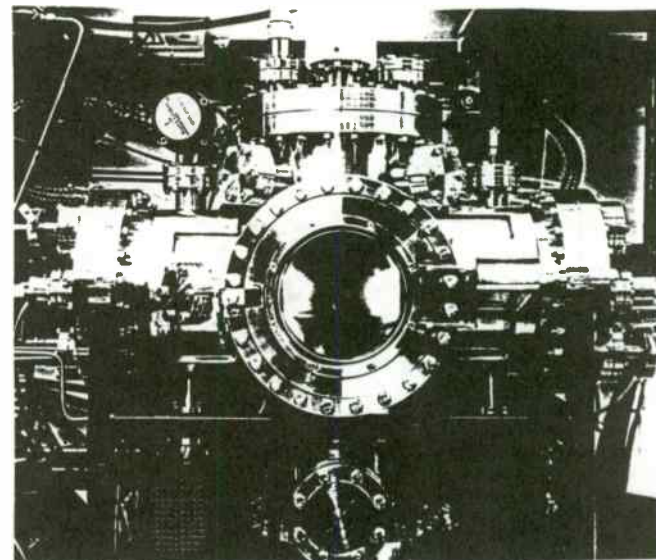


A 25-nanosecond capacitor delay line using a  $TiBaCaCuO$  HTS thin film on a 2-inch diameter aluminum substrate

Figure 11.



Du Pont  
Superconductivity  
Experimental Station  
E304/C116  
Wilmington, DE  
19880-0304



In-situ helium HTS thin film deposition (showing plasma discharge in a helium atmosphere)

Figure 12.



Du Pont  
Superconductivity  
Experimental Station  
E304/C116  
Wilmington, DE  
19880-0304



## Choosing the Optimum Synthesizer Architecture for Your Receiver Application

Larry W. Kimbrough  
Harris Government Communications System Division  
PO Box 91000, MS 25-5844  
Palm Bay , FL 32902

### Abstract

Synthesizer design is driven by 6 basic parameters; 2 are fixed, the other 4 are usually somewhat flexible depending on other System tradeoffs. The fixed requirements are Output Frequency and Frequency Step Size. The flexible parameters are ; Phase Noise, Output Spurious, Settling Time, and Size/Power. The relative flexibility of these 4 parameters will depend on the specific application and the overall System Requirements. For example in a battery powered hand-held Radio, the Size/Power is fixed and Phase Noise and Spurious Specs may simply be "as good as possible". But in a rack mounted Intercept Receiver, the Phase Noise may be spec'ed at state-of-the-art performance with no limits on Size/Power parameters. The complexity of the Synthesizer's implementation is related to the difficulty of each of the 4 flexible parameters. The 2 fixed parameters will typically dictate a complexity that will increase with increasing Output Frequency and decreasing Step Size. This paper takes a Synthesizer Spec and derives the performance for 6 basic synthesizer architectures, all of which meet the fixed requirements but meet the 4 flexible requirements according to implementation. This process may be applied to any Receiver application. For the purpose of comparison, the example used is an HF Receiver's Local Oscillator.

### HF Receiver Local Oscillator Spec

An HF Receiver has as its input, a 3 - 30 MHz input RF spectrum. A typical frequency plan that minimizes spurious in the first conversion and also has excellent

image rejection is an inverting upconversion to a 1st IF of 80 MHz. This 1st IF is then downconverted to a 2nd IF of 10.7 MHz, then again converted to a 455 KHz IF. The Synthesizer design will only encompass the 1st Local Oscillator(LO) . The other LO's are fixed frequency sources that will not be addressed. The 1st LO does all the frequency tuning and dominates the additive phase noise and spurious in the Receiver. Table 1 lists the 6 Synthesizer parameters and the Specs for the HF Receiver 1st LO.

Table 1 - Synthesizer Requirements

<i>Output Frequency</i>	<i>83 - 110 MHz (RF = 3 - 30 MHz, IF = 80 MHz)</i>
<i>Step Size</i>	<i>1 kHz</i>
<i>Phase Noise</i>	<i>Integrated 100 Hz- 3 KHz, 1 deg RMS 10 KHz offset = -90 dBc/Hz SSB</i>
<i>Spurious</i>	<i>&lt; -80 dBc (goal)</i>
<i>Settling Time</i>	<i>&lt; 5 msec from F1 to F2 tuning</i>
<i>Size/Power</i>	<i>TBD</i>

The majority of Synthesizers implementations are Indirect Architectures based on Phased Locked Loops (PLL). The PLL lends itself to moderate to superior performance with relatively less circuit complexity than a Direct Synthesizer Architecture. Applications which require extremely fast switching can best be met with a Direct Architecture. The classic Mix and Divide Direct Synthesizer [1] has its own performance niche that is not likely to change in the near future. Advances in Direct Digital Synthesizers (DDS) have made it practical to implement HF and VHF Synthesizers with a minimum of VLSI chips. The limitations of a DDS are spurious (-65 dBc) and output frequency (< 100 MHz). Their advantages are small step size , with no increase in circuitry, and ultra fast switching speed. The DDS is an excellent choice to implement the fine frequency section of a multi-loop Synthesizer.

### Single Loop PLL Solutions to the HF Receiver LO

The basic building block in an Indirect Synthesizer is a single Loop PLL. Figure 1 shows a PLL and all possible options to extend its basic performance. This diagram is also representative of the output Loop in a multi-Loop Architecture. The first step in designing a Synthesizer is to see if a single loop design will meet the necessary requirements. The output frequency will be met by the design. A VCO must be designed or purchased to tune the full output band. The step size will be equal to the reference frequency (Fref). The system Reference Frequency is typically a ultra-stable source such as a Crystal Oscillator (XCO) or even an Atomic standard ( i.e. Cesium Beam). The System Reference will usually be higher than the required reference of the PLL. A 10 MHz XCO is a readily available source from many vendors.

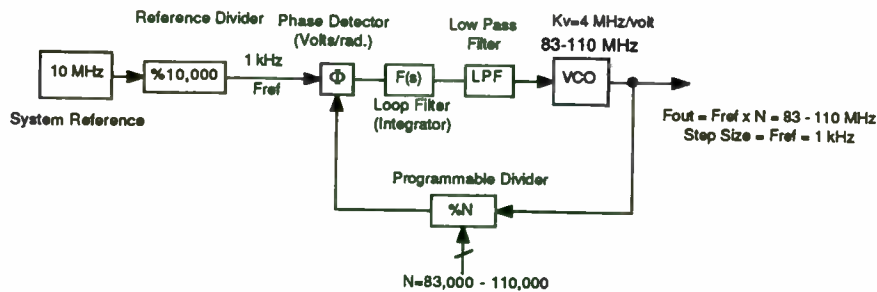


Figure 1. Basic PLL Configuration

Selection of the Loop Bandwidth (Loop BW) in a PLL will determine the output phase noise response. The Loop BW is also the dominant parameter that determines loops dynamics , settling time, and spurious performance. The Noise outside the Loop BW will be dominated by the VCO open loop phase noise. The noise inside the Loop BW will be dominated by Category I (Cat 1) noise [2]. Cat 1 noise is defined as all phase noise inputs to the loop that occur before the Loop integrator. They will consist of the Reference Noise, Phase Detector Noise, and the Logic noise floor. This noise will increase by a factor of  $20 \log N$  at the output frequency: where N is the division ratio of the PLL. Minimizing N will minimize the output phase noise at frequency offsets from the carrier.

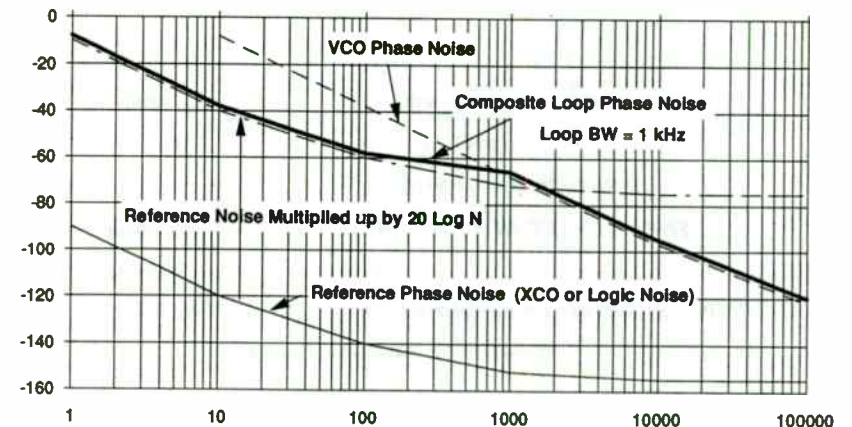


Figure 2. Selection of Optimum Loop BW for Phase Noise

The Loop BW selection is critical to Loop performance. There are conflicting trades that must be considered in its selection. The following factors must be considered when choosing the Loop BW.

- o Switching Speed is directly proportional to the inverse of the Loop BW. The smaller the Loop BW, the slower the switching time will become.
- o Optimum output phase noise is obtained by setting Loop BW at the point where VCO noise and multiplied up Category 1 noise sources intersect.
- o Loop BW will typically should not exceed 1/10 of Fref. Loop phase margin will degrade when this is violated.
- o Loop BW should be as low as possible, so that filtering can be added which attenuates reference sidebands, without affecting phase margin.

These factors must all be weighed and a compromise made between speed and noise, or phase margin and reference spurs, or some other parameter tradeoff in order to implement a working Loop.

### Designing the HF Receiver Single Loop Synthesizer

Figure 3. shows the block diagram of the Single Loop design. The 6 performance parameters and how they are analyzed are discussed below.

**Output Frequency:** The output frequency will be obtained by using a voltage controlled oscillator (VCO). The tuning sensitivity for the example is 4 MHz/volt. This will allow the complete 83-110 MHz band to be tuned by approximately a 7 volt span. The phase noise for all VCO's in this paper is determined by Leeson's model [3] of a VCO. The dominating parameter in a widely tuned VCO is the loaded Q. From experience a rule of thumb is that loaded Q of a VCO is 1/5 of the unloaded Q of the resonator circuit. The resonator Q is usually dominated by inductor Q, which in the VHF region will be 50 to 150. All examples use a loaded Q of 10. See Appendix for VCO design equation.

In order to reduce VCO phase noise, some of the following design changes may be implemented:

- o Raise Resonator Q
- o Use an active device with a low noise Figure and a low 1/f point.  
(JFETs are better than bipolars)
- o Break the VCO into several bands, this reduces tuning constant ( $K_{vco}$ ).
- o Noise on the tuning voltage will modulate the VCO, use low noise driving circuits (low noise op-amps, filtered power supplies).

**Step Size:** The step size of a PLL is equal to the reference frequency input to the phase detector (Fref). The system reference is a 10 MHz XCO. Divide it down to 1 KHz as the Fref input to the PLL.

**Phase Noise:** As stated previously the output phase noise close in to the carrier (inside the Loop BW) will be set by the Category 1 noise multiplied by a factor of  $20 \text{ Log}(N)$ . A source of noise that is often excluded but will dominate many Synthesizer implementations is the Logic noise floor of the programmable divider. The output of a digital divider or digital phase detector will have a phase noise profile. This data is rarely measured or documented by manufactureres of Synthesizer dividers. For large divisors of N, this noise will dominate the output noise. There are some documented cases of phase noise for different logic families [4]. In this example the 10 MHz Reference is divided to 1 kHz. This reduces the reference noise by  $20 \text{ Log}(10,000) = 80$  dB. Figure 3 shows that the phase noise of the 1 kHz Fref is below the Logic noise. The Logic noise is the dominant term for Cat 1 noise sources. The Logic noise for this example is taken from extrapolated data of synthesizer hardware designed with a CMOS programmable divider and digital phase/frequency detector. The ratio of Fout to Fref is greatest at  $110 \text{ MHz}/1\text{kHz} = 110,000$ . The Logic Noise will be multiplied up by  $20 \text{ Log}(110,000) = 101$  dB. By making a Bode plot of the open loop VCO noise and the multiplied up Category 1 noise, it becomes clear where the Loop BW should be set. It

should be set at the crossover point of VCO noise and multiplied up Cat 1 noise. A 100 Hz Loop BW will allow loop phase margin to be good, and also allow modest filtering to be added for reference spurious suppression.

**Spurious:** The spurs that dominate PLL output are the reference spurs that are caused by leakage of the Fref component through the phase detector. This voltage will modulate the VCO and produce spurs at offsets of Fref on the output carrier. The level of these spurs can be determined by methods documented by others [5], [6]. The basic design equation is given in the Appendix. For all examples, a digital Phase/Frequency Detector will be used. From measured hardware data, the output pulse width of the phase detector is approximately 20 nsec for typical Fref below 1 MHz. This is the differential pulse width in a dual output digital Phase/Frequency Detector that drives a differential active integrator Loop Filter. This will determine the magnitude of the Fref component that modulates the VCO. Spurious can be attenuated by placing a low pass filter between the Loop Integrator and the VCO. The limitations of this "fix" is that loop phase margin will degrade. It is useful to look at a family of filter curves of increasing poles that will all have the same phase shift at the 0 dB crossover frequency of the open loop. This is the point at which phase margin is measured on a Bode plot. The filter graph is shown in Figure A-1 in the Appendix. The curves were made for a phase shift of 20 degrees at the 0 dB point of the Loop. In a PLL, the phase margin should be typically set to 45 degrees for good margin. Thus the phase margin must be >65 degrees before filtering is added in order to maintain a good loop response. After the filtering is added, the phase margin, whatever it is, will be 20 degrees less. What should be noted from the curves is that adding more poles to the filter does not always help attenuate reference spurs. For example a 5 pole filter must have its corner frequency set at a larger offset from Loop BW in order to maintain the 20 degree phase shift than a 3 pole filter.

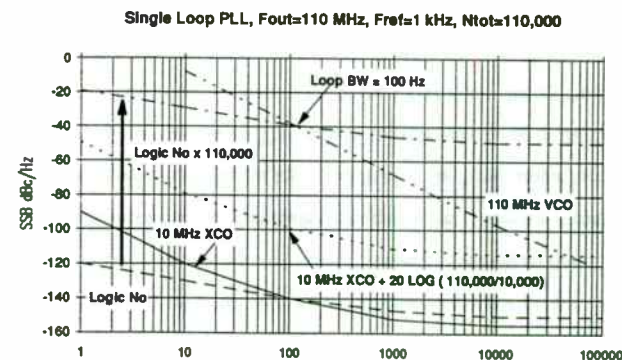
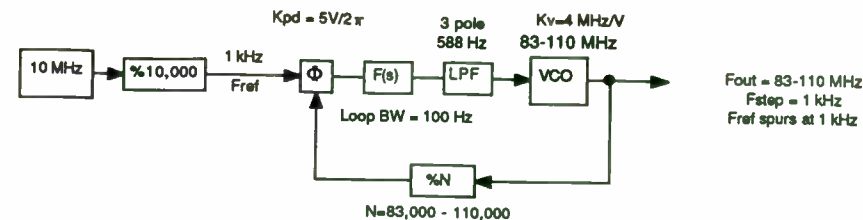


Figure 3. Single Loop Implementation of HF Rx Synthesizer

**Switching Time:** The switching time or settling time of a synthesizer is made of two components; the frequency slewing and the phase settling. The frequency slewing is a non-linear process that is not obtained by the classical steady state analysis of a control system. The phase settling is assumed to be a linear analysis that is described in many textbooks[7]. When changing frequencies the frequency slewing will occur until the frequency is within the capture range of the Loop. This can be approximated to be equal to the Loop BW. A general rule of thumb for approximating Switching Time from F1 to F2 in a Loop is :  $10 * 1/Loop\ BW(rad)$  . This is a conservative rule that has been verified in practice. For initial analysis of a Loop's switching speed, this is sufficient to make tradeoffs. For a more exact calculation see References [8] and [9].



**Size/Power:** In general the higher the output frequency, the higher the power. There are many Synthesizer IC's available which allow synthesis of all architectures presented in this paper. The complexity of a Synthesizer is ultimately related to the number of Loops that must be built.

### Adding a Divider to the Output of a Single Loop Synthesizer

Adding a divider to the output of a PLL will often allow a single Loop to have adequate performance, so that a multi-loop approach is not needed. Figure 4. shows this design. The phase noise of the Loop's Output will improve by  $20\text{Log}(N)$ , when it is divided by  $N$ . The availability of many high speed digital dividers make possible what once was impossible: generating a 83-110 MHz LO from a 664-880 MHz source (divided by 8 in this example). The advantages are:

- o Fref is increased ( to 8 KHz), this may allow a larger Loop BW, and a faster switching time.
- o The Fref spurs at the output of the divide by 8 occur at the 8 kHz rate, not at the step size of 1 kHz. This is perhaps the biggest advantage - filtering of Fref spurs is easier and more effective.
- o The Fref spurs and the Output phase noise will improve by  $20\text{Log}(8) = 18 \text{ dB}$ .

This example uses a VCO with a 32 MHz/volt tuning sensitivity. The phase noise for the VCO is again determined by Leeson's model [3]. The phase noise of the VCO will be worse because of the higher frequency, but the divide by 8 will improve it to the equivalent of a 110 MHz VCO. The size/power will increase due to the high current of high speed dividers.

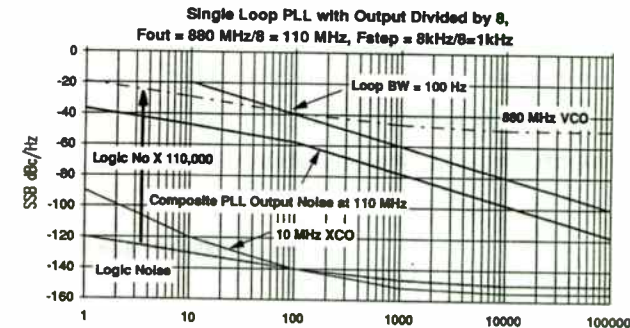
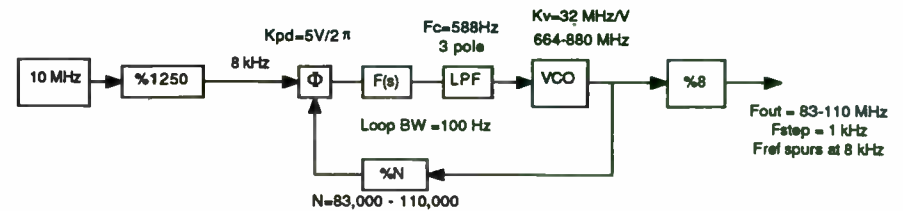


Figure 4. Single Loop Synthesizer with Divide by 8 on Output

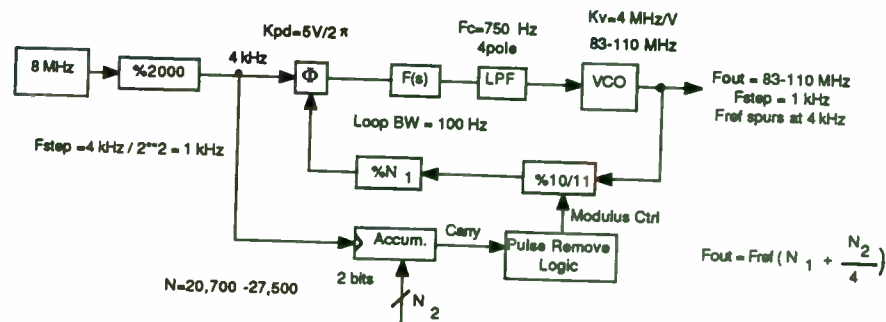
### Designing a Fractional N Single Loop Synthesizer

The Fractional N PLL [8], [10], and [11] is perhaps the most under utilized implementation in Synthesizer Architectures. It has many advantages such as faster switching time and better phase noise than an equivalent single Loop design. Its drawbacks are a much more complex design, which if not understood by the designer, usually is so formidable that it is not even considered as a candidate.

The operational theory of a Fractional N Loop is not difficult to understand if a few basic concepts are known. A normal PLL will move in step sizes that are equal to Fref. A fractional N Loop will generate step sizes that are multiples of a fraction of Fref. For example 2 bits of Fractional N will generate  $2^{**}2$  (or 4) fractional steps. An

accumulator keeps a running total of the fractional content so that an overflow occurs at the desired fractional count. For example an accumulator with two bits can move in 1/4 steps of Fref. Fractional step sizes are 1/4, 1/2, 3/4, and 1 times Fref. If the Accumulator is loaded with 01, and is clocked by Fref, the overflow will occur 1 out of every 4 cycles of Fref. This overflow controls a pulse swallower, which in practice is a dual modulus divider such as a divide by 10/11. An overflow pulse is deleted from the pulse swallowing divider which causes the division to decrease by 1, for one cycle of Fref. If the total divide by N of the Loop is 1000, and 1 of every 4 Fref cycles, N is 999; the divide sequence will be 1000, 1000, 1000, 999. If the VCO is at 100 MHz and Fref is 100 KHz, The output of the programmable divider will be 100 kHz, 100 kHz, 100 kHz, 100.1001 kHz. the average frequency over time is 100.025 kHz. This is a 1/4 step of the 100 kHz Fref.

The advantage of a Fractional N Loop is that small step sizes can be had without the penalty of a large divide by N, which will multiply up Category 1 Noise by a high number. A 2 bit Fractional N Divider will allow Ntotal to decrease by a factor of 4, which decreases the multiplied up noise of Fref and the Logic noise by 12 dB. The rate of Fref will increase so that switching speed can be faster, due to the larger Loop BW. The price for this scheme is spurs at the fractional rate. Filtering must be used, or some method of analog cancellation [11]. There are Synthesizers on the market that in a single Loop move in 10 Hz steps with a Fref of 100 kHz, with spurious of at least -70 dBc. The complexity of the spur cancellation circuitry is greater than rest of the synthesizer. It is composed of analog current cancelling schemes which are too complex to discuss. For the purposes of comparison and simplicity, the example used here will be 1/4 steps of Fref. This will only require a LPF to meet the Fref spur requirement.



$$F_{out} = F_{ref} \left( N_1 + \frac{N_2}{4} \right)$$

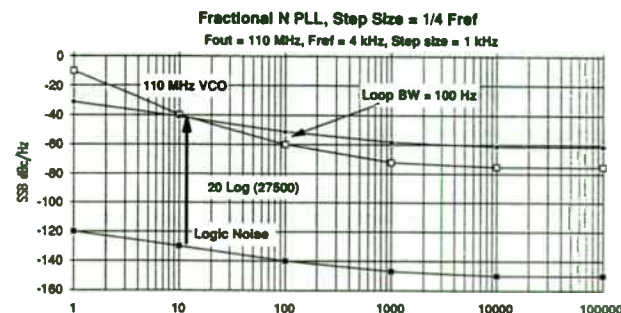


Figure 5. Fractional N Loop for HF Rx Synthesizer

### Fractional N Loop with an Output Divider

An Architecture that will minimize phase noise in a single Loop is a Fractional N Loop that is running at 8 times the desired the output, followed by a divide by 8. Figure 6 shows this block diagram. The Fref is 32 kHz. The fractional steps of 2 bits will make the step size at the output of the VCO, 8 kHz. The divide by 8 will reduce the step size to 1 kHz. The fractional Fref spurs occur at the 8 kHz rate, so the Loop BW can be relatively large at 1 kHz. The total divide by N is 27,500/8 = 3437.5 or a degradation of 71 dB. This is a 30 dB improvement over the single Loop with a 1 kHz Fref. Sometimes the Loop BW may be set smaller in order to get more filtering of the Fref and fractional sidebands.

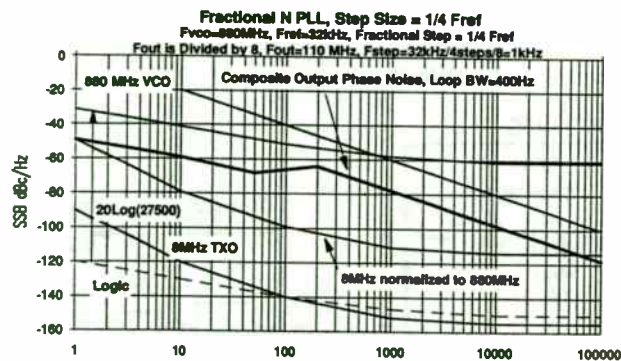
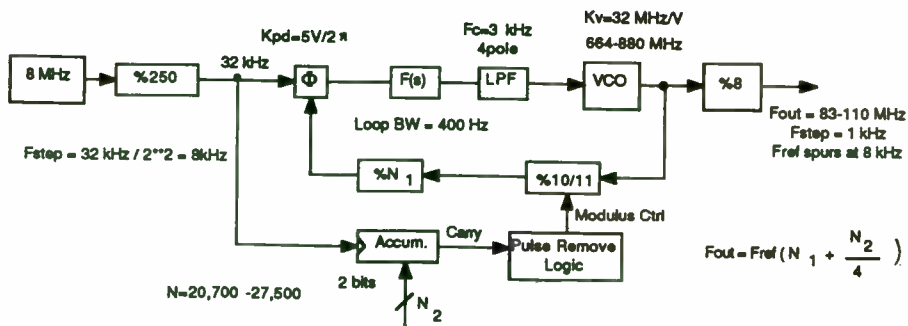


Figure 6. Fractional N Loop with Divide by 8 on Output

### Multi-Loop Solutions to the HF Receiver Synthesizer

When any of the three variable parameters of; phase noise, switching speed, or spurious cannot be met with a single Loop implementation, a multi-Loop Architecture is used. There are two basic output configurations of a Multi-Loop design. Figure 7. shows the two options. The choice is driven by ease of implementation as well as performance issues. The two methods are the Triple Loop method which uses a Summing Loop, and the Dual Loop method. The inputs to both configurations can be other multi-Loop designs, but in their simplest form are other single Loop PLL's. The Sum Loop will sum two separate Loops or frequencies to generate an output frequency.

Both configurations contain a mixing process which introduces the step sizes from another Loop or synthesizer. The Sum Loop has its fine tuning done at the reference input, and the coarse steps are injected into the loop through the mixer. The Dual Loop has its Fine and Coarse step injections at reverse positions from the Sum Loop. The basic features of the two methods of generating an output loop in a multi-loop design are listed below.

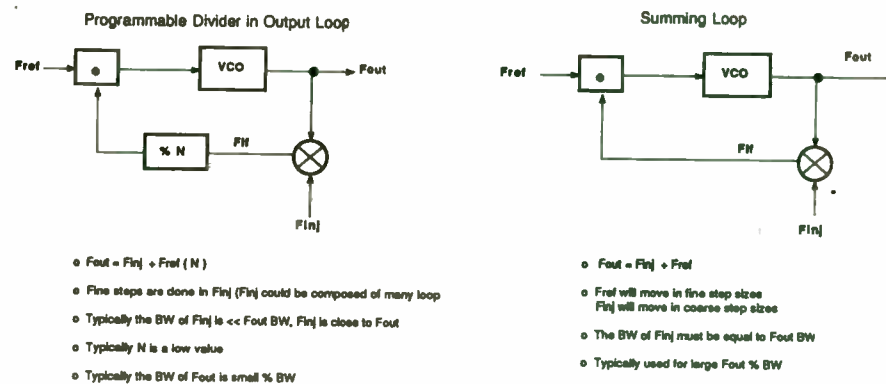


Figure 7. 2 Options of Output Loop for a Multi-Loop Synthesizer

When phase noise of a multi-Loop design must be determined, it is best determined by a logical progression of analyzing each loop separately. Each individual Loop's noise contribution is determined, then it becomes an input to the output Loop. Figure A-2 in the Appendix shows all possible noise source inputs and the transfer function that modifies their power spectral density as they appear at the output.

### Triple Loop Synthesizer

The Triple Loop design is shown in Figure 8. It is composed of ; a Fine Frequency Loop ( 1kHz steps in a 1 MHz bandwidth), a Coarse Frequency Loop ( 1 MHz steps in a 27 MHz bandwidth), and a Summing Loop. The Fine Loop is divided by 100 to realize 1 kHz steps.

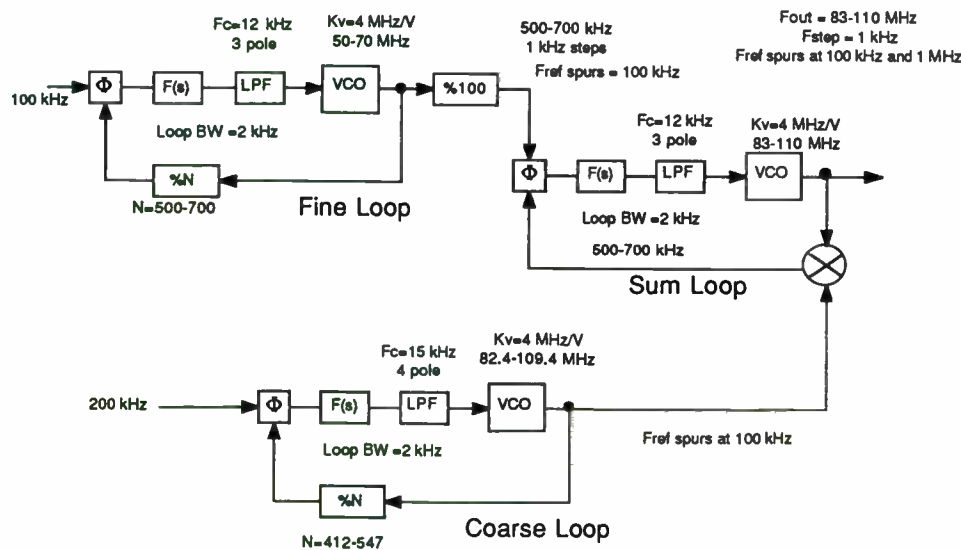


Figure 8. Triple Loop Synthesizer for HF Receiver

The dominant noise term in a Triple Loop design is usually the Coarse Loop noise. The Fine Loop noise is decreased by a large division number that is needed to create the Fine step size. There is no divider noise that is multiplied up to the output in the Sum Loop, so any noise from the Coarse Loop that falls inside the Loop BW of the Sum Loop will be present at the final output. The phase noise of the Fine Loop is negligible. The logic noise of the Sum Loop phase detector is below that of the Coarse Loop. The Coarse Loop noise is passed directly to the Sum Loop's output since the divide ratio of the Sum Loop is 1.

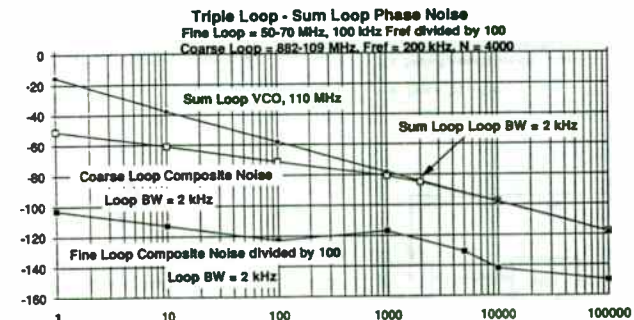


Figure 9. Phase Noise of Triple Loop Synthesizer

The spurious in the Sum Loop will be dominated by Four major contributors:

- o Fref spurs from the Fine Loop
- o Fref spurs from the Coarse Loop
- o Fref spurs from the Sum Loop
- o VCO leakage from the Coarse Loop (same offset as the Fref of Sum Loop)

The spurs caused by the mixer (  $M \times N$  products ) will not be a factor if the IF is low compared to the RF and LO frequencies. A low pass filter is needed to knock down mixer leakage and higher order spurs such as  $1 \times 2$  and  $2 \times 1$  products. The selection of the Loop BW of the Sum Loop will determine the magnitude of each of these spurs. The Fref spurs of the Fine and Coarse Loop do not need to be -80 dBc. The Fine Loop spurs will be attenuated by an additional  $20 \log_{10} 100 = -40$  dB by the divider on its output. The Fine Loop spurs will be at 100 kHz as will the Coarse Loop spurs. The level of the spurs on the Coarse Loop and Fine Loop outputs will be at the same level in the Sum Loop output minus any filtering in the Sum Loop VCO tune line. The Sum Loop Loop BW is 2 kHz, which allows a 3 pole LPF to be added at 12 kHz. The 12 kHz LPF will attenuate the Coarse Loop and Fine Loop Fref spurs of 100 kHz by at least an additional 50 dB. The level of the spurs in either of the tuning loops needs only to be 30 dB.

The settling time of a triple Loop is a composite of the three loops. The Sum Loop cannot settle out to its final frequency until both the Fine and Coarse Loops are their final frequency. Sometimes it is advantageous to pre-tune or steer the Sum Loop VCO close to its final frequency in order to decrease synthesizer settling time.

#### Triple Loop ( Summing Loop) Summary:

- o Fine Resolution is tunable and injected directly into Phase Detector.
- o Coarse Resolution Loop is mixed with the Sum Loop VCO
- o Synthesizer Output Frequency is the sum of Fine and Coarse Loops.
- o Coarse Loop will have to tune same % BW as Sum Loop VCO. Often the same VCO design is used for both Loops.
- o Phase Noise of Coarse Loop is transferred directly to Sum Loop's Noise, inside the Loop BW. The same goes for the Fine Loop Noise.
- o The IF filtering can be a simple LPF, since LO and RF leakage terms are far removed from the IF

#### Dual Loop Synthesizer

The Dual Loop design is shown in Figure 10. The fine steps are generated in a 100 kHz reference Loop running at 300-400 MHz, which is divided by 100 to obtain 1 kHz steps. The divide by 100 will lower the phase noise of this Loop by 40 dB. This will offset the large divisor number of 3000-4000. The output of the Fine Step Loop is mixed with a 70 MHz signal that is derived by multiplying the 10 MHz XCO. The phase noise of the 73-74 MHz signal is the sum of the noise powers of the two inputs. The implementation of the Fine Step Loop could take on many different forms. It could be a DDS at 3-4 MHz, that is mixed with the 70 MHz. It could also be a Fractional N Loop operating at 73-74 MHz with a 4 kHz Fref and 1/4 fractional steps. The Output Loop will move in 1 MHz steps, and downconverts the 73-74 MHz signal to a lower IF. A programmable divider will move the IF in 1 MHz steps. The frequency plan of the conversion in the loop should be done so that spurious mixer products are minimized.

The spurious in the Dual Loop are caused by the same process as those in the Triple Loop; the Fref spurs from both Loops, and the mixer spurs. There is another source of spurs that are unique to a Loop such as the Dual Loop that contain a mixer whose output goes into a digital divider. This will occur when spurious signals at the mixer output are not a multiple of the Fref of the Output Loop. A divider's output is at the Fref rate. In the frequency domain there are spectral lines at all multiples of this Fref rate. These lines will alias any spur present at the input to the divider down to baseband or down to a frequency below the Fref rate. This will show up as a spur on the Loop's output at the same level as the spur level in dBc at the input to the divider. Minimizing mixer Output spurs must be a priority in this design.

The settling time is a composite of the two Loops settling time. The Fine Step Loop will settle to its final frequency before the Output Loop will slew to its final frequency.

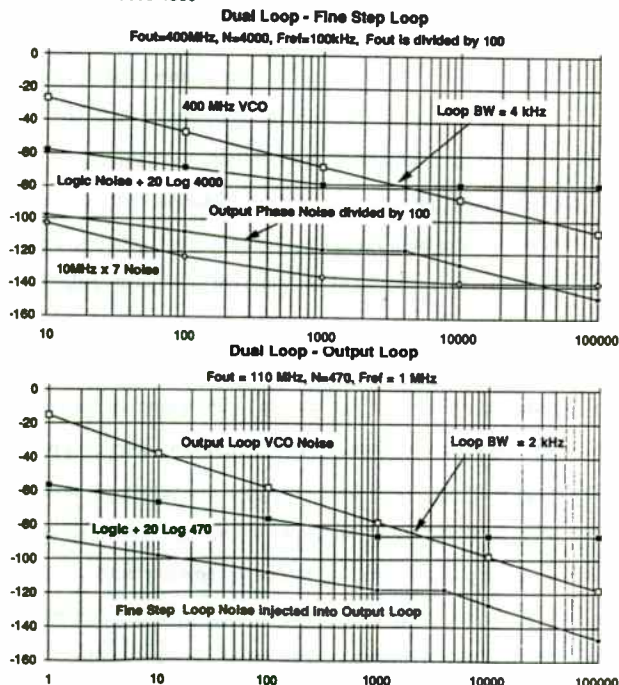
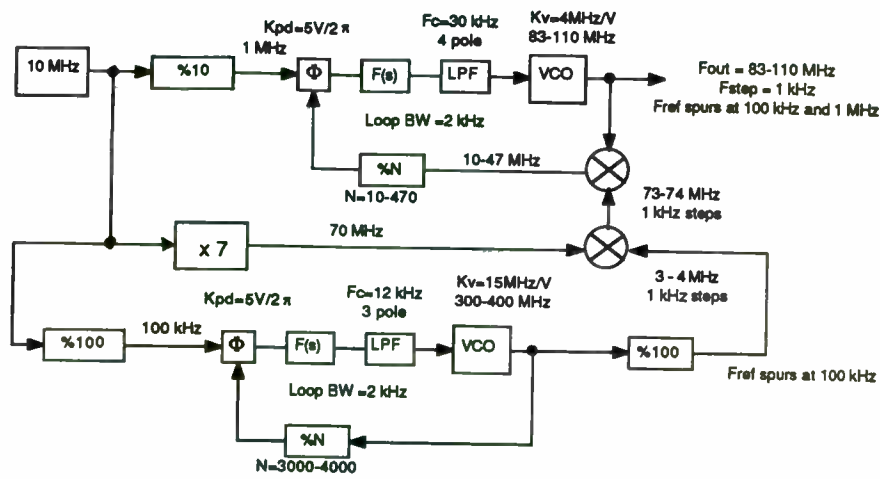


Figure 10. Dual Loop Synthesizer for HF Receiver

The complexity of the Dual Loop versus the Triple Loop will vary according to implementation. For the example presented, the Dual Loop contains a multiplier, a mixer, and a BPF versus another PLL in the Triple Loop design.

#### Dual Loop Summary:

- o In its most simple form, the Fref is fixed and performs the Coarse Steps and the injection into the Output Loop is a narrow band Fine step Loop.
- o The Fine steps may be generated by an unconverted DDS, another PLL, or another multi-loop design.
- o The BW of the injection Loop is narrow.
- o Filtering of the IF in the main Loop is typically a BPF.
- o Phase Noise of the Main Loop will contain a multiplied component of the Logic Noise of  $20\text{Log}N$ .

#### Synthesizer Tradeoff of the 6 Candidates

The 6 Synthesizer Candidates were analyzed so that the 3 variable parameters of phase noise, spurs, and settling time are compared. A computer program developed at Harris was used, that takes all noise sources in Power Spectral Density and all loop Transfer functions, and will calculate phase noise and phase margins. The spurious were estimated from the methods in [6], and the equations in the Appendix. For the Multi-Loop designs the mixer spurious analysis was omitted. The settling times were analyzed using the rule of thumb that  $T_{\text{settle}} = 10(1/\text{LoopBW})$ .

The Bode plots of phase noise of each architecture have been shown with each block diagram. The Logic noise floors is -140 dBc with an increasing 1/F slope at 1 kHz. the VCO noise in all cases is based on Leeson's model;  $\text{SNR} = 140 \text{ dB}$ ,  $Q_1=10$ ,  $1/F = 1 \text{ kHz}$ . These plots are all straight line approximations that are done manually. This type

of manual analysis is usually sufficient to make architecture trades. The results in Table 1 are exact results from the CAE PLL program.

### Summary and Conclusions

The Synthesizer performance can be the limiting factor in many areas of System Performance, such as adjacent channel rejection, dynamic range, and settling time. In order to optimize performance and minimize Size and complexity in a Receiver's Synthesizer, a logical progression of performance analysis of all candidate architectures should be performed. Manual Bode plots of phase noise are often a good starting point to grasp the tradeoffs involved in selecting a design. The rules of thumb for settling time and Fref spurious also offer a quick evaluation of one design versus another. After architecture analysis is complete, a trade study will point to the optimum Synthesizer performance.

The results of each architecture's performance is summarized in Table 1. In each Loop the spurs are calculated before and after a LPF. The attenuation curves in Figure A1. are used to determine filter order. The tradeoff matrix of Table 1 show that in order to meet all performance goals, the Triple Loop architecture must be used. The Fractional N Loop with a divide by 8 on its output provides good performance in all areas except integrated noise. The methods presented in this paper can be applied towards any Receiver application in order to determine the optimum Synthesizer Architecture for given performance requirements.

### References

1. Robert Papaieck and Robert Coe, New Techniques Yield Superior Frequency Synthesis at Lower Cost, EDN, October 20, 1975
2. Larry Martin, Noise-property analysis enhances PLL designs, EDN September 16, 1981
3. Dieter Scherer, Today's Lesson - Learn About Low-Noise Design Part 1, Microwaves, April 1979
4. Mark R. McClure, Residual Phase Noise of Digital Frequency Dividers, Microwave Journal March 1992
5. Mark O'Leary, Practical Approach Augurs PLL Noise in RF Synthesizers, Microwaves and RF, September 1987
6. James W. Maben, Design and Analysis of Fourth and Fifth Order Indirect Synthesizer Loops, Proceedings RF Expo East, November 10-12, 1986
7. Vadim Manassewitsch, Frequency Synthesizers: Theory and Design, John Wiley and Sons, New York 1976
8. William F. Egan, Frequency Synthesis by Phase Lock, Krieger Publishing Company 1990, Malabar FL
9. Roland E. Best, Phase -Locked Loops, Theory, Design, and Applications, McGraw-Hill 1984
10. Excerpts from Hewlett Packard 3335A Signal Generator Technical Manual
11. Roland Hassum, The Common Denominator in Fractional N, Microwaves and RF June 1984

Table 1 Performance Matrix of 6 Candidate Architectures													
	VCO Fo (MHz)	Fref	Loop BW	Phase Margin	Fref Spurs (dBc)	LPF poles/Fc	Fref atten. of LPF	Fref atten. of Div.	Output Fref Spurs	Tsettle	Output Tsettle (msec)	Output L(f) 100Hz-3kHz	Output L(f) 10 kHz offset SSB dBc/Hz
<b>Spec Req't</b>				45 deg					-80		5 msec	<1 degRMS	>-90 dBc/Hz
<b>Single Loop</b>	83-110	1 kHz	100 Hz	58 deg	-20	3pole,588Hz	-15	N/A	-35	16 msec	16 msec	10.3 deg	-94
<b>Single Loop with %8</b>	664-880	8 kHz	100 Hz	74 deg	-2	3pole,588Hz	-68	-18	-88	16 msec	16 msec	5.4 deg	-93
<b>Fract N Loop</b>	83-110	4 kHz	100 Hz	72deg	-8	4pole,750Hz	-14	N/A	-22	16 msec	16 msec	5.8deg	-93
<b>Fract N Loop with %8</b>	664-880	32 kHz	400 Hz	74 deg	10	4pole,3kHz	-34	-18	-42	4 msec	4 msec	2.2 deg	-92
<b>Dual Loop</b>													
Fine Loop	300-400	100 kHz	4 kHz	69 deg	-8.5	3pole,12kHz	-56	-40	>-100	0.4msec		0.04 deg	-117
Output Loop	83-110	100 kHz	2 kHz	76 deg	-20	4pole,30kHz	-42	N/A	-62	0.8msec	1.2ms	0.6 deg	-94
<b>Triple Loop</b>													
Fine Loop	50-70	100 kHz	2 kHz	72 deg	-8.5	3pole,12kHz	-19	-40	-67.5	0.8msec		0.01 deg	-130
Coarse Loop	82.4-109.4	100 kHz	2 kHz	74 deg	-20	4pole,15kHz	-42	N/A	-62	0.8msec		0.63 deg	-92
Sum Loop	83-110	500-700	2 kHz	72 deg	-20	3pole,12kHz	-80	N/A	-100	0.8msec	2.4ms	0.66 deg	-92



## Appendix

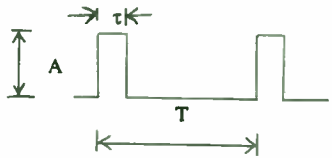
### Leeson's model for phase noise of an Oscillator

$$L(f_m) = \frac{1}{2} \left[ 1 + \frac{1}{f_m^2} \left( \frac{f_o}{2Q_1} \right)^2 \right] \frac{FkT}{P_s} \left( 1 + \frac{f_c}{f_m} \right)$$

- $f_m$  = offset frequency from carrier
- $f_o$  = frequency of oscillator
- $Q_1$  = loaded Q of resonator
- $FkT$  = thermal noise floor + oscillator noise figure, this determines wideband SNR of oscillator which is  $P_s/FkT$ .
- $P_s$  = Output power of oscillator
- $f_c$  = corner frequency of flicker characteristic of active device (i.e. 1 kHz)

### Approximation of Fref spurs in a PLL

Phase Detector Output



$\tau$  = pulse width of PD out  
 $T$  = period of Fref  
 $A$  = amplitude of PD out

$$V_{out} \text{ of Fref component} = \frac{2A}{\pi} \left[ \sin \frac{\pi \tau}{2T} \right]$$

$$V_{in} \text{ to VCO} = V_{out} * \text{Loop Filter(s)} * \text{LPF(s)}$$

$$\text{VCO spurs} = 20 \text{ Log} \left( \frac{V_{in}(\text{volts}) * K_{vco}(\text{Hz/volt})}{2 * f_m(\text{Hz})} \right) \text{ dBc}$$

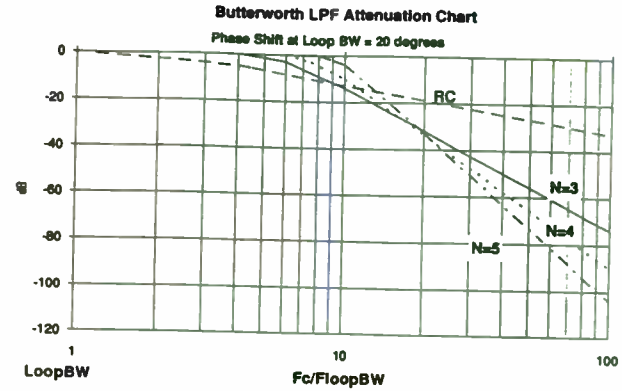


Figure A1. Attenuation Curves for LPF with  $\phi$  shift =  $20^\circ$  at LoopBW

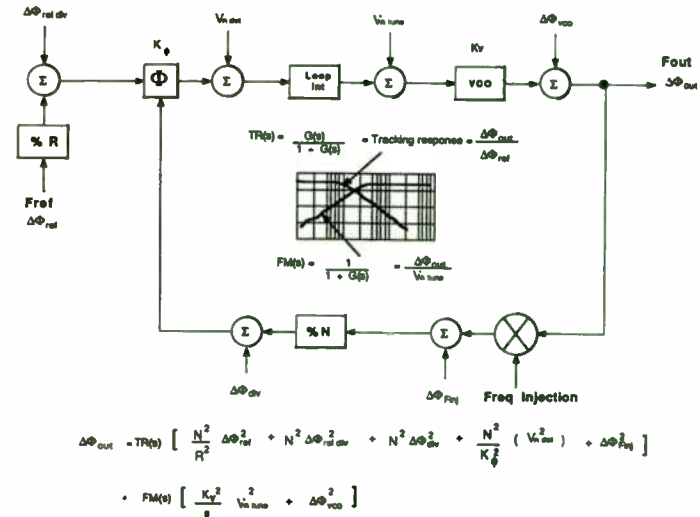


Figure A2. Contribution of Noise Sources in a PLL to Output Phase Noise



IF Transversal Filtering:  
Applications and Implementation

Richard D. Roberts, Ph.D.  
Harris Corporation  
Electronic Systems Sector  
Government Communication Systems Division  
Melbourne, Florida

Abstract

IF transversal filtering, consisting of a periodically tapped delay line, has a wide variety of uses such as gain slope compensation or adaptive equalization of digital signals. At Harris Corporation's Government Communication Systems Division, adaptive IF filtering and data equalization techniques have been implemented at IF frequencies of 6 GHz with QPSK baud rates in excess of 1 Gbaud. This paper looks at a few applications and implementations of IF transversal filters.

Introduction

Signal processing via samples of a time domain waveform is often viewed as belonging to the arena of digital signal processing (DSP). However, we can sample via tapped delay lines and process these samples in an analog manner as long as we are mindful of the sampling constraints. We can call this type of signal processing "analog signal processing" - ASP. This paper describes an analog sampling structure called the transversal filter. In particular, we will be looking at IF implementation of this structure.

In section 1 we will heuristically derive the characteristics of a sampled data system and show the time and frequency domain responses. In section 2 we apply these concepts to doing some nonadaptive ASP and show the relationship between the forward and inverse discrete Fourier transform (DFT and IDFT). In section 3 we show the real power of transversal filtering by letting the filtering process be adaptive, that is the filter response changes dynamically to optimize a particular parameter. Finally, in section 4 we cover a few implementation considerations.

I. The IF Transversal Filter - A Sampled Finite Impulse Response (FIR) System

The sampling of a time domain waveform at intervals of  $T$  seconds causes the frequency response to be repetitive as shown in Figure 1, with the frequency response repetition rate being determined by the sampling frequency  $1/T$ .

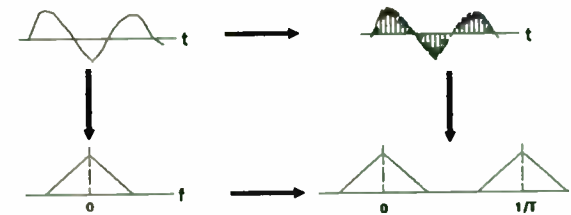


Figure 1  
A Sampled Waveform

We can see for the sampled signals to avoid overlapping of the frequency response the sampling frequency  $1/T$  must be greater than twice the single sided bandwidth of the baseband signal.

We can sample in an "analog manner" by using a tapped delay line. For example, suppose that we have a filter whose sampled impulse response yields the values as shown in Figure 2.

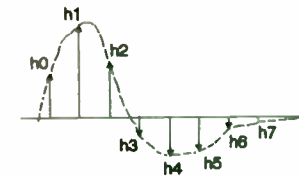


Figure 2  
A Sampled Filter Impulse Response

We can synthesize this same impulse response using a tapped delay line as shown in Figure 3. An impulse applied at the input to the transversal filter propagates down the delay line with a portion of the impulse tapped off at each tap. The tapped signals are each weighted by a value  $h_i$  and collected at the output of the filter. Clearly the impulse response will be the same as that of Figure 2.

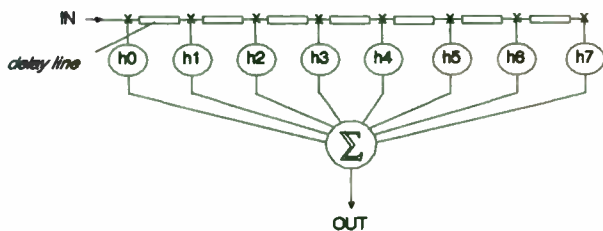


Figure 3  
Synthesized Impulse Response

The frequency response of this filter is periodic with the period determined by the delay of the delay lines. It may be used as either a lowpass filter or a bandpass filter by appropriate zonal prefiltering as shown in Figure 4. Such prefiltering is called antialiasing filtering because it prevents a response by the sampling filter at an "alias" frequency band.

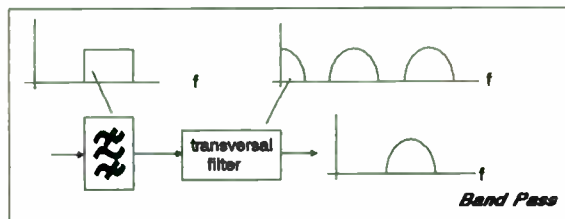
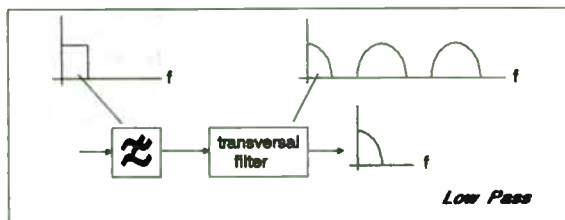


Figure 4  
Antialias Filtering

## II. Nonadaptive Applications

The determination of the value of the tap weights from the sampled impulse response is called the direct method. Using this technique, we describe the frequency response we desire in terms

of bands of attenuation separated by transition regions. We can then take the inverse discrete Fourier transform of this frequency response which yields the time domain weight series. A low pass filter example is shown in Figure 5.

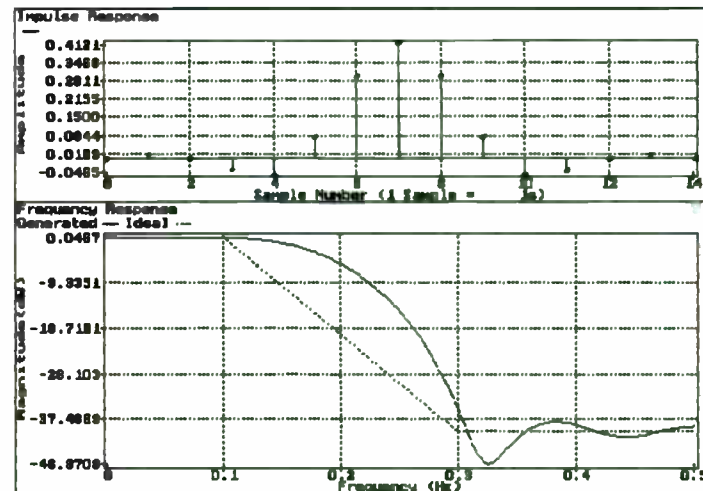


Figure 5  
Impulse Response and Corresponding Frequency Response

Because the impulse response is symmetric with respect to the center sample, it will have a linear phase response in the frequency domain. We can show this by considering the frequency response of this filter which is given as

$$H(\omega) = e^{j\omega n\tau} \left\{ h_0 + 2 \sum_{k=1}^n h_k \cos(k\omega\tau) \right\}.$$

The exponential term represents the phase response while the term in the brackets represents the amplitude response. The frequency response for this filter is plotted in Figure 6.

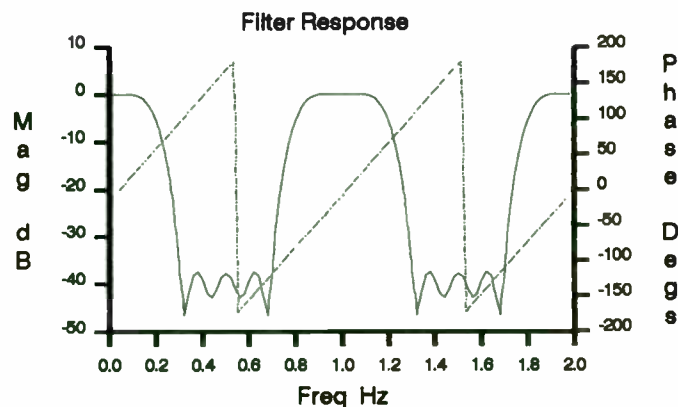


Figure 6  
Periodic Frequency Response

We can clearly see the potential for utilization as a bandpass filter (e.g. in the vicinity of 1.0 Hz). We can also see the periodic response which must be suppressed by a zonal anti-alias filter.

### III. Adaptive Applications

Up until now we have reviewed some applications of transversal filters that were perhaps interesting, but duplicated responses that could be obtained by simpler and cheaper lumped element implementations. However, the real power of transversal filtering is seen in adaptive applications. For example, consider the adaptive slope equalizer shown in the circuit of Figure 7.

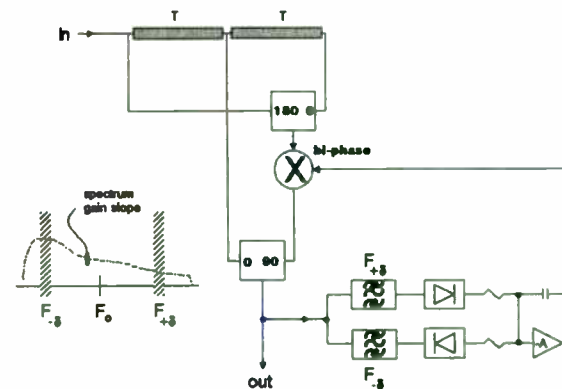


Figure 7  
Adaptive Slope Equalizer

The frequency response of the slope equalizer, with respect to the center tap, is given as

$$H(\omega) = 1.0 - ja(e^{j\omega T} - e^{-j\omega T}) = 1.0 - 2a\sin(\omega T)$$

where  $a$  is the bi-phase multiplication factor ( $-1.0 \leq a \leq 1.0$ ). (Note: we have ignored the linear phase term normally associated with this circuit). We will design our slope equalizer to work over a bandwidth of  $\pm\delta$  Hertz centered about a frequency of  $\omega_0$ . By letting  $\omega_0 T = \pi$  and forcing  $(\omega_0 - \omega) T < \pi/10$  radians we can approximate the frequency response as

$$H(\omega) \approx 1 + 2a\omega T = 1 + 2aF\pi/F_0$$

The value of  $a$  is determined by monitoring the difference between the spectral energy located near  $F_{+\delta}$  and  $F_{-\delta}$  and adjusting  $a$  in such a manner as to achieve energy balance. For example, with a value of  $a=0.1$ , the circuit could be compensating for the gain slope of a long coax line carrying a wideband data signal as shown in Figure 8.

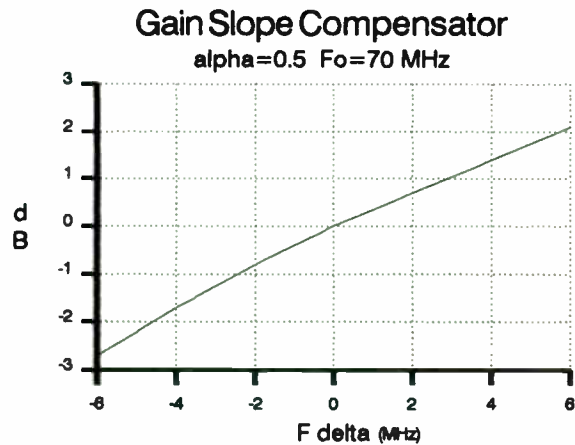


Figure 8  
Slope Compensator Response

Another application is adaptive data equalization. A multiple tap equalizer is shown in Figure 9. This type of equalizer was used at Harris Corporation GCSD for improving the bit error rate of a 2 Gbps QPSK data link by several orders of magnitude. It was implemented at an IF frequency of 6 GHz using microstrip delay lines and custom four quadrant weights.

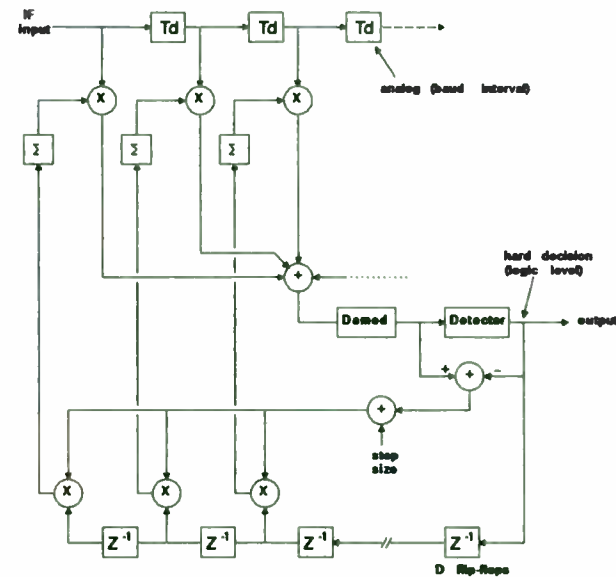


Figure 9  
Zero Forcing Equalizer

This equalizer is controlled by a zero forcing process that attempts to remove data intersymbol interference by destroying the correlation between future and past data symbols with respect to the current data symbol error voltage. This equalizer requires precise delay line values of  $T$  where  $1/T$  is the data symbol rate. The frequency response of this filter is exactly the inverse of the data channel filter. For example, if the channel frequency response is  $H_c(w)$ , then the frequency response of the equalizer is  $H_{eq}(w)=1/H_c(w)$ . If  $H_c(w)$  approaches zero at any frequency (e.g. a spectral notch) then  $H_{eq}(w)$  becomes ill defined at that frequency. This is one of the limitations of this type of adaptive filtering.

#### IV. Implementation Considerations

Transversal filtering using transmission delay lines is often an attractive alternative at very high data rates. At Harris Corporation's Government Communication Systems Division, we have been using adaptive IF filtering and data equalization techniques for baud rates in excess of 1 Gbaud. However, the designer needs to be aware of several implementation considerations.

One consideration concerns use of transversal filters with a digitally modulated carrier. Since the IF transversal equalizer uses delay lines in its construction, and the signal passing through the delay lines is a modulated RF carrier, there is a resulting modulation envelope phase rotation. For example, if we represent the modulated carrier as a quadriphase signal - represented by I/Q vectors - we can see that as the carrier propagates along the transversal filter, a rotation of the basis vectors will occur (as expected).

We can illustrate this via Figures 10 and 11. If the tap spacing is such that there is an integer relationship between the RF carrier rotation frequency and the data baud rate then no apparent carrier rotation occurs at each tap as shown in Figure 10.

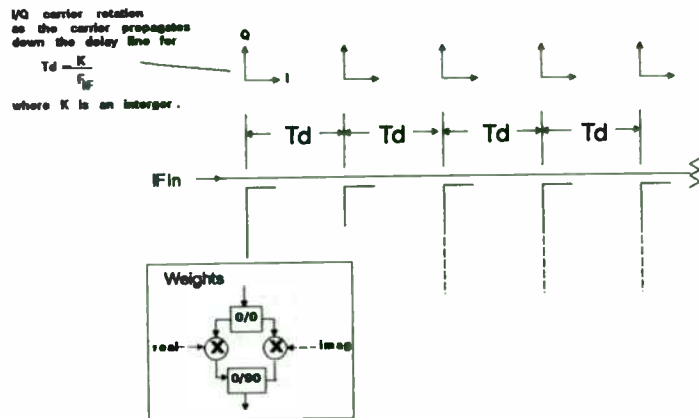


Figure 10  
Tap Spacing for No Rotation

On the other hand, if there is no particular relationship between the data rate and the carrier frequency as shown in Figure 11, then the carrier appears at each tap with a slightly different phase.

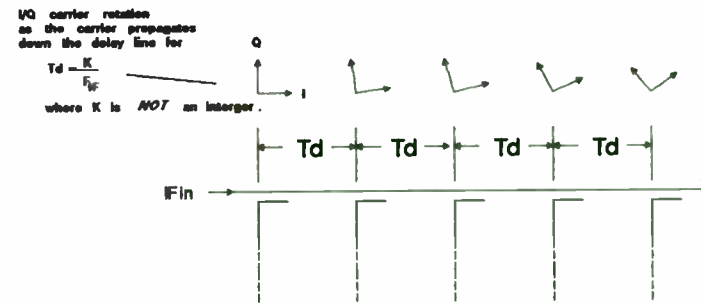


Figure 11  
Tap Rotation

In general, if the weights are complex (i.e. able to scale magnitude and phase) then the delay rotation can be compensated by the weights. However, if the controller is expecting a particular phase of the carrier with respect to the basis vectors, then problems could arise.

Another area of concern is VSWR reflections building up inside the filter structure and causing amplitude and phase ripples in the frequency response. This can be mitigated somewhat by trying to avoid tap spacings where reflections can start adding in phase. However, if this can not be avoided then careful use of isolating directional couplers can help diminish the problem. A practical implementation of a fractionally spaced IF equalizer is shown in Figure 12.

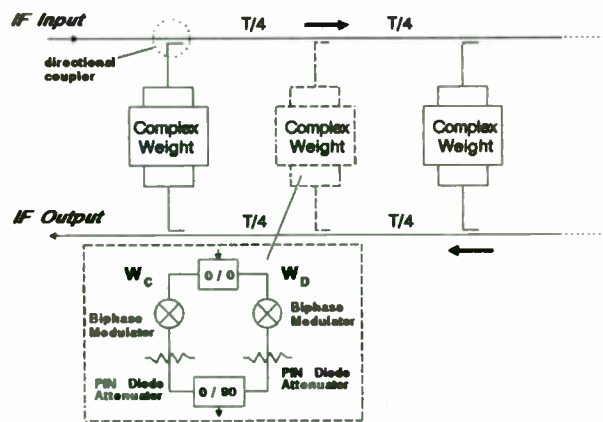


Figure 12  
An Equalizer Implementation

The circuit of Figure 12 is divided into two halves, an output and an input structure. This particular implementation had an IF frequency of 6 GHz with a symbol rate of about one gigabaud. The delay lines were implemented using microstrip transmission line with the taps spaced at  $T/4$  intervals (which when added onto the corresponding  $T/4$  delay on the other side of the circuit yields a  $T/2$  delay).

Directional couplers, formed from edge coupled quarter wavelength transmission line, were used to tap onto the delay line. This allowed a loose coupling value (typically less than -10 dB) so as not to load down the delay line. It also provided a large degree of signal isolation between the weights to prevent weight interaction.

Our final consideration is compensating for the increasing insertion loss as the signal propagates down the tapped delay line by tapering the directional coupler coupling values. The minimum value of directional coupler should be about -10 dB and will be located at the far end of the delay line. The coupler values then increase proportionally as we approach the signal input end of the delay line. This is illustrated in Figure 13 where we have assumed 1 dB loss per section of delay line.



Figure 13  
Tapered Value of Coupling

### Conclusions

In this paper we have discussed the fundamentals of transversal filtering. We started with showing the periodic frequency response of the filter due to the sampling effect. We showed how this periodic response can be used in our favor to construct linear phase bandpass filters. We then discussed adaptive filtering using an adaptive gain slope transversal compensator as an example. We next discussed some implementation considerations, including ways to mitigate VSWR effects via the use of directional couplers.

IF transversal filtering is a cousin to digital filtering and includes many of the unique and powerful features usually only associated with DSP.

### References

1. Proakis, J.G., *Digital Communications*, McGraw-Hill, New York, 1983.
2. Oppenheim, A.V., and Ronald W. Schaffer, *Digital Signal Processing*, Prentice-Hall, New Jersey, 1975.
3. *Monarch User's Manual*, Ver. 2.0, The Athena Group, Gainesville, FL.
4. Chamberlain, J.K., F.M. Clayton, H. Sari, and P. Vandamme, "Receiver Techniques for Microwave Digital Radio", *IEEE Communications Magazine*, November 1986.
5. Harris GCSD IR&D Project 4137, "A 2.4 Gigabit Equalized SQPSK Demodulator - Final Report", Harris Corporation Internal Report, June 1987.
6. Roberts, R.D., "Analysis of a Gradient Based IF Equalizer Control Algorithm with a Pseudoerror Monitor Cost Function", Ph.D. Dissertation, Florida Institute of Technology, 1992.



## RHODES FILTERS

### CATALOG OF RHODES FILTER TRANSFER FUNCTIONS AND ELEMENT VALUES

-----

William B. Lurie, Consultant  
Boynton Beach, Florida

Filters based on the Rhodes linear-phase polynomial have received some small attention in the literature. They are interesting because they offer as a compromise a somewhat flat amplitude and delay characteristic over 75-80% of the passband. This is an improvement over minimum-phase realizations of the Bessel filters, which can not avoid their typical rounded nose, or of the Butterworth types which can not avoid their typical steep rise in delay in the passband. Even the transitional types do not comprise a rigorous consistent set significantly improved in both respects.

A complete set of poles and zeros for the transfer functions of Rhodes-derived low-pass filters is presented, for all the even degrees through  $N=16$ , for the entire useful range of the parameter  $\epsilon$ . Comparisons are made among those of the same  $\epsilon$  in various degrees, and with some other filters such as the Bessel of the same degrees.

None of these filters can be realized without negative elements or mutual coupling. Methods are presented for accommodating these in practical filters.

A class of non-minimum-phase low-pass filters is described, presenting, in the passband, a compromise between constancy of amplitude (or minimization of amplitude variation and roll-off) and constancy of envelope delay (or minimization of delay variation). The fundamental description of the Rhodes polynomials has been published [1] and is quite clear, and the properties of filters based on these have also been presented [1,2], in a qualitative way.

What has not been presented heretofore is a corresponding catalog of transfer functions and low-pass prototype element values. The transfer functions can be generated with relative ease following the methods outlined in [1], and the techniques for the synthesis of filters from transfer functions are 'state of the art'. What remains, then, is to tabulate both the transfer functions and element values for a number of degrees and for a variety of values of the parameter  $\epsilon$  within each degree. In this paper, only even overall degrees are treated, although a complete set for odd degrees is also possible. Also presented are some response comparisons within the framework of the class of filters discussed.

The significant feature of all of these transfer functions is that they are non-minimum phase, that is, the zeros are not restricted to the left half-plane (which in this text will be defined as including the imaginary axis). In this regard, the current tabulation represents the first for a complete class of non-minimum phase networks.

Within those of any one degree, a parameter  $\epsilon$  is allowed to increase from zero, causing the arrays of poles and zeros to have somewhat varying aspects. The poles, of course, must all be in the left half-plane (but NOT on the imaginary axis), and are in complex conjugate pairs, since all the networks here described are even-degree low-pass filters. The zeros, however, need not be (and, for the non-minimum phase filters, MUST not be) all in the left half-plane. In some cases, there are pure imaginary pairs (conjugates, at  $+jb$  and  $-jb$ , for example), and in some cases complex conjugate QUADS (for example at  $+a+jb$ ,  $+a-jb$ ,  $-a+jb$ , and  $-a-jb$ ). In addition they may be on the positive and negative real axis, at  $-a+j0$  and  $+a+j0$ , or may be in various combinations of all of these, as will be seen. In all cases, for even degrees, the total number of zeros will be two less than the number of poles.

When there are only four poles, there are only two zeros, one each on the positive and negative real axis. When there are six poles, there are four zeros, with three possible arrangements:

- For small  $\epsilon$ , less than about .71, a quad of zeros.
- For  $\epsilon$  between about .71 and .8, two pairs on the real axis (two plus, two minus).
- For  $\epsilon$  greater than about .8, one pair on the real axis, one pair on the imaginary axis.

For higher degrees, more pure imaginary pairs and more quads can appear, but never more than two pairs on the real axis.

The amplitude and delay responses for the various transfer functions can be compared in many ways, with each other and with other classes of filters. Those within one degree have been plotted for a number of values of  $\epsilon$ , as seen in the attached figures. Note that, in most cases, a curve has been added which corresponds to the maximally-flat-delay (Bessel) response of the same degree, for comparison. It should be mentioned that the pole-zero arrays have all been normalized so that a frequency of 1 Hz. always corresponds to the -3 db point relative to the maximum amplitude in the passband.

Alternatively, to evaluate the effects of changing from one degree to another for a fixed value of  $\epsilon$ , another series of plots has been supplied.

These are all low-pass filters, of even degree. An entire family of filters with similar characteristics exists, for odd degrees, but no attempt has been made to expand the present set to include those, for a number of reasons. As with all low-pass filters, any of these can be converted to high-pass, but since the constancy of delay is completely destroyed in such a transformation, there would be no point in doing so. Low-pass to band-pass transformations, however, are possible and practical, and these low-pass transfer functions form a good starting point from which to develop band-pass filters with corresponding amplitude and delay characteristics.

It is not the purpose of this paper to make any judgements or comparisons between filters of these types and those of other classes (for example, Elliptic filters, or transitional filters between Bessel and Butterworth, with or without zeros at real frequencies, or conventional filters with delay equalizers added). For every systems application or intended use, all of the available filter design approaches should be considered, and the pros and cons of each evaluated. The actual implementation must also be taken into account. In many cases, the choice among the various possibilities is influenced by the number of coils, the need for realizing negative inductances (provided by tapping coils), the range of element values, and, often by

cost considerations, or, to a lesser degree these days, the availability of suitable means for synthesizing the filter.

Examination of the schematic diagrams and element values can lead to disturbing thoughts regarding physical realization. In each case, a negative element is seen to exist, or a bridged-T with a negative element, or both. For the bridged-T with negative inductor, conventional methods of creating this with tapped coils already exist (see Reference 3). In the case of a negative capacitor, as in Schematic 4, this can be handled in a number of ways, but perhaps the most straightforward is to synthesize the network in dual form. It is beyond the scope of this paper to present all the synthesis techniques, all the alternate pole-removal sequences, and possible network conversions. However, one dual pair of networks with identical transfer functions is shown (Figure 1) to illustrate a way to avoid the negative capacitor. It should be mentioned that, even with its negative capacitor, the network is quite adequate as the prototype for certain low-pass to band-pass transformations, yielding filters with all positive, realizable physical components.

Once the transfer functions are obtained, in order to prepare for their practical use, it is convenient to normalize to some cut-off frequency such as the -3db point. The next step is to synthesize using modern synthesis methods, and here some decisions have to be made with regard to the pole-removal sequences. The transfer function zeroes, as mentioned, can be on the positive and negative real axis, in pairs, on the positive and negative imaginary axis, in conjugate pairs, in quad in all four quadrants symmetrically about both axes, or in various combinations of all these, depending on the order of the transfer function and the parameter  $\epsilon$ .

In synthesizing, the various sections are created sequentially, the sequence being varied at the computer operator's option, yielding a number of entirely equivalent but different schematics for each transfer function. The numbers of elements are essentially the same, but their values may be more desirable in one configuration than in another. In this article, the convention has been adopted that all quads will be removed first, then all zeroes on the real axis (if any), then all zeroes on the imaginary axis (if any), and finally the two zeroes at infinity. The tabulations which follow adhere to this convention, and it is not represented here that these are the optimum configurations. Of course, from the transfer functions given, numerous other sequences can be used. What are given are tabulations for  $M=2$  through  $M=8$ , giving transfer functions of degree 4 through 16 (even values only), for a number of values of the parameter  $\epsilon$ . Curves of amplitude and delay are also given.

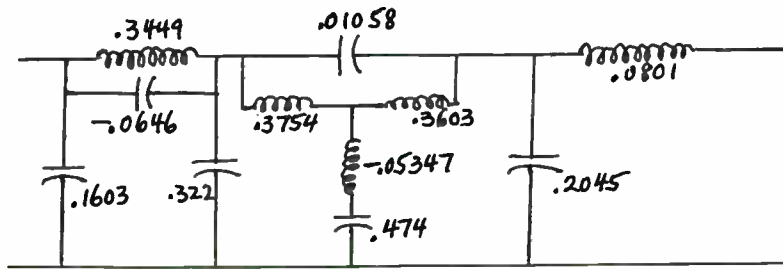
It is difficult to prepare tables of transfer function zeros and poles for so many different degrees of filter and possible values of the parameter  $\epsilon$  for each degree. Unfortunately, these table would be very large if organized to separate those on the real axis, the imaginary axis, and the four quadrants. What is being furnished instead is a tabulation for poles in which each entry lists the poles as real, imaginary real, imaginary, et cetera, and the complex conjugate of each is implied. For the zeros, a somewhat different approach is needed, and no all-encompassing rule is possible. In every zeros case, each pair of numbers must be examined, and then:

1. If both are non-zero, then they describe a quad of zeros ( $a+jb$ ,  $a-jb$ ,  $-a+jb$ ,  $-a-jb$ ), and one or more of these four pairs may be shown.
2. If the first of the pair is a zero, then the pair describes a pair of zeros on the imaginary axis ( $0+jb$ ,  $0-jb$ ), and again one or more of the two pairs may be shown.
3. If the second of the pair is a zero, then the pair describes a pair of zeros on the real axis ( $a+j0$ ,  $-a+j0$ ), and again one or more of the two pairs may be shown.

The assistance of several people in composing and assembling this paper should be acknowledged: Paul Herzig of ECI/E-Systems, Inc., for ground work and encouragement, Mark Mell of Microsonics, Inc., for some of the software used in synthesizing, and Dick Abrahams of Radio Systems, Inc., in preparing the illustrations.

#### REFERENCES

- 
- [1] Rhodes, J. D., IEEE Trans. Microwave Theory & Techniques, MTT-18 No. 6, June 1970, pp. 290-307
  - [2] Herzig, P. A., & Swanson, T. W., 'A Polyolithic Crystal Filter Employing a Rhodes Transfer Function', presented at 32nd Annual Frequency Control Symposium, 31 May -2 June 1978, Atlantic City, N. J.
  - [3] Lurie, W. B., 'Useful Network Transformations in Filter Design', presented at RF EXPO/East, November 13-15, 1990, pp. 477-489



DUAL

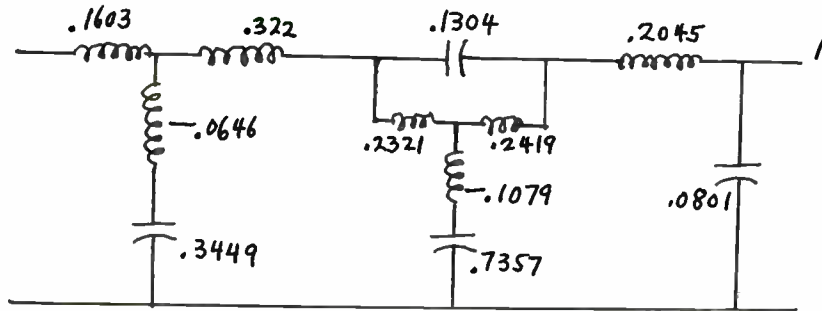
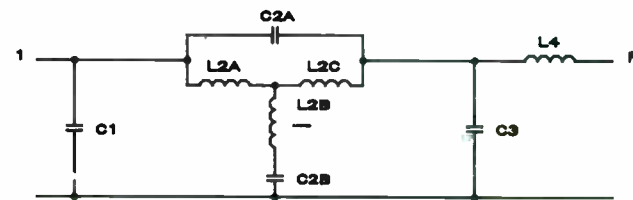


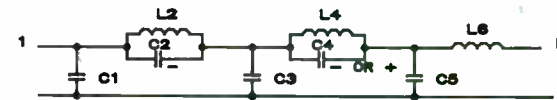
FIGURE 1



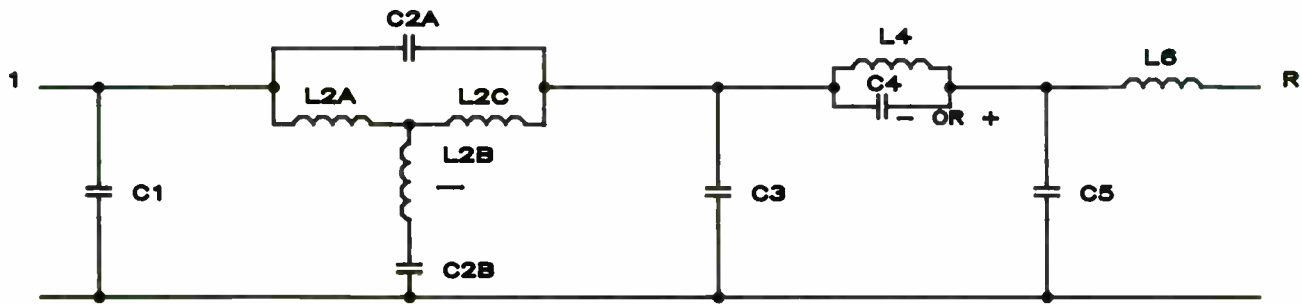
SCHEMATIC 2



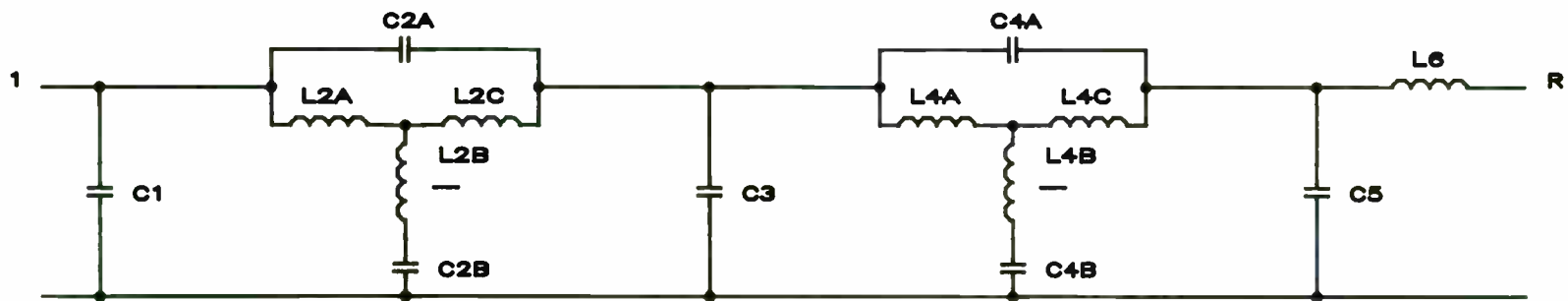
SCHEMATIC 3A



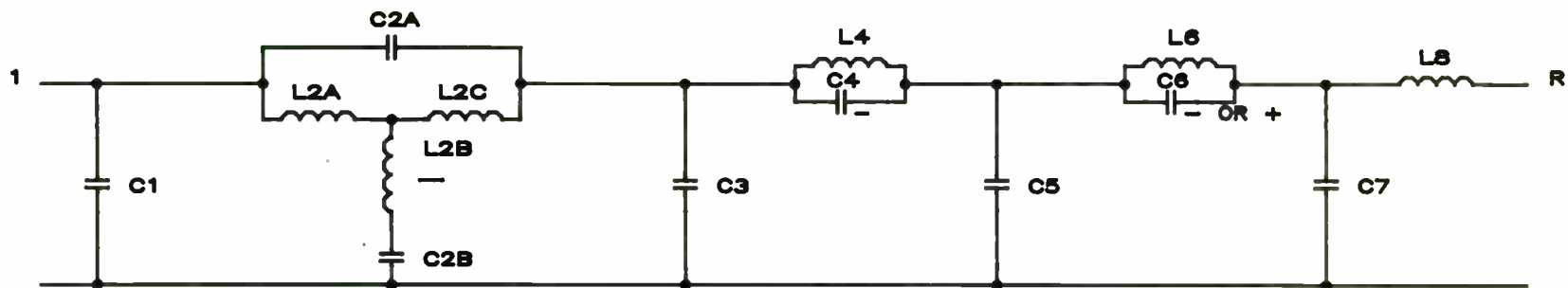
SCHEMATIC 3B



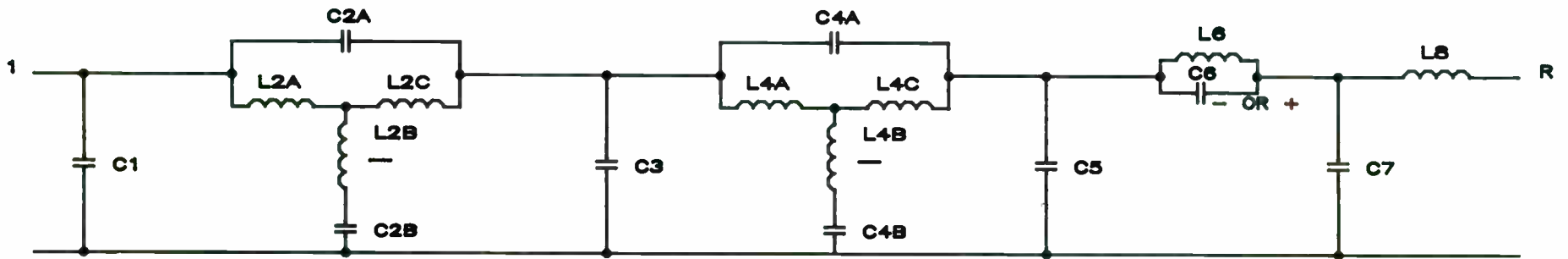
SCHEMATIC 4



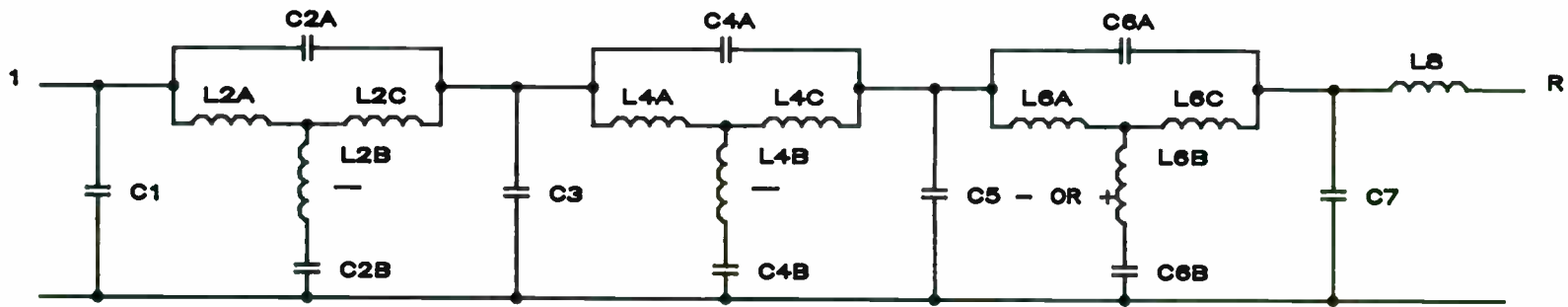
SCHEMATIC 5A



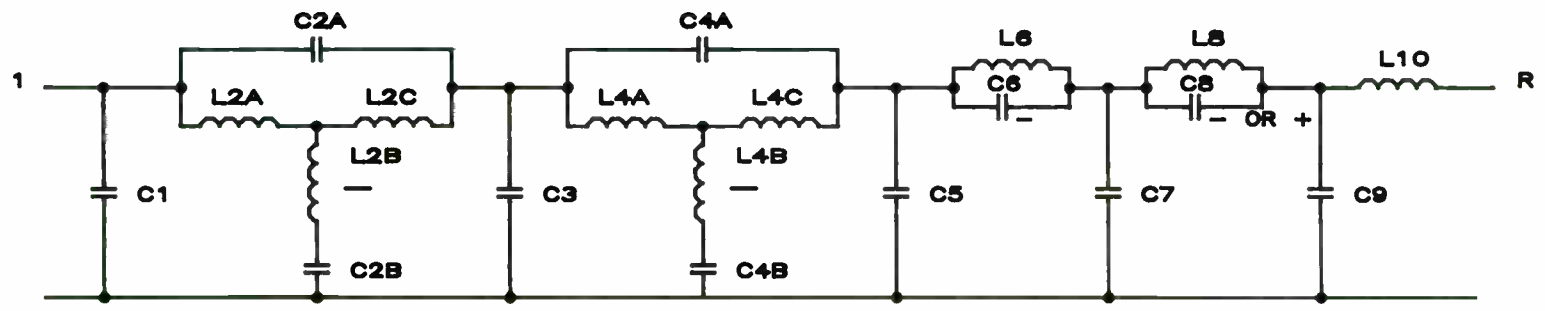
SCHEMATIC 5B



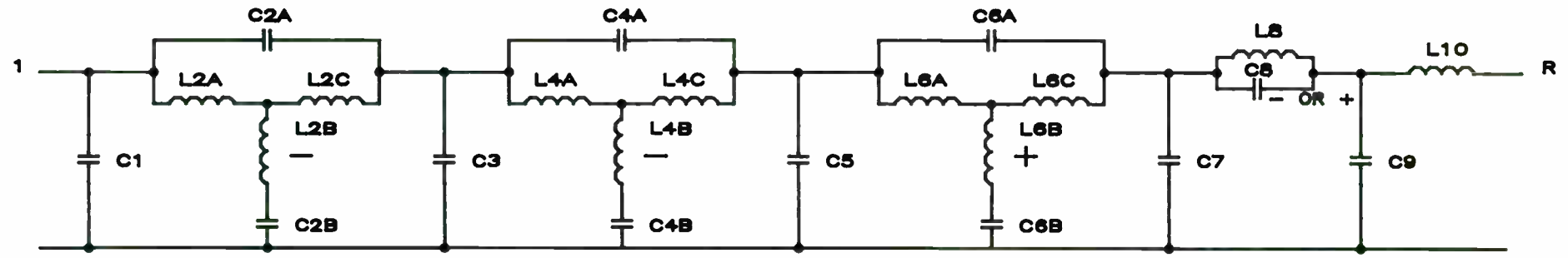
SCHEMATIC 6



SCHEMATIC 7A



SCHMATIC 7B



SCHMATIC 8

EPS	C1	L2	C2	C3	L4	R
.000	.1684	.3842	-.0393	.3526	.1370	1.000
.314	.1814	.3870	-.0382	.3454	.2379	.996
.393	.1861	.3838	-.0379	.3447	.1442	.991
.471	.1920	.3785	-.0374	.3441	.1472	.981
.550	.1995	.3705	-.0365	.3438	.1529	.963
.628	.2088	.3584	-.0350	.3442	.1589	.935
.707	.2203	.3409	-.0325	.3445	.1648	.894
.785	.2345	.3163	-.0283	.3481	.1701	.835
.864	.2522	.2829	-.0209	.3520	.1739	.756
.942	.2743	.2389	-.0069	.3569	.1754	.656

TABLE 2  
REFERENCE: SCHEMATIC 2

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3
.000	.0715	.0100	.3122	-.0693	.4956	.4067	.2565
.314	.0918	.0086	.3383	-.0673	.5018	.4001	.2453
.393	.0964	.0082	.3420	-.0674	.5004	.3989	.2493
.471	.1025	.0077	.3459	-.0677	.4988	.3973	.2541
.550	.1104	.0067	.3500	-.0681	.4971	.3952	.2600
.628	.1210	.0052	.3538	-.0686	.4958	.3924	.2673

EPS	L4	R
.000	.1049	1.000
.314	.1001	1.000
.393	.1043	1.000
.471	.1096	.997
.550	.1162	.993
.628	.1242	.984

TABLE 3A  
REFERENCE SCHEMATIC 3A

EPS	C1	L2	C2	C3	L4	C4	C5
.707	.1534	.2977	-.0146	.3997	.4474	-.0625	.354
.785	.2570	.4545	-.0821	.4001	.2857	.0047	.284
.864	.3049	.4908	-.0997	.4128	.2385	.0281	.276
.942	.3726	.5252	-.1236	.4378	.1831	.0658	.269

EPS	L6	R
.707	.1338	.9690
.785	.1448	.9400
.864	.1567	.8910
.942	.1682	.8120

TABLE 3B  
REFERENCE SCHEMATIC 3B



EPS	C1	C2A	L2A	L2B	C2B	L2C	C3
.000	.0696	.0043	.2292	-.0257	.3390	.3353	.3394
.314	.0439	.0202	.2608	-.0626	.4382	.3540	.3576
.393	.0621	.0156	.2831	-.0580	.4504	.3737	.3618
.471	.0657	.0151	.2912	-.0603	.4585	.3772	.3587
.550	.0701	.0147	.3050	-.0657	.4765	.3848	.3507
.628	.0740	.0160	.3300	-.0776	.5179	.4012	.3312
.707	.0771	.0200	.3610	-.0908	.5734	.4211	.3043
.785	.0814	.0250	.3890	-.1007	.6249	.4380	.2792
.864	.0878	.0310	.4149	-.1081	.6759	.4530	.2547
.942	.0972	.0385	.4383	-.1130	.7300	.4658	.2294
1.021	.1113	.0487	.4556	-.1140	.7907	.4735	.2026

EPS	L4	C4	C5	L6	R
.000	.3450	-.0313	.2354	.0738	1.0000
.314	.4368	-.0558	.3032	.0916	1.0000
.393	.4172	-.0569	.2766	.0769	1.0000
.471	.4122	-.0540	.2778	.0801	1.0000
.550	.3985	-.0476	.2753	.0841	1.0000
.628	.3654	-.0327	.2638	.0892	.9990
.707	.3236	-.0128	.2492	.0956	.9980
.785	.2868	.0065	.2889	.1038	.9940
.864	.2507	.0276	.2309	.1142	.9870
.942	.2116	.0543	.2238	.1271	.9710
1.021	.1668	.0944	.2169	.1430	.9380

TABLE 4  
REFERENCE SCHEMATIC 4

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3	C4A
.000	.0126	.0349	.3227	-.1170	.6142	.3531	.1694	.0137
.314	.0367	.0238	.3327	-.1111	.5896	.4013	.2518	.0151
.393	.0391	.0225	.3317	-.1095	.5834	.4014	.2546	.0149
.471	.0305	.0242	.3189	-.1099	.5731	.3778	.2312	.0138

EPS	L4A	L4B	C4B	L4C	C5	L6	R
.000	.3275	-.0285	.4172	.3425	.2238	.0918	1.0000
.314	.3542	-.0326	.4132	.3038	.1635	.0629	1.0000
.393	.3565	-.0344	.4190	.1310	.1690	.0654	1.0000
.471	.3564	-.0364	.4348	.3398	.2079	.0851	1.0000

TABLE 5A  
REFERENCE SCHEMATIC 5A

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3
.550	.0258	.0341	.2511	-.0634	.4446	.3134	.3210
.628	.0458	.0291	.2821	-.0628	.4713	.3516	.3650
.707	.0478	.0332	.2980	-.0659	.4982	.3595	.3747
.785	.0508	.0380	.3107	-.0653	.5177	.3640	.3866
.864	.0563	.0438	.3203	-.0612	.5303	.3651	.4028

EPS	L4	C4	C5	L6	C6	C7	L8	R
.550	.4584	-.0610	.4308	.3834	-.0417	.2706	.0874	.9996
.628	.4965	-.0735	.4144	.3094	-.0916	.2153	.0722	.9999
.707	.5265	-.0842	.4125	.2721	.0032	.2006	.0773	.9997
.785	.5593	-.0960	.4139	.2405	.0254	.1909	.0839	.9990
.864	.5989	-.1099	.4190	.2105	.0499	.1842	.0925	.9970

TABLE 5B  
REFERENCE SCHEMATIC 5B

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3	C4A
.000	.0030	.0441	.3108	-.1093	.6121	.3206	.1008	.0200
.314	.0068	.0413	.3123	-.1097	.6056	.3309	.1312	.0180
.393	.0095	.0392	.3148	-.1109	.6047	.3386	.1458	.0176
.471	.0244	.0317	.3326	-.1144	.6110	.3802	.2074	.0167
.550	.0234	.0360	.3660	-.1281	.6795	.4081	.1970	.0182
.628	.0111	.0464	.3921	-.1412	.7571	.4145	.1397	.0235
.707	.0229	.0455	.4269	-.1496	.8025	.4616	.1795	.0252
.785	.0207	.0520	.4586	-.1608	.8704	.4882	.1635	.0311

EPS	L4A	L4B	C4B	L4C	C5
.000	.2639	-.0147	.3661	.3275	.3710
.314	.2852	-.0143	.3750	.3316	.3660
.393	.2951	-.0158	.3816	.3341	.3609
.471	.3275	-.0175	.3951	.3298	.3248
.550	.3293	-.0216	.4075	.3403	.3029
.628	.3075	-.0263	.4145	.3450	.2946
.707	.3268	-.0222	.4215	.3460	.2677
.785	.3175	-.0179	.4192	.3397	.2511

EPS	L6	C6	C7	L8	R
.000	.4452	-.0682	.2934	.0783	1.0000
.314	.4344	-.0683	.2833	.0737	1.0000
.393	.4247	-.0656	.2779	.0732	1.0000
.471	.3686	-.0585	.2335	.0575	1.0000
.550	.3109	-.0315	.2049	.0607	1.0000
.628	.2948	-.0049	.2206	.0847	1.0000
.707	.2327	.0193	.1660	.0729	1.0000
.785	.2111	.0431	.1708	.0817	1.0000

TABLE 6  
REFERENCE SCHEMATIC 6

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3	C4A
.000	.0018	.0443	.3530	-.1310	.7009	.3578	.0686	.0247
.236	.0013	.0445	.3522	-.1308	.7007	.3558	.0575	.0273
.314	.0056	.0414	.3498	-.1296	.6849	.3631	.1086	.0196
.393	.0113	.0366	.3415	-.1259	.6556	.3658	.1444	.0174

EPS	L4A	L4B	C4B	L4C	C5	C6A	L6A	L6B
.000	.3292	-.0856	.5366	.3863	.2570	.0216	.3098	.0038
.236	.3229	-.0886	.5404	.3755	.2431	.0196	.3096	.0051
.314	.3572	-.0825	.5436	.4079	.2704	.0248	.3051	-.0018
.393	.3821	-.0835	.5584	.4215	.2698	.0279	.2985	-.0069

EPS	C6B	L6C	C7	L8	R
.000	.3357	.2577	.1557	.0624	1.0000
.236	.3423	.2788	.1853	.0759	1.0000
.314	.3341	.2436	.1348	.0552	1.0000
.393	.3363	.2365	.1213	.0510	1.0000

TABLE 7A  
REFERENCE SCHEMATIC 7A

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3	C4A
.471	.0018	.0493	.3226	-.1107	.6389	.3289	.0768	.0293
.550	.0025	.0521	.3448	-.1185	.6817	.3524	.0823	.0306
.628	.0104	.0512	.3661	-.1233	.7058	.3884	.1386	.0271

EPS	L4A	L4B	C4B	L4C	C5	L6	C6	C7
.471	.2385	-.0098	.3467	.3005	.3710	.4871	-.0708	.4343
.550	.2472	-.0105	.3568	.3043	.3842	.5162	-.0805	.4199
.628	.2819	-.0021	.3650	.3171	.4040	.5370	-.0904	.3938

EPS	L8	C8	C9	L10	R
.471	.3663	-.0455	.2460	.0719	1.0000
.550	.3021	-.0179	.2108	.0715	1.0000
.628	.2195	.0088	.1479	.0555	1.0000

TABLE 7B  
REFERENCE SCHEMATIC 7B

N = 3 POLES, THEN ZEROS

EPS	C1	C2A	L2A	L2B	C2B	L2C	C3	C4A
.000	.00014	.0477	.3465	-.1255	.6925	.3471	.0241	.0402
.079	.00010	.0477	.3466	-.1255	.6928	.3470	.0204	.0418
.157	.00013	.0477	.3467	-.1256	.6929	.3472	.0222	.0411
.236	.00033	.0474	.3465	-.1257	.6919	.3477	.0336	.0367
.314	.00055	.0468	.3461	-.1259	.6902	.3478	.0396	.0349
.393	.00048	.0463	.3519	-.1294	.7023	.3534	.0331	.0380
.471	.00360	.0478	.3852	-.1423	.7607	.3937	.0839	.0279
.550	.00360	.0516	.4138	-.1527	.8181	.4219	.0807	.0317
.628	.00082	.0573	.4417	-.1630	.8811	.4438	.0370	.0461
EPS	L4A	L4B	C4B	L4C	C5	C6A	L6A	L6B
.000	.2871	-.0820	.5211	.3240	.1863	.0212	.2750	.0226
.079	.2856	-.0829	.5248	.3187	.1746	.0216	.2694	.0222
.157	.2868	-.0827	.5240	.3210	.1785	.0215	.2710	.0220
.236	.2934	-.0808	.5177	.3360	.2051	.0210	.2825	.0219
.314	.2997	-.0813	.5209	.3438	.2127	.0211	.2854	.0209
.393	.3042	-.0869	.5416	.3414	.1923	.0219	.2774	.0189
.471	.3473	-.0841	.5546	.3940	.2532	.0238	.2992	.0171
.550	.3587	-.0891	.5819	.4015	.2453	.0245	.2983	.0185
.628	.3466	-.0979	.6207	.3788	.1850	.0283	.2755	.0215
EPS	C6B	L6C	C7	L8	C8	C9	L10	R
.000	.3133	.2881	.3553	.4366	-.0783	.2742	.0626	1.0000
.079	.3125	.2911	.3615	.4433	-.0771	.2821	.0672	1.0000
.157	.3131	.2912	.3614	.4432	-.0771	.2822	.0674	1.0000
.236	.3155	.2859	.3499	.4314	-.0786	.2696	.0604	1.0000
.314	.3177	.2868	.3475	.4266	-.0767	.2676	.0608	1.0000
.393	.3200	.2959	.3500	.4168	-.0647	.2744	.0732	1.0000
.471	.3238	.2756	.2710	.2931	-.0409	.1850	.0747	1.0000
.550	.3278	.2855	.2534	.2326	-.0087	.1515	.0498	1.0000
.628	.3251	.2961	.2663	.2369	.0192	.1773	.0742	1.0000

TABLE 8  
REFERENCE SCHEMATIC 8

0.752505 0.174836 0.649067 0.500752 0.337307 0.905542  
0.75028 0.175056 0.647401 0.501963 0.336108 0.906427  
0.742079 0.175343 0.641073 0.504549 0.331817 0.907211  
0.72907 0.175924 0.631055 0.509212 0.324928 0.909236  
0.710967 0.176658 0.617041 0.51563 0.315265 0.912034  
0.687876 0.177471 0.599048 0.5237 0.302811 0.915569  
0.65993 0.178276 0.577105 0.533292 0.287546 0.919795  
0.627291 0.178977 0.55126 0.544261 0.269454 0.924655  
0.590149 0.179481 0.521582 0.55645 0.24852 0.930091  
0.548723 0.179699 0.488173 0.569706 0.224742 0.936052  
0.503264 0.17956 0.451167 0.583894 0.198128 0.942513  
0.454055 0.179019 0.410745 0.59893 0.168706 0.949509  
0.401414 0.178071 0.36713 0.614823 0.136517 0.9572  
0.345668 0.176763 0.320579 0.631742 0.101603 0.965952  
0.287132 0.175219 0.271366 0.650146 0.0639696 0.976509  
0.226032 0.173681 0.219725 0.670994 0.0235091 0.990274  
0.211573 0.8887 0.20726 0.220304 0.0257502 1.28916  
0.338286 2.26218 0.296302 0.535578 0.211302 3.2266  
0.148299 2.38693 0.0724037 0.538233 0.377238 3.3487

1.04619 0.44566 -1.04619 0.44566  
1.04697 0.445729 -1.04697 0.445729  
1.0473 0.444917 -1.0473 0.444917  
1.04933 0.443549 -1.04933 0.443549  
1.05364 0.440707 -1.05364 0.440707  
1.06215 0.435026 -1.06215 0.435026  
1.0783 0.423676 -1.0783 0.423676  
1.10863 0.399851 -1.10863 0.399851  
1.16747 0.341892 -1.16747 0.341892  
1.29483 0.0441978 -1.29483 0.0441978  
2.41211 0 0.95205 0 -2.41211 0 -0.95205 0  
0.823684 0 -0.823684 0 0 4.3565  
0.719386 0 -0.719386 0 0 1.94524  
0.624562 0 -0.624562 0 0 1.44988  
0.53415 0 -0.53415 0 0 1.20644  
0.445002 0 -0.445002 0 0 1.05839  
0.450118 0 -0.450118 0 0 1.22633  
0.753895 0 -0.753895 0 0 2.77557  
0 2.61379 0 0.268128 0 -2.61379 0 -0.268128

N = 4 POLES, THEN ZEROS

2.62451 0.420994 2.41105 1.3262 1.97716 2.19575 0.987284 3.29055  
2.61757 0.422495 2.40484 1.32882 1.97296 2.20066 0.98406 3.29433  
2.59677 0.426984 2.38618 1.33663 1.96031 2.21536 0.974353 3.30571  
2.56216 0.434426 2.35497 1.34954 1.93907 2.23987 0.958063 3.32476  
2.51378 0.444766 2.31105 1.36735 1.90898 2.27418 0.935021 3.3516  
2.4517 0.457948 2.2542 1.38982 1.86971 2.3183 0.904984 3.38643  
2.37594 0.47392 2.18416 1.41669 1.82085 2.37222 0.867632 3.42945  
2.28651 0.492655 2.10059 1.44762 1.76191 2.43594 0.822559 3.48098  
2.18332 0.514153 2.00314 1.48229 1.69234 2.50948 0.76927 3.54133  
2.06618 0.538455 1.89136 1.52038 1.61153 2.59285 0.707168 3.61092  
1.93481 0.565648 1.7648 1.56158 1.51882 2.68612 0.635548 3.69019  
1.78875 0.59586 1.62287 1.60564 1.41351 2.78943 0.553587 3.77965  
1.62738 0.629268 1.46494 1.65237 1.29484 2.90297 0.460341 3.87987  
1.44988 0.666096 1.29023 1.70168 1.16196 3.02706 0.354736 3.99148  
1.2552 0.706616 1.0978 1.7536 1.01395 3.16217 0.235567 4.11519  
1.04205 0.751153 0.886506 1.80826 0.84979 3.30893 0.10149 4.25182  
0.808831 0.800099 0.668288 3.46819 0.654925 1.86596 0.0489826 4.40233  
0.553601 0.853919 0.468087 3.64108 0.401291 1.92713 0.217504 4.56787  
0.274019 0.913178 0.247592 3.82904 0.123419 1.99236 0.405923 4.74985

3.59168 0 ^-3.59168 0 3.11585 ^-2.21554 3.11585 2.21554 ^-3.11585 ^-2.21554  
^-3.11585 2.21554  
3.59444 0 ^-3.59444 0 3.11869 ^-2.21637 3.11869 2.21637 ^-3.11869 ^-2.21637  
^-3.11869 2.21637  
3.60382 0 ^-3.60382 0 3.12772 ^-2.21831 3.12772 2.21831 ^-3.12772 ^-2.21831  
^-3.12772 2.21831  
3.62356 0 ^-3.62356 0 3.14459 ^-2.21942 3.14459 2.21942 ^-3.14459 ^-2.21942  
^-3.14459 2.21942  
3.66172 0 ^-3.66172 0 3.17232 ^-2.2155 3.17232 2.2155 ^-3.17232 ^-2.2155 ^-3.17232  
2.2155  
3.73566 0 ^-3.73566 0 3.21555 ^-2.19767 3.21555 2.19767 ^-3.21555 ^-2.19767  
^-3.21555 2.19767  
3.88814 0 ^-3.88814 0 3.27922 ^-2.14583 3.27922 2.14583 ^-3.27922 ^-2.14583  
^-3.27922 2.14583  
4.25806 0 ^-4.25806 0 3.35293 ^-2.01137 3.35293 2.01137 ^-3.35293 ^-2.01137  
^-3.35293 2.01137  
5.42609 0 ^-5.42609 0 3.31622 ^-1.74314 3.31622 1.74314 ^-3.31622 ^-1.74314  
^-3.31622 1.74314  
9.33368 0 ^-9.33368 0 3.07756 ^-1.54694 3.07756 1.54694 ^-3.07756 ^-1.54694  
^-3.07756 1.54694  
0 14.1974 0 ^-14.1974 2.81931 ^-1.47266 2.81931 1.47266 ^-2.81931 ^-1.47266  
^-2.81931 1.47266  
0 7.48158 0 ^-7.48158 2.57377 ^-1.44727 2.57377 1.44727 ^-2.57377 ^-1.44727  
^-2.57377 1.44727  
0 5.92046 0 ^-5.92046 2.33243 ^-1.44499 2.33243 1.44499 ^-2.33243 ^-1.44499  
^-2.33243 1.44499  
0 5.17253 0 ^-5.17253 2.08799 ^-1.45646 2.08799 1.45646 ^-2.08799 ^-1.45646  
^-2.08799 1.45646  
0 4.73993 0 ^-4.73993 1.83523 ^-1.47772 1.83523 1.47772 ^-1.83523 ^-1.47772  
^-1.83523 1.47772  
0 4.47416 0 ^-4.47416 1.56935 ^-1.50715 1.56935 1.50715 ^-1.56935 ^-1.50715  
^-1.56935 1.50715  
0 4.31104 0 ^-4.31104 1.28232 ^-1.54464 1.28232 1.54464 ^-1.28232 ^-1.54464  
^-1.28232 1.54464  
0 4.2119 0 ^-4.2119 0.950173 ^-1.59102 0.950173 1.59102 ^-0.950173 ^-1.59102

^-0.950173 1.59102  
0 4.14075 0 ^-4.14075 0.439523 ^-1.64698 0.439523 1.64698 ^-0.439523 ^-1.64698  
^-0.439523 1.64698

N = 5 POLES, THEN ZEROS

3.29013 0.444783 3.13297 1.29564 2.77671 2.21194 2.21804 3.10638 1.08214  
4.2453  
3.28153 0.445455 3.12521 1.29916 2.77003 2.21625 2.21372 3.11265 1.07865  
4.25011  
3.25573 0.447445 3.10192 1.30967 2.74995 2.22916 2.20067 3.13149 1.06816  
4.26456  
3.2128 0.450671 3.06305 1.32707 2.71626 2.25058 2.17871 3.16294 1.05056  
4.28876  
3.1528 0.455004 3.00857 1.35116 2.66867 2.28036 2.14752 3.20709 1.02568  
4.32288  
3.07581 0.460285 2.93839 1.38174 2.60676 2.31829 2.10667 3.26402 0.993265  
4.36719  
2.98187 0.466333 2.85242 1.41858 2.53002 2.36412 2.05559 3.33386 0.952979  
4.42199  
2.87096 0.472958 2.75052 1.46146 2.43786 2.41752 1.99362 3.41674 0.904403  
4.48771  
2.74298 0.479974 2.63249 1.51022 2.32961 2.47818 1.92 3.51278 0.847012 4.56485  
2.59768 0.487206 2.49805 1.56479 2.20457 2.54572 1.8339 3.62213 0.780167  
4.65399  
2.4347 0.494491 2.34681 1.62516 2.06196 2.61981 1.73441 3.74496 0.703101  
4.75582  
2.25346 0.501684 2.17823 1.69147 1.90095 2.70014 1.62056 3.88149 0.614898  
4.87114  
2.05319 0.50866 1.99158 1.76394 1.72066 2.78647 1.49134 4.03201 0.514482  
5.00083  
1.83287 0.515312 1.78592 1.84295 1.52011 2.87869 1.34571 4.19693 0.400595  
5.14588  
1.59122 0.521552 1.56005 1.92896 1.29819 2.97683 1.18253 4.37684 0.271792  
5.30742  
1.32663 0.527313 1.31245 2.02261 1.05365 3.08117 1.0006 4.57254 0.126437  
5.48666  
1.04128 2.12465 1.03711 0.532545 0.798581 4.78511 0.784924 3.19221 0.0372936  
5.685  
0.744278 2.23601 0.720242 0.537218 0.574977 5.016 0.490129 3.31077 0.221412  
5.90403  
0.418731 2.35784 0.373099 0.541314 0.328069 5.26711 0.166884 3.43799 0.428123  
6.14565

4.21769 -1.04981 4.21769 1.04981 3.44126 -3.24485 3.44126 3.24485 -4.21769  
-1.04981 -4.21769 1.04981 -3.44126 -3.24485 -3.44126 3.24485  
4.22078 -1.05018 4.22078 1.05018 3.44445 -3.24616 3.44445 3.24616 -4.22078  
-1.05018 -4.22078 1.05018 -3.44445 -3.24616 -3.44445 3.24616  
4.23157 -1.0507 4.23157 1.0507 3.4545 -3.24922 3.4545 3.24922 -4.23157 -1.0507  
-4.23157 1.0507 -3.4545 -3.24922 -3.4545 3.24922  
4.25529 -1.04906 4.25529 1.04906 3.47279 -3.25116 3.47279 3.25116 -4.25529  
-1.04906 -4.25529 1.04906 -3.47279 -3.25116 -3.47279 3.25116  
4.30373 -1.03906 4.30373 1.03906 3.50143 -3.24565 3.50143 3.24565 -4.30373  
-1.03906 -4.30373 1.03906 -3.50143 -3.24565 -3.50143 3.24565  
4.40373 -1.00156 4.40373 1.00156 3.54154 -3.21966 3.54154 3.21966 -4.40373  
-1.00156 -4.40373 1.00156 -3.54154 -3.21966 -3.54154 3.21966  
4.62532 -0.852146 4.62532 0.852146 3.58466 -3.14781 3.58466 3.14781 -4.62532  
-0.852146 -4.62532 0.852146 -3.58466 -3.14781 -3.58466 3.14781  
5.90186 0 4.46638 0 -3.90186 0 -4.46638 0 3.58238 -3.00227 3.58238 3.00227  
-3.58238 -3.00227 -3.58238 3.00227  
9.81825 0 3.95699 0 -9.81825 0 -3.95699 0 3.45866 -2.83754 3.45866 2.83754

-3.45866 -2.83754 -3.45866 2.83754  
3.64074 0 -3.64074 0 0 26.0608 0 -26.0608 3.25356 -2.74193 3.25356 2.74193  
-3.25356 -2.74193 -3.25356 2.74193  
3.36356 0 -3.36356 0 0 9.97658 0 -9.97658 3.0264 -2.70696 3.0264 2.70696  
-3.0264 -2.70696 -3.0264 2.70696  
3.09562 0 -3.09562 0 0 7.7502 0 -7.7502 2.79041 -2.70877 2.79041 2.70877  
-2.79041 -2.70877 -2.79041 2.70877  
2.82487 0 -2.82487 0 0 6.76815 0 -6.76815 2.54497 -2.73449 2.54497 2.73449  
-2.54497 -2.73449 -2.54497 2.73449  
2.54472 0 -2.54472 0 0 6.22297 0 -6.22297 2.28725 -2.77761 2.28725 2.77761  
-2.28725 -2.77761 -2.28725 2.77761  
2.25065 0 -2.25065 0 0 5.90003 0 -5.90003 2.01435 -2.8347 2.01435 2.8347  
-2.01435 -2.8347 -2.01435 2.8347  
1.9389 0 -1.9389 0 0 5.71558 0 -5.71558 1.72325 -2.90426 1.72325 2.90426  
-1.72325 -2.90426 -1.72325 2.90426  
1.60491 0 -1.60491 0 0 5.62772 0 -5.62772 1.40885 2.98649 1.40885 -2.98649  
-1.40885 2.98649 -1.40885 -2.98649  
1.23767 0 -1.23767 0 0 5.60727 0 -5.60727 1.05468 -3.08341 1.05468 3.08341  
-1.05468 -3.08341 -1.05468 3.08341  
0.783399 0 -0.783399 0 0 5.62148 0 -5.62148 0.569309 -3.19778 0.569309 3.19778  
-0.569309 -3.19778 -0.569309 3.19778

N = 6 POLES, THEN ZEROS

3.9585 0.425215 3.82139 1.31495 3.54979 2.1771 3.08326 3.10747 2.42303 4.02547  
1.16403 5.20537  
3.94836 0.426682 3.81176 1.31767 3.54152 2.18244 3.07626 3.11331 2.4186  
4.03303 1.1603 5.21117  
3.918 0.431066 3.78283 1.32577 3.51668 2.1984 3.05519 3.1308 2.40523 4.05574  
1.14907 5.22863  
3.86748 0.438326 3.73461 1.3391 3.47513 2.22485 3.01978 3.1599 2.3827 4.0937  
1.13025 5.25788  
3.79693 0.448403 3.66703 1.35738 3.41667 2.26158 2.96962 3.20055 2.35066  
4.14708 1.10364 5.29913  
3.70647 0.461229 3.58003 1.3803 3.34104 2.30831 2.90412 3.25261 2.30859  
4.21607 1.06899 5.3527  
3.59621 0.47675 3.47351 1.4075 3.24793 2.36474 2.82259 3.31592 2.25582 4.30094  
1.02597 5.41901  
3.46615 0.49493 3.34729 1.43861 3.13699 2.43056 2.72416 3.39028 2.19157  
4.40197 0.974148 5.49857  
3.31623 0.51577 3.20113 1.47328 3.00781 2.50551 2.60789 3.47542 2.11489  
4.51946 0.912987 5.59203  
3.1462 0.539307 3.03466 1.51123 2.85996 2.58941 2.47275 3.57104 2.02473  
4.65372 0.841836 5.70017  
2.95565 0.56563 2.84736 1.55221 2.69289 2.68221 2.31763 3.67682 1.91993  
4.80508 0.759901 5.82389  
2.74394 0.594873 2.63854 1.59606 2.50595 2.784 2.14139 3.79244 1.79928 4.9739  
0.666226 5.96427  
2.51019 0.627223 2.40728 1.64271 2.29831 2.89506 1.94284 3.91764 1.66148  
5.1606 0.559658 6.12253  
2.2532 0.662921 2.15239 1.69216 2.06892 3.01588 1.72072 4.05227 1.50524 5.3657  
0.438823 6.30007  
1.97147 0.702263 1.87238 1.74449 1.81645 3.14713 1.47373 4.19636 1.32918  
5.58991 0.302088 6.49847  
1.6631 0.745606 1.5654 1.7999 1.53924 3.28972 1.20046 4.35018 1.13191 5.8342  
0.147556 6.71944  
1.32573 0.793374 1.23524 3.44476 1.22917 1.85867 0.911927 6.09985 0.899294  
4.51441 0.0269428 6.9649  
0.956532 0.846071 0.901962 3.61366 0.860929 1.92122 0.667616 6.3886 0.568282  
4.69013 0.223821 7.23698  
0.552024 0.904299 0.536373 3.79812 0.457284 1.9881 0.397155 6.70278 0.204998  
4.879 0.445709 7.5381

5.00922 0 -5.00922 0 4.71933 -2.06088 4.71933 2.06088 3.71652 -4.25834 3.71652  
4.25834 -4.71933 -2.06088 -4.71933 2.06088 -3.71652 -4.25834 -3.71652  
4.25834  
5.01267 0 -5.01267 0 4.72276 -2.06163 4.72276 2.06163 3.72006 -4.26014 3.72006  
4.26014 -4.72276 -2.06163 -4.72276 2.06163 -3.72006 -4.26014 -3.72006  
4.26014  
5.02553 0 -5.02553 0 4.7347 -2.06255 4.7347 2.06255 3.73106 -4.26444 3.73106  
4.26444 -4.7347 -2.06255 -4.7347 2.06255 -3.73106 -4.26444 -3.73106  
4.26444  
5.0572 0 -5.0572 0 4.76074 -2.05855 4.76074 2.05855 3.75052 -4.26752 3.75052  
4.26752 -4.76074 -2.05855 -4.76074 2.05855 -3.75052 -4.26752 -3.75052  
4.26752  
5.13258 0 -5.13258 0 4.8123 -2.03516 4.8123 2.03516 3.77909 -4.26145 3.77909  
4.26145 -4.8123 -2.03516 -4.8123 2.03516 -3.77909 -4.26145 -3.77909  
4.26145  
5.33268 0 -5.33268 0 4.90539 -1.9463 4.90539 1.9463 3.81333 -4.2311 3.81333

4.2311 -4.90539 -1.9463 -4.90539 1.9463 -3.81333 -4.2311 -3.81333 4.2311  
6.11425 0 -6.11425 0 4.96424 -1.64468 4.96424 1.64468 3.83278 -4.15466 3.83278  
4.15466 -4.96424 -1.64468 -4.96424 1.64468 -3.83278 -4.15466 -3.83278  
4.15466  
9.16749 0 -9.16749 0 4.65669 -1.35207 4.65669 1.35207 3.78524 -4.03339 3.78524  
4.03339 -4.65669 -1.35207 -4.65669 1.35207 -3.78524 -4.03339 -3.78524  
4.03339  
24.1811 0 -24.1811 0 4.33433 -1.27986 4.33433 1.27986 3.64336 -3.92828 3.64336  
3.92828 -4.33433 -1.27986 -4.33433 1.27986 -3.64336 -3.92828 -3.64336  
3.92828  
0 13.9272 0 -13.9272 4.04615 -1.26245 4.04615 1.26245 3.44614 -3.87867 3.44614  
3.87867 -4.04615 -1.26245 -4.04615 1.26245 -3.44614 -3.87867 -3.44614  
3.87867  
0 9.94197 0 -9.94197 3.76685 -1.26572 3.76685 1.26572 3.22421 -3.87669 3.22421  
3.87669 -3.76685 -1.26572 -3.76685 1.26572 -3.22421 -3.87669 -3.22421  
3.87669  
0 8.49834 0 -8.49834 3.4839 -1.28044 3.4839 1.28044 2.98598 -3.90903 2.98598  
3.90903 -3.4839 -1.28044 -3.4839 1.28044 -2.98598 -3.90903 -2.98598  
3.90903  
0 7.74964 0 -7.74964 3.19025 -1.30307 3.19025 1.30307 2.73173 -3.96716 2.73173  
3.96716 -3.19025 -1.30307 -3.19025 1.30307 -2.73173 -3.96716 -2.73173  
3.96716  
0 7.31702 0 -7.31702 2.88116 -1.33196 2.88116 1.33196 2.45956 -4.04608 2.45956  
4.04608 -2.88116 -1.33196 -2.88116 1.33196 -2.45956 -4.04608 -2.45956  
4.04608  
0 7.06978 0 -7.06978 2.55294 -1.36629 2.55294 1.36629 2.16714 -4.14282 2.16714  
4.14282 -2.55294 -1.36629 -2.55294 1.36629 -2.16714 -4.14282 -2.16714  
4.14282  
0 6.9507 0 -6.9507 2.20222 -1.40572 2.20222 1.40572 1.85203 -4.25588 1.85203  
4.25588 -2.20222 -1.40572 -2.20222 1.40572 -1.85203 -4.25588 -1.85203  
4.25588  
0 6.92938 0 -6.92938 1.82529 -1.45031 1.82529 1.45031 1.51053 -4.38555 1.51053  
4.38555 -1.82529 -1.45031 -1.82529 1.45031 -1.51053 -4.38555 -1.51053  
4.38555  
0 6.98251 0 -6.98251 1.41492 -1.50064 1.41492 1.50064 1.13044 -4.53462 1.13044  
4.53462 -1.41492 -1.50064 -1.41492 1.50064 -1.13044 -4.53462 -1.13044  
4.53462  
0 7.07921 0 -7.07921 0.936793 -1.55817 0.936793 1.55817 0.639015 -4.70799  
0.639015 4.70799 -0.936793 -1.55817 -0.936793 1.55817 -0.639015 -4.70799  
-0.639015 4.70799

N = 7 POLES, THEN ZEROS

4.62262 0.441602 4.51072 1.29651 4.27034 2.19145 3.90725 3.06563 3.34961  
4.01102 2.60287 4.95131 1.2366 6.16938  
4.61086 0.442275 4.49956 1.30003 4.26 2.19605 3.89864 3.07263 3.34236 4.01826  
2.59834 4.9601 1.23264 6.17618  
4.57564 0.444267 4.46611 1.31056 4.22895 2.20979 3.87273 3.09358 3.32051 4.04  
2.58469 4.98653 1.22071 6.19662  
4.51704 0.447495 4.41039 1.32794 4.17705 2.23246 3.8293 3.12837 3.28375  
4.07623 2.56167 5.03073 1.20071 6.23085  
4.4352 0.451833 4.33243 1.35196 4.10411 2.26376 3.76802 3.1768 3.23159 4.12698  
2.5289 5.09295 1.17246 6.27914  
4.33026 0.457122 4.23228 1.38237 4.00986 2.30329 3.68843 3.23864 3.16333  
4.19224 2.4858 5.17351 1.1357 6.34187  
4.20232 0.463187 4.10991 1.4189 3.89398 2.35059 3.59001 3.31356 3.07807  
4.27198 2.43165 5.2728 1.09009 6.41951  
4.0514 0.469847 3.96528 1.46133 3.75608 2.40517 3.47215 3.40125 2.97475  
4.36613 2.36554 5.39128 1.03521 6.51271  
3.87739 0.476925 3.79822 1.50948 3.5957 2.46659 3.33422 3.50141 2.85217  
4.47457 2.28638 5.52944 0.970503 6.62224  
3.68 0.484256 3.60841 1.56327 3.41229 2.53446 3.17555 3.61381 2.70897 4.59712  
2.19294 5.68778 0.895332 6.74903  
3.45874 0.491691 3.39537 1.62272 3.20516 2.60846 2.99542 3.73834 2.54373  
4.73355 2.08384 5.86685 0.8089 6.89421  
3.21286 0.499093 3.15837 1.68796 2.97345 2.68842 2.79301 3.87508 2.35492  
4.8836 1.95758 6.06719 0.710241 7.05911  
2.94131 0.506347 2.89643 1.75926 2.7161 2.77428 2.56739 4.02433 2.141 5.04702  
1.8126 6.28937 0.598181 7.24531  
2.6427 0.51335 2.60823 1.837 2.43175 2.86614 2.31738 4.18666 1.90039 5.22364  
1.64731 6.53408 0.471284 7.4546  
2.31528 0.520014 2.29208 1.92169 2.11875 2.96425 2.04155 4.36293 1.63152  
5.41344 1.46009 6.80216 0.327804 7.68907  
1.95683 0.526269 1.94585 2.01399 1.77503 3.06903 1.73812 4.55429 1.33277  
5.61667 1.24931 7.0947 0.165629 7.95102  
1.56694 2.1147 1.56461 0.532055 1.40489 4.76224 1.39809 3.18109 1.0133 7.41321  
1.00247 5.83403 0.017735 8.24298  
1.15213 2.22482 1.13531 0.537329 1.03916 4.98861 0.984876 3.30127 0.750319  
7.75971 0.638673 6.06683 0.22514 8.56771  
0.697545 2.34553 0.664886 0.542057 0.637658 5.23567 0.531644 3.43067 0.458454  
8.13687 0.239034 6.31712 0.459721 8.92822

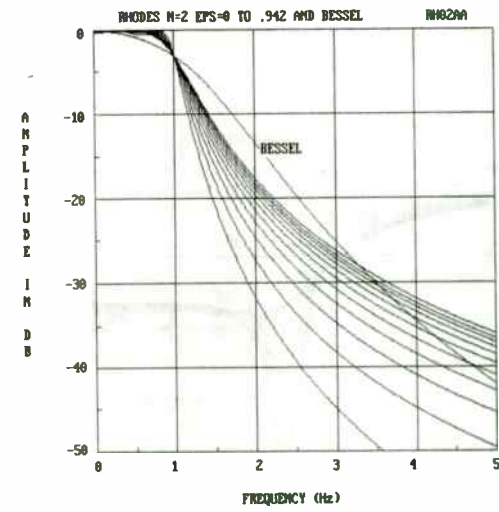
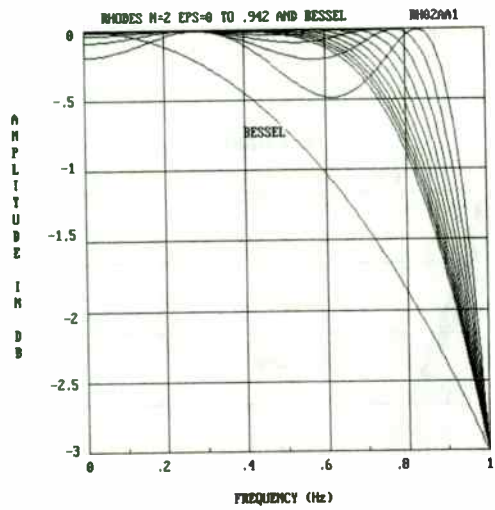
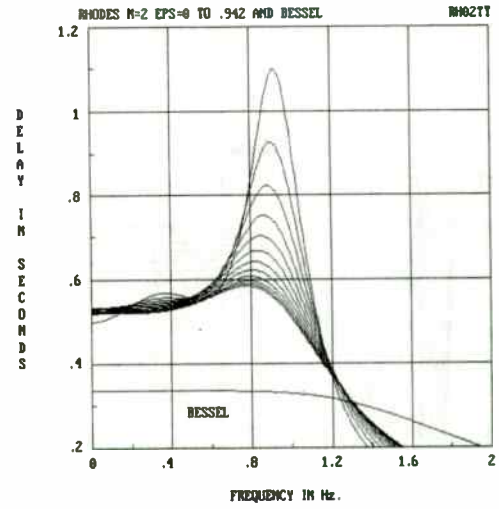
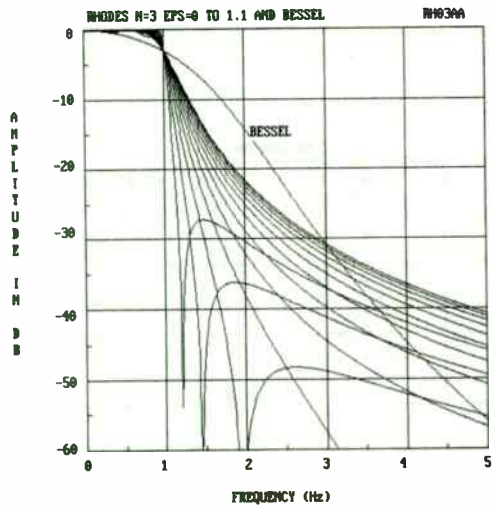
5.23833 4.04956 5.23833 6.06653 0.810244 6.06653 0.810244 5.29334  
2.8925 5.29334 2.8925 4.04956 5.23833 4.04956 5.23833  
5.69741 0 8.19165 0 5.69741 0 8.19165 0 5.21804 2.64515 5.21804 2.64515  
4.04938 5.16619 4.04938 5.16619 5.21804 2.64515 5.21804 2.64515  
4.04938 5.16619 4.04938 5.16619  
5.24351 0 14.5257 0 5.24351 0 14.5257 0 3.97724 5.07128 3.97724 5.07128  
4.95732 2.48631 4.95732 2.48631 3.97724 5.07128 3.97724 5.07128  
4.95732 2.48631 4.95732 2.48631  
4.91422 0 4.91422 0 24.6991 0 24.6991 0 4.67469 2.4364 4.67469 2.4364  
3.82734 5.00484 3.82734 5.00484 4.67469 2.4364 4.67469 2.4364  
3.82734 5.00484 3.82734 5.00484  
4.61018 0 4.61018 0 12.9601 0 12.9601 0 4.39391 2.43401 4.39391 2.43401  
3.63076 4.98863 3.63076 4.98863 4.39391 2.43401 4.39391 2.43401  
3.63076 4.98863 3.63076 4.98863  
4.30706 0 4.30706 0 10.5144 0 10.5144 4.1079 2.45678 4.1079 2.45678  
3.40767 5.01659 3.40767 5.01659 4.1079 2.45678 4.1079 2.45678  
3.40767 5.01659 3.40767 5.01659  
3.99432 0 3.99432 0 9.40105 0 9.40105 3.81017 2.49618 3.81017 2.49618  
3.16409 5.07953 3.16409 5.07953 3.81017 2.49618 3.81017 2.49618  
3.16409 5.07953 3.16409 5.07953  
3.66586 0 3.66586 0 8.78557 0 8.78557 3.49591 2.54833 3.49591 2.54833  
2.90017 5.17105 2.90017 5.17105 3.49591 2.54833 3.49591 2.54833  
2.90017 5.17105 2.90017 5.17105  
3.31733 0 3.31733 0 8.43089 0 8.43089 3.16136 2.61131 3.16136 2.61131  
2.61412 5.28704 2.61412 5.28704 3.16136 2.61131 3.16136 2.61131  
2.61412 5.28704 2.61412 5.28704  
2.9451 0 2.9451 0 8.24405 0 8.24405 2.80317 2.68415 2.80317 2.68415  
2.3036 5.42484 2.3036 5.42484 2.80317 2.68415 2.80317 2.68415  
2.3036 5.42484 2.3036 5.42484  
2.54571 0 2.54571 0 8.1822 0 8.1822 1.96622 5.58295 1.96622 5.58295  
2.41814 2.76649 2.41814 2.76649 1.96622 5.58295 1.96622 5.58295  
2.41814 2.76649 2.41814 2.76649  
2.11547 0 2.11547 0 8.22239 0 8.22239 2.00275 2.85852 2.00275 2.85852  
1.59898 5.76149 1.59898 5.76149 2.00275 2.85852 2.00275 2.85852  
1.59898 5.76149 1.59898 5.76149  
1.64903 0 1.64903 0 8.34593 0 8.34593 1.55138 2.96119 1.55138 2.96119  
1.19252 5.96363 1.19252 5.96363 1.55138 2.96119 1.55138 2.96119  
1.19252 5.96363 1.19252 5.96363  
1.12606 0 1.12606 0 8.52339 0 8.52339 1.03924 3.07707 1.03924 3.07707  
0.685696 6.19621 0.685696 6.19621 1.03924 3.07707 1.03924 3.07707  
0.685696 6.19621 0.685696 6.19621

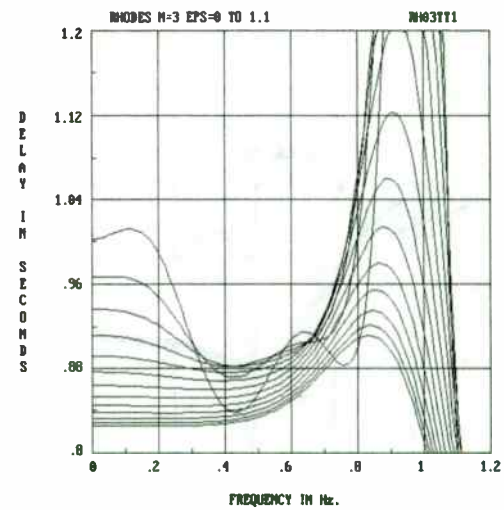
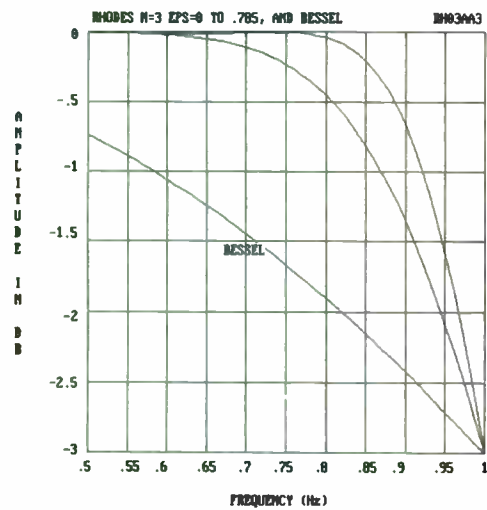
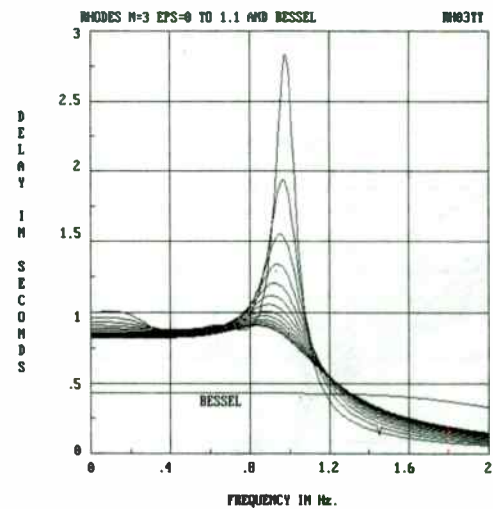
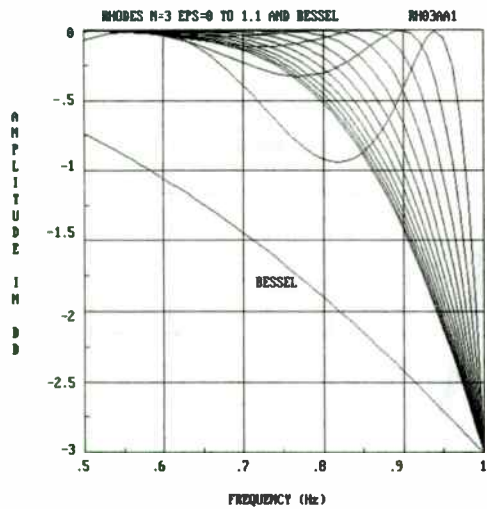
5.64808 1.00519 5.64808 1.00519 5.14481 3.05312 5.14481 3.05312 3.95749  
5.26306 3.95749 5.26306 5.64808 1.00519 5.64808 1.00519 5.14481  
3.05312 5.14481 3.05312 3.95749 5.26306 3.95749 5.26306  
5.65189 1.00552 5.65189 1.00552 5.14857 3.05426 5.14857 3.05426 3.96138  
5.26536 3.96138 5.26536 5.65189 1.00552 5.65189 1.00552 5.14857  
3.05426 5.14857 3.05426 3.96138 5.26536 3.96138 5.26536  
5.66654 1.00549 5.66654 1.00549 5.16143 3.05565 5.16143 3.05565 3.97331  
5.27098 3.97331 5.27098 5.66654 1.00549 5.66654 1.00549 5.16143  
3.05565 5.16143 3.05565 3.97331 5.27098 3.97331 5.27098  
5.70435 1.0005 5.70435 1.0005 5.18838 3.0495 5.18838 3.0495 3.99376 5.27549  
3.99376 5.27549 5.70435 1.0005 5.70435 1.0005 5.18838 3.0495  
5.18838 3.0495 3.99376 5.27549 3.99376 5.27549  
5.79956 0.972852 5.79956 0.972852 5.23647 3.01419 5.23647 3.01419 4.02182  
5.26978 4.02182 5.26978 5.79956 0.972852 5.79956 0.972852 5.23647  
3.01419 5.23647 3.01419 4.02182 5.26978 4.02182 5.26978  
6.06653 0.810244 6.06653 0.810244 5.29334 2.8925 5.29334 2.8925 4.04956

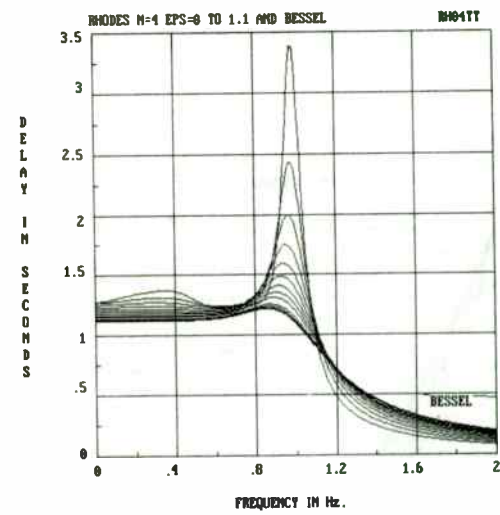
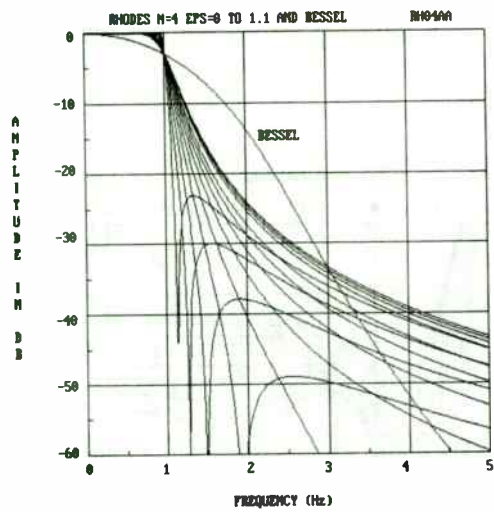
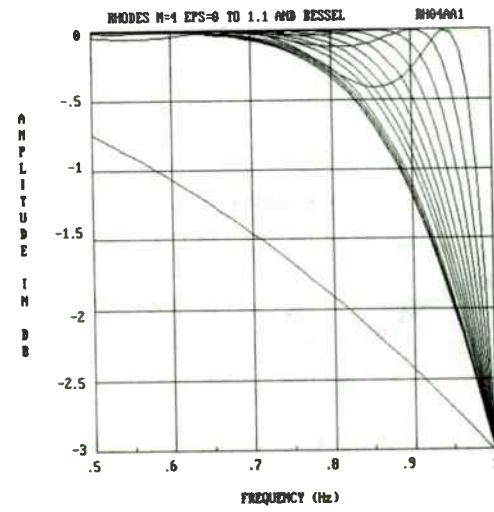
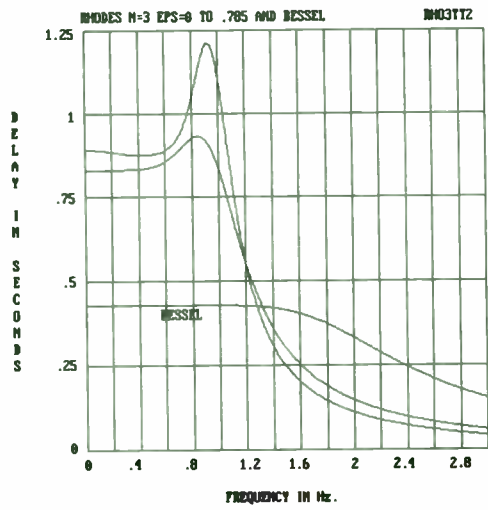
N = 8 POLES, THEN ZEROS

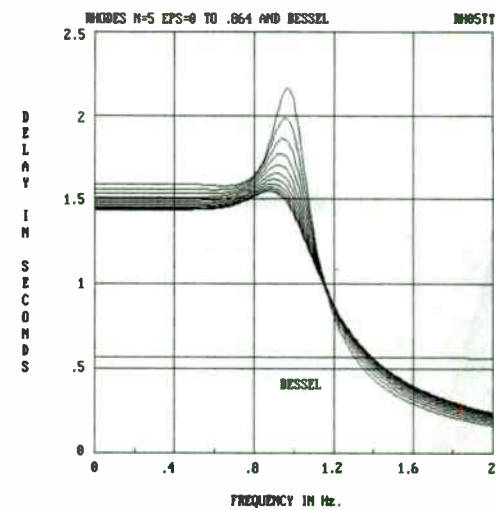
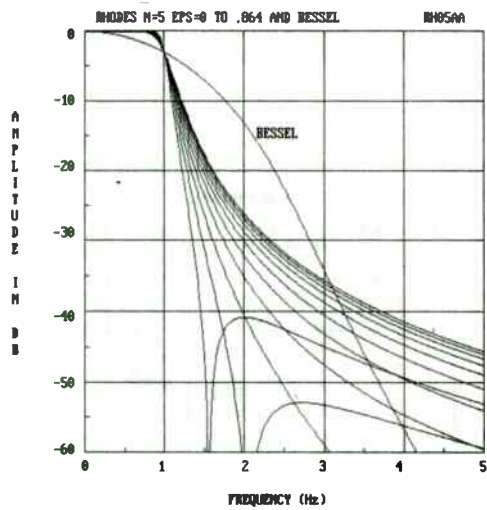
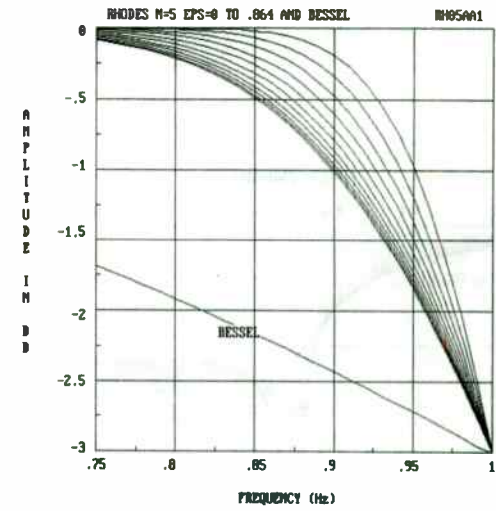
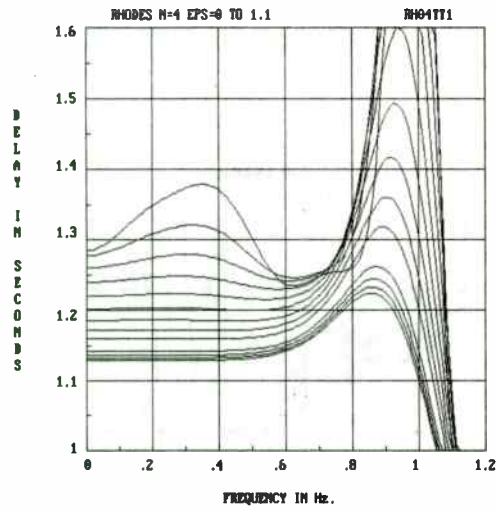
5.28835 0.427352 5.18695 1.31066 4.98698 2.17196 4.66261 3.07424 4.22253  
3.96062 3.58654 4.92122 2.76396 5.88267 1.30209 7.13648  
5.27507 0.428798 5.17404 1.3134 4.9751 2.17741 4.65176 3.08058 4.21366 3.96916  
3.57908 4.92979 2.75934 5.89266 1.2979 7.14426  
5.23527 0.433121 5.13533 1.32155 4.93944 2.19371 4.61914 3.09953 4.18694  
3.99476 3.55657 4.95553 2.74543 5.92269 1.2853 7.16766  
5.16907 0.440277 5.07084 1.33495 4.87995 2.22065 4.56452 3.13091 4.1421  
4.03735 3.51869 4.99852 2.72196 5.97295 1.26418 7.20685  
5.07665 0.450206 4.98064 1.35329 4.79651 2.25792 4.48755 3.1744 4.07868  
4.09678 3.46486 5.05884 2.68853 6.04376 1.23436 7.26213  
4.95818 0.46284 4.86475 1.37624 4.68901 2.30515 4.38775 3.2296 3.99604 4.17287  
3.3943 5.13663 2.64452 6.13554 1.19558 7.33394  
4.81381 0.478122 4.72316 1.40341 4.55725 2.36193 4.26458 3.29603 3.89347  
4.26538 3.30597 5.232 2.58916 6.24883 1.1475 7.42283  
4.64358 0.496017 4.55579 1.43444 4.40101 2.42789 4.11738 3.37315 3.77011  
4.37403 3.19865 5.34505 2.52145 6.38424 1.08969 7.52954  
4.44739 0.516522 4.36242 1.46898 4.21997 2.50271 3.94545 3.46047 3.62509  
4.49855 3.07088 5.47585 2.44019 6.54248 1.02162 7.65496  
4.22496 0.539678 4.14267 1.50676 4.01371 2.5862 3.74801 3.55751 3.45747  
4.63868 2.92103 5.62438 2.344 6.72428 0.942644 7.80019  
3.97574 0.56557 3.89591 1.54755 3.78162 2.67832 3.52417 3.66393 3.2663 4.79431  
2.74734 5.79057 2.23127 6.93039 0.851979 7.96654  
3.6989 0.594335 3.6213 1.59124 3.52289 2.77919 3.27287 3.7795 3.05059 4.9655  
2.54791 5.97427 2.10028 7.16158 0.748676 8.1556  
3.39329 0.626164 3.31763 1.63775 3.23643 2.88914 2.99289 3.90421 2.80924  
5.15256 2.3208 6.17531 1.94917 7.41866 0.631576 8.36926  
3.05735 0.661301 2.98335 1.68713 2.9208 3.00868 2.6827 4.0382 2.54096 5.35614  
2.06405 6.3935 1.77604 7.70246 0.499238 8.60975  
2.6891 0.700053 2.61649 1.73948 2.57417 3.13853 2.34049 4.18188 2.24422  
5.57723 1.77575 6.62879 1.579 8.01399 0.349866 8.87964  
2.28605 0.742787 2.21456 1.79503 2.19425 3.27965 1.96402 4.33588 1.91714  
5.81724 1.45404 6.88137 1.35615 8.3545 0.181209 9.1819  
1.8451 0.789945 1.77819 3.43322 1.77454 1.85408 1.55738 6.07799 1.55058 4.5011  
1.10565 8.72568 1.0971 7.15188 9.51093E-3 9.51978  
1.36251 0.842049 1.3225 3.60072 1.29269 1.91704 1.16208 6.36177 1.09687 4.6788  
0.825596 9.12978 0.703078 7.4417 0.225604 9.89678  
0.833693 0.899717 0.822932 3.78392 0.764498 1.98445 0.727765 6.67141 0.598885  
4.87057 0.514021 9.56985 0.269796 7.75325 0.470803 10.3166  
  
6.40234 0 -6.40234 0 6.19232 -1.98821 6.19232 1.98821 5.51785 -4.03501 5.51785  
4.03501 4.17326 -6.26245 4.17326 6.26245 -6.19232 -1.98821 -6.19232  
1.98821 -5.51785 -4.03501 -5.51785 4.03501 -4.17326 -6.26245 -4.17326  
6.26245  
6.40655 0 -6.40655 0 6.19647 -1.98886 6.19647 1.98886 5.52192 -4.03655 5.52192  
4.03655 4.17748 -6.26529 4.17748 6.26529 -6.19647 -1.98886 -6.19647  
1.98886 -5.52192 -4.03655 -5.52192 4.03655 -4.17748 -6.26529 -4.17748  
6.26529  
6.42387 0 -6.42387 0 6.21249 -1.98858 6.21249 1.98858 5.53554 -4.03853 5.53554  
4.03853 4.1903 -6.27229 4.1903 6.27229 -6.21249 -1.98858 -6.21249  
1.98858 -5.53554 -4.03853 -5.53554 4.03853 -4.1903 -6.27229 -4.1903  
6.27229  
6.47376 0 -6.47376 0 6.25353 -1.97709 6.25353 1.97709 5.56245 -4.03084 5.56245  
4.03084 4.21166 -6.27846 4.21166 6.27846 -6.25353 -1.97709 -6.25353  
1.97709 -5.56245 -4.03084 -5.56245 4.03084 -4.21166 -6.27846 -4.21166  
6.27846  
  
6.62457 0 -6.62457 0 6.35034 -1.91315 6.35034 1.91315 5.60325 -3.98738 5.60325  
3.98738 4.23901 -6.27387 4.23901 6.27387 -6.35034 -1.91315 -6.35034  
1.91315 -5.60325 -3.98738 -5.60325 3.98738 -4.23901 -6.27387 -4.23901  
6.27387  
7.35288 0 -7.35288 0 6.47288 -1.58437 6.47288 1.58437 5.6208 -3.85861 5.6208  
3.85861 4.26054 -6.2436 4.26054 6.2436 -6.47288 -1.58437 -6.47288  
1.58437 -5.6208 -3.85861 -5.6208 3.85861 -4.26054 -6.2436 -4.26054  
6.2436  
11.1292 0 -11.1292 0 6.07907 -1.23509 6.07907 1.23509 5.4976 -3.6728 5.4976  
3.6728 4.2452 -6.17969 4.2452 6.17969 -6.07907 -1.23509 -6.07907 1.23509  
-5.4976 -3.6728 -5.4976 3.6728 -4.2452 -6.17969 -4.2452 6.17969  
27.397 0 -27.397 0 5.71312 -1.1793 5.71312 1.1793 5.25552 -3.56943 5.25552  
3.56943 4.15793 -6.10822 4.15793 6.10822 -5.71312 -1.1793 -5.71312  
1.1793 -5.25552 -3.56943 -5.25552 3.56943 -4.15793 -6.10822 -4.15793  
6.10822  
0 18.508 0 18.508 5.39161 -1.17054 5.39161 1.17054 4.98528 -3.54454 4.98528  
3.54454 4.0024 -6.07148 4.0024 6.07148 -5.39161 -1.17054 -5.39161  
1.17054 -4.98528 -3.54454 -4.98528 3.54454 -4.0024 -6.07148 -4.0024  
6.07148  
0 13.1378 0 -13.1378 5.07733 -1.17838 5.07733 1.17838 4.70403 -3.56372 4.70403  
3.56372 3.80396 -6.08358 3.80396 6.08358 -5.07733 -1.17838 -5.07733  
1.17838 -4.70403 -3.56372 -4.70403 3.56372 3.80396 -6.08358 -3.80396  
6.08358  
0 11.3003 0 -11.3003 4.75544 -1.19551 4.75544 1.19551 4.4095 -3.61077 4.4095  
3.61077 3.57759 -6.1397 3.57759 6.1397 -4.75544 -1.19551 -4.75544  
1.19551 -4.4095 -3.61077 -4.4095 3.61077 -3.57759 -6.1397 -3.57759  
6.1397  
0 10.3739 0 -10.3739 4.41816 -1.21922 4.41816 1.21922 4.09759 -3.67831 4.09759  
3.67831 3.32782 -6.23285 3.32782 6.23285 -4.41816 -1.21922 -4.41816  
1.21922 -4.09759 -3.67831 -4.09759 3.67831 -3.32782 -6.23285 -3.32782  
6.23285  
0 9.84932 0 -9.84932 4.06042 -1.24826 4.06042 1.24826 3.76457 -3.76266 3.76457  
3.76266 3.05445 -6.3579 3.05445 6.3579 -4.06042 -1.24826 -4.06042  
1.24826 -3.76457 -3.76266 -3.76457 3.76266 -3.05445 -6.3579 -3.05445  
6.3579  
0 9.55567 0 -9.55567 3.67829 -1.28205 3.67829 1.28205 3.40709 -3.86189 3.40709  
3.86189 2.75548 -6.51135 2.75548 6.51135 -3.67829 -1.28205 -3.67829  
1.28205 -3.40709 -3.86189 -3.40709 3.86189 -2.75548 -6.51135 -2.75548  
6.51135  
0 9.42096 0 -9.42096 3.26819 -1.32034 3.26819 1.32034 2.42834 -6.69083 2.42834  
6.69083 3.02198 -3.97507 3.02198 3.97507 -3.26819 -1.32034 -3.26819  
1.32034 -2.42834 -6.69083 -2.42834 6.69083 -3.02198 -3.97507 -3.02198  
3.97507  
0 9.41149 0 -9.41149 2.8266 -1.36309 2.8266 1.36309 2.0704 -6.8948 2.0704  
6.8948 2.60603 -4.10196 2.60603 4.10196 -2.8266 -1.36309 -2.8266 1.36309  
-2.0704 -6.8948 -2.0704 6.8948 -2.60603 -4.10196 -2.60603 4.10196  
0 9.50989 0 -9.50989 2.34966 -1.41047 2.34966 1.41047 1.67893 7.12313 1.67893  
7.12313 2.15576 -4.24288 2.15576 4.24288 -2.34966 -1.41047 -2.34966  
1.41047 -1.67893 -7.12313 -1.67893 7.12313 -2.15576 -4.24288 -2.15576  
4.24288  
0 9.70159 0 -9.70159 1.83249 -1.46286 1.83249 1.46286 1.66624 -4.39906 1.66624  
4.39906 1.24658 -7.37899 1.24658 7.37899 -1.83249 -1.46286 -1.83249  
1.46286 -1.66624 -4.39906 -1.66624 4.39906 -1.24658 -7.37899 -1.24658  
7.37899  
0 9.95865 0 -9.95865 1.26081 -1.52122 1.26081 1.52122 1.11838 -4.57386 1.11838  
4.57386 0.720573 -7.67088 0.720573 7.67088 -1.26081 -1.52122 -1.26081  
1.52122 -1.11838 -4.57386 -1.11838 4.57386 -0.720573 -7.67088 -0.720573  
7.67088

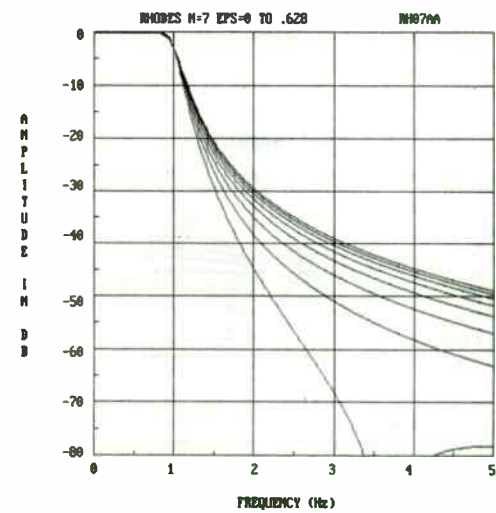
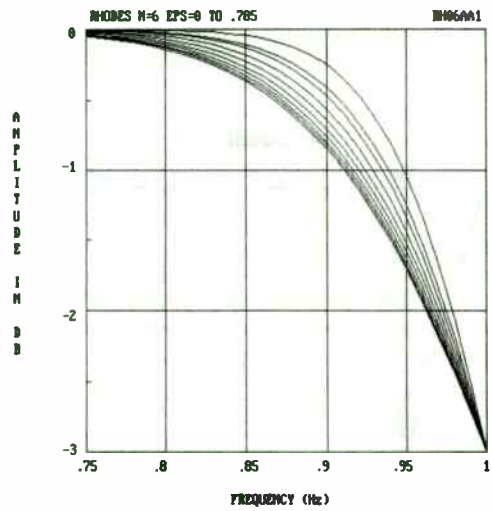
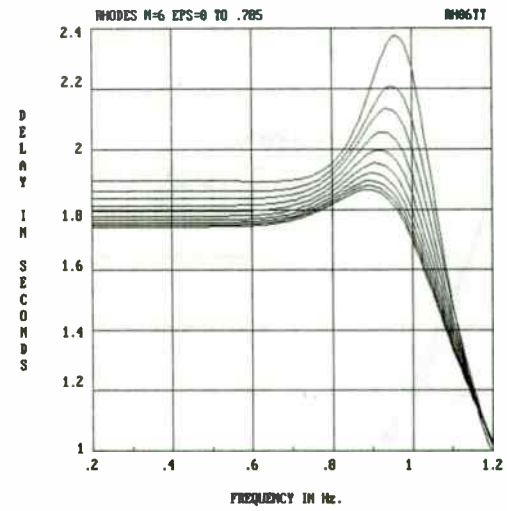
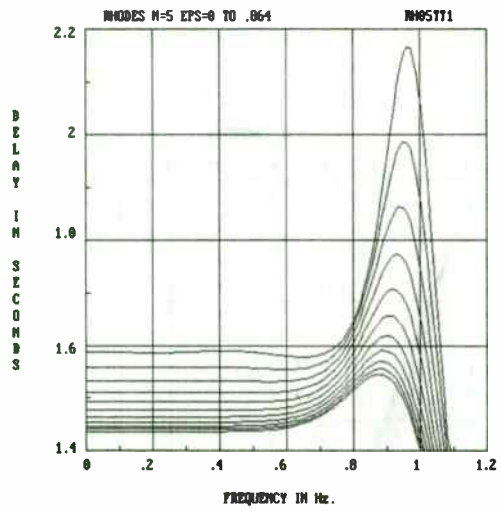


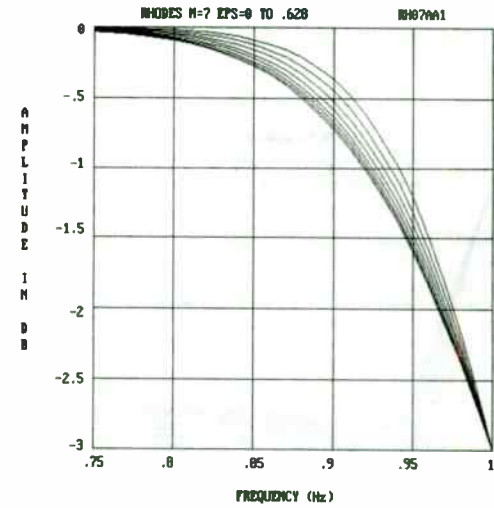
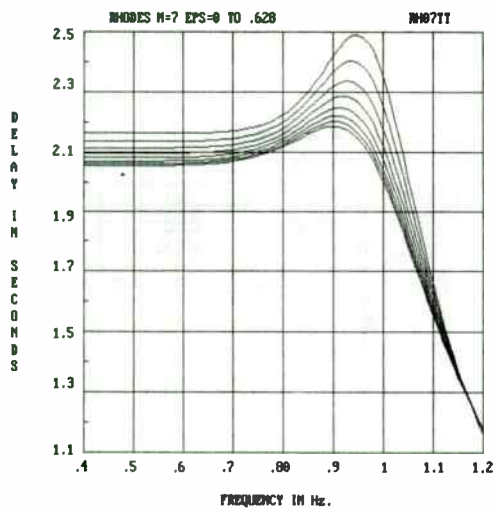
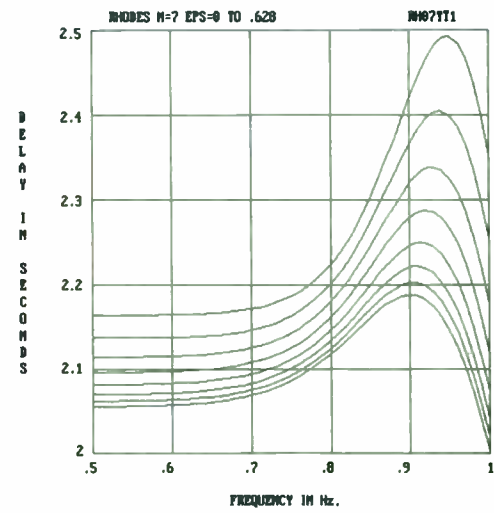
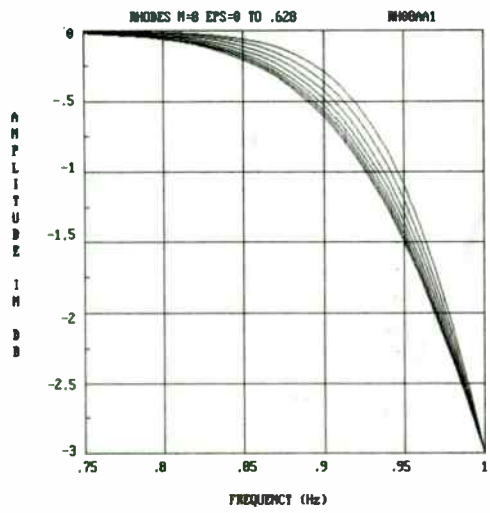


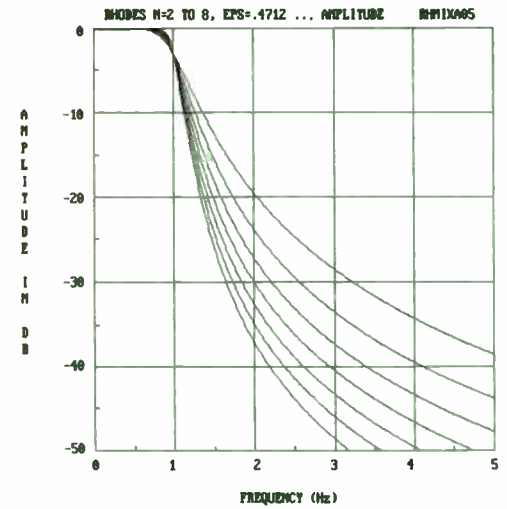
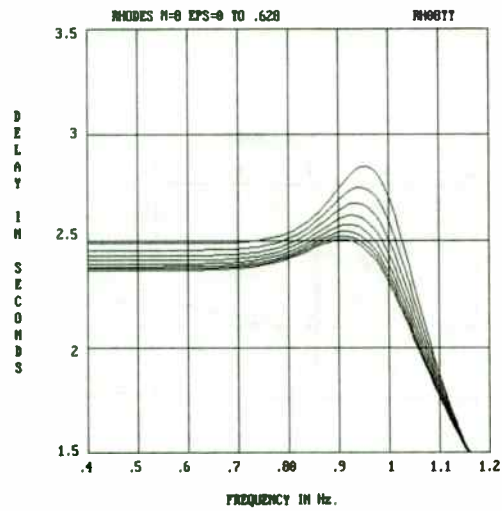
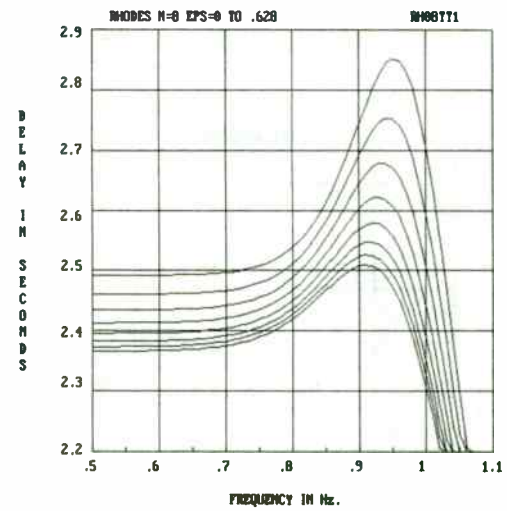
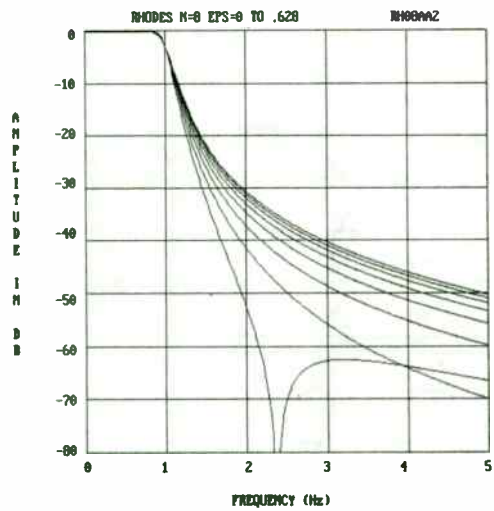


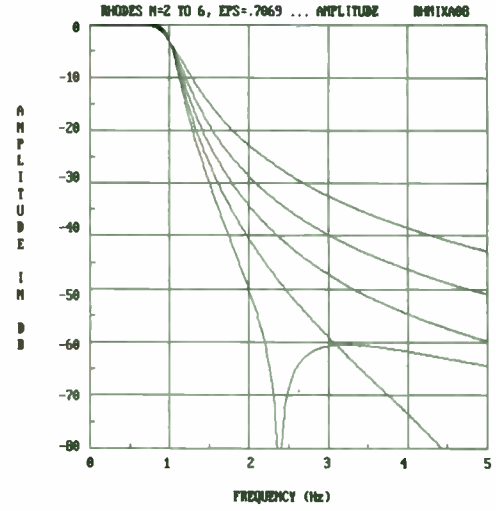
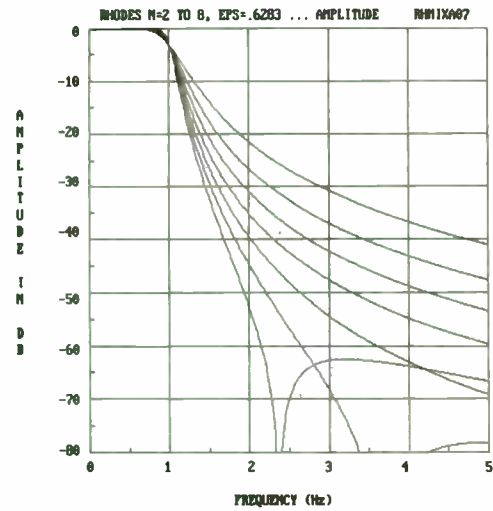
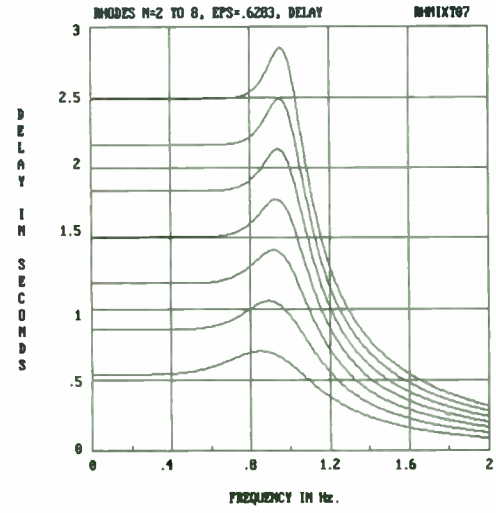
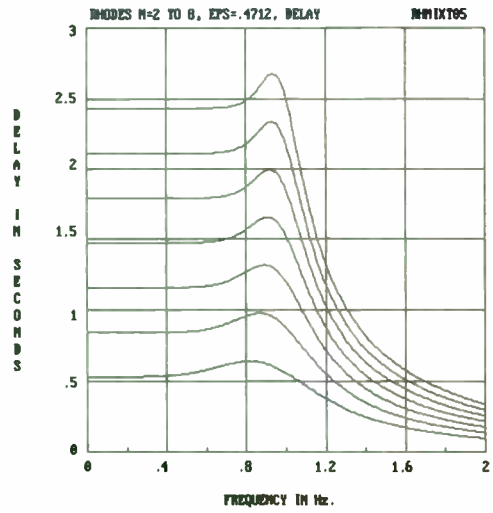




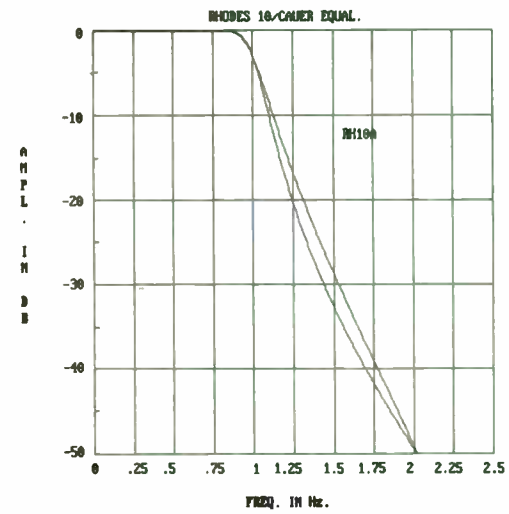
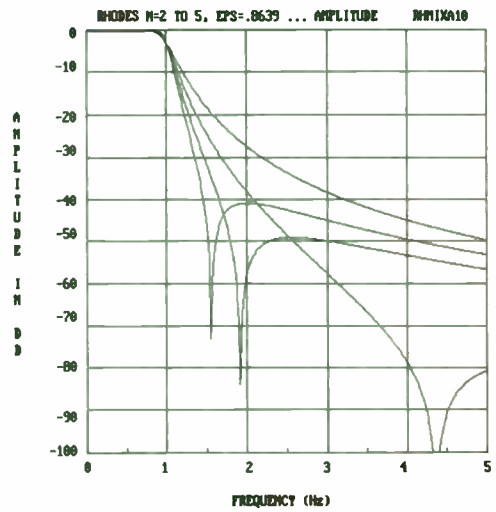
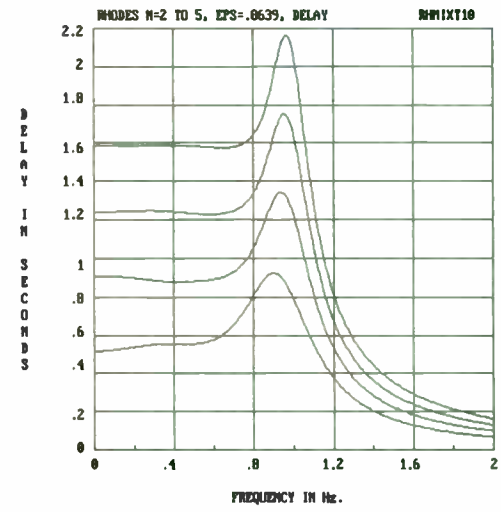
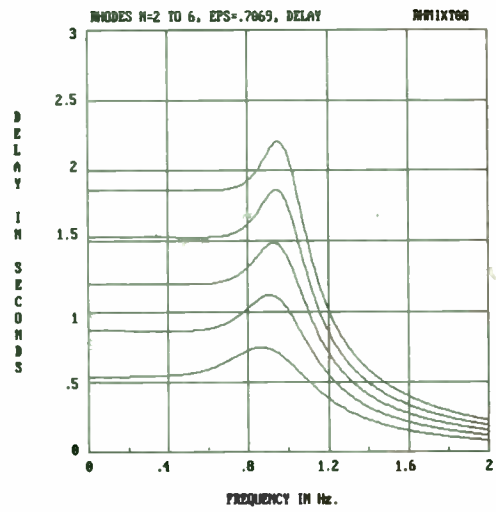


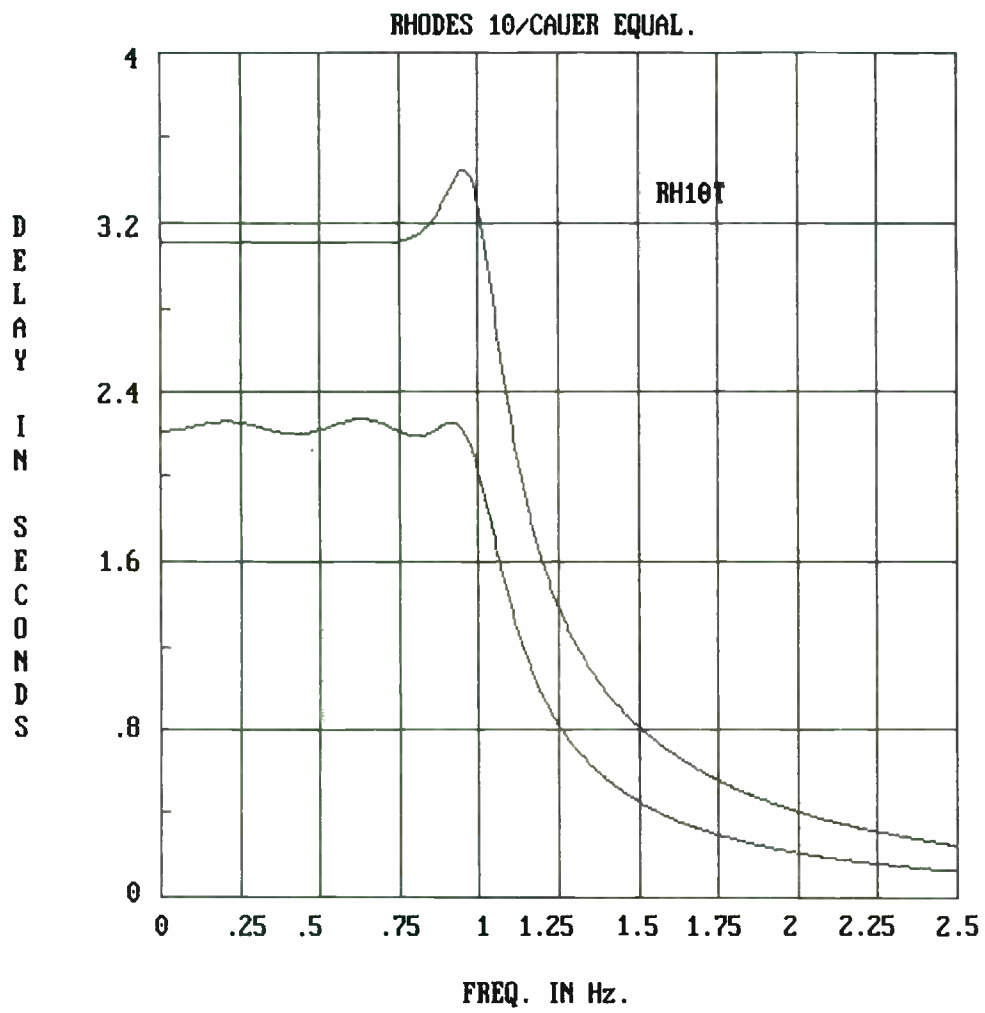












## SYNTHESIS OF LOW-PASS FILTERS CONTAINING QUADS OF ZEROS

Mark Mell  
ST Microsonics Inc.  
60 Winter Street  
Weymouth, MA 02188

### ABSTRACT

The article below details the calculations required to synthesize a special type of low-pass filter, namely one that contains quads of zeros. Including a quad of zeros with any existing set of pole locations has the effect of modifying only the amplitude response. The delay variation of the network remains consistent with the pole locations alone, while the selectivity of the filter can be improved.

In Darlington's<sup>1</sup> method of synthesis, an impedance function, which represents the input impedance of a network, is reduced and the network synthesized by removing partial sections. Youla<sup>2</sup> developed a procedure, based on Darlington's method, that generates the partial networks and reduces the impedance expression. If the network contained a quad of zeros, a Darlington D section (shown in figure 1) was used to realize the partial network.

Later Scanlan and Rhodes<sup>3</sup> developed their own procedure. They extended it to include the cascade synthesis of an admittance function which included a transformerless realization of the D section. This partial network has a very significant limitation. It contains a negative

element value whenever the real component of the zero is greater than the imaginary. This negative element value, a capacitor, cannot be absorbed within the network.

It is the purpose of this article, utilizing information contained within both previously referenced articles, to synthesize a network containing complex zeroes and to present an additional partial section realization of a Darlington D section.

For this example, a third order Rhodes<sup>4</sup> polynomial, which results in a transfer function of a sixth order with one quad of zeros, was selected. The poles and zeros of the transfer function are tabulated in table 1. In analyzing the pole and zero locations, the attenuation at one radian is approximately 3 dB with respect to the minimum attenuation frequency of 0.404626 radians. The resulting magnitude at this frequency is 16.20485. A multiplying constant equal to the reciprocal will reduce the magnitude to 1 and result in a network with a termination ratio as close to 1 as possible. The resulting transfer function  $H(s)$  is also included in table 1.

The characteristic function  $K(s)$  can now be determined as described by Saal and Ulbrich<sup>5</sup>. The impedance function  $Z(s)$  is then formed utilizing both the transfer and characteristic functions. The above calculations are summarized in table 2.

The next step in our synthesis is to remove partial sections of the network and reduce the impedance expression. The equations to calculate the element values of the partial sections contain variables which are determined by analyzing the impedance function and its derivative at the zero to be removed. In the procedure outlined by Scanlan and Rhodes, these variables are coefficients of the polynomials contained within the transmission matrix of the type section to be realized. In the case of the Darlington D section, the equivalences between the coefficients of the transmission matrix and the variables in the equations developed by Youla

are tabulated in table 3.

The admittance function is formed by taking the inverse of the impedance function and the derivative of the admittance function by the quotient rule for differentiation. The functions are then evaluated at  $s = 1.294724 + j.04419413$ . The results of the analysis are substituted for the variables in Youla's equations. The coefficients of the transmission matrix polynomials are then found from the equations in table 3. The above calculations are summarized in table 4.

The schematic of the partial network used to realize the Darlington D section is shown in figure 2, and the equations required to determine the element values in table 5. The negative capacitor resulting from this realization can usually be absorbed. The negative inductor is equal to the mutual inductance of a tapped coil represented by L3, L4, and L5.

The admittance function is then simplified employing the procedures described by Youla as follows. The first admittance remainder function  $W(s)$  is formed. This function contains, within its numerator and denominator, the quad of zeros removed by the partial network.  $W(s)$  is easily reduced and the second admittance remainder function  $K(s)$  is formed. Again, the quad of zeros is found within the numerator and denominator and is easily removed. The above calculations are summarized in table 6.

The admittance function  $K(s)$  can then be reduced and the remaining network generated by the methods described by Saal and Ulbrich for networks whose zeros occur at infinity. These remaining calculations are summarized in table 7, and the complete network is shown in figure 3.

## Appendix A

The following procedure was utilized to develop the expressions required to calculate the component values for this realization of a Darlington D section.

The transfer characteristic of any reactive network can be represented by a chain matrix as shown in figure 4. Variables A, B, C, and D represent functions of the complex frequency variable  $s$ . All four functions share a common denominator consisting of zeros removed by the section. These four functions are related to the input and output admittances of the network they describe by the equivalences shown in table 8.

In the case of the Darlington D section realization shown in figure 2, the input admittances at port 1 and port 2 are listed in table 8.

At this point a few clarifications should be made with regard to the admittance equations. When the input admittance is calculated at port 1 with port 2 short circuited, the constant term in polynomial D is equal to one. In order to agree with the design procedure developed by Youla, both the numerator and denominator must be multiplied by a constant equal to  $|PO|^4$ . The same is true for the input admittance at port 2 with port 1 short circuited.

When the input admittance at port 1 is calculated with port 2 open circuited, the constant term in polynomial D is equal to the reciprocal of the product of C1, C2, C3, and C5. A multiplying constant is again used to correct this value. The resulting B polynomial is of a fifth order with the coefficient multiplied by the highest term equal to zero.

Another useful relationship between the functions contained within the transmission matrix for

a reactive network is that the product of A and D less the product of B and C is equal to one. Utilizing this relationship and the equations developed above, the expressions listed in table 5 can be developed.

#### Acknowledgment

I wish to acknowledge the assistance from William B. Lurie, an independent filter consultant residing in Boynton Beach FL, in preparing this paper. Also appreciated is the contribution by fellow ST Microsonics engineer Paul Mailloux for reducing the rather tedious task of polynomial manipulation by utilizing commercially available software.

#### References

- <sup>1</sup> S. Darlington, "Synthesis of reactance 4-poles which produce prescribed insertion loss characteristics", *Journal of Mathematics and Physics*, vol 18, pp.257-353; September, 1939
- <sup>2</sup> D. C. Youla, "A new theory of cascade synthesis", *IRE Transactions on Circuit Theory*, CT-8, pp. 244-260; September, 1961
- <sup>3</sup> J. O. Scanlan, J. D. Rhodes, "Unified theory of cascade synthesis", *Proceedings IEE*, Vol. 117, No 4, April 1970
- <sup>4</sup> J. D. Rhodes, "A low-pass prototype network for microwave linear phase filters", *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-18, No. 6, June 1970
- <sup>5</sup> R. Saal, E. Ulbrich, "On the design of filters by synthesis", *IRE Transactions on Circuit Theory*, Vol. CT-5, pp. 284-317, December 1958

Poles: -0.5486778 +/-j 0.1796838  
 -0.4881322 +/-j 0.5696584  
 -0.2247229 +/-j 0.9359748

Zeros: 1.2947245 +/-j 0.04419549  
 -1.2947245 +/-j 0.04419549

$$H(s) = \frac{E(s)}{P(s)mc}$$

$$E(s) = s^6 + 2.52307s^5 + 3.82595s^4 + 3.74856s^3 + 2.43434s^2 + 0.95805s + 0.1738158$$

$$P(s) = s^4 - 3.34871s^2 + 2.81657$$

$$mc = 0.0617099$$

TABLE 1

$$K(s) = \frac{F(s)}{P(s)mc}$$

$$F(s)F(-s) = E(s)E(-s) - P(s)P(-s)mc^2$$

$$E(s)E(-s) = s^{12} + 1.28605s^{10} + .590893s^8 + .0888547s^6 + .0734439s^4 - .0715987s^2 + .0302119$$

$$P(s)P(-s) = s^8 - 6.69743s^6 + 16.8470s^4 - 18.8638s^2 + 7.93308$$

$$F(s)F(-s) = s^{12} + 1.28605s^{10} + .587085s^8 + .114359s^6 + .00928843s^4 + .000236750s^2 + 0.0000185978$$

$$F(s)F(-s) = (s+j.134384)(s-j.134384)(s+j.404625)(s-j.404625)(s+j.134384)(s-j.134384)(s+j.404625)(s-j.404625)(s+j.679149)(s-j.679149)(s+j.679149)(s-j.679149)$$

$$F(s) = (s+j.134384)(s-j.134384)(s+j.404625)(s-j.404625)(s+j.679149)(s-j.679149)$$

$$F(s) = s^6 + .643025s^4 + 8.68020s^2 + .00136374$$

$$Z(s) = \frac{E(s)+F(s)}{E(s)-F(s)}$$

$$Z(s) = \frac{2s^6 + 2.52307s^5 + 4.46898s^4 + 3.74856s^3 + 2.52114s^2 + 0.958046s + 0.175180}{2.52307s^5 + 3.18293s^4 + 3.74856s^3 + 2.34754s^2 + 0.958046s + 0.172452}$$

TABLE 2

$$PO = a_o + jw_o$$

$$| PO |^2 = a_o^2 + w_o^2$$

$$a_1 = | PO |^2 (I_4 + J_4) + I_2 J_2$$

$$b_1 = | PO |^2 (I_2 + J_2)$$

$$c_1 = | PO |^2 (I_3 + J_3)$$

$$d_1 = | PO |^2 (I_1 + J_1) + I_3 J_3$$

$$a_2 = I_4 J_4$$

$$b_2 = I_4 J_2 + I_2 J_1$$

$$c_2 = I_3 J_4 + I_1 J_3$$

$$d_2 = I_1 J_1$$

TABLE 3

$$Y'(s) = \frac{-5.04613s^{10} - 12.7317s^9 - 19.2466s^8 - 18.7803s^7 - 13.0872s^6 - 7.00238s^5 - 3.01111s^4 - .852402s^3 - .135648s^2 - .0470711s + .00261304}{4s^{12} + 10.0922s^{11} + 24.2418s^{10} + 37.5453s^9 + 48.9721s^8 + 50.0587s^7 + 42.1207s^6 + 28.3482s^5 + 15.1044s^4 + 6.14408s^3 + 1.80116s^2 + .335660s + .0306878}$$

$$Y(PO) = Y_o$$

$$Y'(PO) = Y_o'$$

$$Y_o = R_o + j X_o$$

$$Y(1.294724 + j .04419549) = .7032868 - j .01228522$$

$$Y'(1.294724 + j .04419549) = -.2780805 + j .00458396$$

$$I_1 = I_4^{-1} = \frac{R_o - X_o}{\frac{a_o}{a_o} \frac{w_o}{w_o}} \quad I_2 = \frac{2 | Y_o |^2}{\frac{R_o}{a_o} + \frac{X_o}{w_o}} \quad I_3 = \frac{2}{\frac{R_o}{a_o} + \frac{X_o}{w_o}}$$

$$X = \frac{a_o X_o}{w_o} + j \frac{w_o R_o}{a_o} \quad \overline{Y_o} = \text{Complex conjugate of } Y_o$$

$$W_o = U_o + j V_o = \frac{-I_1 PO Y_o' \overline{Y_o} - Y_o \overline{X}}{PO Y_o' - X} \quad J_1 = J_4^{-1} = \frac{\frac{U_o}{a_o} - \frac{V_o}{w_o}}{\frac{U_o}{a_o} + \frac{V_o}{w_o}} \quad \text{ect.}$$

$$I_1 = 3.096183 \quad I_2 = 3.730969 \quad I_3 = 2.435553 \quad I_4 = .3229783$$

$$X = -.35990035 + j .02400673 \quad U_o + j V_o = .36387704 - j .00959050$$

$$J_1 = 7.776645 \quad J_2 = 4.137729 \quad J_3 = 4.015679 \quad J_4 = .1285902$$

$$a_1 = 28.32514 \quad a_2 = 24.0779 \quad b_1 = 10.82688 \quad b_2 = 12.74646$$

$$c_1 = 13.20575 \quad c_2 = 30.3508 \quad d_1 = 15.74022 \quad d_2 = .0415318$$

TABLE 4

$$C1 = \frac{d_2 - 1}{c_2} \quad C2 = \frac{1 - d_2}{d_2 c_2} \quad C5 = \frac{1}{c_2}$$

$$L2 = \frac{c_2}{a_1 + d_1 - 4(w_0^2 - a_0^2) - \frac{c_1 (d_2 - 1)^2}{d_2 c_2}}$$

$$L3 = \frac{c_1 [a_1 - 2(w_0^2 - a_0^2) - \frac{c_1 (1 - d_2)}{d_2 c_2}]}{PO^4 [a_1 + d_1 - 4(w_0^2 - a_0^2) - \frac{c_1 (1 - d_2)^2}{d_2 c_2}]}$$

$$L4 = \frac{c_1 [d_1 - 2(w_0^2 - a_0^2) - \frac{c_1 (d_2 - 1)}{d_2 c_2}]}{PO^4 [a_1 + d_1 - 4(w_0^2 - a_0^2) - \frac{c_1 (1 - d_2)^2}{d_2 c_2}]}$$

$$L5 = -\frac{L3 - L4}{L3 + L4}$$

TABLE 5

$$W(s) = \frac{(I_1 s^2 + |PO|^2) Y(s) - I_2 s}{I_4 s^2 + |PO|^2 - I_3 s Y(s)}$$

$$W(s) = \frac{.34994s^7 + .44146s^6 - .83303s^5 - 1.3756s^4 - .14893s^3 + .89929s^2 + .95426s + .28942}{.64596s^8 + .8149s^7 - 1.345s^6 - 2.30712s^5 - .81541s^4 + .88295s^3 + 1.9543s^2 + 1.1878s + .294}$$

$$W(s) = \frac{(s - .49985)(s - 1.2947 + j.044195)(s - 1.2947 - j.044195)(s + 1.2947 + j.044195)(s + 1.2947 - j.044195)(s - .38084 + j.66515)(s - .38084 - j.66515)}{(s - .41256 + j.2717)(s - .41256 - j.2717)(s - 1.2947 + j.044195)(s - 1.2947 - j.044195)(s + 1.2947 + j.044195)(s + 1.2947 - j.044195)(s - .2182 + j.78394)(s - .2182 - j.78394)}$$

$$W(s) = \frac{.349936s^3 + .441456s^2 + .338804s + .102756}{.645957s^4 + .81489s^3 + .817977s^2 + .421732s + .104381}$$

$$K(s) = \frac{(J_1 s^2 + |PO|^2) W(s) - J_2 s}{J_4 s^2 + |PO|^2 - J_3 s W(s)}$$

$$K(s) = \frac{.0485343s + .0612276}{.0830637s^2 + .104788s + .062196}$$

TABLE 6



$$B(s) = \frac{1}{K(s)} \quad C = \frac{B(s)}{s} = \frac{.0830637s^2 + .104788s + .062196}{.0485343s^2 + .0612276s} = 1.71144$$

$$B'(s) = B(s) - Cs = \frac{.0830637s^2 + .104788s + .062196 - 1.71144s(.0485343s + .0612276)}{.0485343s + .0612276}$$

$$X(s) = \frac{1}{B'(s)} \quad L = \frac{X(s)}{s} = \frac{.0485343s + .0612276}{.062196s} = .78034$$

$$X'(s) = X(s) - Ls = \frac{.0485343s + .0612276 - .78034s(.062196)}{.062196}$$

$$X'(s) = \frac{.0612276}{.062196} = R = .9844299$$

TABLE 7

Input admittance at port 1 when port 2 is short circuited.

$$y = \frac{D}{C} = \frac{|PO|^4 + d_1 s^2 + d_2 s^4}{c_1 s + c_2 s^3}$$

Input admittance at port 2 when port 1 is short circuited.

$$y = \frac{A}{C} = \frac{|PO|^4 + a_1 s^2 + a_2 s^4}{c_1 s + c_2 s^3}$$

Input admittance at port 1 when port 2 is open circuited.

$$y = \frac{B}{D} = \frac{b_1 s + b_2 s^3 + b_3 s^5}{|PO|^4 + d_1 s^2 + d_2 s^4}$$

$$a_1 = \left[ \frac{L_5 + L_2 + L_3}{C_3} + \frac{L_3 + L_4}{C_2} + \frac{L_3 + L_4}{C_5} \right] |PO|^4$$

$$a_2 = \left[ \frac{1}{C_2} + \frac{1}{C_5} \right] \left[ \frac{L_3(L_5 + L_2)}{C_3} + \frac{L_4(L_5 + L_2)}{C_3} + \frac{L_4 L_3}{C_3} \right] |PO|^4$$

$$b_1 = [C_1 + C_2 + C_3] |PO|^4$$

$$b_2 = [ (L_3 + L_4) (C_1 C_2 + C_1 C_5 + C_2 C_5 + C_3 C_5) + C_2 C_3 (L_2 + L_4 + L_5) + C_1 C_3 (L_2 + L_3 + L_5) ] |PO|^4$$

$$b_3 = [ L_3 L_4 + L_3 (L_2 + L_5) + L_4 (L_2 + L_5) ] [ C_1 C_3 C_5 + C_1 C_2 C_3 + C_2 C_3 C_5 ] |PO|^4$$

$$c_1 = [L_3 + L_4] |PO|^4$$

$$c_2 = \left[ \frac{L_3(L_2 + L_5)}{C_3} + \frac{L_4(L_2 + L_5)}{C_3} + \frac{L_3 L_4}{C_3} \right] |PO|^4$$

$$d_1 = \left[ \frac{L_5 + L_2 + L_4}{C_3} + \frac{L_3 + L_4}{C_1} + \frac{L_3 + L_4}{C_5} \right] |PO|^4$$

$$d_2 = \left[ \frac{1}{C_1} + \frac{1}{C_5} \right] \left[ \frac{L_3(L_5 + L_2)}{C_3} + \frac{L_4(L_5 + L_2)}{C_3} + \frac{L_4 L_3}{C_3} \right] |PO|^4$$

TABLE 8

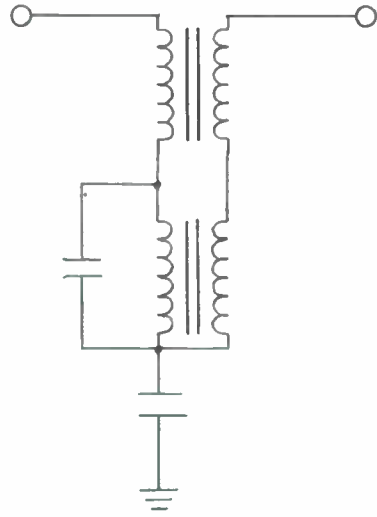


FIGURE 1

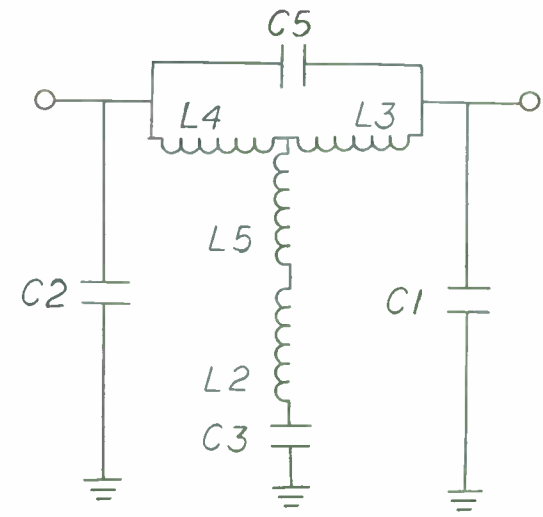
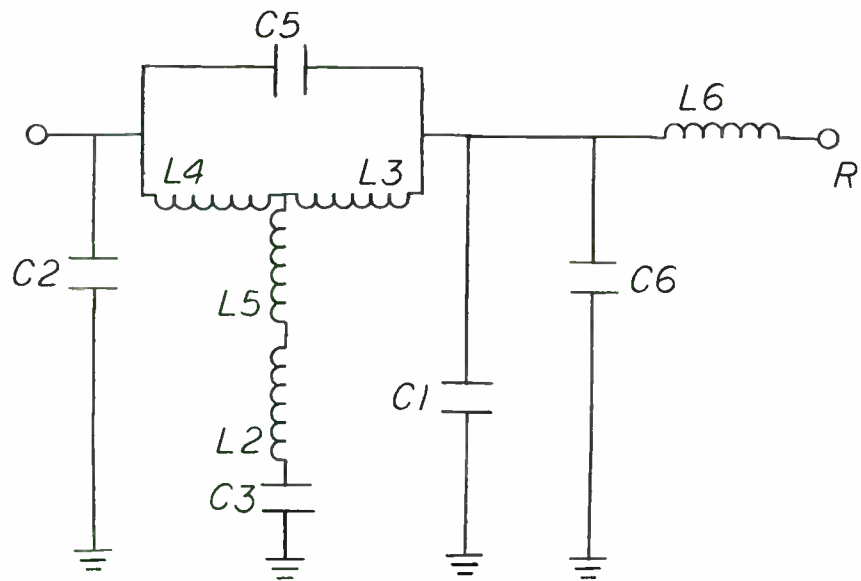
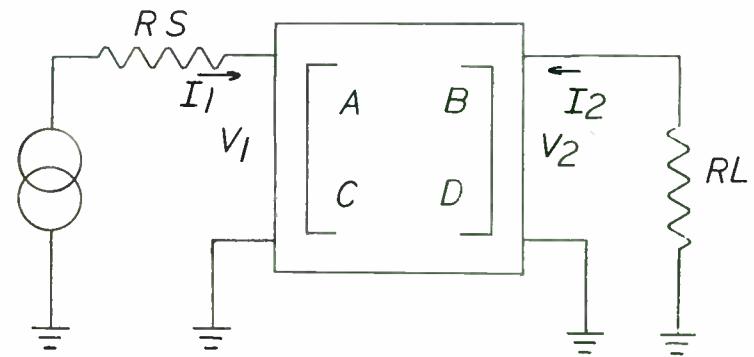


FIGURE 2



C1 = -.031577    C2 = .760372    C3 = 3.115197    C5 = .032948  
 L2 = .73777    L3 = 2.46548    L4 = 2.223111    L5 = -1.16902  
 C6 = 1.71144    L6 = .780345    R = .98443

FIGURE 3



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

$$V_1 = A V_2 + B(-I_2)$$

$$I_1 = C V_2 + D(-I_2)$$

FIGURE 4



1000 Ω = 1000 Ω  
 1000 Ω = 1000 Ω  
 1000 Ω = 1000 Ω  
 1000 Ω = 1000 Ω

FIG. 10-10

## Modern Methods of Frequency Control

David J. Brunell

Department of Electrical Engineering

Florida Atlantic University

Boca Raton, FL 33431

### 1 Introduction

All communications systems have requirements for the generation and control of signals at specific frequencies. Depending on the nature of the system, the accuracy with which these signals must maintain their nominal frequency can vary widely, from a few percent to less than one part per billion. This paper reviews the major methods of frequency control in use today.

### 2 Oscillators

An electronic oscillator can be defined as any circuit which produces a periodic output signal without an input signal. A sinusoidal oscillator produces a sine wave output which may include harmonics of the primary or *fundamental* frequency of oscillation. Since all waveforms may be represented through Fourier analysis as a series of sine waves, only oscillators of this type will be considered. The signal power in an oscillator is generated by converting power in a dc supply to an alternating ac waveform at the desired frequency. Most oscillators in use today are designed to provide fairly low power output levels (less than one watt).

### 2.1 Basic Feedback Amplifier Theory

Originally, oscillators were made by connecting a spark gap, which is a source of high-voltage white noise, to an L-C circuit which was resonant at the desired frequency of operation. Most modern oscillators utilize an amplifying device with an appropriate feedback network. Figure 1 is a block diagram illustrating the basic feedback amplifier configuration. An amplifying circuit of gain  $A$  produces an output signal  $V_o$ , which is proportional to the sum of the input signal  $V_i$  and a fraction of the output signal,  $V_o\beta$ , where  $\beta$  is the transfer function of the *feedback network*. The overall transfer function of this circuit can be shown to be [1]

$$\frac{V_o}{V_i} = \frac{A(j\omega)}{1 - A(j\omega)\beta(j\omega)}. \quad (1)$$

In order to have an output signal with no input signal, the denominator must go to zero, i.e.,

$$A(j\omega)\beta(j\omega) = 1 \angle 0^\circ. \quad (2)$$

In other words, a signal must be fed back at the input which is the same amplitude and in phase with (or  $360^\circ$  lagging) the original signal in order for oscillation to occur. In most implementations of the feedback oscillator, the amplifying device is a transistor and the feedback network consists of reactive components designed to resonate at the desired frequency.

## 2.2 LC Oscillators

### 2.2.1 Circuit Configuration

By using a feedback network consisting of an inductor and capacitor in series, as illustrated in Figure 2, it can be shown that the conditions necessary for oscillation will occur only at a frequency given by [2]

$$f_{osc} = \frac{1}{2\pi\sqrt{LC}}. \quad (3)$$

There are many different realizations of the LC oscillator, depending on the active device used and the particular circuit configuration. A typical example is shown in Figure 3, a Pierce oscillator using a bipolar junction transistor. This circuit will oscillate at a frequency of

$$f_{osc} = \frac{1}{2\pi\sqrt{LC_1C_2}} \sqrt{\frac{C_1+C_2}{C_1C_2}}. \quad (4)$$

It should be pointed out that the accuracy in the frequency of oscillation is only as great as the tolerance of the components used, which is seldom greater than 1 percent for inductors and capacitors used at radio frequencies. For this reason, LC oscillators usually include a variable reactance in the circuit to permit adjustment of operating frequency.

### 2.2.2 Stability of LC Oscillators

The stability of any oscillator can be broken down into two categories: frequency drift and phase noise. Drift is a long-term phenomenon and causes the operating

frequency to digress from its original value. The three main causes of drift are component variations due to temperature, supply voltage variation, and load impedance variation. Modern integrated voltage regulators can produce tightly-regulated outputs that eliminate this consideration as a source of error. Through the use of intermediate buffer stages, the effect of load variations can be nullified.

Temperature variations can have a significant impact on the stability of LC oscillator circuits. Ceramic capacitors for use at RF frequencies have temperature coefficients on the order of 200 ppm/°C. Air-wound inductors are fairly stable, but the overall temperature stability of free-running LC oscillators makes them unusable for most communications circuits.

The short-term stability of an oscillator in general refers to the presence or absence of phase noise in its output. This noise appears as spurious sidebands in the output signal and can be determined by taking the differential  $d\phi/df|_{f=f_0}$ . For the LC circuit it can be shown that the phase noise is directly proportional to the Q or quality factor of the circuit, with higher Q values providing higher short-term stability [3]. This also turns out to be a shortcoming of LC oscillators, since Q values higher than 100 are difficult to obtain. For a more complete treatment of oscillator noise, the reader is referred to an excellent paper by Edson [4].

## 2.3 Crystal Oscillators

It is because of the stability limitations of LC oscillators that crystal oscillators were developed, and they are by far the most widely used type of frequency standard. Before the merits of the crystal oscillator can be understood, the crystal itself must be

understood.

### 2.3.1 Quartz Crystal Resonators

A crystal is nothing more than a piece of quartz, cut and mounted in such a way as to cause a mechanical resonance to occur at the desired frequency. The equivalent circuit of the crystal is given in Figure 4. The equivalent impedance of the crystal can be found with elementary network analysis and is given by

$$Z_p = \frac{\frac{2\pi f L_1 - 1/2\pi f C_1}{2\pi f C_0} - \frac{jR_1}{2\pi f C_0}}{R_1 + j\left(2\pi f L_1 - \frac{1}{2\pi f C_1} - \frac{1}{2\pi f C_0}\right)} \quad (5)$$

There are two resonant frequencies, called series and parallel, for this impedance and they are given by [5]

$$f_{series} = \frac{1}{2\pi} \sqrt{\frac{1}{L_1 C_1}} \quad (6)$$

$$f_1 = \frac{1}{2\pi} \sqrt{\frac{1}{L_1 C_1} \left(1 + \frac{C_1}{C_0}\right)} \quad (7)$$

Since  $C_1$  is typically three orders of magnitude smaller than  $C_0$ , the two resonance frequencies differ by a very small amount, on the order of 200 ppm. Typical values for the model components are  $R_1=20 \Omega$ ,  $L_1=11 \text{ mH}$ ,  $C_0=6 \text{ pF}$ , and  $C_1=2.6 \text{ fF}$ , yielding a Q of around  $10^5$ .

### 2.3.2 Crystal Oscillator Circuits

A great variety of circuit configurations exist for the crystal oscillator, but the principle is the same as for the LC oscillator, i.e., the feedback network must provide a phase shift of 360 degrees at the frequency of operation. Figure 5 shows a Pierce oscillator using a JFET as the active device. Aside from parasitic reactances, the crystal is the only frequency-determining component in the circuit, which leads to high stability. Other possible circuit topologies include Colpitts, Hartley, Clapp, Butler, grounded base, and impedance-inverting Pierce. The following table provides a comparison among some of the various types [5]:

Oscillator Type	Recommended Frequency Range	Relative Stability	Immunity to Stray/Parasitic Circuit Effects
Pierce	0.1 to 20 MHz	High	Very good
Colpitts	1 to 20 MHz	Moderate	Good
Clapp	2 to 20 MHz	Moderate to High	Good
Impedance-inverting Pierce	20 to 75 MHz	High	Fair
Grounded-base	20 to 150 MHz	Moderate	Poor

### 2.3.3 Stability of Crystal Oscillators

Since the phase noise of an oscillator is inversely proportional to resonator Q, and the Q of most crystals is very high, the phase noise of a crystal oscillator will be very low. Through the proper choice of circuit topology, the most significant source of long-term instability will be the crystal itself. There are two factors contributing to this: aging and

temperature effects.

#### *Aging*

Aging in a crystal is caused by shifts in the mechanical mounting arrangement or by the absorption of atmospheric gasses over a period of time. Typical values for aging range from 1-5 parts in  $10^8$  per week for standard crystals to 1 part in  $10^{11}$  for precision vacuum-sealed types. Also, the rate of aging is greatest just after a crystal is manufactured and slows down after 1-2 months.

#### *Temperature Effects*

The variation with temperature of a crystal's frequency is primarily a function of the way in which it was cut from the mother lode. There are two cut types in common use for frequencies above 1 MHz: AT and SC. For any given cut type, the exact angle of cut also affects its temperature coefficient. Figure 6 illustrates the temperature-frequency characteristics of an AT-cut crystal for varying cut angles, normalized to 25°C. The best choice of cut angle would be one which gives the smallest variation over the intended range of operating temperatures. Over a -40 to +85°C range, an AT-cut crystal will show a frequency variation of around  $\pm 20$  ppm. Typical temperature coefficients near the median temperature are 4 ppb/°C and 20 ppb/°C for SC and AT types, respectively. There is also a certain amount of hysteresis in the  $f$ - $T$  curves, i.e., the frequency at a particular temperature depends whether the temperature is rising or falling [6]. This hysteresis is more pronounced at lower temperatures and can reach 0.5 ppm.

The easiest and oldest method of stabilizing a crystal oscillator is to accurately

control the temperature of the crystal. This is usually done with a crystal oven, a device wherein the crystal is mounted along with an electric heating element and a temperature measuring circuit. Stability with this type of arrangement varies from 1 ppm to 0.1 ppb, the latter figure for an oven which houses the entire oscillator circuit. The disadvantages of this method include bulky size, long warmup times, and maximum ambient temperature limited by oven temperature; furthermore, the crystal will still continue to age at the normal rate.

#### **2.3.4 Temperature-Compensated Crystal Oscillators**

It is because of the variation of frequency with temperature in crystals and the unsuitability of crystal ovens for many applications that a means of adjusting the operating frequency of a crystal oscillator as the crystal varies in temperature was sought. The basic technique involves measuring the temperature of the crystal and then *pulling* the frequency of the oscillator with a reactive component such as a varactor diode to counteract the effects of temperature.

#### *Analog temperature compensation*

In analog compensation methods, one or more temperature-sensitive components directly alters the operating frequency of the oscillator. Consider the network in Figure 7, designed for compensating AT-cut crystals [5]. The bias on the varactor diode can be shown to be



$$\frac{V_o}{V_i} = \frac{RT_3(R_1 + RT_1)(R_2 + RT_2)}{(R_1 + RT_1)(R_2 + RT_2)(R_3 + RT_3) + R_2 RT_2 [(R_1 + RT_1) + (R_3 + RT_3)]} \quad (8)$$

Computer-aided-design techniques may be used to optimize the component values of such a network [7], and an improvement in frequency stability by a factor of four can be had. Many other configurations with a greater or lesser number of components are possible and will give performance commensurate with the complexity level. The disadvantages of analog compensation networks stem from the cost of the thermistors, limits in the performance of circuits of reasonable complexity, and the difficulty in optimizing the design.

#### Digital temperature compensation

Digital compensation techniques are becoming quite popular, reflecting advancements in digital devices which continually bring system costs down. A correction voltage is applied to a varactor diode as in the analog technique. This error voltage is obtained in the following manner. First, a thermistor measures the crystal temperature and produces a proportional voltage. This voltage is then applied to an analog-to-digital converter, which produces a binary word used to address a read-only memory. At this address in the read-only memory is the digital word corresponding to the compensation voltage necessary to keep the oscillator on frequency. Next, this word is applied to a digital-to-analog converter which produces the final error voltage to bias the varactor. The following equation may be used to design this type of system [5]:

$$\Delta f = \frac{S(T)}{2} + \frac{F}{2} \left( \frac{1}{2^b} \right) \quad (9)$$

where

- $\Delta f$  is the maximum frequency error allowed in ppm
- $S$  is the maximum temperature coefficient of the crystal in ppm/°C
- $T$  is the total temperature range of operation in °C
- $n$  is the number of words in the memory
- $F$  is the maximum peak-to-peak frequency excursion of the crystal in ppm
- $b$  is the number of bits in each correction word

By adding a microprocessor and real-time clock to the system, even effects such as thermal hysteresis and aging [8] can be neutralized. The following table compares the relative stability of the various crystal oscillators:

Oscillator type	Stability w/temp. (ppm -40 to +70)	Aging (ppm/yr)	Cost to OEM, \$
Average Crystal	25	5	2
Good Crystal	10	0.5	5-10
Average Analog TCXO	2	0.5	10
Good Analog TCXO	0.5	0.5	20
Average Digital TCXO	0.5	0.5	20
Good Digital TCXO	0.01	<0.5	50+

#### 2.4 Other Oscillator Types

So far, the use of LC networks and crystals to determine the oscillation frequency has been discussed. These approaches provide good results up to the low VHF range

without frequency multiplication, and up to the low GHz range with. There are other components, such as dielectric resonators, surface acoustic wave devices, and YIG spheres which can take the place of the crystal in an oscillator circuit. Although method of excitation and operating principles vary, these resonators can also be analyzed in terms of Q and temperature stability.

### 3 Frequency Synthesis and Control

As the complexity and frequency stability of communications systems continually increases, a need for more accurate and flexible signal sources emerges. The simple crystal oscillator no longer suffices alone, but merely becomes a reference as part of a larger signal source.

#### 3.1 Phase-Lock Loop Principles

The heart of the modern frequency source is the phase-lock loop (PLL). Figure 8 shows the basic configuration consisting of a phase detector, loop filter, and voltage-controlled oscillator (VCO). The input signal has a phase angle of  $\theta_i$ , and the VCO has a phase angle of  $\theta_o$ . The output of the phase detector is equal to some constant times the difference in phase between the VCO and input signal. This output signal,  $v_d$ , is filtered by the loop filter and applied to the VCO. When the circuit is in "lock," any deviation from the input phase by the VCO will be detected and corrected. The VCO will therefore be perfectly locked in frequency and phase to the incoming signal. The basic loop equations are [9]

$$\frac{\theta_o(s)}{\theta_i(s)} = H(s) = \frac{K_o K_d F(s)}{s + K_o K_d F(s)} \quad (10)$$

$$\frac{\theta_i(s) - \theta_o(s)}{\theta_i(s)} = \frac{\theta_o(s)}{\theta_i(s)} = \frac{s}{s + K_o K_d F(s)} = 1 - H(s) \quad (11)$$

$$V_d(s) = \frac{s K_d F(s) \theta_i(s)}{s + K_o K_d F(s)} = \frac{s \theta_i(s)}{K_o} H(s) \quad (12)$$

The response of the PLL to noise is beyond the scope of this paper, but the interested reader is referred to the excellent papers listed in the bibliography. In essence, proper design of the loop filter response  $F(s)$  will allow the loop to maintain lock in the presence of most types of noise encountered in communications systems.

#### 3.2 PLL Frequency Synthesizers

It is often necessary to generate many frequencies from a single reference source, or allow the frequency to be changed in discrete steps or channels. In these applications, the PLL frequency synthesizer is often used. As shown in Figure 9, several components have been added to the basic PLL. One divider has been placed in front of the reference input, and another between the VCO output and the phase detector input. In order for the loop to be locked,  $f_r/N$  must equal  $f_o/M$ , so that the reference frequency is a rational fraction of the output frequency. The reference frequency is typically derived from a stable crystal oscillator, and the long-term stability of the output is equal to the stability of the reference times  $M/N$ . By incorporating programmable dividers, many different frequencies can be generated in increments of  $f_r/N$ .

### 3.3 Direct Digital Synthesis

One of the shortcomings of the PLL frequency synthesizer is excessive noise as compared with a simple crystal oscillator. An analysis of this noise can be done by taking the reference oscillator noise, multiplying it by the division ratio  $N$ , then bandpass filtering the result by the loop filter response  $F(s)$ . Even with a very narrowband loop filter, the close-in phase noise is going to be many times that of the reference oscillator.

A frequency generation method which is getting increased attention is direct digital synthesis (see figure 10). At the heart of the DDS approach is a sine lookup table which contains the digitized values of a sine wave. The output signal is generated by passing the digital values through a digital-to-analog converter and lowpass filter. Undesired signal components at the output are largely a function of word size and DAC linearity, but the performance achievable with state-of-the-art DDS chips sets is less than -150 dBc/Hz of phase noise and -65 dBc of spurious products.

### 3.4 Automatic Frequency Control

It is often required that the local oscillator in a receiver must accurately track a transmitted carrier for demodulation purposes, or the size and cost restraints of a portable unit do not permit it to have its own stable frequency source. Under these conditions, some method of automatic frequency control (AFC) must be used. In an AFC system, two oscillators, one local and one distant, are synchronized in phase and frequency. Two methods commonly used in AFC are the synchronous oscillator (SO) and the PLL.

In the SO, Figure 11, a free-running LC oscillator comprised of  $Q_1$ ,  $L$ ,  $C_1$ , and  $C_2$

is "coaxed" into following an input signal by adding another transistor ( $Q_2$ ) which injects current into the emitter of  $Q_1$  at a frequency of  $f_{in}$ . The output amplitude at the collector of  $Q_1$  is independent of the amplitude or noise content of the input signal. The free-running frequency, given by equation (4), must be sufficiently close to the input frequency for the circuit to work. Further discussion can be found in the paper by Uzunoglu [10].

Another AFC implementation is shown in Figure 12 [11]. Here, the received carrier is first down-converted to an intermediate frequency which makes it possible to use a standard phase-lock loop to achieve synchronization. The block labelled FDD is a frequency difference detector. In the short loop, Figure 12a, the local oscillator is outside the loop and any variations in its frequency will appear as variations in the frequency of the input signal. In the long loop, Figure 12b, the local oscillator has been incorporated into the loop and its variations can be effectively nulled out. This arrangement lends itself to portable communication products that must maintain accurate control of frequency but cannot afford the size or cost of a precision reference.

## 4 Conclusion

There are many different techniques of frequency generation and control, some of which have been elucidated herein. The mathematics were kept simple in this paper in an effort to limit it to a reasonable size; however, the many references given allow the reader to pursue any topic to the desired level of analysis.

## References

- [1] J. Millman and A. Grabel, *Microelectronics*. New York: McGraw-Hill, 1987.
- [2] F. E. Terman, *Electronic and Radio Engineering*. New York: McGraw-Hill, 1955.
- [3] J. Smith, *Modern Communication Circuits*. New York: McGraw-Hill, 1986.
- [4] W. A. Edson, "Noise in oscillators," *Proc. IRE*, vol. 48, pp. 1454-1466, Aug. 1960.
- [5] M. E. Frerking, *Crystal Oscillator Design and Temperature Compensation*. New York: Van Nostrand, 1978.
- [6] R. L. Filler, "Measurement and analysis of thermal hysteresis in resonators and TCXO's," *Proc. IEEE 42nd Annual Frequency Control Symposium*, pp. 380-388, 1988.
- [7] P. Faerber and A. Victor, "CAD simplifies TCXO design," *RF Design*, Aug. 1988.
- [8] M. R. Miljkovic, et. al., "Aging prediction of quartz crystal units," *Proc. 42nd Annual Frequency Control Symposium*, pp. 404-411, 1988.
- [9] F. M. Gardner, *Phaselock Techniques*. New York: John Wiley & sons, 1979.
- [10] V. Uzunoglu and M. H. White, "Synchronous and the coherent phase-locked synchronous oscillators: new techniques in synchronization and tracking," *IEEE Trans. on Circuits and Systems*, vol. 36, no. 7, pp. 997-1003, July 1989.
- [11] F. M. Gardner, "Characteristics of frequency-tracking loops," *Phase-Locked Loops*. Ed. W. C. Lindsey and C. M. Chie. New York: IEEE Press, pp. 3-25, 1986.

## Bibliography

- E. J. Baghdady, R. N. Lincoln and B. D. Nelin, "Short-term frequency stability: characterization, theory and measurement," *Proc. IEEE*, vol. 53, pp. 704-722, July 1965.
- J. A. Barnes et. al., "Characterization of frequency stability," *IEEE Trans. Instrum. Meas.*, vol. 1M-20, pp. 105-120, May 1971.
- W. A. Edson, "Noise in oscillators," *Proc. IRE*, vol. 48, pp. 1454-1466, Aug. 1960.
- D. B. Leeson, "A simple model of feedback oscillator noise spectrum," *Proc. IEEE*, vol. 54, pp. 329-330, Feb. 1966.
- K. Kurokawa, "Noise in synchronized oscillators," *IEEE Trans. Microwave Tech.*, vol. MTT-16, pp. 234-240, Apr. 1968.
- K. F. Schunemann and K. Behm, "Nonlinear noise theory for synchronized oscillators," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, pp. 452-458, May 1979.
- W. C. Lindsey and C. M. Chie, "Frequency multiplication effects on oscillator instability," *IEEE Trans. Instrum. Meas.*, vol. 1M-27, pp. 26-28, Mar. 1978.
- 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. Instrum. Meas.*, vol. 1M-24, pp. 15-20, Mar. 1975.
- V. F. Kroupa, "Noise properties of PLL systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 2244-2252, Oct. 1982.
- P. Faerber and A. Victor, "CAD simplifies TCXO design," *RF Design*, Aug. 1988.
- S. Fry, "Digital temperature compensation for oscillators," *RF Design*, Feb. 1989.
- R. Matthys, "A high performance VHF crystal oscillator circuit," *RF Design*, Mar. 1987.
- R. Matthys, "Crystal oscillator circuits for VHF," *RF Design*, May/June 1983.
- D. Belcher, "Designing a high-stability VHF oscillator," *RF Design*, Jan./Feb. 1983.
- F. Brown, "Stable LC oscillators," *RF Design*, Mar. 1987.
- F. L. Walls, "The influence of pressure and humidity on the medium and long-term frequency stability of quartz oscillators," *Proc. IEEE 42nd Annual Frequency Control Symposium*, pp. 279-283, 1988.

T. Miyayama, Y. Ikeda, and S. Okano, "A new digitally temperature compensated crystal oscillator for a mobile telephone system," *Proc. IEEE 42nd Annual Frequency Control Symposium*, pp. 327-333, 1988.

B. Parzen, "Clarification and a generalized restatement of leeson's oscillator noise model," *Proc. IEEE 42nd Annual Frequency Control Symposium*, pp. 348-351, 1988.

R. L. Filler, "Measurement and analysis of thermal hysteresis in resonators and TCXO's," *Proc. IEEE 42nd Annual Frequency Control Symposium*, pp. 380-388, 1988.

M. R. Miljkovic, et. al., "Aging prediction of quartz crystal units," *Proc. 42nd Annual Frequency Control Symposium*, pp. 404-411, 1988.

A. Blanchard, *Phase-Locked Loops: Application to Coherent Receiver Design*. New York: John Wiley & Sons, 1976.

F. M. Gardner, *Phaselock Techniques*. New York: John Wiley & sons, 1979.

C. M. Chie and W. C. Lindsey, "Phase-locked loops: applications, performance measurement, and summary of analytical results," *Phase-Locked Loops*. Ed. W. C. Lindsey and C. M. Chie. New York: IEEE Press, pp. 3-25, 1986.

V. F. Kroupa, "Noise properties of PLL systems," *IEEE Trans. Commun.*, vol. COM-30, pp. 2244-2252, Oct. 1982.

J. E. Ohlson, "Statistics of the first-order phase-locked loop with fluctuating signal amplitude," *IEEE Trans. Commun.*, vol. COM-26, pp. 1472-1474, Oct. 1978.

R. W. D. Booth, "Carrier phase and bit sync regeneration for the coherent demodulation of MSK," *National Telecommunications Conference Record*, pp. 6.1.1-6.1.5, vol. 1, Dec. 1978.

F. D. Natali, "AFC tracking algorithms," *IEEE Trans. Commun.*, vol. COM-32, pp. 935-947, Aug. 1984.

A. S. Sedra and K. C. Smith, *Microelectronic Circuits*. New York: Holt, Rinehart and Winston, 1987.

J. Millman and A. Grabel, *Microelectronics*. New York: McGraw-Hill, 1987.

F. E. Terman, *Electronic and Radio Engineering*. New York: McGraw-Hill, 1955.

U. L. Rohde and T. T. N. Bucher, *Communications Receivers: Principles and Design*. New York: McGraw-Hill, 1988.

J. Smith, *Modern Communication Circuits*. New York: McGraw-Hill, 1986.

V. Uzunoglu and M. H. White, "Synchronous and the coherent phase-locked synchronous oscillators: new techniques in synchronization and tracking," *IEEE Trans. on Circuits and Systems*, vol. 36, no. 7, pp. 997-1003, July 1989.

A. Benjaminson and S. C. Stallings, "A microcomputer-compensated crystal oscillator using a dual-mode resonator," *Proc. IEEE 43rd Annual Symposium on Frequency Control*, pp. 20-26, 1989.

V. Manassewitsch, *Frequency Synthesizers*. New York: John Wiley, 1987.

*Phase-Locked Loop Systems Data Book*. Motorola, 1973

M. E. Frerking, *Crystal Oscillator Design and Temperature Compensation*. New York: Van Nostrand, 1978.

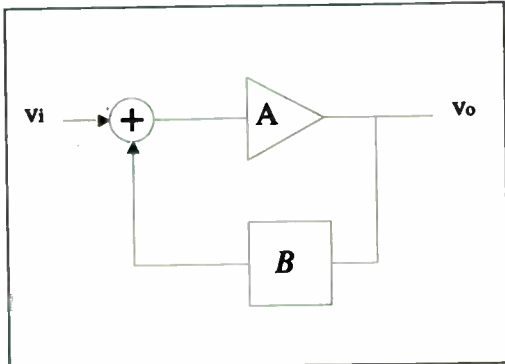


Fig. 1. Basic feedback amplifier.

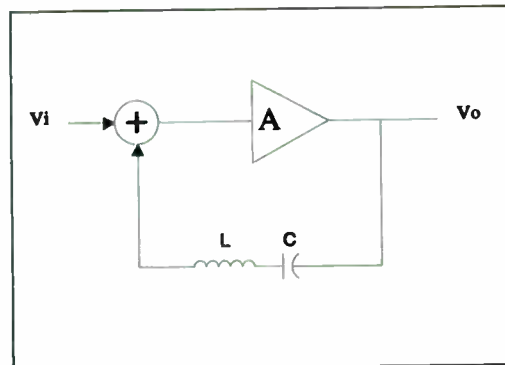


Fig. 2. LC oscillator.

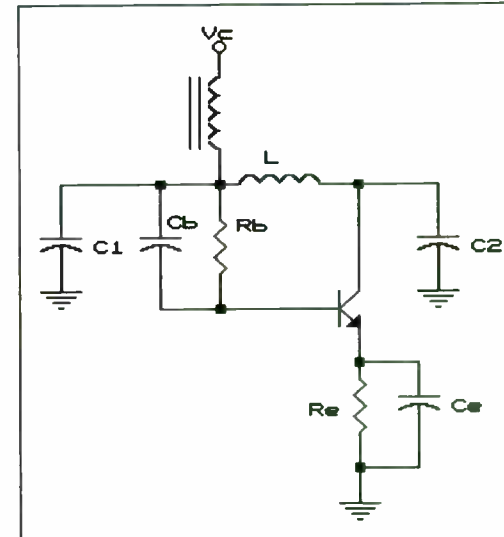


Fig. 3. Pierce BJT Oscillator.

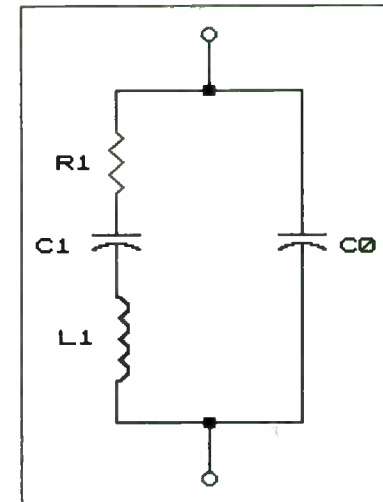


Fig. 4. Quartz crystal equivalent circuit.

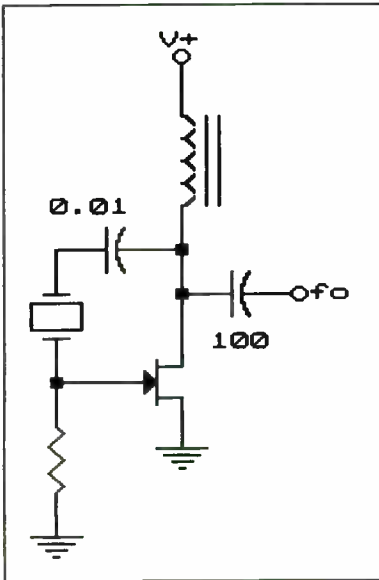


Fig. 5. Pierce JFET crystal oscillator.

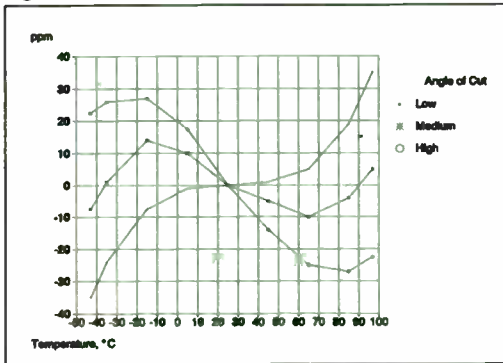


Fig. 6. Temperature-frequency characteristics of AT crystal.

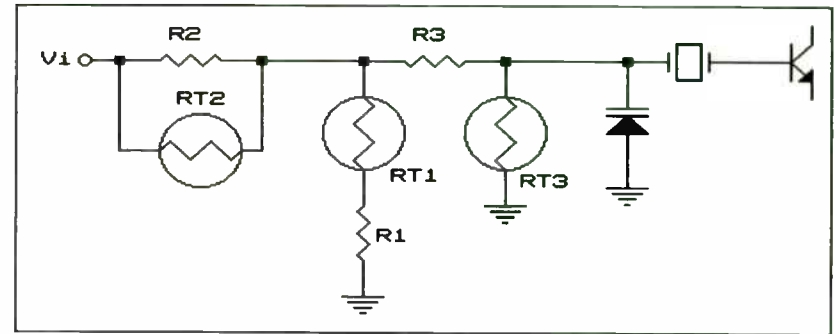


Fig. 7. Analog temperature compensation network [5].

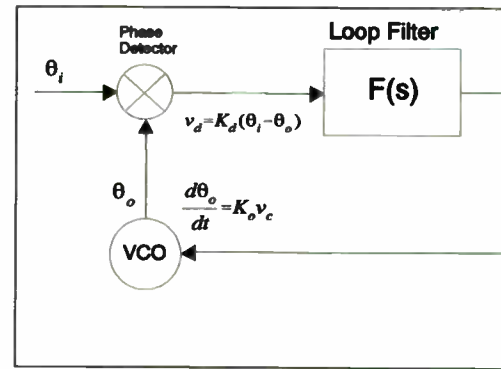


Fig. 8. Diagram of basic PLL [9].

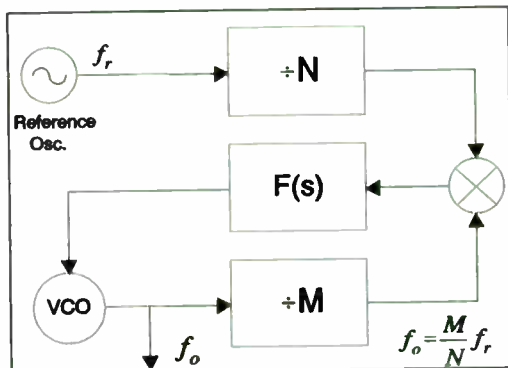


Fig. 9. PLL Frequency Synthesizer [9].

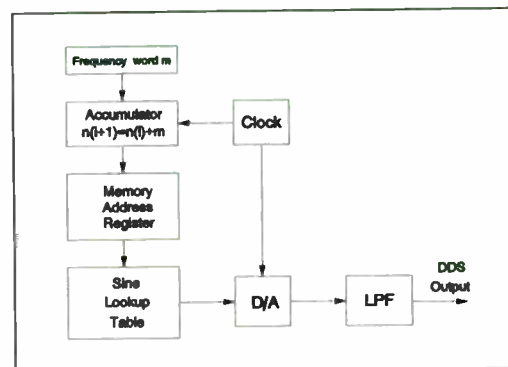


Fig. 10. Direct Digital Synthesizer.

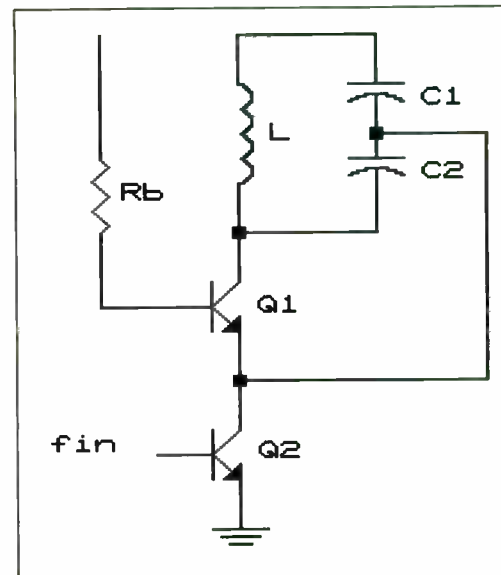


Fig. 11. Synchronous oscillator.

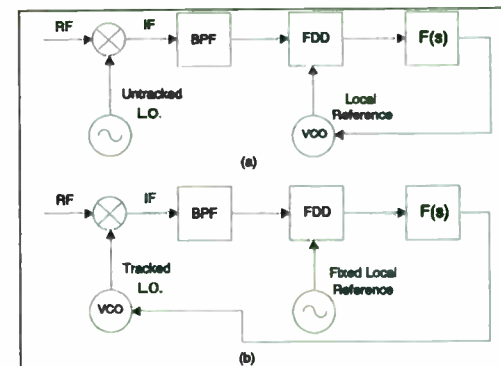


Fig. 12. Short loop (a) and long loop (b) PLL AFC diagrams [11].



## SELF-ADAPTIVE VCO

Wojciech Klimkiewicz, Theodore Grosch

SSDD Research Corporation

315 South Allen Street, Suite 421

State College, PA 16801

### ABSTRACT

The growing popularity of digital signal processing solutions with obvious advantages over analog solutions like temperature stability and algorithm flexibility are invading the RF and microwave market in the form of synthesizers, digitally compensated oscillators, DDS excitors, e.t.c. The dramatically falling prices of high speed digital devices will soon make the price of a sophisticated Digital Signal Processor (DSP) competitive with a single RF circuit. In this paper, the authors would like to present an FM VCO design that utilizes a DSP to achieve high quality modulation over changing environmental conditions and aging. The research on this matter was started with practical experiments on VCOs with digital feedback. The effort to explain why it worked so well has resulted in this paper.

### 1. INTRODUCTION

The emergence of fast inexpensive DSPs, A/D, and D/A converters creates an opportunity to take a different look at classical feedback control systems. The fast processing capability of these devices makes it possible to build iterative feedback control systems where the control strategy  $v_n(t)$  at iteration  $n$  is based on the previous observations of the controlled process  $x_k(t)$  and the previous controls  $v_k(t)$ ,  $k < n$ . A large class of control systems can be found where such algorithms will uniformly converge. If these algorithms do not depend on process parameters, then a robust and adaptive control system can be built.

The authors of this paper report on one example of such a system with a Voltage Controlled Oscillator (VCO) that has a nonlinear voltage tuning characteristic. A novel VCO/PC-DSP test system has been developed as a response to difficulties in measuring and reducing the FM nonlinearity error of RF VCOs under dynamic conditions. The test system operation is based on an interferometer FM-CW concept. When a mixer is excited with a frequency modulated VCO output signal and a portion of this output that is time delayed, there should be a constant frequency at the mixer's output for a linear FM modulation. A simple algorithm can be written to calculate how to correct the modulation function to reach the desired signal at the mixer IF output port. The dynamic iteration process uniformly converges to the ideal solution, whose accuracy depends on the system phase noise, time delay, modulation bandwidth, and A/D, D/A converters used.

The system has been used to test an automotive proximity sensor developed by the authors. Independently, the Self-Adaptive VCO/DSP concept can be applied to automated testing of VCO tuning curves, realizing low cost

signal generators, signal simulators, FM DDS exciters, and improving the performance of FM/CW and chirp radars. The concept of Digital Iterative Feedback Control Systems (DIFCS) presented in this paper can be generalized for the class of Repetitive Control Systems (RCS). The reader interested in this matter is referred to [1].

## 2. VCO SPECIFICATIONS AND MODELING

The performance of a VCO can be described in many ways, but in general, only a few specifications are usually used [2]. Each measure of performance impacts the system differently depending on the application. Digital control methods have been reported [3] that improve one or more of these specifications. The main focus of the results reported here was to create a very linear voltage-to-frequency tuning, but the general feedback technique can be applied to other aspects of VCO performance.

Nonlinearities in tuning the frequency of a VCO can create signal drift and distortion in communication systems and measurement errors in automated test equipment. The voltage-to-frequency tuning characteristics can be made stable and linear by circuit design [4] or component design [5]. It is particularly difficult to achieve linearity at low voltages. If these voltages are avoided, the effective bandwidth of the VCO will decrease. Experimental results have shown that iterative feedback is effective in compensating for nonlinearities in tuning, dwell time, tuning speed, and settling time of VCOs where the frequency is tuned periodically. These applications include FM communications and radar.

## 3. SELF-ADAPTIVE VCO WITH ITERATIVE FEEDBACK

Let us consider a Frequency Modulated Voltage Controlled Oscillator with digital feedback as shown in Figure 1, where the DSP controls the VCO by

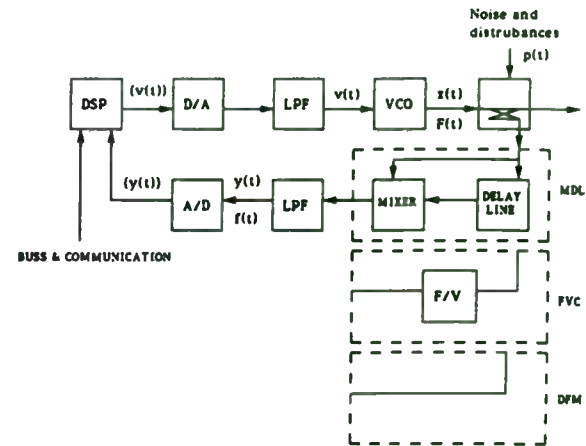


Figure 1. FM VCO with digital feedback: Mixer-Delay Line (MDL) Freq.-Voltage Converter (FVC), and Direct Freq. Measurement DFM)

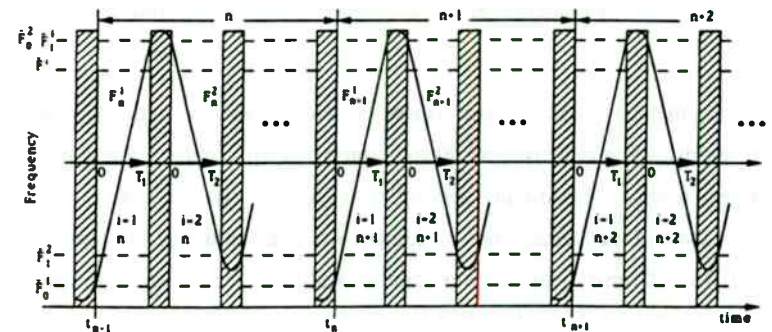


Figure 2. High frequency chirp signals

producing the voltage function  $v(t)$ ,  $0 \leq t \leq T$ , according to a coded command and correcting this function based on the measurement results of the mixer's output  $y(t)$ ,  $0 \leq t \leq T$ . In Figure 1, the dotted line encloses a mixer and delay line, MDL subsystem, a frequency-to-voltage converter, FVC subsystem, to stress that these devices are necessary to convert the high frequency output of the VCO to the region where it can be sampled by the A/D converter. If the VCO frequency is low enough to be sampled and processed directly, DFM, then these devices are not needed and the procedure is a little simpler.

The following notation is assigned to the signals in Figure 1:

- $v(t)$  - voltage control signal of the VCO,
- $x(t)$  - output signal of the VCO,
- $F(t)$  - instantaneous frequency of  $x(t)$ ,
- $y(t)$  - measured output signal of the mixer,
- $f(t)$  - instantaneous frequency of  $y(t)$ .

Also, terms enclosed in brackets,  $\{v(t)\}$ ,  $\{y(t)\}$ , and  $\{f(t)\}$ , will denote digital signals which correspond to  $v(t)$ ,  $y(t)$ , and  $f(t)$ ,

$$\{v(t)\} = \{v(k \cdot \Delta t)\}, \quad k = 0, 1, 2, \dots, N \quad (1)$$

where  $N$  is the number of samples,  $\Delta t$  is the time between consecutive samples, and  $T$ ,

$$0 \leq t \leq T, \quad \Delta t \cdot N = T, \quad (2)$$

is the A/D and D/A total sampling time.

Let us assume that the task of the DSP is to produce a repetitive, series of arbitrary wideband chirp signals  $x_n(t)$ , whose frequency  $F_n(t)$  is shown in Figure 2. These types of signals are commonly used in pulse and FM/CW

radars to match a dispersive delay-line filter or create a linear frequency sweep. Also assume that the processor does not need to create a perfect signal at each event, but instead, the series of events

$$F_n(t), \quad n = 0, 1, 2, \dots \quad (3)$$

with the same task (the same indicator  $i$ , Figure 2), has to converge to some desired solution with increasing  $n$ ,

$$E_n = \|F_n - G\| \leq C \|F_{n-1} - G\|. \quad (4)$$

The term inside double vertical bars  $\|\cdot\|$  is the norm of the function,  $C$  is a constant describing how fast the process (iteration error  $E_n$ ) is converging to the ideal solution  $G$  where  $0 < C < 1$ . In fact the parameter  $CS$ , the conversion speed,

$$CS = 100 \|1 - C\| \% \quad (5)$$

tells the percentage of improvement that will occur from the  $(n-1)^{th}$  to  $n^{th}$  iteration. It should be noted that the indicator "i" from Figure 2 is omitted since the considerations for the other series of events are the same. Also, note that the lower case "t" in the equations denote a specific time of the event. In other words, we have shifted all the event in time (functions describing those events) to 0 in order to simplify the notation as shown in Figure 3.

In reality, when only a low frequency  $f(t)$  signal at the output of a mixer or frequency-to-voltage converter can be measured, Figure 1, task (4) may not be achievable. The VCO control is based on the low frequency iteration  $f_n(t)$  and the goal is to make  $f_n(t)$  converge to an ideal signal  $g(t)$ ,

$$e_n = \|f_n - g\| < c \|f_{n-1} - g\|, \quad (6)$$

where  $e_n$ , is the low frequency error, and  $c$  is the convergence coefficient. The

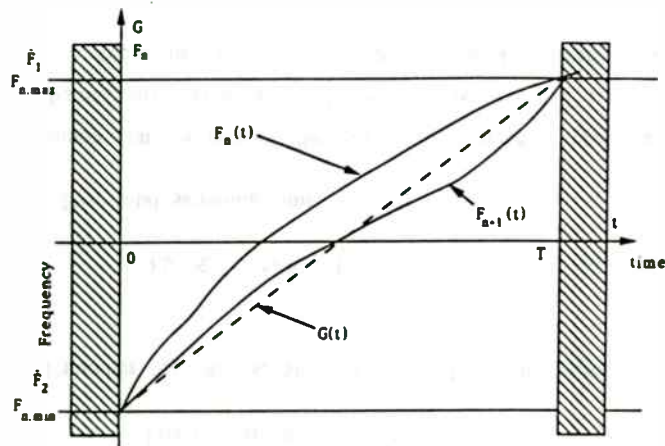


Figure 3. Series of frequency sweep signals  
i-constant, n=1,2,....

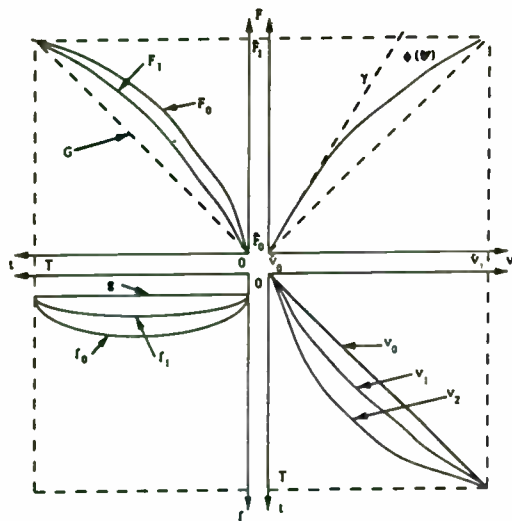


Figure 4. Iterative feedback for VCO with mixer-delay line

term  $g$  is related to the original desired signal  $G$  by the mixer and delay line characteristic

$$f_n(t) = \text{MIX}(F_n(t), h), \quad g(t) = \text{MIX}(G(t), h) \quad (7)$$

For simplicity, it is assumed that the mixer and delay line are ideal, so (7) is in the form:

$$f(t) = h F(t) \quad (8)$$

or

$$F(t) = \tilde{F}_0 + \frac{1}{h} \int_0^t f(\tau) d\tau, \quad F(T) = \tilde{F}_1 \quad (9)$$

where  $h$  [sec] is the delay time, and  $\tilde{F}_0$  and  $\tilde{F}_1$  are known, but cannot be measured or controlled within the feedback MDL structure of Figure 1. This way, by replacing (4) with (6), we can iteratively control only the shape of the chirp signals  $F(t)$ ,  $0 \leq t \leq T$ , and guarantee that

$$F_n(t) \Rightarrow G'(t) = \frac{g(t)}{h} \quad (10)$$

The type of control described above is called repetitive control and the feedback

$$v_n = L[g, v_{n-1}, v_{n-2}, \dots, y_{n-1}, y_{n-2}, \dots] \quad (11)$$

which describes how to produce the control  $v_n$  at iteration  $n$  based on previous controls and measurements  $v_k$  and  $y_k$ ,  $k < n$ , and (6) or (4) holds is called iterative feedback. A more formal definition of these terms can be found in [1], and for now we assume that they are obvious.

Generally, the problems associated with finding the appropriate algorithm  $L$  result from the nonlinear characteristics of the devices used (VCO, Mixer,

delay line, A/D, D/A, etc.), their unknown dynamics, noise present in the system, and changing environmental conditions ( $p(t)$  in Figure 1). We will address these issues one at a time.

First, consider the nonlinear frequency tuning characteristic of the VCO,

$$F(t) = \phi[p, v(t)] = \phi[v(t)], \quad (12)$$

where  $p(t)$  is the unknown disturbance parameter, constant or varying slowly in time, and  $\phi$  is the nonlinear tuning characteristic of the VCO that is fixed during the iteration  $p_n(t) = \text{constant}$ ,  $0 \leq t \leq T$ , Figure 4, continuously differentiable on  $[0, T]$  and

$$\phi(\tilde{v}_0) = \tilde{F}_0, \quad \phi(\tilde{v}_1) = \tilde{F}_1, \quad 0 < \phi' < \gamma, \quad (13)$$

where  $\gamma$  [Hz/V] is the maximum slope coefficient.

The high frequency signals  $F_n(t)$  resulting from the VCO's voltage control signals  $v_n(t)$ ,  $0 \leq t \leq T$ ,  $v_n(0) = \tilde{v}_0$ ,  $v_n(T) = \tilde{v}_1$ ,

$$F_n(t) = \phi[v_n(t)], \quad 0 \leq t \leq T \quad (14)$$

are illustrated in Figure 4.

### 31. DIRECT HIGH FREQUENCY MEASUREMENT (DFM)

In the case when the VCO output can be sampled directly, we have  $y(t) = x(t)$  and  $f(t) = F(t)$ . Consider the following iterative feedback algorithm

$$v_n(t) = L[G, v_{n-1}(t), F_{n-1}(t)] = v_{n-1}(t) - \frac{\alpha[F_{n-1}(t) - G(t)]}{S_V} \quad (15)$$

where  $B = \tilde{F}_1 - \tilde{F}_0$  is the fixed bandwidth,  $S_V$  is the average slope of the VCO characteristic  $S_V = B/V_R$ , and control voltage range  $V_R = \tilde{v}_1 - \tilde{v}_0$ . The series of functions  $F_n(t)$ ,  $0 \leq t \leq T$ , are converging uniformly to a given function  $G(t)$ ,

$$\begin{aligned} E_n &= \|F_n(t) - G\| = \|\phi[v_n(t)] - G\| = \|\phi[v_{n-1} - \frac{\alpha[F_{n-1} - G]}{S_V}] - G\| \\ &= \|1 - \frac{\alpha \phi'(\xi)}{S_V}\| \|F_{n-1} - G\| \leq C \|F_{n-1} - G\|, \end{aligned} \quad (16)$$

for an  $\alpha$ ,  $0 < \alpha < S_V/\gamma$ , such that  $0 < C < 1$ , and

$$0 < \|1 - \frac{\alpha \phi'(\xi)}{S_V}\| \leq C. \quad (17)$$

The iterative algorithm for direct high frequency measurement, (15) is simple. To calculate the VCO control signal  $v_n(t)$ , the following steps are taken:

- I. Sample the signal  $x_{n-1}(t)$ ,  $0 \leq t \leq T$ ,  $\{x_{n-1}(t)\}$ .
- II. Estimate the frequency characteristics, by detecting the extremums, performing a nonlinear polynomial regression and calculating,  $\{F_{n-1}(t)\}$ .
- III. Calculate the correction function,  $\{F_{n-1}(t)\} - \{G(t)\}$ , and multiply it by a small parameter,  $\alpha/S_V < 1/\gamma$
- IV. Subtract the result from the last control,  $\{v_{n-1}(t)\}$ , to get  $\{v_n(t)\}$ .

Step II above is probably the most difficult operation to perform because of the difficulty in detecting the extremums of the signal  $\{x_{n-1}(t)\}$ . One can get particularly bad results if the procedure is missing some of them and the program is not signalling these events.

### 3.2 MIXER-DELAY LINE SYSTEM (MDL)

In this type of system, only the mixer's output signal  $y(t)$  can be sampled. Because the feedback, Figure 1, is working as a differentiator due to the mixer-delay line (8), the algorithm must insure that the functions in the feedback loop

are much smoother. Let us assume that the VCO frequency characteristic is not only continuously differentiable (13), but it also has a continuous second derivative, and

$$|\phi''| < \delta \text{ [Hz/v}^2\text{]}, \quad (18)$$

together with

$$T \delta < \gamma. \quad (19)$$

Consider the following iterative feedback algorithm:

$$v_n(t) = v_{n-1}(t) - \alpha \Delta v_{n-1}(t), \quad (20)$$

where

$$\Delta v_{n-1} = \frac{1}{S_V \cdot h} \int_0^1 [f_{n-1}(\tau) - g(\tau)] d\tau, \quad (21)$$

and the average values  $\bar{f}_n$  and  $\bar{g}$  of the functions  $f_n$  and  $g$  are equal,

$$\bar{f}_n = \frac{1}{T} \int_0^T f_n(\tau) d\tau = \frac{1}{T} \int_0^T g(\tau) d\tau = \bar{g} = h \cdot B/T = S_F \cdot h, \quad (22)$$

where  $B$  is the bandwidth  $B \doteq \bar{F}_1 - \bar{F}_0$ , and  $S_F$  is the time frequency slope  $S_F = B/T$ . The estimation of the low frequency error  $e_n$  at iteration  $n$  is as follows:

$$\begin{aligned} f_n - g &= h \frac{d}{dt} \phi[v(t)] - g \\ &= \left(1 - \frac{\alpha \phi'(\xi)}{S_V}\right) (f_{n-1} - g) - \frac{\alpha}{S_V} \left[\frac{d}{dt} \phi'(\xi)\right] \int_0^1 [f_{n-1}(\tau) - g(\tau)] d\tau, \end{aligned}$$

and

$$e_n = \|f_n - g\| \leq \|(1 - \alpha \phi'/S_V) + \frac{\alpha \delta \beta V_R}{S_V}\| \|f_{n-1} - g\|,$$

where  $0 \leq \left| \frac{d}{dt} (v_n(t)) \right| \leq \beta \frac{V_R}{T}$ , and  $V_R$  is the voltage range  $V_R = \bar{v}_1 - \bar{v}_0$ . Hence, there exists an  $\alpha$ ,  $0 < \alpha < S_V/\gamma$ , small enough so that

$$\|f_n - g\| \leq c \|f_{n-1} - g\|, \quad 0 < c < 1 \quad (23)$$

and  $f_n$  is converging uniformly to  $g$ . At the same time, (8) and (9) implies

$$F_n - G = \frac{1}{h} \int_0^T [f_n(\tau) - g(\tau)] d\tau \quad (24)$$

and therefore

$$\|F_n - G\| \leq \frac{T}{h} \|f_n - g\|, \quad (25)$$

which guarantees the convergence of  $F_n$  to  $G$ , with  $C \leq c$ . The iterative process  $v_n$ ,  $f_n$ , and  $F_n$  is illustrated in Figure 4, for

$$G(t) = \bar{F}_0 + S_F t, \quad g(t) = h \cdot S_F = \text{constant} \quad (26)$$

where the iterative feedback works as a dynamic linearizer and is trying to keep the modulation process linear, (26), even when the voltage tuning characteristic of the VCO is changing slowly with time. However, the process cannot prevent a drift of the DC values of  $G(t)$ ,  $\bar{F}_0$ , and  $\bar{F}_1$ , since the only task realized by the algorithm is to keep  $g(t)$  constant. The algorithm is similar to the one in section 3.1 with the exception that the calculations are now performed on  $f_n(t)$  instead of on  $F_n(t)$ , and step III is done according to the formulas (20), (21), and (22).

### 3.3 SYSTEM DYNAMICS

The iterative feedback for a nonlinear VCO tuning characteristic described above is often not sufficient since every element in the loop, Figure 1, has its

own dynamic response with a specific response times. Only when this response time is very small in comparison to T can the formulas from sections 3.1 and 3.2 be used to build a practical digital iterative feedback. In the case where the response time is significant, the transition time between consecutive events (the shaded areas in Figure 2 and 3) can be used to quiet the system response. This is usually done by insuring that the control voltage at 0 and T is smooth enough,

$$v_n^{(k)}(0+) = v_n^{(k)}(0-), \quad v_n^{(k)}(T+) = v_n^{(k)}(T-) \quad (27)$$

for a high enough k<sup>th</sup> order derivative. Generally, however, some other means are necessary. The reader is referred to [1] for a unified approach. We will only consider here the simple DFM case when the feedback from Figure 1 contains only one major time constant. The mixer-delay line loop MDL, Figure 1, is similar.

Suppose that the time constant of the VCO is h [sec], and the dynamic equations of the system are:

$$F(t) = f(t) = \phi[u(t)], \quad (28)$$

$$\dot{v} = D(u) = h \frac{du}{dt} + u, \quad (29)$$

$$u(0) = v(0) = \tilde{v}_0 \quad (30)$$

where D is a differential operator. The general solution of (29) and (30) is in the form

$$u(t) = \tilde{v}_0 + K(v)(t) = \tilde{v}_0 + (k*v)(t) = \tilde{v}_0 + \int_0^t k(t-\tau)v(\tau)d\tau, \quad (31)$$

where

$$k(t) = (\exp(-t/h))/h. \quad (32)$$

It is easy to show that the operations K and D are inverse,

$$K D (u) = u, \text{ for any } u. \quad (33)$$

Now, consider the following iterative feedback

$$v_n = v_{n-1} - \alpha D (F_{n-1} - G)/S_V, \quad (34)$$

and the iteration error is found by

$$\begin{aligned} F_n - G &= \phi [\tilde{v}_0 + k*v_n] - G = \phi [\tilde{v}_0 + k*(v_{n-1} - \frac{\alpha}{S_V} D(F_{n-1} - G))] - G \\ &= F_{n-1} - G - \frac{\alpha}{S_V} \phi'(\xi) k*D (F_{n-1} - G) \\ &= [1 - \alpha \phi'(\xi)/S_V] (F_{n-1} - G). \end{aligned}$$

As we see, it is exactly the same formula as (16) for the direct high frequency measurement (DFM) iterative feedback discussed in section 3.1. The digital signal processing algorithm for section 3.1 now includes;

$$\text{III'. Compensate for system dynamics, } h\{F_{n-1}(t)\}' + \{F_{n-1}(t)\}.$$

If the system dynamics are more complicated and contains more time constants, but (31) still holds, then D=K<sup>-1</sup>, if it exists, will still work. Generally, however, we can only hope that in practical applications the iterative error can only be reduced, since the exact model of the system is not known or is too complicated to consider. Some other problems may arise when the system dynamics are not only unknown, but are changing with time and fluctuations of

temperature. Experimentation may be the only way the designer can insure the feedback method will work.

### 3.3 ERRORS AND LIMITATIONS

There are many sources of errors in the digital iterative feedback from Figure 1:

- TN - thermal noise,
- PN - phase noise,
- AM - unwanted amplitude modulation,
- LSB - least significant bit error,
- IFM - instantaneous frequency measurement error,
- CST - convergence speed limitation error,
- ITE - iteration error.

Most of these errors can be significantly reduced by proper component selection and DSP design. From the DSP standpoint, over-sampling is a powerful tool to reduce TN, PM, AM, LSB, and IFM errors. Faster A/D and D/A converters allow some additional digital filtering on the data. Averaging over several chirps could also reduce these errors, but at the cost of a higher convergence speed error. Iteration error (ITE) can be decreased by setting the parameter  $\alpha$  to the highest value possible, but at the risk of causing the process to become unstable. It is also a good practice to choose the right initial iteration  $v_0(t)$  to avoid the initial iterative response.

The biggest challenge comes from the error during the IFM measurement and estimation. It is necessary to insure that the number of cycles during the period  $T$  is significantly larger than the order  $P$  of the regression polynomial used,  $T \bar{F} > qP$ ,  $q = 5$  to  $10$ , where  $\bar{F}$  is the center frequency,  $\bar{F} = (\tilde{F}_0 + \tilde{F}_1)/2$ . It imposes a limitation on the bandwidth,  $B > q P S_F/\bar{F}$ , (DFM), and the A/D and

D/A conversion process,  $N > q P f_s/\bar{F}$  (where  $N$  is the number of samples and  $f_s$  is the sampling frequency). Equations for the mixer-delay line are  $B h > q P$ , (MDL), and  $N > q P f_s/(h S_F)$ , because  $\bar{f} = h S_F$  according to (8).

The quantization error (LSB) caused by the limited number of bits  $M$  used for the control voltage,  $\Delta v/V_R \approx 1/2^M$ , results in a error for high frequency,  $\Delta F/B \approx 1/2^M$ , and the low frequency from the mixer-delay line structure,  $\Delta f/f \approx \lambda N/2^M$ , where the  $\lambda$  depends on how fast the error of the control voltage  $\Delta v(t)$  is changing with time. The experience of the authors have shown that  $\lambda \approx 0.1$ .

### 4. TEST RESULTS

The digital iterative feedback system with a mixer-delay line, Figure 1, MDL, has been tested using a chirp FM/CW radar transceiver with two microstrip antennas (transmit and receive) developed by the authors [6]. A TMS320E15 from Texas Instruments was used as the DSP. A National semiconductor ADC1061, 10 bit, 500 kHz A/D converter was used to sample the mixer output. We selected a AD7870, 12 bit, 100 kHz D/A converter for the VCO tuning voltage interface. We used a basic sampling rate of  $f_s = 40$  kHz with 8 times over-sampling, and tried averaging over 8 to 128 consecutive chirps. The number of samples was  $N = 128$ , so that  $T = 128/f_s = 3.2$  ms. The bandwidth was 100 MHz, and the modulation slope was  $S_F = 36$  GHz/sec. The theoretical nonlinearity error of the iterative feedback was  $\Delta F/B \leq 1/(2^M) = 0.2\%$ . The VCO and mixer were developed and built by the authors and they were incorporated into the radar. A delay line of  $h = 200$  ns was used. In Figure 5 and 6, the computer data is shown before and after the linearization process performed with a test target at 50 ft.



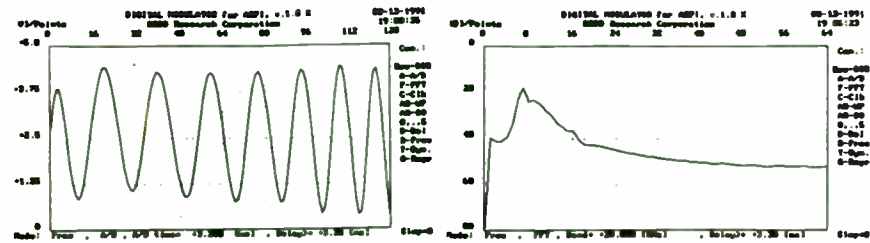


Figure 5. A/D and FFT from 50 ft. target return signal before linearization. Modulation with linear voltage.

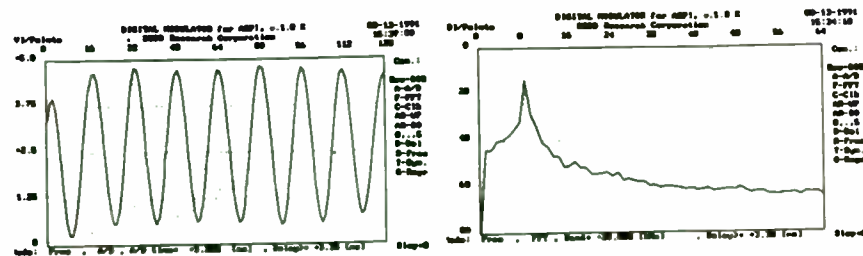


Figure 6. A/D and FFT from 50 ft. target return signal after linearization process of 10 iterations

## 5. CONCLUSIONS

In this article, we presented the basic approach for controlling the instantaneous frequency of a VCO using a novel digital iterative feedback method. This method can also be successfully used to control other aspects of increasingly complex signals (AM, PM, ...). Details of the general iterative feedback theory and other possible applications can be found in [1]. It is hoped that this paper will stimulate further interest in applying the DSP methods in the RF and microwave field in general, and in investigating the digital iterative feedback concept in particular.

## 6. REFERENCES

- [1] Klimkiewicz, W., "Digital Iterative Feedback", ICSPAT International Conference, Cambridge, Massachusetts, November 1992.
- [2] Stickney, R., "System Requirements Set Specifications of Microwave VCOs," *Microwaves & RF*, June 1990.
- [3] Fry, S., "Digital Temperature Compensation for Oscillators," *RF Design*, February 1989.
- [4] Cupidio, L., "A Linear Driftless VCO," *RF Design*, January 1991.
- [5] Niblack, W., "Programmed Diodes Yield Linear Tuning in Broadband VCOs," *Microwaves & RF*, February 1990.
- [6] Grosch, T. O., and Klimkiewicz, W., "Digital FM/CW Radar", ICSPAT International Conference, Cambridge, Massachusetts, November 1992.



Faint, illegible text located below the first diagram.



Faint, illegible text located below the second diagram.

# A CAD AND OPTIMIZATION PROGRAM FOR MICROWAVE OSCILLATORS

A. Heitbrink, B. Roth, A. Beyer  
Duisburg University

DEPARTMENT OF ELECTRICAL ENGINEERING AND SFB 254  
BISMARCKST. 69 , D-4100 DUISBURG 1, FRG

## ABSTRACT

A CAD method will be presented, which is able to calculate all important parameters such as the operating frequency and the power of the harmonics for integrated microwave oscillators. The algorithm is based on several models for microwave transistors in combination with modern numerical methods for linear and nonlinear networks. Results are presented showing the performance capability of this CAD-tool for the design and optimization of such oscillators by means of numerical and experimental examples.

## 1. INTRODUCTION

At the beginning of the integration of microwave circuits the oscillator design was mastered mostly by empirical methods [1]. After a period, in some cases linearized models were used for the active elements to determine the frequency of oscillation and then the power of fundamental oscillation is estimated with the help of empirical derived equations[2,3,4]. But in our days there are some very powerful modeling and simulation methods for the description of GaAs FETs [5]. In combination with modern numerical methods it is possible to implement a full nonlinear analysis of such circuits. A complete concept is now discussed.

## II. FET-MODELING

Several different procedures for the modeling of microwave transistors have already been published. The model chosen here uses the well-known equivalent-circuit shown in Fig. 1.1. For extracting the model parameters the transistor is divided into an inner and outer part. Only the inner part describes the nonlinear characteristics. The calculation of the equivalent circuit components is carried out exclusively by means of measured scattering parameters, no time consuming optimization is needed. By the help of the measurements at different bias conditions this procedure leads to a unique analytical description for the FET [5]. Each element can be described as a two-dimensional function such as

$$f_{ij}(U_1, U_2) = \sum_{k=0}^3 \sum_{l=0}^3 A_{ijkl}(U_1 - U_{1i})^k (U_2 - U_{2i})^l. \quad (1.1)$$

This is a formulation as a two dimensional bicubic - spline - function, which guarantees a continuous function up to the second derivation. For the calculation of the coefficients  $A_{ijkl}$  in equation (1.1) a power algorithm is used. Fig.1.2 and Fig. 1.3 shows examples for the nonlinear elements.

## III. HARMONIC BALANCE METHOD FOR OSCILLATOR DEVELOPMENT

With the numerical description of the FET the nonlinear analysis of microwave and millimeterwave oscillators can be started. A modified Harmonic - Balance approach is used for the calculation of the steady state [6]. If autonomous networks should be used, it must be taken into account, that the frequency is determined by the system itself and may change during the calculation process. Thus, the frequency is an additional state variable. The well known subdivision (see Fig. 1.4) of the network into a nonlinear part to be calculated in time-domain and a linear part to be calculated in frequency-domain can be carried out as shown in Fig. 1.4. At first a linear analysis brings up the frequency of fundamental oscillation and the condition of oscillation is checked. Next, a vector of state variables is defined, consisting

of the port voltages. The first index marks the number of intrinsic (or state) voltages, the second the harmonic, zero means DC.

$$\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_N]^T \quad (1.2)$$

$$\mathbf{U}_k = [U_{k0}, U_{k1}, \dots, U_{kM}], k = 1, 2, \dots, N \quad (1.3)$$

After a transformation of equation (1.2) into time domain, the resulting time-vector leads to the currents at the ports, with the help of the elements of the nonlinear network in Fig. 1.4.

$$\mathbf{U} \rightarrow \mathbf{u}, i = f_{ni}(\mathbf{u}) \quad (1.4)$$

If the currents are transformed back into frequency domain

$$i \rightarrow \mathbf{I}, \mathbf{U} = f_i(\mathbf{I}) \quad (1.5)$$

it is possible to determine a new vector of state voltages by use of the linear network. This can be done with complex AC-analysis. If we turn back to equation (1.4) and close the loop, a fixpoint problem depending on the vector of state-voltages can be defined.

$$\mathbf{U} = f(\mathbf{U}) \quad (1.6)$$

with

$$f(\mathbf{U}) = [f_1(\mathbf{U}), f_2(\mathbf{U}), \dots, f_N(\mathbf{U})]^T \quad (1.7)$$

$$f_k(\mathbf{U}) = [f_{k0}(\mathbf{U}), f_{k1}(\mathbf{U}), \dots, f_{kM}(\mathbf{U})]^T \quad (1.8)$$

There are several methods for solving equation (1.6). A very fast algorithm was presented by Camacho-Penalosa for mixers and upconverter applications [2]. The iteration scheme for the regula-falsi method is like

$$\mathbf{U}^{[n+1]} = f(\mathbf{U}^{[n]}) + \mathbf{K}(\mathbf{U}^{[n]}, \mathbf{U}^{[n-1]}) \quad (1.9)$$

with

$$\mathbf{K}(\mathbf{U}^{[n]}, \mathbf{U}^{[n-1]}) = \frac{[f(\mathbf{U}^{[n]}) - f(\mathbf{U}^{[n-1]})][\mathbf{U}^{[n]} - f(\mathbf{U}^{[n]})]}{[f(\mathbf{U}^{[n]}) - \mathbf{U}^{[n]}] - [f(\mathbf{U}^{[n-1]}) - \mathbf{U}^{[n-1]}}] \quad (1.10)$$

Equation (1.10) shows, that only divisions and multiplications took place. No derivatives must be builded as for example in the standard newton - method. For this algorithm it is necessary to formulate equation (1.6) as a zeropoint problem. This can easily be done by:

$$\mathbf{g}(\mathbf{U}) = \mathbf{U} - f(\mathbf{U}) = \mathbf{0} \quad (1.11)$$

This leads to the well known iteration scheme:

$$\mathbf{U}^{[n+1]} = \mathbf{U}^{[n]} - \mathbf{A}^{-1} \cdot \mathbf{g}(\mathbf{U}^{[n]}) \quad (1.12)$$

The elements  $a_{ij}$  of the Jacobian  $A$  in equation (1.12) are given by

$$a_{ij} = \frac{\partial g_i}{\partial U_j} \quad (1.13)$$

Unfortunately this causes a large numerical expense to calculate this Jacobian in each iteration step. Another disadvantage of this method is, that the convergence of equation (1.12) is only guaranteed in a small surrounding of the solution. This means that it is necessary to have a good initial guess for starting the iterations. Another possibility for solving equation (1.11) is to find the solution by optimization strategies such as gradient or evolution algorithms. This can easily be implemented for searching the fundamental frequency and

the complex amplitudes of equation (1.12). When regarding higher harmonics of the state-voltages, the amount of optimization parameters rises. This can lead to problems in the convergence, because all harmonics are treated as independent free parameters. So local minima of equation (1.12) can occur.

The regula-falsi iteration scheme mentioned above allows an easy way of numerical handling and leads to short CPU-times. On the other hand, this correction term does not include the coupling between the harmonics. But in practice a wide range of microwave circuits have been calculated with this method without any problems of convergence.

The complete computation process is shown in Fig. 1.5.

During the iteration, the real part of  $U_{11}$  is set to a fixed value  $U_s$ . Then the process changes between two iteration loops. In the inner loop the frequency of oscillation is calculated by an one-dimensional newton algorithm under the assumption that the imaginary part of  $U_{11}$  is set to zero in this autonomous system. In the second loop the iteration of the network function is done by the help of the mentioned regula-falsi scheme. If convergence can be obtained in both loops, the up to now fixed voltage  $U_s$  is corrected in the outermost loop by a regula-falsi iteration. The result of all these iterations is the determination of the vector of state voltages  $U$ . All other results of interest (e.g. the output power) can be expressed by elementary network analysis. All this is done typically in 20-40 seconds on a 25 MHz PC. By the help of this approach, another outer loop can be implemented in order to optimize e.g. the linear circuit parameters.

#### IV. RESULTS

As a concrete application the oscillator shown in Fig. 1.6. was calculated by means of the harmonic-balance approach described above. The parameter of the circuit are as followed:

$$\begin{aligned} R_{in} &= 1 \text{ Ohm} & R_s &= 1 \text{ Ohm} & R_{out} &= 50 \text{ Ohms} \\ L_{in} &= 4 \text{ nH} \\ C_{in} &= 200 \text{ fF} & C_s &= 200 \text{ fF} \end{aligned}$$

In this context, the time dependence of the nonlinear elements is of interest. The first diagram (Fig. 1.7) shows the nonlinear behaviour of Cgs of a FET in an oscillator circuit. The elliptic curve marks the values of the element during oscillation. A similar behaviour of the drain current is shown in Fig. 1.8. In both cases, it can be recognized that an extremely nonlinear range is covered.

#### V. CONCLUSION

A complete CAD method for the design of microwave oscillators has been presented. Based on the parameter estimation for the nonlinear circuit elements a nonlinear calculation process can be carried out. Some examples of realized oscillators have been investigated. In future work the noise behaviour of these circuits has to be predicted.

#### REFERENCES

- [1] Roth, B., Sledzik, H., Joseph, M., Beyer, A.: An Advanced Method for the Design and Simulation of Integrated Microwave Oscillators. 14th International Conference on Infrared and Millimeterwaves, Wrzburg, FRG, October 1989, pp. 342-343.
- [2] Kitchen, J.: Octave bandwidth varactor-tuned oscillators. Microwave Journal, May 1987, pp. 347-353.
- [3] Pengelly, R., S.: Microwave Field-Effect Transistor-Theory, Design and Applications (Second Edition), Research Studies Press, Ltd., Letchworth, Herfordshire England, 1986.
- [4] Pucel, R.A., Bera, R., Masse, D.: Experiments on integrated Gallium-Arsenide FET Oscillators at X-band. Electronic Letters, vol. 11..May 15 1975, pp. 219-220.
- [5] Sledzik, H., Wolff, I.: Analytical derivation of a nonlinear FET-equivalent circuit for MMIC' applications. International Journal of Infrared and Millimeter Waves, Vol. 10. No.

3, March 1989, pp. 310-319.

[6] Roth, B., Beyer, A.: Computer Aided Design Techniques for Integrated Microwave Oscillators, Journal on Communications, Vol. XLII, Sept. 1991, pp. 22-34.

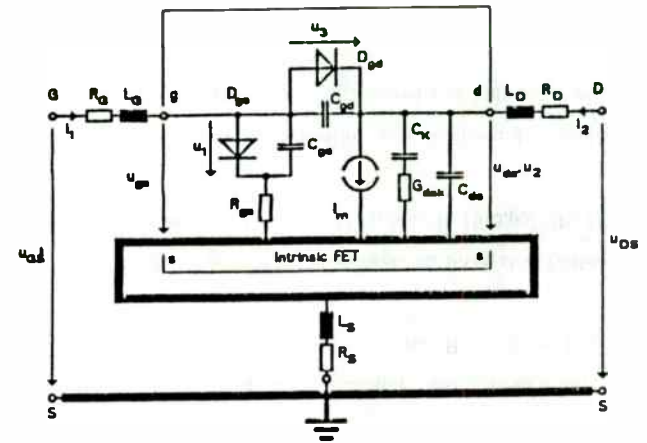


Fig. 1.1: Equivalent circuit of a FET.

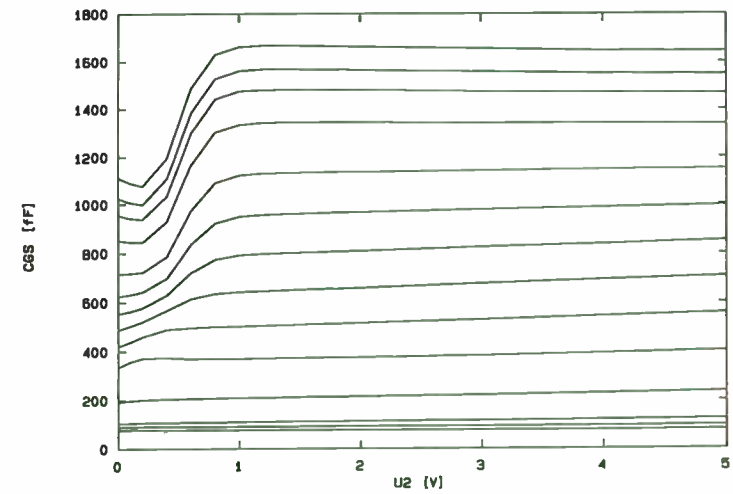


Fig. 1.2: Large signal behaviour of Cgs.

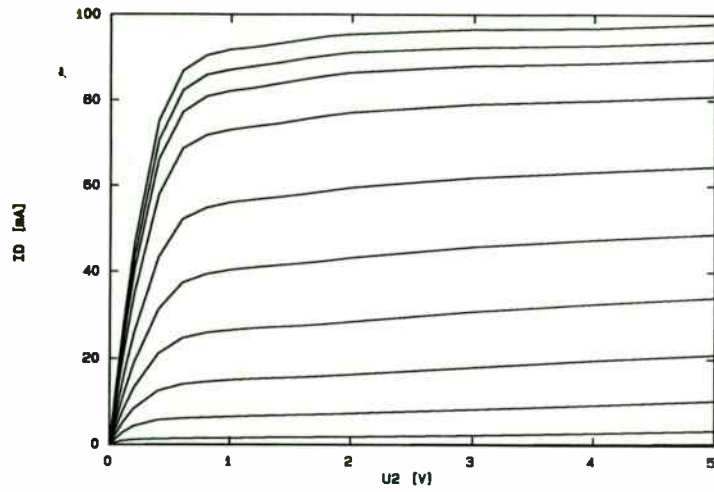


Fig. 1.3: Large signal behaviour of  $I_d$ .

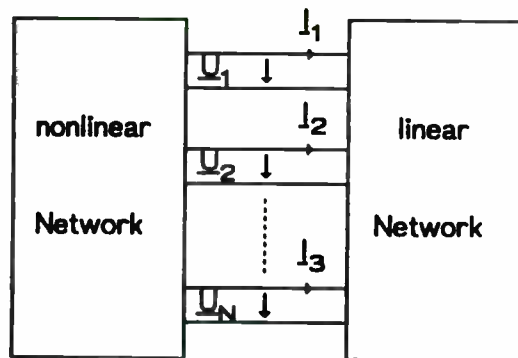


Fig. 1.4: Subdivision of a network into a nonlinear and a linear subnetwork.

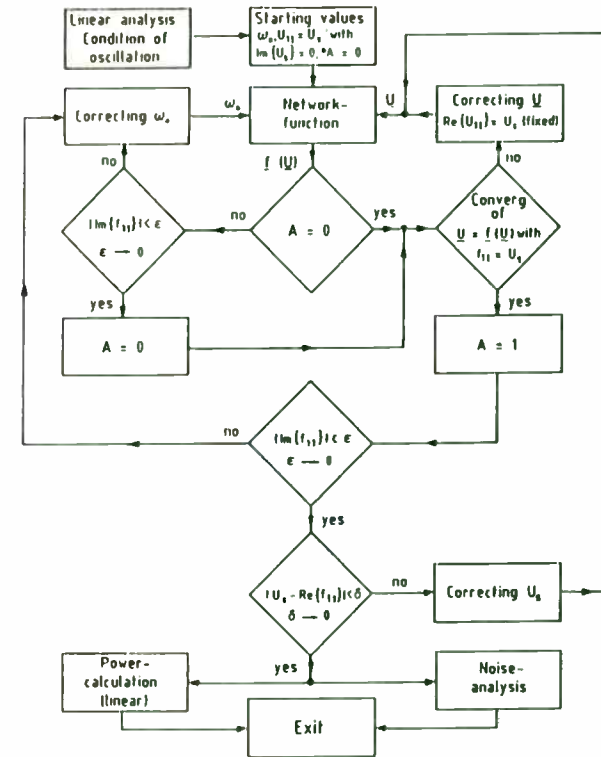


Fig. 1.5: Block diagram of the numerical computation process

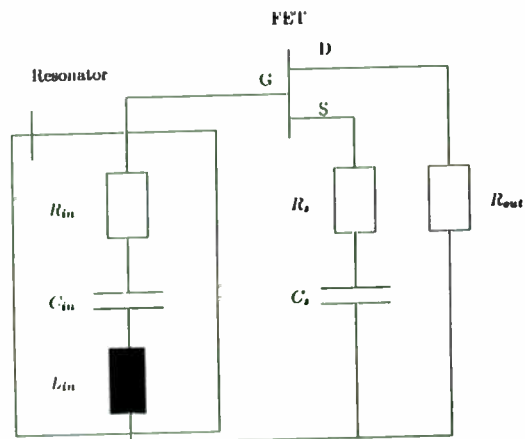


Fig. 1.6: Equivalent circuit of the oscillator

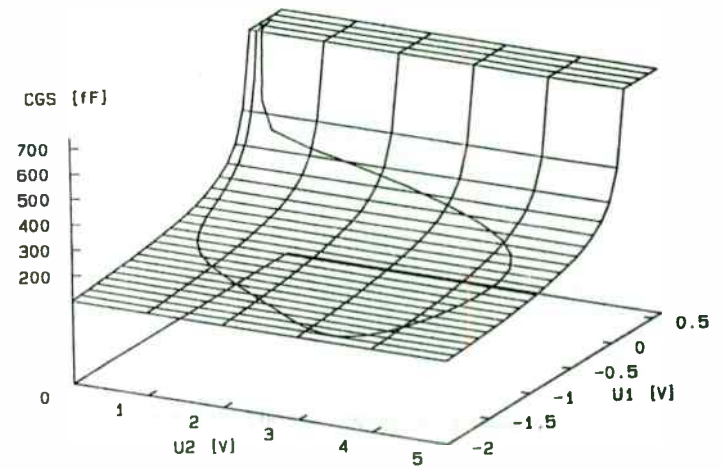


Fig. 1.7: Nonlinear behaviour of  $C_{gs}$  of a FET in an oscillator circuit

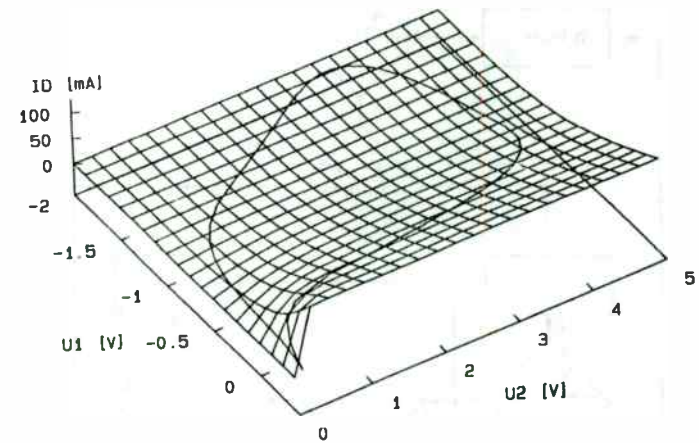


Fig. 1.8: Nonlinear behaviour of  $I_d$  of a FET in an oscillator circuit



## A Simple Road Information Transmitter

Thomas P. Hack  
Comlinear Corporation  
4800 Wheaton Drive  
Fort Collins, CO 80525  
303-226-0500

### Introduction

As an avid Colorado skier, I frequently pass over the Rocky Mountains. Winter driving can be unpredictable and hazardous. Rain changes to snow when ascending and descending, weather moves in quickly and can be radically different on opposite sides of a mountain pass. Systems for advising drivers of adverse and emergency road conditions (flashing lights, electronic signs, etc.) exist but are inadequate.

The Eisenhower tunnel (at an elevation of approximately 2 miles above sea level) is the gateway to the mountains for Colorado's front range residents and would be a great place to transmit road information. The tunnel is long enough to provide sufficient time for a great deal of road information to be passed to motorists. The tunnel also acts as a natural RF barrier between broadcast stations outside the tunnel and road information broadcasts inside the tunnel. Systems are already in place for rebroadcasting commercial radio in tunnels at other locations. Why not transmit road information?

To make the system most effective, it is desirable to transmit road and emergency information on all AM and FM broadcast band channels simultaneously. By transmitting on all channels at once, motorists wouldn't have to flip channels. This eliminates a potentially hazardous distraction, gives more time for receiving road information, and increases the likelihood motorists would use the system. Other locations such as the Lincoln and Holland tunnels in the New York metropolitan area would also benefit from the technique by providing traffic and emergency information.

The following article describes a technique for simultaneously transmitting traffic and road information on all broadcast band channels.

### Design alternatives

The AM broadcast band extends from 535 to 1605 KHz with carrier spacings of 10KHz starting at 540KHz. The extra 5KHz on the ends of the band account for modulation sidebands. A straight forward way to modulate all channels at the same time would be to create a pulse-amplitude modulated signal (with a very low duty cycle pulse to achieve a reasonably flat comb across the broadcast band) and filter the results (Figure 1). Unfortunately an extremely sharp and difficult to implement filter is

required. Assuming filters restrict the audio bandwidth to 4KHz, the upper sideband of the highest desired channel could extend to 1604KHz and the lower sideband of the lowest undesired signal (above the desired signals) could extend to 1606KHz. The ratio of stopband to cutoff frequency is 1.001! Similar problems occur in filtering the undesired comb lines below 535KHz. Filtering can be relaxed by working with the unmodulated components first (Figure 2) but the filter requirements are still intimidating. Bandpass filter 1 requires  $f_s/f_c=1610/1600=1.0063$

Another approach that looks promising at first is shown in Figure 3. In this case a composite comb is generated by amplitude modulating the bandlimited comb from bandpass filter 1 with the lowpass filtered comb. If bandpass filter 1 passes M comb lines, the total number of comb lines produced is equal to:

$$1) \text{ comb lines} = (2N+1)M.$$

The problem with using this approach is that the AM broadcast band has 107 channels--and 107 is prime. Equation 1 tells us we can't get there with this approach. We either end up with too many comb lines--which makes bandpass filter 2 tough to design--or we must be content with covering something less than the full broadcast band. With a

derivative of this approach we can get 107 channels, but it doubles circuit complexity.

At this point, a direct digital synthesis (DDS) approach starts to look quite appealing (Figure 4). Create a lookup table with a representation of the bandlimited comb signal, send it to a digital-to-analog converter, modulate it, add a modest (by comparison) amount of filtering, and you're done. A look at the modulator

We want to generate 107 tones spaced 10KHz apart with the first at 540KHz and the last at 1600KHz. The easiest way to do this is to create a lookup table that is repeated at a 10KHz rate. 54 complete sinusoidal cycles represent the 540KHz tone and 160 complete cycles represent the 1600KHz tone. The table will be clocked at 4MHz. This frequency was selected because it is high enough to keep the first sampling image at well-removed from principal band, yet low enough to allow commonly available EPROMs to be used for the lookup table. A 4MHz clock and a 10KHz table repetition frequency translates into a 400 sample lookup table. This means that a 9 bit counter is required to address the lookup table. Three readily available ICs give us a 12 bit counter chain which will be cycled through only 400 states.

It is important that the counter including all unused flip flops repeats every 400 cycles. This insures that any electrical noise (from TTL shoot through currents) which might couple to power supplies lands at harmonics of 10KHz. Without this precaution it is easy to develop spurious tones that appear as modulation.

Another undesired product which was reduced by design was quantization noise. Quantization noise and virtually all DAC non-idealities produce products at harmonics of 10KHz (in this design)--those that land out-of-band and close to the desired comb spectrum can impose tighter filter requirements on the design. To reduce the quantization error as much as possible, the crest factor of the carrier stored in the lookup table should be as low as possible. Figure 5 shows a non-optimum combination of 107 sinewaves (each with 2 unit-pp amplitude). Figure 6 shows the same 107 sinewaves with the amplitude of each unchanged after the phase relationships have been altered to reduce the crest factor (see appendix for the computer program). The peak-to-peak value went from 168.8 to 34.9 units p-p or a 4.8 to 1 reduction in crest factor.

Figure 7 shows the completed design. Standard noise reduction techniques were used wherever possible to reduce noise. Component placement, use of a ground plane, careful supply decoupling, and differential analog signal paths were used wherever possible.

#### The results

Figure 8 is a broadband look at the spectrum. Because this is a rather broad sweep, only the envelope of the spectrum is visible. The spectrum, as expected has the distinctive  $\sin(x)/x$  shape characteristic of zero-order hold devices such as DACs. Figures 9 and 10 show the spurs that reside just outside the broadcast band. Because of their close proximity, these spurs usually set external filter's order. Quantization error in the lookup table is probably limiting the present design. A deeper and wider lookup table could reduce spurs further. The DAC, a CLC912 is capable of much better spurious performance at megaHertz output frequencies. A 12-bit wide table with 8192 samples would be a good choice.

Figure 11 shows the flatness of the comb spectrum. Most of the flatness error is due to  $\sin(x)/x$  rolloff and could easily be corrected for in the lookup table. Adding a 2MHz

filter (which was already available in the lab) gives the results Figure 12. Note the additional ripple in the comb and the absence of the first image. If a 1.6MHz filter of the same type had been used spectrum above the desired signals would have looked virtually ideal. Figure 13 shows the modulation envelope with a 1.8 Vpp, 1KHz sinewave applied to the audio input. Although the envelope looks noisy, this is really the "comb" carrier (refer back to Figure 6). The spectrum of this signal is shown in Figure 14. In order to make the modulation sidebands clearly visible, only a small part of the total spectrum (near the center of the comb band) is shown. At approximately 90% modulation, U7 (a CLC522) has low distortion-approximately 0.54%.

#### What about FM?

A related technique can be applied to the FM broadcast band(Figure 15). Generate a bandlimited comb using a lookup table and a DAC. Several 100 megasamples/second ECL-input DACs are available that would work quite well here. Frequency modulation is developed in a separate path. The mixer simultaneously places the modulation on all comb lines and translates the output to the final frequency.

#### Summary

Unique modulators have been described that allow simultaneous modulation of all channels in the AM and FM broadcast bands. These devices could be used for improving the safety and efficiency of road travel in a variety of ways.

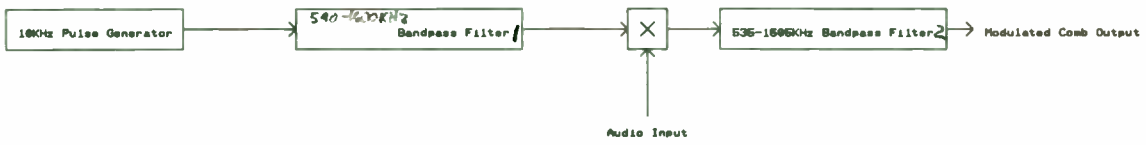


Fig 2

A slightly improved comb modulator			
T.P. Mack			
DATE:	REV	SHEET	OF

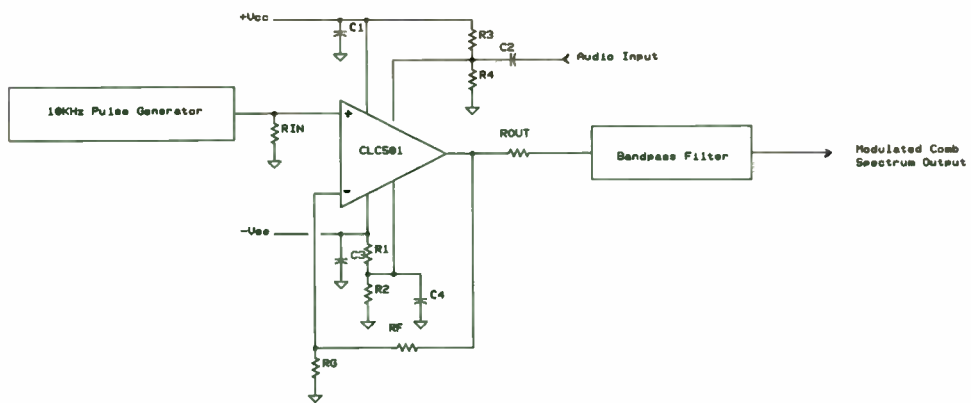
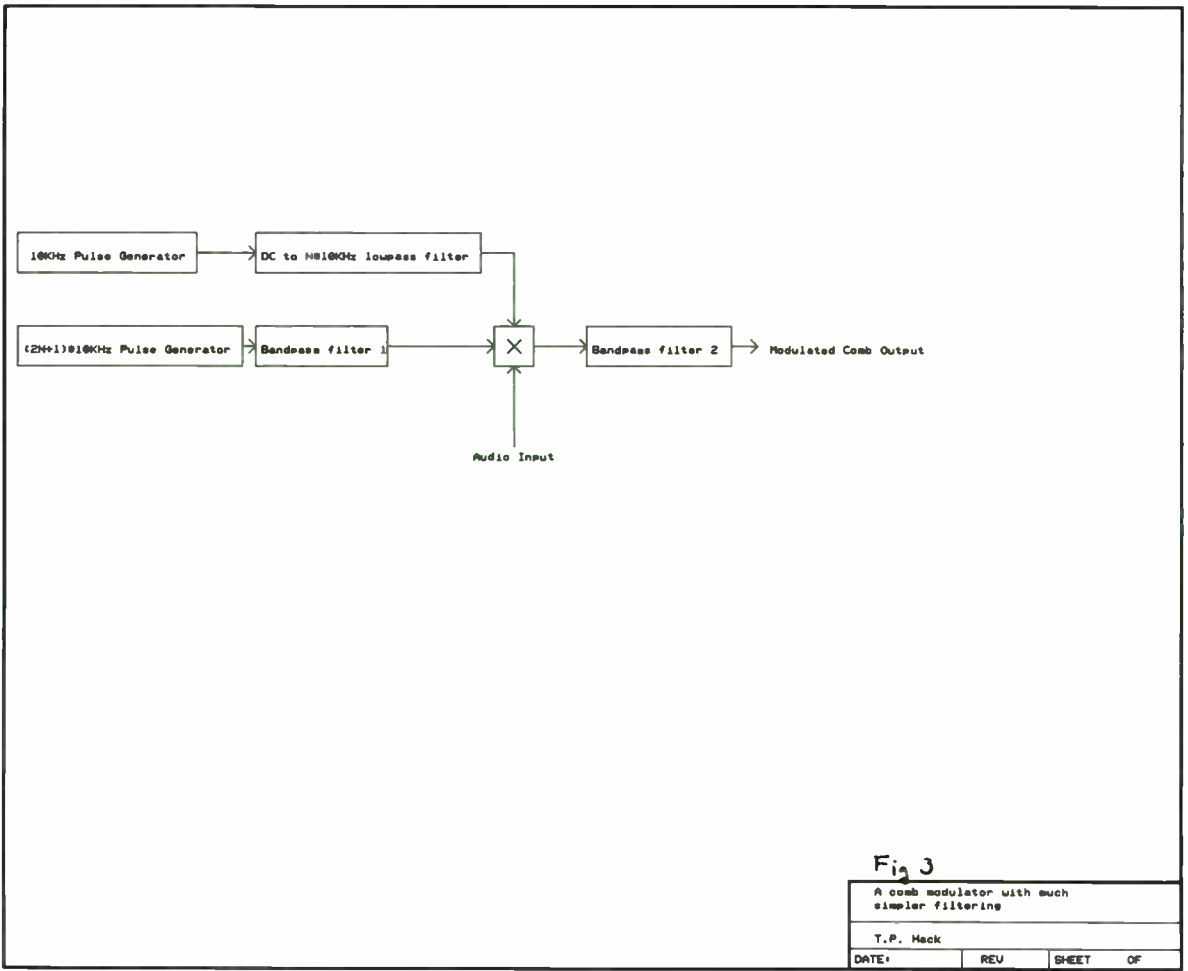
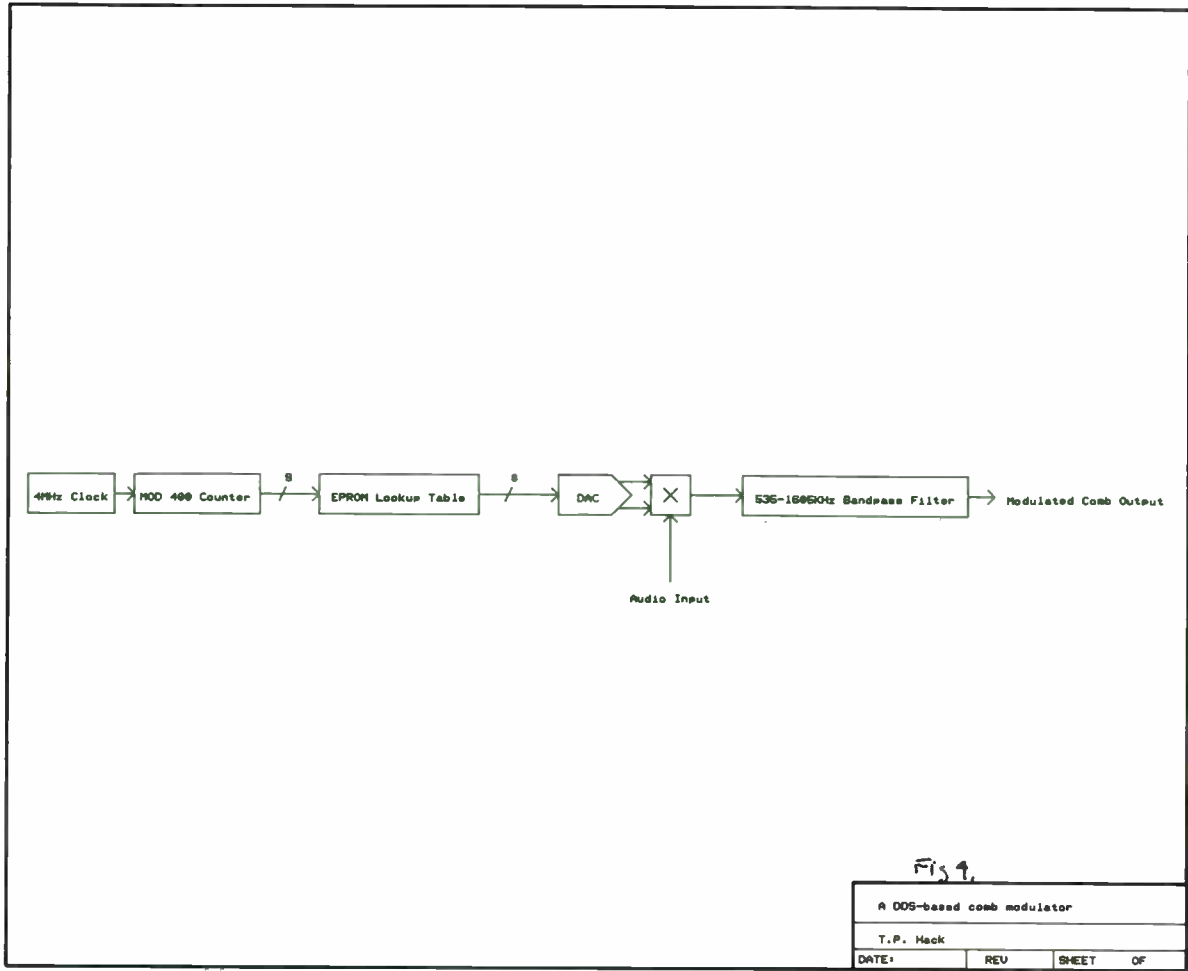


Fig 1 Pulse Amplitude Modulator



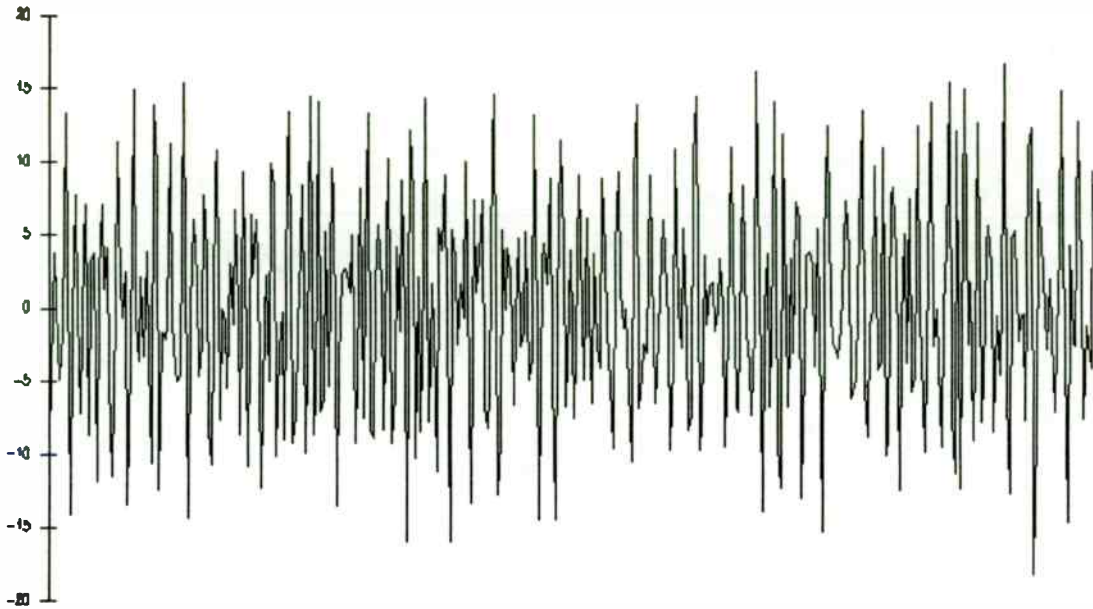


Figure 6. Optimum waveform map.

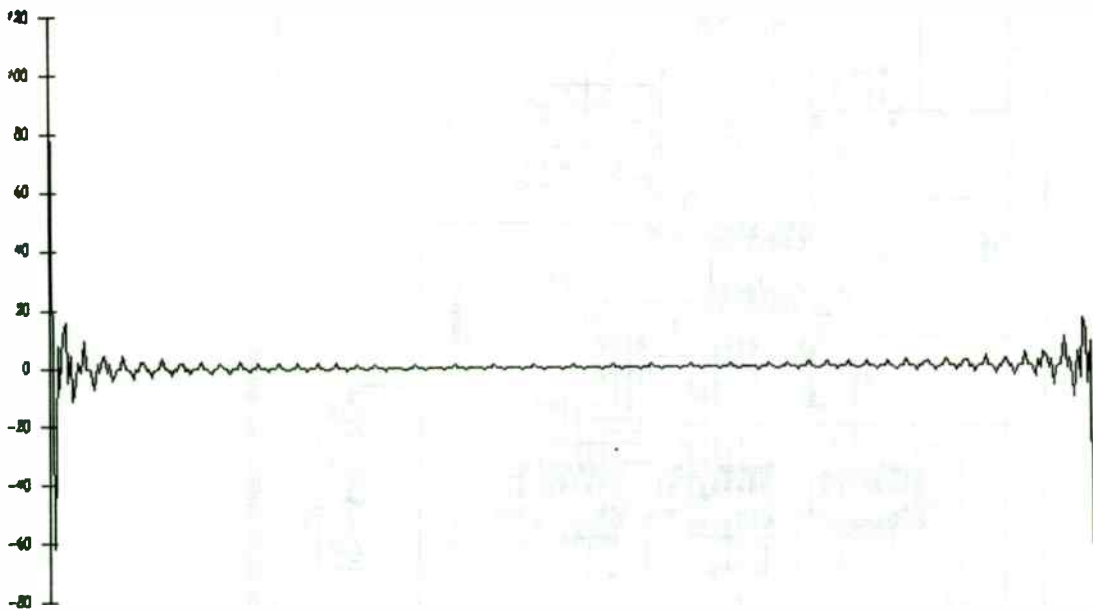


Figure 5. Non-optimum waveform map.

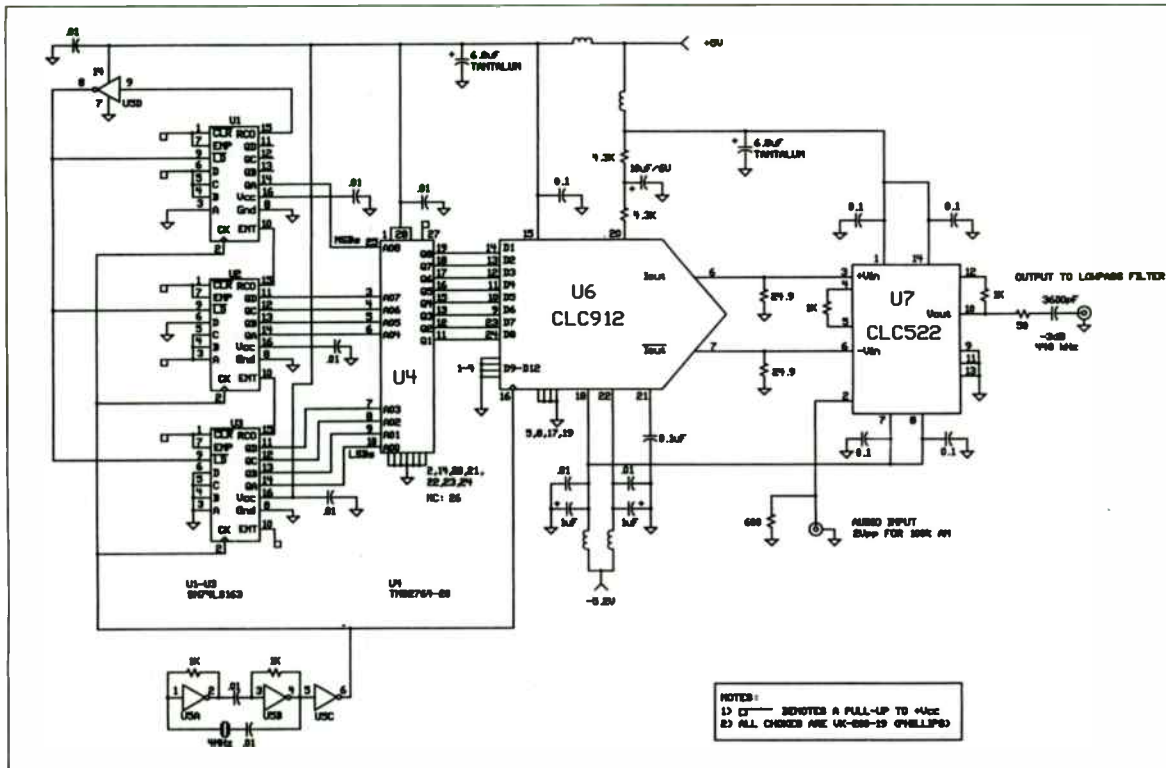


Figure 7. Comb modulator schematic.



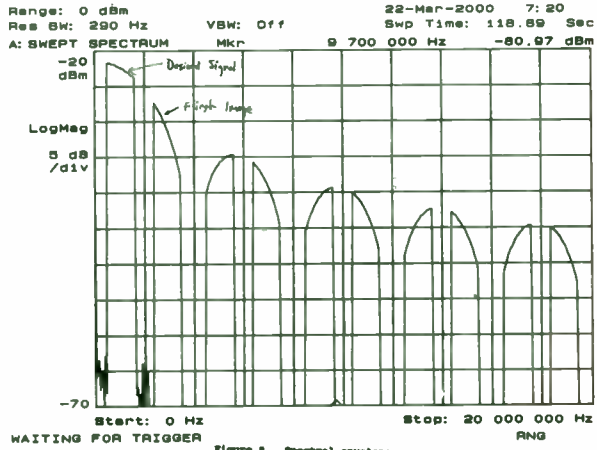


Figure 8. Spectral envelope.

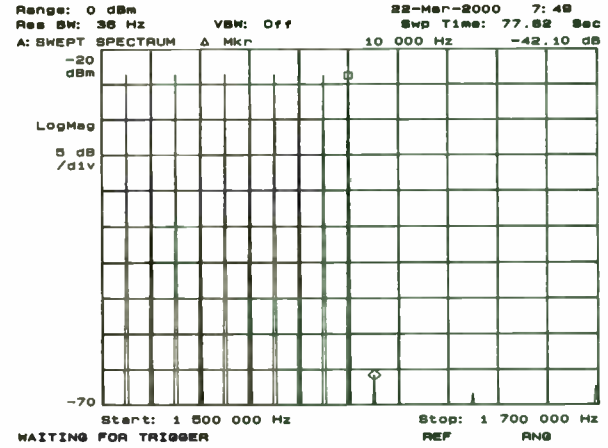


Figure 10. Band edge spur — upper edge.

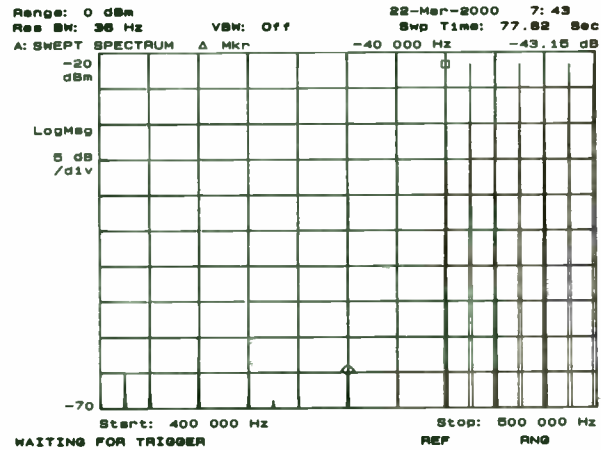


Figure 9. Band edge spur — lower edge.

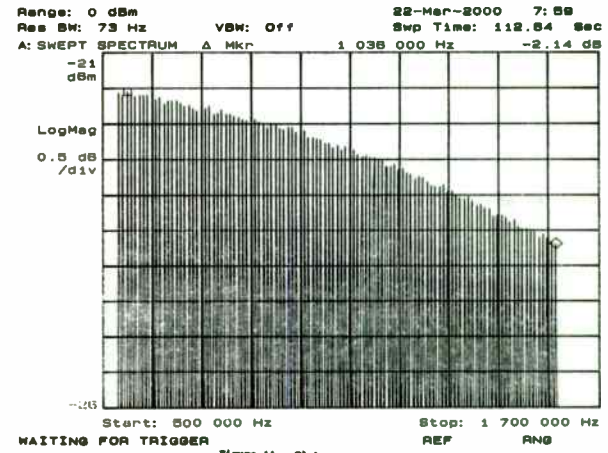


Figure 11. Flatness measurement.

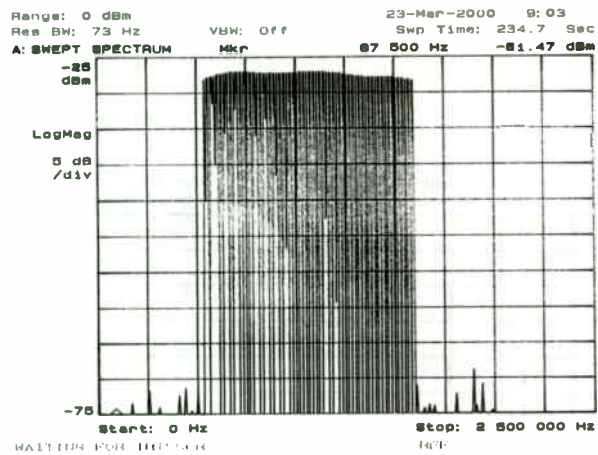


Figure 12. Filtered comb spectrum.

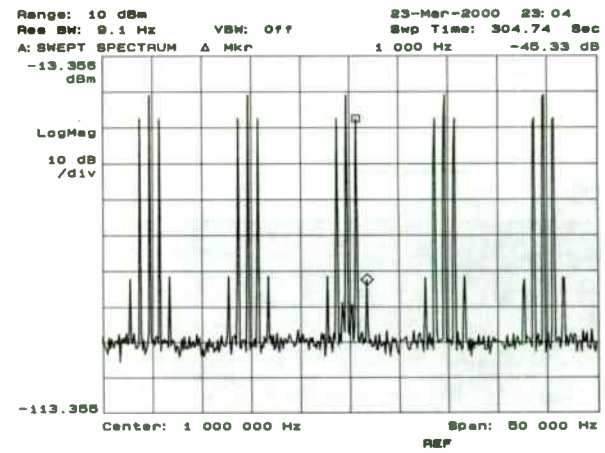


Figure 14. A portion of the modulated comb spectrum.

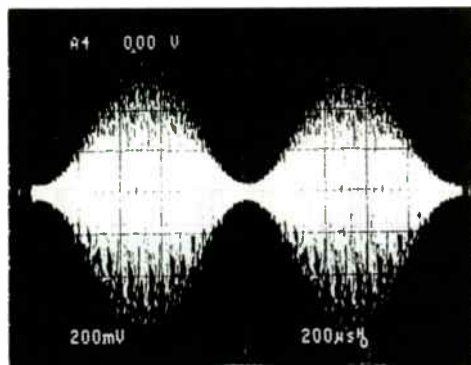
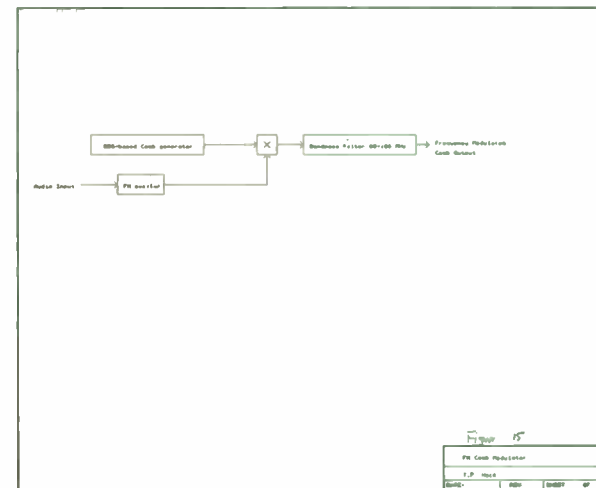


Figure 13. Modulator output.



```

CLS
SCREEN 12
WINDOW (0, -20)-(399, 20)
VIEW (0, 50)-(639, 430)
RLM wave#(399)
pi# = 4# * ATN(1#)
twopi# = 2# * pi#
dphi# = twopi# / 400# 'for the fundamental

'initialize working array
FOR i% = 0 TO 399: wave#(i%) = 0: NEXT i%

FOR freq% = 54 TO 160 STEP 1
LOCATE 1, 1: PRINT freq%
dphil# = freq% * dphi#

peak# = -1E+30
'find largest peak value
FOR i% = 0 TO 399
IF ABS(wave#(i%)) >= peak# THEN
    peak# = ABS(wave#(i%))
    peakpoint% = i%
ELSE
END IF
NEXT i%

'add in new frequency component
'adjust phase to reduce largest excursion
phase# = -dphil# * peakpoint%
IF wave#(peakpoint%) >= 0 THEN phase# = phase# + pi#
FOR i% = 0 TO 399
wave#(i%) = wave#(i%) + COS(dphil# * i% + phase#)
NEXT i%

CLS
'do graph
COLOR 15: PSET (0, wave#(0))
FOR i% = 0 TO 399
LINE -(i%, wave#(i%))
NEXT i%
'do labels
LOCATE 2, 1: PRINT "Comb generator waveform"
NEXT freq%

'find peak to peak value
max# = -1E+30: min# = 1E+30
FOR i% = 0 TO 399
IF wave#(i%) >= max# THEN max# = wave#(i%)
IF wave#(i%) < min# THEN min# = wave#(i%)
NEXT i%
LOCATE 28, 1: PRINT "peak to peak="; max# - min#

'save data to disk
OPEN "c:\wavea.dat" FOR OUTPUT AS #1
FOR i% = 0 TO 399
PRINT #1, wave#(i%)
NEXT i%
CLOSE #1

'generate data for eprom

```

```

'normalize the data
k# = 255 / (max# - min#)
FOR i% = 0 TO 399
wave#(i%) = k# * (wave#(i%) - min#)
wave#(i%) = INT(wave#(i%) + .5)
NEXT i%

OPEN "c:\eprom.dat" FOR OUTPUT AS #1

'pad the first 112 locations
FOR i% = 1 TO 112
PRINT #1, CHR$(0);
NEXT i%

'store the data
FOR i% = 0 TO 399
PRINT #1, CHR$(wave#(i%));
NEXT i%
CLOSE #1

```



AN IMAGE IMPROVEMENT OF MICROWAVE DIFFRACTION TOMOGRAPHY  
FOR 2-DIMENSIONAL INHOMOGENEOUS DIELECTRIC CYLINDER  
BASED ON PROJECTION FUNCTION

K.W. Suh\*, D.Y. Lee\*, Y.C. Han\*,  
C.Y. Park\*, S.Y. Kim\*\*, and J.W. Ra\*\*\*

\* RF Systems Laboratory, Communication System R & D Center,  
Information System Business, SAMSUNG Electronics, 8-2,  
Karak-Dong, Songpa-Ku, Seoul, Korea

\*\* Applied Electronics Laboratory, Korea Institute of Science  
and Technology, P. O. Box 131 Cheongryang, Seoul, Korea

\*\*\* Department of Electrical Engineering, Korea Advanced  
Institute of Science and Technology, Kusung-Dong,  
Yusung-Ku, Taejeon, Korea

Abstract

An alternative to the conventional Fourier diffraction tomography, the Fourier diffraction slice-projection algorithm based upon the projection function is developed under the Born approximation. In order to verify the usefulness and advantage of the presented method, a number of results obtained by numerical simulations and laboratory experiments are shown for a 2-dimensional dielectric cylinder and the limiting factors in the degrading reconstructed image due to Born approximation are also identified in terms of projection function. The experimental results performed at X-band provide a good agreement with the simulated ones. In addition, the possibility of image improvement for degraded images by modifying the projection function is also discussed.

1. Introduction

In comparison with the extremely high frequency range of X-ray, however the relatively lower frequency range of microwave restricts straightforward application of the tomography to microwave imaging of dielectric objects. It comes from the fact that the electromagnetic field inside a dielectric object can not propagate along straight lines but travels in multi-path according to the refraction and diffraction. There is no rigorous reconstruction algorithm accounting the effects of refraction and diffraction precisely. A classical method for the perturbation analysis of electromagnetic or acoustic wave propagation in slightly inhomogeneous media is called the Born approximation. Under this approximation with measurement of scattered fields in far-field region, a tomography-like reconstruction algorithm called the Fourier diffraction projection theorem has been developed by many researchers[1,2,3].

The scattered fields measured in far-field region by using frequency and angular diversity take the form of spatial-frequency data in polar format. In particular, it is well recognized that these data after the range correction are directly equal to the Fourier transform of the permittivity profile of the dielectric scatter. In consequence, the reconstruction process on the Fourier diffraction projection theorem consists of interpolation and inverse fast Fourier transform[IFFT]. However the interpolation process becomes key factors of the degradation in retrieved image and the computation time in image processing[4].

Recently in the microwave imaging, a spatial slice theorem has been introduced and the computerized tomography(CT) algorithm was directly employed in the reconstructing image without interpolation processing[5]. In addition, it was shown that the result obtained by this procedure led to the better image quality

in comparison with that obtained by such algorithms as direct Fourier inversion[6] and circular convolution integral[7]. As a extensive application of this concept, it was found that the same method can be applied to the reconstruction of dielectric objects in the sense of the first-order Born approximation. Also the limiting factors of degradation in the reconstructed image can be identified in terms of the projection function and the necessary condition on the first-order Born approximation[8,9].

In this paper, as a way of overcoming the limitation of the conventional microwave imaging under the Born approximation, an alternative to the Fourier diffraction projection theorem is developed, which is called a diffraction slice-projection theorem. The proposed method based on a projection function is derived as an extension of the projection tomography for inducting the diffraction effect under the first-order Born approximation.

Two of the interesting features on the presented method are that first, the interpolation processing is not necessitated and the conventional CT algorithm can be directly used to retrieve the microwave imaging of 2-dimensional dielectric objects. Second, the physical reason for the degraded image, resulting from the violation of the necessary condition on the first-order Born approximation, can be readily explained by means of projection function.

In order to verify the usefulness and adequacy of the presented method, various effects of reconstructed images are investigated by numerical simulations and laboratory experiments for a 2-dimensional dielectric cylinder as a function of object size, dielectric constant, and the bandwidth of transmitting frequencies.

Furthermore, the possibility of image improvement for degraded image by modifying the projection function is also discussed

here.

## II. Formulation

### A. The integral equation for a 2-dimensional dielectric cylinder

Consider the 2-dimensional dielectric cylinder, whose permittivity and permeability are  $\epsilon_0 \epsilon_r(x,y)$  and  $\mu_0$ , respectively, distributed over arbitrary cross-sectional shape S as shown in Fig. 1. The scalar wave equation is given by

$$\nabla^2 u^t(x,y) + k^2 u^t(x,y) = -k^2 [\epsilon_r(x',y')-1] u^t(x,y), \quad (1)$$

where  $u^t$  denotes the total field,  $k$  is the free-space wavenumber and  $\epsilon_r$  represents the relative dielectric constant of object.

The two-dimensional free-space Green's function of  $G(x,y;x',y')$  satisfies

$$\nabla^2 G(x,y;x',y') + k^2 G(x,y;x',y') = -\delta(x,y;x',y'). \quad (2)$$

The solution of eq. (2) is given as

$$G(x,y;x',y') = -j/4 H_0^{(2)}(k \sqrt{(x-x')^2+(y-y')^2}), \quad (3)$$

where  $H_0^{(2)}$  denotes the zeroth-order Hankel function of the second kind.

When a known incident field of  $u^i$  propagating along the x-axis illuminates the object, then the scattered field  $u^s$  can be exactly calculated with an integral representation given by[2]

$$u^s(x,y) = -jk^2/4 \iint_{-\infty}^{\infty} dx' dy' [\epsilon_r(x',y')-1] u^t(x',y') \times H_0^{(2)}(k \sqrt{(x-x')^2+(y-y')^2}), \quad (4)$$

where the total field  $u^t$  is a sum of incident field  $u^i$  and scattered field  $u^s$ .

### B. The projection function based on the Born Approximation

If the dielectric profile of object has a slight perturbation of  $\epsilon_r(x,y)$  over the background medium, one may think that the magnitude of scattered field is smaller than that of the incident field. For the plane wave incidence propagating along the x-axis, the first-order Born approximation in eq. (4) can be simplified by representing the  $H_0^{(2)}$  into the dominant term of its asymptotic expansion and by replacing the  $u^t$  into  $u^i = e^{-jkx}$ , respectively. Compared with the object extent, the distance between the object and the mono-static antenna is taken large enough to regard iso-range contours in S as straight lines normal to the incident wave vector. Under the above situation and Fig. 1, the backscattered fields are measured by using the wave number diversity. Then eq. (4) can be rewritten[9]

$$\begin{aligned} T(\theta; \alpha) &= u^s(\rho_0) \times \sqrt{8\pi\rho_0} / \sqrt{k^3} e^{j(k\rho_0 + \pi/4)} \\ &= \iint_{-\infty}^{\infty} dudv [\epsilon_r(u,v) - 1] e^{-j\alpha u} \end{aligned} \quad (5)$$

where  $\rho_0$  is the distance between the antenna and the origin of the coordinates system at the object,  $T(\theta; \alpha)$  denotes the backscattered field normalized by the amplitude and phase of a reference line target located at the origin,  $\theta$  is the rotation angle of the object measured from the u-axis, and  $\alpha$  is equal to  $2k$ .

To evaluate the integral for variable  $v$  in eq. (5), let us define the projection function of  $P(\theta; u)$  given by :

$$P(\theta; u) = \int_{-\infty}^{\infty} dv [\epsilon_r(u,v) - 1], \quad -a < u < a, \quad (6)$$

where  $a$  is the radius of the circular cylinder. It should be noted that  $P(\theta; u)$  in eq. (6) is simply the tomographic projection of the relative dielectric constant distribution onto the u-axis. Substituting eq. (6) into eq. (5) leads to a 1-dimensional inverse scattering problems as follows :

$$T(\theta; \alpha) = \int_{-\infty}^{\infty} du P(\theta; u) e^{-j\alpha u}. \quad (7)$$

If one may know all the spectrum for the backscattered field  $T(\theta; \alpha)$  as a function of spatial frequency  $\alpha$ , then eq. (7) can be analytically evaluated by the Fourier transform relation. It can be expressed as following :

$$P(\theta; u) = 1 / 2\pi \int_{-\infty}^{\infty} d\alpha T(\theta; \alpha) e^{j\alpha u} \quad (8)$$

In practice, it is impossible to get the backscattered field  $T(\theta; \alpha)$  of object for all the spatial frequency. The use of finite bandwidth with discrete spatial frequency leads to the backscattered fields corresponding to each spatial frequency. In this case the solution of eq. (8) can be obtained directly by using the numerical integration or inverse fast Fourier transform (IFFT) from the discrete data of  $T(\theta; \alpha)$ .

In order to retrieve the tomographic image of object, the object should be rotated so as to gather the projection data as a function of aspect angle. Therefore repeating the above process for different aspect angle  $\theta$ , one may obtain the projection data in all the scanning angle. In consequence, the relative dielectric constant distribution  $\epsilon_r(x,y)$  can be readily reconstructed only by applying the conventional X-ray CT algorithm to all projection data. The reconstruction algorithm adopted in this article is the convolution back-projection[10]

with the Shepp and Logan filter[11].

C. The necessary condition on the first-order Born approximation

As previously stated, the first-order Born approximation is valid only when the magnitude of the scattered field,

$$|u^t - u^i| = |u^s| < |u^i|, \quad (9)$$

is smaller than that of the incident field. If the object is the 2-dimensional cylinder of constant refractive index, the expression for the necessary condition on the first-order Born approximation can be derived as a function of the object size, refractive index, and transmitting frequency[2,9].

First let us consider the imaging geometry for the forward scattered line measurement as shown in Fig. 2-(a). This geometry is widely used in microwave diffraction tomography based on the Fourier diffraction projection theorem[1,2,3]. In order to define the problem state more clearly, the propagation of incident plane wave is taken in the direction of the wave vector  $k_0 (= k_x)$ . If the object to be tested is a large, the field inside the object can not be directly expressed as following :

$$u_{obj}^t = A e^{j k_0 \cdot r} \quad (10)$$

where  $r$  is the position vector in the object space and  $A$  denotes the amplitude of total field inside the object. On the other hand, the phase term of eq. (10) will has a form of the change in refractive index of  $\delta\sqrt{\epsilon_r}$ . Taking into consideration of a ray through the object center and parallel to the propagation vector of incident plane wave, it is possible to deduce that the total field inside the object becomes a slow or fast version of the

incident wave such a

$$u_{obj}^t = A e^{j (1 + \delta\sqrt{\epsilon_r}) k_0 \cdot r} \quad (11)$$

where  $\delta\sqrt{\epsilon_r}$  represents the difference in refractive index between the object and its surrounding medium.

From this fact, one may directly drive the phase difference between the incident field and the total field inside the object which is approximately equal to the integral through the object of the change in refractive index.

Therefore, the total phase difference,  $\delta\varphi$ , through the 2-dimensional dielectric cylinder is approximately given by

$$\delta\varphi = 2\pi \delta\sqrt{\epsilon_r} d_{max} c / f, \quad (12)$$

where  $d_{max}$  denotes the maximum size of object diameter,  $f$  indicates the highest frequency of source used as transmitting signals, and  $c$  is the light velocity in vacuum. In order to find the criterion for the first-order Born approximation to be valid, a necessary condition becomes that the change in phase,  $\delta\varphi$ , between the incident field and total field inside the object should be less than  $\pi$  radian. A mathematical expression for this condition can be represented as follows:

$$\delta\sqrt{\epsilon_r} d_{max} < c / 2f. \quad (13)$$

The physical meaning of eq. (13) stems from the Nyquist sampling criterion so that sampled data can provide a recovery of the original signal without the aliasing effect in the reconstructed image. Therefore, it is very interesting to note that eventually



the necessary condition on the first-order Born approximation has a something to do with the Nyquist sampling criterion, which depends on the phase difference between the incident field and total field propagating through the largest size of object.

As a second case, let's take into consideration of driving the criterion for the geometry of the backward scattered line measurement, as shown in Fig. 2(b). In this case, the same theory mentioned in first case is used for obtaining the tomographic image. One may know that well known radar imaging techniques such a synthetic aperture radar(SAR) and inverse SAR(ISAR), extensively used in microwave imaging and remote sensing, are a special case of this geometry in view of mono-static measurement. Therefore, the same procedure for derivation of the criterion can be utilized in this case for the previous one while keeping the same physical meaning.

As above discussed, the parameters involved in the criterion is related to the phase difference depending upon the line integral of wave propagation through the largest size of object. For the mono-static measurement case, the total propagating distance inside the object for Fig. 2(b)' is a twice than that of the first case which considered the ray propagating along the object center. However, for the other diffracted rays except the above case, the propagating distance is less than that of mono-static measurement. Therefore, one may directly express the criterion of the necessary condition for Fig. 2(b) given by[9]:

$$\delta \sqrt{\epsilon_r} d_{\max} < c / 4f . \quad (14)$$

The constant factor 4 in eq. (14) comes from the round-trip distance compared with eq. (13).

#### D. The criterion on the angular increment

As a similar set of criterion required in imaging process, one should also take into account the constraint on the angular increment of rotation angle. In the imaging geometry of Fig. 1 such as monostatic case, the necessary condition on the angular increment of  $\delta \theta$  over  $360^\circ$  is directly expressed as follows:

$$\delta \theta = \sin^{-1} \{ c / 2d_{\max} ( 1 + \delta \sqrt{\epsilon_r} ) f \} , \quad (15)$$

In general, since  $\delta \sqrt{\epsilon_r}$  is not known, it can be conveniently approximated as zero under the Born approximation. Eq. (15) also comes from the sampling criterion in order to obtain the image without aliasing effect.

### III. Results and Discussions

To investigate how the retrieved image of object is degraded under the Born approximation, a number of projection functions and reconstructed images for the 2-dimensional dielectric object are presented here.

#### A. The Projection Functions

For a fixed rotation angle  $\theta$ , let us first investigate the projection function of a homogeneous dielectric object. The wave number diversity for backscattered measurement is exploited in the frequency range from 0.1 to 4.0 GHz with frequency increment of 0.1 GHz.

Fig. 3 shows the normalized projection functions for the dielectric cylinder of radius  $a=10\text{cm}$  as a function of  $\epsilon_r =1.01, 1.1, 1.5, 2.0, 3.0,$  and  $4.0$ . One may find that the extents of projection functions for  $\epsilon_r =1.01$  and  $1.1$  give nearly the same size of actual diameter and also show symmetric forms with

respect to  $u=0$ , respectively. But, as  $\epsilon_r$  increases, the extents of projection functions except those for  $\epsilon_r = 1.01$  and  $1.1$  are shown to be greater than the extent of  $\epsilon_r = 1.01$  and also have asymmetric forms. Therefore one may know that the extents of these projections have something to do with the relative dielectric constant  $\epsilon_r$ . It is very interesting to note that projection functions extend on an interval  $[-a, 2a\sqrt{\epsilon_r} - a]$ , optical length which is different from the real one  $[-a, a]$ . This is the consequence of the difference between wave velocities inside and outside the object [9,12]. As was expected, the formulation of eq. (5) leads to a good agreement in case that the necessary condition on the first-order Born approximation satisfies eq. (14).

Next, in order to check the same effect for an inhomogeneous dielectric object, Fig. 4 presents the projection function of the hollowed dielectric cylinder, with inner radius 5cm and thickness 15cm between inner and outer radius, as a function of  $\epsilon_r = 1.01, 1.1, 1.5, 2.0,$  and  $3.0$ . The reconstructed projection functions have the same characteristic as Fig. 3 except the concaved forms around the object center. The extents of the projection functions for  $\epsilon_r = 1.01$  and  $1.1$ , which satisfy the condition of eq. (14), give the actual object size of 40cm, but the other extents for  $\epsilon_r = 1.5, 2.0,$  and  $3.0$  do not provide the actual size due to the excessive violation of eq. (14). As  $\epsilon_r$  increases, it is also shown that the projection functions are remarkably extended because the larger object size is chosen in comparison to Fig. 3.

Finally, to confirm the similar effects for the offset frequency signals such the X-band, the same procedure is performed. Fig. 5 depicts the reconstructed projection functions for the homogeneous dielectric cylinder of radius 5cm and  $\epsilon_r = 1.1, 2.1,$  and  $2.8$ . The frequencies from 7.5 to 11.5 GHz with

increment of 0.08 GHz were used. The projection functions for each dielectric constant have two peak profiles, which corresponds to the front and rear surfaces of object along the  $u$ -axis, respectively. Differing from the projection functions of Figs. 3 and 4, the peak profiles are obtained because of the loss for scattered data corresponding to the frequencies below 7.5 GHz. The width between two peaks of the projection function for  $\epsilon_r = 1.1$  provides almost the actual size of object because the condition of eq. (14) is still conserved. On the other hand, for the cases of  $\epsilon_r = 2.1$  and  $2.8$ , they do not show the actual size because of the violation of eq. (14). But it is also shown that the width between two peaks is coherently increased in proportion to the refractive index of object.

From all the results, it is concluded that the projection function providing the actual size of object can be reconstructed if only the condition on the first-order Born approximation is satisfied. On the other hand, when the necessary condition is violated, it will require some remedy to modify the extent of projection function from optical length to real one. This procedure will be introduced and deeply discussed in the chapter IV.

#### B. The Reconstructed Images

In order to explore the retrieved images obtained by using the above reconstructed projections, the angular diversity is also accounted by rotating the object with angular increment of  $1.5^\circ$  over  $360^\circ$  to satisfy the Nyquist sampling criterion of eq. (15).

Figs. 6(a), (b), (c), and (d) are presented for the cases of  $\epsilon_r = 1.01, 1.1, 1.5,$  and  $2.0$ , respectively. Since the extents of projection functions for  $\epsilon_r = 1.01,$  and  $1.1$ , as shown in Fig. 3, have nearly the symmetric forms over  $u=0$  and present the

actual diameter of object, Figs. 6(a) and (b) provide fairly good results. On the other hand, Figs. 6(c) and (d) whose projection functions, as shown in Fig. 3, have the asymmetric forms for  $u=0$  show a severe degradation compared with Figs. 6(a) and (b).

Since eq. (14) implies that the quality of the reconstructed image for dielectric cylinder depends only upon the involved parameters  $\epsilon_r$ ,  $a$ , and  $f$ , so a similar set of simulation is also performed in order to see the various effects of retrieved images. The results for the different values of parameters  $\epsilon_r$ ,  $a$ , and  $f$  are shown in Figs. 7(a), (b), and (c). Figs. 7(a) and (b) show the reconstructed images of dielectric cylinder of  $a=15\text{cm}$  for  $\epsilon_r = 1.1$  and  $1.2$ , respectively, while keeping the same condition of frequency bandwidth used in Fig. 6. Unlike the Figs. 6(a) and (b), these provide poor resolution because the values of the necessary condition on eq.(14) are greater than those of Figs. 6(a) and (b). From this result, one may find that the retrieved images get worse as the product of refractive index and radius gets larger. On the other hand, Fig. 7(c) shows the effect of reconstructed image for the center frequency shift under the same condition as Fig. 6(b), which has the same bandwidth of 3.9 GHz from 1.1 to 5.0 GHz while keeping the condition on eq. (14). This image leads to a contour profile which provides nearly an actual circumference of object, which results from the loss for low frequency scattered fields in the range from 0.0 to 1.0 GHz.

Next, let us look into the reconstructed images for inhomogeneous dielectric cylinder of a hollowed object. The images depicted in Figs. 8(a), (b), (c), and (d) are obtained from the projection functions of Fig. 4 for the cases of  $\epsilon_r = 1.01, 1.1, 1.5,$  and  $2.0$ , respectively. The extents of projection functions for  $\epsilon_r = 1.01$  and  $1.1$ , as shown in Fig. 4, provides the almost symmetric forms over  $u=0$  and show the actual diameter of

object. Therefore, Figs. 8(a) and (b) provide fairly good results. On the contrary, one can readily deduce that Figs. 8(c) and (d) show the severely degraded images due to the violation of eq. (14) or asymmetric forms of projection functions.

Finally, in order to examine the effect of reconstructed image obtained from the higher frequency signals, the simulations are also performed at the X-band for the projection functions of Fig. 5. The frequency range from 7.5 to 11.5 GHz with frequency increment of 0.08 GHz is exploited. Figs. 9(a), (b), and (c) are revealed for the cases of  $\epsilon_r = 1.1, 2.1,$  and  $2.8$ , respectively. Fig. 9(a) shows the edge detected image whose size has the same circumference of object because the width between two peaks of the projection function provides almost the actual diameter of the object. On the other hand, Figs. 9(b) and (c) which correspond to the cases of  $\epsilon_r = 2.1$  and  $2.8$ , respectively, show the enlarged size of actual diameter because of the violation in eq. (14).

To testify the inhomogeneous case, the simulated result for a hollowed dielectric cylinder is shown in Fig. 9(d) in case of the relative dielectric constant 2.8, inner radius 5.0cm, and thickness 0.2cm, which is between inner and outer radius. Although the test object is the hollowed dielectric cylinder, it still satisfies the necessary condition in eq. (14). In consequence, one may deduce that the reconstructed image provides the characteristic of edge detection of object as predicted in Fig. 7(c) and Fig. 9(a).

From all the results, it is concluded that in order to retrieve the reconstructed image without any degradation under the Born approximation, it should keep the necessary condition on the first-order Born approximation in eq. (14).

### C. The Experimental Results

In order to verify the proposed method by means of an actual measurement, an experimental work was also performed in an anechoic chamber at the X-band. The blockdiagram of experimental set-up for data acquisition is shown in Fig. 10(a) and the photograph for the transmitting, receiving antennas and turntable including object is also shown in Fig. 10(b). The backscattered fields with the frequency diversity are measured by using HP-8510A network analyzer in the frequency range from 7.5 to 11.5 GHz with frequency increment of 0.08GHz. The angular diversity is also implemented by rotating object discretely in the  $\theta$ -direction with interval  $1.5^\circ$  over  $360^\circ$ , which also satisfies the Nyquist sampling criterion enough to avoid the aliasing effect in the reconstructed images.

First, let us investigate the effect of image reconstructed by a single frequency for the homogeneous dielectric object. If the wave length of the source is much greater than the diameter of object, one may assume the object as a line target. Fig. 11(a) shows the intensity of the reconstructed image for two line objects, which are the nylon ( $\epsilon_r=3.5$ ) and acethal ( $\epsilon_r=4.5$ ) with diameter of 1.7cm, separated by 30cm for the cross-range direction. Due to the use of the single frequency of 10 GHz, one may find that the high sidelobe exists in the retrieved image as a background noise. this effect can be quite removed by introducing the multi-frequency. The difference of relative dielectric constants between the nylon and acethal yields the different reflectivity in the retrieved image.

Next, consider the effect of the retrieved image for the 2-dimensional object by introducing the multi-frequency. Fig. 11(b) reveals the reconstructed image for the hollowed p.v.c pipe under the same condition of Fig. 9(d). It provides a very similar result in comparison with Fig. 9(d) obtained via computer simulation. Comparing to Fig. 11(a), the use of multi-frequency

produces the high image quality in view of quite low sidelobe and high resolution. In general the resolution of microwave image depends only on the frequency bandwidth of transmitting source. Hence the resolution of Fig. 11(b) is given by  $3.75\text{cm}(=c/2B$  :  $c$  is the light velocity in vacuum and  $B$  denotes the signal bandwidth).

From these results, it is particularly interesting to note that the edge profile of dielectric object can be successfully reconstructed at the X-band if only the necessary condition on the first-order Born approximation in eq. (14) becomes a valid.

#### IV. A consideration of image improvement

##### A. A Modified Projection Function

In order to improve the degraded image caused by the violation of the necessary condition on the first-order Born approximation, the procedure for modifying the projection function consists of two steps. one is to find the average refractive index of object from the extent of projection function according to the variable  $\theta$ . It can be expressed as follows :

$$\sqrt{\epsilon_r^{\text{avg}}(\theta)} = d_{\text{opt}}(\theta) / d(\theta), \quad (16)$$

where  $d_{\text{opt}}$  is the extent of optical length and is known from the projection function and  $\sqrt{\epsilon_r^{\text{avg}}}$  is the average refractive index of object when the incident wave propagates in the  $\theta$  direction. If we know the object size of  $d$  as a pre-information, then average refractive index of object can be easily calculated by means of eq. (16).

The other is to find the modified projection function of  $P_{\text{mod}}(\theta;u')$  related by  $P(\theta;u)$  and the average refractive index.

It is shown that the extent of this modified projection function provides approximately the same diameter of object. It can be defined as follows:

$$P_{\text{mod}}(\theta; u') \simeq P(\theta; u), \quad -a < u < 2a\sqrt{\epsilon_r^{av}} - a, \quad (17)$$

where

$$u' = -a + 1 / \sqrt{\epsilon_r^{av}} [u + a], \quad -a < u' < a. \quad (18)$$

When  $\sqrt{\epsilon_r^{av}}$  is nearly 1.0,  $P_{\text{mod}}(\theta; u')$  is almost reduced to  $P(\theta; u)$  which is the result given by the Born approximation. The modifying procedure of above equations is approximately valid to some extent although the necessary condition on the first-order Born approximation is violated. However the exact analysis on the limitation of this method is not specified yet and currently being performed by means of more quantitative method.

## B. The Simulated Results

First consider the improved image for homogeneous dielectric cylinder by the presented method. Figs. 12(a) and (b) show the improved images reconstructed from the new projection data for Figs. 6(c) and (d) with  $\epsilon_r = 1.5$  and  $2.0$ , respectively, and also provide remarkable results. From these images, one may know that Fig. 12(a) presents the better image compared with Fig. 12(b) since the effect of refraction inside the object for Fig. 6(c) is less than that for Fig. 6(d).

Next let us investigate the improvement for the inhomogeneous cylinder. A rectangular dielectric cylinder with square cross-section has been studied for the numerical simulation. The cross-section area of dielectric cylinder is  $2.4$

$\lambda \times 2.4\lambda$  where  $\lambda (=7.5\text{cm})$  is the wave length of  $4.0$  GHz. Its area is equally divided into two subdivision of  $S_1$  and  $S_2$  and their relative dielectric constant are  $\epsilon_{r1}$  and  $\epsilon_{r2}$ , respectively. The variations of  $\epsilon_{r1}$  and  $\epsilon_{r2}$  are given in the ranges from  $1.01$  to  $1.5$  and from  $1.02$  to  $2.0$ , respectively.

The reconstructed image represents the relative dielectric constant of object, which is normalized by a maximum real value of reconstructed profile. The frequency range from  $0.1$  GHz to  $4.0$  GHz with increment of  $0.1$  GHz is chosen. The number of angular increment is  $100$  over  $360^\circ$  with increment of  $3.6^\circ$  which satisfies the Nyquist sampling criterion.

In order to explore the images reconstructed from the projection data taken by a function of angle  $\theta$  over  $360^\circ$ , Figs. 13(a), (b) and (c) are shown for the cases of  $\epsilon_{r1} = 1.01, 1.1, 1.5$  and  $\epsilon_{r2} = 1.02, 1.2, 2.0$ , respectively. Since Figs. 13(a) and (b) satisfy the condition of eq. (14), the retrieved images provide fairly good results. On the other hand, Fig. 13(c) shows a severely degraded image in comparison with Fig. 13(a). This results from the extended interval of projection function over each rotation angle  $\theta$  and its asymmetric form with respect to  $u = 0$ . Fig. 13(d) represents the improved images reconstructed from the modified projection data for Fig. 13(c). One may know that it also provides a remarkable result compared with degraded image of Fig. 13(c).

From these results, it is concluded that the presented method leads to a fairly good improved image for even the inhomogeneous dielectric profiles. In the future the analysis on the limitation of this improved method should be performed by means of more quantitative method.

## V. Conclusions

In this article, it was demonstrated that the proposed

method based upon projection function can be applied to obtaining the imaging of dielectric objects irrespective of homogeneous and inhomogeneous profiles and the experimental results performed at the X-band are also in a good agreement with the simulated results. Furthermore, the limiting factors of degradation in the reconstructed image under the Born approximation are extensively identified and explored in terms of not only the projection function, but also the necessary conditions on the first-order Born approximation and the angular increment.

From all the results, it is concluded that dielectric image can be successfully reconstructed if only the necessary condition on the first-order Born approximation is satisfied. On the contrary, when the necessary condition on the Born approximation is violated, the retrieved images reveal the severe degradation. In order to overcome the limitation of the Born approximation, the presented method based on the coherent tomographic projection function has shown that the fine improved images can be obtained by employing the procedure for modification of projection data.

The interesting features on the presented method are the follows: first since it does not require the interpolation processing needed to convert the scattered data in polar coordinate to ones in cartesian coordinate, one can directly make use of the conventional CT algorithm such a convolution-backprojection in order to retrieve the microwave diffraction tomography. Second the physical reason for degraded image can be readily explained in terms of a coherent projective function. Finally, it has the possibility of image improvement for degraded image by modifying the projection data.

Therefore the presented method will be expected to play an important role in removing the limitation of the first-order

Born approximation, especially in applying the nondestructive testing and the diagnosis of the human body in the future. Further work is being developed in order to analyze the criterion on the limitation of this method.

#### Acknowledgement

The authors would like to thank Dr. S. G. Kim for fruitful discussions and experimental work during the preparation of this article.

#### References

- [1] A. J. Devaney, "Reconstructive tomography with diffracting wave-fields," inverse problems, vol. 2, pp. 161-183, 1986.
- [2] M. Slaney, A. C. Kak, and L. E. Larsen, "Limitation of imaging with first-order diffraction tomography," IEEE Trans. Microwave Theory Tech., vol. MTT-32, pp. 860-873, 1984.
- [3] T. H. Chu and K. Y. Lee, "Wide-band microwave diffraction tomography under Born approximation," IEEE Trans. Antennas Propagat., vol. AP-37, pp. 515-519, 1989.
- [4] R. M. Mersereau and A. V. Oppenheim, "Digital reconstruction of multidimensional signals from their projections," Proc. IEEE, vol. 62, pp. 1319-1338, 1978.
- [5] K. W. Suh, S. Y. Kim, and J. W. Ra, "A new spatial slice theorem for microwave imaging," Microwave and Optical Technology Lett., vol. 3, pp. 362-365, Oct. 1990.
- [6] N. H. Farhat, C. L. Werner, and T. H. Chu, "Prospect for three-dimensional projective and tomographic imaging radar networks," Radio Sci., vol. 19, Set.-Oct. pp. 1347-1355, 1984.

- [7] D. L. Mensa, S. Halevy, and G. Wade, " Coherent Doppler tomography for microwave imaging, " Proc. IEEE, vol. 70, pp. 254-261, Feb. 1983.
- [8] K. W. Suh, S. G. Kim, and J. W. Ra, " Improvement of the Born approximation in the reconstruction of dielectric object by using the projection function, " Inter. IEEE/AP-S Symp., pp. 1620-1623, Jun. 1991.
- [9] K. W. Suh, S. G. Kim, J. W. Ra, and S. Y. Kim, " Microwave imaging for dielectric objects under the Born approximation based on projection, " Microwave and Optical Technology Lett., vol. 5, pp. 184-188, April 1992.
- [10] G. T. Herman, Imaging reconstruction from projections, Academic, New York, 1980.
- [11] L. A. Shepp and B. F. Logan, " The Fourier reconstruction of a head section, " IEEE Trans. Nucl. Sci., vol. NS-21, pp. 21-43, 1974.
- [12] W. Tabbara, " Reconstruction of permittivity profiles from a spectral analysis of the reflection coefficient, " IEEE Trans. Antennas Propagat., vol. AP-17, pp. 241-244, Mar. 1979.

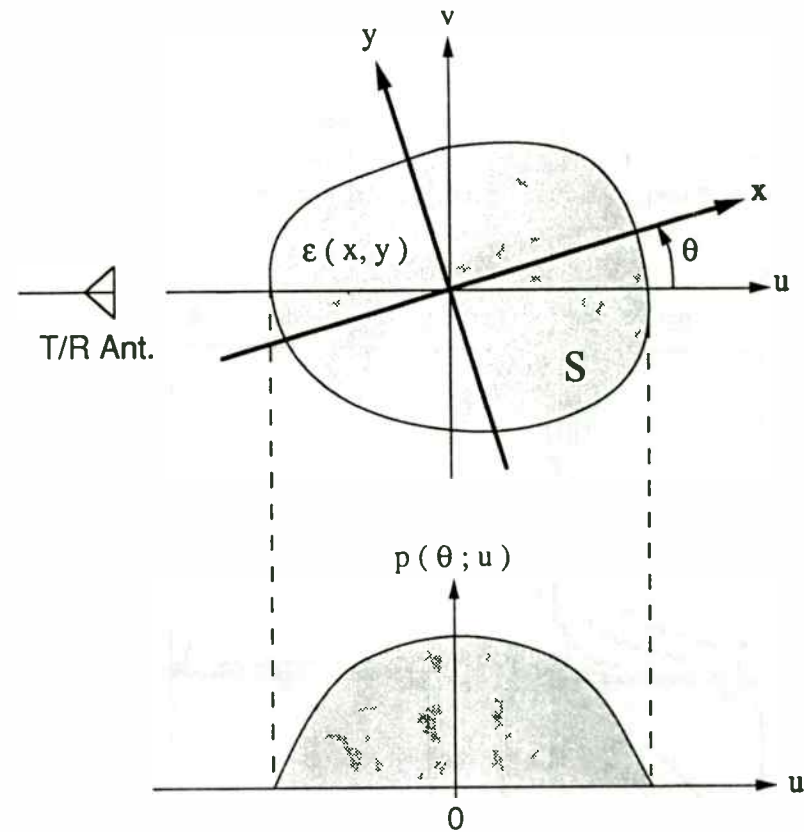


Fig. 1 The object geometry and its projection for microwave imaging.

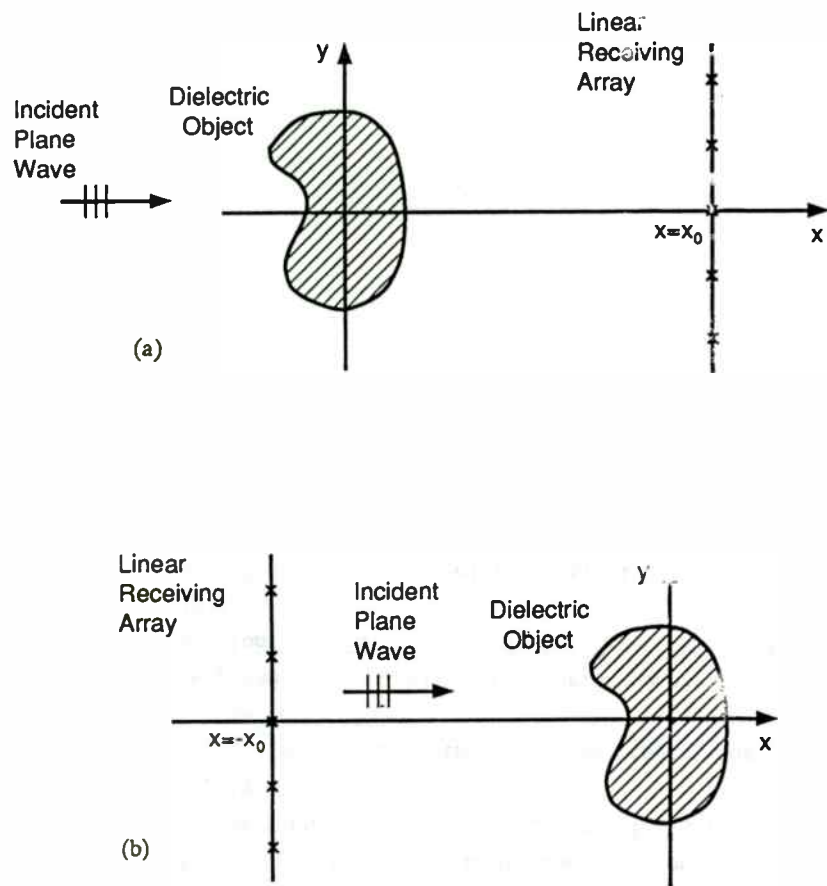


Fig. 2 The imaging geometry for microwave diffraction tomography: a) forward line measurement, b) backward line measurement.

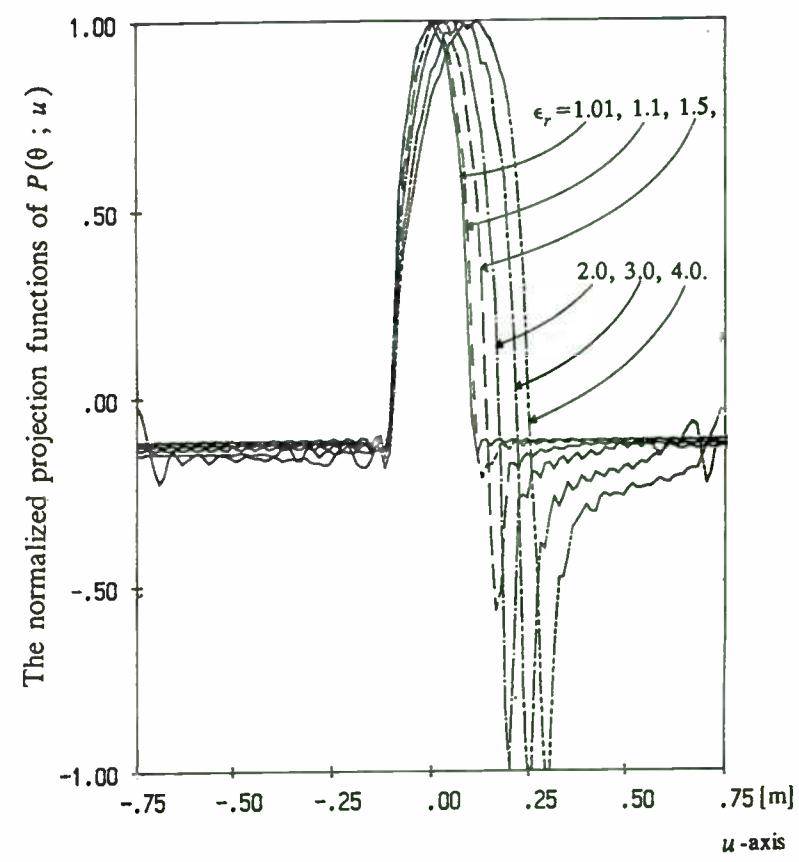


Fig. 3 The normalized projection functions of homogeneous dielectric cylinder of  $\epsilon_r = 1.01, 1.1, 1.5, 2.0, 3.0,$  and  $4.0$  for  $a = 10\text{cm}$  and  $f = 0.1 - 4.0$  GHz.



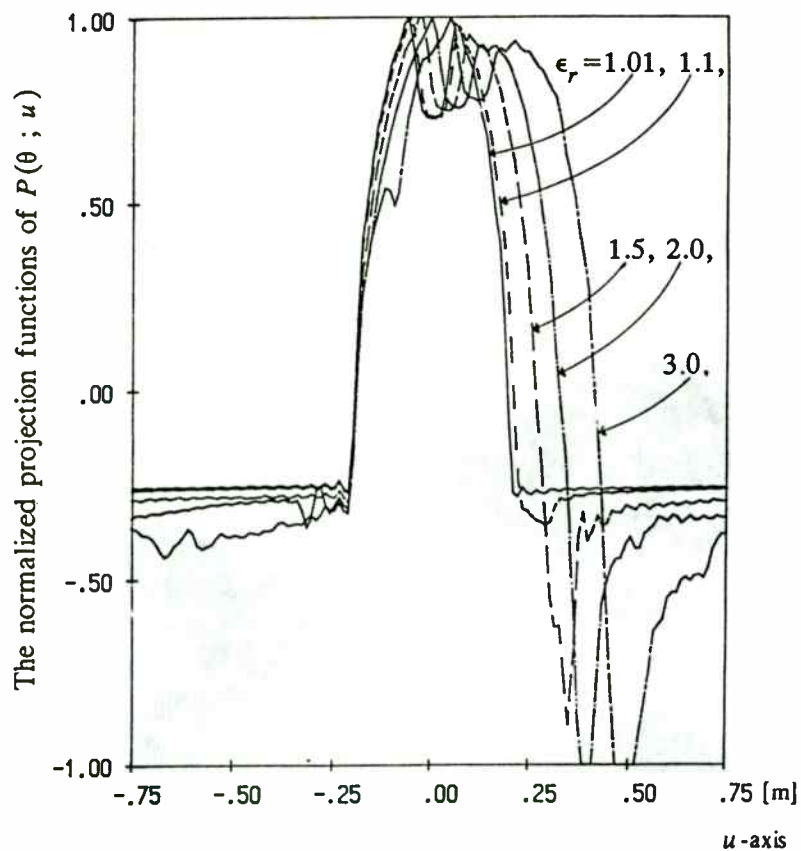


Fig. 4 The same as Fig. 3, but the hollowed dielectric cylinder of  $\epsilon_r = 1.01, 1.1, 1.5, 2.0,$  and  $3.0$  for inner radius 5cm, outer radius 20cm, and  $f=0.1 - 4.0$  GHz.

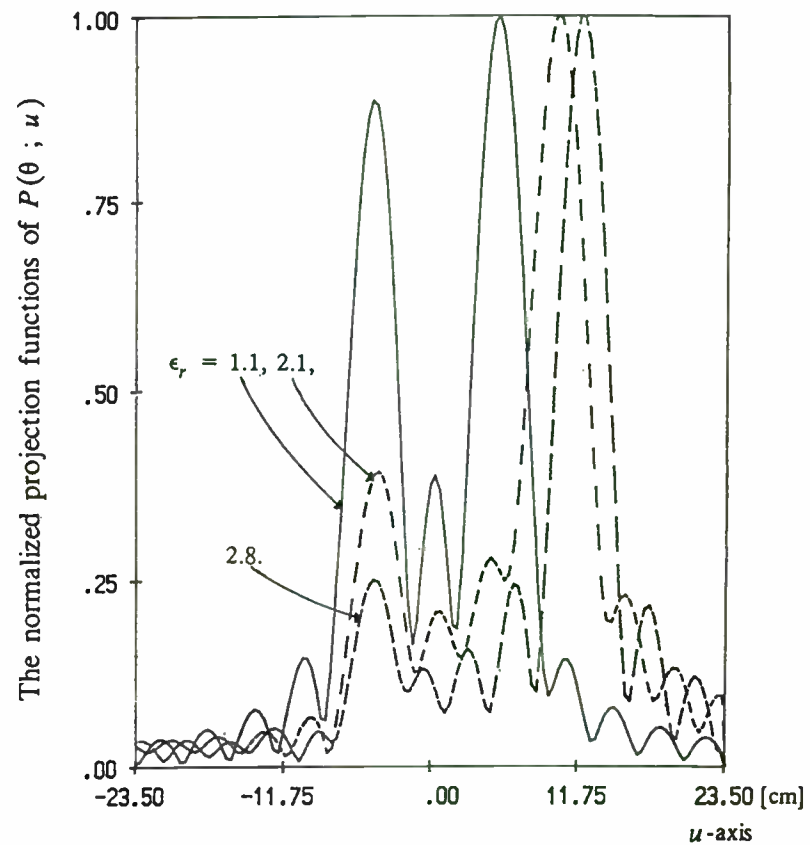


Fig. 5 The same as Fig. 3, but  $\epsilon_r = 1.1, 2.1,$  and  $2.8$  for  $a=5$ cm,  $f=7.5 - 11.5$  GHz.

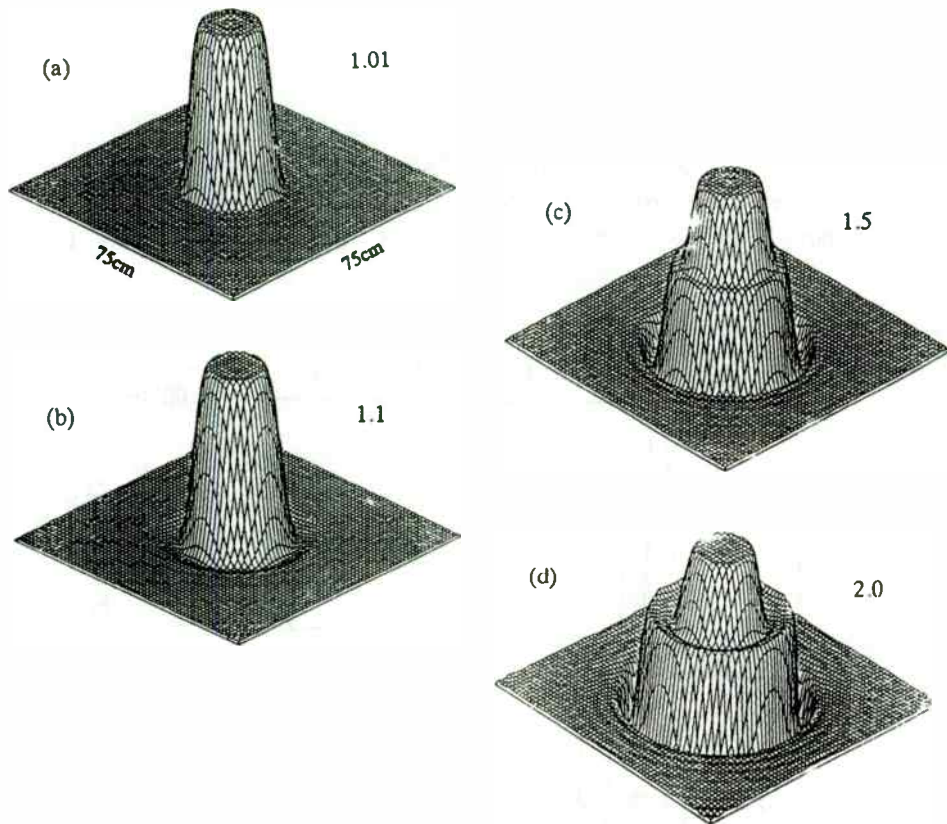


Fig. 6 The simulated results of reconstructed images from Fig. 3 for: (a)  $\epsilon_r = 1.01$ , (b)  $\epsilon_r = 1.1$ , (c)  $\epsilon_r = 1.5$ , (d)  $\epsilon_r = 2.0$ .

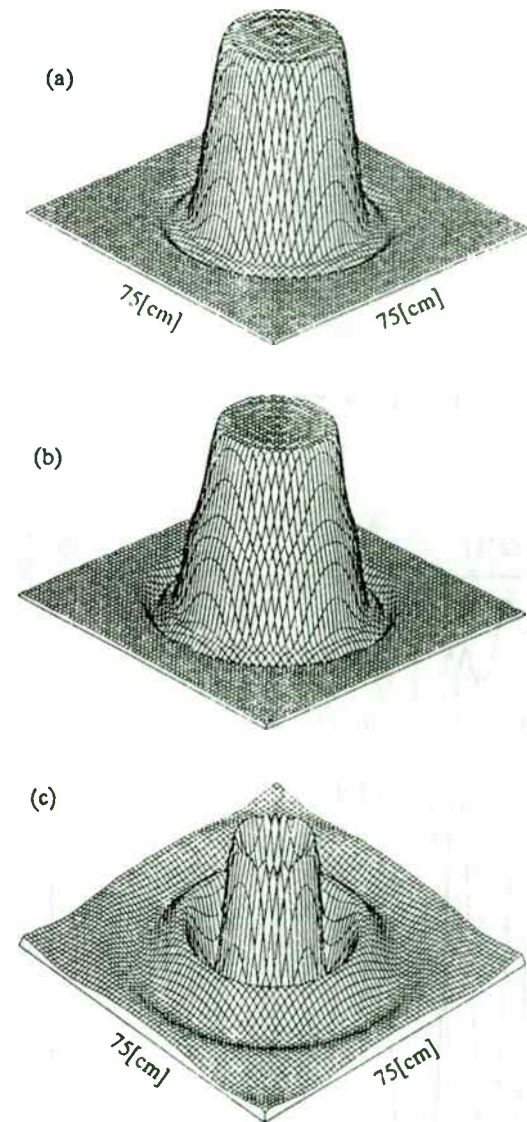


Fig. 7 The same as Fig. 6, but  $a=15\text{cm}$  and  $f=0.1 - 4.0\text{ GHz}$  for: (a)  $\epsilon_r = 1.1$ , (b)  $\epsilon_r = 1.2$ , (c) same as (a), but  $a=10\text{cm}$  and  $f=1.1 - 5.0\text{ GHz}$ .

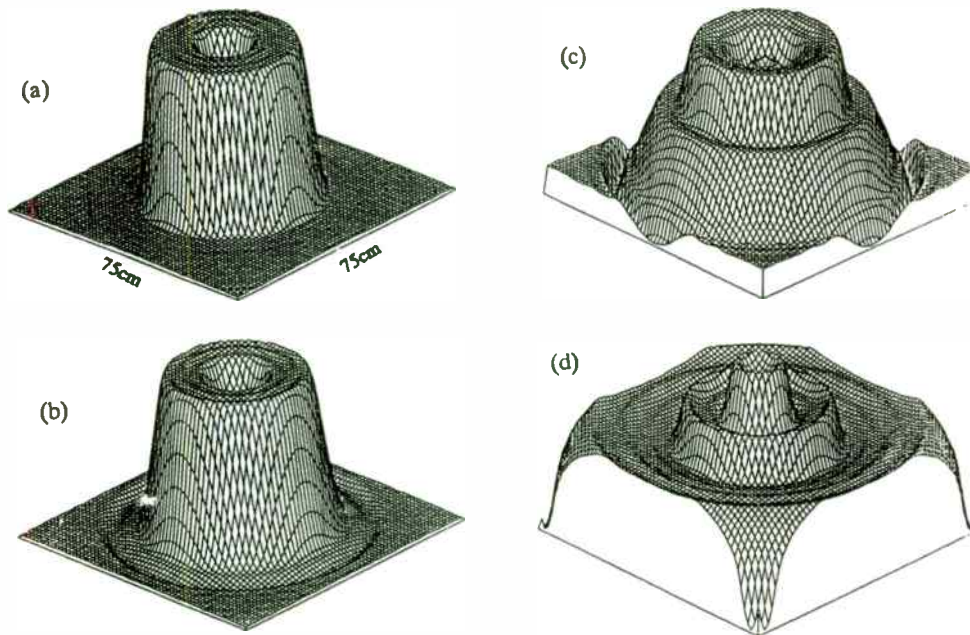


Fig. 8 The simulated results of reconstructed images from Fig. 4 for: (a)  $\epsilon_r = 1.01$ , (b)  $\epsilon_r = 1.1$ , (c)  $\epsilon_r = 1.5$ , (d)  $\epsilon_r = 2.0$ .

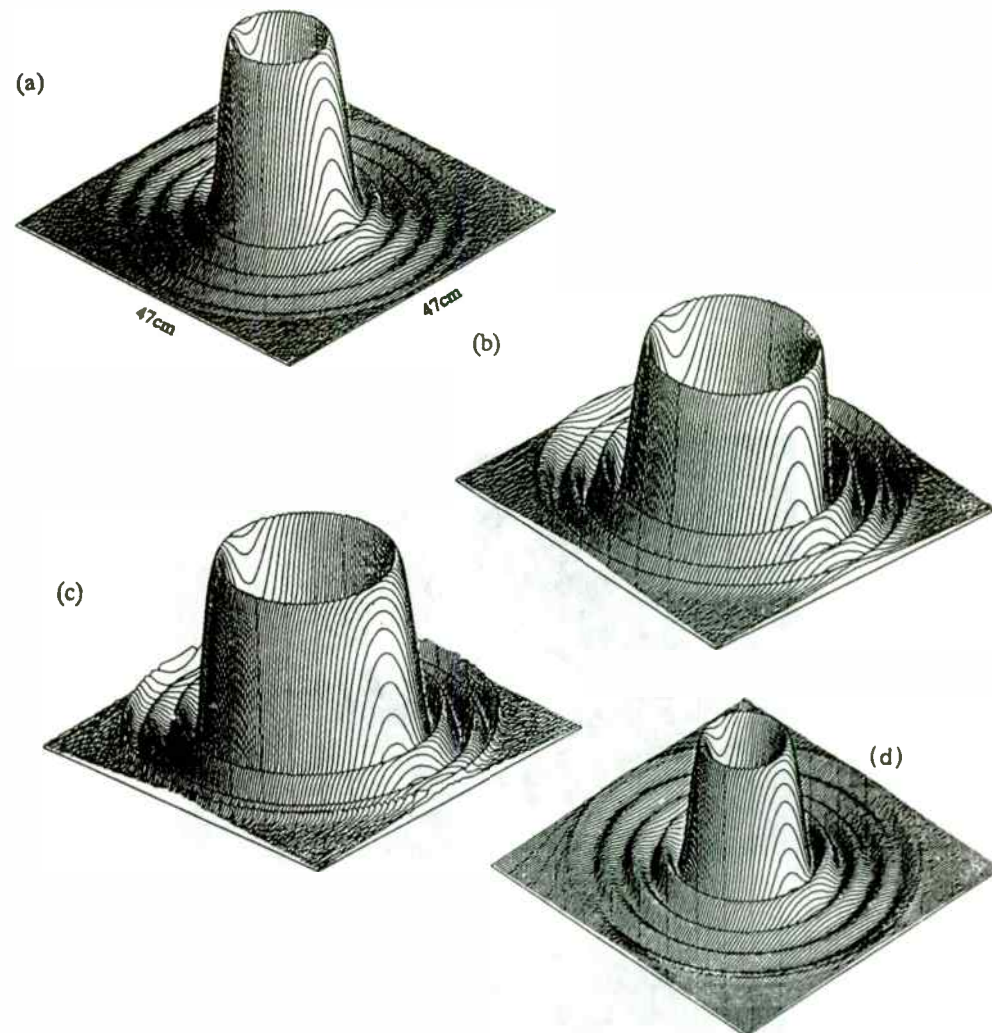


Fig. 9 The simulated results of reconstructed images from Fig. 5 for: (a)  $\epsilon_r = 1.1$ , (b)  $\epsilon_r = 2.1$ , (c)  $\epsilon_r = 2.8$ , (d) the same as (c), but a hollowed dielectric cylinder of inner radius 5.0cm, thickness 0.2cm.

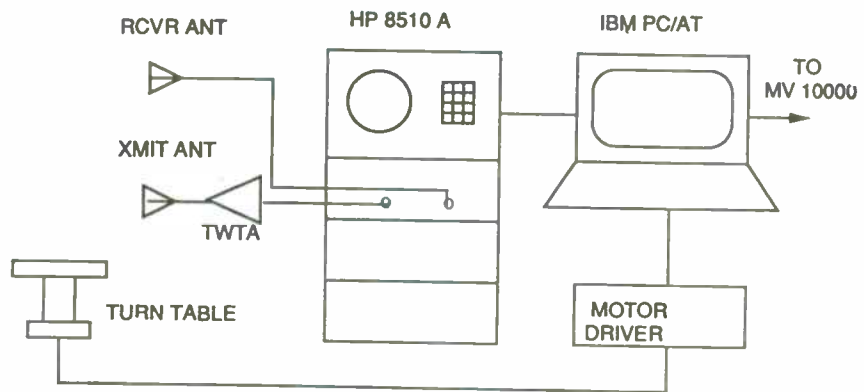
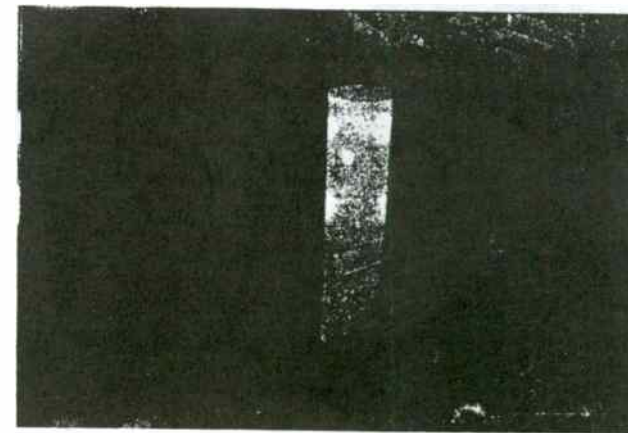
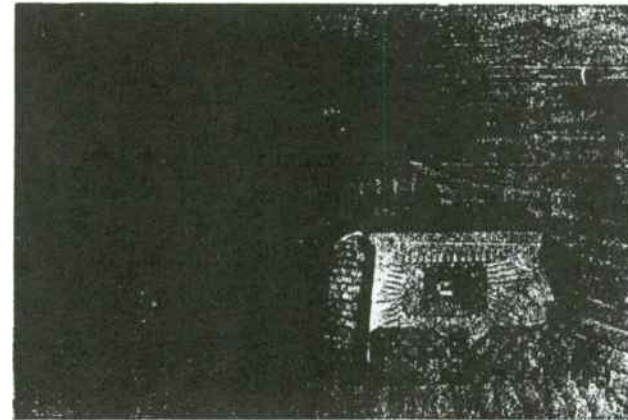


Fig. 10 (a) The blockdiagram of experimental set-up for data acquisition,



(b) The photograph for the transmitting, receiving antennas and turntable including object in the anechoic chamber.

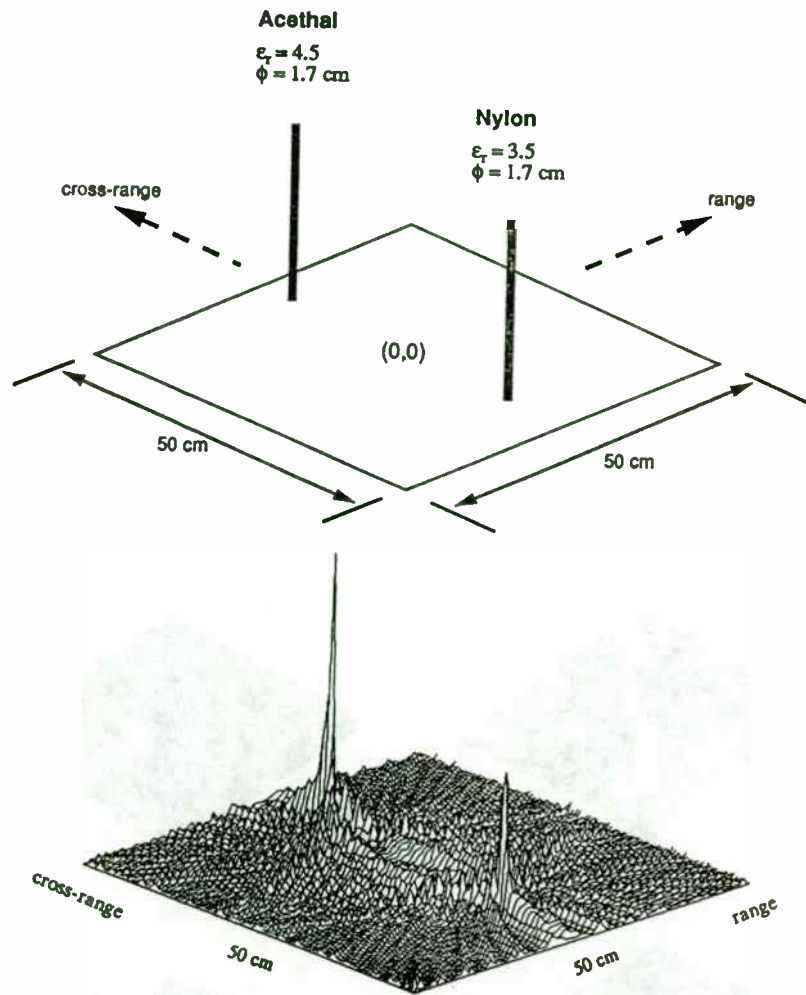
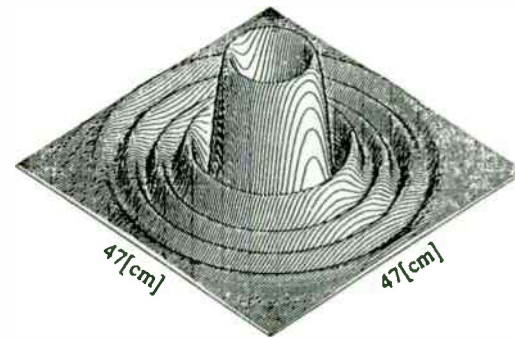


Fig. 11 (a) The intensity of retrieved image for two line objects from a single frequency of 10 GHz.



(b) The retrieved image from experiment for Fig. 9(d).

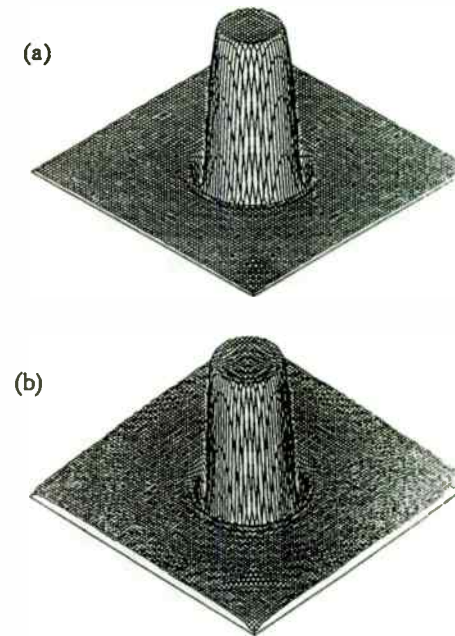


Fig. 12 The improved images of homogeneous dielectric cylinder for: (a) Fig. 6(c), (b) Fig. 6(d).

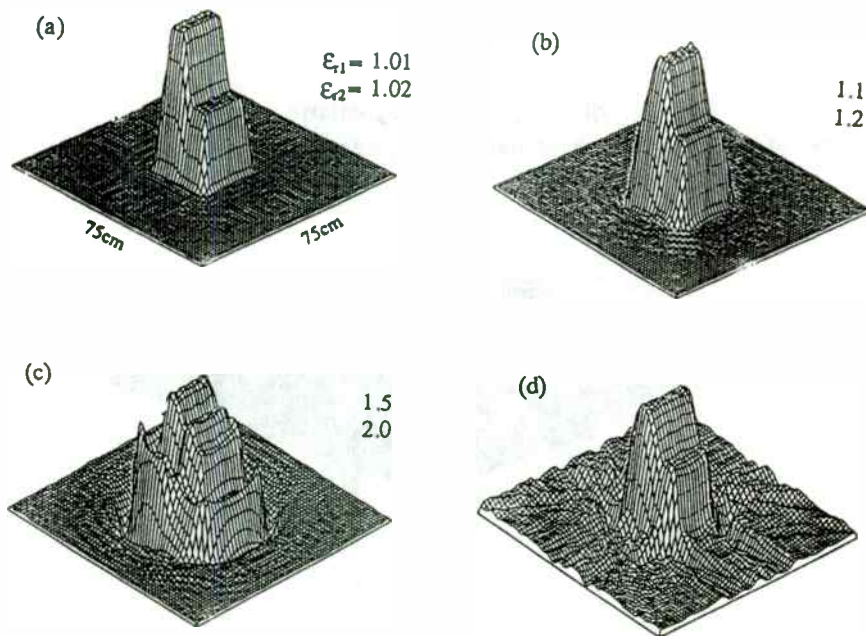


Fig. 13 The reconstructed images of inhomogeneous square dielectric cylinder for: (a)  $\epsilon_{r1} = 1.01$  and  $\epsilon_{r2} = 1.02$ , (b)  $\epsilon_{r1} = 1.1$  and  $\epsilon_{r2} = 1.2$ , (c)  $\epsilon_{r1} = 1.5$  and  $\epsilon_{r2} = 2.0$ , (d) the improved image for (c).

CRYSTAL FILTERS HAVING SUPERIOR INTERMODULATION  
CHARACTERISTICS AND THE MEASUREMENT OF LOW LEVEL INTERMODULATION  
IN A SWEEP FREQUENCY MODE

M.D. Howard and W.C.F. van den Akker  
Piezo Technology Inc.  
Orlando, Florida

Introduction

Many modern receivers and some radar systems in particular require crystal filters having levels of intermodulation performance ( I.M.) that exceed that of previously available product. PTI has successfully developed and is manufacturing product for such applications. The required I.M. performance is such that it stretches the capability of existing measurement systems. One of the factors contributing to the success of this program was the development of an I.M. measurement system capable of detecting I.M. products for outof-band signals below -140 dBm while retaining the dynamic range needed to measure in-band signals. The system also has the capability to sweep the incident test tones and provide plots of the generated I.M. product. This paper describes the design and evolution of a product that satisfies stringent I.M. requirements. Results on the filter performance are presented, with some interesting observations on the effects of spurious responses. The concepts underlying the measurement system will be presented and discussed.

Some Comments On  $IP_3$

When comparing the intermodulation performance of devices the comparison is often confused by differences in the applied power levels. Devices operating at higher power levels will often have worse intermodulation ratios than their counterparts operating at lower power levels. If account is taken of the increased drive level the device operating at the higher power level may in actual fact be offering the more linear performance. For this reason it is often convenient to use a normalized measure of intermodulation performance. One such measure is the third order intercept point or  $IP_3$ .  $IP_3$  is defined as the calculated output power level at which the generated third order intermodulation products are equal in amplitude to the output level of the test tones. The use of  $IP_3$  is not without pitfalls. It must be understood that the  $IP_3$  point is hypothetical and is obtained by extrapolation assuming a cubic nonlinearity. It is the point at which the device would have an intercept point if, as power is increased, the device continued to behave in the same simple, rational third order manner as it does at the power level at which actual measurements are being made. If a device is behaving in a simple third order manner then a 1 dB change in input power creates a 3 dB increase in the generated intermodulation products (a 2 dB reduction in the intermodulation ratio). Often, especially with crystal filters, there may be two or more non-linear mechanisms active at the same time. In this situation, as the phasing of the mechanisms changes, we may have addition or cancellation of third order products. The device behavior is not simple and often appears irrational as IM products

may even decrease over certain ranges of increasing power (or vice versa). Clearly, in this situation any projection of the  $IP_3$  is meaningless. Also, as power levels and frequency ranges change, different non-linear mechanisms may come into play particularly when testing filters both in-band and outof-band. For an  $IP_3$  number to have any meaning then, the test conditions, particularly the power levels and in-band or outof-band test, must be specified. Since different mechanisms come into play at different power levels, for example non-linear coupled modes in crystal filters, direct measurement of  $IP_3$  is meaningless and the  $IP_3$  measure of performance must remain a mathematically derived value (equation. 1) based on measurements taken at power levels and frequencies of the actual application.

$$IP_3 = P_{TTout} - IM_r/2 \quad (\text{dBm}) \quad (1)$$

where :

$P_{TTout}$  = Test Tone Power output (dBm)

$IM_r$  = Measured intermodulation ratio (dB)

#### Comparing Prior Art to Current Requirements

The prior art for the intermodulation performance of crystal filters was established to meet the requirements of high-performance mobile radio receivers in which good outof-band performance is necessary. Receivers are often required to detect a weak, desired in-band signal in the presence of a clutter of strong

outof-band signals. Other performance standards have been set by the requirements of military VHF communications systems. More recently Pti has produced filters with superior intermodulation (SIM filters) for an "Over the Horizon" radar system. The returned radar signals are of especially low amplitude while local interfering signals may be of substantial amplitude and unknown relative frequency. If such a system is to successfully detect a desired but weak signal while rejecting strong interfering signals then the system (and hence system components) must behave in a highly linear manner or third order intermodulation products generated from the strong interfering signals may prevent the detection of the desired signals. The SIM filter has IM performance that substantially exceeds the prior state of the art. The requirement also created new demands on measurement systems. The outof-band IM test is to be performed over a range of frequencies (not just at fixed test tone frequencies) and the in-band IM level specification is -137 dBm, a signal that is equal to or below the sensitivity of most spectrum analyzers.

The  $IP_3$  concept is used in table 1 to compare the specifications of the SIM filter to prior product specifications.

The requirement for the outof-band IM performance of the SIM filter is seen to be 20.5 dBm in excess of previously accepted standards while, as discussed above, the power level for in-band measurement is seen to be -137 dBm.



	Frequency (MHz)	BW <sub>3</sub> (kHz)	Type	P <sub>OUT</sub> dBm	IM <sub>r</sub> dB	IM dBm	IP <sub>3</sub> dBm
Mobile	21.4	14	outof-band	-29	-86	-115	+14
Mobile	21.4	14	outof-band	-27	-93	-120	+19.5
VHF. Com	98	27	in-band	-12.5	-60	-72.5	+18.5
SIM Filter	50	8	outof-band	-12	-105	-117	+40
		8	in-band	-47	-90	-137	-2

Table 1

Design Concepts

Prior work on the intermodulation performance of crystal filters [references 1 and 2], points out that there are two principal types of non-linearity associated with crystal filters, namely low level effects and high level effects. Low level effects are caused by surface finish, electrode adhesion and surface cleanliness. Normally, outof-band intermodulation is a low level effect. Signal levels are most often low, in the region of -25 dBm, and most of the incident signal is reflected by the filter, so that high level non-linearities are not of concern.

The principal causes of high level effects are the anelastic properties of quartz; that is, the material fails to deform in a linear manner with increasing applied force. The nonlinearity is quite small but not insignificant in terms of modern-day requirements. The anelastic effect is subject to analysis; analytical predictions of in-band intermodulation were first presented in reference [3]. In the case of signals within the filter passband the filter is required to pass much of the incident power which is, most often, stronger than the incident outof-band

signals. The in-band intermodulation ratio can be calculated using equation (2) [references 3 and 4].

$$IM_r = 130 + 20 \log_{10} ( A_e \cdot N / P_{IT} \cdot (BW_3/Fon)^2 ) \text{ dB} \quad (2)$$

where :

IM<sub>r</sub> is the calculated intermodulation ratio

A<sub>e</sub> is the electrode area in mils<sup>2</sup>

N is the overtone of operation

P<sub>IT</sub> is the incident test tone power in mW

BW<sub>3</sub> is the 3 dB design bandwidth in Hz

Fon is the nominal center frequency in Hz

This equation has been derived for one and two pole Butterworth filters for two test tones close to center frequency. However, it has proved valuable in predicting the comparative performance of filters with more than two poles and with transfer functions that are not Butterworth.

When first evaluating the requirements for the SIM filter it is evident that the in-band requirement is not of a high enough power level for the anelastic effects of quartz, as predicted by equation 2, to be of concern.

A third overtone monolithic designed for a 3 dB bandwidth of 8 KHz, having electrodes 60 mils by 60 mils and an input test tone power of -40 dBm would be expected to have an intermodulation ratio of :

$$IM_r = 130 + 20 \cdot \log_{10} ( 3 \cdot 3600 / 0.0001 \cdot (8 \cdot 10^3 / 50.006 \cdot 10^6)^2 )$$

$$= 138.8 \text{ dB}$$

This ratio is well in excess of the required 90 dB ratio. A fallacious conclusion would be that, given proper resonator processing to achieve low level outof-band performance, a filter constructed as a six pole monolithic from resonators of the type described above would satisfy the requirements . The schematic of such a filter is shown in figure 1. Results on a filter of this type show in-band intermodulation products well within specification limits but difficult to measure. Typical outof-band intermodulation results, although excellent in terms of prior art, were marginal for this requirement typically running at 105 to 107 dB, very close to the specified ratio of 105 dB. Extensive experimentation with processes yielded no consistent improvement in performance.

A method of calculating the IM products generated by a non-linear network performance is presented in reference [4]. An equivalent circuit that includes current-controlled voltage sources at the frequency of IM product is used as a model. Phasor equation (3) relates the voltage at the IM frequency,  $V_{IM}$  to the currents  $I_1$  and  $I_2$  flowing at the test tone frequencies  $F_1$  and  $F_2$ ,

$$V_{IM} = j \cdot \alpha \cdot I_1^2 \cdot I_2^* , \quad (3)$$

where the asterisk denotes complex conjugate.

The constant  $\alpha$  is dependent upon resonator properties. The analytical relationship between resonator properties and the constant  $\alpha$  can be derived from reference [3]. An alternate is to establish a value for  $\alpha$  empirically. The IM performance of actual devices is measured (preferably well-behaved devices, having good performance). A value for alpha is then chosen such that analysis

of the IM model described above reproduces the measured IM values. The established value of  $\alpha$  can then be used to predict IM performance at other drive levels and frequencies. This method has limited use as it requires a filter section incorporating the appropriate resonators to be measured before prediction of performance can be made. A more useful variation of this approach is to use equation (2) to predict IM performance close to filter center frequency and then establish a value for  $\alpha$  such that the network analysis gives the same result. The network analysis can then be extended to other frequencies. In this manner the outof-band IM performance resulting from the anelastic properties of quartz can be predicted.

The results of such an analysis on a two-pole as used above are shown in figure 2. At the frequency of interest ( +12 kHz ) the test tones are significantly attenuated by the two-pole. Consequently filter components other than the first two-pole section should not offer any noticeable contribution to the IM performance. As can be seen, the projected IM performance is in the region of 108 dB. This result agrees closely with the measured results on filters the best of which were in the range 105 to 108 dB.

This analysis shows that, for this requirement, the anelastic properties of quartz are contributing to outof-band IM performance and that improvements should be gained by improving the power handling capabilities of the resonators. Equation (2) gives insight to those resonator parameters that control IM performance. When designing monolithic resonators the acoustic coupling requirements of the two resonators impose restrictions on resonator design

parameters that are not consistent with maximizing IM performance. Consequently, the input section of the filter was redesigned to use discrete resonators. The remaining two filter sections are protected from the high level outof-band signals and in the interest of economy are left as monolithic devices. The circuit schematic of this filter is given in figure 3. Figure 2 includes projections of the IM performance of this section and compares it with the original monolithic two-pole performance. As can be seen the new circuit offers  $\approx 5$  dB improvement. In practice this improvement is surpassed. We hypothesize that this is because the two-pole discrete section has finite transmission zeros and attenuates at a faster rate than the model, which has all transmission zeros at infinity. Typical results for outof-band IM for the new filter are in the region of 117 dB compared to the projected 113 dB.

#### Results on the Improved Filter

The outof-band performance of a typical unit is given in the swept frequency plot of figure 4. The frequency shown is that of the test tone closest to the filter center frequency,  $F_1$ . The intermodulation product is measured at the center frequency of the filter. As can be seen the filter readily satisfies the specified requirement of 105 dB for IM products over the measured range.

Of interest are the returns measured at 50.053 and 50.065 MHz. These returns are coincident with spurious responses. As part of the resonator design described above, it was found that the design must be simultaneously optimized for both spurious performance and IM performance. An improper resonator design would result in a strongly excited spurious response and an associated IM return. As

result of design optimization of the input section, spurious responses and IM returns are of a low level and correlation is difficult to illustrate. However, the output sections are not optimized in this manner. Correlation is more readily illustrated by a studying the IM and spurious responses generated in the output section of the filter. Figure 5 illustrates the outof-band IM performance and spurious response performance of filter with the input and output reversed for the purpose of illustration. Clearly, when either of the two test tones is coincident with a spurious response, the IM products increase. As equation (3) indicates, the IM generated below the test tones is proportional to the  $F_1$  current squared and linearly to the  $F_2$  current. For high side IM products,  $F_2$  and  $F_1$  are interchanged. Some of the smaller IM returns are generated when  $F_2$  passes through a spurious response (located at twice the indicated frequency deviation) while the larger returns are associated with  $F_1$ . On reflection, it comes as no surprise that IM returns occur when either of the incident test tones is at the frequency of a spurious response. A high Q spurious response can be considered as a filter of very narrow bandwidth. Equation (2) tells us that IM products caused by the anelastic properties of quartz will increase as the inverse of the filter bandwidth squared. Hence, when test tones are coincident with a spurious response, the IM products can be expected to increase.

#### The Measurement System

These results would not have been observed without the capability to sweep the outof-band IM test over a range of frequencies. In fact all of the enhanced product performance

features would be, at best, difficult to attain without the capability to perform adequate measurements of IM products. A part of the development program for this product was the construction of such a measurement system. A block diagram of the system used by Pti is shown in figure 6 and in many respects resembles conventional IM measurement systems. The system is computer controlled which allows easy and repeatable set up. The output levels to the DUT and the gain of the switched, pre-detection amplifier are both calibrated and the calibration stored on disc. One of the novel features of this system is the use of switched notch filters on the output of the programmable synthesizers. The purpose of these filters is to suppress synthesizer noise at the frequency of measurement. In the outof-band IM measurement the test tone frequencies are adjusted so that the IM product is always at the center frequency of the DUT. A filter that suppresses noise at this frequency will enhance measurements. More commonly a bandpass filter centered at the test tone frequency would be used. In this instance, since the test tones are to be swept, a band pass filter that would pass the range of swept frequencies while simultaneously attenuating noise at the frequency to be measured is not feasible. The notch filter also suppresses any synthesizer spurious outputs that may become coincident with the test frequency. The test system has a noise floor of -145 dBm and a detector  $IP_3$  of +12 dBm. This measurement is referenced to the output of the DUT and is based on readings taken with test tone power levels of -55 dBm to -20 dBm. The system has adequate dynamic range and sensitivity to measure the in-band requirements. The required outof-band  $IP_3$  is + 40 dBm which exceeds the detector capability. However, the DUT is of

course a filter and offers the detector over 60 dB of protection from the test tones resulting in an apparent detector  $IP_3$  well in excess +40 dBm for the outof-band measurement. The system response with the DUT replaced with a short circuit is shown in figure 7. The source power is the power level of the two test tones at the input to the DUT. The flat portion of the response measured with a source output power in the range of -55 dBm to approximately -45 dBm represents the detector noise floor. The portion of the response for source output power from -30 dBm to -20 dBm shows a 3.0:1 increase in measured power at the IM frequency and represents IM products generated in the detector. The section of the response between -45 dBm and -30 dBm source output represents the transition between noise limitations and IM limitations of the detector, probably with the addition of some source noise (which should reduce at a 1:1 ratio with the source power).

#### Conclusions

Some insight into the design of filters having superior outof-band IM performance has been presented.

A relationship between outof-band IM and the anelastic properties of quartz has been established and shown to be a factor in the generation of IM products when the outof-band signals have moderately high power levels.

Swept outof-band IM data has been presented showing IM returns that have some correlation to the location of crystal spurious responses.

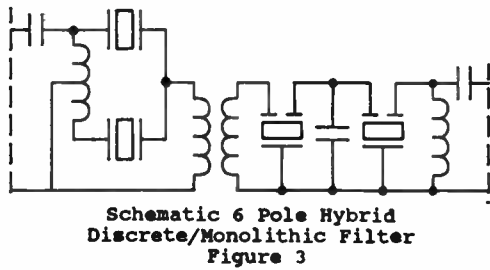
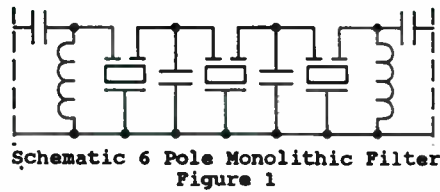
The use of narrowband crystal notch filters to clean up synthesizer output while making a swept IM measurement was presented.

#### Acknowledgement

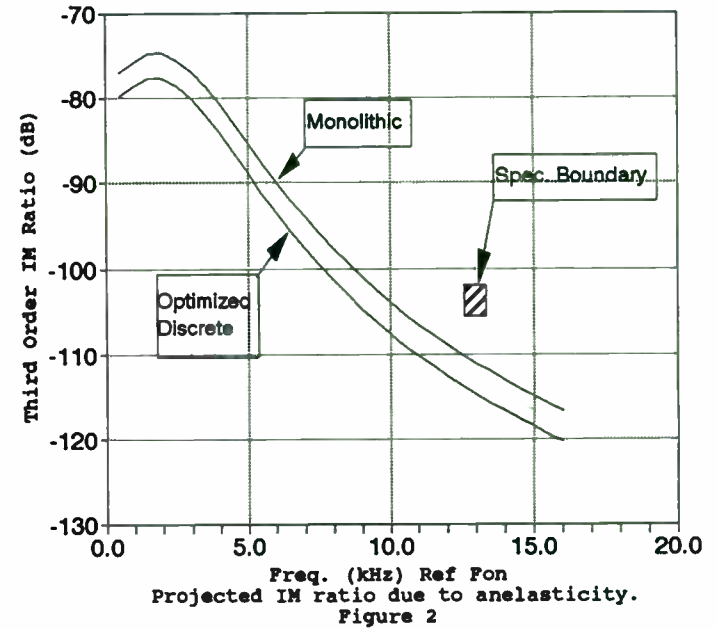
To Dr. W.H. Horton, R.C. Smythe and B.H. Wolfe for their contribution to the development of the IM test system.

#### References

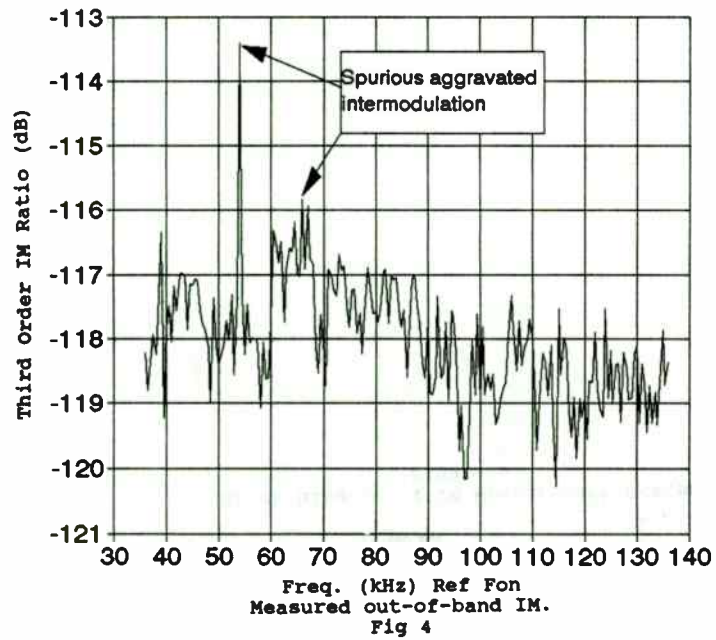
- [1] Horton, W.H. and R.C. Smythe, "Experimental Investigations of Intermodulation in Monolithic Crystal Filters," Proc. 27th Annual Symposium on Frequency Control, pp. 243-245; 1973.
- [2] Smythe, R.C., "Bulk Acoustic Wave Filters," in Precision Frequency Control, Vol. 1, (E.A. Gerber and A.D. Ballato, ed.s), pp. 216-219, Academic Press, New York; 1985.
- [3] Smythe, R.C., "Intermodulation in Thickness-Shear Resonators," Proc. 28th Annual Symposium on Frequency Control, pp. 5-7; 1974.
- [4] Howard, M.D., R.C. Smythe, and P.E. Morley, "Monolithic Crystal Filters Having Improved Intermodulation and Power Handling Capability," Proc. 39th Annual Symposium on Frequency Control, pp. 491-503; 1985.



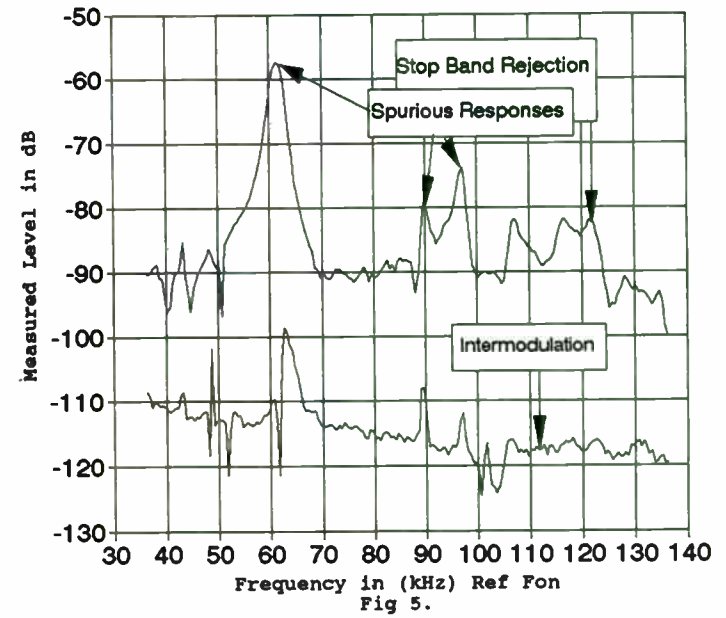
Piezo Technology Inc.  
Comparison of Discrete and Monolithic

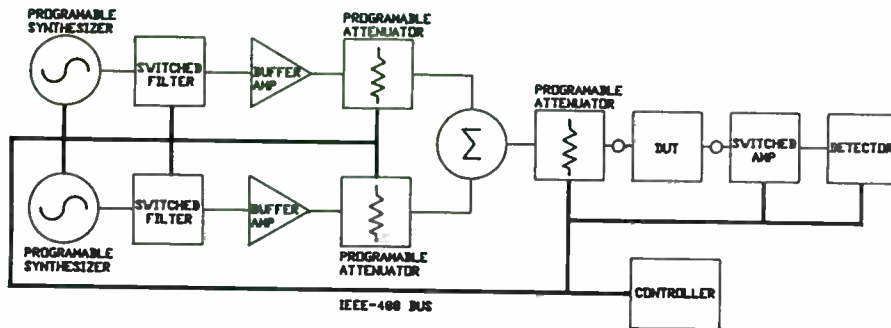


Piezo Technology Inc  
Filter Out of Band I.M



Piezo Technology Inc.  
Filter I.M and Spurious Responses





Block Diagram of IM Measurement System  
Figure 6

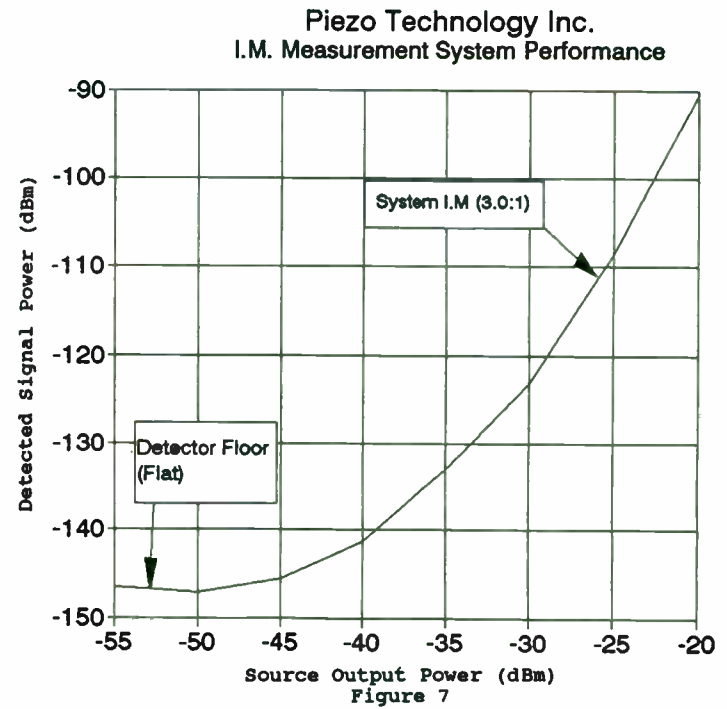


Figure 7



## EXTENSIONS OF THE HOLT & GRAY CRYSTAL FILTER DESIGN TECHNIQUE

-----

William B. Lurie, Consultant  
Boynton Beach, Florida

The Holt & Gray transformation was developed to make bandpass crystal filters as cascaded lattices from a low pass prototype such as an Elliptic filter. For that application it provided a practical way to design exact multi-section lattice filters, but without modification the element values tend to have a wide spread. A technique and program are described to narrow that spread to a practical minimum, with an example being furnished.

The same technique can be used (with some adaptation) to transform a high-pass filter into a band-reject filter having (for Elliptic prototypes) two equal-ripple passbands, and stopbands with equal returns. The rationale is presented, along with an example illustrating the input and the results, for a multi-section lattice filter with only two quite similar crystals per section

The inputs need not be Elliptic filter prototypes, and an illustration is presented of the use of a Rhodes low-pass filter as a prototype, containing negative elements, but yielding a practical filter with all positive values.

The original Holt & Gray transformation [1] was a method of designing bandpass crystal lattice filters as a cascade of two-crystal sections. As such it is rigorous, complete, and quite effective and useful. The basic article suggests a method of modifying the inverter so that the spread of crystal motional parameters can be minimized, and also the use of Dishal's transformation [2] to realize a restrict class of asymmetric (single-sideband) filters.

Briefly, the method consists of the following steps:

1. An even-degree lowpass ladder with two zeros at infinity is selected as a prototype (Figure 1).
2. The first section of the filter is converted to its lattice equivalent, which includes a narrow-band impedance inverter (Figure 2).
3. The conversion to lattice is repeated for each succeeding section. (Figure 3).
4. If an asymmetric response is desired, the bilinear transformation is performed (Figure 4) on each lattice branch.
5. Capacitors are now substituted to approximate the constant reactances in and near the passband, to convert the lowpass to a bandpass (Figure 5).
6. The impedance inverters are now varied up or down (see below) and then absorbed into the lattice arms.
7. The appropriate steps are iterated manually or automatically, to change the range of motional parameters in the finished filter. This involves primarily this step and the preceding step.

From an economic viewpoint, it is sometimes convenient to have one crystal in each section have some preselected motional capacitance value. If the impedance levels of two successive inverters are both multiplied by some constant  $K$ , then the impedance of the network between those inverters is multiplied by a factor  $K^2$ , but the overall response of the network is unaltered. This is the mechanism for varying the motional parameters of each section, one section at a time, to narrow their range.

As an illustration, an Elliptic filter (Cauer C 12 20 65 B) was used as a prototype, to design a symmetrical bandpass centered at 10 MHz with a bandwidth of 10 KHz. Before modifying the inverters, the range of motional capacitances of the crystals is about 11:1. After modification, the range becomes 5.5:1, as seen in Figures 6 and 7, where the impedance levels have been scaled so that the largest motional capacitance in each filter has become 10 fF. Note also that, in the final design, the capacitance value of 10 fF appears in seven out of twelve crystals.

Another distinct advantage can be realized from having the larger of the two motional capacitances in each section the same for all sections. If the design is converted to a form which can be built using a cascade of two-resonator monolithic structures, then all of the resonators in the filter have identical motional capacitances, but not set at identical frequencies, nor with the coupling between resonators the same from section to section. This would seem to imply that identical masks can not be used from section to section, even though the 'spot sizes' are the same, but this is not the case; a technique exists to make use of the fact that coupling between resonators on any one piece of quartz can be adjusted as a part of the plating process (but that is a subject for a separate article).

The Holt & Gray design of bandpass filters can accommodate for its prototype lowpass any even-degree pole-zero filter with the familiar configuration as originally described in the classic Saal and Ulbrich Elliptic filter article [3]. It need not be an Elliptic design, although that is probably the most conventional. For example, the constant-delay pole-zero set described by Feistel and Unbehauen [4]

can be used to provide a bandpass with constant delay response, as is illustrated in Figure 8. An interesting case is presented in Figure 9, which is a low-pass filter derived from a Rhodes non-minimum phase polynomial [5]. Even though this filter has a negative capacitor, inasmuch as the transfer function has zeroes on the real axis, the Holt & Gray transformation yields a network, Figure 10, with all positive values, and an interesting response curve, Figure 11.

Another application of the Holt & Gray transformation, not envisioned in the original paper, is of some interest. The lowpass ladder prototype was converted to a lowpass lattice, above, using inverters and constant reactance elements, and then a lowpass to bandpass conversion was applied. If the lowpass were converted to a highpass before that last conversion, however, then that highpass would become a band-reject. This is indeed the case, and the process is reasonably simple: elements which have a constant reactance remain a constant reactance, but all inductors are replaced by their dual, i.e., capacitors. Figures 12 and 13 illustrate these steps.

The process proceeds through a number of steps, parallel-resonating the capacitors at the reference reject frequency, and then replacing the constant-reactance elements by capacitors with that reactance at the reference frequency, and then the resulting branches are converted to their Foster equivalents in the familiar crystal form, and finally the inverters are absorbed into the lattices. Again, as with the bandpass filter, the inverters can be raised or lowered to make the motional capacitances of the crystals as near equal as possible.

With the band-reject case, however, some oddities appear. It takes a rather large value of the inverter multiplier to make the crystal values attractive, and, finally, the inverters can not be fully absorbed without leaving some negative values behind. For example, see Figure 14. A practical solution is to insert, at each node where a negative capacitance to ground is seen, a circuit in parallel with it, consisting of a coil and capacitor in parallel, resonating at the notch frequency. The positive capacitance must be at least large enough to swamp the negative capacitance, with enough left over to provide capacitance to tune the coil, including its distributed capacitance. Actually, the filter must have two real hybrid transformers, which are conveniently located at the input and output, and so the added requirement is for only one additional coil and capacitor. In practice, it is usual to have the inductances of the added coils as large as is practical, to minimize the disturbances they introduce.

The effect of these tuned circuits on filter performance in the notch and nearby passband is negligible, but obviously they will prevent the filter from behaving like a band-reject far from the notch, for example at very low and very high frequencies. In Figure 15 is shown the filter of Figure 14 with practical coils added, and Figures 16 and 17 illustrate its response. Note that the notch is about 10KHz wide, and the passbands extend roughly 1 MHz each side.

The assistance of several people in preparing this paper should be acknowledged: Mr. Mark Mell of Microsonics, who furnished valuable advice into the mechanization of the notch filter synthesis, and Mr. Dick Abrahams of Radio Systems, inc., in preparing the illustrations.

#### REFERENCES

- 
- [1] Holt, A.G.J. and Gray, R. L., IEEE Trans. Circuit Theory, CT-15, Dec. 1968, pp. 492-494
  - [2] Dishal, M., Proc. IEEE, vol. 53, pp. 1205-1216, Sept. 1965
  - [3] Saal & Ulbrich, IRE Trans. on Circuit Theory, CT-5, Dec. 1958, pp. 284-317
  - [4] Feistel, K. H. & Unbehauen, R., FREQUENZ, Bd. 19/1965 Nr. 8 (in German), Lowpasses with Tchebycheff Characteristic in the attenuation and maximally flat delay.
  - [5] Rhodes, J. D., IEEE Trans. Microwave Theory & Techniques, MTT-18 No. 6, June 1970, pp. 290-307

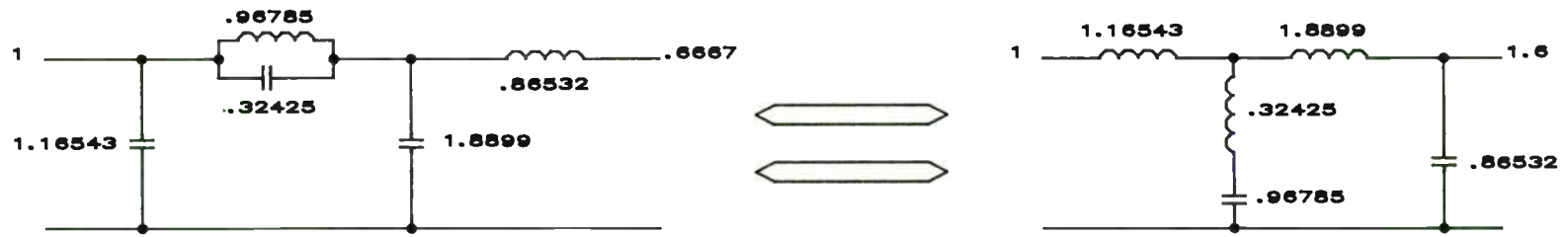


FIGURE 1

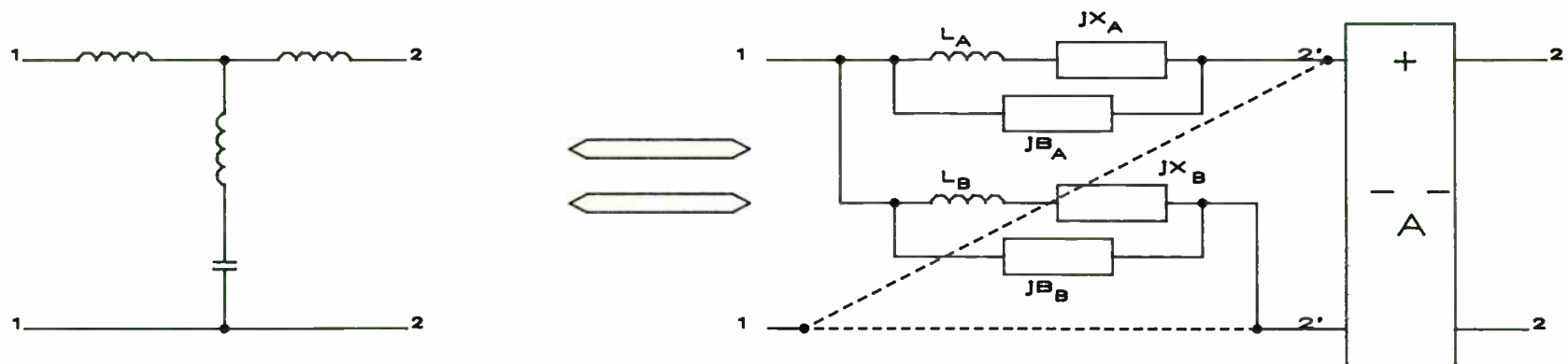


FIGURE 2

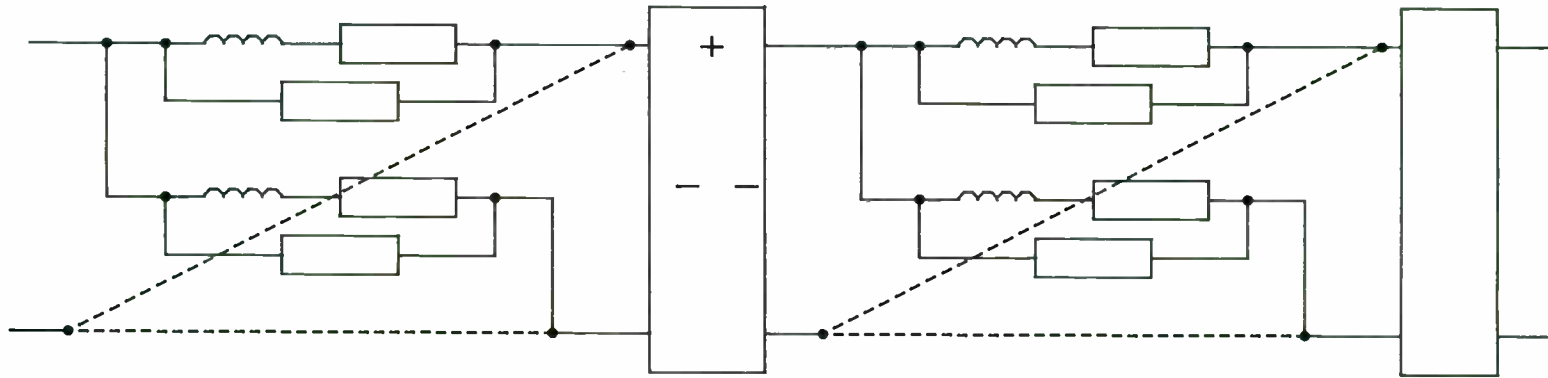


FIGURE 3



$$L = \frac{\omega_0^2}{\omega_0^2 - 1} L'$$

$$X = \frac{\omega_0}{\omega_0^2 - 1} L'$$

$$B = \frac{1}{\omega_0 L'}$$

$$\Omega \rightarrow \frac{\Omega, \omega_0 + 1}{\Omega, + \omega_0}$$

FIGURE 4

DISHAL (BILINEAR) TRANSFORMATION

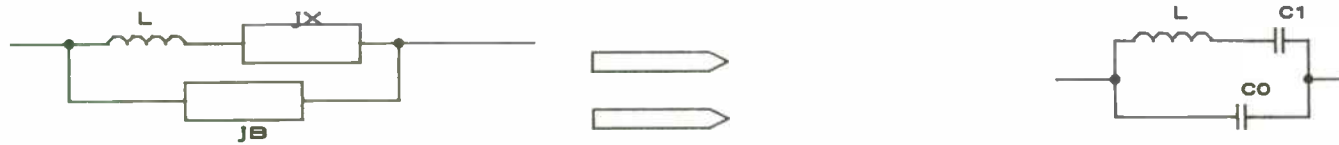
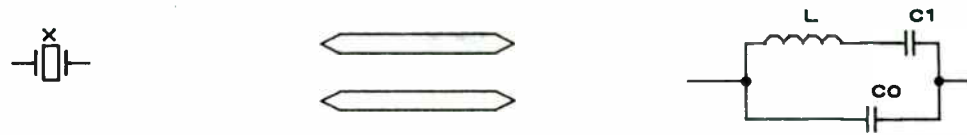
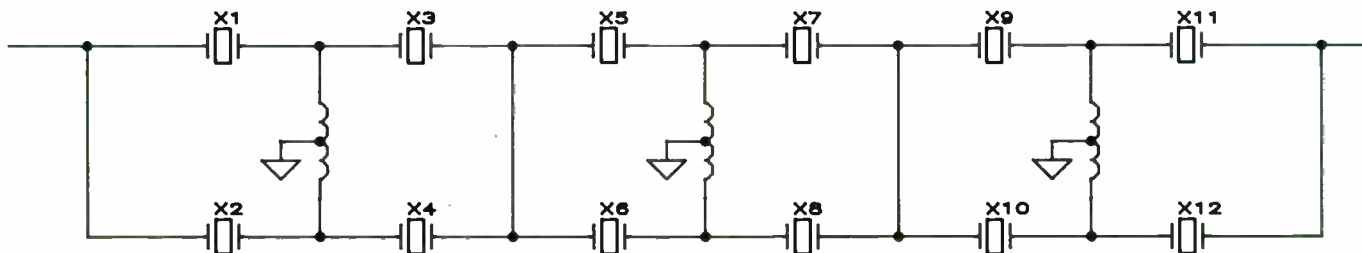


FIGURE 5  
LOWPASS TO NARROWBAND  
BANDPASS CONVERSION



TYPICAL CRYSTAL



CRYSTAL	MOTIONAL CAPACITANCE (fF)	CRYSTAL	MOTIONAL CAPACITANCE (fF)
X1	1.97	X2	1.97
X3	4.35	X4	1.83
X5	2.85	X6	.884
X7	10.0	X8	.967
X9	2.56	X10	1.04
X11	3.68	X12	2.29

FIGURE 6  
N=12 ELLIPTIC BANDPASS  
BEFORE ADJUSTING INVERTERS

CRYSTAL	MOTIONAL CAPACITANCE (fF)	CRYSTAL	MOTIONAL CAPACITANCE (fF)
X1	9.99	X2	10.0
X3	9.99	X4	3.96
X5	10.0	X6	2.10
X7	9.99	X8	1.81
X9	9.99	X10	2.67
X11	9.99	X12	5.55

FIGURE 7

N = 12 ELLIPTIC BANDPASS  
AFTER ADJUSTING INVERTERS

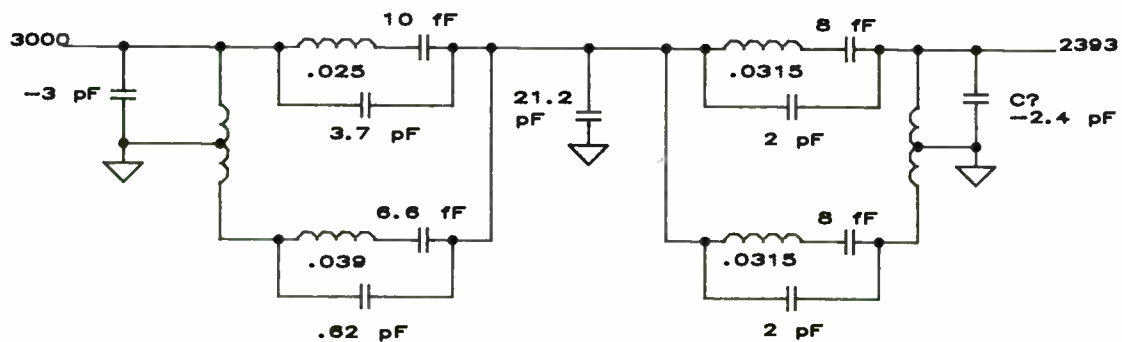


FIGURE 8

BANDPASS DERIVED FROM  
FEISTEL FLAT DELAY LOWPASS

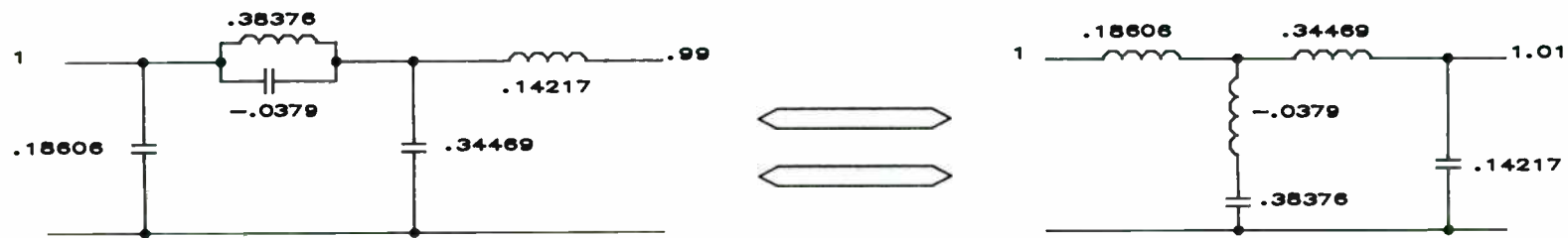


FIGURE 9

N=4 LOWPASS FROM  
RHODES POLYNOMIAL  
(NORMALIZED)

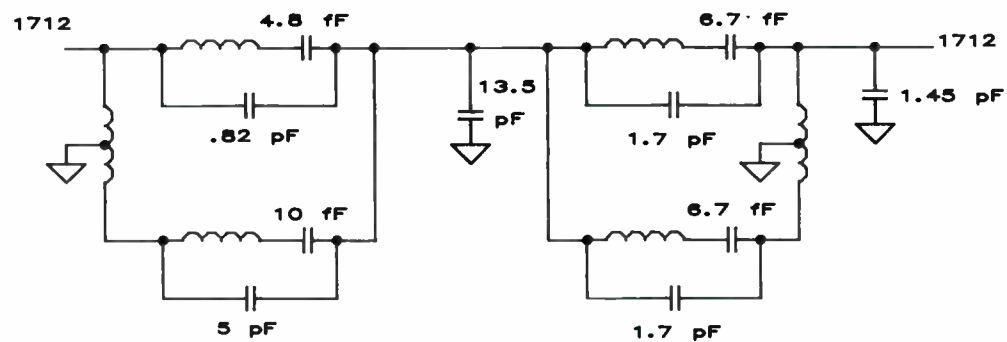
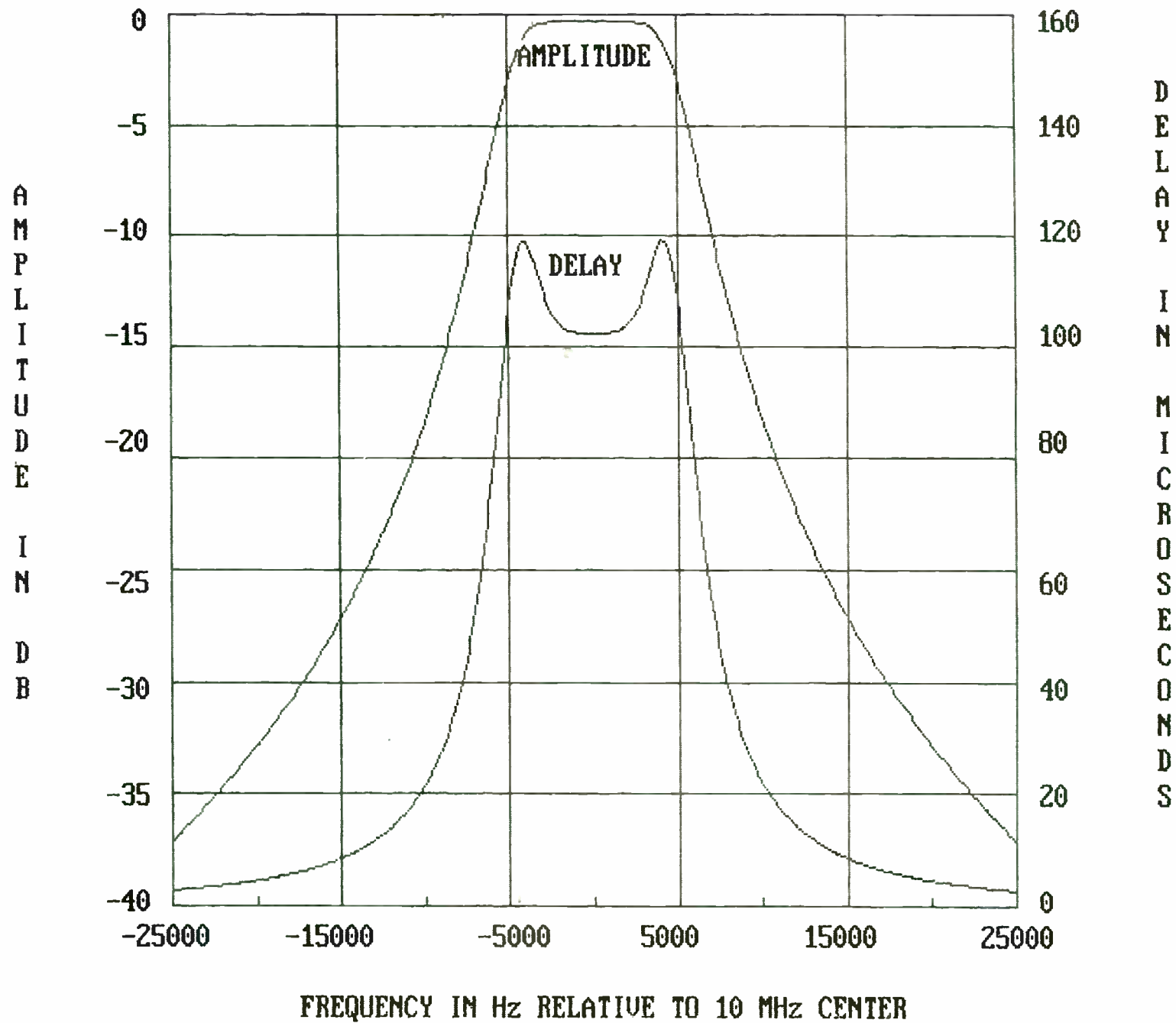


FIGURE 10

BANDPASS DERIVED FROM  
RHODES LOWPASS



FIGURE 11 -- RHODES BANDPASS, AMPLITUDE AND DELAY



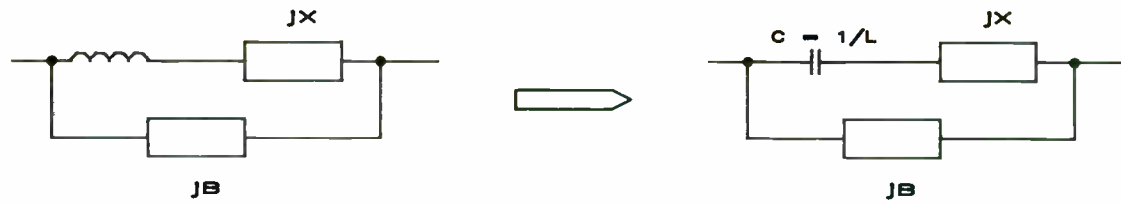


FIGURE 12

LOW PASS TO HIGH PASS  
CONVERSION

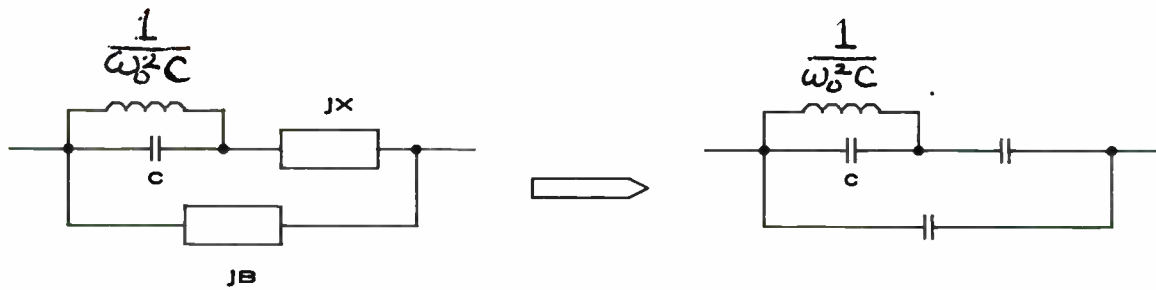


FIGURE 13

DERIVATION OF BAND REJECT FROM  
HIGH PASS OF FIGURE 12

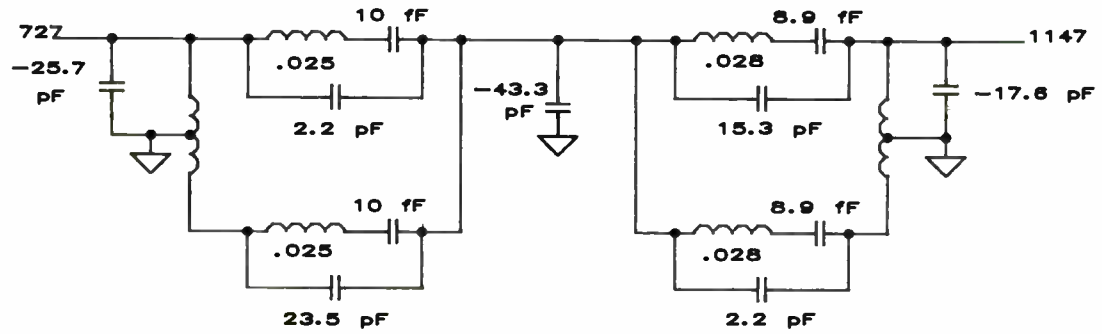


FIGURE 14

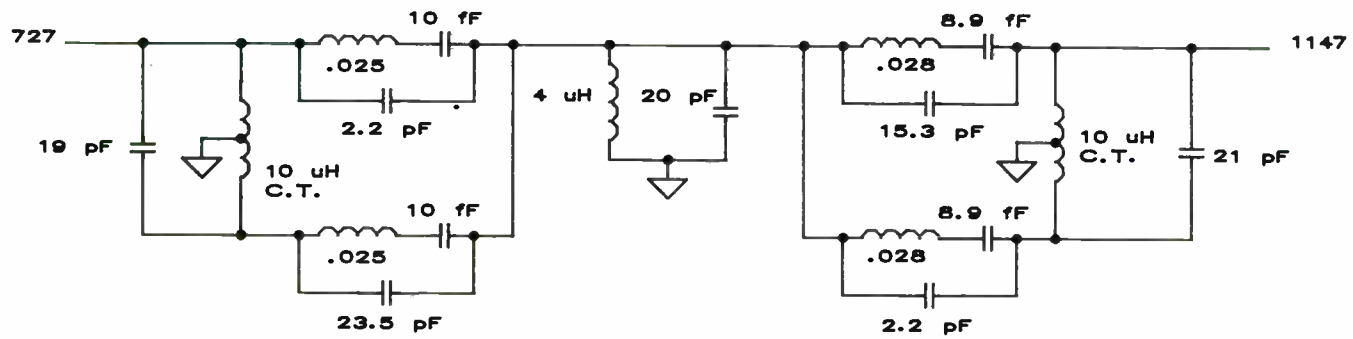
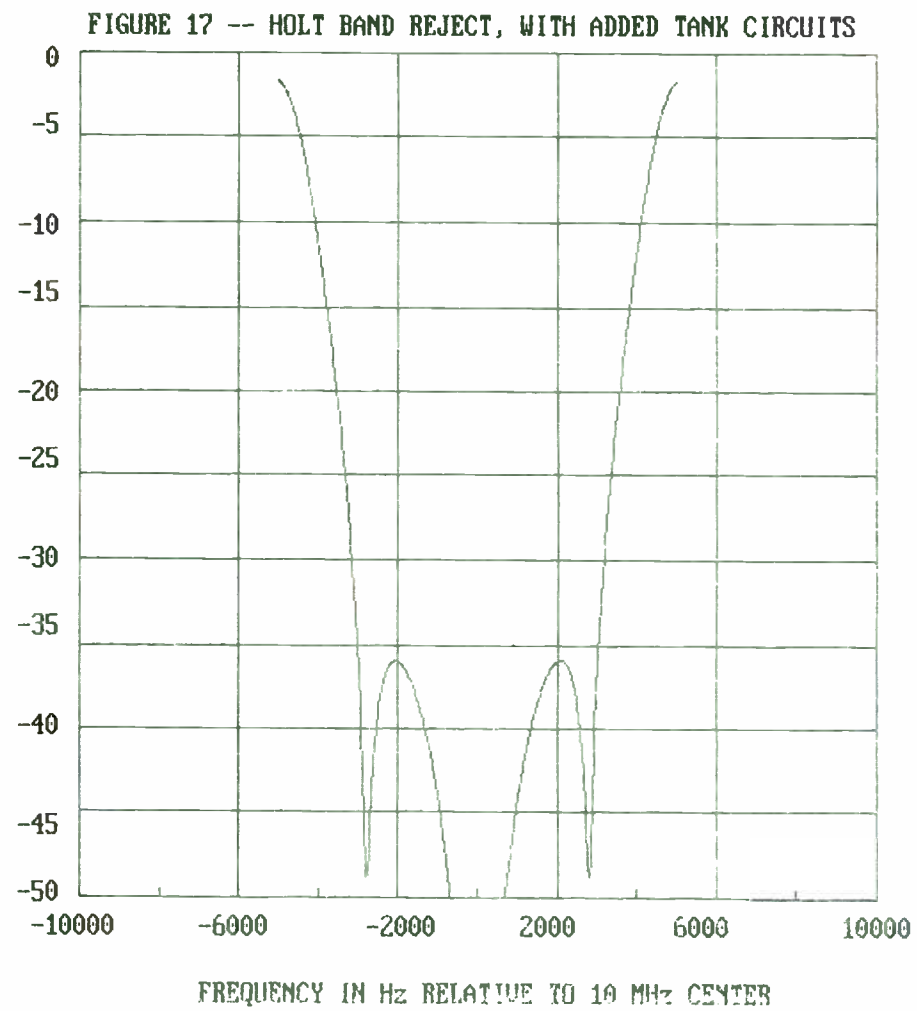
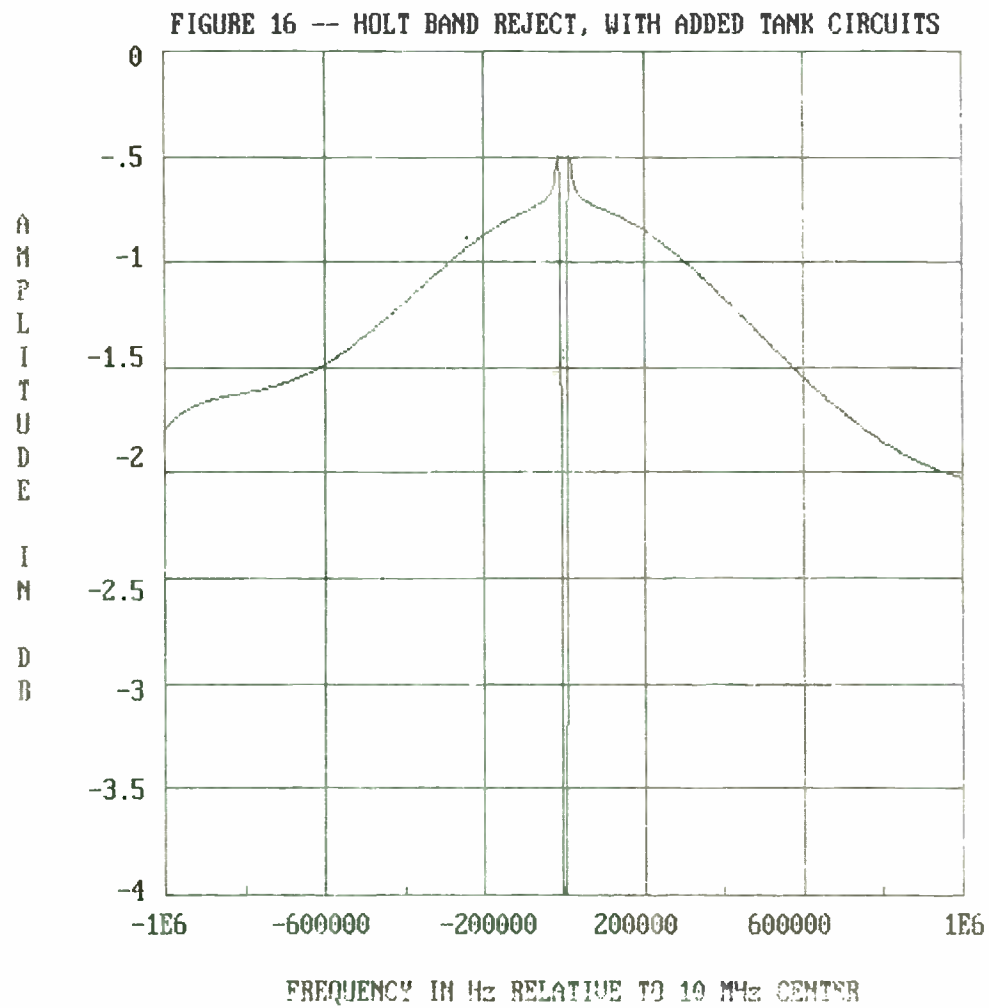


FIGURE 15



## Modular Implementation of the Coupling of Modes Analysis for Surface Acoustic Wave Resonators

Sam M. Richie, and Robert P. Burke

University of Central Florida  
Dept. of Electrical and Computer Engr.  
Orlando, Florida 32816-2450

### ABSTRACT

Surface acoustic wave (SAW) resonators can be used to form high Q, extremely narrowband resonators and resonator filters for application in the 20 MHz to 2 GHz range. The overall performance of the SAW resonator is a very sensitive function of geometry and material parameters and therefore must be modelled accurately. Recently a great deal of attention has been given to the application of the coupling of modes (COM) formalism, better known for its contribution in microwaves and optical systems than to the world of ultrasonic devices, for the analysis of SAW structures. The basic theory of the COM formalism will be discussed via its application to the SAW reflective grating and the SAW transducer to obtain a mixed-matrix representation of the SAW transducer. The S-parameters for these structures will be shown. These two acoustic elements, which represent basic building blocks, may then be combined in different configurations to form either a one port SAW resonator or an inline two port SAW resonator. By cascading the mixed-matrix representations, the resultant device response may be calculated. A discussion of the modular nature of the device analysis will be given. An example one port and two port device will be shown for comparison of theory and experiment.

### 1.0 INTRODUCTION

Surface acoustic wave (SAW) Resonators can be used to form high Q, extremely narrowband resonators and/or resonator filters, for applications in the 20 MHz to 2 GHz range. The lower frequency limitations are imposed by device size, while the upper frequency is limited by photolithography techniques. The overall transfer function of the

SAW resonator is a very sensitive function of geometry and material parameters, which must be accounted for accurately. However, a rigorous analysis of arbitrary resonator structures on various substrates that fully accounts for all piezoelectric and mechanical effects is extremely difficult and mathematically cumbersome from a design point of view. In recent years, however, a great deal of attention has been given to the application of the Coupling of Modes (COM) formalism, better known for its contribution in microwave and optical systems than to the world of ultrasonic devices. Many authors have demonstrated its usefulness as a simpler, less cumbersome analysis and design tool. It is the intention of this thesis to present an example of how the COM analysis techniques allow a straightforward, efficient approach to the analysis of a specific family of SAW devices, namely SAW resonators and SAW resonator filters. First, the basic theory of the COM formalism will be discussed, via its application to the simplest of SAW structures, the SAW reflective grating. Next, the concepts of piezoelectric transduction will be discussed, and used to obtain the solution to a mixed-matrix representation for an interdigital transducer (IDT). Next these two basic SAW components will be put together as a one-port resonator, and then a two-port resonator. As a final example of the many applications of the COM analysis, a multistrip coupler (MSC) will be analyzed, and included in the design of a SAW resonator filter. To verify the validity of the analysis technique, a computer program is presented to implement the analysis, and the predicted insertion loss for a two-port resonator is compared with experimental results.

### 2.0 COUPLING OF MODES EQUATIONS FOR A GRATING

The coupling of modes formalism describes the motion of the surface acoustic wave propagating through the substrate. We assume the surface wave  $u(z)$  can be broken into two components, a forward travelling wave,  $a(z)$ , and a reverse travelling wave  $b(z)$ , with slowly varying amplitudes  $R(z)$  and  $S(z)$  respectively:

$$u(z) = a(z) + b(z) \quad (1)$$

$$a(z) = R(z) e^{-j\frac{k_g z}{2}} e^{jkz} \quad (2)$$

$$b(z) = S(z) e^{+j\frac{k_g z}{2}} e^{jkz} \quad (3)$$

$$k_g = \frac{2\pi}{p} \quad (4)$$

$k_g$  = wave number under the grating

Then, the first order wave equations can be written (dropping the  $e^{jkz}$  factor and the  $z$  dependency) in the absence of grating as<sup>[1]</sup>:

$$\frac{da}{dz} = -jk_R a \quad (5)$$

$$\frac{db}{dz} = +jk_R b \quad (6)$$

$k_R$  is the Rayleigh wave number in the absence of any grating. The grating changes the propagation constant by  $\Delta k$  and couples the forward and reverse travelling waves:

$$\frac{da}{dz} = -j(k_R + \Delta k)a + jk_{z1}b \quad (7)$$

$$\frac{db}{dz} = +j(k_R + \Delta k)b + jk_{z2}a \quad (8)$$

Using the differential form of the equations assumes the change of amplitude per period of the grating is small. Since we have as yet modelled no losses, conservation of power requires:

$$\frac{d}{dz}(R^2 - S^2) = 0 \quad (9)$$

$$k_{12} = -k_{21}^* \quad (10)$$

For the periodic grating,  $k_{12}$  is a function of  $z$ . It can be decomposed as:

$$k_{12} = \sum_{n=-\infty}^{\infty} K(n) e^{jn\frac{k_g z}{2}} \quad (11)$$

Substituting (2) and (3) into (7) and (8) gives:

$$\frac{dR}{dz} = -j\left(k_R - \frac{k_g}{2} + \Delta k\right)R + j \sum_{n=-\infty}^{\infty} K(n) S e^{j(n-1)\frac{k_g z}{2}} \quad (12)$$

$$\frac{dS}{dz} = +j\left(k_R - \frac{k_g}{2} + \Delta k\right)S - j \sum_{n=-\infty}^{\infty} K^*(n) R e^{-j(n-1)\frac{k_g z}{2}} \quad (13)$$

Of all  $n$  Fourier components, only  $n = -1$  produces coupling between the forward and reverse travelling wave that is independent of  $z$ . To simplify notation, we now define the detuning parameter  $\delta$  as:

$$\delta = k_R - \frac{k_g}{2} + \Delta k \quad (14)$$

Using this definition and writing  $K(-1)$  as just  $K$ , the COM equations simplify to:

$$\frac{dR}{dz} = -j\delta R + jKS \quad (15)$$

$$\frac{dS}{dz} = j\delta S - jK^*R \quad (16)$$

To account for propagation loss, we must add an attenuation coefficient  $\alpha$ . This modifies the COM equations to:

$$\frac{dR}{dz} = (-\beta - \alpha)R + jK^*S \quad (17)$$

$$\frac{dS}{dz} = (+\beta + \alpha)S - jK^*R \quad (18)$$

The solution to these equations is straightforward, and given by Elachi<sup>[2]</sup> as:

$$R(z) = c_1 e^{\gamma z} + c_2 e^{-\gamma z} \quad (19)$$

$$S(z) = c_1 \left[ \frac{\gamma - (j\kappa - \delta)}{K^*} \right] e^{\gamma z} + c_2 \left[ \frac{\gamma + (j\kappa - \delta)}{K^*} \right] e^{-\gamma z} \quad (20)$$

The propagation constant  $\gamma$  is defined by:

$$\gamma = \sqrt{K^2 - (\delta - j\kappa)^2} \quad (21)$$

The constants  $c_1$  and  $c_2$  can be evaluated by applying the boundary conditions. The total wave  $u(z) = a(z) + b(z)$  and its derivative with respect to position must be continuous.

This gives boundary conditions as follows:

$$R(0) = 1 \quad (22)$$

$$S(np) = 0 \quad (23)$$

Solving for  $c_1$  and  $c_2$  and substituting in to (20) and (21), we get the amplitudes  $R(z)$  and  $S(z)$ :

$$R(z) = \frac{\gamma \cosh[\gamma(z - np)] - j(\delta - j\kappa) \sinh[\gamma(z - np)]}{\gamma \cosh[\gamma np] + j(\delta - j\kappa) \sinh[\gamma np]} \quad (24)$$

$$S(z) = \frac{-jK^* \sinh[\gamma(z - np)]}{\gamma \cosh[\gamma np] + j(\delta - j\kappa) \sinh[\gamma np]} \quad (25)$$

The distributed reflector nature of the grating can be seen if one examines the amplitudes

of the forward and reverse travelling waves<sup>[2]</sup>. These relationships are shown in Figure 3, which will also aid in visualizing the boundary conditions. By examining the solutions, it can be seen that the incident wave  $R(z)$  decays exponentially along the grating as it scatters, giving its power to the reverse travelling wave.

The exponential decay of the incident wave only occurs when the propagation constant,  $\gamma$ , is real. Therefore, the stopband of the reflector occurs when<sup>[2]</sup>

$$|K^2 > (\delta - j\kappa)^2 \quad (26)$$

With the wave equations determined, we can now determine the scattering matrix characterization of the grating<sup>[2]</sup>:

$$S = \begin{bmatrix} b(o) & a(np) \\ a(np) & b(o) \end{bmatrix} \quad (27)$$

Therefore, from (24) and (25) we obtain:

$$S_{11} = S_{22} = \frac{jK^* \sinh(\gamma np)}{\gamma \cosh(\gamma np) + j(\delta - j\kappa) \sinh(\gamma np)} \quad (28)$$

$$S_{12} = S_{21} = \frac{(-1)^n \gamma}{\gamma \cosh(\gamma np) + j(\delta - j\kappa) \sinh(\gamma np)} \quad (29)$$

These solutions are plotted for a typical grating in Figures 4 and 5.

### 3.0 COUPLING OF MODES EQUATIONS FOR THE IDT

To model the transducer, we must add to the grating equations a transduction term, and develop a third equation to describe the relationship between the current and voltage at the electrical port.

Modifying the grating equations as described above gives<sup>[1]</sup>:

$$\frac{dR}{dz} = (-\beta - \alpha)R + jKS + j\xi V \quad (30)$$

$$\frac{dS}{dz} = (\beta + \alpha)S - jK^*R - j\xi^*V \quad (31)$$

The parameter  $\xi$  is defined as the transduction coefficient. The electrical port can be described by:<sup>[1]</sup>

$$\frac{dI}{dz} = -2j\xi^*R + 2j\xi S + j\omega CV \quad (32)$$

$$C = C_s W \quad (33)$$

$C_s$  is the static capacitance per finger pair, and  $W$  is the IDT aperture.

The solution to (30) - (32) is well documented, but previous to Abbot's<sup>[4]</sup> solution to the bus bar current, the evaluation of the transducers admittance was performed indirectly by calculating the acoustic conductance, taking the Hilbert Transform, and getting the acoustic susceptances<sup>[5]</sup>. This technique required that propagation loss and finger resistance effects be neglected, and also consumed a great deal of time to calculate acoustic conductance over a broad bandwidth. Abbott<sup>[4]</sup> obtained a complete solution to the bus-bar current, allowing direct calculation of the transducer admittance. The P-Matrix<sup>[6]</sup> representation of the transducer solution is:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ V_3 \end{bmatrix} \quad (34)$$

The advantage of using this representation is that it allows easier manipulation of the parallel electrical circuit and acoustically cascaded connection for multiple transducer/grating circuits. The elements  $P_{11}$ ,  $P_{12}$ ,  $P_{21}$ ,  $P_{22}$ , are identical to the scattering parameters developed for the grating in Chapter 2.  $P_{33}$  is the admittance of the IDT as seen at the electrical port if no incoming waves are present, and  $P_{13}$ ,  $P_{23}$  are voltage to SAW transfer elements with units of (amps/volts)<sup>0.5</sup><sup>[5]</sup>

Applying Abbott's<sup>[4]</sup> solutions for the IDT COM Equations, and expressing them in P-Matrix form, we get:

$$P_{11} = S_{11} = \frac{jK^* \sinh(\gamma np)}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \quad (35)$$

$$P_{12} = S_{12} = \frac{(-1)^n \gamma}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \quad (36)$$

$$P_{13} = np \left[ \frac{\sinh(\frac{\gamma np}{2})}{\frac{\gamma np}{2}} \right] \left[ \frac{\xi^* j \gamma \cosh(\frac{\gamma np}{2}) - [K^* \xi + (\delta - j\alpha) \xi^*] \sinh(\frac{\gamma np}{2})}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \right] \quad (37)$$

$$P_{23} = (-1)^n np \left[ \frac{\sinh(\frac{\gamma np}{2})}{(\frac{\gamma np}{2})} \right] \left[ \frac{\xi j \gamma \cosh(\frac{\gamma np}{2}) - [K \xi^* + (\delta - j\alpha) \xi] \sinh(\frac{\gamma np}{2})}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \right] \quad (38)$$

$$P_{33} = 2 \left[ \frac{K^*(\xi)^2 + K(\xi^*)^2 + 2(\delta - j\alpha)(\xi)^2}{\gamma^3} \right] \left[ j \gamma np - \frac{j \gamma \sinh(\gamma np) + (\delta + j\alpha)(1 - \cosh \gamma np)}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \right]$$

$$+ 2 \left[ \frac{(\delta - j\alpha)[K(\xi^*)^2 + K^*(\xi)^2] + 2K^2 \xi^2}{\gamma^3} \right] \left[ \frac{1 - \cosh(\gamma np)}{\gamma \cosh(\gamma np) + j(\delta - j\alpha) \sinh(\gamma np)} \right] + j \left[ \frac{3\omega C_s np}{L(3 + j\omega C_s R_A)} \right] \quad (39)$$

From reciprocity and conservation of energy requirements, the remaining elements are found to be<sup>[4]</sup>.

Typical shapes for the various IDT P-Matrix elements are shown in Figures 7 and 8.



$$P_{22} = P_{11} \quad (40)$$

$$P_{21} = P_{12} \quad (41)$$

$$P_{31} = -2P_{13} \quad (42)$$

$$P_{32} = -2P_{23} \quad (43)$$

The solution to  $P_{33}$  is a very sensitive function of reflectivity and attenuation coefficient. Any errors in these parameters will introduce large errors in  $P_{33}$ . Notice that the first and second terms of the  $P_{33}$  equation both have the cube of  $\gamma$  in the denominator.

Additionally, as frequency is increased to the natural resonant frequency for the structure,  $\gamma$  goes from imaginary to real and back to imaginary as the resonant frequency is exceeded. At each transition, the magnitude of  $\gamma$  approaches zero, and the hyperbolic function terms of  $P_{33}$  become indeterminate. This behavior allows for little error in the reflectivity and attenuation parameters, since they control the amplitude of the individual terms.

#### 4.0 COUPLING OF MODES PARAMETERS

In order to use the COM analysis technique, we must now determine the various parameters used in the COM equations.

##### Detuning Parameter ( $\delta$ )

From Equation (15),  $\delta$  was defined as:

$$\delta = k_R - \frac{k_g}{2} + \Delta k \quad (46)$$

where  $k_R$  is the unperturbed propagation constant at the unperturbed Bragg frequency  $\omega_0$ , so that

$$\delta = \frac{\omega}{v_f} - \frac{\pi}{p} + \Delta k \quad (47)$$

The free-wave velocity,  $v_f$ , is measurable. Typical values are<sup>[1]</sup> 3488 m/sec for Y,Z LiNbO<sub>3</sub>, and 3158 m/sec in ST,X Quartz. The surface wave velocity is perturbed by the presence of grating strips. Morgan<sup>[1]</sup> has summarized the results of a Floquet analysis for an infinite array of electrodes, which ultimately allows an approximation for the surface wave velocity in the presence of open-circuited or short circuited electrodes. The approximations are presented below:

$$v_{oc} = v_f + \frac{1}{2} (v_m - v_f) \left[ 1 - \frac{P_s(\cos\Delta)}{P_{s-1}(\cos\Delta)} \right] \quad (48)$$

$$v_{sc} = v_f + \frac{1}{2} (v_m - v_f) \left[ 1 + \frac{P_s(\cos\Delta)}{P_{s-1}(\cos\Delta)} \right] \quad (49)$$

$$\text{where } s = \frac{k_f p}{2\pi} - n \quad (50)$$

$$\text{for } n \leq \frac{k_f p}{2\pi} \leq n+1 \quad (51)$$

$$\Delta = \pi \frac{g}{p} \quad (52)$$

##### Attenuation Coefficient ( $\alpha$ )

The attenuation coefficient,  $\alpha$ , models the effects of the electrode strips on the surface wave (electrical loading) as well as the various mechanical loadings. The electrical

loading effect can be explained by considering that the conducting strips short out the tangential electric field associated with the surface wave. A surface wave travelling under the strip induces a voltage which acts like a source, generating reflected waves. This loading looks like a characteristic impedance mismatch between the electrode and gap regions, the magnitude of which is a function of the piezoelectric coupling of the substrate.

The mechanical loading effects stem from several mechanisms. First, the weight of the electrode must be "moved" by the surface wave, which slows the wave down. Next, the presence of the electrodes affects the elastic properties of the substrate, making it appear stiffer, increasing the surface wave velocity. A final mechanical loss is due to "air loading", which is attributed to excitation of compressional waves in air (above the substrate) at higher frequencies, carrying power away.

A final means of attenuation stems from the scattering of the surface wave energy into bulk modes at the discontinuities created by the strips. Due to the evanescent nature of these bulk waves, they decay away quickly, resulting in the energy being "stored" in the vicinity of the discontinuity<sup>[9]</sup>.

An analytic expression for  $\alpha$  that concisely models the various contributions is given by Thorvaldsson<sup>[9]</sup> as:

$$\alpha = jk_l [\theta_1 (\frac{k^2}{2}) + \theta_2 (\frac{H_m}{L}) + \theta_3 (\frac{H_m}{L})^2] \quad (53)$$

$$k_r = \frac{\omega}{v_l} \quad (54)$$

The variable  $k$  is the piezoelectric coupling and  $(H_m/L)$  is the height of the grating strips in wavelengths. The constants  $e_1$ ,  $e_2$ , and  $e_3$  are dependent on the substrate material and are given in Table 1.

TABLE 1  
CONSTANTS FOR CONTRIBUTIONS TO ATTENUATION

CONSTANT	QUARTZ	LiNbO <sub>3</sub>
$e_1$	0.7	0.7
$e_2$	0.02	0.3
$e_3$	7.9	Assumed Negligible

#### Coupling (Reflection) Coefficient (K)

The various mechanical and electrical mechanisms that contributed to attenuation also contribute to the reflection of waves. Thorvaldsson<sup>[9]</sup> has considered the various reflection mechanisms, both mechanical and electrical, and developed the following relationship:

$$K(\omega) = k_l [\theta_4 (\frac{k^2}{2}) + \theta_5 (\frac{H_m}{L})] \quad (55)$$

The variable  $k$  is the effective piezoelectric coupling coefficient, and as for the case of  $\alpha$ ,  $H_m/L$  is the height of the discontinuity in wavelengths. The coefficients are given in Table 2<sup>[9]</sup>:

TABLE 2  
CONSTANTS FOR CONTRIBUTIONS TO REFLECTION

CONSTANT	ST-X QUARTZ	LiNbO <sub>3</sub>
$e_4$	0.1337	0.1337
$e_5$	0.1592	0.0764

Although this analytic approximation does not give the best results for some of the more exotic cuts of substrate material, it will serve as a good first approximation for the two substrate materials of interest.

#### Transduction Coefficient ( $\xi$ )

Abbott<sup>[4]</sup> obtained an analytical formulation for the transduction by comparing the Impulse

Response Model solution to  $P_{13}$  with the COM solution for  $P_{13}$  with zero reflectivity (an assumption of the impulse response model). Comparing the two solutions yielded

$$\xi = k(\epsilon_s(\omega) \frac{V_F(k)}{L_T} \sqrt{\frac{\omega W \Gamma_s}{2}} e^{i\phi_T}) \quad (56)$$

$k$  = wave number

$\epsilon_s(\omega)$  = electrostatic permittivity of the piezoelectric substrate

$\Gamma_s$  = SAW coupling coefficient defined by Morgan<sup>88</sup>

$V_F(k)$  = Fourier Transform of the elemental potential of the transducer structure

$L_T$  = transduction period

$\phi_T$  = transduction phase (zero for the case of uniformly spaced gratings/electrodes)

Morgan<sup>88</sup> has shown that the Fourier transform of the elemental potential for an infinite array of regular electrodes can be expressed as:

$$V_F(k) = (\epsilon_o + \epsilon_p) \frac{2 \sin(\pi s)}{P_{-s}(-\cos \frac{\pi a}{p})} P_n(\cos \frac{\pi a}{p}) \quad (57)$$

$$\text{For } n \leq \frac{k_1 p}{2\pi} \leq n+1 \quad (58)$$

$$\epsilon_p = (\epsilon_{11} \epsilon_{33} - \epsilon_{13}^2)^{1/2} \quad (60)$$

$$\text{where: } s = \frac{k_1 p}{2\pi} - n \quad (59)$$

Finally, the transduction phase can be expressed as<sup>89</sup>:

$$\phi_T = -\arg[V_F(k)] \quad (61)$$

For the case of uniformly spaced electrodes, the transduction phase will be zero.

#### Static Capacitance ( $C_s$ )

Engan<sup>194</sup> gives an expression for capacitance per finger pair, which was found by calculating total charge per unit length of each electrode:

$$C_s = (\epsilon_o + \epsilon_p^T) \frac{K(\xi)}{K(\sqrt{1-\xi^2})} \quad (62)$$

$\epsilon_o$  = dielectric permittivity

$\epsilon_p^T$  = zero stress permittivity

$$\xi = \cos\left(\frac{\pi}{2}\left(1 - \frac{a}{p}\right)\right) \quad (63)$$

(geometry factor)

$$K(\xi) = \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{1-\xi^2 \sin^2 \theta}} d\theta \quad (64)$$

Thorvaldsson<sup>195</sup> applied this technique and obtained results of  $C_s = 50.338$  pF/m for ST-Quartz and  $C_s = 464.38$  pF/m for YZ-LiNbO<sub>3</sub> assuming a metallization ratio  $\eta = a/p$  of 0.5<sup>89</sup>.

#### Finger Resistance ( $R_f$ )

The finger resistance is the parallel combination of all electrodes attached to the positive side of the voltage port in series with the parallel combination of all electrodes attached

to the negative side of the voltage port:

$$R_F = \left[ \sum_{\text{electrodes}} \frac{1}{R_n} \right]^{-1} + \left[ \sum_{\text{electrodes}} \frac{1}{R_n} \right]^{-1} \quad (65)$$

where  $R_n$  is the end to end resistance of the nth electrodes, given by:

$$R_n = R_s \frac{W}{a_n} \quad (66)$$

where  $W$  is the transducer's aperture (electrode length),  $a_n$  is the width of the  $n^{\text{th}}$  electrode, and  $R_s$  is the resistivity of the electrode material.

## 5.0 CASCADING COMPONENTS

In order to make efficient use of computer time and memory, different techniques were used to cascade the various SAW components for each specific type of device analyzed. The following section delineates these techniques. For future applications, other more flexible techniques could be used. Especially attractive is the conversion of the various components to two-port ABCD Matrices, and then cascading, as demonstrated by Brown, et. al.<sup>[11]</sup>. Another viable option would be to calculate the S-Parameters for each component of the device to be modelled, store them in files, then import them to a standard microwave circuit analysis program<sup>[12]</sup>.

### One-Port Resonator

A one-port resonator can be constructed from a single IDT and two reflective gratings. All components can be modelled by 3 X 3 P-Matrices, with the appropriate terms forced to zero for the grating and the delay (which models the distance travelled by surface acoustic wave between the IDT and each grating). Since the terms  $P_{11}$  through  $P_{22}$  are effectively the scattering parameters for the forward and reverse travelling acoustic waves, they are identical to  $S_{11}$  through  $S_{22}$  for the grating.

Thus, the P-Matrix for the grating becomes:

Additionally, the S-Matrix for the delay is<sup>[13]</sup>

$$[P]_{\text{grating}} = \begin{bmatrix} S11 & S12 & 0 \\ S21 & S22 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (67)$$

$$[S]_{\text{delay}} = \begin{bmatrix} 0 & e^{-jkD_m} \\ e^{-jkD_m} & 0 \end{bmatrix} \quad (68)$$

and thus the P-Matrix for the delay is

$$[P]_{\text{delay}} = \begin{bmatrix} 0 & e^{-jkD_m} & 0 \\ e^{-jkD_m} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (69)$$

For this simple configuration, the basic elements can each be considered as separate three-port devices, each with acoustic ports (ports 1 and 2) and a single electrical port (port 3). When cascaded to similar three-port devices, the acoustic ports are effectively in series, while the electrical ports are in parallel.

Under these conditions, the equivalent P-Matrix for the cascaded pair becomes<sup>[14]</sup>

$$P11 = P11_A + P11_B \left[ \frac{P21_A P12_A}{1 - P11_B P22_A} \right] \quad (70)$$

$$P12 = \frac{P12_A P12_B}{1 - P11_B P22_A} \quad (71)$$

$$P13 = P13_A + P12_A \left[ \frac{P13_B + P11_B P23_A}{1 - P11_B P22_A} \right] \quad (72)$$

The remaining terms are obtained from the reciprocity and power conservation

$$P22 = P22_B + P22_A \left[ \frac{P12_B P21_B}{1 - P11_B P22_A} \right] \quad (73)$$

$$P23 = P23_B + P21_B \left[ \frac{P23_A + P22_A P13_B}{1 - P11_B P22_A} \right] \quad (74)$$

$$P33 = P33_A + P33_B + P32_A \left[ \frac{P13_B + P11_B P23_A}{1 - P11_B P22_A} \right] + P31_B \left[ \frac{P23_A + P22_A P13_B}{1 - P22_A P11_B} \right] \quad (75)$$

requirements summarized by equations 40-43.

The above technique is used to cascade first the left most grating to the delay, then the equivalent P-Matrix to the P-Matrix for the IDT, and so on, until a single equivalent P-Matrix is converted to S-Parameters by<sup>[4]</sup>

$$S_{11} = P_{11} - \frac{Z_3 P_{31} P_{13}}{1 + Z_4 P_{33}} \quad (76)$$

$$S_{12} = P_{12} - \frac{Z_3 P_{32} P_{13}}{1 + Z_4 P_{33}} \quad (77)$$

$$S_{13} = \frac{2 \sqrt{R_5} P_{13}}{1 + Z_4 P_{33}} \quad (78)$$

$$S_{21} = P_{21} - \frac{Z_3 P_{31} P_{23}}{1 + Z_4 P_{33}} \quad (79)$$

$$S_{22} = P_{22} - \frac{Z_3 P_{32} P_{23}}{1 + Z_4 P_{33}} \quad (80)$$

$$S_{23} = \frac{2 \sqrt{R_5} P_{23}}{1 + Z_4 P_{33}} \quad (81)$$

$$S_{31} = \frac{-\sqrt{R_5} P_{31}}{1 + Z_4 P_{33}} \quad (82)$$

$$S_{32} = \frac{-\sqrt{R_5} P_{32}}{1 + Z_4 P_{33}} \quad (83)$$

$$S_{33} = \frac{1 - Z_4 P_{33}}{1 + Z_4 P_{33}} \quad (84)$$

Since port three is the electrical port,  $S_{33}$  represents the electrical reflection coefficient, and so on.

### Two-Port Resonator

For the two-port resonator, the IDT's acoustic ports are in series through a delay element representing the distance between the transducers, but their electrical ports are not connected, so the equivalent P-Matrix technique of the previous section does not apply. If the outer acoustical port of each IDT is terminated in its characteristic impedance, their effects will be inconsequential, and the device can be reduced to an equivalent two-port, with the two ports of interest being the electrical port of each IDT.

First, the effects of the gratings and their associated delay are accounted for by cascading each IDT to its grating and delay as in the last section. Next, the two equivalent P-matrices are converted to an equivalent two-port Admittance Matrix (Fig 11). The resulting admittance matrix elements are<sup>[4]</sup>

$$Y_{11} = P_{33_A} + P_{11_B} \left[ \frac{P_{32_A} P_{23_A}}{1 - P_{11_B} P_{22_A} e^{-j2kD_1}} \right] e^{-j2kD_1} \quad (85)$$

$$Y_{12} = \frac{P_{32_A} P_{13_B}}{1 - P_{11_B} P_{22_A} e^{-j2kD_1}} e^{-j2kD_1} \quad (86)$$

$$Y_{21} = \frac{P_{13_A} P_{32_B}}{1 - P_{11_B} P_{22_A} e^{-j2kD_1}} e^{-j2kD_1} \quad (87)$$

$$Y_{22} = P_{33_B} + P_{22_A} \left[ \frac{P_{31_B} P_{13_B}}{1 - P_{11_A} P_{22_B} e^{-j2kD_1}} \right] e^{-j2kD_1} \quad (88)$$

Next, the Y-Parameters are converted to S-Parameters by<sup>[7]</sup>

$$S_{11} = \frac{(1 - Z_1^* Y_{11})(1 + Z_2 Y_{22}) + Y_{12} Y_{21} Z_2 Z_2^*}{(1 + Z_1^* Y_{11})(1 + Z_2 Y_{22}) - Y_{12} Y_{21} Z_1 Z_2} \quad (89)$$

$$S_{12} = \frac{-2 \sqrt{R_1 R_2} Y_{12}}{(1 + Z_1 Y_{11})(1 + Z_2 Y_{22}) - Y_{12} Y_{21} Z_1 Z_2} \quad (90)$$

$$S_{21} = \frac{-2 \sqrt{R_1 R_2} Y_{21}}{(1 + Z_1 Y_{11})(1 + Z_2 Y_{22}) - Y_{12} Y_{21} Z_1 Z_2} \quad (91)$$

$$S_{22} = \frac{(1 - Z_2^* Y_{22})(1 + Z_1 Y_{11}) + Y_{12} Y_{21} Z_1 Z_1^*}{(1 + Z_1 Y_{11})(1 + Z_2 Y_{22}) - Y_{12} Y_{21} Z_1 Z_2} \quad (92)$$

## 6.0 COMPUTER MODEL

A computer program utilizing the C programming language was written to implement the COM analysis technique. The program models each of the individual components necessary to model a resonator device. Two main modules tie together the various components necessary to model both a one-port and a two-port resonator. The basic structure of the program is outlined in Figure 22.

To run the program, at the DOS prompt for the drive containing the program, simply type "res", short for resonator. A menu will appear, asking the user which of five devices he wishes to analyze: grating, IDT, MSC, one-port resonator, or two-port resonator. After the device is specified, the user is prompted to specify a file to hold the data, the geometry of the device and the analysis range. The model sets an upper limit of 10,000 data points, which is more than sufficient for most applications. The notation used for the user prompts is consistent with that used throughout this paper.

The output is sent to a text file whose name is specified by the user. Each column in the file has a heading describing the data below it. The suffixes .M, .P, .R, and .I are used to denote magnitude, phase, real part, and imaginary part respectively. The text file is easily imported into any commercially available spreadsheet to allow plotting. All magnitudes are absolute magnitudes, all phase data are in degrees, and real/imaginary data are in the appropriate fundamental units.

The modular design of the program allows for easy modification of the methods used for calculating the COM parameters. The use of the C programming language enhances the

portability of the program, and ultimately will allow it to be incorporated into SAWCAD.

## 7.0 EXPERIMENTAL RESULTS

To test the COM theory and the analysis program, a two port resonator was modelled. The dimensions of the device modelled are given in Table 3.

TABLE 3  
DEVICE DIMENSIONS

$P_{\text{grating}}$	10.5 $\mu\text{m}$
$a_{\text{grating}}$	5.25 $\mu\text{m}$
$N_{\text{grating}}$	650
$P_{\text{IDT}}$	10.5 $\mu\text{m}$
$a_{\text{IDT}}$	5.25 $\mu\text{m}$
$N_{\text{IDT}}$	45
$H_m$	1800 $\text{\AA}$
$W$	1260 $\mu\text{m}$
$D_t$	225.75 $\mu\text{m}$
$D_m$	13.125 $\mu\text{m}$
Substrate	ST,X Quartz

Figures 23 through 25 show a comparison of Thorvaldsson's<sup>[5]</sup> results with those predicted by the computer model. Excellent agreement between the two is shown.

## CONCLUSIONS

An accurate, efficient analysis program has been developed for SAW resonators utilizing the coupling of modes model. The analysis method was applied specifically to uniform grating and transducer structures, but with only slight modifications can be applied to more complex structures, including apodization and reflectivity weighting. This work provides verification of Abbott's<sup>[4]</sup> contribution to the coupling of modes model, specifically

his closed form expression for transducer admittance, and his analytical expression for the transduction parameter.

To verify both the analysis method and the computer program implementation, the predicted insertion loss of a two-port resonator was compared with experimental measurements made by Thorvaldsson<sup>[3]</sup>. Excellent agreement between the theoretical calculations and experimental measurement was obtained.

A large area of difficulty involved the determination of the best means of representing the various coupling of modes parameters, such as reflectivity and attenuation. Small errors in these parameters cause large variations in the solutions to the wave equations, which is to be expected when dealing with the numerous coupled parameters associated with this technique. One loses a great wealth of intuition when choosing the coupling of modes model, since its representation of devices is not in terms of parameters necessarily familiar to electrical engineers. However, once some familiarity with the COM equations is obtained and the parameters required are known with some degree of certainty, it presents an attractive alternative to the equivalent circuit model, in terms of both simplicity and numerical efficiency.

## REFERENCES

- [1] Chen, D.P. and H.A. Haus. "Analysis of Metal-Strip SAW Gratings and Transducers." IEEE Transactions on Sonics and Ultrasonics, vol. SU-32, pp. 395-408, May 1985.
- [2] Elachi, C. "Waves in Active and Passive Periodic Structures: A Review." Proceedings of the IEEE, vol. 64, pp. 1666-1698, December 1976.
- [3] Thorvaldsson, Thor. "Design and Analysis of Surface Acoustic Wave Resonators and Resonator Filters." Ph.D. Dissertation, Swiss Federal Institute of Technology, Zurich, 1986.
- [4] Abbott, B.P. "A Coupling-of-Modes Model for SAW Transducers with Arbitrary Reflectivity Weighting." Ph.D. Dissertation, University of Central Florida, Orlando Florida, 1989.
- [5] Tobolka, G. "Mixed Matrix Representation of SAW Transducers." IEEE Transactions on Sonics and Ultrasonics, vol. SU-26, pp. 426-428, November 1979.
- [6] Abbott, B.P. and D.C. Malocha. "Closed Form Solution for Multistrip Coupler Operation Including the Effects of Electrode Resistivity." 1990 Ultrasonics Symposium Proceedings, pp. 25-30.
- [7] Pozar, David M. Microwave Engineering. Reading, Mass.: Addison-Wesley, 1990.
- [8] Morgan, D.P. Surface-Wave Devices for Signal Processing. New York: Elsevier, 1985.
- [9] Li, R.C.M. and J. Meingalis. "The Influence of Stored Energy at Step Discontinuities on the Behavior of Surface Wave Gratings." IEEE Transactions on Sonics and Ultrasonics, vol. SU-22, pp. 189-198, May 1975.
- [10] Engan, H. "Excitation of Elastic Surface Waves by Spatial Harmonics of Interdigital Transducers." IEEE Transactions on Electron Devices, vol. ED-16, pp. 1014-1017, 1969.
- [11] Brown, R.B., J.H. Hines, and B.H. Horine. "Practical Implementation of Coupling-of-Modes Theory for SAW Device Modelling." 1989 Ultrasonics Symposium Proceedings, pp. 153-158.
- [12] Gerd Scholl, Andreas Christ, Werner Ruile, Peter H. Russer, and Robert Weigel. "Efficient Analysis Tool for Coupled-SAW-Resonator Filters." IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 38, pp. 243-251, May 1991.
- [13] Gonzalez, Guillermo. Microwave Transistor Amplifiers, Analysis and Design. New Jersey: Prentice Hall, Inc., 1984.
- [14] Hartman, C.S. and B.P. Abbott. "Generalized Impulse Response Model for SAW Transducers Including Effects of Electrode Reflections." 1988 Ultrasonics Symposium Proceedings, pp. 29-34.

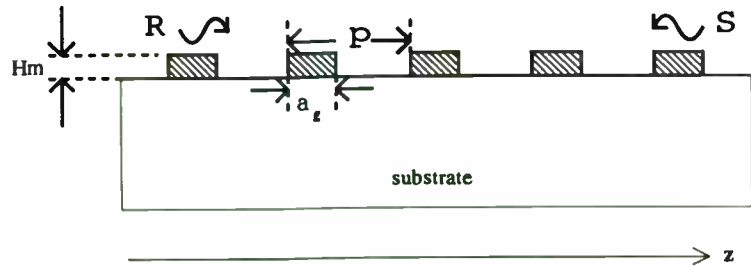


Figure 1. A grating consisting of  $n$  periodic strips of period  $p$ , strip width  $a_p$ , and height  $H_m$ .

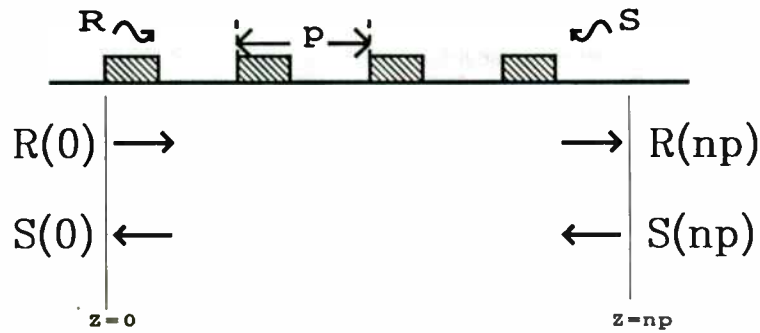


Figure 2. Forward wave amplitude  $R(z)$  and backward wave amplitude  $S(z)$  at the device boundaries.

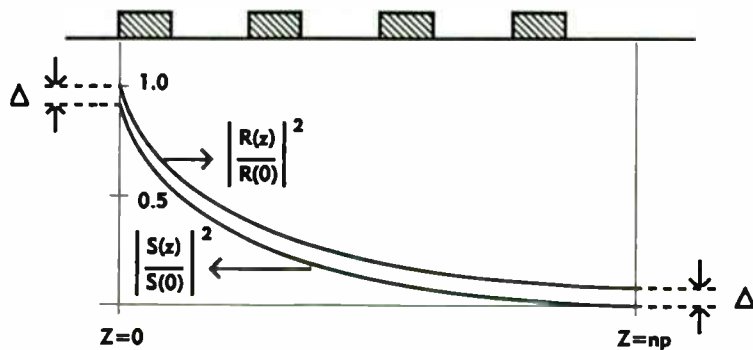


Figure 3. Power transfer from incident to reflected wave.

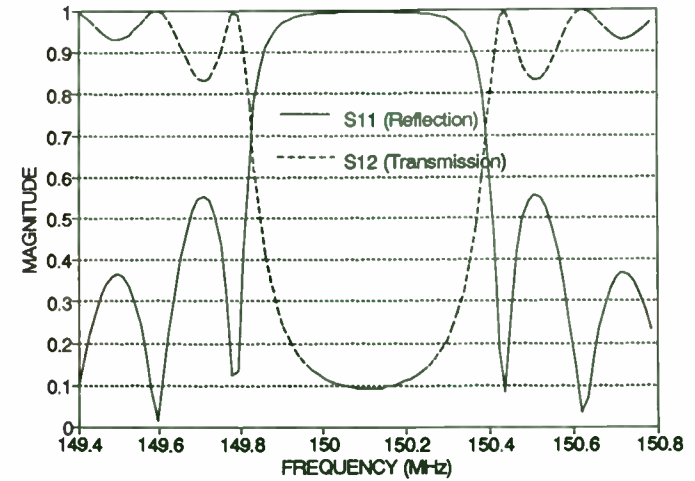


Figure 4. Reflection and Transmission coefficient magnitudes for a typical grating ( $p=10.5 \mu\text{m}$ ,  $a=5.25 \mu\text{m}$ ,  $n=650$ , ST-X Quartz).

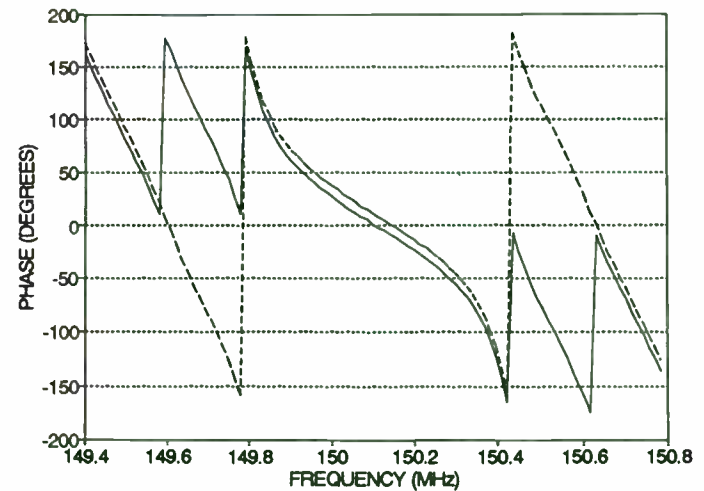


Figure 5. Reflection and Transmission phase for same grating as Figure 4. Phase for  $S_{12}$  has been shifted up by ten degrees for visual clarity.



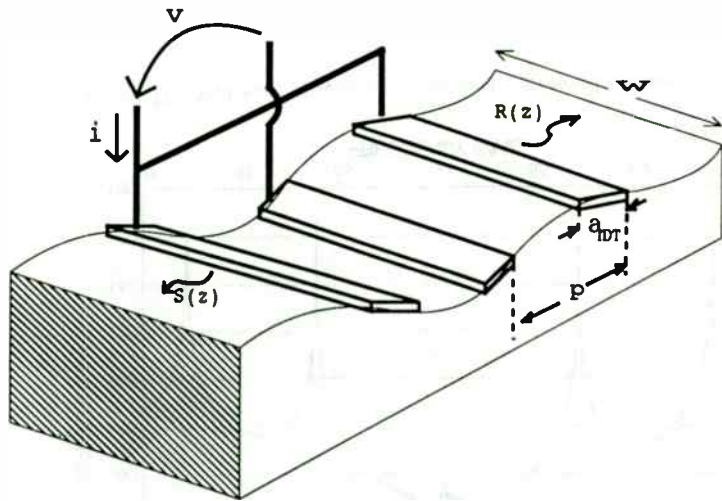


Figure 6. Physical arrangement of interdigital transducer.

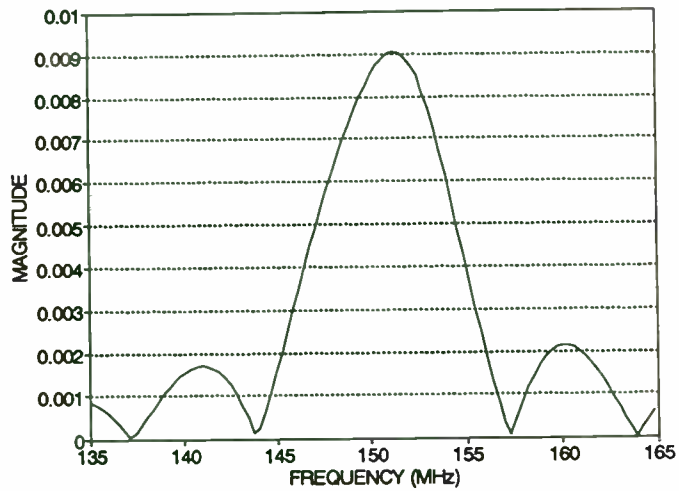


Figure 7.  $P_{31}$  and  $P_{32}$  for a typical IDT ( $p=10.5 \mu\text{m}$ ,  $a= 5.25 \mu\text{m}$ ,  $W=1260 \mu\text{m}$ ,  $n=45$ ).

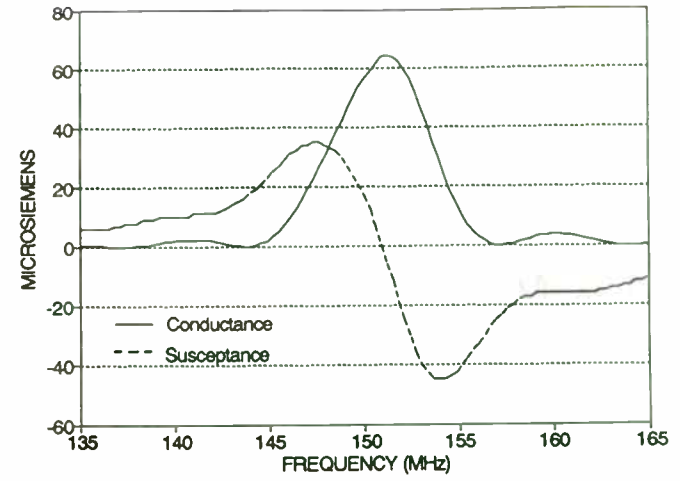


Figure 8 IDT Admittance,  $P_{33}$  for same transducer as Figure 7. The static capacitance has been omitted from the acoustic susceptance term for clarity.

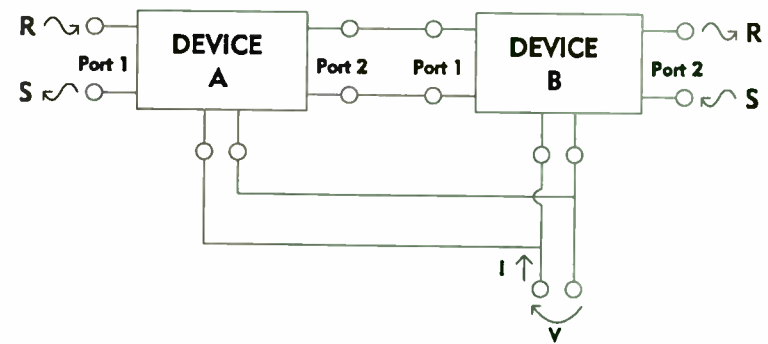


Figure 9. Assumed configuration for cascading of P-matrices.

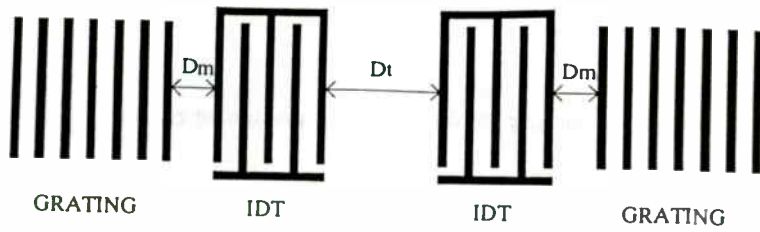


Figure 10. Configuration for two port resonator.

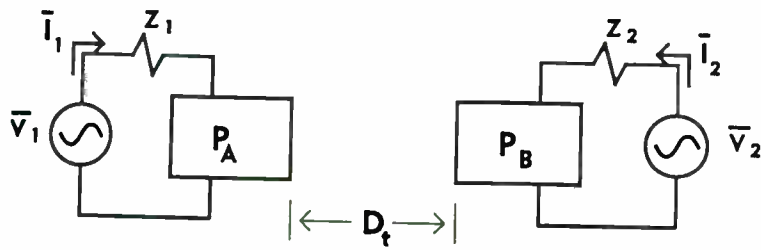


Figure 11. Configuration for conversion to Y-parameters.

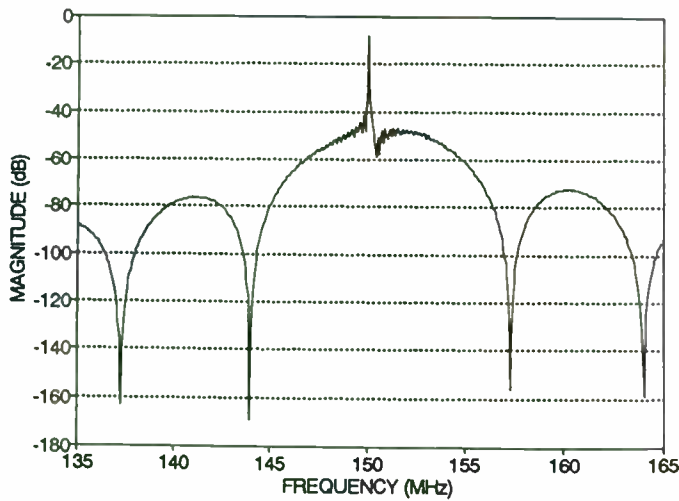


Figure 12. Wideband view of computer model predicted insertion loss.

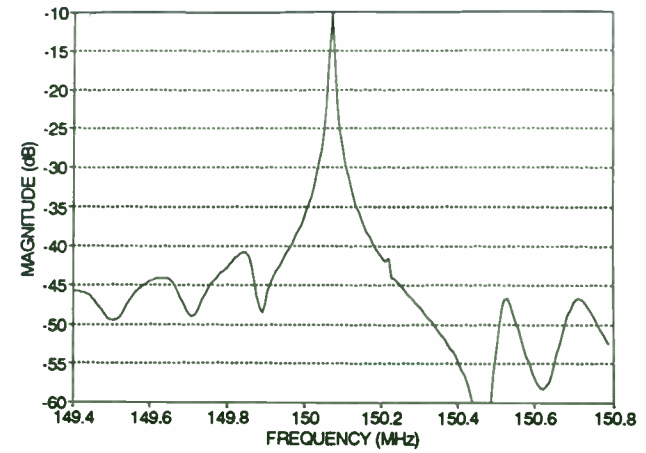


Figure 13. Narrowband view of computer model predicted insertion loss.

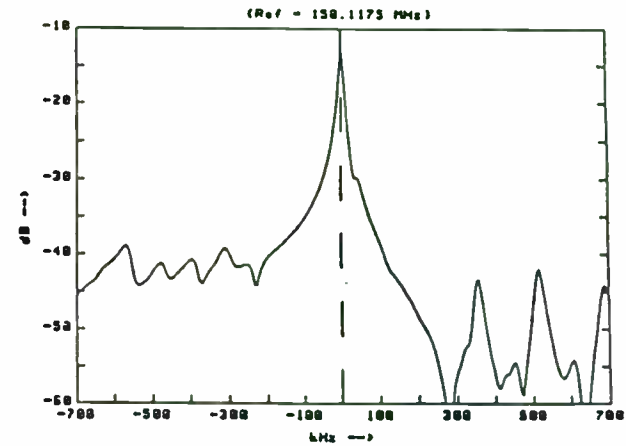


Figure 14. Measured insertion loss for two-port resonator. Reprinted, with permission, from Thorvaldsson<sup>[3]</sup>, Figure 12.4.

**Multiple Receiver Configurations Realizable  
With One Set of Standard Stanford Telecom ASICs**

Herman A. Bustamante  
Stanford Telecommunications, Inc.  
2421 Mission College Boulevard  
Santa Clara, CA 95054

The use of advanced digital signal processing techniques has become the preferred method of modern receiver and modem design for two reasons. First of all, the performance characteristics of a digital circuit can be clearly defined with statistically predictable exactness and without having to perform a possibly tedious adjustment procedure to balance component characteristics as may have to be performed for an analog design. Secondly, once a digital design is well defined, tested and proven, it can be counted on to always perform in the same manner. It is true that there are many types of receivers, and many types of systems such that one digital receiver design is incapable of performing in all cases. However, it is possible to define a relatively small number of basic receiver building blocks from which many types of receivers can be implemented. This note will describe some of these basic blocks and will show how Stanford Telecom, or any one else, can use them to design a very broad range of receiver, or modem, types.

### Introduction

Stanford Telecom has been designing receivers since its inception. Originally its receiver designs were primarily in the form of analog circuits except for that portion processing the reconstituted digital data. Figure 1 shows a typical receiver in a simple block diagram form consisting of an RF front end and downconverter, carrier tracking circuits, data demodulator, data matched filter, bit synchronizer, and decoder. Originally everything except the outputs of the data matched filter and

the decoder were implemented in analog form. It soon became clear that there were digital techniques which could be used to replace portions of the analog circuitry. Not only could the analog circuitry be replaced but some of the design problems of adjusting, balancing, or matching analog circuit characteristics would be eliminated by the use of the digital circuit techniques. Many of these blocks have now been designed in digital form and are used in all of Stanford Telecom's modems.

Stanford Telecom now has a complete set of off-the-shelf standard ASIC's which permit the implementation of wide variety of modem designs. These various modem types span a wide range of modulation techniques, and data rates in both clear and spread spectrum configurations. Use of this family of ASIC's achieves implementation of the receiver in an all digital form all the way from the IF to the baseband data output.

### STEL-2130 Downconverter Functional Characteristics

Four chips are described here. These consist of the STEL 2110 Bit Sync/PSK Demodulator, the STEL-3310 64-Tap Matched Filter, the STEL-2120 DPSK Demodulator, and the STEL-2130 Downconverter and NCO ASIC. The STEL-2130 is the latest addition to the Stanford Telecom set of digitally implemented signal demodulation ASIC's. With the addition of the STEL-2130, the only receiver functions performed in an analog form are the front end high frequency RF signal downconversion and first IF amplification.

### General Chip Characteristics

The new STEL-2130 ASIC continues the march towards integration of more and more receiver functions into one single chip. The chip block diagram is shown in Figure 2, and its performance features are summarized in Table I. Note that these five major functions are all contained in the device:

1. IF sampling and complex multiplication (SSB mixing)

RPT/M 09z

2. Dual integrate and dump accumulators
3. Post processing signal scaling and output control
4. Self contained dual (quadrature) output NCO for I/Q carrier tracking
5. Controllable loop filter to be used for either phase or frequency carrier tracking

This represents another major step in providing the best available ASICs containing major receiver functions in a single chip. This has been done for all the good reasons of lower cost, more compact hardware implementation, and enhanced equipment reliability.

By combining the STEL-2130 with several other of Stanford Telecom's modem design ASICs, the implementation of a variety of both clear and spread spectrum modem configurations can be realized almost entirely in digital form. BPSK, QPSK, and DQPSK modulation can all be accommodated. The interfaces of the chip are specifically designed to permit "snap together" operation with other chips without the need of any interconnecting circuitry, or "glue logic". The STEL-2130 performs the final down conversion from IF to baseband and provides the carrier frequency correction capability for signal tracking operations. Note that this achieves not only the integration of several functions, it also achieves the elimination of high speed costly DACs and filters required between the carrier tracking NCO and the complex multiplier (SSB) function when performed in an analog form.

The digital IF sampling can operate at up to a 40 MHz sample rate either as I/Q inputs for wider bandwidth signals, or as a single sampled input in the case of a narrow band input. In this mode the circuit can be operated in a reduced power down mode, providing all the performance required, but with reduced power requirements.

The complex multiplier operates with 8-bit complex data inputs at up to 40 MHz rates. The operation of the complex multiply results effectively in a SSB mixer operation producing

RPT/M 09z

significantly reduced levels of unwanted components. Truncated 10-bit complex products are provided as inputs to the succeeding 19-bit I/Q integrate and dump circuits. The maximum throughput rate of the complex multiplier circuits is 40 MHz.

The integrate and dump filters are used for baseband filtering and can be operated at a variable dump rate which can be synchronous with the input sampling rate or can be controlled by a separate external clock. The effective filter bandwidth is a direct function of the selected dump rate. A maximum accumulation rate of 40 MHz can be used.

The NCO, which provides the LO for the final signal down conversion, can also operate with a maximum clock rate of 40 MHz. It has a 32-bit accumulator, contains a 10-bit phase/8-bit amplitude look up table and provides sine and cosine waveform generation with unwanted spurs down more than 40 dB. Note that the NCO can be driven from the loop filter, from a separate external input, or by the combination of the two. The inclusion of the NCO within the chip permits driving the complex multipliers directly in a digital signal format without the need for high speed DACs and analog filters.

The loop filter can provide either a 1st or 2nd order filter function, and controls the effective gain in the loop. High update rates can be accommodated and the input to the filter can be either phase or frequency discriminator information such as that available from Stanford Telecom STEL-2120 or STEL-2110A demodulator ASICs.

The input and output interface circuits have been designed to be directly compatible with other Stanford Telecom ASICs. The input circuits can be operated either in a wide band mode with dual A/D inputs in an I/Q format, or in a narrow band mode with a single A/D input. When operating in the narrow band/single A/D mode, the unused internal circuits can be powered down to achieve a reduced power operating mode.

The output circuits operate in a fixed point signal format. Also the output signal is provided in a saturated mode thereby preventing amplitude rollover so as not to provide erroneous output signal magnitudes caused by the cyclic nature of two's complement arithmetic.

#### STEL-2120 Digital DPSK Demodulator Functional Characteristics

A block diagram for the STEL-2120 Digital DPSK Demodulator is shown in Figure 2. This unit can be configured for the demodulation of DPSK, DMSK, DQPSK and  $\pi/4$  DQPSK modulated data. A  $\pm\pi/4$  rotation of each symbol can be activated for demodulation of QPSK signals, and this rotation may be disabled for  $\pi/4$  QPSK modulation used in many systems, including digital cellular telephone systems. The device provides 4-bit soft decision demodulated data for all the modulation formats.

The STEL-2120 demodulates In-Phase and Quadrature data at rates up to 8 Mbps and incorporates extensive programmability to permit demodulation in many environments and applications. The device is configured to connect easily to other Stanford Telecom ASICs to provide an essential function block within a communication system. The device incorporates all of the necessary logic to accommodate timing, latency, and interface considerations for use with the STEL-3310 Digital Matched Filter for Acquisition/Correlation in Spread Spectrum applications, and with the STEL-5269 Viterbi Decoder for Forward Error Correction.

Data demodulation is accomplished by complex multiplication of the current symbol with the previous symbol to generate Dot and Cross product terms. Automatic Frequency Control information is generated for use in carrier acquisition and timing through the use of complex decision directed processing. Automatic Level Control (ALC) sensors providing adaptive control of the soft decision values of the I and Q output data signals are also incorporated within the device. The ALC sensors not only provide a digital means to operate with dynamically varying

signal strengths, but achieve an optimum output signal format for interfacing with other Digital Signal Processing ASIC's.

#### STEL-3310 Expandable Digital Matched Filter/Correlator Characteristics

The STEL-3310 is a dual high-speed digital matched filter/correlator circuit which can be used in many data communications applications. The circuit is designed in  $1.5 \mu$  CMOS and will operate at chipping rates up to 11 Mcps. The device requires a Master Clock (MCLK IN) input at four times the chipping rate, i.e. up to 44 MHz. The internally generated System Clock (SCLK) is provided as an output and runs at half the MCLK IN rate. It can be used as a sample clock at twice the chip rate. The device is designed to be expandable up to 256 taps by cascading up to four devices. Cross-term inputs and outputs are provided for QPSK applications where different spreading sequences are applied to the I and Q data channels. The dual channels allow the device to be used directly after the baseband down-converter, the magnitude of the complex signal being computed internally with an approximation algorithm. The built-in threshold comparator allows the user to select the level at which the match is detected, permitting optimum operation over a wide range of signal conditions.

The optional Front-End Processor (FEP) function is a sliding window filter which adds the previous signal sample to the current sample. This allows the use of non-coherent sampling at two samples per chip. For operation at one sample per chip the input signals should be duplicated so that each sample is loaded twice per chip. If the FEP is turned off only one of these samples will be processed since the samples are moved down the filter register twice per chip.

#### STEL-2110A Bit Synchronizer/PSK Demodulator Characteristics

RPT/M 09z

The STEL-2110A Bit Synchronizer/PSK Demodulator provides bit timing to control the sampling of the signal in a receiver as well as a feedback signal to control the frequency of the local oscillator. It can be used in high speed coherent PSK, QPSK and DPSK modems in either continuous carrier or burst carrier (TDMA) environments. Soft decision output data is provided to facilitate the inclusion of Forward Error Correction (FEC) using convolutional coding and Viterbi decoding into the system.

The bit synchronizer (clock recovery) portion of the STEL-2110A is a digital phase locked loop which operates by integrating the input signals in both the I and Q channels over one symbol period. This is done 3 times; in addition to the nominally "on time" integration, "quarter period early" and "quarter period late" integrations are also carried out. The difference between the last two gives an indication of the timing error, since the averaged difference will be zero when the timing is correct. This signal is used to drive a numerically controlled oscillator which produces the clock signals to drive the entire circuit as well as sampling the incoming signals. The signals can also be integrated prior to the main timing integrators. This can simplify operation over a wide range of data rates by maintaining a constant sampling rate, thereby eliminating the need to change the filters before the A/D converters. The number of accumulations per symbol in this section is controlled by the pre-accumulation control factor,  $A_i$ . This pre-accumulation factor can be set to one, two, four, eight or sixteen times. The optimally integrated IOPT and QOPT signals are used to derive a DPSK output and a feedback signal to control the local oscillator for carrier tracking. This signal can be selected to be a phase locked loop (PLL) control signal or an automatic frequency control (AFC) signal. The former is intended for coherent demodulation in continuous carrier systems, when the IOPT and QOPT signals will be the data outputs, and the latter is intended for use where fast acquisition is required, e.g., in burst carrier systems, such as TDMA, where the acquisition time of a PLL would cause significant loss of data at the beginning of each burst. In such cases either differential demodulation should be used or an STEL-2210 Block Phase Estimator can be used to compensate for the residual phase roll in burst applications. An output

RPT/M 09z

signal derived from the I and Q data can be processed to give an indication of when the timing recovery circuit is in lock. The STEL-2110A is designed for maximum flexibility and can be used in Offset QPSK as well as regular QPSK systems.

Several options are available with the STEL-2110A; the internal bit timing can be bypassed allowing external timing to be used, allowing two or more devices to be slaved together, and the timing discriminator loop filter can also be internal or external, giving the maximum flexibility for loop design.

#### Receiver Design Examples

The devices described here have been used by Stanford Telecom, as well as many of its customers, in a wide variety of products ranging from burst type spread spectrum receivers, to correlator type spread spectrum receivers depending on the special needs of the particular application. The totality of Stanford Telecom ASIC's now available makes a complete set which not only provide signal downconversion and demodulation but error correction coding and decoding, and adaptive signal equalization. All new Stanford Telecom demodulation product ASIC's are designed so as to make them compatible with the previously existing chips in the set and permit "snap together" implementation of complex receivers. A number of typical applications and how they evolved are described below.

#### Application Example 1: Low Power Low Data Rate Receiver

A low power S-band receiver, as might be used to receive 9.6 kbps data with either BPSK or QPSK modulation, is illustrated in Figure 6. Utilization of the Stanford Telecom STEL-2110A Bit Synchronizer/PSK Demodulator chip together with the new STEL-2130 Down-converter chip provides a complete demodulator function. Adding to this a simple RF down-converter, a single A/D and a crystal oscillator completes the entire receiver with the minimum of design required for the major functions. It should be noted that only nominal values are shown for the received

frequency, receiver LO and sampling clock frequency. The exact values are not important for the discussion to follow.

The fact that this unit is receiving a low data rate, narrow band signal, permits use of a single A/D converter. In this case alternate input signal samples represent samples of the I and Q input signal functions. When operating in this mode, not all of the internal STEL-2130 circuitry need be used and it can be operated in a reduced power-down mode.

The data rate in this receiver design is low relative to the sampling rate which can be accommodated by the STEL-2130. This being so we can refer to this design as being representative of an "over sampled" signal case. It is of interest to examine the resultant signal spectra occurring within the STEL-2130 circuitry. These are illustrated in Figure 3 for successive processing points within the STEL-2130.

The input data rate is 9.6 kbps modulated onto an IF carrier at a frequency  $F_1$ , on the order of 800 KHz. This is illustrated in Figure 7(a) as a two sided spectrum for both positive and negative frequencies. The data spectrum is purposely shown to have a slope so that the changes in the successive spectra can be followed more easily. The sampling signal spectrum is represented as a single frequency in Figure 7(b), and is here selected to be 4 MHz.

Once sampled the resultant input IF spectrum appears as shown in Figure 7(c). Components appear about  $F_{Sa}$  and DC. This spectrum is now mixed digitally within the STEL-2130 SSB mixer function with the self contained NCO unit output. The NCO output is shown in Figure 7(d), and the mixing operation produces the spectrum shown in Figure 7(e). Note that we now have a data spectrum at baseband, i.e., at DC. This signal is processed by the integrate and dump filters in the STEL-2130 to produce the output spectrum containing the desired baseband signal and a residual

of the other signals incompletely filtered. An ideal low pass filter would, of course, result in an output containing only the desired baseband signal.

#### Application Example 2: Spread Spectrum Receiver

Another receiver design example is shown in Figure 8. The signal is a QPSK modulated spread spectrum signal with a PN sequence overlaid on each data bit. The data occurs in short bursts. When the data occurs, the signal must be quickly detected and acquired, and the data must then be extracted and stored. The implementation shown is a conventional design achieved without the use of the new STEL-2130 down-converter. It performs well using an I/Q mixer implementation to derive I and Q inputs to the succeeding circuitry. This circuitry consists of a pair of amplifiers, a pair of low pass filters to remove higher order mixer components, and a pair of A/D converters to derive the required quantized input signals to the STEL-3310 matched filter chip used to detect a spread spectrum signal. The STEL-3310 derives a matched filter detected I and Q data outputs and provides them to the following STEL-2120 DPSK demodulator chip which performs the DPSK function by determining the phase difference between successive symbols, provides a carrier frequency discriminator signal, and outputs the processed data and clock signals. The carrier frequency discriminator signal is used to drive an STEL-1175 NCO to the correct frequency. This NCO output, after D/A conversion, amplification and low pass filtering, is used as the LO to drive the input I/Q mixers. A micro controller is used to orchestrate the operational sequence of events as the receiver proceeds through the various phases of signal acquisition, tracking, data detection and collection. A crystal oscillator is provided from which all receiver frequencies are derived.

A second design version for this same exact receiver is shown in Figure 9. A significant overall circuit simplification can be seen as follows. The STEL-2130 down-converter chip has been used to replace all the previously required IF front end hardware. Gone are the I/Q mixers, the dual amplifiers, the dual low pass filters and the two A/D converters. This is desirable for a variety of reasons. Not only is the design easier to understand, change and made to operate as the designer

RPT/M 09z

wishes, it is smaller, lower cost, lower power, and free of the analog imbalance concerns resulting from the I/Q paths, which may require adjustment and balancing during the hardware production testing phase, and can degrade receiver performance. In addition note that the external NCO, D/A, amplifier and NCO used to generate the LO for the input mixers are gone. These have been replaced by the internal NCO, divider circuits and low pass filter all contained within the STEL-2130 down-converter chip. Only a single front end amplifier and A/D converter remain of all the circuitry used in the original design.

The existing frequency spectra within the STEL-2130 are the same as shown for example 1, and illustrated in Figure 7.

This implementation, with a single A/D converter, is limited to receiving signals up to 5 Mbps. If a spread signal with a chip rate of up to 11 Mcps were to be received, the single A/D converter would be replaced with a pair of A/D converters as shown in Figure 10.

#### STEL-9230 Modem Chip Set Evaluation Board

A modem chip set evaluation board is available from Stanford Telecom which can be used to demonstrate and evaluate the performance of the STEL-2130 Down-converter in conjunction with the STEL-2110A Bit Synchronizer and the STEL-2030A Viterbi Decoder. The design is a simple solution to a typical modem design which minimizes the use of analog components. Only the first IF down-converter is implemented in analog form; all remaining signal processing operations are performed digitally realizing the benefits of high precision, repeatable performance, low power and reliability.

#### Block Diagram Description

The block diagram of the STEL-9230 is shown in Figure 11. The unit is designed to accept the input so as not to saturate the receiver front end on broadband noise. The single sideband mixer

RPT/M 09z

function is implemented using a pair of conventional mixers, two 90° splitters and one splitter. A crystal oscillator with a stability of  $\pm 10$  ppm is used to translate the 70 MHz input to an IF of about 3.33 MHz. The use of the single sideband mixer has excellent image rejection thereby simplifying the receiver design by avoiding the need for high resolution filters and multiple IF stages.

The IF signal is then filtered using an active filter centered at 3.3 MHz which provides 40 dB of isolation above 5 MHz. The filtered IF signal is amplified and sampled with a single A/D converter clocked at 4.167 MHz. This is sufficiently high to avoid noise aliasing. The choice of IF frequency permits operation with data symbol rate up to about 256 Ksps. AGC logic at the output of the A/D converter senses whether the signal is at the nominal operating point of the receiver. The digital level is applied to an analog integrator whose output is fed back to control the gain of the IF amplifier which has a 40 dB dynamic range. The AGC loop maintains the input signal properly scaled to the A/D converter range over the entire 40 dB input signal range.

The NCO within the STEL-2130 is tuned to the nominal IF center frequency. The IF signal is then digitally mixed with the quadrature signals from the NCO, and the resultant components are filtered using a digital integrate-and-dump filter. The filtered results are provided to the STEL-2110A bit synchronizer.

Two functions are provided by the STEL-2110A. The first is a complete digital phase locked loop used for the bit synchronization function. An NCO in the STEL-2110A provides the dump clock (data bit clock) which samples the integrate and dump filter of the STEL-2130. In addition, the STEL-2110A contains the symbol timing discriminator and the programmable loop filter to complete the digital phase locked loop. The second function provided by the STEL-2110A is the carrier discriminator function. This function operates on the synchronized I and Q signals and



performs either a frequency or phase discriminator operation whose output is fed back to the carrier loop filter in the STEL-2130 to complete the carrier tracking function.

Once bit sync and carrier phase lock have been achieved, the soft decision data output and the bit clock from the bit sync chip are provided to the STEL-2030A Viterbi decoder. The Viterbi decoder self synchronizes to the symbols and provides a decoded data bit stream. Logic in the STEL-2030A following the Viterbi decoder performs differential decoding to provide an NRZ-L output data format.

The entire signal acquisition process is managed by a simple microcontroller which also manages the interfaces. Inputs are provided to the controller either via switches or an RS-232C interface, to select the mode (BPSK or QPSK), the data rate and whether Viterbi decoding is to be used or not. Based on these control inputs, the controller sets up the bit sync NCO and the interface logic appropriately. During operation the controller determines when signal lock has occurred by monitoring signals within the STEL-2110A and provides status signal to the external interface.

Table 2 summarizes the characteristics of the STEL-9230. Figure 12 illustrates several of the spectra within the receiver.

**Table 1. STEL-2130 Down-converter and NCO Performance Features**

<b>Digital IF Sampling:</b>	Two Input Signal Sampling Modes 1. Dual input I/Q sampling 2. Single input sampling mode provides reduced power consumption mode
<b>Complex Multiplier:</b>	8-bit Complex Data Inputs 40 MHz throughput Truncated 10-bit complex summed products Provides SSB mixer characteristics with minimum magnitude unwanted components
<b>Complex Integrate and Dump Filters:</b>	19-bit Integration Accumulator 40 MHz accumulation rate Internal automatic or external dump control options. Dump operation can be performed asynchronously with respect to input clock timing Automatic dump rate variable in binary multiples of input rate
<b>Numerically Controlled Oscillator:</b>	32-bit phase accumulator 10-bit phase/8-bit amplitude look-up table Sine and Cosine waveform generation Frequency control can be performed in three ways: 1. By the internal loop filter output 2. By an external control input 3. By a combination of both 1 and 2 Microprocessor bus loading of carrier frequency control word 40 MHz maximum clock rate
<b>Versatile Loop Filter:</b>	Selectable first or second order loop filter operation Variable gain setting of frequency or phase discriminator data provided for NCO control 31-bit phase/frequency NCO control signal High carrier recovery loop update rates Internal saturation protection preventing accumulator rollover resulting in erroneous control signals magnitudes
<b>Versatile Input/Output Interface Configuration:</b>	Direct compatibility with other Stanford Telecom devices Single or dual A/D conversion input options Fixed or pseudo-floating point discriminator input signal format Saturated output signal operation prevents rollover and eliminates the possibility of erroneous output signal levels Microprocessor bus carrier frequency reference loading

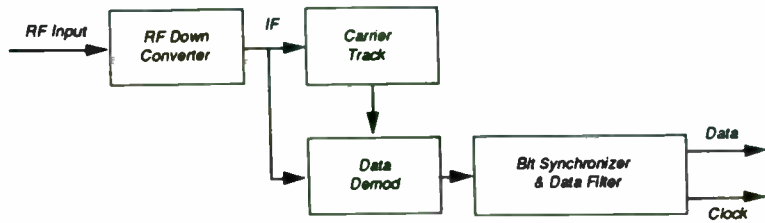


Figure 1. Analog Receiver Implementation

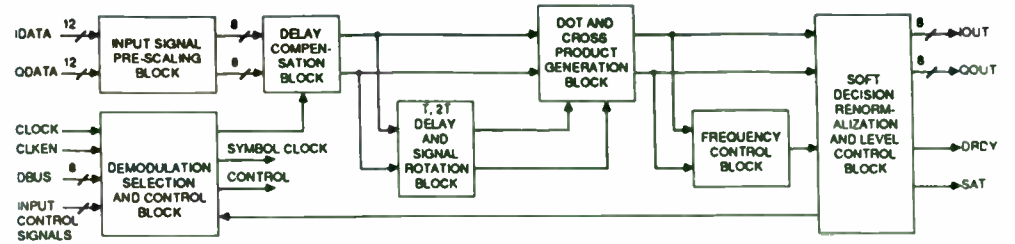


Figure 3. STEL-2120 Digital DPSK Demodulation Block Diagram

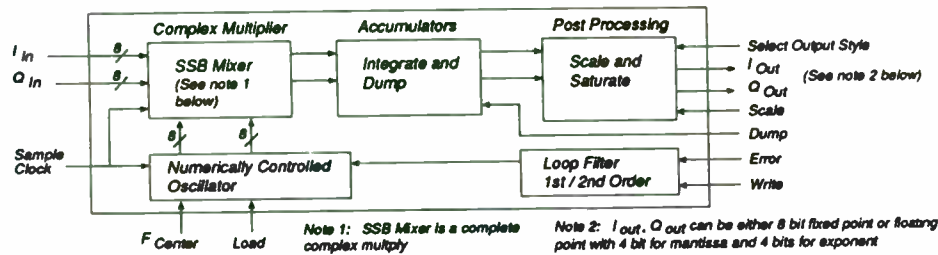


Figure 2. STEL-2130 Down-converter and NCO Block Diagram

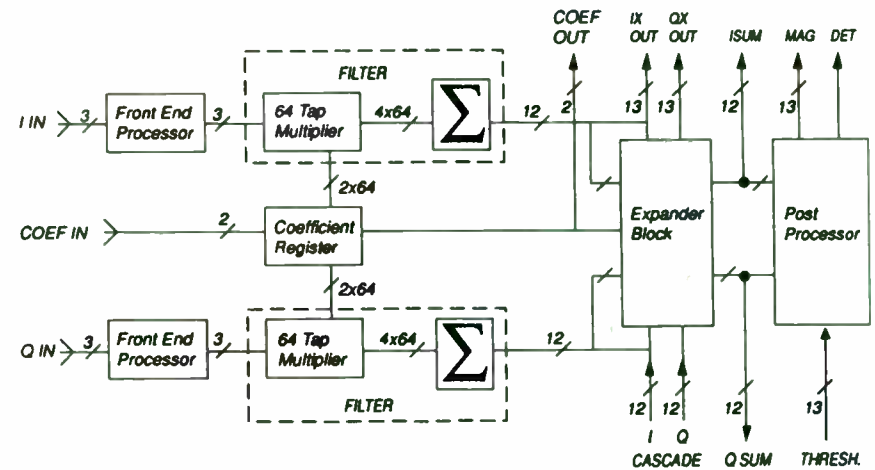


Figure 4. STEL-3310 Dual 64 Tap, 11 Mcps Digital Matched Filter/Correlator Block Diagram

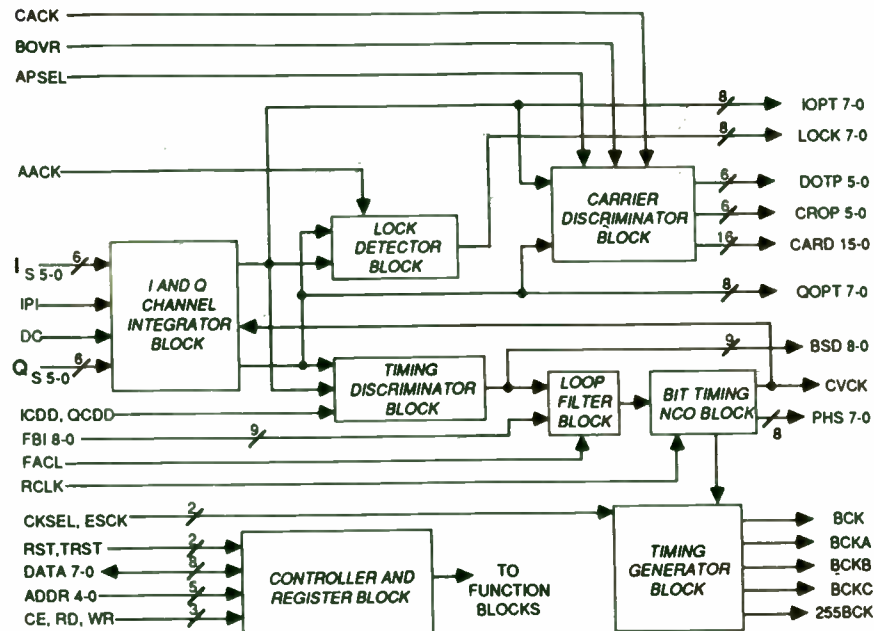


Figure 5. STEL-2110A Bit Synchronizer/PSK Demodulation Block Diagram

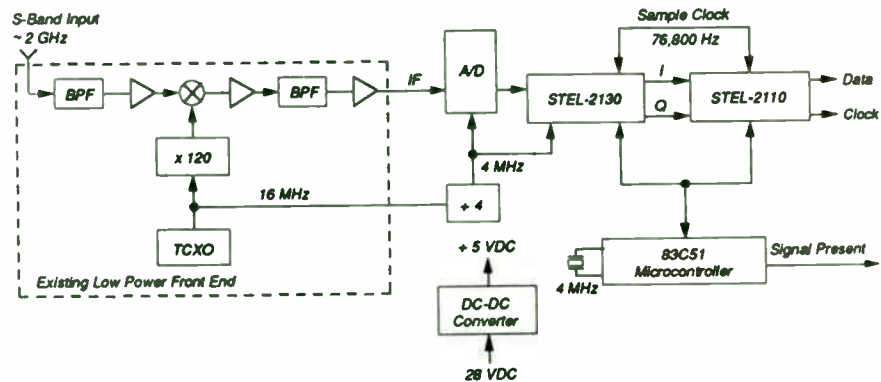


Figure 6. 9600 BPS Command Data Receiver

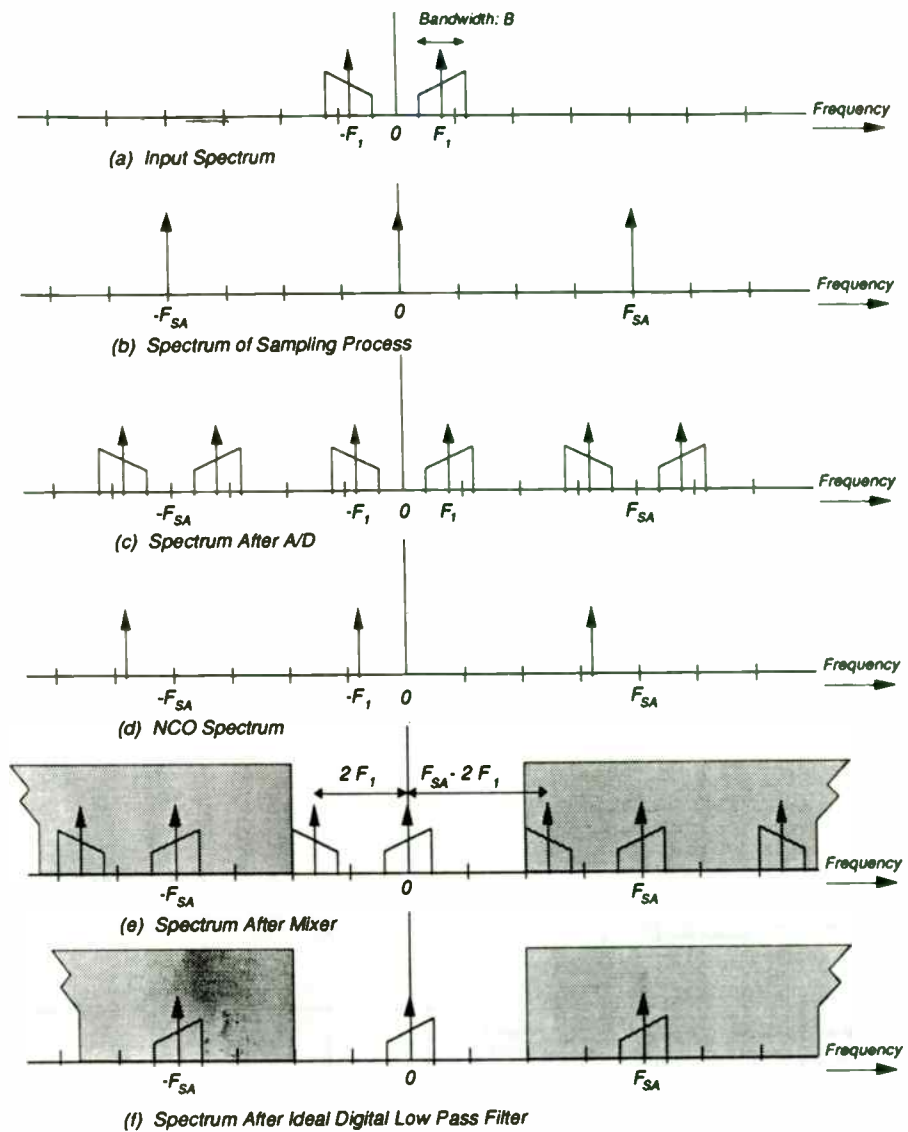


Figure 7. Spectra within STEL-2130 for a narrow band oversampled signal

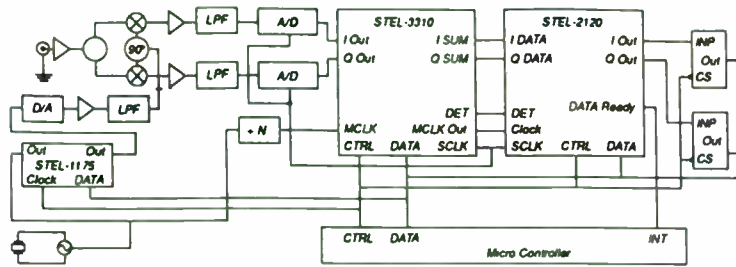


Figure 8. Spread Spectrum Receiver with conventional Down-converter and LO design

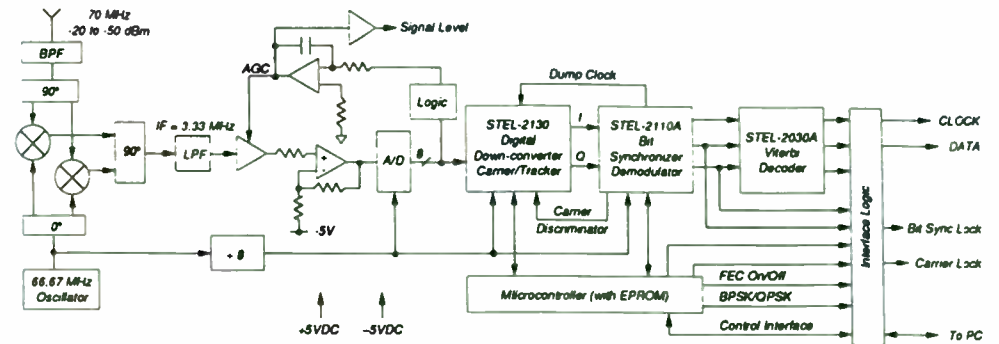


Figure 11. Digital Receiver Block Diagram

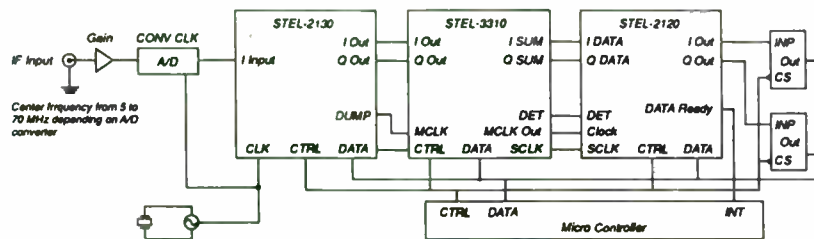


Figure 9. Spread Spectrum Receiver design with new STEL-2130 Down-converter ASIC

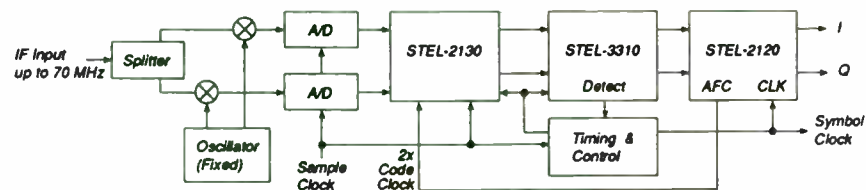
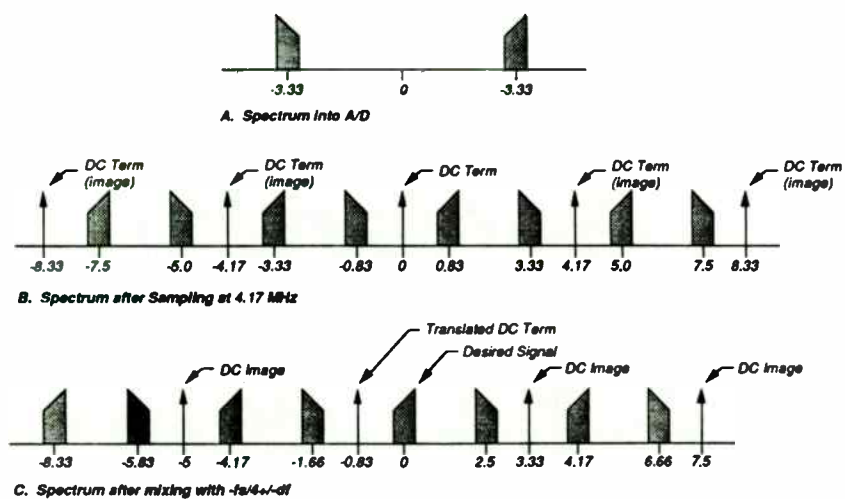


Figure 10. Low Cost Spread Spectrum Demodulator up to 11 Mcps

**Table 2. STEL-9230 Modem Chip Set Evaluation Board Characteristics**

Input Frequency: 70 MHz  $\pm$  100 kHz  
 Input Signal Level: -10 dBm to -50 dBm  
 Input IF-Bandwidth: 1 MHz  
 Modulation: BPSK, QPSK  
 Data Rates: Programmable up to 256 Kbps BPSK or 512 Kbps QPSK  
 Other data rates can be supported with a minimum of modification  
 Coding: 1. No coding  
 2. Viterbi decoding,  $k = 7, r = 1/2$   
 Self Synchronizing  
 5.2 dB coding gain @ BER = 1E-5  
 Carrier Loop Bandwidth: Microcontroller programmable  
 Bit Sync Loop Bandwidth: Microcontroller programmable  
 Data Output Interface: TTL, NRZ-L data format



**Figure 12. Spectrum Waveforms for IF Translation Process**



PUSHING LOW QUIESCENT POWER OP AMPS TO GREATER  
THAN 55dBm 2-TONE INTERCEPT, AND AN AUTOMATED, VERY  
WIDE DYNAMIC RANGE SYSTEM TO MEASURE THESE EXCEPTIONALLY  
LOW SPURIOUS LEVELS.

Michael Steffes  
Comlinear Corporation  
4800 Wheaton Rd.  
Ft. Collins, Co. 80525

#### ABSTRACT

It is commonly expected that very low distortion amplifiers must dissipate considerable quiescent power to achieve their high linearity. With the advent of intrinsically low distortion current feedback op amps, along with the linearity improvements achieved by the negative feedback used in these devices, wideband low distortion amplifiers have become available at much lower quiescent power levels. This discussion will focus on the 2-tone, 3rd order intermodulation distortion of current feedback amplifiers. Following a brief review the harmonic distortion mechanisms of current feedback op amps, a simple means to further improve an already high intercept will be described. Having pushed the 3rd order spurious levels into the noise, an automated means of measuring these very low distortion levels will be described.

#### Harmonic Distortion in a Current Feedback Amplifier

Figure 1 shows a simplified internal circuit for a current feedback operational amplifier. Note that the structure is very symmetric with complementary NPN and PNP devices. The input buffer stage, Q1-Q4, forms an open loop voltage buffer from the non-inverting input to the inverting pin. Transistors Q3 & Q4 provide a means to simultaneously drive the inverting node voltage and cascode an error current signal through their collectors to a current mirror stage. The outputs of the two symmetric current mirror stages are fed back together to form the high transimpedance node for the amplifier. This is the high gain node for the amplifier. Small changes in the error current fed back through Q3 and Q4 will have a significant impedance gain to a voltage at the outputs of the two current mirrors. This voltage is buffered to the output pin by another open loop voltage buffer, transistors Q5-Q8. This buffer's high input impedance contributes to achieving a high forward impedance gain through the amplifier while providing a low impedance output drive. Both the input buffer and the output buffer are essentially Class AB buffer stages (see ref.1 for a more complete description of a current feedback op amp).

To this point, the amplifier's internal elements have been treated from an open loop standpoint. When the output is connected back to the inverting input through a feedback resistor, with a gain setting resistor to ground on the inverting node, we get the closed loop op amp configuration. Figure 2 shows the closed loop current

feedback op amp block diagram along with the resulting transfer function. Looking at the transfer function, reproduced below, the numerator expression is our desired signal gain in Volts/Volts.

$$\frac{V_o}{V^*} = \frac{1 + \frac{R_f}{R_i}}{R_f + R_i \left(1 + \frac{R_f}{R_i}\right) + \frac{1}{Z(s)}}$$

The denominator expression represents the error terms due to a finite forward gain in the amplifier. If the forward transimpedance gain  $Z(s)$  were infinite, this error term would drop out and the amplifier would produce exactly the gain shown in the numerator. The forward transimpedance is a frequency dependent gain, having a very large value at DC with a dominant low frequency pole and higher frequency poles. When the magnitude of the  $Z(s)$  has rolled off to equal the value of the  $R_f + R_i(1+R_f/R_i)$  term, the loop gain has dropped to one and the overall amplifier frequency response begins to roll off. The  $R_i(1+R_f/R_i)$  part of the feedback transimpedance is the principal parasitic effect limiting the amplifier's bandwidth as higher closed loop  $(1+R_f/R_i)$  gains are desired ( $R_i$  is the output impedance of the buffer driving out of the inverting node). For the remainder of this discussion this term will be neglected with the feedback transimpedance being set by only  $R_f$ .

One of the basic advantages to the current feedback topology is that with the loop gain, and hence the frequency response, set

externally by  $R_f$ , the desired signal gain may be set by  $R_g$  with minimal impact on the frequency response. It is through this mechanism that the current feedback op amp is said to offer a Gain - Bandwidth product independence. Please see Comlinear application note OA-14 for a more complete discussion of the current feedback transfer function and frequency response control.

#### Distortion Mechanism's in Current Feedback Amplifiers

From an open loop standpoint, harmonic distortion arises from any non-linearities in going from the the inverting error current signal to the output voltage. Although we have shown this transimpedance gain to be a frequency dependent linear gain,  $Z(s)$ , at any particular frequency,  $Z$  can actually be represented by a (polynomial expression from the error current to the output voltage. This polynomial will have a very high linear gain term ( at low frequencies) with relatively small coefficients for the higher order terms.

The signal path from the inverting input to the output follows two symmetric paths as shown in Fig. 1. The open loop 2nd order coefficient is set by any transfer mismatch's between the upper and lower signal paths to the output. The open loop 3rd order coefficient is principally set by the 3rd order curvature (crossover distortion) in the transfer function of the Class AB output buffer. (See Ref.2, page 694 for a discussion of Class AB buffer distortion)



With the 2nd order distortion arising from mismatch effects, it is often observed that this distortion is strongly dependent on the DC operating point at the output. Changing the relative voltages across the two halves of the forward gain path will effect the balance of the parasitic effects (voltage dependent base-collector capacitance and output impedances) that give rise to this non-linearity. Similarly, for a ground centered, sinusoidal, output swing, imbalancing the power supplies can be used to null the 2nd harmonic at a specific frequency.

The magnitude of the open loop 3rd harmonic distortion term, at a given frequency, is principally a function of the output load current vs. the biasing current,  $I_{b2}$ , in the output stage. Hence, as signal swings go up, load resistors go down, or quiescent biasing current goes down, this third order distortion will increase. Conversely, as the load impedance increases, the signal swing decreases, or the quiescent biasing current increases, this 3rd order distortion will be decreased.

The intrinsic symmetry of the forward gain path, with well matched PNP and NPN signal paths, along with the fully complementary Class AB output buffer, yields low open loop distortion. This distortion is further reduced by the action of the negative feedback when the loop is closed (as shown in Fig.2). At a particular frequency,  $Z(s)$ , can be taken to have specific values for the coefficients of a polynomial approximation to the transimpedance gain from the

inverting error current to the output voltage. Figure 3 steps through a transfer function development using a polynomial expression for the forward transimpedance gain. Although this approach does not yield a closed form solution for the output voltage polynomial vs an input signal, it does illustrate the loop gain dependence of the higher order terms.

A more common way to show the effect of negative feedback on forward gain distortion effects is to introduce an error signal at the output of the forward gain block. Figure 4 shows this control theory approach with a similar result to Fig. 3 - open loop distortion effects are reduced by the loop gain in a negative feedback closed loop configuration. This approach also does not reach a closed form solution for  $V_o$ , ( $V_d$  should actually depend on  $V_i$ ). For a single frequency input, it is the loop gain at the fundamental frequency that is applicable to determining the loop gain's impact on the distortion. Some of the literature seems to imply that it is the loop gain at the harmonics that is acting to linearize the closed loop performance to decrease the distortion (ref 2, page 418). However, testing with a loop gain tailored to be higher at the fundamental than at the harmonics has shown a direct dependence on the loop gain at the fundamental, but not at the harmonics.

For the 3rd order terms then, the achievable harmonic and 2-tone, intermod distortion levels are set by the intrinsic 3rd order distortion of the output buffer and the loop gain at the

fundamental frequency of operation. Figure 5 shows a Bode plot of the frequency dependent forward transimpedance,  $20 \cdot \log(|Z(s)|)$ , for a typical current feedback amplifier. The solid horizontal line intersecting this plot at about 100MHz is  $20 \cdot \log(R_f)$ , the feedback transimpedance. The vertical distance between the forward  $Z(s)$  and this solid horizontal line is the loop gain,  $|Z(s)|/R_f$ .

As is apparent from Fig.5, this loop gain decreases with increasing frequency as the forward transimpedance gain rolls off. The intersection of  $Z(s)$  and the feedback transimpedance is of critical importance in determining the closed loop frequency response flatness. This decreasing loop gain with frequency is the dominant cause for an increase in harmonic distortion with increasing frequency for negative feedback amplifiers. One of the key contributions of the current feedback amplifier is the ability to use higher loop gains to higher frequencies than for equivalent voltage feedback parts. Typically, however, we still see the 3rd order distortion terms starting to increase, for a given output power level and gain, for frequencies above 5MHz. Conversely, as operating frequencies go below about 5MHz, the distortion performance remains constant.

As a point of comparison, this discussion will focus on the 2-tone 3rd order intermodulation distortion at 10MHz. Figure 6 shows a plot of 3rd order intercept (at 10MHz) vs quiescent power dissipation (in dBm) for two current feedback amplifiers along with several other high linearity amplifiers. Many of these other

amplifiers use a Class A output which requires significantly higher quiescent power to achieve low distortion. Also, minimal feedback, and hence minimal distortion improvement due to loop gain, is generally used. This yields a distortion performance that is not nearly as frequency dependent. Basically, these Class A type amplifier's have driven the forward path non-linearities down with high quiescent currents and used minimal feedback to keep their distortion performance constant over a wide frequency range.

The data of Fig. 6 shows that, for HF frequencies, considerably higher intercepts per mW of quiescent dissipation can be achieved with current feedback op amps through the use of a very linear forward gain path and high loop gain in the feedback network. The principal drawbacks to using the current feedback op amp in an HF or RF application is a steadily decreasing intercept above 5MHz, useable bandwidths limited to about 100MHz, and relatively poor noise performance. Intercepts have typically dropped to below 30dBm by 50MHz for the op amps shown in Fig. 6.

#### Improving Distortion by Shaping the Loop Gain

Since a current feedback amplifier allows the loop gain to be set separately from the signal gain, it should be possible to adjust the loop gain to yield an improved distortion performance without changing the signal gain. The simplest way to do this would be to scale the resistor values down, keeping the same ratio for  $R_f/R_g$ . Decreasing  $R_f$  will increase the loop gain over the full frequency range, but runs the risk of inadequate phase margin at the

crossover, where  $|Z(s)| = R_f$ . It would be preferable to decrease the feedback transimpedance at lower frequencies but return to the nominal design value for  $R_f$  where this feedback  $R$  is intended to equal  $Z(s)$ .

Figure 7 shows one possible circuit that achieves this loop gain shaping. This circuit was originally reported in the literature as a means to improve the equivalent input noise. (ref. 3). At low closed loop gain settings, the relatively large inverting input current noise for a current feedback amplifier can dominate the overall noise performance. This noise current shows up at the output pin multiplied by the feedback resistor. The circuit of Figure 7 reduces the feedback resistor by using a parallel combination of the two feedback resistors as the low frequency gain for this noise current. The coupling inductor is set to remove this parallel feedback path before the crossover frequency (where  $Z(s) = R_{f1}$  with approx. 60 degree phase margin).

Figure 8 shows a generalized transfer function for the circuit of Figure 7. In general, this circuit could be used to shape both the loop gain and the forward signal gain. At low frequencies, the gain is set by the parallel combination of the two  $R_f$ 's divided by the parallel combination of the two  $R_g$ 's. At high frequencies, once the inductor has opened up the connection between the two feedback's, the gain is simply  $(1 + R_{f1}/R_{g1})$ . Similarly, the feedback transimpedance at low frequencies is the parallel combination of the two feedback  $R$ 's while at high frequencies it has increased to

equal  $R_{f1}$ . In going from low to high frequencies, this circuit shows a zero/pole pair for both the signal gain and the feedback transimpedance.

For this analysis, we will take the ratio of  $R_{f1}/R_{g1} = R_{f2}/R_{g2}$ . Under this condition, the numerator of the transfer function in Fig. 8 simplifies to equal  $1+R_{f1}/R_{g1}$  indicating a flat frequency response up to the rolloff frequency. The principle concern here is for the frequency dependence of the feedback transimpedance. The loop gain has been increased due to the decreased feedback transimpedance at low frequencies which should provide an improved harmonic distortion performance. The feedback transimpedance now shows a zero/pole pair due to the inductor coupling the two feedback paths together. Figure 9 shows an analysis of the loop gain for this parallel, inductor coupled, feedback.

To use this approach to increasing the low frequency loop gain the following steps would be followed -

1. The desired signal gain would be set
2. A current feedback amplifier appropriate for this gain be selected
3.  $R_{f1}$  is set to the selected op amp's nominal recommended value
4. The pole frequency for the feedback transimpedance is set to be less than nominal crossover frequency for the selected op amp
5. The desired reduction in feedback transimpedance is set.

This will also determine the zero frequency. The pole/zero ratio is equal to the  $R_{f1}/(\text{DC feedback transimpedance})$  ratio

6. The coupling inductor is resolved from the above ratio,  $R_{f1}$  and the desired signal gain (see Fig.9).
7.  $R_{f2}$  and  $R_{g2}$  are resolved from the above ratio and the desired signal gain.
8. The additional loading on the output due to the additional feedback network should be checked to see that it is not significantly changing the intended load.

Figure 10 shows a Bode plot of the same forward transimpedance gain of Fig. 6 with a 4:1 reduction in the DC feedback transimpedance using this paralleled, inductor coupled, feedback. Note that the targeted pole frequency was 80MHz which forces the zero frequency to be at 20MHz. This yields a 12dB (4X) increase in the low frequency loop gain. This loop gain has decreased by 3dB at 20MHz and continues on up to only a 3dB improvement from the nominal  $R_{f1}$  value at 80MHz. The goal here was to be approximately back to an  $R_{f1}$  feedback impedance by the 100MHz unity gain x-over point on the forward transimpedance curve. This 12dB improvement in the loop gain at 10MHz should translate directly into a 6dB increase in the 2-tone, 3rd order intermodulation intercept (a 12dB decrease in the spurious levels for a given output power level).

#### Test Circuit Implementing Enhanced Low Frequency Intercept

Figure 11 shows an example circuit using the paralleled feedback

approach to increasing the loop gain at low frequencies. This circuit also uses a 1:4 step up transformer at the input to improve the Noise Figure (ref. 4). This input stage presents a 50 ohm input impedance in the passband of the transformer and reduces the overall Noise Figure to 7.2 dB for this test circuit. Although this transformer will AC couple the signal path, it is important to remember that the amplifier itself is a true DC coupled device. Since the input is already AC coupled, a 1uF blocking capacitor at the output has been added to strip off any amplifier DC offsets that may be present. The op amp itself presents a very low output impedance. To get into a 50 ohm system, a series, discrete, 50 ohm resistor must be added. The defined measurement point for both gain and intercept is at the 50 ohm load.

The overall gain for this circuit is 40 Volt/Volt (32dB). Figure 12 shows the measured frequency response for this test circuit. Although the CLC221 is specified to provide >165Mhz -3dB bandwidth in the gain of +20 configuration used here, the transformer and the input RC filter limit the bandwidth to approx. 90 Mhz. Significant latitude in trading off gain, bandwidth, and noise are possible when using op amps in these types of applications. Please see Comlinear application note OA-11 for additional discussion of interpreting and using op amp specifications in RF applications.

At 10MHz, the CLC221 without any loop gain shaping, has a 49dBm 3rd order intercept while dissipating only 900mW quiescent power. Using +/-15 volt supplies, the full scale output swing is approximately

+/-5 volts in order to satisfy the 50mA maximum output current into the 100 ohm load. For 2-tone, 3rd order intercept testing, this translates into a maximum 12dBm test power level for each tone of the two tone test at the 50 ohm load. Given a maximum peak to peak swing at the amplifier output, from either a voltage swing or current limit standpoint, the maximum single tone power level at the load is for a voltage swing 1/4 this level. This accounts for the 6dB loss in going through the matching network and the fact that the full voltage envelope for a two tone test is the sum of the peak to peak swings for the individual test frequencies. Pushing a full 12dBm in each test tone at the load puts the amplifier into a slightly higher distortion mode. The 49dBm intercept for the amplifier by itself is not observed until the single tone power at the load has dropped to 8dBm. Using this as a test condition will yield a 3Vpp swing for the voltage envelope at the load.

This test circuit shows bipolar supplies for the amplifier. True DC coupled devices typically use balanced bipolar supplies. However, since most current feedback amplifiers don't actually use a ground reference in their design, single supply operation is perfectly acceptable. Comlinear application note OA-11 describes this in detail.

#### **Measuring 2-Tone, 3rd Order Intercepts above 50dBm**

The 3rd order spurious levels for 8dBm test power levels and 50dBm intercept would be -76dBm. This 84 dB dynamic range requirement is

right at the edge of what most spectrum analyzers can measure. Figure 13 shows a typical test setup for measuring intercepts where the dynamic range requirement is less than 90dB. One of the principle requirements for this measurement are to provide clean input test tones, with no intermodulation of the sources. The post amplifier's after the sources and the hybrid power combiner achieve this quite well. These amplifiers, the CLC142, can provide 27dBm output power through 100MHz and are particularly useful in getting enough power to the DUT for testing low gain devices. They also isolate any leveling loops in the output stage of the sources from mixing together and would themselves provide in excess of 50dBm intercepts through 10MHz.

The 2nd primary requirement for the test setup of Fig. 13 is to attenuate the fundamental power levels at the input of the spectrum analyzer mixer to a low enough power level to eliminate any analyzer produced intermodulation products. This is typically a -36 to -40 dBm power level at the mixer.

The test set up of Fig. 13 would begin to have trouble making this measurement as the required dynamic range extends beyond 85dB. This would occur as the test power level is decreased (to evaluate if intercept performance is observed) or if higher intercepts were to be measured. The anticipated 55dBm intercept for the test circuit described earlier would exceed the measurement range for the test configuration of Fig. 13. With 8dBm test power levels at the load, the anticipated spurious power levels would be at -86dBm. This 94dB

dynamic range is probably beyond most spectrum analyzers. Averaging could be used to lower the noise floor in an attempt pull this very low spurious out of the noise. As the test power is decreased, the required averaging would significantly extend the test time. An alternative technique would be to filter out one or both of the two test frequencies before the signal is applied to the analyzer. This realistically requires a relatively broad spacing of the test input frequencies (to avoid filtering the spurious levels) and is not particularly useful for swept frequency measurements.

Figure 14 shows a modification to the basic test setup of Fig. 13 to extend the dynamic range of this intermodulation measurement system. In this approach a 3rd signal source, phase locked to one of the other two input signal sources, is used to null out that one test frequency at the output of the DUT. The same type of power combiner used at the input is used here to combine the output signal with a nulling signal from a 3rd signal generator. The Fluke 6080A sources offer a programmable phase capability that allows the nulling source to be tuned to very nearly 180 degrees out of phase with one of the test signals at the DUT output. The output of this 2nd power combiner will then have one of the test frequencies plus the original intermodulation tones at the output of the DUT. This allows the system to use less attenuation to the analyzer mixer since no intermodulation terms will be generated in the analyzer. This technique also offers the advantage of being fully programmable over a wide range of frequencies and test powers.

#### **Test Results for the Improved Intercept Amplifier**

The intermodulation intercept for the circuit of Figure 11 was first measured using the basic setup of Figure 13. This measurement resulted in approximately a 52dBm intercept. Figure 15 shows the test signals at +/- 100Khz around 10MHz. Note that the measured power is -27dBm or -37dBm at the mixer considering the internal 10dB attenuator. Figure 16 shows the measurement for the upper spurious signal at 10.3Mhz. Note the very narrow resolution bandwidth, video bandwidth, and span to make this measurement. The low phase noise of the sources and phase locking all of the sources and the analyzer together are critical to this measurement.

Continuing on to the same measurement with the test set up of Figure 14 did allow a slightly increased measurement range. After nulling out the lower test frequency the test signal plot of Figure 17 shows the widely disparate power levels going into the analyzer. It is important to remember that equal test power levels are still being generated at the DUT output. Note that the measured power on the un-cancelled test tone has increased to approximately -11 dBm from the earlier -27dBm level. This reflects the reduced (19dB) attenuation used at the output signal path. Also note that the lower test tone has been attenuated by 40 dB from the upper tone with this cancelling technique. Although the mixer is seeing a fairly high absolute power level for the upper tone, -21dBm, since the lower tone is significantly lower, no analyzer generated intermodulation terms should interfere with this measurement.

Figure 18 shows the measured lower spurious at 9.7MHz using the test setup of Fig. 14. This measured spurious level is about 8dB more out of the noise than the earlier measurement. However, the noise floor has clearly come up along with this reduced attenuation from the DUT output to the analyzer input. This indicates that the DUT output noise is a significant part of the total noise at the analyzer. Taking the measured test tone power to be at -11dBm, while the DUT output power at the load was 8dBm, this yields a  $(-11 - (-98.5))/2 + 8 = 51.75\text{dBm}$  intercept.

The loop gain shaping network of Figure 11 appears to have improved the intercept by only 3dBm instead of the 6dBm expected. Further investigation revealed that this can be attributed to the non-zero inverting input impedance of the current feedback amplifier. With this impedance greater than 0, some of the feedback error signal splits off and is wasted through the gain setting resistors. This reduces the loop gain. Simulations including this effect revealed only a 6dB improvement in loop gain which is consistent with the 3dBm increase in intermodulation intercept.

However, the intercept does continue to improve as the test frequency is decreased. The following tables shows the measured intercepts from 5MHz to 10MHz for the circuit of Fig.11 using the test setup of Fig. 14.

Frequency	Intercept
5Mhz	55.8 dBm
6MHz	54.9 dBm
7Mhz	54.0 dBm
8Mhz	52.8 dBm

9Mhz	52.4 dBm
10Mhz	51.5 dBm

Below 5MHz a 55dBm intercept is achieved.

#### Conclusions:

High speed current feedback amplifiers can offer exceptionally 2-tone 3rd intercept performance at relatively low quiescent powers. This intercept performance does decrease with frequency due to the decreasing loop gain as frequency is increased. Although the example device used here, the CLC221, is a high performance hybrid amplifier, similar results at lower maximum output power levels can be achieved with monolithic current feedback amplifiers (such as the CLC409 and CLC401 particularly). A simple loop gain shaping network can be used to further increase the intercept at low frequencies. And finally, a simple means to extend the measurement dynamic range through output test signal cancellation has been described and demonstrated. This approach offers the principle advantage of being easily programmable over a wide range of frequencies.

#### References:

1. "Current Feedback Amplifiers", Sergio Franco; reprinted in Comlinear 1991 Databook as application note AN
2. Integrated Electronics: analog and digital circuits and systems, Millman & Halkias; McGraw/Hill 1972
3. "T Network Quiets Current-Feedback Op Amps", Howard Bandell; EDN, Aug.20,1990 pp152
4. "Improving Amplifier Noise Figure for High 3rd Order Intercept Amplifiers", Michael Steffes; Comlinear application note OA-14

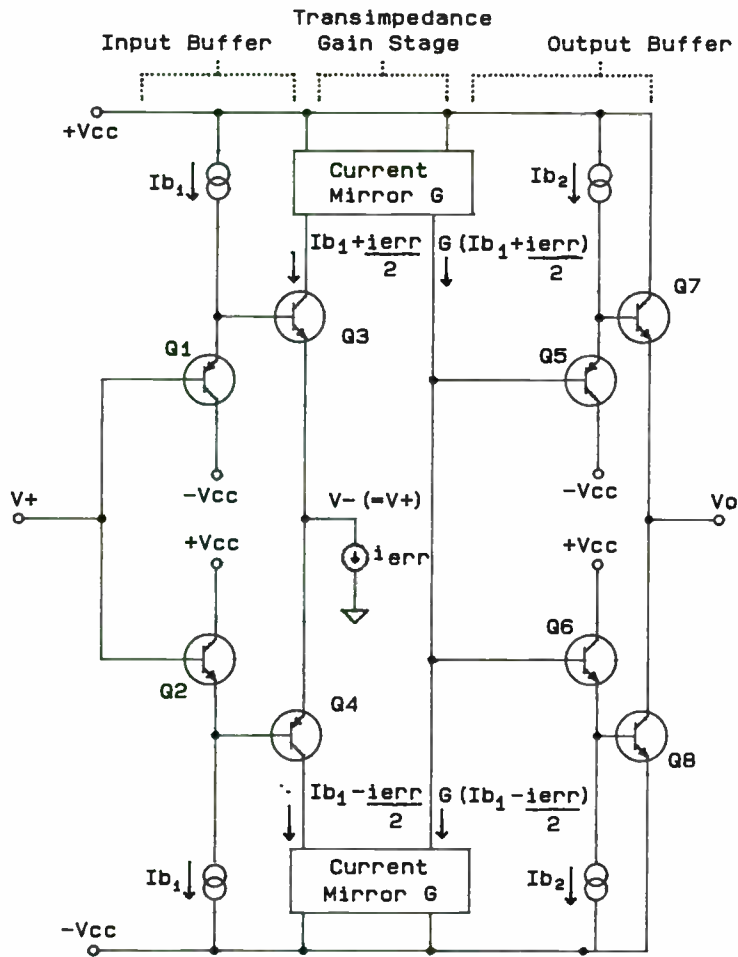
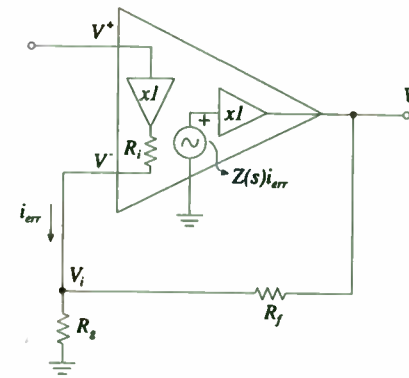


Figure 1. Simplified Current Feedback Topology



$$\frac{V_o}{V^+} = \frac{1 + \frac{R_f}{R_s}}{R_f + R_s \left(1 + \frac{R_f}{R_s}\right) + \frac{1}{Z(s)}}$$

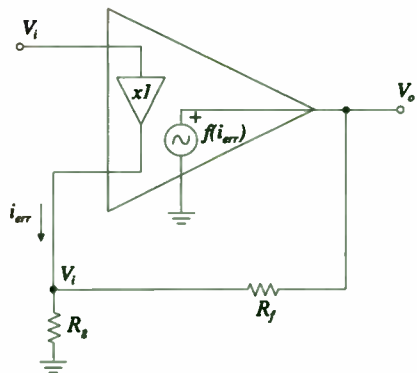
Where

$$1 + \frac{R_f}{R_s} \quad \text{desired non-inverting gain}$$

$$\frac{Z(s)}{R_f + R_s \left(1 + \frac{R_f}{R_s}\right)} \quad \text{frequency dependent loop gain}$$

Figure 2. Closed Loop Transfer Function





$$f(i_{err}) = Z_1 i_{err} + Z_2 i_{err}^2 + Z_3 i_{err}^3 + \dots$$

$$= V_o$$

$Z_1 = \text{linear transimpedance}$   
 $Z_2 = \text{2nd order gain}$   
 $Z_3 = \text{3rd order gain}$

Summing currents at the inverting node

$$\left[ \frac{V_i}{R_s} = i_{err} + \frac{V_o - V_i}{R_f} \right] R_f \quad , \text{ multiplying through by } R_f \text{ and grouping terms}$$

$$V_i \left( 1 + \frac{R_f}{R_s} \right) = R_f i_{err} + V_o$$

From the above expression for  $V_o$  (using first three terms)

$$i_{err} = \frac{V_o}{Z_1} - \frac{Z_2}{Z_1} i_{err}^2 - \frac{Z_3}{Z_1} i_{err}^3$$

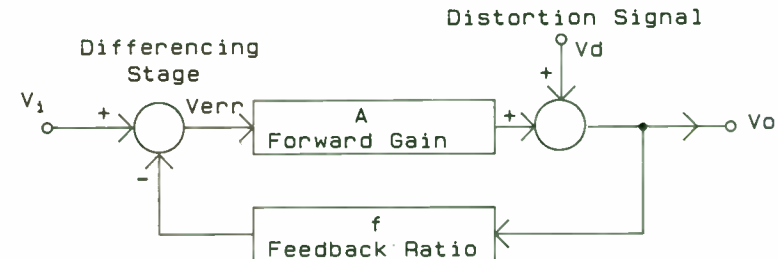
Should isolate and solve for  $i_{err}$  polynomial here, but this doesn't yield a very clear result. Simply putting this into above equation

$$V_i \left( 1 + \frac{R_f}{R_s} \right) = V_o \left( 1 + \frac{R_f}{Z_1} \right) - R_f \frac{Z_2}{Z_1} i_{err}^2 - R_f \frac{Z_3}{Z_1} i_{err}^3$$

Then, solving for  $V_o$

$$V_o = V_i \frac{1 + \frac{R_f}{R_s}}{1 + \frac{R_f}{Z_1}} + \frac{Z_2}{Z_1} \frac{R_f}{R_f} i_{err}^2 + \frac{Z_3}{Z_1} \frac{R_f}{R_f} i_{err}^3 \quad \text{where } \frac{Z_1}{R_f} = \text{Loop Gain}$$

Figure 3. Loop Gain Effect on Non-Linearity



$$V_o = A * V_{err} + V_d$$

$$V_{err} = V_i - f * V_o$$

$$V_o = A * V_i - A * f * V_o + V_d$$

$$(1 + A * f) * V_o = A * V_i + V_d$$

$$V_o = A * V_i / (1 + A * f) + V_d / (1 + A * f)$$

$$A * f \equiv \text{Loop Gain}$$

Figure 4. Control Theory Model of Distortion

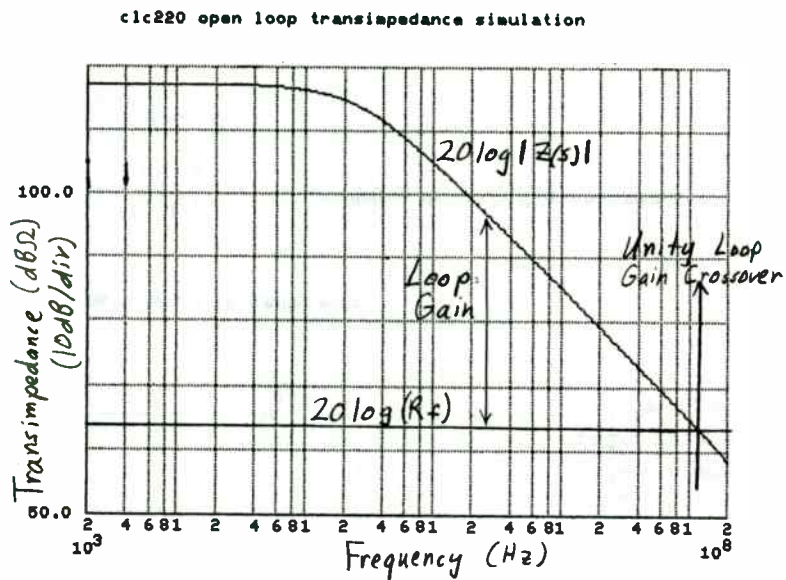
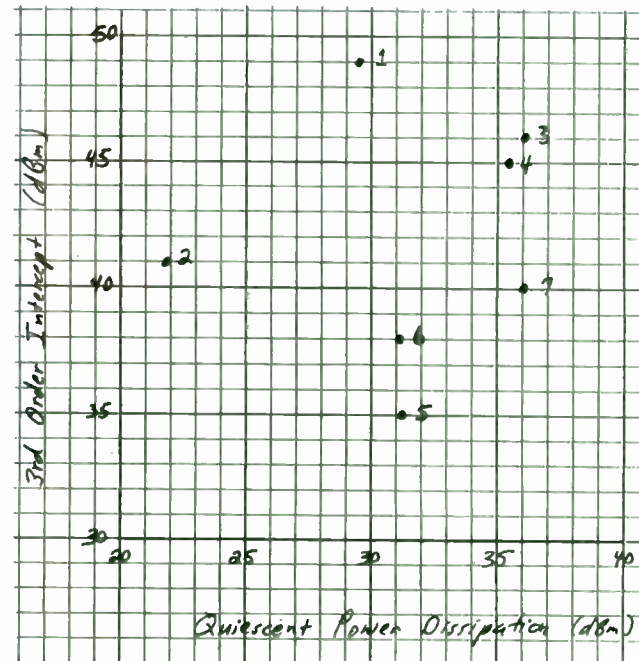


Figure 5. Forward + Feedback Transimpedance



- |                           |            |
|---------------------------|------------|
| 1. Comlinear              | CLC221     |
| 2. Comlinear              | CLC401     |
| 3. Advanced Milliwave Lab | AR01003252 |
| 4. Adams Russell          | AM-109     |
| 5. Avantek                | UTO-509    |
| 6. Watkins Johnson        | WJ-A59     |
| 7. Q-BIT                  | QB-210     |

Figure 6. 2-tone, 3rd Order Intercept vs. Quiescent Power (in dBm)

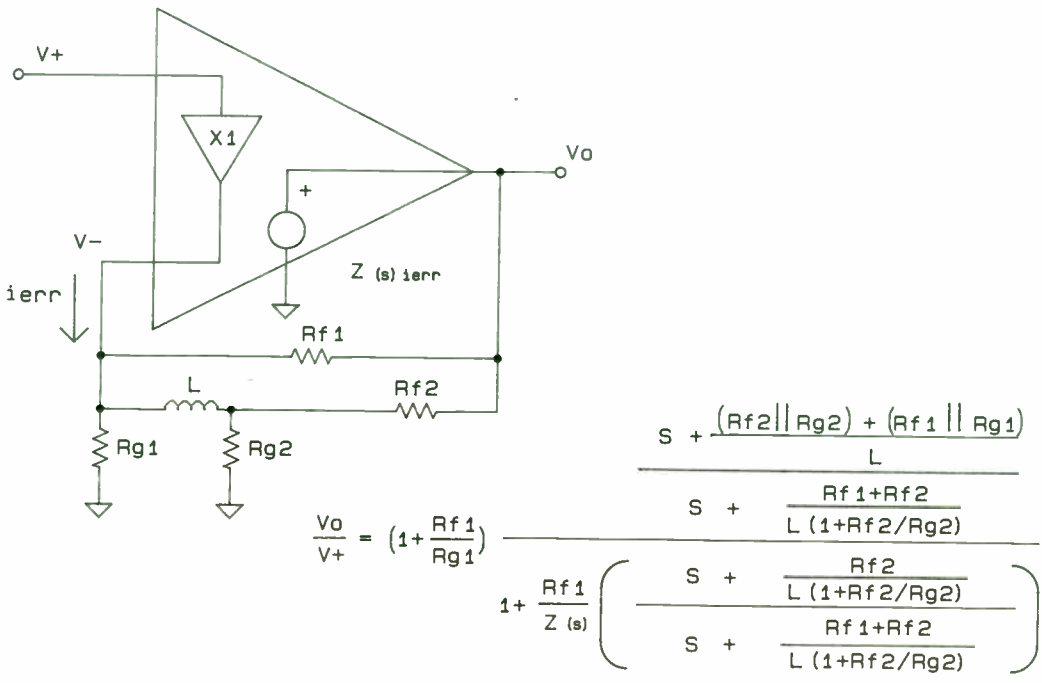


Figure 8. Transfer Function for Parallel, Inductor Coupled, Feedback

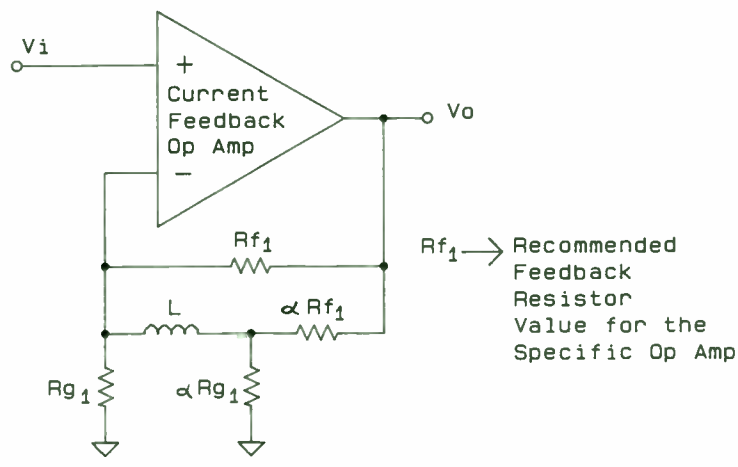


Figure 7. Loop Gain Shaping Network

With  $1 + \frac{R_f}{R_a} = 1 + \frac{R_f}{R_b} = A_v$ ,  $R_f = \alpha R_a$   
 Normally  $\alpha \leq 1$

$$\text{Loop Gain} = \frac{Z(s)}{R_f} \left( \frac{s + \frac{R_a(1+\alpha)}{L_a}}{s + \frac{R_a \alpha}{L_a}} \right)$$

Let desired ratio of D.C. feedback transimpedance to  $R_{f1}$  be  $\beta$

$$\text{required } \alpha = \frac{\beta}{\beta - 1}$$

$$\text{Resulting } \frac{\text{zero}}{\text{pole}} \text{ ratio} = \frac{\alpha}{\alpha + 1}$$

Typically would target  $\beta$  and pole frequency to be less than nominal crossover frequency for  $R_{f1}$  (recommended value for amplifier)

$$\text{Then } L = \frac{R_f \alpha}{\text{Pole frequency} * A_v}$$

Figure 9. Loop Gain Analysis for Parallel, Inductor Coupled, Feedback

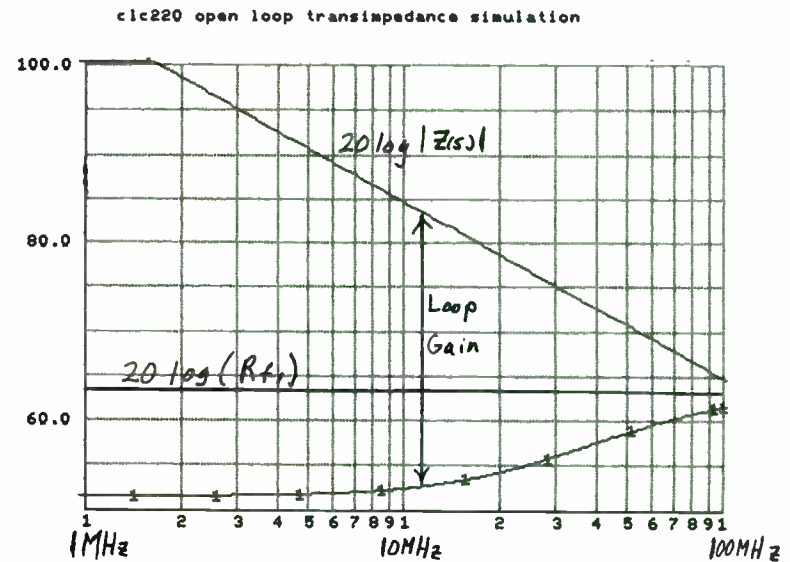


Figure 10. Forward Transimpedance with Shaped Feedback showing Increased Loop Gain

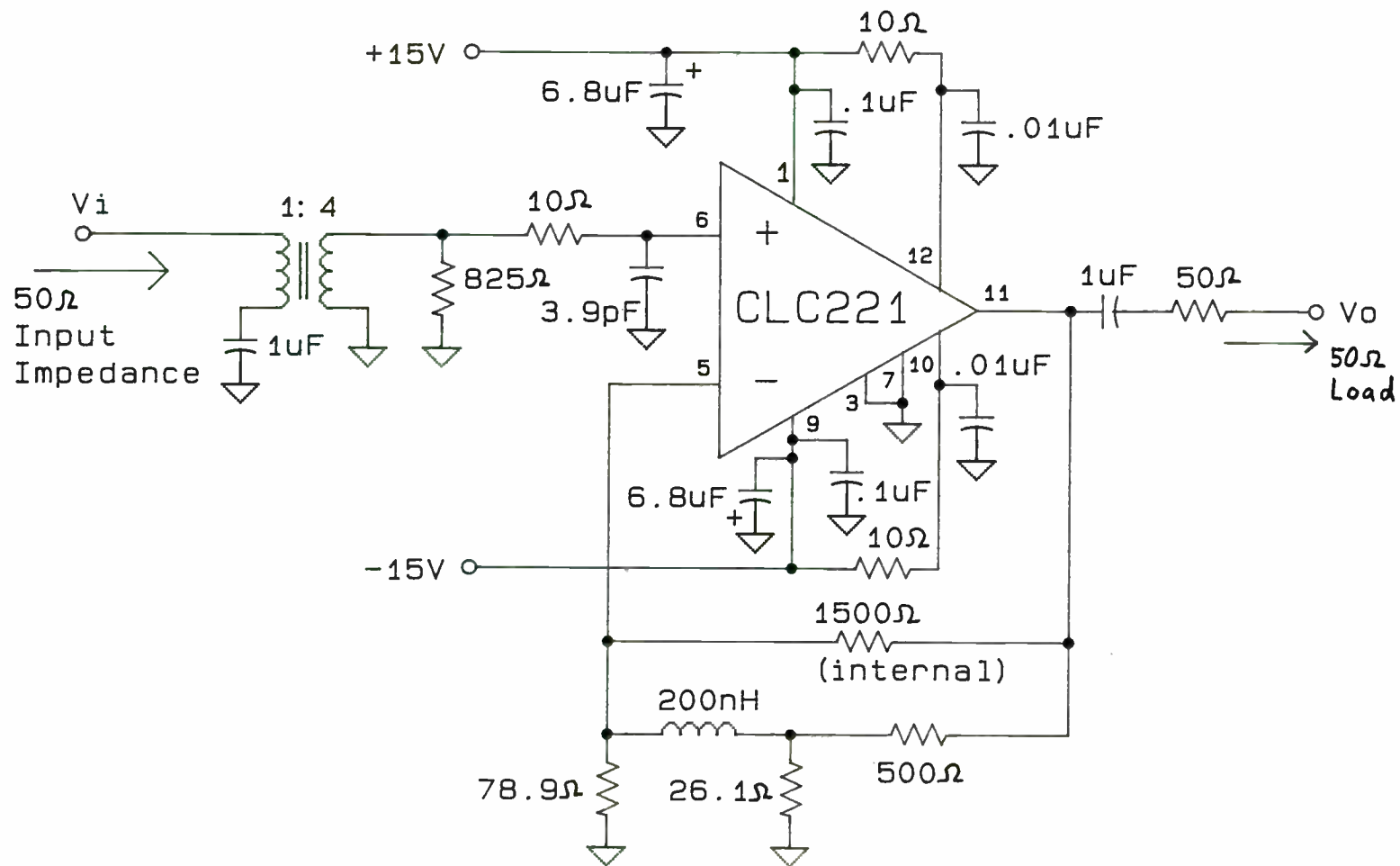


Figure 11. Test Circuit for Increasing Loop Gain

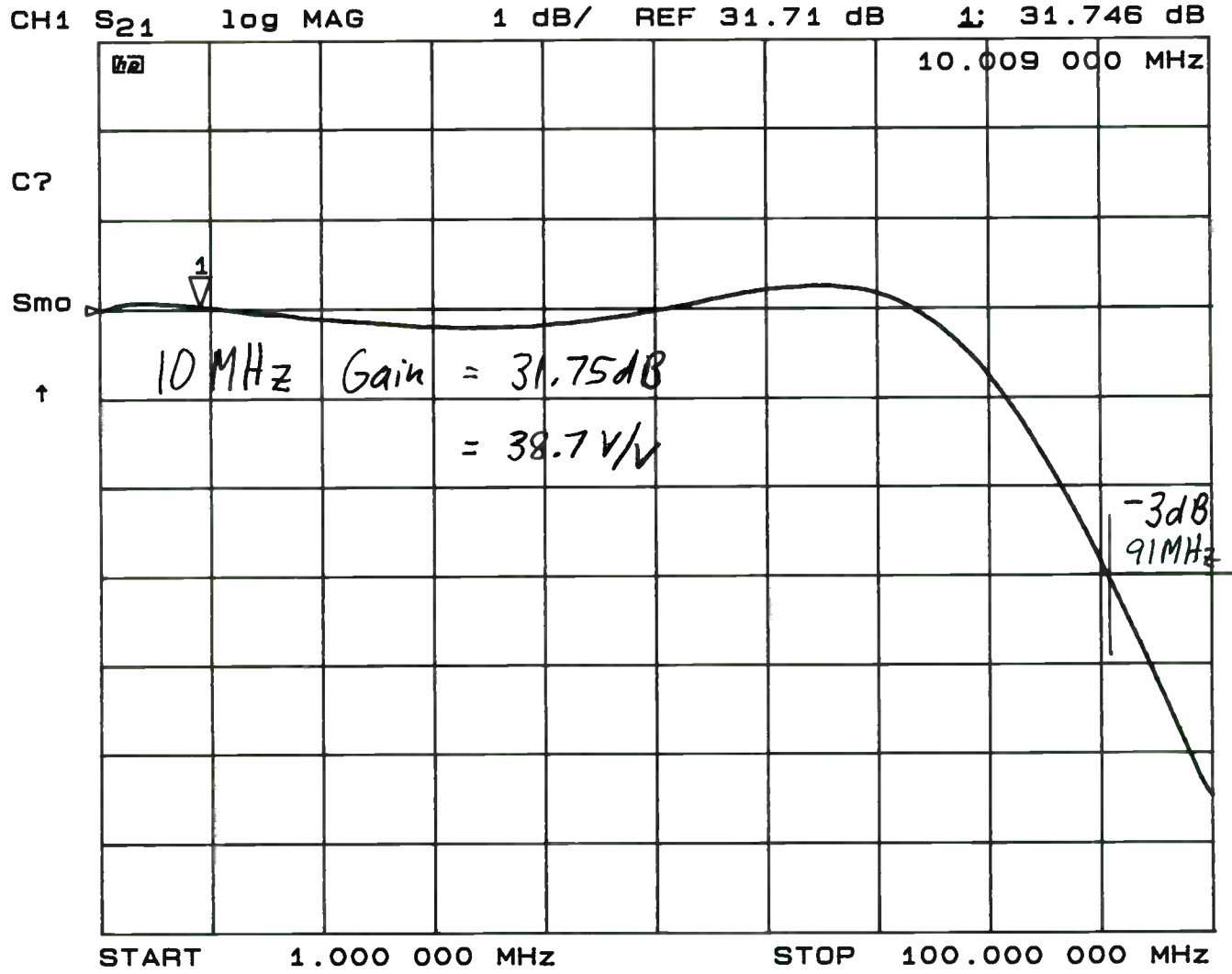


Figure 12. Measured Frequency Response for Wide Dynamic Range Test Circuit using a CLC221

Low Frequency (100KHz to 10MHz) 3rd Order Intermodulation Intercept Test Setup Template

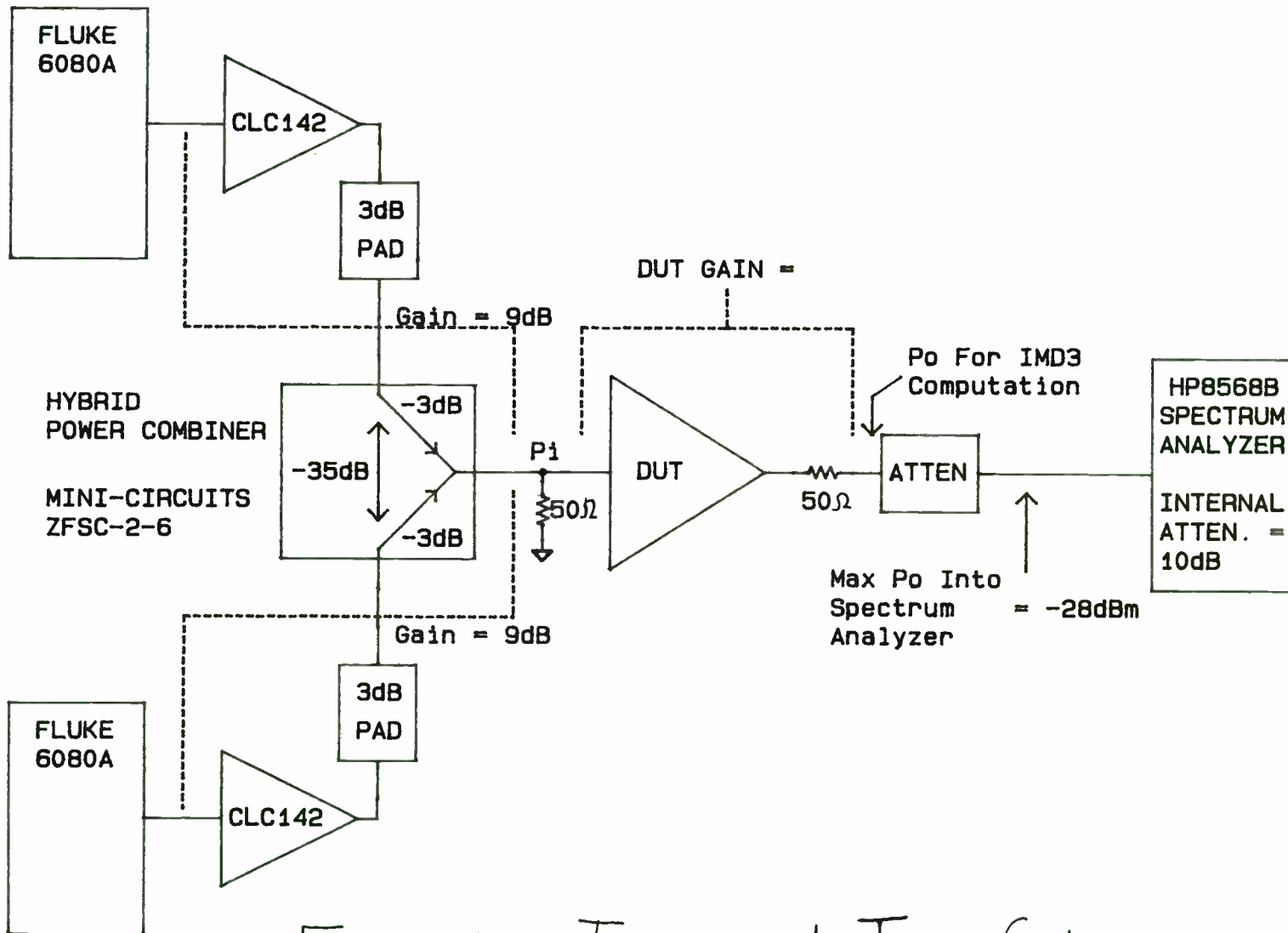


Figure 13. Intermod Test Setup

Low Frequency (100KHz to 10MHz) 3rd Order Intermodulation Intercept Test Setup Template  
With Output Single Tone Cancellation

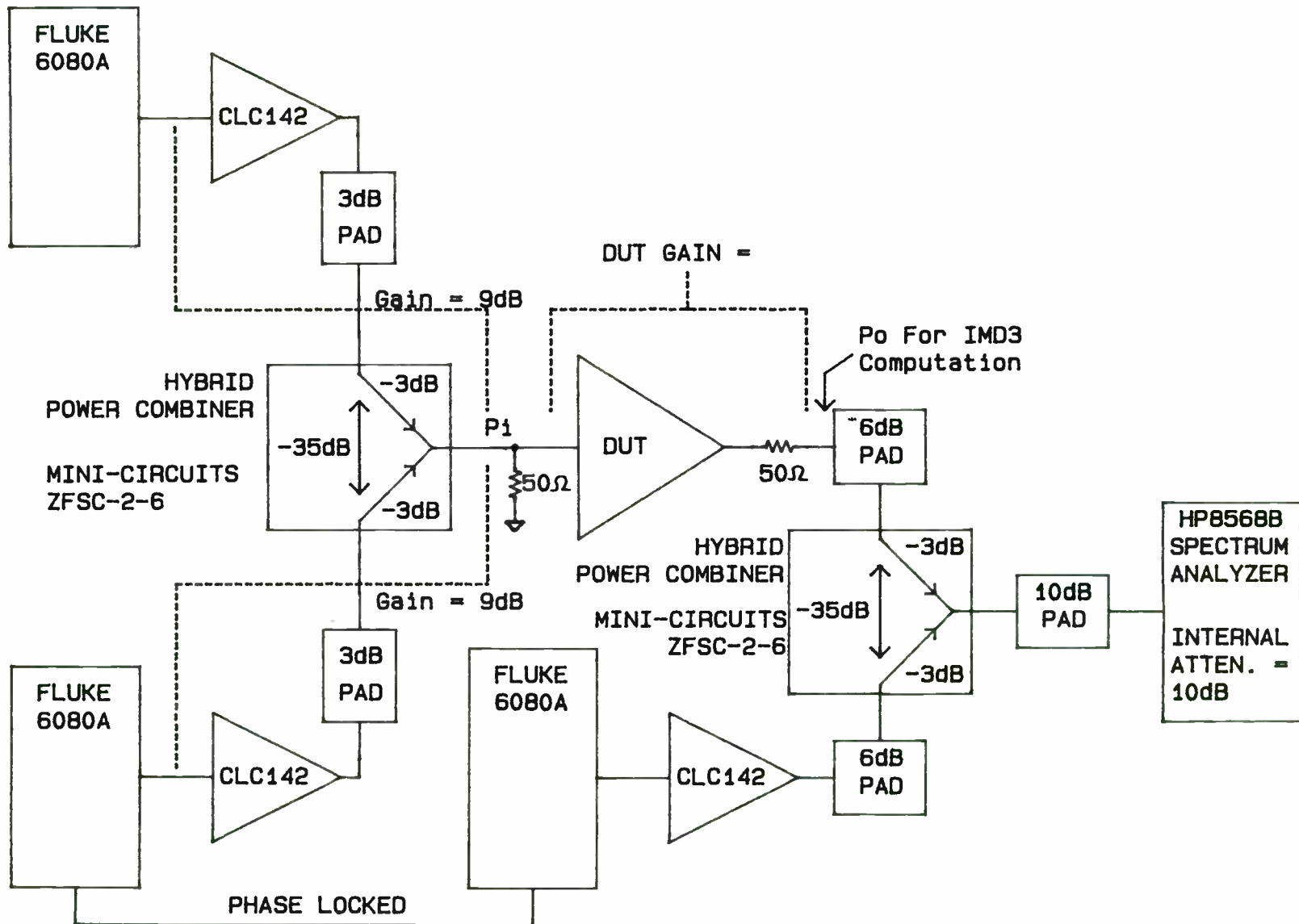


Figure 14. Improved Dynamic Range Intermod Test Setup



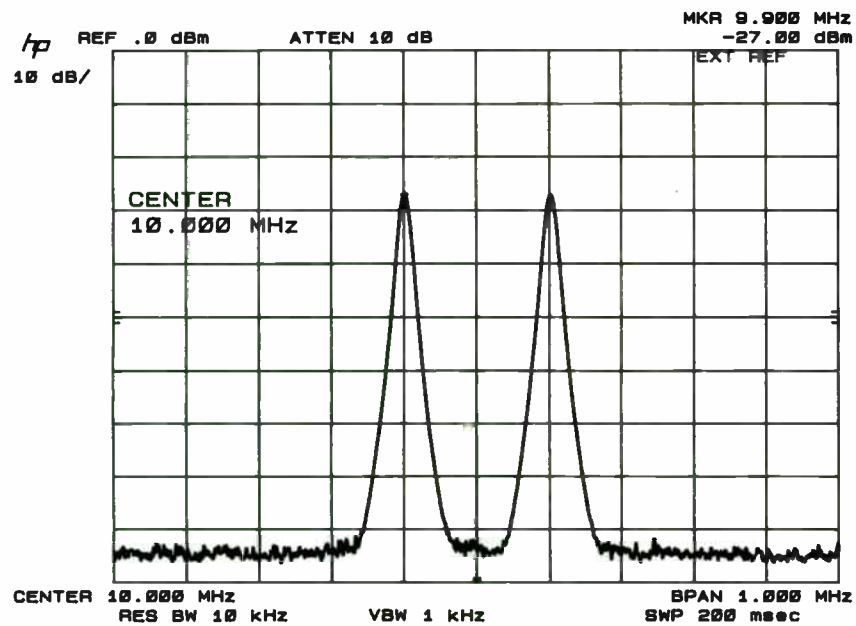


Figure 15. Test Power Levels for 10MHz Test

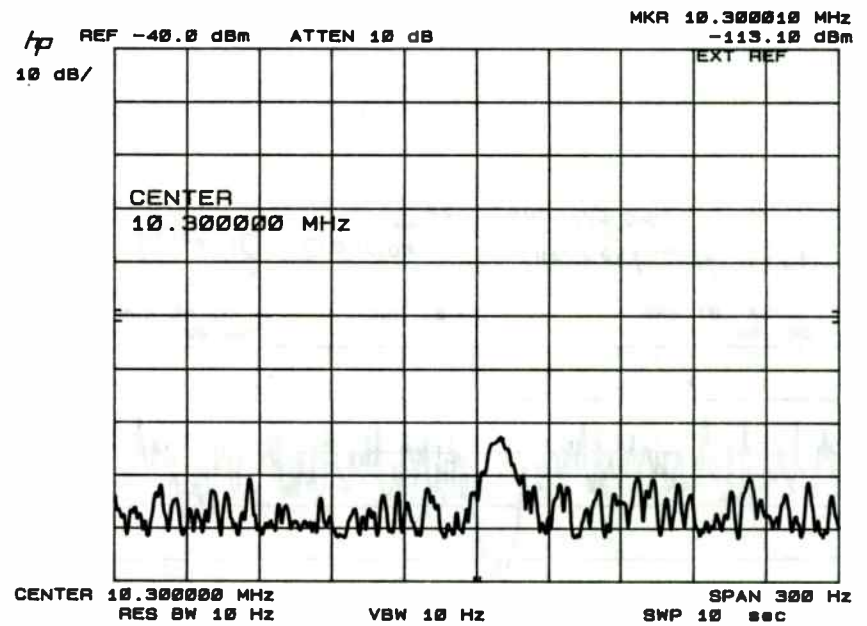


Figure 16. Spurious Measurement at 10.3MHz

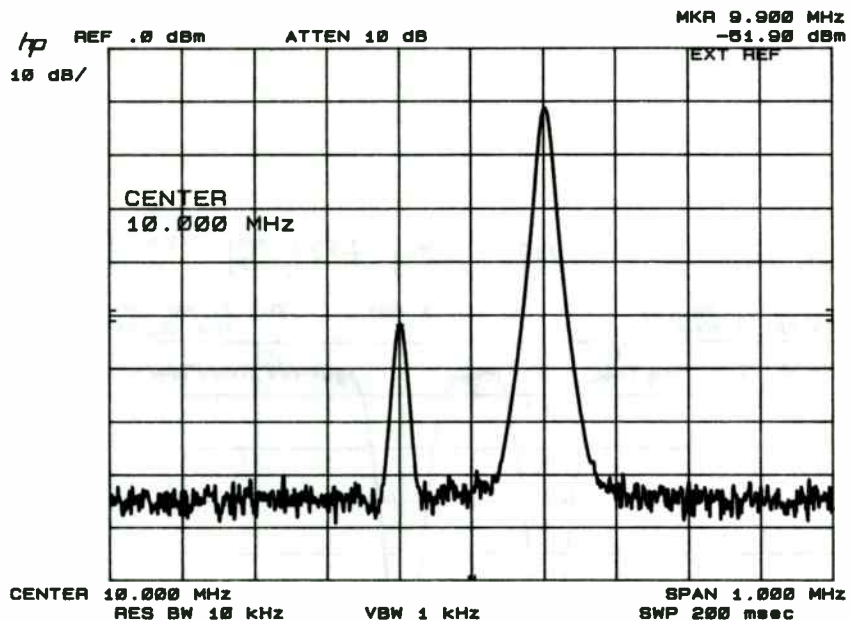


Figure 17. Test Power Levels at the Analyzer with Single Tone Cancellation

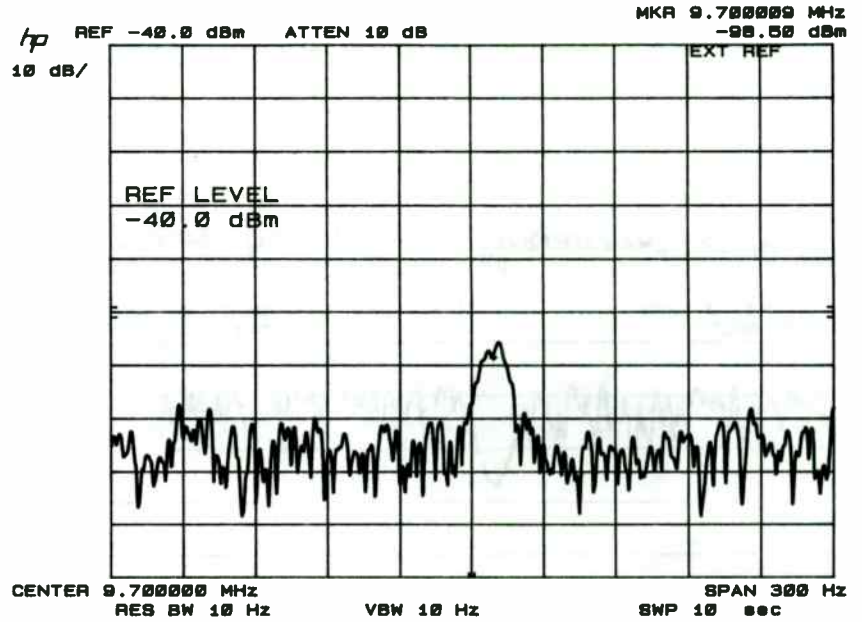


Figure 18. Spurious Measurement with single Test Tone Cancellation.

## A 1Kw. RF. Amplifier for UHF 470-860Mhz

### Television Applications

Michael J. Culling  
Harris Allied  
515, Coldhams Lane  
Cambridge, CB1 3JU.  
United Kingdom

#### Abstract

This paper will describe the evolution of a solid-state 1Kw. broad-band amplifier for television applications. Previous design and manufacturing techniques will be re-visited and refined where necessary, in order to produce a cost effective modular solution.

Ratings and parameters given in the paper will assume the use of a negative modulation system where the television line synchronising pulse tips are at maximum (peak) power, 'black level' is at 56-58% of carrier power depending on the transmission system and the video information is modulated downward onto the RF. carrier with 'peak-white' near cut-off.

#### Introduction

The potential benefits of all solid state transmitters have been well considered and documented over the last five years however, prior to around mid 1989 it was thought in Europe that launching into such designs would not be commercially viable. Two factors were central to this view:

- 1) The highest transistor power in the market place useable at the high frequency end of the UHF television band was only 50 - 60W
- 2) The device cost of around 8.5 \$/Watt of output power.

A combination of these factors would have produced amplifiers having a high component count and which would not have been cost effective alternatives to the thermionic solutions of the day.

As activity in the semiconductor industry increased, it was decided to enter the market with a modular design concept, furthermore, our view was that the minimum capability of such a module should be 1Kw of peak sync power and, in order to minimise combining losses, this power should be sourced from not greater than eight transistors.

To have some margin for losses, such a target imposes a requirement of about 150W peak sync power per transistor.

Between early to mid 1989 and following efforts by various semiconductor manufacturers, it became evident that such power levels were achievable from both FETs and bipolars with power gains in the range 6 - 11dB.

In the first instance FETs were favoured by Harris because of their higher gain and more simple bias requirements. Regretably however, it became apparent that with increasing frequency, the low level linearity performance compared to that of a bipolar, prohibited their use.

A further required attribute at the start of design was the need for full UHF band coverage without the need for retuning.

Some transistor manufacturers were able to realise output powers in excess of 200W peak sync but as a consequence the terminal impedances of these devices were so low as to make impedance matching over the full UHF band near impossible.

The final choice for output transistor went to Motorola who, with their internal package matching were able to demonstrate full band capability at 150W peak sync.

Prior to this development, the highest power UHF RF amplifier developed at Harris was a unit producing 150W peak sync from four transistors and again having full UHF band coverage. The printed circuit for this amplifier was

fabricated on 0.032in. double clad PTFE laminate which was screwed down onto the flat upper surface of a conventional heatsink extrusion. Differential expansion and warp and twist in the laminate provoked RF eddy currents in the ground plane of the printed circuit resulting in amplitude disturbances in the amplifier response. From this experience, for the new development, metal backed laminates were employed to remove any such tendency and to provide mechanically rigid circuit assemblies.

#### **General construction**

To minimise combining losses and to provide redundancy within the module, the output circuitry was arranged into two groups of four transistors with two separate pre-amplification chains, the RF input signal being split between the two groups. See fig.1 Such an approach creates two electrically identical circuit assemblies each having two outputs and one input and which are mounted together with heatsinks in a plug-in mainframe assembly.

At the output, four high power self mating connectors were used per module and the RF input signal was brought in via a standard 'Mini-quick' connector. Having laid out some basic circuit requirements, it was then necessary to consider the means of interconnecting the individual amplifying sections.

One approach is to design and build each half of the module as a series of sub-amplifiers and passive hybrids and to provide hardwiring and RF connections between them. Such an approach would lend itself to flowline amplifier manufacturing where each sub unit could be tested against some specification. In the event of catastrophic failure of or damage to a sub-assembly it could be removed from the module and replaced with another.

Applying some value engineering principles to this philosophy however, it can be shown to have some negative aspects:

1) 'Building block' assembly practices are less efficient in terms of raw materials, parts handling and allocation.

2) RF interconnections are labour intensive to manufacture and require accurate phasing.

3) Heatsinking of separate sub-assemblies may become costly.

4) Mechanical fixing of individual parts may become costly.

If however, the prospect of damage to a one-piece printed circuit by whatever cause could be removed and sufficient monitoring provided for fault finding, it then becomes possible to incorporate each group of four transistors with their pre-amplifiers onto a more cost effective one-piece amplifier assembly. It was decided to opt for this course.

The laminate chosen for the design was 0.020in glass loaded PTFE on a 1/16in copper backing, adopted for the reasons given earlier. Initially, it was intended to use aluminium backed laminate, but the advice of the time from the processing houses was that plated through holes into aluminium was a costly process and had been shown to be unreliable.

#### **Output power generation**

Before starting on the details of circuit design, it was necessary to fully evaluate the transistors developed for the output stages. These devices had only been recently developed and it was necessary to determine their true capability and therefore their limitations. In this design, some signal pre-correction was available, so the transistors could be driven close to compression. One of the aims of evaluation is to determine and define some criteria of performance which can be fixed whilst the device is exercised over the design frequency range. This was achieved by determining from the manufacturers impedance data, equivalent parallel resistance values representing the input and source of the transistor and using these values to design first iteration matching circuitry (1). The ratio of the equivalent parallel resistances to the 50ohm system impedance determines the necessary number of impedance transforming stages for a given bandwidth coverage. In this example where output matching is described, three stages were used. See

fig.2. Initially the network was considered as unbalanced and then converted to a balanced microstrip structure. Touchstone (2) was used to optimise the network elements for best performance.

Having achieved at the frequency of interest, a return loss of 16dB at the input, thus ensuring 97% of power transfer into the device, attention was then focussed on the output performance of the transistor/circuit combination. By interposing a mis-match between the first iteration output circuit and the termination, the effective loading on the transistor can be modified in a controlled manner whilst observing the effect on a modulated TV signal. See fig.3. By such a technique, a more accurate series of complex impedances can be determined depending on the required criterion. e.g. optimum power output, optimum efficiency or optimum linearity. This data was then collected at several frequencies in the band and re-entered into Touchstone (2) and, as in this example used to synthesise a full band circuit. Fig.4 shows the elemental make-up of the output circuit and its practical realisation.

A similar technique was used for the design of the input network excepting that, the optimisation criterion was the tailoring of insertion loss to offset the rising gain of the device at low frequencies whilst preserving a good match at high frequencies.

The nature of the television signal is such that amplification is required with a reasonably linear characteristic. In order to maximise output power with reasonable efficiency, Class AB biasing was used.

Several full band prototypes were thus built up to check the repeatability of the design, and according to these principles, each of the output stages was identically built. The stages were arranged in pairs, quadrature coupled by parallel line 3dB hybrids, with two pairs per sub-assembly. Class AB biasing was achieved using a temperature sensing emitter follower circuit which was also steered according to an externally derived function of mean collector current. (3)

A Class AB stage having identical circuitry is used to drive two pairs of output transistors. The output signal from this driver stage is divided between the output pairs by means of a Wilkinson (4) splitter, modified for increased bandwidth by the inclusion of an additional impedance transformer. (5)(6)

#### Pre-amplification

In order to minimise distortion to a suitably high power level, three stages of Class A ultra-linear amplification were employed to raise the low level module input signal to that required at the input of the Class AB driver stage. Manufacturers impedance data was used for the synthesis of load and interstage matching circuitry in this area followed by optimisation by CAD.

At the input to the amplifier assembly a gain stabiliser is used to level the gain around the Class A pre-amplifiers and the Class AB driver. The signal sample used to steer this unit is derived from a microstrip directional coupler (7) at the output of the driver stage. See fig.1

#### Protection and monitoring

To monitor and protect the amplifier assemblies, thermal sensors are employed on the heat sink and collector currents are monitored on all Class AB stages. Forward and reverse power is also sensed by microstrip directional couplers in each of the module outputs. These are of the same form as that described above and their performance and realisation can be seen in fig.5

Fault conditions are signalled back to the gain stabiliser where an electronic attenuator causes the RF. input signal to be muted.

#### Manufacturing issues

Most UHF RF power amplifiers have their circuitry printed in microstrip form but which still have a significant 'lumped component' count. Applying surface mount technology (SMT) to these components is seen as a necessary step to lower the manufacturing cost of RF amplifiers. If en-route more RF components can be included in the same process then so much the better.

Some RF components have physical aspects that are not compatible with SMT, a ferrite or coaxial component for example, but as RF circuit techniques become more adapted to SMT then lower costs will follow. In view of the still slightly mixed technologies, and the high capital cost of 100% pick and place equipment, at Harris we have opted for a semi-automatic pick and place workstation where the operator is presented with the correct component and location information. All dc components and a high proportion of RF components are assembled by this process. Where high placing accuracy is required, some in-house etch resist 'footprints' have been employed to ensure consistency of build and performance.

Prior to component placing, the printed circuit is solder 'pasted' via a screen which masks most of the printed circuit area whilst allowing solder paste to pass through the screen down onto pre-defined areas. Components are then placed as described above, and the circuit fed to the reflow workstation.

#### Thermal issues

An area requiring special attention in this development was cooling. The favoured coolant was forced air and the worst case heat load is around 750W per half module. Conventional extruded heat sinks were unsuitable for this application even with the use of forced air, since only a small part of the airflow picks up any heat energy.

In order to improve this situation therefore, it is necessary to introduce more heat-bearing surfaces into the airflow. This is achieved with a matrix where the extra heat bearing surfaces are thin to minimise air flow resistance but of a sufficiently large area to maximise heat transfer to the air. To further improve efficiency, a longitudinal 'weave' is given to the internal vanes of the matrix to produce air turbulence and thus aid heat transfer. Fig.6 shows the heatsink employed for this application.

A matrix as described above is assembled to each amplifier sub-assembly with the RF transistors mounted on its upper surface. The two sub-assemblies are

then fitted back to back in the module housing which is built in the form of a duct in order to contain the airflow around the matrix heat sinks. To maximise reliability, 250c.f.m. of airflow is provided in the transmitter to limit the heatsink temperature rise to 25 degrees centigrade.

#### Combining issues

In parallel with this development, various multi-port RF splitter and combining structures were designed in order to satisfy modular transmitter requirements of varying output powers up to 30kW peak-sync. The isolating nature of the signal combiners also facilitates servicing by allowing for the on-air removal of a module without interruption of the broadcasting service.

#### Conclusions

After optimisation, a black level power gain greater than 9dB was achieved in the output stages where black level is 87W for system I.

The corresponding collector efficiency at this power is 40% minimum and the worst case sync. pulse compression is 2dB at 150W peak. Allowing for imbalance and transmission losses, 250W peak is obtainable at each of the module output ports.

The class A pre-amplifier gain is 34dB and has a power capability of 16W minimum.

Each RF amplifier assembly has a gain in excess of 50dB from the input to the class A pre-amplifier to each of the output ports, and the complete module covers the full UHF television band without the need for retuning.

The design described here has been in full volume production for almost one year, and is used in the 'Sceptre' range of UHF television transmitters manufactured by Harris Allied.

#### References

[1] G.L.Matthaei "Tables of Chebyshev Impedance-Transforming Networks of Low Pass Filter Form" Proc.IEEE. Aug.1964

[2] Touchstone software is copyright of EEsof, Inc., Westlake Village CA. USA.

[3] D.A.Culling "Solid State Television Transmitters, Performance and Correction" International Broadcasting Convention 1992 Pub.No.358 pp.72-80

[4] E.J.Wilkinson "A N-way Hybrid Power Divider" IRE. Trans.MTT. Vol.MTT-8 No.1 Jan.1960 pp.116-118

[5] L.I.Parad & R.L.Moynihan "Split-Tee Power Divider" Trans.MTT. Vol.MTT-13 No.1 Jan.1965 pp.91-95

[6] S.B.Cohn "A Class of Broadband Three-Port TEM Mode Hybrids" Trans.MTT. Vol.MTT-16 No.2 Feb.1968 pp.110-118

[7] A.Podell "A High Directivity Microstrip Coupler Technique" IEEE Cat.No.70C 10-MTT May 1970 pp.33-36

The author would like to thank the management of Harris Allied for granting their permission to publish this paper, and to his colleagues for their assistance and comments during its preparation.

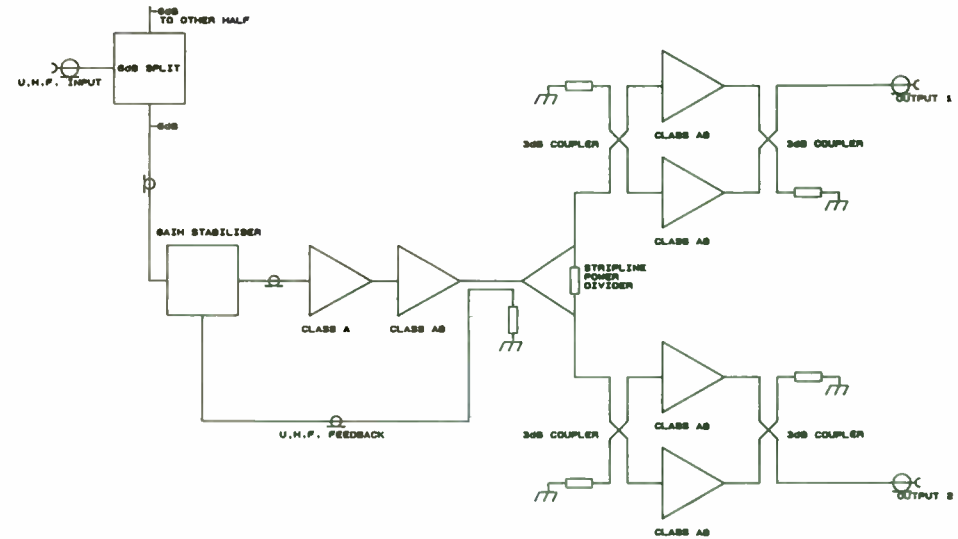


Figure 1 - Half of P.A. Module

FIG. 2

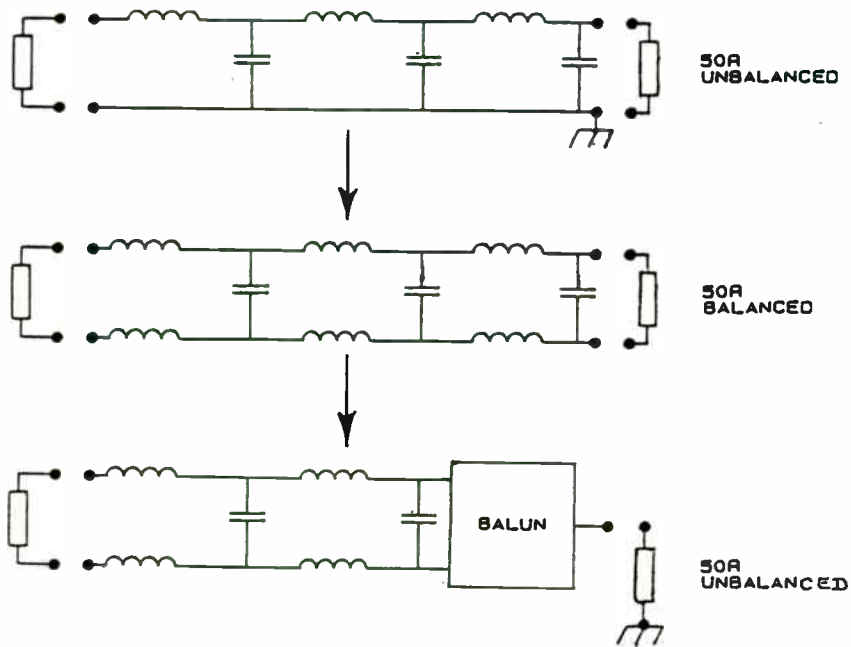


FIG. 3

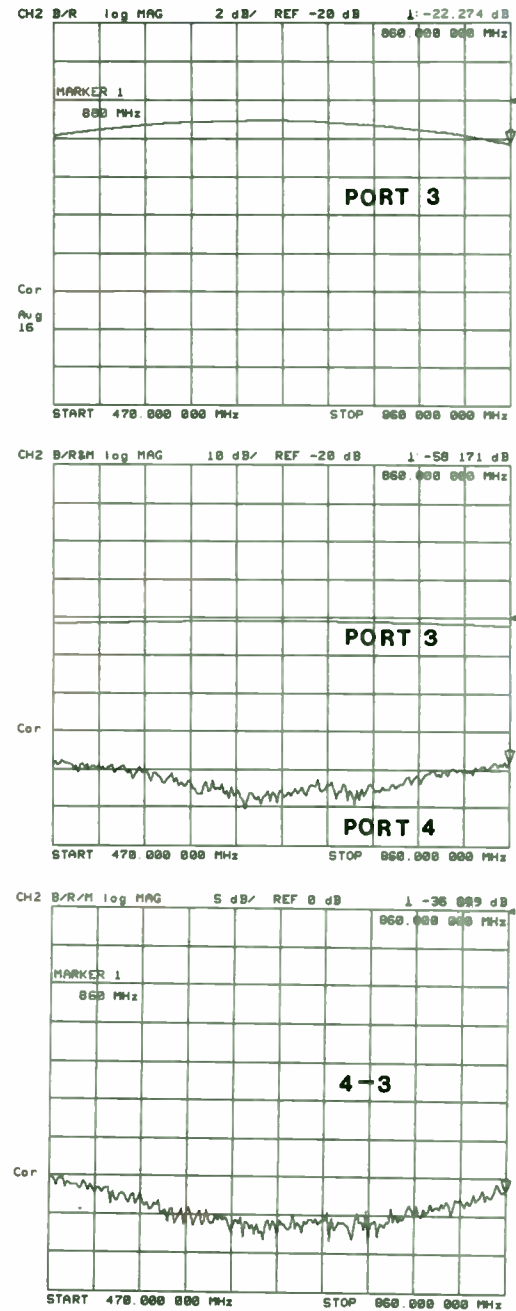
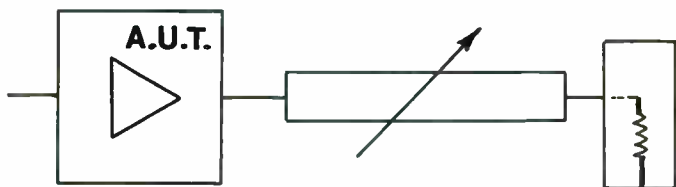
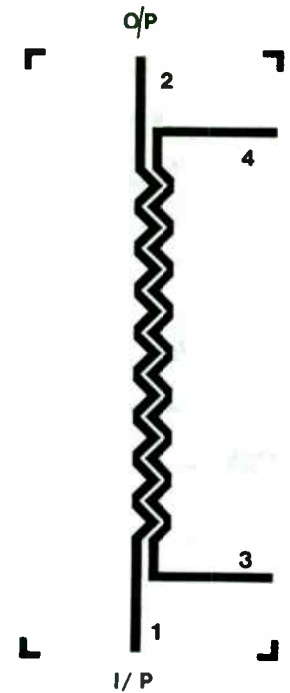


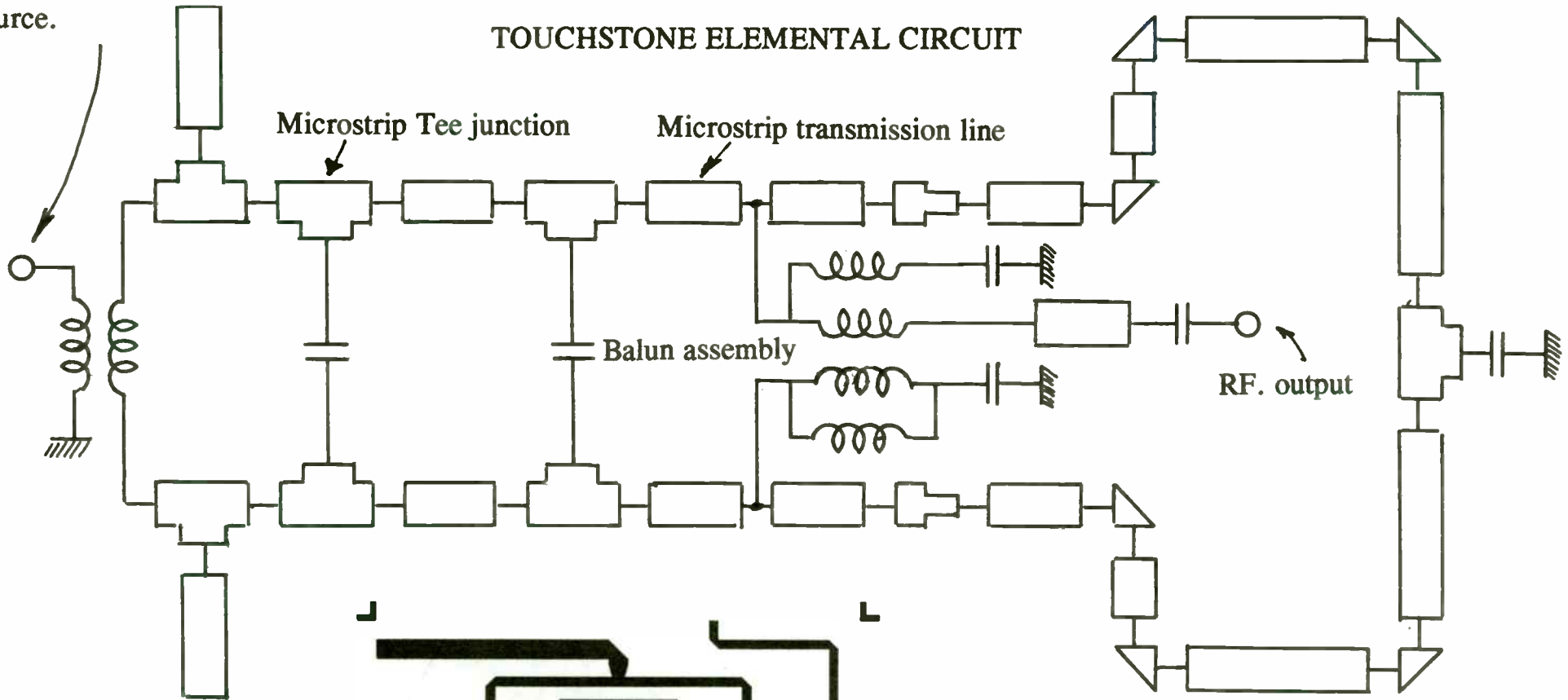
FIG. 5





Connection to equivalent  
single port transistor  
source.

### TOUCHSTONE ELEMENTAL CIRCUIT



### PRACTICAL CIRCUIT

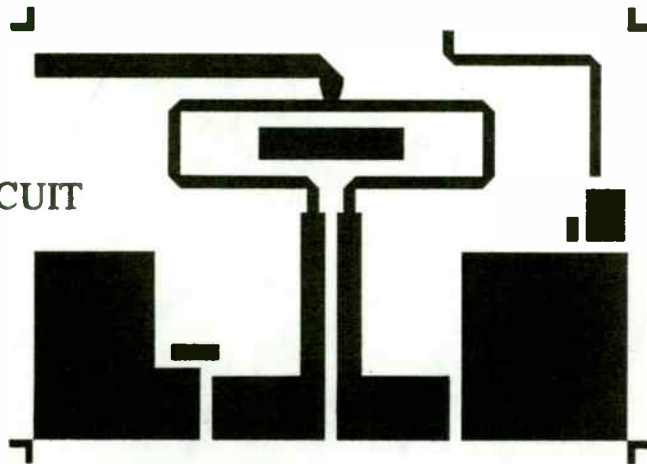
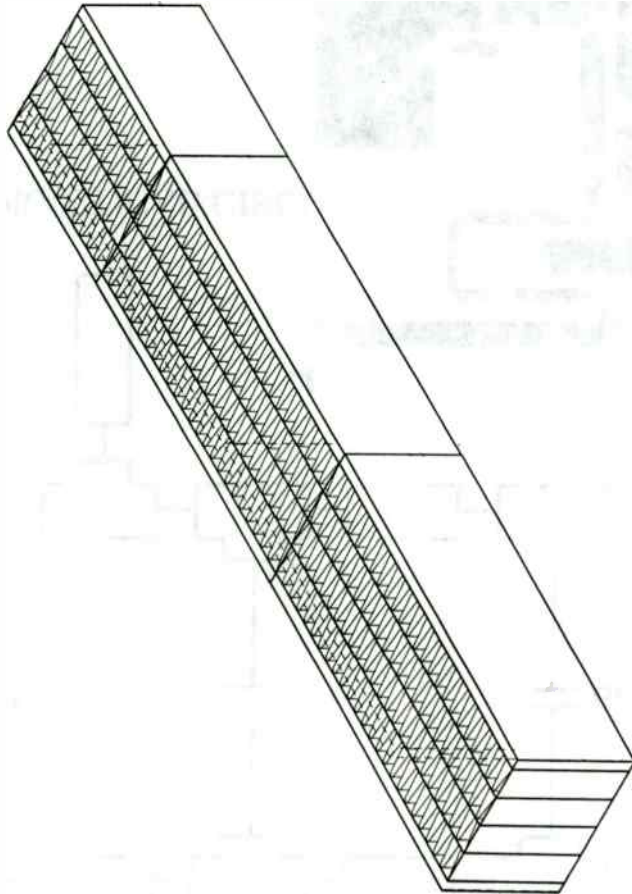


FIG 4

**FIG.6**

**HEAT SINK MATRIX**



## DESIGN CONSIDERATIONS FOR SOLID-STATE HIGH POWER AMPLIFIERS USED IN RADAR SYSTEMS

Michael Hanczor and Mahesh Kumar  
Paramax Systems Corporation  
365 Lakeville Road, Great Neck, N.Y. 11020

### ABSTRACT

This paper presents the design techniques for Solid State High Power Amplifiers using Silicon Bipolar Transistors. The design considerations include: transistor selection/development, amplifier circuit topology, matching circuit design, substrate selection, manufacturing techniques, etc. for providing low-cost, highly reliable amplifiers. A design example of a 2.7-2.9 GHz 300-watt (W) amplifier is presented. This amplifier was developed as a building block for a 12 Kilo-Watt (KW) transmitter used in Air-Traffic-Control Radar.

### INTRODUCTION

Current and future radar systems requires performance that far exceeds the capabilities of Klystron or Tube-Type transmitters. Solid-state devices have demonstrated the performance and reliability that exceeds the performance achieved by tubes.

Modern surveillance and air traffic control radars are complex systems that require field personnel to dedicate a majority of their time to the operation of the radar. Dedication to the radar mission dictates that these field personnel

divorce themselves (to a large degree) from the maintenance, repair and adjustment of the radar equipment. The use of solid state transmitters (SSTs) and solid state high power amplifiers (HPAs) can accomplish this since they require low maintenance and can provide high reliability and stable operation. This paper addresses the key issues associated with the HPA design, such as: transistor selection, impedance matching and D.C. biasing, maximizing output power, module cooling, device derating, junction temperature limits, pulse stability, and performance monitoring. The technology and procedures for designing, fabricating and testing S-Band high power amplifier modules with high efficiency and reliability, and low fabrication costs will be explained. A design example of a 300-watt (W) amplifier operating over 2.7-2.9 GHz is presented.

### HIGH POWER AMPLIFIER DESIGN

The high power amplifier design begins with the selection of the architecture to meet the overall power, gain, efficiency, and thermal requirements. The circuit topology is chosen based on the availability of the power transistor such that the minimum number of transistors are required to be combined to provide overall output power for the HPA. This is important since the loss in the combiner increases with the number of transistors to be combined. The efficiency of the high power transistor is inversely proportional to the output power, since many cells are combined within a transistor to provide the required output power. Therefore, a trade-off study is necessary to determine the number of devices to be combined and to obtain the required output power and efficiency.

An example of a 300-W HPA design is presented in the following paragraphs. The specifications of a 300-W HPA are presented in Table 1.

Table 1: Specifications of a 300-Watt HPA

Frequency:	2.7-2.9 GHz
Output Power:	300 Watts Peak minimum
Gain	17dB min.
Input Power	5W Peak
Pulse Width	100µsec & 10µsec
Duty Cycle	11%
Collector Voltage:	36V
Collector Current	3-6 Amperes
Efficiency	25% minimum
$\Delta P_o/\Delta V_c$	0.2dB/Volt
$\Delta \theta/\Delta V_c$	4.5°/V
Unit to Unit tracking	
Amplitude	0.5dB RMS
Phase	10° RMS

The architecture for the HPA design, shown in Figure 1 consists of three stages of amplification that amplifies a 5-W input signal to a 300-W output signal. The first stage is a 20-W power transistor that supplies 17 watts minimum to the second stage. The second stage transistor is an 85-W transistor having a minimum gain of 6.5 dB followed by a third stage of four paralleled 85-W transistors. As seen in Figure 1 a four way power divider/combiner consisting of 2 levels of two branch branchline microstrip couplers is used to excite and collect power from the four paralleled output stage transistors.

The highest output power of 85-W could be achieved from a transistor. Therefore,

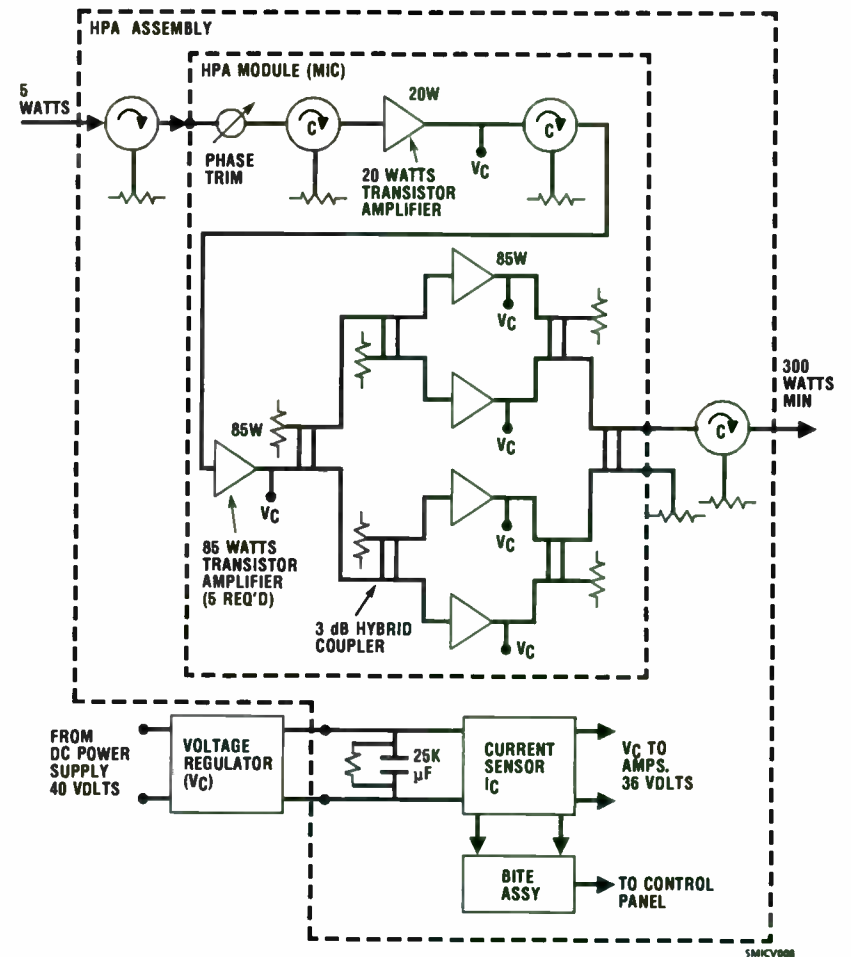


Figure 1 HPA Assembly Block Diagram

a transistor with an output power of 85 W was selected to be combined in a 4-way combiner to obtain 300-W output power.

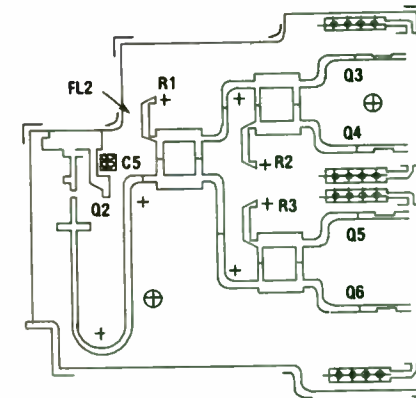
The 4-way power divider/combiner provides a good match, good isolation and low loss as seen in Figure 2. Extreme care was taken to keep the load impedances of the cascaded MIC circuitry to the output of the final stages low. A circulator is used on the output of the HPA module to maintain a constant load VSWR. Further isolation from load VSWR variations is provided by the power combiner. At the input 4-way power divider the hybrid branchline couplers provide an input match (output match of second stage) that is degraded only by the isolation of the couplers when the four output ports are terminated in equal mismatches. Reflected power is terminated in the terminations of the isolated ports of the hybrid couplers.

The module contains an input isolator and an isolator between the first and second stage. These isolators are very low loss (<.2dB) which were designed as above resonance units to handle various levels of peak power without performance degradation. The isolator between the first and second stage provides interstage isolation of at least 25 dB. The isolator at the input of the first stage provides isolation from mismatches of the loaded line phase shifters (See Figure 1) that are used to phase track HPA modules for use in solid state transmitters. The performance of the loaded line phase shifter can be seen in Figure 3.

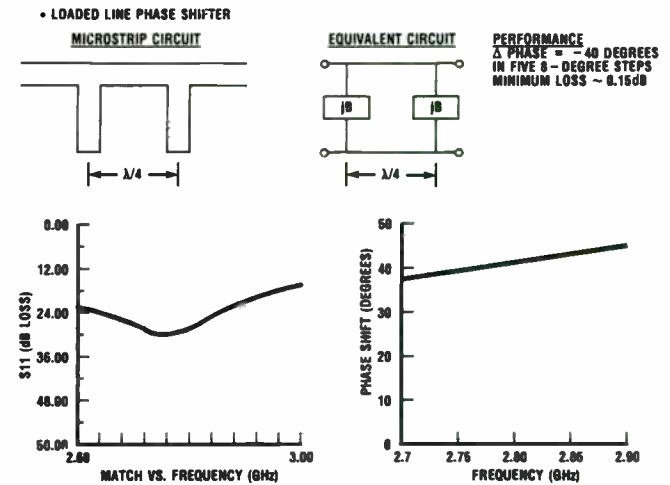
**TRANSISTOR DEVELOPMENT**

A high power silicon bipolar NPN transistor specifically needed for S-Band High Power Amplifiers was developed. The transistor is a 12 cell device with peak output power of 85-Watts at 100 microsecond pulse width and 11 percent duty cycle

FREQUENCY (GHz)	S11(dB)	S21(dB)	S31(dB)	S41(dB)	S51(dB)	LOSS(dB)
2.7	23.9	6.36	6.37	6.19	6.30	-0.284
2.8	30.9	6.30	6.36	6.25	6.35	-0.294
2.9	20.9	6.21	6.49	6.49	6.38	-0.370



**Figure 2 Layout and Performance of Four-Way Power Divider**



**Figure 3 HPA Phase Compensation Network**

over a 2.7 to 2.9 GHz bandwidth. The device was specifically designed to operate with a maximum flange to junction temperature rise ( $\Delta T_{j-f}$ ) of 70°C. A low RF thermal resistance was achieved with a transistor chip that has a high density overlay geometry with emitter site ballasting. The 12 cells of the transistor were internally matched, using bondwires as inductors and MOS capacitors, in an MSC AMPAC\* package to provide an input/output impedances of ~5-15 ohms. A photograph of the transistor is shown in Figure 4. Table 2 lists the characteristics of this transistor.

Table 2: Electrical Characteristics of 85W Transistor

Frequency	2.7 to 2.9 GHz
Output Power	85 Watt peak minimum
Pulse Width	100 microseconds
Duty Cycle	11%
Collector Voltage	36 Volts
Gain	6.5dB minimum
Collector Efficiency	40% minimum
Pulse Droop	0.30dB maximum
$\Delta\theta/\Delta V_{cc}$	6.5 degrees/volt max
$\Delta P/\Delta V_{cc}$	0.4dB/volt

\* Microwave Semiconductor Corporation, Somerset, New Jersey

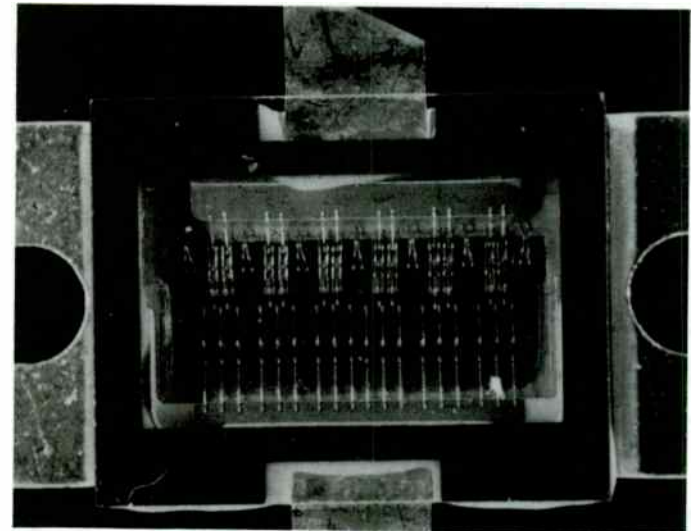


Figure 4 MSC AMPAC Transistor Developed for High-Power Amplifier

#### MATCHING CIRCUIT DESIGN

The external matching circuits were developed on 0.050" thick soft substrate (dielectric constant 10.2). The thickness of substrate material was chosen to be .050" to minimize the losses in the matching circuits. The procedure for matching the transistor input/output impedances to 50 ohms was an iterative process that involved starting with input and output terminal impedances of the device and multiple-stub tuner optimization with computer aided design tools (Supercompact) to converge on a solution that met the requirements of Table 2. The resulting matching and bias circuits are shown in Figure 5 and the resulting transistor model and input and output matching impedances are shown in Figure 6.

The transistor was chosen to operate in class C and in a common base configuration. The common base configuration was chosen because it delivers

higher power gain than common emitter at S-Band. Class C operation was chosen for high efficiency and high output power. In class C operation, the transistor is biased below cut-off and is turned on only when the negative input voltage swing exceeds the base emitter diode drop. The transistor is turned on for less than 180 degrees of the RF cycle thus yielding high efficiency operation.

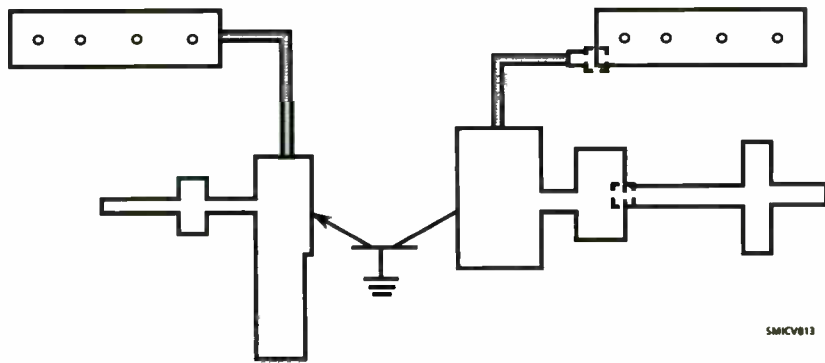


Figure 5 Input and Output Transistor Matching Circuits

**THERMAL/RELIABILITY CONSIDERATIONS**

An extensive thermal analysis was performed on the transistor in the HPA environment. The results (Figure 7) indicate that with an air intake temperature of 40°C the maximum junction temperature will be 135°C. The 135°C temperature translates to an MTBF of over 1 million hours for the transistor.

During the development of the HPA module 20 high power devices were delivered. The statistics of these 20 devices are shown in Table 3.

- COMMON BASE
- CLASS C SATURATED
- PIN = 5.1 WATTS PK



Z <sub>in</sub>	F(GHz)	Z <sub>in</sub> (Ω)
	2.7	8.0 + J18.5
	2.8	15.0 + J21.0
	2.9	17.5 + J12.0

Z <sub>CL</sub>	F(GHz)	Z <sub>CL</sub> (Ω)
	2.7	3.75 - J9.00
	2.8	4.50 - J9.50
	2.9	5.00 - J8.00

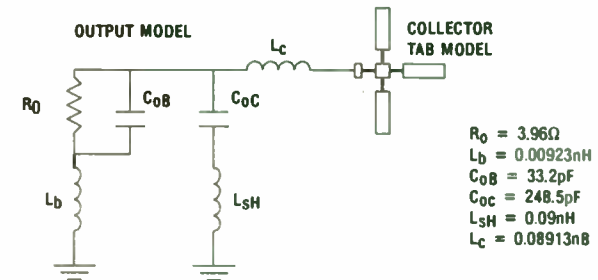
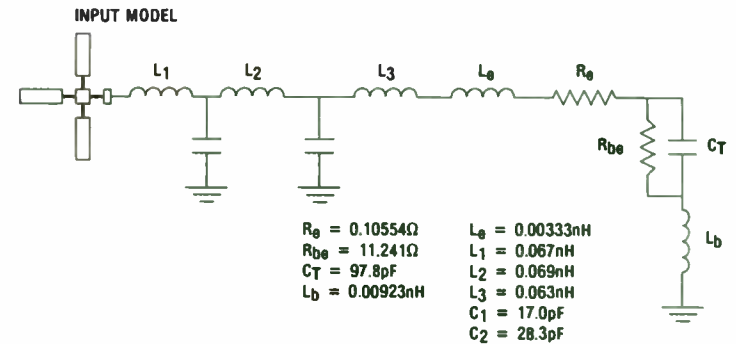


Figure 6 High-Power Transistor Model/Equivalent Circuit

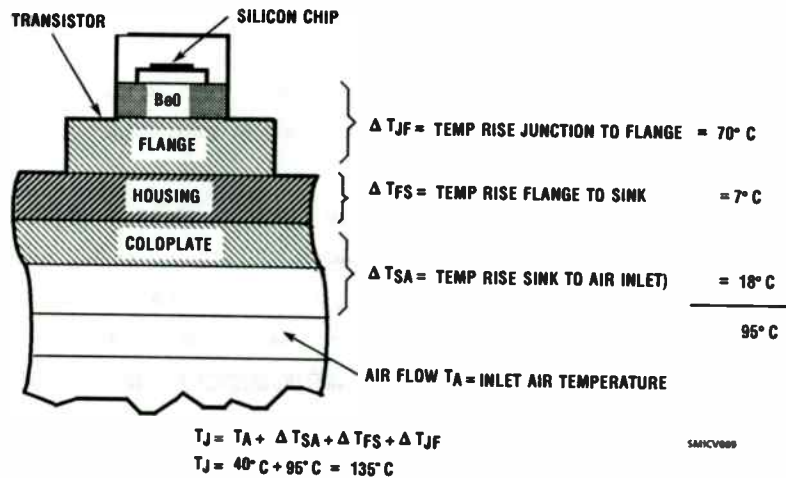


Figure 7 HPA Thermal Analysis

Table 3: Performance Statistics of 20 High Power S-Band Transistors

Output Power	85.6 Watts
$\theta_{TJF}$	0.431°C/Watt
$\Delta T_{JF}$	50.6°C
Power Dissipation	117.7 Watts
Collector Efficiency	46%
Gain	7.47dB

Observe that the performance of these transistors exceeds the requirements initially placed on these devices which leaves room for a higher power HPA than originally required.

**RELIABILITY AND THE REGULATOR/BITE ASSEMBLY**

The mean-time-between-failures (MTBF) of the High Power Amplifier is 156,000 hours minimum when calculated in accordance with MIL-HDBK-217. The prediction assumed a 50 degrees centigrade ambient temperature and a ground fixed environment ( $G_f$ ).

In solid state transmitters many HPAs are needed to generate the power levels needed in radar systems. In the event of failures it is necessary to replace failed HPA's when their number represents a significant drop in output power. As part of performance monitoring and fault isolation (PM/FI) needed for HPA's in SST's bite capacity is essential to the HPA design.

Bite circuitry is included on a regulator printed circuit board which monitors the performance of the HPA. A current sense transformer measures the amplitude of the current pulses drawn by the HPA from its storage capacitors. The height of these pulses is compared to a fixed reference by a comparator. The height of the pulses is proportional to the power output of the HPA. If the output of the HPA falls below the reference level, the bite circuitry sends a signal to a control panel in the SST and a LED is lit, pinpointing the failed unit.

**VOLTAGE REGULATION AND PULSE-TO-PULSE STABILITY**

In a solid state transmitter the limiting factor in MTI performance is the pulse to pulse stability of the high power amplifier. The limit on the improvement factor due to phase instability is

$$I = 20 \text{ Log } (I/\Delta\phi); \Delta\phi = \text{Interpulse phase change (Radians)}$$

The upper limit on the improvement factor due to amplitude instability is

$$I = 20 \text{ Log } (V/\Delta V); V, \Delta V = \text{Voltage, Change (Volts)}$$

With the pushing factors achieved for this HPA

$$\Delta\phi/\Delta V_{cc} = 5.0^\circ/\text{Volt max}$$



and

$$\Delta P_o / \Delta V_{cc} = 0.5 \text{ dB/Volt max}$$

and a 3 millivolt regulation gotten from our voltage regulator, the theoretically predicted combined amplitude and phase improvement factor (for subclutter visibility) was - 76.1 dBc. The HPA described in this paper measured - 57.53 dBc.

This value when projected up to a complete solid state transmitter containing 56 HPA's in the output stage, the projected MTI improvement for an 8 pulse cancellation scheme is  $I = 81.68 \text{ dBc}$ .

The approach used to measure the pulse to pulse stability of the HPA consisted of recording the HPA output in-phase and quadrature voltage referenced to the input on a pulse-to-pulse basis. The next phase in the data reduction presents the results as a pulse pair vector residue, amplitude residue, phase residue, phase and power runout, and a spectral analysis.

The measurement equipment contains a homodyne detector and two waveform recorders to digitize both I and Q channel voltages with a dynamic range of 60 dB. The pulse-to-pulse, in-phase and quadrature components of the output signal relative to the input signal are digitized and recorded.

Data on successive pulses is then cataloged and stored by a computer. The stored data is then analyzed to show pulse-to-pulse amplitude and phase instabilities as well as pulse spectrum.

#### FABRICATION AND PERFORMANCE

The HPA module as seen in Figure 8 is constructed out of a machined aluminum housing. The microwave circuits as well as the transistors are soldered to the floor of the housing. The circuits are fabricated out of low loss high dielectric constant soft substrate material. Bias to the transistors is achieved by the use of small and large capacitors that feed up through the substrates via feed through filters. The filters connect to the transistor collectors through a quarter wave section of high impedance transmission lines terminated in chip capacitors to ground. Grounds for the bias networks on the top of the substrates are brought up from the metal housing via plated through via holes in the substrates. The small capacitors mounted in the underside module body (Figure 8) provide current for the leading portion of the 100 microsecond pulse with a rise time of less than 100 nanoseconds (Figure 9). The large capacitor supplies current to support the waveform to within .1 volt droop for 80 percent of the remaining portion of the pulse. The large capacitor mounted on the cover of the module is connected to the small caps via a low resistance low inductance flex print circuit (Figure 8).

The performance of the HPA assembly is shown in Table 4. Phase shift measurements and maximum junction temperature measurements in the module are shown in Figure 9. The transistor junction temperature measurements were taken with a Barnes RM 2A infrared scanning microscope.

#### CONCLUSIONS

Design considerations for a solid-state High Power Amplifier have been presented. Various considerations were illustrated using a design example of a 300-watt amplifier operating over 2.7 - 2.9 GHz. This amplifier was designed, fabricated and integrated into a 12 KW transmitter used in a Air-traffic Control Radar.

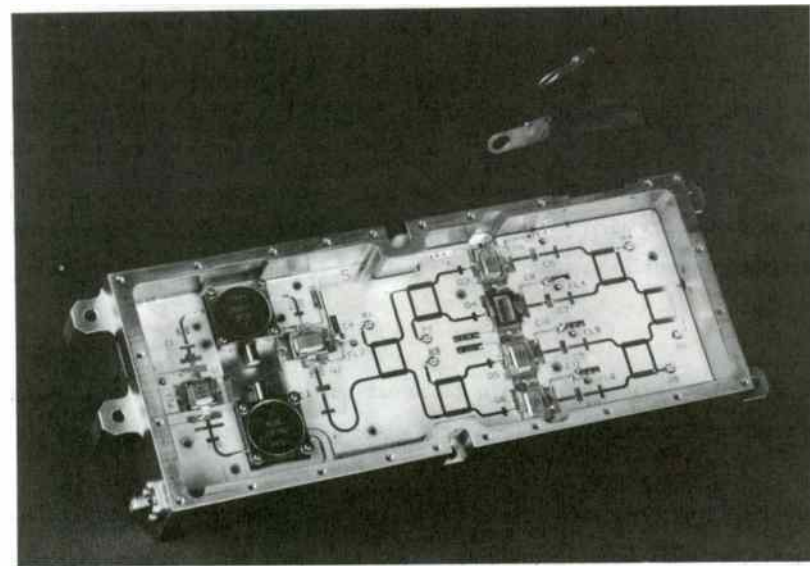
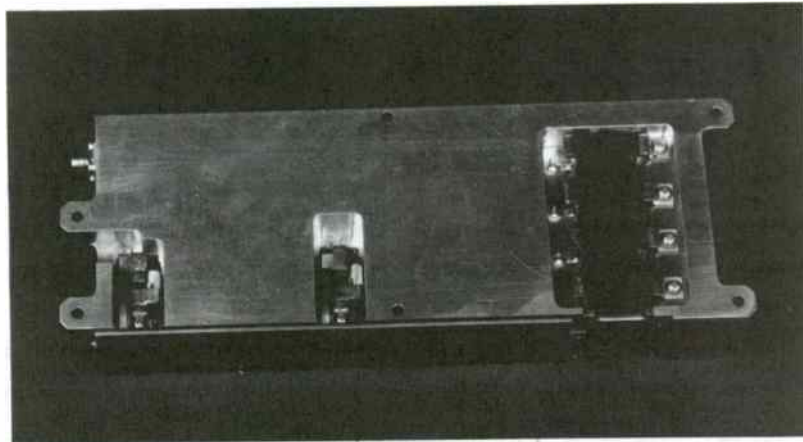
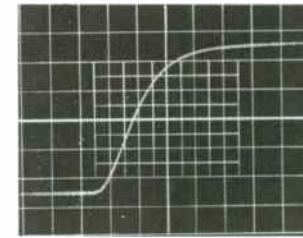


Figure 8 300-Watt S-Band High-Power Amplifier Module (Top and Bottom View)



HORIZONTAL = 20 NS/DIV.  
RISETIME < 60 NANOSEC

HPA Module Output Pulse - Leading Edge

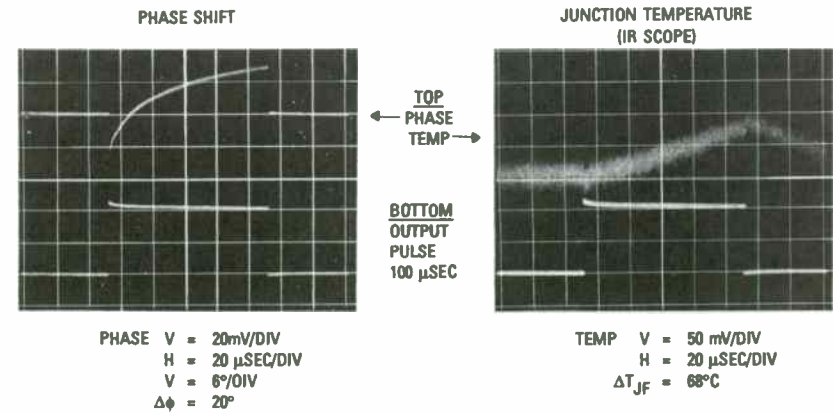


Figure 9 Waveform Test Data of 300-Watt HPA Module

**TABLE 4: HPA PERFORMANCE**

<b>FREQUENCY</b>	<b>2.7 TO 2.9 GHZ</b>
<b>PULSE WIDTH</b>	<b>100 μSEC AND 10 μSEC</b>
<b>DUTY CYCLE</b>	<b>11%</b>
<b>OUTPUT POWER</b>	<b>300 WATTS PK MIN.</b>
<b>INPUT POWER</b>	<b>5 WATTS PK</b>
<b>GAIN</b>	<b>17 dB MINIMUM</b>
<b>COLLECTOR VOLTAGE</b>	<b>36 V (AMPLIFIER) 39.5V (REGULATOR)</b>
<b>COLLECTOR CURRENT</b>	<b>3 . 6 A M P S . A V E . (REGULATOR INPUT)</b>
<b>EFFICIENCY</b>	<b>26%</b>
<b>ΔPo/ΔVc</b>	<b>0.2 dB/VOLT</b>
<b>Δ/ΔVc</b>	<b>4.5°/VOLT</b>
<b>Vc REGULATION</b>	<b>3 MILLIVOLTS</b>
<b>UNIT TO UNIT TRACKING</b>	
<b>AMPLITUDE</b>	<b>0.5 dB RMS</b>
<b>PHASE</b>	<b>10 DEGREES RMS.</b>



## Quasi-complementary class-D HF power amplifier

by

Frederick H. Raab, Ph.D., and Daniel J. Rupp  
Green Mountain Radio Research Company  
50 Vermont Avenue, Fort Ethan Allen  
Colchester, Vermont 05446

### ABSTRACT

This quasi-complementary power amplifier (PA) directly connects a pair of  $n$ -channel MOSFETs in a totem-pole configuration, eliminating the need for a push-pull output transformer. The PA operates in class D for maximum efficiency and power output. The 25-W version achieves an efficiency of 74 percent at frequencies up to 30 MHz, and an efficiency of 50 percent at 60 MHz. The 100-W version uses MRF148 MOSFETs and achieves an efficiency of 70 percent at frequencies up to 20 MHz and an efficiency of 50 percent at 30 MHz. The amplitude-modulation linearity is equivalent to -43-dB IMD products. It offers superior efficiency for low-amplitude signals and reactive loads.

### 1. INTRODUCTION

Class-D power amplifiers (PAs) employ two transistors in a push-pull configuration. The transistors are driven to act as switches and generate a square-wave voltage. The fundamental-frequency component of the squarewave is passed to the load through a filter. The class-D PA is ideally 100-percent efficient. In practice, it is significantly more efficient than a similar class-B PA, especially for lower amplitudes and reactive loads.

Virtually all modern broadband RF PAs (with two exceptions [1], [2]) are based upon transformer-coupled topologies. In contrast, virtually all modern

AF PAs are based upon complementary (transformerless) topologies. The potential advantages of a transformerless RF PA include:

- Greater bandwidth (limited only by the active devices),
- Elimination of power loss in the output transformer,
- Reduced size and weight, and
- Elimination of IMD due to core saturation.

Such a PA might be fabricated in integrated-circuit form for considerably less than is currently possible for PAs of the same capability.

A true complementary configuration offers great simplicity but requires both  $n$ - and  $p$ -channel MOSFETs. While  $p$ -channel devices are available for power-switching applications, none are currently made for RF applications. Experiments with complementary pairs of power-switching MOSFETs (e.g., IRF610 and SMP2P20) show that they can be used at frequencies no higher than 3.5 or 7 MHz.

The quasi-complementary power amplifier (Figure 1) directly connects a pair of  $n$ -channel MOSFETs in a totem-pole configuration. Drive is supplied through a specially designed input transformer. An important advantage of this topology is that it can be implemented with currently available RF devices.

The voltage and current ratings of currently available devices may not be adequate for delivery of the desired power output directly to a 50- $\Omega$  load, hence the use of an output transformer may be required. However, this transformer is external to the PA itself and can therefore be a single-ended, equal-delay configuration. As a result, it is smaller and less complex than the push-pull transformer normally required for class-B or -D operation, and also has a larger bandwidth.

## 2. PRINCIPLES OF OPERATION

The complementary voltage-switching configuration (Figure 1) is the least complicated class-D power amplifier (PA) [3]. Transformer T1 causes the driving signals applied to the gates of Q1 and Q2 to be of opposite phases. The driving signal is more than sufficient to support the drain current. Consequently, Q1 is on when Q2 is off and vice versa. The pair of FETs therefore forms the equivalent of an SPDT switch and generates the rectangular voltage waveform  $v_{D_2}(\theta)$  shown in Figure 2.

### Ideal Operation

Maximum fundamental-frequency output is obtained with a 50-percent duty ratio. Expansion of the square-wave drain voltage into a Fourier series yields

$$v_{D_2}(\theta) = V_{DD} \left( \frac{1}{2} + \frac{2}{\pi} \sin \theta + \frac{2}{3\pi} \sin 3\theta + \frac{2}{5\pi} \sin 5\theta + \dots \right), \quad (1)$$

where  $\theta = \omega t$ .

The output of the PA is determined by the response of the load network to each of the components in the applied voltage waveform  $v_{D_2}(\theta)$ . The load network consists of the load resistance  $R$  and a series-tuned tank circuit (L2-C3). The tank circuit is assumed to have a reasonable loaded  $Q$  (i.e.,  $> 3$ ), to have negligible loss, and to be resonant at the fundamental (switching) frequency. Under such conditions, the input impedance of the load network at dc and the harmonics of the switching frequency is very high. The amplitudes of the harmonic currents are therefore very low and their effect upon PA operation is negligible. (Additional filtering may be necessary for other purposes, however, such as prevention of radiation of the harmonics by

an antenna.)

If the tank circuit is tuned to resonance at the fundamental frequency, its net reactance is zero at that frequency. The fundamental-frequency component of the switching waveform (3) is then applied directly to load  $R$ , producing the output voltage and current waveforms

$$v_o(\theta) = v_{om} \sin \theta = \frac{2}{\pi} V_{DD} \sin \theta, \quad (2)$$

and

$$i_o(\theta) = v_o(\theta)/R = I_{om} \sin \theta. \quad (3)$$

The corresponding power output is

$$P_o = \frac{v_{om}^2}{2R} = \frac{2}{\pi^2} \frac{V_{DD}^2}{R} \cong 0.203 \frac{V_{DD}^2}{2R}. \quad (4)$$

The sinusoidal output current flowing into the load network is drawn through whichever FET is on. When the tank circuit is tuned to have zero reactance at the fundamental frequency, both drain current waveforms are half sinusoids, as shown in Figure 2 ( $V_{DD} = R = 1$ ).

The dc input current is obtained by averaging  $i_{D_1}(\theta)$  over one RF cycle, thus

$$I_{dc} = \frac{1}{\pi} I_{Dm} = \frac{1}{\pi} I_{om} = \frac{2}{\pi^2} \frac{V_{DD}}{R}. \quad (5)$$

The dc input power is therefore

$$P_i = V_{DD} I_{dc} = \frac{2}{\pi^2} \frac{V_{DD}^2}{R} = P_o. \quad (6)$$

Since the dc input power and RF output power of this PA are equal, its efficiency is 100 percent ( $\eta = 1$ ). The 100-percent efficiency is consistent with

zero dissipation in the FETs; for both FETs, either the drain voltage or the drain current is zero at all times.

### Losses

The efficiency of a class-D PA is less than unity because of three principal factors:

- On resistance of the MOSFETs,
- Switching time of MOSFETs, and
- Drain capacitance of the MOSFETs.

The effects of on resistance  $R_{on}$  can be incorporated into the equations of an ideal amplifier by using the effective supply voltage

$$V_{eff} = \frac{R}{R + R_{on}} V_{DD} \quad (7)$$

in place of  $V_{DD}$  except in the calculation of  $P_i$ .

Finite switching time reduces the output voltage by the factor  $\tau^2/6$ , where  $\tau$  is the rise/fall time converted into radians (Figure 14-6, [3]). Its effects are incorporated into design calculations by including this factor in (7).

The dc-input current required to deliver power to the output is given by (5). The dc-input current required to charge a fixed total drain capacitance  $C_F$  is

$$I_{dc1} = f C_F V_{DD}^2 \quad (8)$$

The dc-input current required to charge the total voltage-variable drain capacitance [4] is

$$I_{dc2} = 2f C_{V_0} \phi [(1 + 2V_{DD} / \phi)^{1/2} - 1] \quad (9)$$

where  $C_{V_0} + C_F$  is the capacitance at  $v_{DS} = 0$  and  $\phi = 1$  is the barrier potential.

It is important to note that load reactance ideally does not reduce efficiency in class-D operation. In practice, the larger peak currents that circulate between the power supply and the load (Figure 14-4 of [3]) cause larger losses in the FET on resistances. These losses are, however, a second-order effect and generally considerably smaller than those in a class-B PA.

### Bandwidth

The bandwidth of a class-D PA is limited by the bandwidth of its output filter. Since the filter must reject the second harmonic of the lowest frequency while passing the fundamental of the highest frequency, it can have a bandwidth ratio no larger than 2. In practice, filters with bandwidth ratios of up to 1.5 or so can be synthesized.

### Envelope Elimination and Restoration

In class-D operation, the FETs are overdriven and linear amplification of the driving signal is impossible. However, class-D amplifiers have excellent amplitude-modulation linearity and can be used as part of a linear-amplifier system through Envelope Elimination and Restoration (EER) [5]. In EER, the hard-limited carrier (which contains phase information) is amplified by highly efficient but nonlinear RF power amplifiers. The signal envelope (amplitude information) is amplified by a highly efficient class-S audio amplifier [3] and remodulated onto the carrier at the final power amplifier.

### 3. CIRCUIT

#### Transistors

Field-effect transistors are preferred to bipolar-junction transistors in a class-D PA because they can pass current in the reverse direction, which is essential for operation with a reactive load. In practice, this characteristic means that FETs are reliable in HF class-D PAs, whereas BJTs are not. MOSFETs have been used successfully in a number of class-D PAs operating at HF and above [6] - [9].

Equation (5) shows that delivery of 25 W to a 12.5- $\Omega$  load requires FETs with ratings of 39.3 V and 2.0 A. Delivery of 100 W similarly requires ratings of 78.5 V and 4.0 A. Allowance for loss in  $R_{on} = 2 \Omega$  increases the voltage ratings by about 16 percent to 45.6 and 91.1 V, respectively. Since increased voltage ratings are associated with larger values of  $R_{on}$  and increased current ratings are associated with larger values of drain capacitance, it is desirable to select FETs with ratings only slightly larger than necessary.

The MRF136 and MRF148 RF-power MOSFETs are therefore used for the 25-W and 100-W amplifiers, respectively. The standard SOE package is intended for transformer-coupled, grounded-source applications and is therefore somewhat awkward in this circuit. Interconnection is accomplished by mounting one MOSFET at a 45° angle with respect to the other MOSFET so that the source lead of Q1 meets the drain lead of Q2.

#### Gate Swamping

The MOSFET gate is a rather inhospitable load. As the MOSFET approaches saturation,  $v_D$  drops below  $v_G$  and the drain-gate capacitance abruptly inflates by a factor of eight or so. Resistors R1 and R2 swamp this effect and provide

the driver with a reasonably well behaved load. They are also essential for maximum bandwidth of the input transformer.

#### Input Transformer

The most difficult and critical design problem in the implementation of the quasi-complementary PA is the input transformer. The reasons include:

- Large bandwidth,
- Unusual arrangement of secondary windings,
- Capacitance to ground must be charged during switching, and
- Common-mode reactance diverts drain current from output.

Four approaches to the input transformer are:

- Twisted windings on a single toroid,
- Three baluns configured to drive equal current into each MOSFET,
- Three baluns configured to apply equal voltage to each MOSFET, and
- Brass-tube transformer with two secondary windings.

The brass-tube transformer (Figure 3) provides better overall performance. It achieves a relatively large bandwidth by choking common-mode current. The use of separate windings provides a higher common-mode impedance than is possible in a balun with cores of the same size.

The brass-tube transformer is built from two 0.48-cm (3/16-in) brass tubes connected in series through 10 Ceramic Magnetics CMD5005 cores (9.525 × 4.763 × 4.763 mm) with two 2-turn windings of 24-AWG enamel-coated wire. Its common-mode impedance of about 1000  $\Omega$  at 3.5 MHz results in a one-percent increase in supply current. The driving impedance is about 12.5  $\Omega$  when used with the gate-swamping resistors shown in Figure 1. It keeps gate-to-gate phase error below 20° for frequencies from 1 to 60 MHz. (A comparable transformer with a 50- $\Omega$  input impedance can be fabricated by using only a single



turn for each secondary winding and adding one or two more cores to each stack).

#### Power-Supply Decoupling

Power-supply decoupling is achieved by C1, C2, and L1. The bandwidth of this filter determines the maximum amplitude-modulation bandwidth. It is essential that C2 provide a good RF ground. It must also be able to store enough charge to feed the output current for half of a cycle (at the lowest operating frequency) without creating voltage drop on the drain of Q1.

#### Series-Tuned Output Filter

The output filters are simple series-tuned circuits with  $Q = 5$ . Capacitor C3 is a combination of an air variable and mica capacitors. The inductors (L2) are wound on various toroids and have unloaded  $Q$  in the range of 70 to 200.

#### Output Transformer

The output transformer (Figure 4) is an equal-delay configuration [10] of two baluns that provides a 4:1 reduction of the load impedance. Both T2 and T3 employ 25- $\Omega$  transmission line formed by connecting two RG-196/U 50- $\Omega$  lines in parallel. Ten Ceramic Magnetics CMD5005 beads are placed over the transmission line of T2. The bandwidth of the T2-T3 combination is sufficiently large that it can be located either before or after the series-tuned filter.

### 4. PERFORMANCE

The drain-voltage waveforms (MRF136) at 2, 14, and 30 MHz are shown in

Figure 5. The waveform is quite rectangular at the lower frequencies, but its top contains significant ringing at the upper frequencies. The ringing is the result of interaction between the drain capacitance and the input-transformer inductance, and varies with the load reactance.

The output power and efficiency are shown in Figure 6. With MRF136 MOSFETs, the PA delivers 25 W from 1 to 60 MHz. Its efficiency is 75 percent or better from 1 to 30 MHz and then drops to about 50 percent at 60 MHz. With MRF148 MOSFETs, the PA delivers 85 to 100 W from 1 to 20 MHz and about 75 W at 30 MHz. Its efficiency is 70 percent or better until 20 MHz and then drops to about 50 percent at 30 MHz. (Power-output measurements are corrected for the losses in the output filter and transformer).

When the MRF148 is used, the maximum operating frequency is limited by the FETs. However, when the MRF136 is used, the maximum operating frequency is limited by the input transformer. When the input transformer is replaced with one tailored for operation in the VHF range, operation can be extended to 100 MHz (Figure 6).

The amplitude-modulation linearity of the 25-W PA at 7 MHz is shown in Figure 7. The rms deviation of the transfer function from a straight line is only 0.7 percent, which is roughly equivalent to IMD products at -43 dBc. With  $V_{DD} = 0$ , drive feedthrough produces an output voltage of 0.5 V (-40 dBc).

The variation of efficiency with output voltage for 7 MHz is shown in Figure 8. Comparison of the measurements to the curve for an ideal class-B PA shows the advantage of class-D for signals with large peak-to-average ratios. The apparent increase in efficiency near zero output is believed to be due to drive feedthrough.

The variation of efficiency with load series reactance at 7 MHz is shown in Figure 9. For these measurements, MRF136 MOSFETs are used and  $V_{DD}$  is re-

adjusted to avoid exceeding the voltage ratings ( $P_o = 10$  W for  $-100 < X < +80$   $\Omega$  and  $P_o = 5$  W for  $-150 < X < +150$   $\Omega$ ). The benefit of class-D operation over class-B operation into reactive loads is apparent.

The experimental PA is driven by an ENI 525L laboratory power amplifier. The drive level is set to produce maximum efficiency and the most square waveforms; no attempt is made to minimize the driving power. For MRF136s, the drive power varies from 5 W at 1 MHz to 7 W at 30 MHz. The resultant sine-waves of 20-V peak on the gates cause dissipation of 4 W of the drive power in the gate-swamping resistors. Little additional power is required to produce a 100-W output with the MRF148 MOSFETs.

#### 5. CONCLUSIONS AND RECOMMENDATIONS

The experimental quasi-complementary class-D PA described here demonstrates the ability of a MOSFET class-D PA and the quasi-complementary configuration in particular to deliver useful power with high efficiency across the entire HF and lower VHF bands. The modulation linearity is more than adequate for implementation of a linear-amplifier system via envelope elimination and restoration.

An obvious area for future development is packaging of the MOSFETs for quasi-complementary operation. Such a package would eliminate the somewhat cumbersome interconnection of standard SOE packages. The resultant reduction in lead lengths should also improve the VHF characteristics of the PA.

Improving the input transformer could also lead to an increased maximum frequency of operation. Development of MOSFETs with higher voltage ratings would eliminate the need for the output transformer, resulting in further simplification and higher efficiency.

A most interesting area for further development is a *p*-channel MOSFET for true-complementary operation. The use of a true complementary pair would eliminate the need for the input transformer, since the same driving signal could be coupled to both gates through capacitors. This configuration would also allow the use of gate bias, which would reduce the drive power by about 30 percent. Such a PA might be driven by logic-type circuits and fabricated in a single package.

#### 6. REFERENCES

- [1] H. Granberg, "A complementary symmetry amplifier for 2 to 30 MHz SSB driver applications," EB-32, Motorola Semiconductor Products, Phoenix, 1975.
- [2] F. H. Raab, "High efficiency RF power amplifiers," *Ham Radio*, vol. 7, no. 10, pp. 8 - 29, October 1974.
- [3] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*. New York: Wiley, 1980.
- [4] F. H. Raab, "Class-D power-amplifier load impedance for maximum efficiency," *Proc. RF Technology Expo*, Anaheim, CA, pp. 287 - 295, January 23 - 25, 1985.
- [5] F. H. Raab, "Envelope-elimination-and-restoration system concepts," *Proc. RF Expo East '87*, Boston, MA, pp. 167 - 177, November 11 - 13, 1987.
- [6] F. H. Raab, "Get broadband, dual-mode operation with this FET power amplifier," *EDN*, vol. 23, no. 19, pp. 117 - 124, October 20, 1978.
- [7] F. H. Raab, "MOSFET power amplifier for operation from 160 through 6 meters," *Ham Radio*, vol. 11, no. 11, pp. 12 - 17, November 1978.
- [8] H. O. Granberg, "FETs versus BJTs: Making the proper choice for RF power," *Microwaves & RF*, vol. 28, no. 8, pp. 125 - 135, August 1989.
- [9] H. O. Granberg, "Applying power MOSFETs in class D/E RF power amplifier design," *RF Design*, vol. 8, no. 6, pp. 42 - 47, June 1985.
- [10] R. K. Blocksome, "Practical wideband RF power transformers, combiners, and splitters," *Proc. RF Expo East*, Boston, MA, pp. 1 - 21, November 10 - 12, 1986.

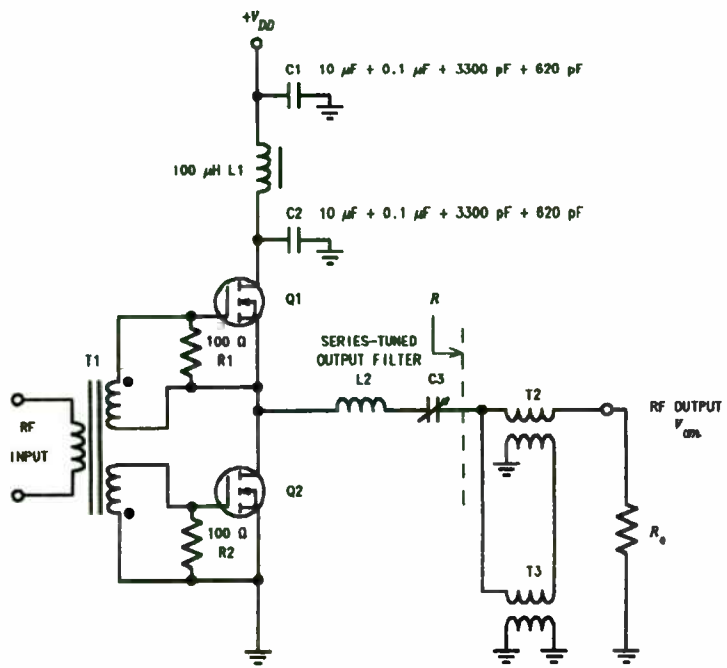


Figure 1. Quasi-complementary PA.

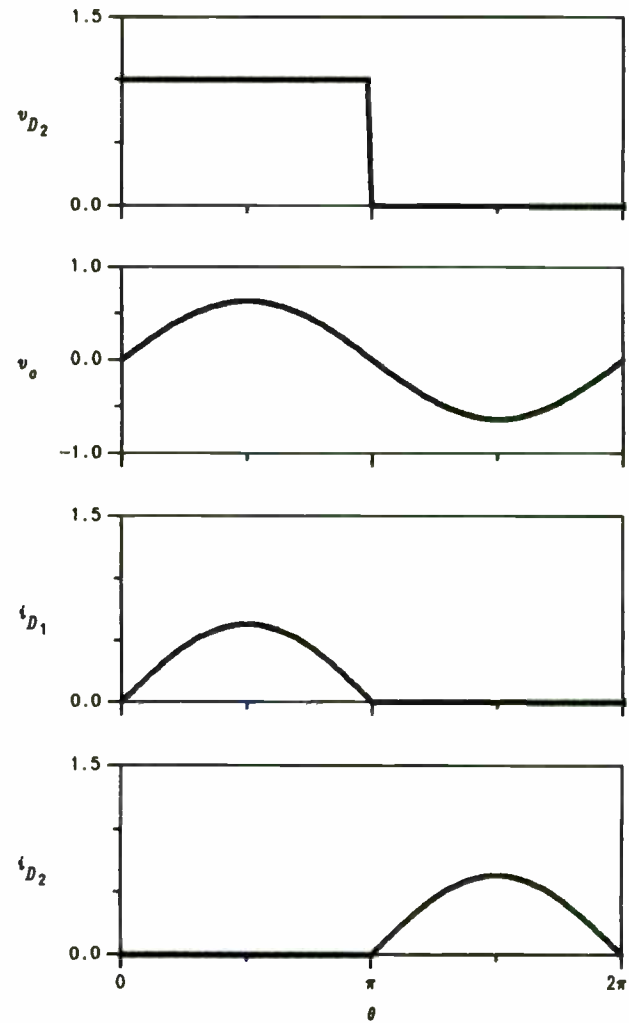


Figure 2. Waveforms of ideal complementary class-D PA.

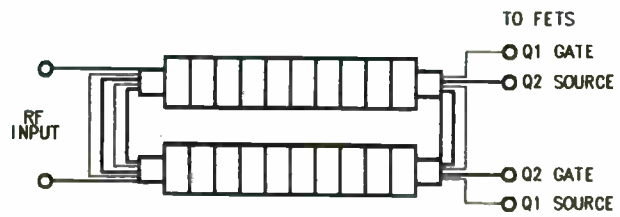


Figure 3. Input transformer.

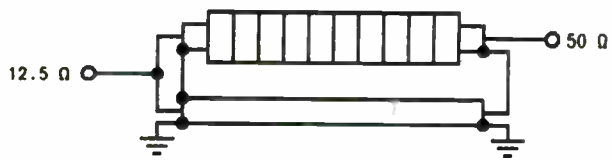


Figure 4. Output transformer.

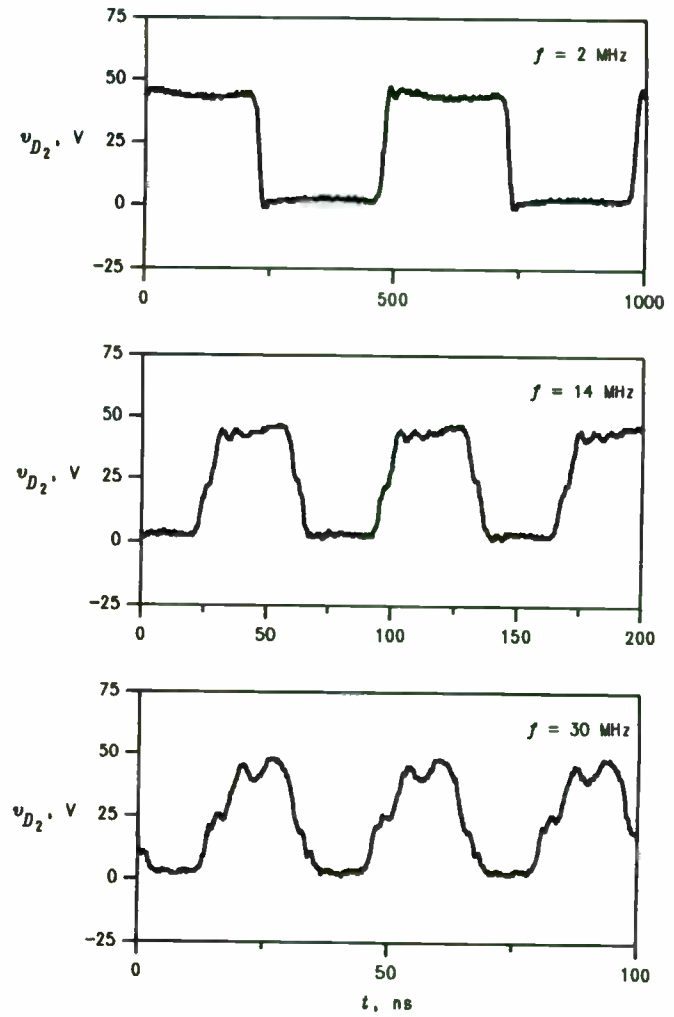


Figure 5. Drain-voltage waveforms.

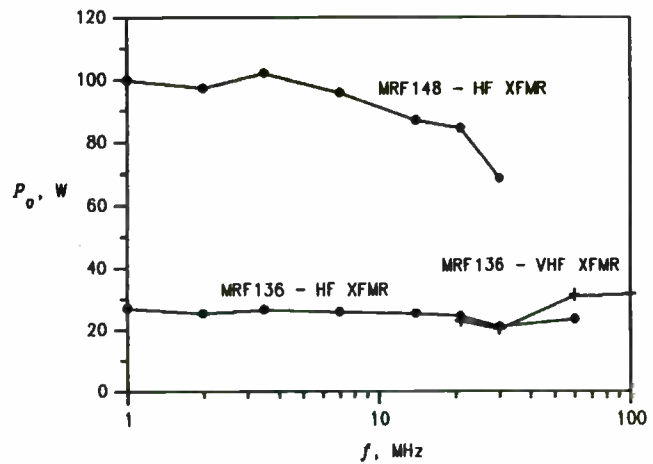
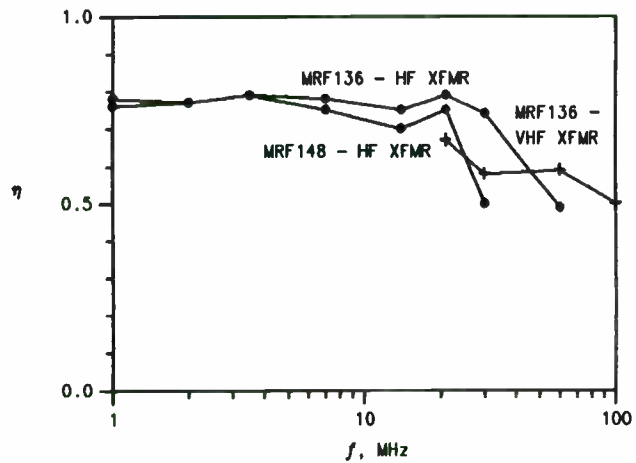


Figure 6. Output power and efficiency.

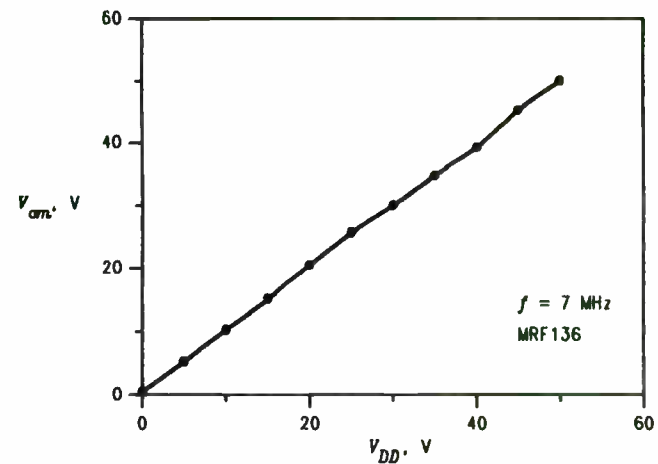


Figure 7. Amplitude-modulation linearity.

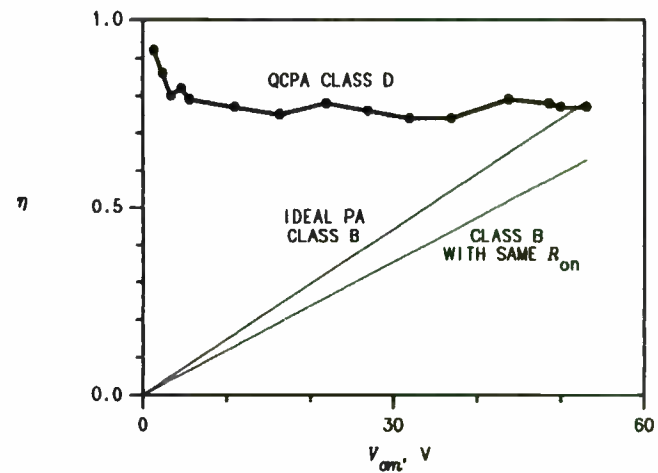


Figure 8. Variation of efficiency with output voltage.

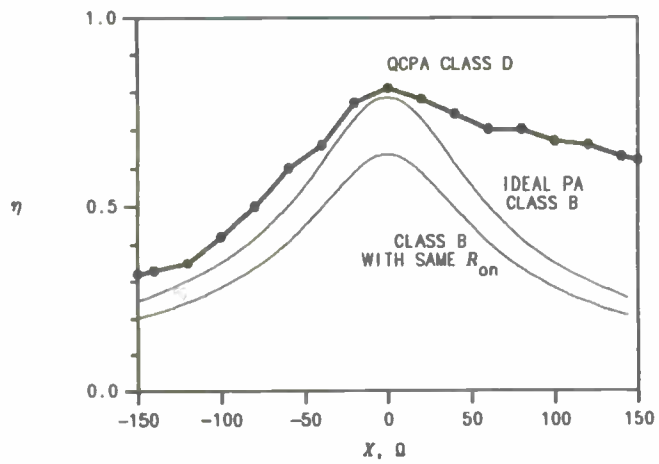


Figure 9. Variation of efficiency with series load reactance.

## DEVCALC – An RF Amplifier Design Aid

Donald J. Miller  
MCC Panasonic  
2001 Westside Parkway Suite 260  
Alpharetta, Georgia 30201

### **Abstract**

The operation and applications of a software program automating some of the more routine calculations involved in the design of small signal RF amplifiers is described. Several application examples are given, the results of which are verified using a commercial circuit analysis package.

### **Introduction**

Numerous complex calculations are necessary to optimally design a small signal RF amplifier. Today, it is fairly common for a designer to arbitrarily select a device, throw a few glue components in for matching, and let an optimizer crank to a solution. If the optimization converges, the design is accepted and the stability is analyzed later. If the optimization does not converge, another device is tried, or a collector resistor is thrown in to "tame" the device. What results may be a design that meets the specification, but the design may not be optimum.

The general purpose computer analysis and optimization tools that exist today are very important engineering tools, but we must avoid being seduced into throwing "just any old input" into these programs and then turning the crank. If we are working on a design that requires optimization, are we not better off starting with a device that we know is capable of meeting the gain objectives? Should we not start with matching

element values that are close to appropriate? The answer, of course is Yes!, but the obstacle that stands in our way is the tedium of the necessary calculations.

Programmable calculator programs that compute stability factors and available gain have been available for some time. What has been lacking is the ability to quickly load device parameters in from data files -- much like the optimization packages do, for quick analysis. Instead of repeatedly keying in S parameters at discrete frequencies for each device under consideration, a PC version could maintain an S parameter library that could be accessed in seconds. An instant summary of the stability and gain characteristics across the S-data band could be generated. "Shopping" for the right device could be fast and painless. Of course, given the frequency of operation, how much effort could it take to have the program tell you how the device should be terminated, what the gain and stability is for arbitrary terminations, or what a port impedance is for a given termination at the other port? What if you could analyze the effects of simple feedback?

DEVCALC evolved as a personal software tool to perform many of the device-oriented calculations required in the design of RF amplifiers. The software is not a circuit analysis tool: it is a DEvice CALCulator. Simple circuits involving collector-base and emitter feedback are handled by DEVCALC by calculating new S parameters. A new virtual device is created which may be analyzed or saved to disk like any other device.

## RF Amplifier Design Methodology

A flowchart of the classical RF amplifier design strategy is shown in figure 1. DEVKALC may be used for all the calculations outlined in the rectangular blocks except for the matching network design. The decision blocks are user decisions that are based on the design requirements. There are fine treatments of this process in the literature, so the treatment here will be brief.

Given the designer's gain objectives, frequency of operation, and other considerations, a candidate device is selected. The Linville stability factor, C is calculated using the device's small signal parameters. If  $C < \text{unity}$  then the device is inherently stable at the frequency of interest and optimum terminations exist. The available gain is calculated, again using small signal parameters. If the gain is sufficient, the optimum terminations are calculated and a suitable matching network is designed to terminate the device properly. If the available gain is not sufficient, then another device must be selected.

If the Linville stability factor for the device is greater than unity, optimum terminations do not exist for the device. It will be impossible to achieve a conjugate match at both ports at the same time. The device may oscillate if you try. To design a stable amplifier, you must mismatch the amplifier on purpose. In general, you can achieve a greater level of stability by a greater level of mismatch. The Stern's stability factor is employed as a measure of a small signal device's stability, given its terminations. Instability results when K is less than unity. Values greater than unity are an indication of successively greater stability. Many designers start with a K of greater than 4 or 5 to insure comfort.

Design of an RF amp using a potentially unstable device is a trade-off between stability and gain. K is calculated from the small signal device parameters and the device

terminations. For a given K, terminations can be calculated. If the mismatch is split up equally between the input and output of the device, the real part of the terminations can be calculated in a straightforward manner. The imaginary parts require complex calculation or iterative methods.

Once the terminations for a given K are calculated, the transducer gain is computed and compared with the gain objective. If the gain is too low, you can reduce K and try again if you have comfort room to do so (if not, pick another device!). If the gain is too high, you may arbitrarily set K as high as necessary. Once the terminations are computed, matching networks are designed as in the inherently stable case.

## Theory Of Operation

DEVKALC must have small signal parameter data as input for its calculations. Some of the calculations employ the Y parameter model, while some of them employ S parameters. Either S or Y data may be supplied by the user, and the software will calculate the other set. The equations employed for parameter conversion follow:

$$\begin{aligned}
 S_{11} &= \frac{(Z'_i - 1)(Z'_o + 1) - Z'_i Z'_o}{(1 + Z'_i)(1 + Z'_o) - Z'_i Z'_o} & Z'_i &= \frac{(1 + S_{11})(1 - S_{22}) + S_{12} S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12} S_{21}} \\
 S_{22} &= \frac{(Z'_i + 1)(Z'_o - 1) - Z'_i Z'_o}{(1 + Z'_i)(1 + Z'_o) - Z'_i Z'_o} & Z'_o &= \frac{(1 - S_{11})(1 + S_{22}) + S_{12} S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12} S_{21}} \\
 S_{12} &= \frac{2Z'_i}{(1 + Z'_i)(1 + Z'_o) - Z'_i Z'_o} & Z'_i &= \frac{2S_{12}}{(1 - S_{11})(1 - S_{22}) - S_{12} S_{21}} \\
 S_{21} &= \frac{2Z'_o}{(1 + Z'_i)(1 + Z'_o) - Z'_i Z'_o} & Z'_o &= \frac{2S_{21}}{(1 - S_{11})(1 - S_{22}) - S_{12} S_{21}}
 \end{aligned}$$

$$Z' = \frac{Z}{50}$$



Both sets of parameters are available for printout, though only S parameter file I/O is supported. After the parameter calculation, the software immediately does a gain and stability analysis at each S parameter frequency. The Linville stability factor is computed:

$$C = \frac{|Y_r| |Y_l|}{2 \operatorname{Re}\{Y_l\} \operatorname{Re}\{Y_o\} - \operatorname{Re}\{Y_r Y_l\}}$$

If  $C > 1$ , then the device is potentially unstable and a warning is issued next to the frequency of calculation on the device summary screen. If  $C < 1$ , then the device is inherently stable, and the optimum terminations -- those that produce a conjugate match -- are computed:

$$G_{l(\text{OPT})} = \frac{\sqrt{((2g_l g_o - P)^2 - (P^2 + Q^2))}}{2g_l}$$

$$B_{l(\text{OPT})} = \frac{Q}{2g_l} - b_o$$

$$G_{o(\text{OPT})} = \frac{\sqrt{((2g_l g_o - P)^2 - (P^2 + Q^2))}}{2g_o}$$

$$B_{o(\text{OPT})} = \frac{Q}{2g_o} - b_l$$

Where  $y_r y_l = P + jQ$ .

The available gain, or maximum transducer gain, as well as the maximum available gain, is computed at each S parameter frequency and summarized on the screen:

$$G_{T_{\text{MAX}}} = G_A = \frac{|Y_r|^2 G_o}{G_l |Y_l + Y_o|^2}$$

$$\text{MAG} = \frac{|Y_r|^2}{4g_l g_o}$$

For frequencies where the device is not inherently stable, the maximum transducer gain and optimum terminations are not computed and displayed on the device summary screen.

DEVCALC provides an "optimum termination" screen where the best terminations can be calculated for any input frequency between the minimum and maximum frequency for which small signal parameter data is available. The small signal parameters for calculation are determined by linear interpolation on the magnitude and angles of the S parameter data. If the device is inherently stable, optimum terminations are computed in the same way they are for the device summary screen. If the device is potentially unstable, the program prompts the user for a Stern's stability factor as a design parameter. The real parts of the terminations are computed from the following equations:

$$G_o = \sqrt{\frac{K[|Y_r Y_l| + \operatorname{Re}\{Y_r Y_l\}] g_l}{2g_o}}$$

$$G_l = \sqrt{\frac{K[|Y_r Y_l| + \operatorname{Re}\{Y_r Y_l\}] g_o}{2g_l}}$$

The imaginary parts are not as straightforward. Direct calculation would be quite complex, so an iterative method is employed. First, it is assumed that the device is terminated with  $Y_L = G_L - jB_O$ , where  $B_O = \operatorname{Im}\{Y_O\}$ . The admittance looking into the input of the device with the  $Y_L$  termination on the output is computed by:

$$Y_I = Y_l - \frac{Y_r Y_o}{Y_o + Y_l}$$

The input of the device is now assumed to be terminated with  $Y_S = G_S - jB_I$ , where  $B_I = \operatorname{Im}\{Y_I\}$ . The admittance looking into the output is now calculated:

$$Y_O = Y_o - \frac{Y_r Y_n}{Y_i + Y_s}$$

Now, the output is assumed to be terminated with  $Y_L = GL - jBO$ , where  $BO$  is the imaginary part of the computed output admittance. This process is iteratively and alternately applied to the input and output of the device until the difference between successively computed terminations is suitably small.

The power gain analysis screen takes in frequency and terminations as data, and computes the Stern's stability factor, the alignability, the available gain, the operating power gain, the transducer power gain, and the maximum available gain. The Stern's stability factor is a measure of stability margin for a particular device termination and is computed with the following equation:

$$K = \frac{2 (g_i + G_s) (g_o + G_l)}{|Y_r Y_n| + R_c (Y_r Y_n)}$$

The alignability, which is a measure of how tuning on one amplifier port affects the tuning of the other one, is computed:

$$d = \frac{|Y_L| |Y_n Y_r|}{|Y_o + Y_L| |Y_i (Y_o + Y_L) - Y_n Y_r|}$$

The available gain is the gain realized with conjugate terminations on both ports:

$$GA = \frac{|Y_r|^2 G_s}{R_c [ (Y_i Y_o - Y_n Y_r + Y_o Y_s) (Y_i + Y_s)^* ]}$$

The power delivered to the load divided by the power into the input port of the device, or the operating power gain, is computed and displayed:

$$G_p = \frac{|Y_r|^2 G_L}{|Y_L + Y_o|^2 G_i}$$

The transducer power gain computed by the power gains menu option is different from that presented by the device characteristic summary in that optimum terminations are not assumed, the user's termination data input is used:

$$G_T = \frac{4 |y_r|^2 G_L G_s}{| (y_i + Y_o) (y_o + Y_L) - y_r y_l |^2}$$

MAG is computed using the same method as in the device characteristic summary screen.

DEVCALC provides a menu option for computing the device input admittance or impedance when the termination on the other port is determined. The equations used are:

$$Y_I = Y_i - \frac{Y_r Y_n}{Y_o + Y_L}$$

$$Y_O = Y_o - \frac{Y_r Y_n}{Y_i + Y_s}$$

Simple feedback may be added to the device for analysis. DEVCALC models collector to base feedback by adding the feedback admittance to  $Y_r$  and  $Y_l$  of the device.

Emitter feedback is modeled by temporarily converting the small signal parameters to "T" parameters, adding the emitter impedance in series with the emitter port, and then re-computing the S and Y parameters. After feedback is added to a device, a new virtual device is created with its own parameters.

Output stability circles are calculated with the following operations:

$$D = S_{11} S_{22} - S_{12} S_{21}$$

$$C_{22} = \frac{(S_{22} - S_{11}^* D)^*}{|S_{22}|^2 - |D|^2}$$

$$R_{2S} = \frac{|S_{12}S_{21}|}{|S_{22}|^2 - |D|^2}$$

$C_{2S}$  is the reflection coefficient of the center of the circle, and  $R_{2S}$  is the radius. Input stability circles are computed with the 1 and 2 subscripts reversed in the above equations.

### **Application Examples**

#### **Inherently Stable Device**

Suppose we would like to design an amplifier at 1 GHz with a gain of at least 10 dB. In our application, input and output return loss shall be important, so both amplifier ports must be conjugately matched. Our task is to find a device which is inherently stable at 1 GHz with sufficient gain, and then to terminate the device ports to realize this gain.

The DEVCALC main operation menu appears in figure 2. If we want to load device data into the program manually, we will select item #1, "Load device parameters". If an S parameter data file already exists for the device we want to analyze, menu item #8, "Get Device Params from Disk" will read the data into the program for us.

The first device we will try will be an MRF536 bipolar transistor. We select menu item #8, type in the filename of the S parameter file, MRF536MO.005, and the program displays a summary of the S and Y parameter data (figure 3). After pushing <Enter> to continue, a gain and stability analysis of the device is presented (figure 4). At 800 MHz, the device is stable, with maximum transducer gain of 10.8 dB. At 1200 MHz, the device only has an available gain of 7.5 dB. We are not sure if the gain at 1 GHz is sufficient or not, so we will need to investigate further.

We can analyze the device performance at 1 GHz by selecting menu item #4,

"Optimum Terminations", having DEVCALC compute the best port terminations, and then analyzing the amplifier gain with menu item #5, "Determine Power Gains". Figure 5 is the result of the program's termination calculation. The computed terminations are entered as data into the power gain computation screen (figure 6) where we see that a gain of only 8.94 dB is attainable at this frequency. By now, we have spent a total of maybe three minutes investigating this device and we know we must discard it.

Our next candidate device is a BFR91 transistor. The gain and stability analysis screen (figure 7) shows us instantly that we have enough stable gain at 1 GHz. The maximum transducer gain is 11.4 dB and  $C < 1$ . We need only figure out the terminations. Figure 8 shows DEVCALC's computed terminations, which we know should give us a gain of 11.4 dB, but we go on to the power computation part of the program anyway, just to verify the results (figure 9). Suitable matching networks are computed (using other software tools or a Smith chart), and the circuit in figure 10 results. For verification, an RF circuit simulator was used to check the results. The circuit file and output plots appear in figure 11.

#### **Potentially Unstable Device**

In this example, we will wish to design a 500 MHz amplifier with a gain of 15 dB. The amplifier is to be used in a system with good broadband 50 ohm terminations on each side and return loss on both ports is not of particular concern. The stockroom has a surplus of MRF536 transistors, so we decide to try using one. We load or enter the S parameter file into DEVCALC for the device, and find that the device may not be inherently stable at 500 MHz, at least we know that potential instability exists at 400 MHz (figure 12). This does not turn us away (purchasing is begging us to use these transistors!), so we attempt to use the "Optimum Terminations" menu item. We enter

our frequency of interest of 500 MHz, and the program computes a Linville stability factor of greater than unity. DEVCALC queries: Device potentially unstable:  $K = \infty$ . The software is asking us for a target Stern's stability factor. Terminations will be calculated with just enough mismatch to achieve this level of stability. The higher the K value we choose, the lower the gain, but the more stable the circuit. A value of 4 or 5 is typical, but we are perhaps a bit paranoid and we choose 10 (figure 13). We take the computed terminations, use them as input to the program to compute gain, and we find we have realized a transducer power gain of 15.9 dB (figure 14).

Since 15.9 dB is slightly more gain than we need, we have the option of increasing our degree of mismatch and lowering our gain. We try a K of 20 (figure 15), and realize a gain of 15 dB (figure 16). Matching network elements are computed, and the circuit in figure 17 results. A circuit file and simulation output appear in figure 18. Again, there is excellent agreement between the software tools.

### Constrained Design

In the final example, we wish to design a low noise amplifier at 1 GHz. The only design restrictions are that we use an MRF572 with a collector current of 5 mA, and that we try to achieve the lowest noise figure possible. We load the MRF572 S parameters into DEVCALC and the gain and stability analysis (figure 19) indicates that the device is potentially unstable at 1 GHz. We will not be able to use the optimum terminations section of the software to find the best terminations for this device. Even if the device was inherently stable, computing the terminations for conjugate match would not produce the desired results.

In general, the source impedance that will result in the lowest noise figure for a device will not provide the best impedance match or highest gain. The tact is often to terminate the input with an impedance recommended by the manufacturer, and then

attempt to conjugately match the output of the device. The Stern's stability factor is then evaluated, and if the design has sufficient stability we are done. If K is too low, then we must mismatch the output as well.

Figure 20 is a page from the Motorola data sheet for the MRF572 bipolar transistor. The source reflection coefficient for optimum noise figure at 1 GHz is specified. We can calculate the desired source conductance to be  $0.0167 - j0.0245$  mhos. DEVCALC will calculate the output conductance of the device with this source termination if we use menu item #6. The results of the program's computed output conductance (and impedance) is shown in figure 21. There is now enough information to design the circuit, but first we should find out how much gain we can expect. The terminating conductances are keyed into the gain computation screen, and the output of figure 22 results. We should be able to get approximately 11.8 dB of gain from this low noise amp. A schematic of this amp and a Superstar simulation verifying this result appear in figure 23 and 24.

### Other Features

DEVCALC has the ability to compute input and output stability circles and display them on the screen. This feature is handy for getting a feel for the possibility of out-of-band oscillations. A network analyzer polar plot of a device's termination impedance can be compared to the stability circle display at various frequencies of interest to insure that the terminations are always in the "safe" region.

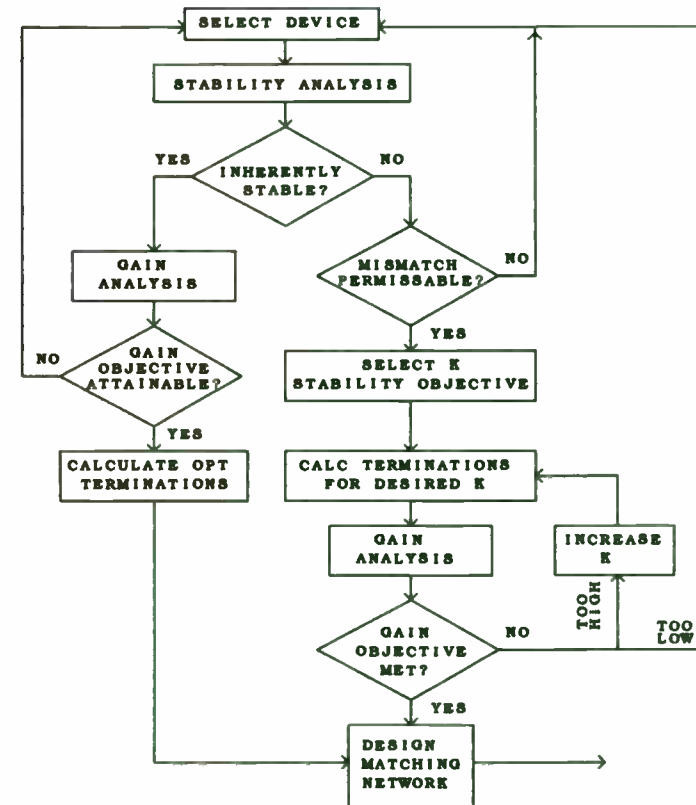
Simple feedback may be added from the collector to base of a device, or in series with the emitter (or both!) and the effects analyzed. The new S parameter set can be analyzed like any other device.

### Summary

DEVCALC is a handy device oriented analysis tool that can be used to perform many of the tedious calculations involved in the design of RF amplifiers. The program can be used to "shop" for appropriate devices for various applications or further optimization. DEVCALC, in conjunction with appropriate matching network design tools, can be used directly to design narrowband RF amplifiers. For wideband designs, DEVCALC can be employed to determine the starting point for optimization tools.

### References

1. Carson, Ralph S., *High Frequency Amplifiers*, John Wiley and Sons, 1975.
2. Krauss, Herbert L. et al., *Solid State Radio Engineering*, Wiley, 1980.
3. Hertling, D. R. and Feeney, R. K., *Design of RF Amplifiers*, RF Design, Cardiff Publishing, March/April and May/June 1982.



AMPLIFIER DESIGN PROCEDURE

Figure 1

Operation Menu

- 1) Load device parameters.
- 2) Summarize Dev Param Input.
- 3) Summarize Device Characteristics.
- 4) Optimum terminations.
- 5) Determine Power Gains.
- 6) Determine Port Impedance.
- 7) Save Dev Params to Disk.
- 8) Get Device Params from Disk.
- 9) Add C-B feedback impedance.
- A) Add emitter impedance.
- B) Input Stability Circle.
- C) Output stability circle.
- D) Disk Directory.
- E) Exit to DOS.

Figure 2

```
Freq magS11 angS11 magS21 angS21 magS12 angS12 magS22 angS22
+.40E+09+.48E+00-.54E+02+.51E+01+.11E+03+.86E-01+.58E+02+.54E+00-.39E+02
+.80E+09+.28E+00-.66E+02+.29E+01+.88E+02+.14E+00+.53E+02+.42E+00-.55E+02
+.12E+10+.17E+00-.97E+02+.21E+01+.73E+02+.18E+00+.48E+02+.39E+00-.62E+02
+.16E+10+.15E+00-.12E+03+.17E+01+.64E+02+.21E+00+.44E+02+.33E+00-.72E+02
+.20E+10+.11E+00-.16E+03+.14E+01+.55E+02+.24E+00+.39E+02+.31E+00-.90E+02

Freq magY11 angY11 magY21 angY21 magY12 angY12 magY22 angY22
+.40E+09+.69E-02+.56E+02+.87E-01-.44E+02+.14E-02-.95E+02+.46E-02+.61E+02
+.80E+09+.97E-02+.48E+02+.62E-01-.61E+02+.30E-02-.96E+02+.81E-02+.67E+02
+.12E+10+.14E-01+.43E+02+.53E-01-.74E+02+.47E-02-.99E+02+.11E-01+.72E+02
+.16E+10+.17E-01+.42E+02+.51E-01-.81E+02+.63E-02-.10E+03+.14E-01+.69E+02
+.20E+10+.19E-01+.39E+02+.51E-01-.89E+02+.86E-02-.10E+03+.19E-01+.70E+02
```

Press any key to continue...

Figure 3

MRF536 Common Emitter Vce=10V Ic=5mA

Gain and Stability Analysis

Freq (Hz)	Opt load Conduct (mhos)	Opt load Suscept (j $\mu$ hos)	Opt Src Conduct (mhos)	Opt Scr Suscept (j $\mu$ hos)	C	Max. Trans. Gain (dB)	MAG (dB)
0.400E+09	Potentially Unstable! C = 0.113E+01						
0.800E+09	0.809E-02	-.129E-01	0.164E-01	-.182E-01	0.9	10.8	16.6
0.120E+10	0.949E-02	-.121E-01	0.290E-01	-.146E-01	0.8	7.5	13.2
0.160E+10	0.122E-01	-.125E-01	0.315E-01	-.105E-01	0.7	5.3	10.0
0.200E+10	0.146E-01	-.144E-01	0.341E-01	-.436E-02	0.7	3.9	8.3

Press any key to continue...

Figure 4

MRF536 Common Emitter Vce=10V Ic=5mA

Freq: 1E9	C= 8.24E-01
Gs(opt)= 2.30E-02	Bs(opt)= -1.71E-02
G1(opt)= 9.24E-03	B1(opt)= -1.24E-02

Figure 5

MRF536 Common Emitter Vce=10V Ic=5mA

Termination Dependant Parameters

Freq: 1E9

Source Termination --- Gs= 2.3E-2 Bs= -1.71E-2 (Mhos)

Load Termination --- G1= 9.24E-3 B1= -1.24E-2 (Mhos)

K= 1.02E+02 Align= 7.06E-01

Ga= 8.94E+00 Gop= 8.94E+00 Gt= 8.94E+00 MAG= 1.48E+01 (dB)

Figure 6

BFR91 Common Emitter Vce=10V Ic=5mA

Gain and Stability Analysis

Freq (Hz)	Opt load Conduct (mhos)	Opt load Suscept (jμhos)	Opt Src Conduct (mhos)	Opt Scr Suscept (jμhos)	C	Max. Trans. Gain (dB)	MAG (dB)
0.200E+09	Potentially Unstable!				C = 0.189E+01		
0.500E+09	Potentially Unstable!				C = 0.108E+01		
0.800E+09	0.262E-02	-.580E-02	0.124E+00	0.636E-02	0.9	12.7	22.5
0.100E+10	0.130E-02	-.607E-02	0.493E-01	0.103E+00	1.0	11.4	19.9
0.150E+10	0.375E-02	-.711E-02	0.392E-01	0.496E-01	0.9	7.2	10.9

Press any key to continue...

Figure 7



BFR91 Common Emitter Vce=10 Ic=5mA

Freq: 1E9	C= 9.86E-01
Gs(opt)= 4.93E-02	Bs(opt)= 1.03E-01
G1(opt)= 1.30E-03	B1(opt)= -6.07E-03

Figure 8

BFR91 Common Emitter Vce=10 Ic=5mA

Termination Dependant Parameters

	Freq: 1E9		
Source Termination ---	Gs= 4.93E-2	Bs= 1.03E-1	(Mhos)
Load Termination ---	G1= 1.3E-3	B1= -6.07E-3	(Mhos)
	K= 1.12E+01	Align= 1.74E+00	
Ga= 1.14E+01	Gop= 1.14E+01	Gt= 1.14E+01	MAG= 1.99E+01 (dB)

Figure 9

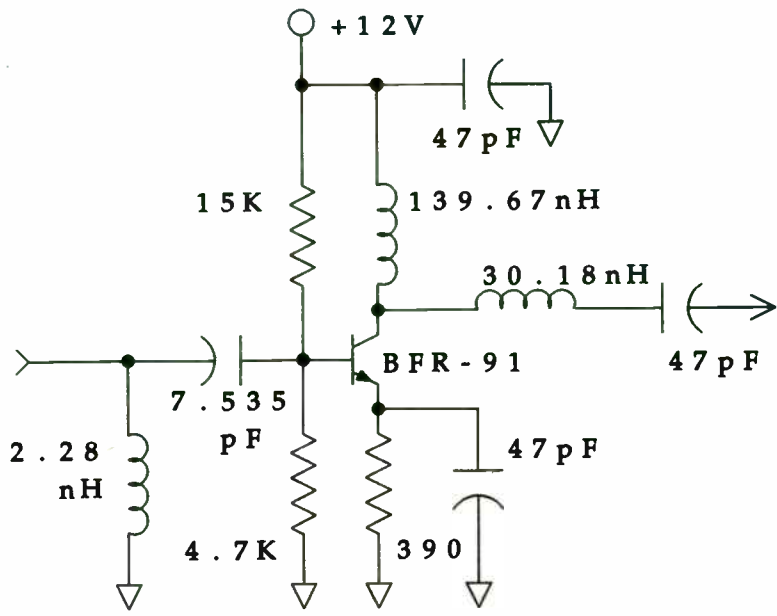
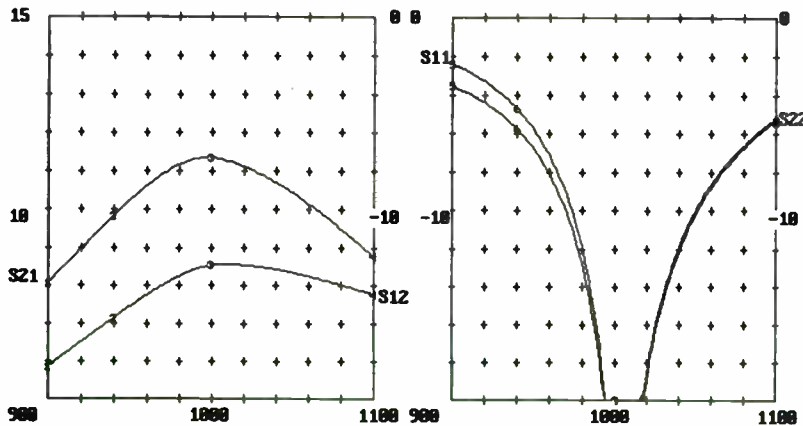


Figure 10



This file is : EXAMPLE1

CIRCUIT  
 IND AA PA 2.28  
 CAP BB SE 7.535  
 RES CC PA 3579  
 TWO DD SP 50 '\SS\DATA\BFR91  
 IND EE PA 139.67  
 IND FF SE 30.176  
 CAP GG SE 47  
 CAX AA GG

NUM OF EXAMPLE1  
 900 940 =1000 1100 900 940 1000 1100  
 8.06976 9.8381 11.3783 8.77899 -2.37285 -4.67835 -38.3389 -5.2523  
 -18.3859 -15.6839 -12.8912 -14.414 -3.52973 -5.71466 -38.8771 -5.36175  
 ↑ Mkr Freq ↓ ← Select Mkr → F5=MkrToggle Alt-F7=ScrnPrt F10=Main Menu  
 F2=Redistribute New Markers F6=Add/Update MARKER Block F8=Enter Value

OUTPUT  
 GPH AA S21 50 5 15  
 GPH AA S12 50 -20 0  
 GPH AA S11 50 -20 0  
 GPH AA S22 50 -20 0  
 FREQ  
 SWP 900 1100 101

Figure 11



MRF536 Common Emitter Vce=10V Ic=5mA

Termination Dependant Parameters

Freq: 500E6

Source Termination ---	Gs= 1.16E-2	Bs= -1.87E-2	(Mhos)
Load Termination ---	G1= 6.5E-3	B1= -1.19E-2	(Mhos)
	K= 1.00E+01	Align= 7.56E-01	
Ga= 1.61E+01	Gop= 1.61E+01	Gt= 1.59E+01	MAG= 2.16E+01 (dB)

Figure 14

MRF536 Common Emitter Vce=10V Ic=5mA

Freq: 500E6

C= 1.05E+00	
Gs( K )= 1.82E-02	Bs( K )= -1.54E-02
G1( K )= 1.02E-02	B1( K )= -1.00E-02
K= 20	
Error func = 7.35851E-07	

Figure 15

MRF536 Common Emitter Vce=10V Ic=5mA

Termination Dependant Parameters

Freq: 500E6

Source Termination ---	Gs= 1.82E-2	Bs= -1.54E-2	(Mhos)
Load Termination ---	G1= 1.02E-2	B1= -1.0E-2	(Mhos)
	K= 2.00E+01	Align= 6.10E-01	
Ga= 1.54E+01	Gop= 1.54E+01	Gt= 1.50E+01	MAG= 2.16E+01 (dB)

Figure 16

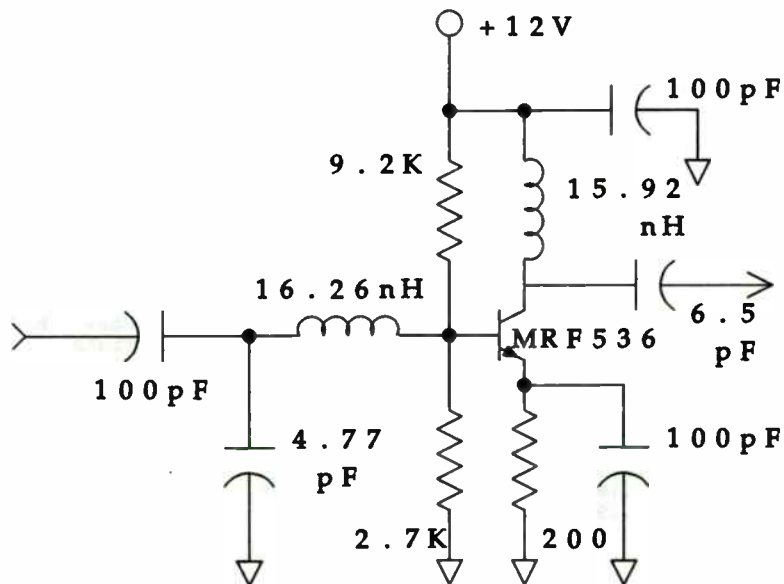
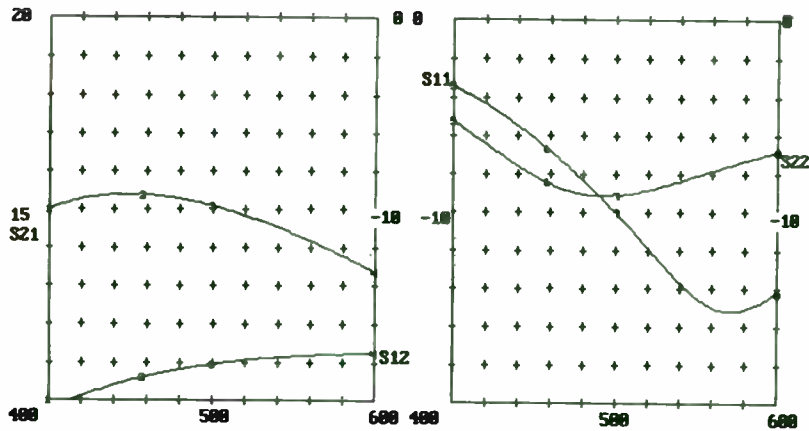


Figure 17



This file is : EXAMPLE2

```
CIRCUIT
CAP AA SE 100
CAP BB PA 4.77
IND CC SE 16.262
RES DD PA 2087
TWO EE SP 50 '\SS\DATA\MRF536
IND FF PA 15.919
CAP GG SE 6.4948
CAX AA GG
```

```
NUM OF EXAMPLE2          ===SUPERSTAR===          Thu Jul 82 22:11:53 1992
400      450      *500      600      400      450      500      600
14.9821  15.3893  15.8935  13.3714  -3.38265 -6.61427 -10.8133 -14.2117
-28.5472 -18.7621 -18.8755 -17.4859 -5.17981 -8.39798 -9.18595 -6.84668
? Mkr Freq ↓ * Select Mkr * F5=MkrToggle Alt-F7=ScrnPrt F10=Main Menu
F2=Redistribute New Markers F6=Add/Update MARKER Block FB=Enter Value
```

```
OUTPUT
GPH AA S21 50 10 20
GPH AA S12 50 -20 0
GPH AA S11 50 -20 0
GPH AA S22 50 -20 0
FREQ
SWP 400 600 101
```

Figure 18

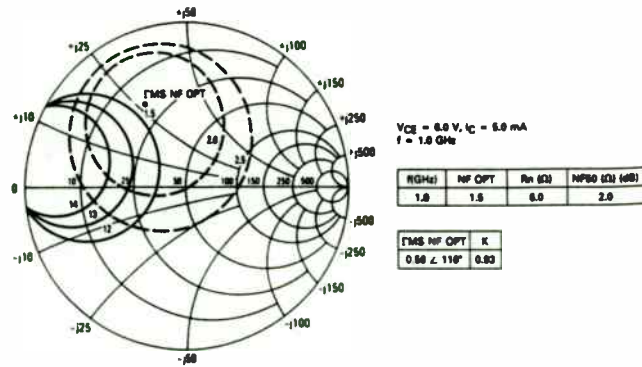
MRF572 Common Emitter Vce=8V Ic=5mA

Gain and Stability Analysis

Freq (Hz)	Opt load Conduct (mhos)	Opt load Suscept (jmhos)	Opt Src Conduct (mhos)	Opt Scr Suscept (jmhos)	C	Max. Trans. Gain (dB)	MAG (dB)
0.200E+09		Potentially Unstable!	C = 0.532E+01				
0.500E+09		Potentially Unstable!	C = 0.245E+01				
0.100E+10		Potentially Unstable!	C = 0.121E+01				
0.150E+10	0.367E-02	-.186E-01	0.245E+00	0.206E+00	1.0	12.7	22.4
0.200E+10	0.830E-02	-.190E-01	0.646E-01	0.934E-01	0.8	9.4	9.6

Press any key to continue...

Figure 19



Reprinted with permission of Motorola, Inc.

Figure 20

MRF572 Common Emitter  $V_{ce}=8V$   $I_c=5mA$

Freq	Gs	Bs	Go	Bo	Ro	Xo
1e9	.016688	-.024474	1.78E-02	1.51E-02	3.26E+01	-2.77E+01

Figure 21

MRF572 Common Emitter Vce=8V Ic=5mA

Termination Dependant Parameters

Freq: 1e9

Source Termination --- Gs= .0166881 Bs= -.02447416 (Mhos)

Load Termination --- Gl= 1.78e-2 Bl= -1.51e-2 (Mhos)

K= 1.23E+01 Align= 6.72E-01

Ga= 1.18E+01 Gop= 1.49E+01 Gt= 1.18E+01 MAG= Undefined (dB)

Figure 22

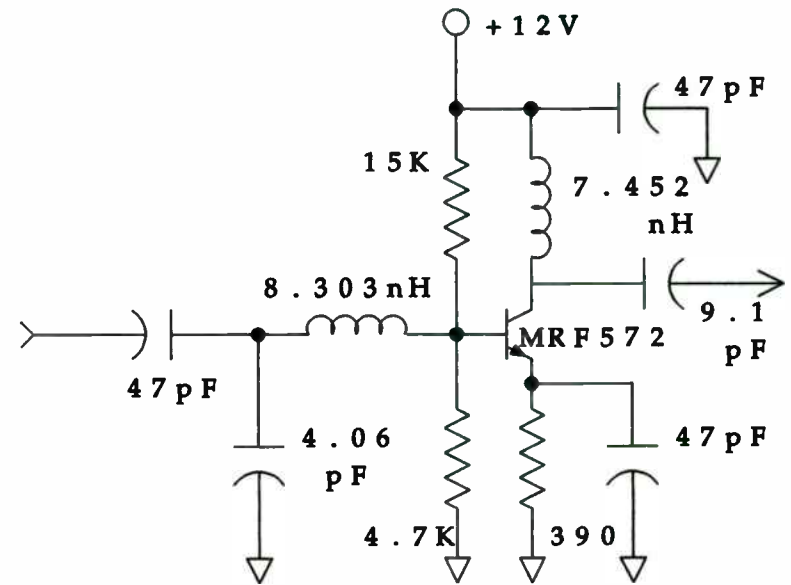
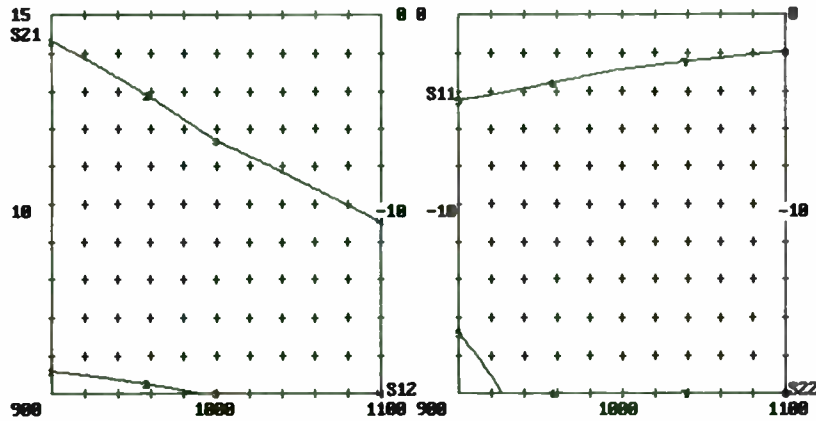


Figure 23





This file is : example3

```

CIRCUIT
CAP AA SE 47
CAP BB PA 4.063
IND CC SE 8.303
RES DD PA 3579
TWO EE SP 50 '\SS\DATA\mrf572
IND FF PA 7.452
CAP GG SE 9.053
CAX AA GG
  
```

```

RUN OF example3          ==SUPERSTAR==          Tue Jul 28 20:44:24 1992
900      950      *1000      1100      900      950      1000      1100
14.3351  12.8932  11.6987  9.5289  -4.47768 -3.55307 -2.45477 -1.98836
-10.8825 -19.4682 -20.1858 -21.4654 -16.7766 -25.5113 -20.9129 -23.7983
↑ Mkr Freq ↓  ← Select Mkr →  F5=Mkr Toggle  Alt-F7=ScrnPrt  F10=Main Menu
F2=Redistribute New Markers  F6=Add/Update MARKER Block  F8=Enter Value
  
```

```

OUTPUT
GPH AA S21 50 5 15
GPH AA S12 50 -20 0
GPH AA S11 50 -20 0
GPH AA S22 50 -20 0
FREQ
SWP 900 1100 101
  
```

Figure 24



**Noise Measurements on Microwave Devices  
Using a Scalar Analyzer**

Stuart A. Fox  
Giga-tronics, Inc.  
Power Measurements Division  
488 Tasman Drive, Sunnyvale, California 94089

**Abstract**

Noise measurements on microwave devices such as amplifiers are generally made using a noise figure meter and some form of down converter. While providing accurate results, the procedure is slow and expensive for manufacturing requirements.

Noise figure can be measured accurately with a power meter using the Y-Factor technique. This paper will describe a Noise Figure Test Set for use with a power meter accurate Scalar Analyzer that provides direct and accurate measurement of noise figure. Techniques that allow a single connection to a 2 port test set to measure noise figure, gain, output power and return loss will be described. Good measurement practices for accurately measuring noise figure and power will be reviewed.

**Introduction**

Manufacturers of high performance amplifiers face several problems resulting from using traditional measurement techniques. One example is increased competition from new companies using the latest equipment to reduce cost or

increase performance of amplifiers. Another problem is that purchasers of amplifiers are demanding more test data but without expecting to pay for it. For example, noise figure might be required at many different frequencies in a broadband amplifier. This will increase testing time over the previously accepted sparsely spaced measurement points. A survey indicated that 60% of the cost of an amplifier in the \$1,000 price range was due to alignment and test, see Table1.

At the same time, manufacturers are striving to improve the quality of their devices without incurring additional costs. To achieve the goals of more test data and improved quality without an increase in cost requires significant improvements in productivity.

The same set of measurements is made on modern amplifiers as on previous generations. Examples are: gain; isolation; VSWR; 1 dB compression power; noise figure and third order intercept. Typically, these measurements take two or more test stations. One test station may have a scalar analyzer and CW power meter. Very often the noise figure meter and spectrum analyzer will be on a second bench.

Productivity improvements can be made by reducing the amount of equipment used to make the measurements and making all measurements at one test station.

By improving the power measurement accuracy of the scalar analyzer to be equal to that of a CW power meter, this latter instrument would be made redundant. The solution developed at Giga-tronics consists of incorporating a

thermistor based calibrator into a scalar analyzer. Under microprocessor control, the user can calibrate any compatible diode sensor to have the same linearity as the thermal based sensors used with CW power meters. The calibration scheme includes the ability to set a 1mW reference, traceable to NIST.

Now a scalar analyzer can be used to make accurate power measurements, either CW or swept, using the same sensors as for the traditional scalar measurements. This capability improves productivity, costs less than separate instruments and is less expensive to maintain and calibrate. In addition, swept power measurements inherently provide more data.

#### **Noise Figure Measurements Require A Narrow Bandwidth System.**

It has been shown that Noise Figure can be calculated by measuring two output power levels corresponding to the hot and cold temperatures of a noise source on the input of a DUT. Some people use a conventional CW power meter and a calculator to determine the Noise Figure.

Just as power, gain and VSWR are measured on a frequency-by-frequency basis, the same requirement is true for noise figure. However, unlike signal generators and sweepers which generate a single frequency at any given time, noise sources produce continuous energy over their entire frequency range - hence the need for a measurement receiver. Noise Figure meters are measurement receivers specifically designed to measure Noise Figure. They are measuring the two output power levels corresponding to the hot and cold temperatures of a noise source on the input of a DUT. Then a microprocessor is

used to calculate Noise Figure. The obvious disadvantage to this approach is that additional equipment is required to measure gain, VSWR and power.

Since Noise Figure measurements are based on accurate power measurements, a precision scalar analyzer could be used for this purpose with the advantage that the same instrument could be used to measure gain, VSWR and power. However scalar detectors are broadband devices so a down converter is required for the noise signal. Then, measurements can be made one frequency at a time while sweeping over a desired range.

#### **Test Set Adds Noise Figure to Scalar Display**

Measurement of Noise Figure at microwave frequencies can be very expensive. Giga-tronics approach was to build a test set specifically for measuring broadband microwave devices such as MMIC amplifiers. Then, double sideband measurements would be as accurate as single sideband measurements. Furthermore, there would be no need for a tracking filter on the input (or multiple conversion) to eliminate the image response. This test set approach improves measurement accuracy by maintaining a low system noise figure - a parameter which can have a significant impact when measuring the noise figure of low gain devices.

The Giga-tronics precision scalar analyzer is used to measure the test set IF output signal levels,  $N_1$  and  $N_2$  corresponding to the  $T_c$  and  $T_h$  noise power levels. The microprocessor in the scalar analyzer is then used to calculate Noise Figure in the same way as in a Noise Figure meter.

To make accurate measurements of noise requires using detectors that are true RMS responding. This is achieved with the Giga-tronics scalar analyzer by ensuring that the noise power levels are in the true RMS region of the diode's response, that is, below - 20 dBm.

Table 2 shows a comparison of measured results using a precision scalar analyzer based noise figure measurement system and a conventional Noise Figure meter. It can be seen that the differences are in the order of a few hundredths of a dB - well within the measurement uncertainty for either system.

Figure 1 is a block diagram of the original test set. All signals are converted down to an IF of 30 MHz. IF bandwidth is 5 MHz giving a frequency resolution of  $2 \times (30 + 5) = 70$  MHz. This test set made possible the first truly scalar based measurement of the 4 key amplifier characteristics. Figure 2 is a plotter output of gain, Noise Figure,  $P_{c1dB}$  and return loss, all measured in the swept mode using the Giga-tronics amplifier test system. However, from a speed of measurement viewpoint, it suffered from the disadvantage of requiring the DUT to be transferred between two test ports depending on the parameter being measured.

This problem was eliminated by making the test set internally switch the input to the DUT between the noise source and the sweeper. Since it is common practice to have devices such as isolators between the noise source and the DUT to improve the match, it was found to be perfectly acceptable to switch the noise source and make allowance for the path loss between the noise source and the DUT. This path loss is accurately measured and subtracted from the Noise Source ENR to give corrected ENR values at the DUT.

In addition, this approach permits a one-time-connect-and-measure approach to device characterization thereby saving considerable time. Another advantage of this approach is much better repeatability when switching between Noise Figure and gain measurements while tuning the DUT. All switching is done electro-mechanically in the test set. The DUT is left in-situ which preserves cable positioning and connector orientation/torque.

The test set approach makes it much easier to measure all four device parameters under GPIB control since it is necessary to program only the scalar analyzer.

#### Calibration Accuracy Considerations

Conventional calibration techniques are used to calibrate the system. Like a Noise Figure meter, the scalar analyzer can calculate the Noise Figure of the test set. This second stage component is then backed out of the measured value of the DUT plus test set to give the Noise Figure for the DUT alone.

The many factors that can contribute to the measurement uncertainty of a Noise Figure measurement have been well documented in various application notes<sup>1</sup>. The same variables apply when using the Giga-tronics amplifier test system. At microwave frequencies, mismatch uncertainties are generally one of the more significant contributors. The advantage of the scalar measurement system is that it can be used to measure the actual VSWR values of the Noise Source and Test Set. This saves using worst case values which may only apply over a small frequency range.

A simple spreadsheet can then be constructed to quickly and easily calculate the Noise Figure measurement uncertainty using the measured parameters of the test set as well as those of the DUT.

### Summary

It has been shown that the Giga-tronics precision scalar analyzer can be used to make accurate measurements of Noise Figure with the addition of a down converter test set. The advantages over a dedicated Noise Figure meter are lower cost, faster measurements, better repeatability and easier programming. Manufacturers can now measure the 4 key amplifier parameters at one test station instead of the traditional two or even three.

### References

1. Noise Figure Measurements, Principles and Measurements, Hewlett Packard, January 1989.

Tables and figures follow for a paper on Noise Measurements on Microwave Devices Using a Scalar Analyzer.

<u>Analysis of Costs to Manufacture a Connected Hybrid Amplifier</u>	
Materials	30%
Assembly	10%
Alignment and Test	60%

(Assumes an amplifier selling price in the  
\$1,000 price range)

Table 1

Comparison of Measured Results Between  
Scalar-Based Noise Figure System  
and a Noise Figure

Frequency (GHz)	Scalar System	Noise Figure Meter
4	2.88	2.83
5	2.71	2.69
6	2.73	2.75
7	2.58	2.64
8	2.72	2.78

Device measured was small signal wideband amplifier.

Table 2

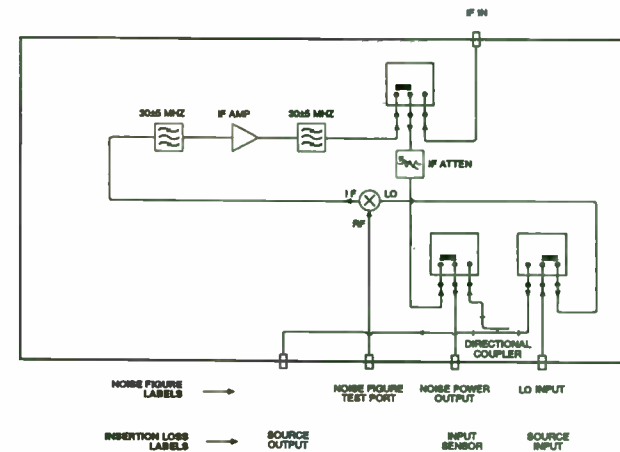


Figure 1

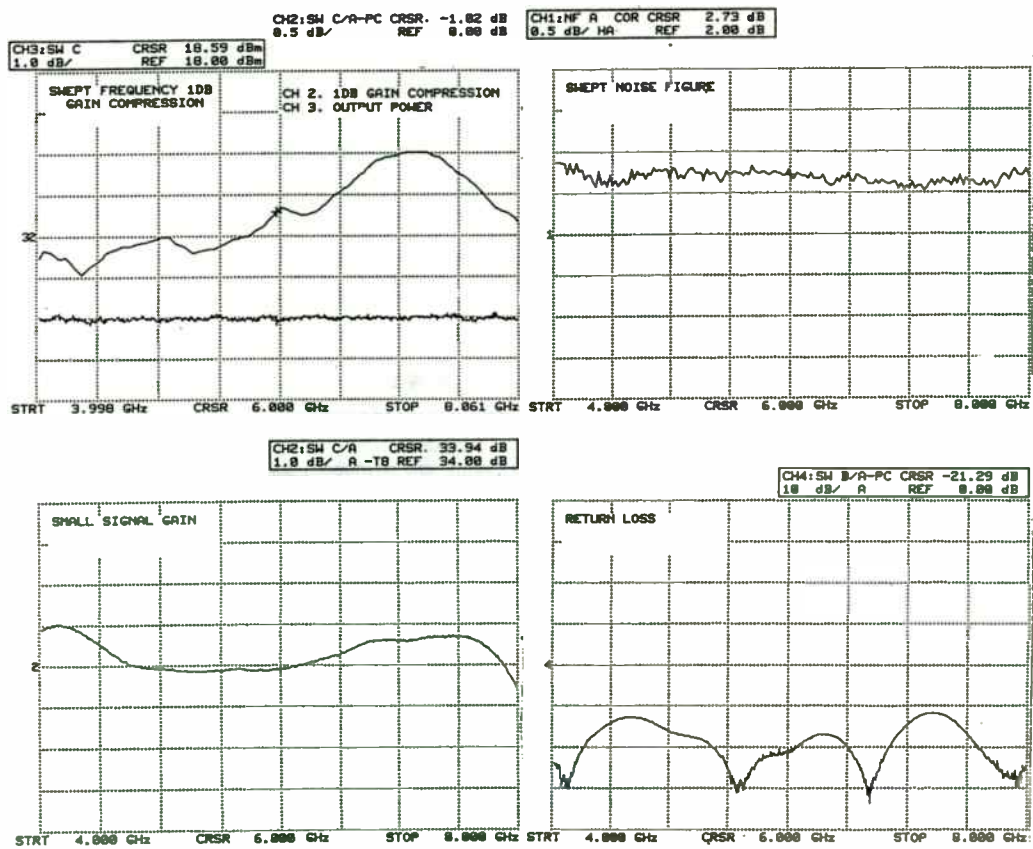


Figure 2



## Predicting Noise Contributions in RF Circuits

Dan Pleasant  
Application Engineer, High Frequency CAE Tools  
Hewlett-Packard Company  
1400 Fountaingrove Parkway  
Santa Rosa, CA 95403

### **Abstract**

Recent improvements to start-of-the-art RF and microwave CAE tools include the capability to calculate noise effects in harmonic balance analyses. This paper discusses the general mathematical techniques used to calculate noise effects in nonlinear circuits. The discussion includes practical examples of noise calculations for mixers as well as phase noise in oscillators.

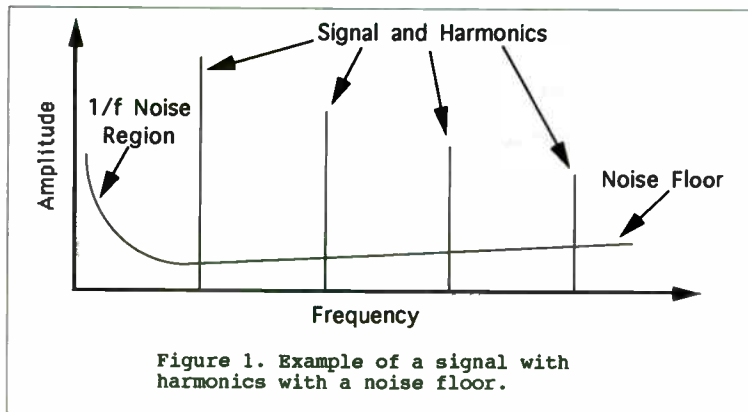
### **Nonlinear Noise Contributions**

For several years, RF designers using high-frequency CAE tools have had the ability to predict small-signal noise effects in linear circuits. While linear noise analysis is an indispensable tool for designing low-noise amplifiers and other small-signal circuits, linear noise analysis cannot handle a whole class of important design problems. Mixers, oscillators, frequency multipliers and other nonlinear circuits cannot be

analyzed using linear noise analysis. For these circuits, a new type of noise analysis that properly accounts for frequency mixing effects is needed. In the HP RF Design System, Hewlett-Packard recently introduced such a small-signal noise analysis capability for nonlinear circuits. This paper discusses the use of nonlinear noise analysis and presents some results of the HP circuit simulator.

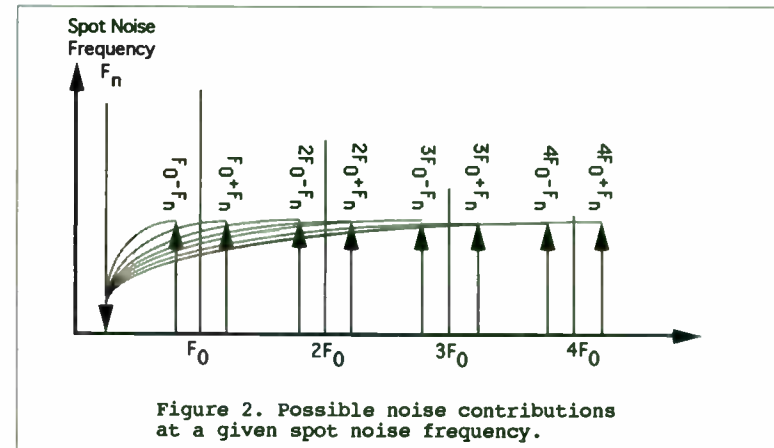
One assumption that is usually made with linear noise analysis is that there is only a single frequency of interest. That is, a circuit is stimulated with a signal at a given frequency; its output signal is not shifted in frequency; and all noise analysis can be performed (within a given bandwidth) at that same frequency.

With nonlinear circuits, the circuit simulator cannot make this simplifying assumption. Consider a broadband signal with several harmonics as shown in Figure 1. This example signal might be typical of the output of an amplifier which is being driven into saturation. A noise floor, much as it might be viewed on a spectrum analyzer display with a narrow resolution bandwidth, is shown throughout this wide signal bandwidth.



To simulate this type of noise output, a simulator must be able to calculate the noise voltage (in volts per square root of frequency) at any given frequency. This frequency is called the **spot noise frequency**. The simulator must automatically find all possible noise contributions at the spot noise frequency. Since frequency translation effects are important, this includes noise which is translated to the spot noise frequency from any other frequency.

Figure 2 illustrates the possible noise contributions at a given spot noise frequency. For each signal harmonic, two sidebands of noise can be mixed down to the spot noise frequency. The simulator must calculate the noise which is produced directly at the spot noise frequency and add to it the contributions of all possible mixing products.



### Noise Correlations

A typical circuit may contain many different sources of noise. Many of these noise sources operate independently of one another, producing **uncorrelated** noise.

However, some devices (especially BJTs and FETs) contain several separate, but dependent, sources of noise. The noise contributions of these sources is **correlated**.

Uncorrelated noise is easily calculated. If each source of noise at a given spot noise frequency produces a noise voltage  $V_N$ ,  $N=1,2,3,\dots$ , then the total noise voltage at that frequency is given by the simple root-sum-square relationship:

$$N_{\text{total}} = \sqrt{\sum_N V_N^2} \quad [1]$$

If the noise is correlated, this simple relationship does not hold. A typical circuit contains some sources of noise which are uncorrelated and others which are correlated. Uncorrelated sources are added as before; correlated sources must have their correlation coefficients taken into account.

For simplicity, assume for the moment that a circuit contains several uncorrelated noise sources  $V_N$  and two correlated sources  $V_{C1}$  and  $V_{C2}$  with a (complex) correlation coefficient  $C_{12}$ . Their total noise contribution would then be

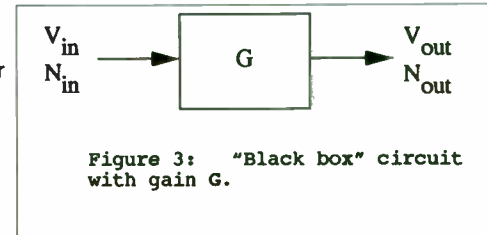
$$N_{\text{total}} = \sqrt{V_{C1}^2 + V_{C2}^2 + 2V_{C1}V_{C2}\text{Re}(C) + \sum_N V_N^2} \quad [2]$$

Other correlated noise sources are added similarly.

### Noise Figure Of Mixer Circuits

Consider a "black box" circuit with a single input and a single output, as shown in

Figure 3. The circuit is excited with a signal  $V_{in}$  and noise power (within a given measurement bandwidth)  $N_{in}$ . It has a gain  $G$ , and the output waveform



contains the amplified input signal  $V_{out}$  and a noise power  $N_{out}$ . The output noise voltage  $N_{out}$  consists of the amplified input noise  $GN_{in}$  and some extra noise which was added by the circuit itself. For a linear circuit with input and output at identical frequencies the noise figure is defined as

$$NF = 10 \log \frac{N_{out}}{GN_{in}} \quad [3]$$

With linear circuits the input and output signals are at the same frequency, and equation [3] is not ambiguous.

Now consider the more complicated case of a mixer (or any other frequency-

translation circuit) as shown in Figure 4. For nonlinear circuits with more than two ports like this, equation [3] is incomplete. The noise figure calculation for a mixer may be perfectly obvious to a human, but a computer needs much more information. The input and output signals of

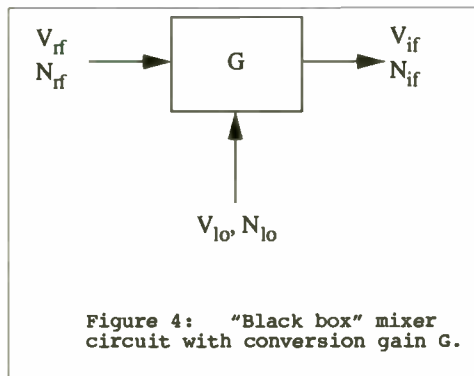


Figure 4: "Black box" mixer circuit with conversion gain G.

interest may be at entirely different frequencies, and there is an added local oscillator signal. There may be many spurious signals present at the output of a mixer, at the LO input, and even at the RF input, so the specific frequencies of interest must be explicitly specified and the simulator must correctly account for noise contributions from all sources. Also, mixers may be used as either single-sideband or double-sideband modulators, and the intended type of usage must also be specified. All of this information must be specified by the user. Figure 5 shows a simple mixer circuit as it would appear in the HP RF Design System and the results of a noise analysis of that circuit.

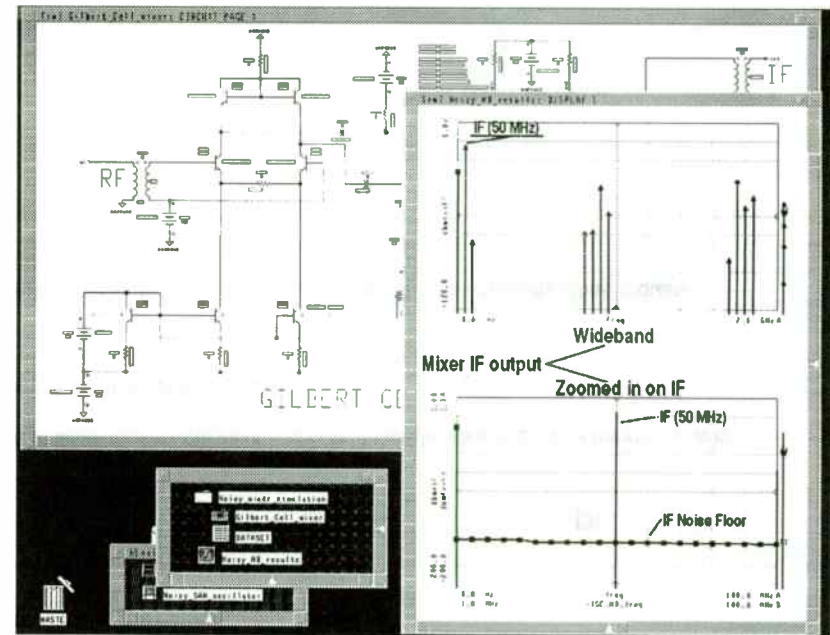
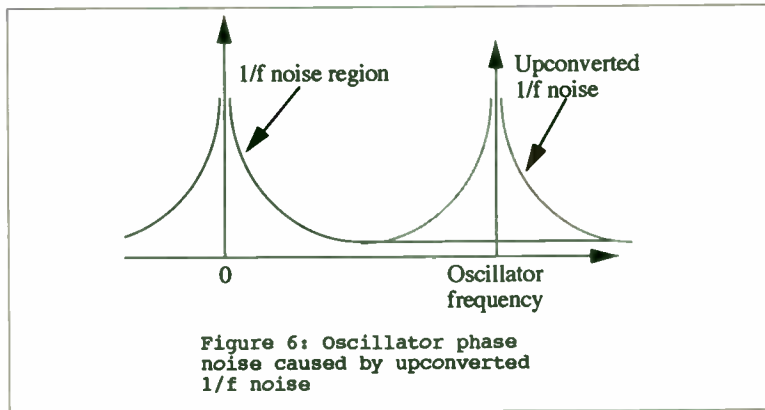


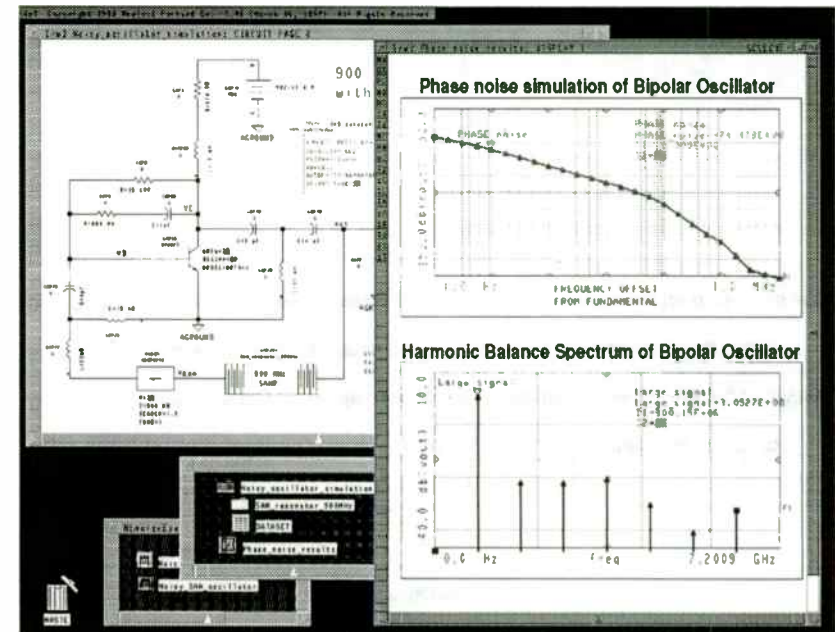
Figure 5: Gilbert cell mixer with noise analysis from the HP RF Design System.

### Oscillator Phase Noise

Phase noise in oscillators is difficult to measure and is of particular interest to users of high-frequency CAE software. Of the several possible causes of oscillator phase noise, the most prevalent is upconversion of  $1/f$  noise to the oscillator frequency as depicted in Figure 6.



This type of oscillator phase noise is analyzed in exactly the same way as mixer noise, with the caveat that the oscillator itself must be analyzed first. The simulator simply calculates the noise contribution from the  $1/f$  noise region and translated it to the oscillator's frequency of operation. Then the oscillator's phase noise can be plotted as shown in Figure 7.



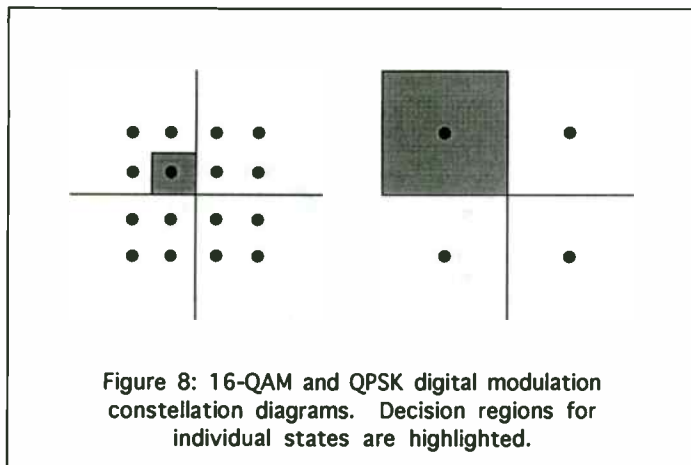
Besides upconverted  $1/f$  noise, oscillator phase noise comes from a variety of sources. For instance, nonlinear FET parameters may fluctuate due to temperature variations or by signal modulation. DC bias supplies may also contain  $1/f$  noise which is not considered by the simulator. These sources of phase noise are not calculated by HP's simulator. Finally, transistor bias points may fluctuate because of circuit

noise. These small changes in the bias may be translated into phase noise. This source of phase noise, however, is correctly accounted for by the HP simulator.

### **Bit Error Rate Calculations**

When both the signal and noise in a circuit are known, it is a fairly simple matter to calculate the expected bit error rate of a digital signal communications channel.

However, the required calculations depend on the specific modulation scheme being used. For instance, consider the 16-QAM and QPSK modulation schemes illustrated in Figure 8. Each state in the constellation diagram has a particular decision region associated with it. Also, each state has a particular noise level associated with it.



With digital modulation schemes, errors are caused by both noise and intersymbol interference. These error sources cause the sample points in the constellation diagram to "smear" toward the edge of the decision threshold. Intersymbol interference is calculated with nonlinear circuit analysis techniques in the frequency domain (harmonic balance). With nonlinear noise analysis, the statistical distribution of the noise waveform (which is assumed to be Gaussian) can be added to the intersymbol interference. From this, the probability that a given sample point will cross the decision threshold can be calculated. This calculation depends on the shape and extent of the decision region for each bit, and therefore, bit error rate calculations are different for every different modulation scheme.

### **Conclusion**

State-of-the-art CAE software now enables designers to calculate noise effects in nonlinear circuits, opening up whole new classes of circuit analyses. Adding relatively simple calculations to well-established harmonic balance simulations, simulations of nonlinear noise analysis allows analysis of mixer noise, phase noise for oscillators and bit error rates for digital modulation schemes.

# WHATEVER HAPPENED TO THE Q-METER?

Albert Helfrick  
Associate Professor  
Embry-Riddle Aeronautical University

## Introduction

Recently a colleague, visiting my office at the University, spotted my Q-meter there and inquired as to where I got the "museum piece". I informed him that the equipment in question was in regular use and that I had spare vacuum tubes to insure it's continued usefulness. I went on defending the insulted instrument by explaining that I was measuring the very same parameters of inductors made the very same way they were 50 years ago when this particular Q-meter was manufactured. "But the modern method would be a network analyzer", my guest replied. "And how much does this network analyzer cost?", I asked; end of discussion.

Many RF engineers have used the Q-meter for as long as we can remember and, if a Q-meter is still available to us, we continue to use it to verify coil designs before the inductors are placed in the circuit. Many manufactured inductors are still specified "as measured on a Boonton Radio model XXX Q-meter", in spite of the fact that there hasn't been a Q-meter made in 25 years. I believe the time is now for the Q-meter to return to the laboratory.

This paper describes Q, Q measurement, a short history of the Q-meter, and a modern Q-meter design.

## What Is Q?

If one were to ask a college student to define the parameter Q, most likely you would get an answer of  $X_p/R$  or  $R/X_p$ . If one were asking about a specific parallel or series resonant circuit, the answer would have a 50 percent chance of being right. Seldom is the answer what this author prefers, i.e. Q is the  $2\pi$  times the ratio of the total energy stored in a resonant system to the energy dissipated each cycle of oscillation.

Notice that this definition does not refer to an electrical circuit. It is important to understand that resonance is more than just electrical. There are nuclear resonances, lattice resonances and mechanical resonances of all sorts. Certainly the RF engineer is familiar with the quartz crystal, which is, of course, a mechanical oscillation.

In a resonant system, the energy of the system is transferred from one energy storage element to another. All the while, energy is being lost by the dissipative elements of the system.

In an electrical circuit, the energy storage elements are the inductor and capacitor. The dissipative elements, represented on a circuit diagram by a resistance, can be rather complex. First, energy may be lost in the resistance of the wire of the inductor. It is a well-known fact, that at higher frequencies, this resistance, because of the skin effect, can be significantly higher than the DC resistance.

The inductor has other ways of losing energy. If the inductor is wound around a core such as ferrite or powdered iron, energy may be lost in that core material. Additionally, the inductor will allow some of the energy to dissipate because of radiation. The magnetic field around the inductor may not be totally contained, and, consequently, the inductor becomes a miniature loop antenna and some of the energy is lost through the radiation of electromagnetic energy.

Another source of energy loss is in the capacitor. The majority of energy lost in the capacitor is lost in the dielectric from heating. However, like the inductor, the capacitor, too, can radiate. The electric field may not be totally contained and the capacitor can become a miniature electric field antenna.

Finally, there are the parasitic circuit elements; the wires, the printed circuit board dielectric, terminals, etc. that have their own radiation characteristics and losses.

Regardless of the source of the energy loss, the resonant circuit can be shown with one resistance which represents the equivalent of all the energy losses. Thus the resonant circuit of fig. 1. shows the inductor and capacitor as perfect elements

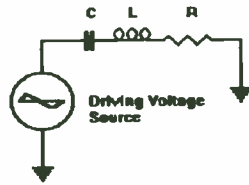


Figure 1. A resonant circuit showing a driving source and the equivalent resistance due to a finite Q.

To derive the equation for Q using the definition stated, an expression for the energy stored in the circuit must be written. For the series resonant circuit shown, it is easier to consider the energy stored in the inductor. Both the current through the circuit as well as the voltage across any circuit element is sinusoid. Therefore, at the peak of the current waveform, all of the energy of the circuit is contained in the inductor. Using this fact, the following relationship is written

$$\text{Energy} = \frac{1}{2} L I^2 \quad (1-1)$$

where L is the circuit inductance and I is the peak instantaneous current.

Consequently the circuit current as a function of time is

$$I(t) = I_m \sin \omega t \quad (1-2)$$

where I(t) is the time dependent current and  $\omega$  is the angular resonant frequency.

The energy dissipated in the resistor per cycle is

$$\int_0^{\frac{2\pi}{\omega}} R I^2(t) dt = R I^2 (\pi \sqrt{LC}) \quad (1-3)$$

Taking the ratio between the energy stored to the energy dissipated per cycle and multiplying by 2 the following result is obtained.

$$2\pi \frac{\frac{1}{2} L I^2}{R I^2 (\pi \sqrt{LC})} = \frac{\omega L}{R} = \frac{X_L}{R} \quad (1-4)$$

This is the familiar equation for Q in an electrical circuit.

The parallel resonant circuit may be analyzed in a similar fashion by considering the energy stored in the capacitor and the circuit voltage.

## Measuring Q

The first Q meter was made available to the Radio Industry in 1934 by the Boonton Radio Corporation of Boonton, N.J. This company grew out of Radio Frequency Laboratories, RFL, which spawned Aircraft Radio Corporation and figured indirectly with many other companies in the northern New Jersey area. Interestingly, of the famous Boonton, N.J. companies; Measurements, Inc., Ballantine Laboratories, Ferris Instruments, RFL, and Aircraft Radio Corporation, only Ferris was in Boonton. The companies were in adjacent towns that used the Boonton post office. One electronics company that bears the Boonton name never had a Boonton address.

The Boonton Radio Corporation produced Q-meters for nearly 30 years. Boonton increased their product line to include RF bridges, signal generators and other test equipment. They moved from their original buildings to a new building in a beautiful bucolic setting in Rockaway N.J., and shortly thereafter were acquired by Hewlett-Packard. Apparently H-P wanted the building as the Boonton products were never improved or expanded. The Q-meter along with the Boonton R-X meter appeared in the H-P catalog to about 1970 looking the way they did in the 1940s still using vacuum tubes that were designed in the 1930s. About 1970, the Boonton products disappeared and a new Q-meter bearing only the H-P name appeared.

Boonton Q meters had been measuring Q and inductance to 250 MHz since the late 1930s but the new Q-meter measured to only 70 MHz. In the mid 1980s the H-P Q-meter was dropped from the catalog index and was given a small space among the RLC measuring systems, where it remains today.

Boonton Radio was not the only source of Q-meters. There was at least two other sources. The Marconi Corporation in England made a Q-meter well into the 1960s. This instrument was unique in that it was called a "circuit magnification meter".

In the 1950s, the Heath Company made a Q-meter kit that was patterned after the Boonton Q-meter. It had the same characteristic sloping front panel with the unknown terminals on the top. Heath at that time made other inexpensive look-alike versions of expensive laboratory equipment.

The principal of Q measurement is very simple. An unknown inductor is placed in a series resonant circuit and driven from an AC voltage source as shown in fig. 2. Either the signal generator frequency or the series capacitance is adjusted to give resonance. The voltage across the capacitor is measured using a high impedance voltmeter and the ratio of the capacitor voltage to the driving voltage is the circuit



Q. It is the "magnification" of the driving voltage that gave the Marconi Q-meter its name.

To prove this voltage ratio is the inductor's Q consider the circuit of figure 2.

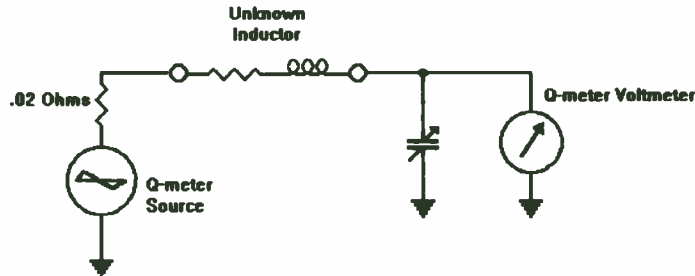


Figure 2. The basic Q-meter configuration.

The first requirement for an accurate Q measurement is that the voltage source resistance be small relative to the equivalent series resistance of the unknown. Additionally, the resistance of the voltmeter must be very large relative to this equivalent series resistance.

At resonance, a series resonant circuit has an impedance which is the circuit's resistive component. Therefore, the circuit current may be calculated by the relationship

$$I = \frac{V}{R} \quad (1-5)$$

where V is the source voltage.

Assuming the capacitor is perfect, that is, has no parallel resistances, the voltage across the capacitor is

$$IX_c = \frac{X_c V}{R} \quad (1-6)$$

where Xc is the capacitive reactance.

Because the inductive and capacitive reactances are the same, at resonance, the equation above may be rewritten for the desired result.

$$V_c = \frac{X_c V}{R} = VQ \quad (1-7)$$

This relationship holds only for a lossless capacitor and is nearly correct when the loss of the capacitor is much smaller than the inductor under test. The circuit shown in fig. 2 used for the derivation of equation 1-7 is the technique used in the Q-meter.

The Boonton Q-meter's voltage source had an internal resistance of 0.02 ohms. This internal resistance was achieved with a 0.02 ohm resistor which was a short length of Ni-chrome wire. The coupling coil from the oscillator fed the 0.02 ohm resistor. The current through the resistance was measured using a thermocouple assembly. The thermocouple DC output drove the "multiply Q by" meter on the front panel of the instrument. The operator was required to set the oscillator power, typically to "multiply Q by 1", but certainly below the red band on the meter. Failure to observe the red band would result in a destroyed thermocouple assembly and a costly repair.

Measuring the current through the resistance to deduce voltage assumes the Ni-chrome resistor is a pure resistance and has no significant inductance or transmission line effects. For the 70 MHz range of the Q-meters employing this technique, this was a safe assumption.

It is necessary, for an accurate Q measurement, that the voltmeter used to measure the RF voltage across the variable capacitor have a very high input impedance so that it will not affect the Q of the circuit. The Boonton Q-meters used an AM detector circuit called the "infinite impedance" detector. This is simply a vacuum tube that was biased near conduction. The cathode circuit had an RC circuit with a time constant considerably longer than the period of the RF that was being measured. Therefore, each RF cycle that exceeded the conduction threshold of the vacuum tube would increase the charge on the capacitor in the cathode circuit and push the vacuum tube bias back to near cutoff. Because the vacuum tube operated continually near cutoff the input impedance was represented by a very small amount of leakage current and the input grid's capacitance.

## A Modern Q-Meter

It was desired to recreate the Q-meter using the same basic measurement topology but modern techniques. Some differences would be incorporated to decrease the cost of the instrument. Some of the goals for the Q-meter were:

1. The Q-meter will not contain a signal source. This is an expensive item and any laboratory where serious RF design work is taking place and a Q meter would be needed, has at least one signal generator.
2. The Q-meter will be able to operate with very low drive levels so that very small geometry devices may be measured without significant heating.
3. The Q-meter should be capable of operation from 100 kHz to 100 MHz.

An interesting difference between the modern Q-meter and the classic meter is that the Q measurement has been designed to be a logarithmic representation. This makes it easier to read a wide variation of Qs but, more importantly, it allows a measuring circuit with a very wide dynamic range. The voltage across the resonating capacitor is measured using a logarithmic amplifier preceded with a high impedance buffer amplifier, as shown in figure 3.

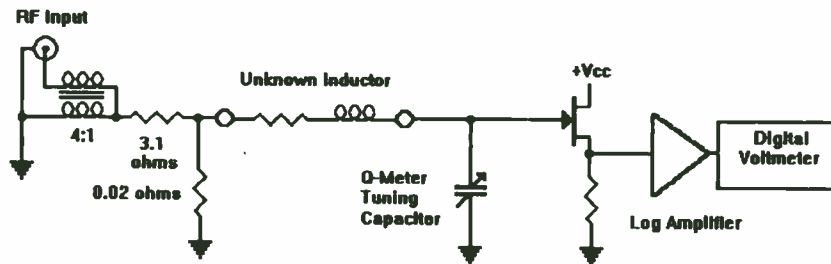


Figure 3. Block diagram of the modern Q-meter.

The biggest hurdle in designing the modern Q-meter was finding an acceptable alternative to the variable capacitor. Mechanical variable capacitors are just not readily available and high-Q, precision variables will be expensive. Varactor diodes are not viable as their capacitance ratio is not sufficient for a Q meter. Also, for the higher capacitance diodes, the diode Qs are not high enough to allow accurate

measurements of high Q inductors. A digitally-switched variable capacitor was designed for the Q-meter. Five binary weighted capacitors are used plus a varactor diode for the remaining capacitor resolution.

The resonating capacitor consists of the relay-switched capacitors, the varactor diode and the all of the parasitic capacitances in parallel. The relays chosen are specially designed for high frequency signal switching and have very low parasitic capacitance. The voltmeter buffer amplifier adds a few pF to the parasitic capacitance. The varactor diode, as it cannot produce zero capacitance, adds another few pF. The design was chosen that the varactor diode would add 10 pF to the total capacitance. Since a comfortable tuning range for a varactor diode is about 4 to one, a 3.33 pF to 13.33 pF diode will be suitable for the Q meter. However, this adds 3.33 pF of residual capacitance to the Q meter.

It is important that the voltage impressed on the varactor diode does not cause conduction in the diode. The usual technique for preventing conduction in varactor diodes is to use two diodes back-to-back. This technique, however, reduces the Q of the varactor diode as it requires the use of two diodes of twice the capacitance and the larger geometry diodes have greater losses.

The technique used in the modern Q-meter is to prevent the varactor diode from conducting. The peak diode voltage should be less than about 500 mV peak or 354 mV RMS to prevent conduction at zero diode voltage. Assuming that the maximum Q ever measured by the modern Q-meter is 1000 this requires a maximum driving source of 354  $\mu$ V.

It was decided that a source generator level could be reduced to as low as 10 $\mu$ V for the following reasons. First, it would insure that the varactor diode would never conduct or operate in the non-linear region. Second, there are some very small geometry parts such as small inductors, and inherently high Q devices such as piezoelectric resonators that could be handled by the modern Q-meter if the driving power was kept to a minimum.

The classic Boonton Q-meter had a voltage source of 20 mV with an internal resistance of 20 milliohms. This voltage source was created by conducting a current of 1 ampere through a 20 milliohm resistor. The accuracy of the open circuit voltage was insured by measuring the current through the 20 milliohm resistor using a thermocouple RF ammeter. The energy dissipated in the resistor is only 20 mW but the energy coupling from the oscillator to the source resistor was very inefficient. At RF, impedances of significance, relative to 20 milliohms, come from everywhere, inductance of wires, copper losses, mutual inductance, etc. The source oscillators for the Boonton Q-meters were capable of providing an output of considerable power and the vast majority of the energy was wasted in the coupling.

A similar problem exists, even for a modern Q-meter. The design goal of the modern Q-meter was to interface with a laboratory signal generator with a 50 ohm output impedance. The impedance transformation required to provide a 20 milliohm source is 2500 to one.

It was desired that the modern Q-meter be accurate to 100 MHz and useful to 200 MHz. On the other end of the spectrum, the Q-meter should be useful to areas where inductors are still being used, or around 100 kHz. Providing a large impedance transformation over this broad frequency range with a single transformer is nearly impossible. Therefore it was decided that two impedance transformation networks be provided each with its own LO terminal. The frequency range of 100 kHz to 1 MHz is transformed by one network while a second network provides the 1 MHz to 100 MHz range.

For the high frequency region, a 4 to 1 transformer giving an impedance reduction of 16 to 1 is used followed by a voltage divider. The impedance looking back into the transformer is 3.125 ohms which requires a 156 to one voltage divider.

When transforming from 50 ohms to 3 ohms, great care must be taken to insure tight coupling and low leakage inductance. Transformers of these characteristics are regularly used for high frequency power amplifiers.

If the Q-meter is to read Q directly, it is necessary that the driving source as well as the measuring circuits have a flat frequency response. This is particularly difficult for a broad frequency range from 100 kHz to 100 MHz. Every effort was made to provide as flat a frequency response as possible but a different operating technique is used for the modern Q-meter.

For many measurements, precise Q is not important and a Q indication of 20% is sufficient. The frequency response is sufficiently flat to allow a direct calibration of Q to about 20%.

A more precise measurement of Q requires a calibration routine to be performed. The unknown inductor is tuned for resonance and the Q is noted. The inductor is replaced with a shorting bar, the capacitance reduced to the minimum and the signal generator level is increased until the Q indicator reads the same value as the inductor at resonance. The ratio of the two signal generator voltages is the inductor Q.

As an example, an inductor reads a Q of 75 at resonance and the signal generator was set to -30 dBm. The shorting bar is applied and the signal generator must be increased to +6 dBm or an increase of 36 db, to achieve the same 75 reading. This indicates a Q of the inductor of 63.

With the use of a laboratory signal generator with the capability of precise

frequency and level control, the modern Q-meter was used to make some Q measurements that would be difficult on the old Boonton Q-meter. The modern Q-meter was modified to measure the Q of very high Q devices such as ceramic resonators and quartz crystals. These devices, however, are not inductors but resonant circuits and cannot be measured by resonating in a series resonant circuit. The only hardware modification of the Q-meter is the addition of a ground terminal near the unknown terminals of the Q-meter. This will allow providing specific termination resistances to the voltmeter circuit. The design of the voltmeter for making Q measurements was for a very high input impedance but for making measurements of resonators a much lower impedance such as 50 ohms is desired.

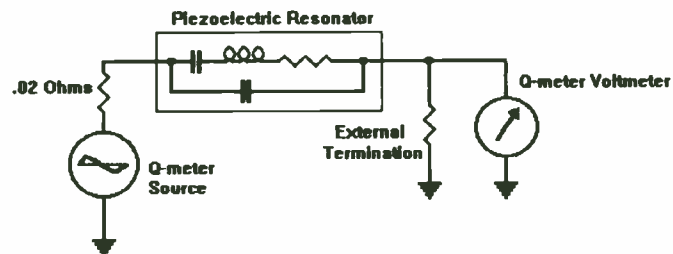


Figure 4. Schematic of the technique for measuring piezoelectric resonators using a Q-meter.

To measure the resonant frequency of a piezoelectric resonator, such as a ceramic resonator or quartz crystal, a 50 ohm resistor is connected from the voltmeter terminal to ground as shown in the schematic of fig. 4 and the resonating capacitor is set to minimum. The series resonant frequency is determined by adjusting the signal generator frequency to the resonant frequency. To determine the equivalent series resistance, the resonator is removed and replaced with the shorting bar and the signal generator level is reduced to produce the same indication from the voltmeter. From this level reduction, the series resistance may be calculated. As an example, if the signal generator was reduced 6 dB to produce the same voltmeter reading as a crystal under test at resonance when terminated in a 50 ohm resistor, the series resistance of the crystal is 50 ohms. By placing a capacitor equal to the desired load capacitance, in series with the resonator, the modern Q-meter may be used to measure the parallel resonant frequency of a quartz crystal.

Using this technique, the modern Q-meter was used to measure the parallel and series resonant frequencies and the resistance of a quartz crystal with 1 PPM accuracy on frequency and 10% on resistance. Thus the modern Q-meter will serve as a crystal impedance, C-I, meter.

Devices that require a defined input and output impedance may be tested by inserting a resistor between the input of the device and the source terminal of Q-meter. If the external signal generator is capable of swept frequency operation, the frequency response of the device may be displayed.

What makes this modern Q-meter capable of performing these high Q measurements that the older Boonton Q meter could not is the use of an external signal generator with very fine frequency resolution and a calibrated and adjustable output. Also, the use of an amplified voltmeter which allows very low driving power to be used prevents damage to these very fragile parts.

There are classic procedures for measuring the distributed capacitance of inductors and for measuring capacitors at RF with the Q-meter. These techniques may be learned from the references at the end of this paper.

Capacitance may be measured using the Q-meter by supplying a good quality inductor and resonating the inductor using the internal capacitor at the desired measurement frequency. The unknown capacitor is applied to the HI terminal and ground and the Q-meter is resonated again. The difference in the Q meter capacitor is the capacitance of the unknown capacitor. The dissipation factor or Q of the unknown capacitor may be calculated from the differences in the before and after Q measurements. These formulae are available in the references.

## Conclusion

The modern Q-meter was conceived to make simple Q measurements of inductors rather than the more expensive network analyzer. However, by providing various source resistances and terminations, the Q meter was capable of some network measurements.

The Q-meter can still be a valuable tool for the RF laboratory when used with a signal generator capable of accurate frequency and amplitude settings. With the use of a sensitive voltmeter the range of Q-meter measurements can be greatly expanded.

## BIBLIOGRAPHY

- V. V. K. Rao, "The Q-Meter and its Theory", Proc. IRE, Nov. 1942  
R. E. Lafferty, "Extended Q-Meter Measurements", Electronics, Nov. 1951  
I. Bady, "Measurement of Q of High Capacity Condensers by the Series-Parallel Method", Signal Corps Report, March 1943.  
A. D. Helfrick and William Cooper, "Modern Electronic Instrumentation and Measurement Techniques", third edition, Prentice-Hall, 1990

## SELECTING THE RIGHT FREQUENCY DOMAIN MEASURING INSTRUMENT

Robert Witte  
Hewlett-Packard Company  
P.O. Box 2197  
Colorado Springs, CO 80901

### Abstract

Frequency domain measurements have been used by engineers for decades. With the increased use of the Fast Fourier Transform (FFT) in spectrum analyzers and digitizing oscilloscopes, choosing the right frequency domain measuring instrument has become more complicated. Frequency domain instruments capable of measuring a signal's frequency content include Swept Spectrum Analyzers, FFT Spectrum Analyzers and Digitizing Oscilloscopes with FFT. Choosing between these instruments involves compromises in terms of cost, frequency range, transient capture capability and dynamic range.

### The Frequency Domain

Signals are most commonly represented in the time domain, usually voltage as a function of time. The Fourier Transform or the Fourier Series can be used to transform a time domain signal into the frequency domain. Both of these mathematical operations involves breaking a signal into its individual frequency components. A graphical way of looking at this is shown in Figure 1. Some typical signals are shown in the time domain and frequency domain in Figure 2.

### Bank-of-filters Analyzer

Although not used often in electronic measurements, the bank-of-filters analyzer is an idealized instrument which will aid in understanding more common spectrum analyzer implementations. This analyzer consists of multiple bandpass filters, each tuned to a different frequency (Figure 3). These filters are very steep-walled and overlap in frequency only slightly as shown in Figure 4. Ideally, infinitely steep "brick wall" filters are used to provide no overlap between adjacent filters. Any frequency component present at the input of the instrument passes through only one of the filters. A bank of detectors following the filters converts the frequency components into a DC value to drive the meter associated with that filter/detector. Of course, the frequency content can be displayed with a variety of display technologies including cathode-ray tubes and liquid crystal displays.

Frequency components which fall within the same bandpass filter cannot be resolved. Stated differently, the instrument cannot measure individual frequency components which reside in the same filter. Thus, the frequency resolution of the bank-of-filters analyzer is equal to the width of the bandpass filter, shown as  $BW_{RES}$  (RESOLUTION BANDWIDTH). The resolution bandwidth can also be expressed as

$$BW_{RES} = f_{max} / N$$

where  $f_{max}$  = maximum frequency of the analyzer

N = the number of filters

This equation points out a serious practical problem with the bank-of-filters analyzer, namely that the number of filters required is often quite large. For example, consider a 10 MHz analyzer with frequency resolution equal to 1kHz. The number of filters required is  $N = (10 \text{ MHz}/1\text{kHz}) = 10000$ .

#### Wave Analyzer

The wave analyzer uses a single tunable filter to measure the frequency content of a signal (Figure 5). The wave analyzer works very similar to the bank-of-filters analyzer except that only one frequency can be measured at a time. This solves the problem of needing a large number of filters, but has the disadvantage of greater total measurement time since the only one frequency can be measured at a time.

The main practical problem in implementing a wave meter is building a narrowband filter which is tunable over a wide range of frequencies. Take the previous example of a 10 MHz analyzer with a resolution bandwidth of 1 kHz. The tunable filter must be capable of tuning from near DC to 10 MHz while maintaining a constant filter shape with a passband of 1 kHz. (When the filter is tuned to 10 MHz, the filter has a Q of 10000, a high value for a tunable filter.)

The standard way of implementing the wave analyzer is to keep the filter fixed in frequency and move the input signal in frequency using heterodyne receiver techniques (Figure 6). The input signal is attenuated or amplified as needed, low pass filtered and delivered to the input of a mixer. A local oscillator signal mixes with the input signal to produce the intermediate frequency (IF). The IF filter removes all frequency components except the frequency the wave analyzer is tuned to.

The block diagram shown in Figure 6 is highly simplified. A practical implementation typically has several stages of conversion using several mixers, local oscillators and intermediate frequency filters.

#### Swept Spectrum Analyzer

The swept heterodyne spectrum analyzer improves on the wave analyzer block diagram by automatically sweeping the local oscillator over the required frequency range and plotting the response on a display. Figure 7 shows an analog implementation of such a block diagram with a ramp generator creating a sweep voltage that drives a voltage-controller oscillator (VCO) used as the LO. This block diagram is also highly simplified and is conceptual in nature. Note the addition of the video filter which acts as a smoothing filter for the detected signal. Most modern swept spectrum analyzers digitize the signal (typically at the output of the detector) and store the frequency domain

representation in digital memory to eliminate display flicker during slow sweeps. Like the wave analyzer, a practical implementation may require multiple conversion stages. The swept spectrum analyzer can be thought of as a bandpass filter which sweeps through the frequency range of interest, plotting the detected output of the filter on the analyzer display. Like the wave analyzer, the frequency resolution of the swept analyzer is equal to the bandwidth of its IF filter. The narrowest resolution bandwidth of a typical swept analyzer is about 10 Hz, providing good but not outstanding frequency resolution. The dynamic range<sup>1</sup> of the swept analyzer varies with the instrument design but is typically in the range of 70 to 100 dB.

If the swept analyzer is swept arbitrarily fast, the IF filter will not have time to settle out and correctly track the signal. To prevent oversweeping the IF filter, the sweep rate (measured in hertz/sec) of a swept analyzer must be less than the maximum sweep rate given by the following equation.

$$\text{sweep rate (max)} = k (BW_{RES})^2$$

where k is a factor dependent on the IF filter shape

The maximum sweep rate is proportional to the square of the resolution bandwidth so for very narrow resolution bandwidths the sweep rate must be decreased dramatically. Obtaining fine

1. Dynamic range is defined as the difference (in dB) between the largest signal and smallest signal that can be reliably measured simultaneously.

frequency resolution carries a measurement time penalty. For example, sweeping a 1 kHz span with a 1 Hz resolution bandwidth requires a sweep time of 30 minutes.

The swept spectrum analyzer is clearly the workhorse of the RF and microwave design and test business due to its wide frequency coverage. Swept analyzers have the ability to zoom in on arbitrary frequency spans. This measurement capability is usually limited by the available resolution bandwidth. That is, the analyzer can be tuned to such a narrow span that the resolution bandwidth no longer provides sufficient frequency resolution.

Since the swept analyzer takes a finite time to complete its sweep, the signal must be stationary in order to ensure that all frequency components are captured. This implies that transient phenomena cannot be reliably measured with a swept analyzer. Later we will see that FFT-based instruments do not have this limitation.

#### FFT Spectrum Analyzer

Digital signal processing techniques, along with advancements in analog-to-digital converter (ADC) technology, provide the means to measure the frequency content of a signal using the Fast Fourier Transform (FFT). Conceptually, these analyzers simply digitize the signal and perform an FFT on the time domain data to generate the frequency domain data. A first cut at the FFT

analyzer block diagram is shown in Figure 8. The input signal passes through an attenuator and a lowpass filter which removes high-frequency components. The signal is sampled and converted to digital form by the ADC to produce the time domain data. The FFT is performed on the time domain data by either a microprocessor or dedicated digital logic. To prevent aliasing<sup>2</sup> and meet the requirements of the Sampling Theorem, the sample rate ( $f_s$ ) must be greater than twice the highest frequency component present after the lowpass filter ( $f_{max}$ ). For a typical instrument,  $f_s = 2.5 f_{max}$ .

The mathematics of the FFT produces a frequency domain representation that is analogous to the bank-of-filters analyzer. That is, the FFT analyzer produces an array of frequency domain values which correspond to the different frequency components present in the input signal. These frequency domain values are valid simultaneously (that is, without the sweeping filter effect of the swept spectrum analyzer).

A typical FFT algorithm in a spectrum analyzer might operate on 1024 time domain data points, producing 1024 complex frequency domain points. But the frequency domain points contain redundant information with only 513 unique points present in the frequency spectrum. For a variety of reasons, even fewer points might be retained by the spectrum analyzer for display. For the purposes

---

2. Aliasing occurs when a frequency component appears as a different, erroneous frequency due to insufficient sample rate.

of this paper, we will consider the FFT output record length to be  $N$  points for a time domain record of  $2N$  points.

The FFT is inherently a baseband transform. In other words, the frequency range of the FFT always starts at 0 Hz and extends to  $f_s/2$ . Most FFT spectrum analyzers can vary the sample rate so that the user has control over the frequency span being measured. Similar to the bank-of-filters analyzer, the frequency resolution (or resolution bandwidth) is given by<sup>3</sup>

$$BW_{RES} = f_s / (2N).$$

So a side-effect of reducing the sample rate (frequency span) is finer frequency resolution. Of course, varying the sample rate introduces the possibility of violating the sampling theorem. The low pass filter in Figure 8 is a fixed-frequency analog filter. If the sample rate is reduced, the cutoff frequency of the analog filter must also be reduced. There are many frequency spans and sample rates required in a full-featured FFT analyzer with each one requiring a separate analog lowpass filter. A more reasonable alternative is to have only one analog filter in the input path followed by a decimating digital filter after the ADC (Figure 9). This digital filter simultaneously limits the bandwidth of the signal and reduces the sample rate appropriately. The analog filter provides the primary alias protection with the digital filter providing further bandwidth and sample rate reduction without any alias effects.

---

3. This approximation ignores the effect that a time domain windowing functions will have in the frequency domain.



Another block diagram improvement for the FFT analyzer is to put a digital mixer in the input path driven by a digital sinusoid (Figure 10). This is the digital equivalent of the heterodyne receiver block diagram which allows an arbitrary frequency span to be mixed down to baseband where the FFT can handle it. This feature, called Band Selectable Analysis or Zoom Capability, frees the FFT analyzer from baseband-only measurements and lets the user zoom in the particular frequency range of interest.

FFT spectrum analyzers typically have been limited in frequency range to about 10 MHz, consistent with a sample rate of around 25 MHz. ADC resolution is usually between 12 and 16 bits. Since the digital filter and digital mixer must run at the highest sample rate, these devices also limit the instrument sample rate.

Since the FFT analyzer samples the input waveform with a sample rate that is at least twice the highest frequency present, the entire frequency content of the input is captured in the time record. Thus, a transient waveform can be acquired and its frequency spectrum displayed.

Another important advantage of the FFT analyzer is its measurement speed. The FFT analyzer does not have the same sweep rate limitations that a swept analyzer has because there is no sweeping filter effect. For an equivalent frequency span and resolution bandwidth, the FFT analyzer can complete a spectrum

measurement much faster than a swept analyzer. This is especially true for measurements requiring very small resolution bandwidths such as narrowband measurements of close-in sidebands around a carrier.

The most notable shortcoming of the FFT analyzer is its limited frequency range. One way to obtain the advantages of the FFT without giving up the frequency range is to combine the two instruments. This can be accomplished by digitizing and FFTing the IF signal of a swept analyzer (Figure 11). The analyzer's sweep is disabled so that it performs like a downconverter for the FFT function. This technique is useful mostly for narrowband measurements where the FFT's increased speed and frequency resolution has a clear advantage. Note that the ADC sample rate limits the frequency span that can be measured. Figure 12 shows a narrowband measurement made using this technique. Note that the resolution bandwidth of 180 millihertz is much narrower than a swept analyzer can produce.

#### Digitizing Oscilloscope with FFT

Many of today's digital scopes offer waveform analysis capability including FFT. Figure 13 shows an oscilloscope measurement of a square wave and the spectrum produced by the FFT. The oscilloscope can simultaneously display wide bandwidth time domain and frequency domain representations of the signal.

It makes sense that digital scope manufacturers would include an FFT function to allow the scope to be used for frequency domain measurements. The next plausible conclusion is that a digitizing scope with FFT is the same thing as an FFT spectrum analyzer, which is NOT the case. Let's look at some of the differences between the scope implementation and the spectrum analyzer implementation.

A digital oscilloscope operating in FFT mode looks much like our first cut block diagram for the FFT Spectrum Analyzer in Figure 8. Scopes may not have a distinct lowpass filter for alias protection, but the front end of the scope will roll off above the scope bandwidth. Recall that the limitation of this block diagram is that a different lowpass filter is required for each sample rate if aliasing is to be prevented. Digital oscilloscopes normally change the sample rate when the time/division setting changes, so clearly this block diagram (with only one filter) does not protect against aliasing at all time/division settings. The other limitation of this block diagram is the lack of band selectable analysis. (Some digital scopes allow a certain amount of zooming in between frequency points but none approach the flexibility of the FFT analyzer.)

So why don't scope manufacturers add the digital filter and digital oscillator just like the FFT spectrum analyzer and have the best of both worlds? Probably the main reason is that scopes are optimized for time domain measurements and adding these

digital circuits running at the sample rate of the scope adds a significant cost penalty to the product or may not even be realizable in today's technology. Note that the latest generation of digital scopes can sample at 4 GSa/sec.

Lack of complete alias protection means that the instrument user must know something about the frequency content of the signal and make sure that the sample rate is high enough to prevent aliasing. Another square wave measurement is shown in Figure 14, but this time the sample rate is not high enough to prevent aliasing. The 1 MHz square wave has high-order harmonics extending out beyond 25 MHz so with a sample rate of 50 MSa/sec, the Sampling Theorem is violated. This causes the high-order harmonics of the square wave (corresponding to frequencies which should have fallen off the right side of the display) to fold back onto the display and appear as aliases. One way to eliminate these aliases is to use a lowpass filter to bandlimit the input signal. The resulting measurement, using a 10 MHz lowpass filter, is shown in Figure 15. It's easy to imagine that for some applications using an external lowpass filter is quite reasonable (e.g. a manufacturing test system where the frequency content of a signal is fairly predictable). For measuring an unknown signal this issue becomes a more serious liability since a drawer full of lowpass filters may be required to cover all sample rate possibilities.

The digital scope with FFT has certain advantages as well.

Having a high-quality time domain measuring instrument along with some frequency domain capability in one box is desirable. Many digital scopes have significantly higher sample rate than available FFT spectrum analyzers, giving them a larger frequency range and the ability to capture high speed transients. This higher sample rate usually has the drawback of fewer ADC bits (perhaps 8 instead of 12) and smaller dynamic range (maybe 50 dB instead of 70 dB).

One important use for the FFT feature on a scope is SIGNAL IDENTIFICATION. For some complex signals, it is difficult to sort out the frequency content of the waveform in the time domain. For example, the time domain measurement shown in Figure 16 (upper trace) has more than one frequency in it, but the number of tones and the frequencies of these tones cannot be easily determined. The lower trace is the FFT of the time domain signal, which uncovers the presence of two main spectral lines, at 623 kHz and 977 kHz.

#### Summary

Table 1 summarizes the differences in the key specifications and features of the frequency domain instruments. The specs shown are typical and may be significantly different for any particular instrument in the marketplace. The main point of the table is to highlight the relative strengths and weaknesses of each instrument type.

#### References

Witte, Robert A., Spectrum and Network Measurements, Prentice-Hall Inc., 1991.

"The Fundamentals of Signal Analysis", Hewlett-Packard Company, Application Note 243, February 1985.

"Spectrum Analysis", Hewlett-Packard Company, Application Note 150, November 1989.

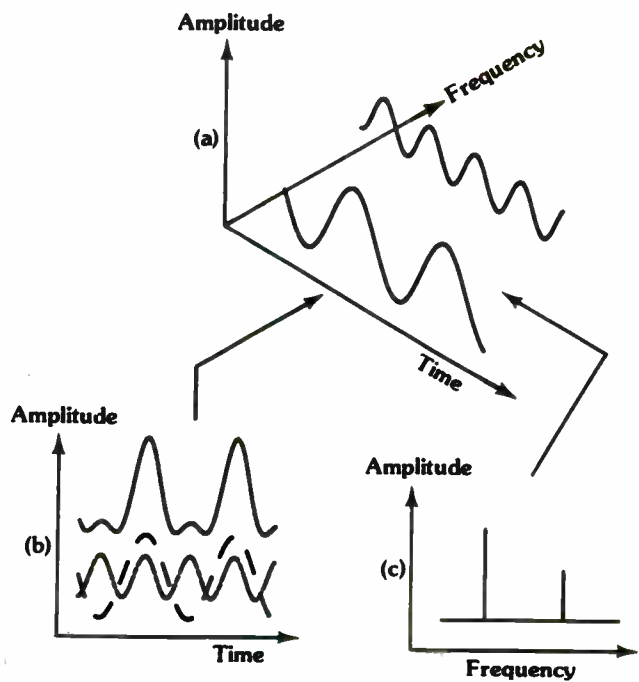


Figure 1. The relationship between the time and frequency domains.

- a) Three-dimensional coordinates showing time, frequency and amplitude
- b) Time domain view.
- c) Frequency domain view.

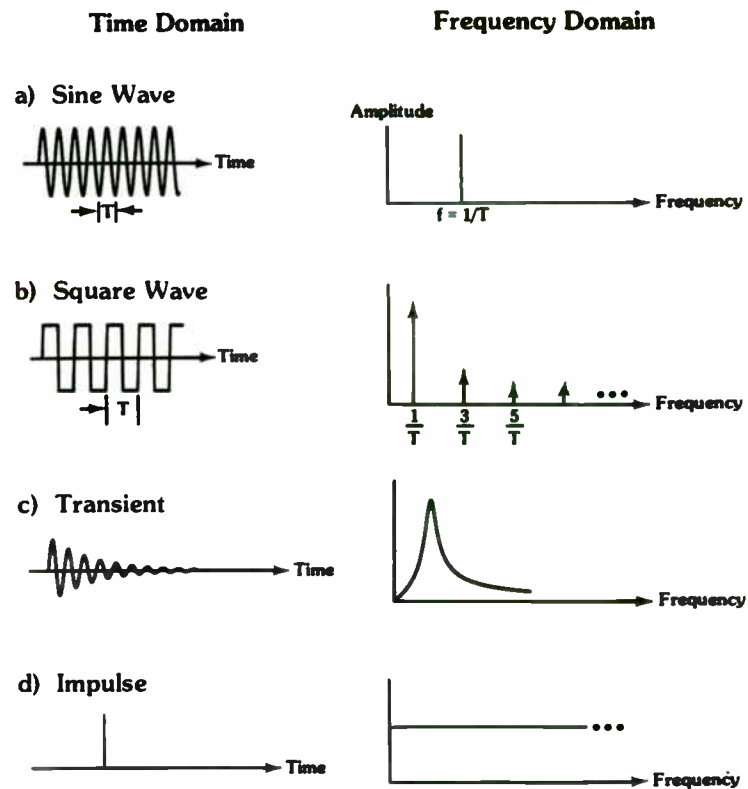


Figure 2. Frequency spectrum examples.

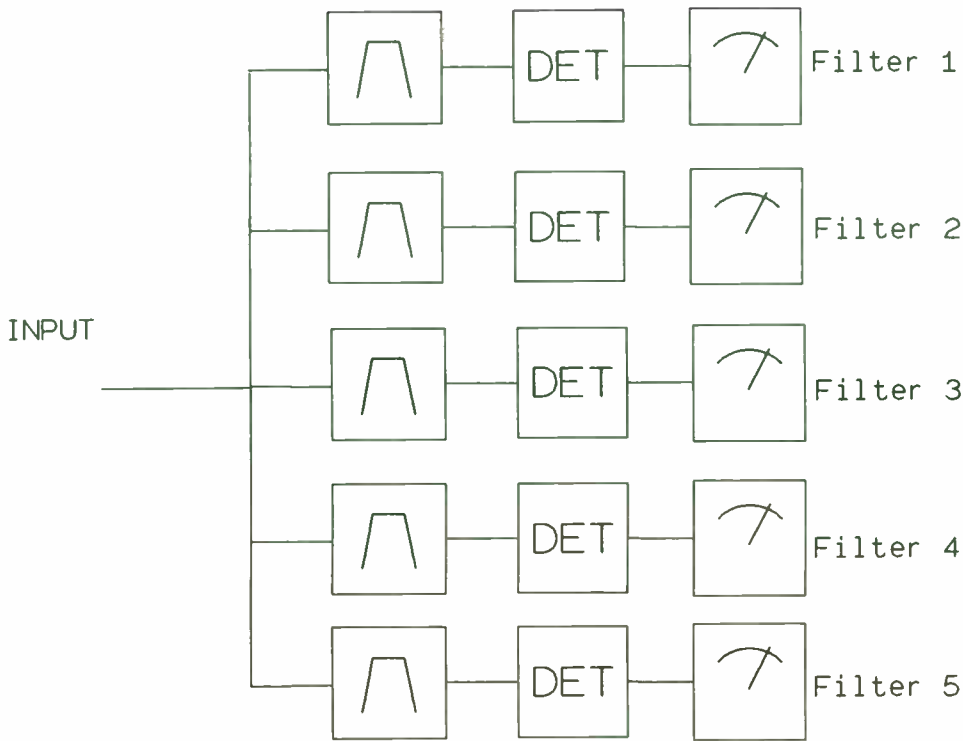


Figure 3. The bank-of-filters spectrum analyzer uses a set of filters to measure the frequency content of a signal.

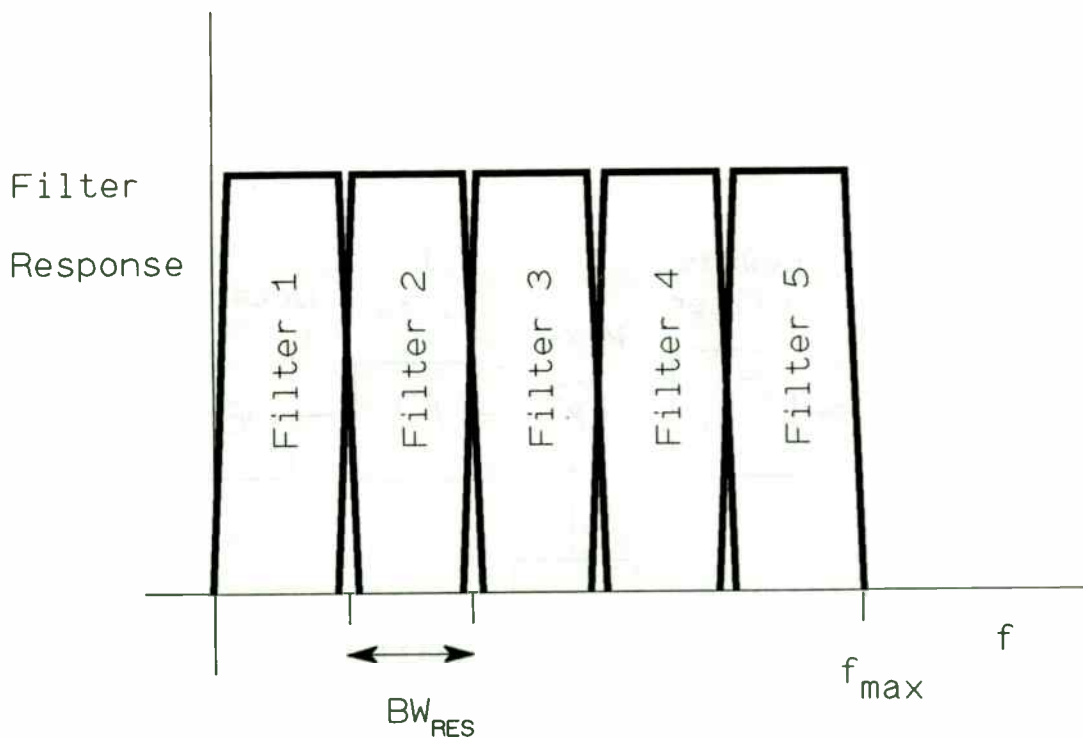


Figure 4. The filters are adjacent in the frequency domain, aligned for minimum overlap.

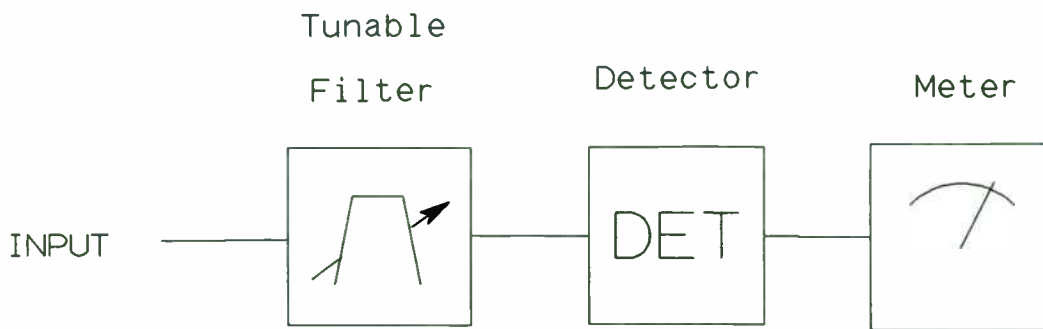


Figure 5. A conceptual block diagram of the wave analyzer.

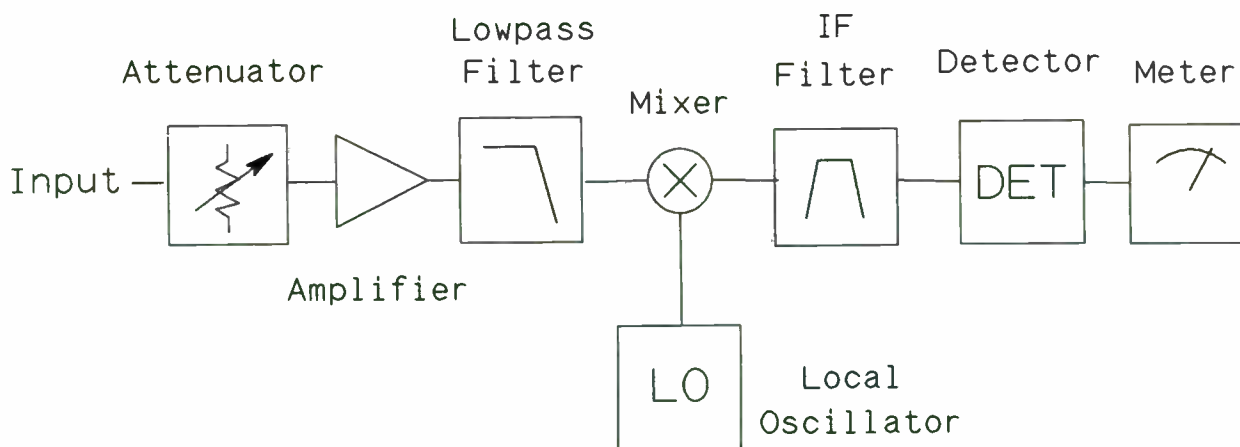


Figure 6. The block diagram of the wave analyzer.

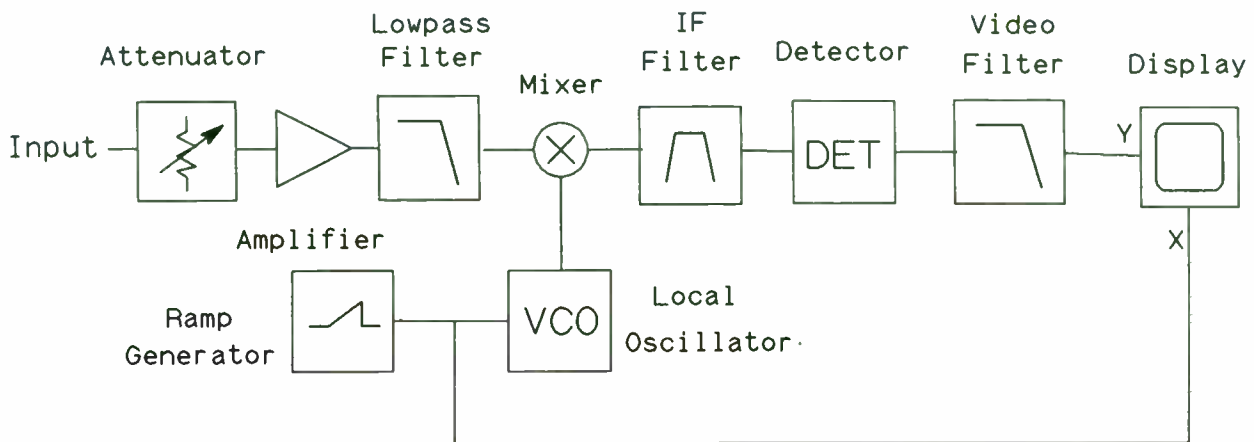


Figure 7. The simplified block diagram of a swept spectrum analyzer.

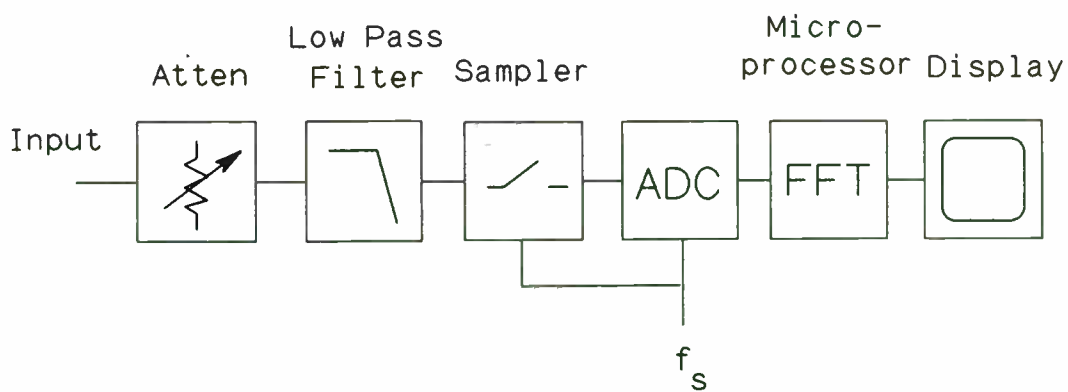


Figure 8. The simplified block diagram of the FFT spectrum analyzer. This diagram also applies to the digital scope with FFT.

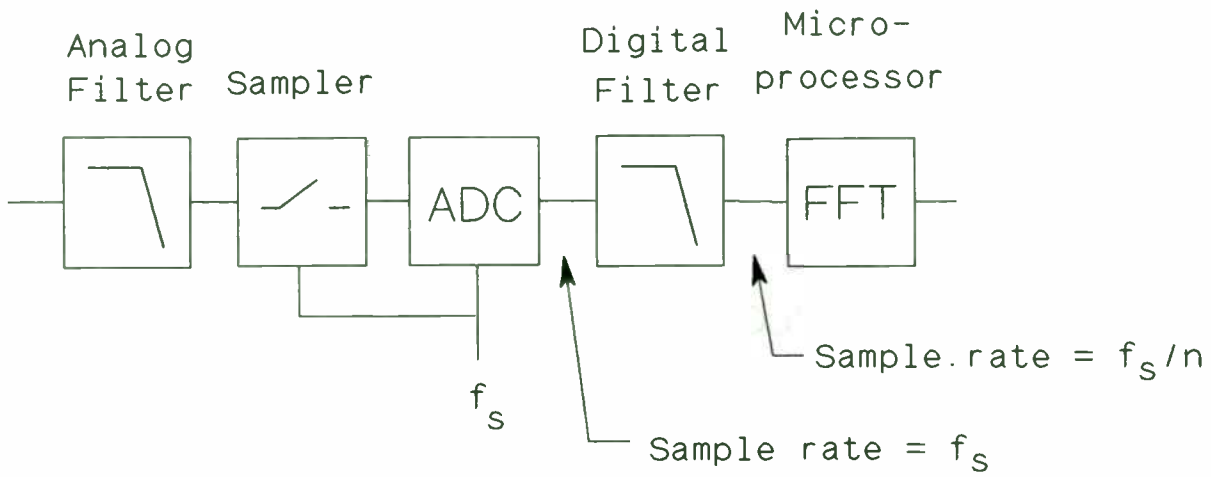


Figure 9. Decimating digital filters are used to reduce the sample rate into the FFT block.

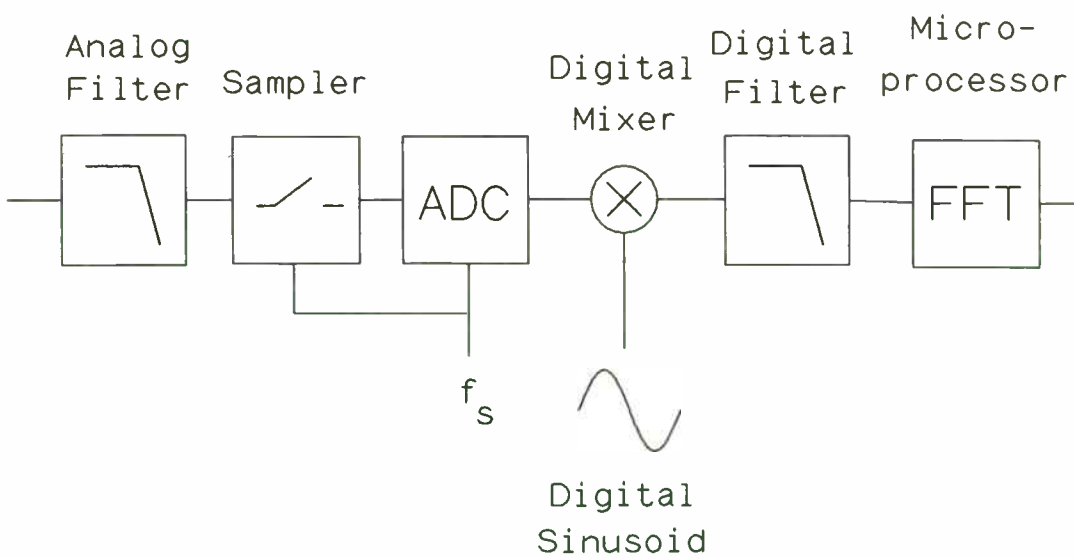


Figure 10. A digital mixer provides band selectable analysis.





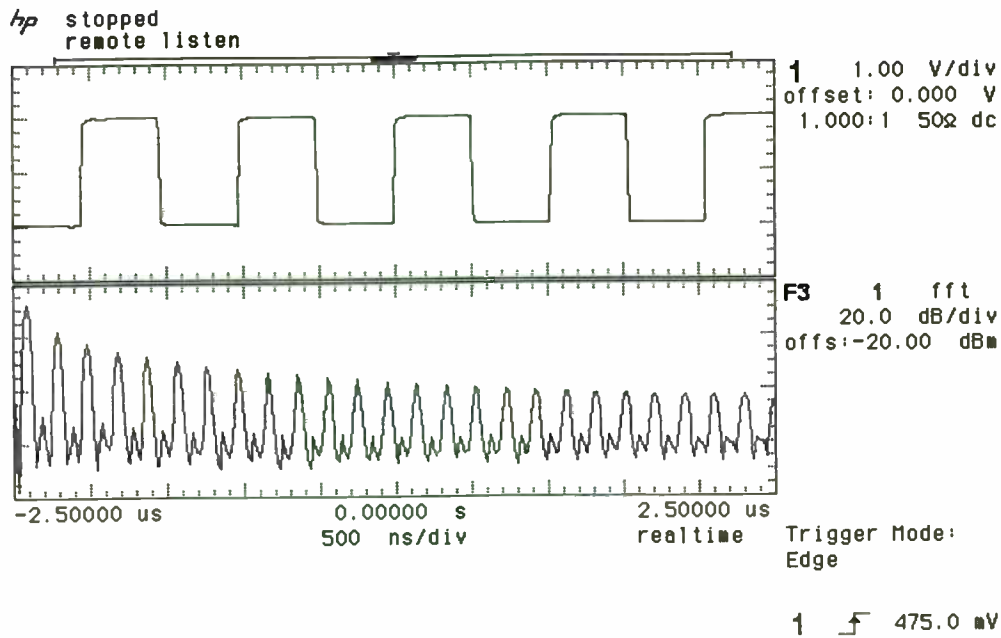


Figure 13. Simultaneous display of the time domain and frequency domain representation of a square wave, using an oscilloscope.

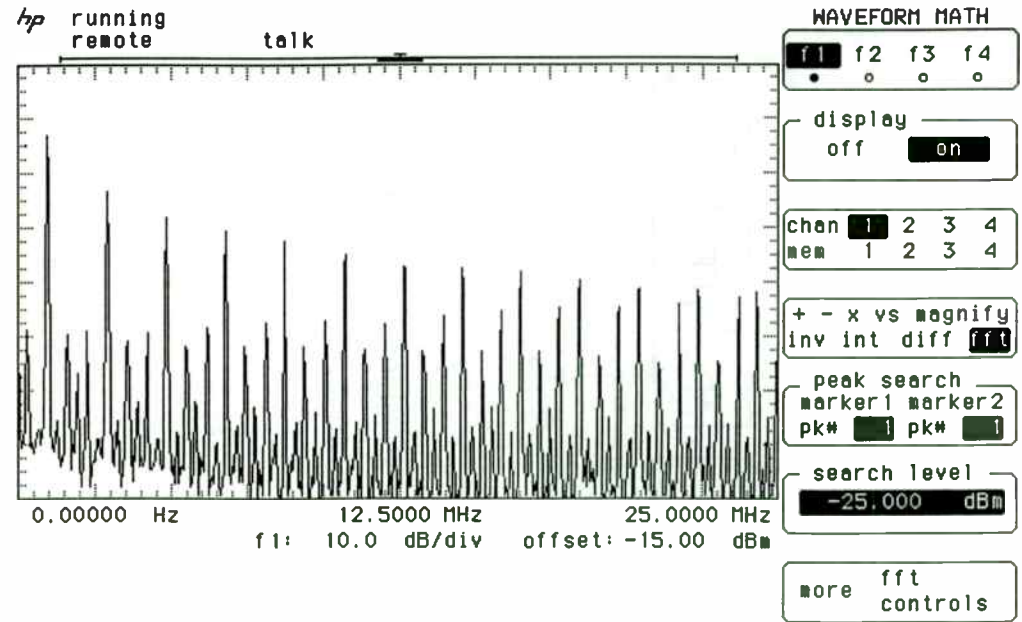


Figure 14. Insufficient sample rate causes this FFT measurement of a square wave to be aliased.

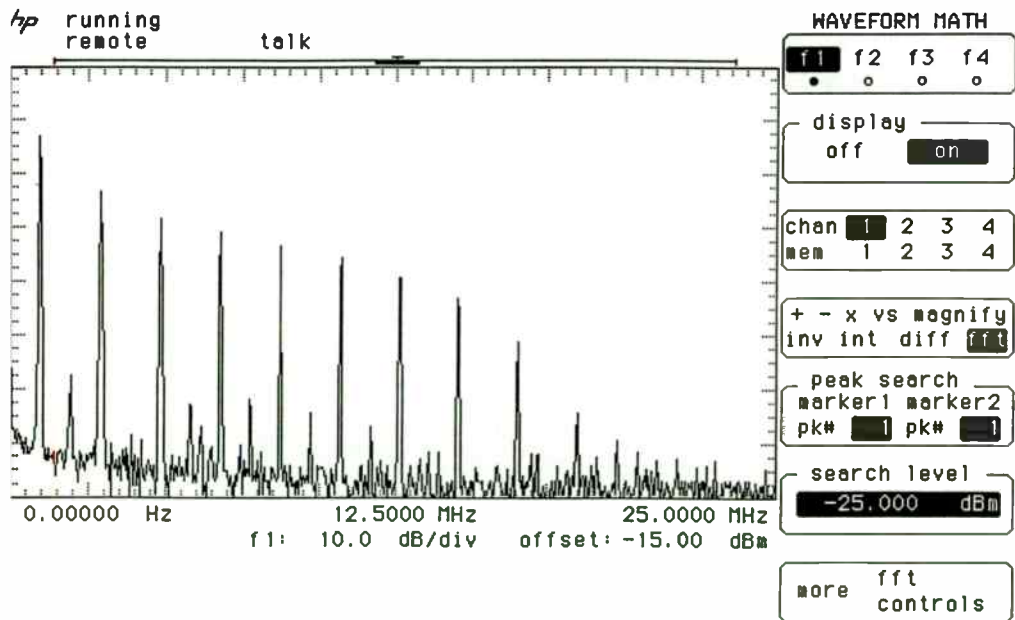


Figure 15. A lowpass filter installed in front of the scope prevents aliasing.

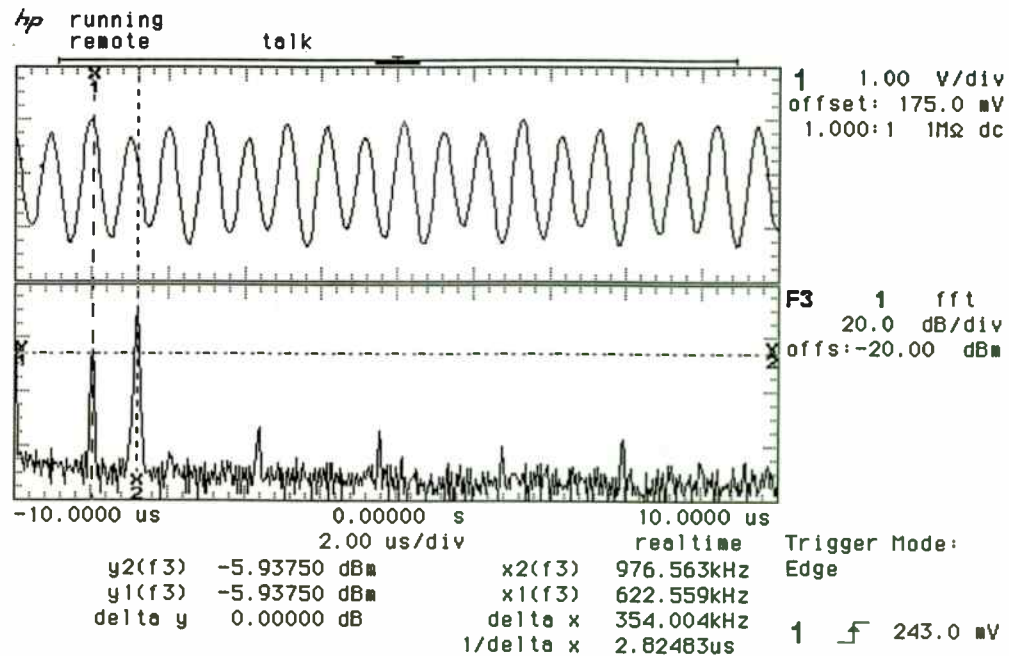


Figure 16. Two tones are difficult to identify in the time domain (upper trace), but are clearly visible in the frequency domain (lower trace).

Table 1. Typical Specifications and Features

	<u>Swept Analyzer</u>	<u>FFT Analyzer</u>	<u>Scope w/FFT</u>
Frequency Range (upper freq limit)	100 MHz to 50 GHz	100 kHz	100 MHz to 1 GHz
Maximum Sample Rate	N/A	250 kHz	4 GHz
Transient Capture?	No	Yes	Yes
ADC Resolution	N/A	12 to 16 bits	8 to 10 bits
Dynamic Range	70 to 80 dB	70 to 80 dB	50 to 55 dB
Frequency Resolution	Good	Excellent	Excellent
Zoom/Narrowband Sweeps?	Yes	Yes	Limited
Time domain capability	No	Yes	Yes
Protection against aliasing or spurious responses	Yes	Yes	Limited

**Design and Optimization of  
Lumped-Element LCM Directional Couplers**

Edmund (Joe) Tillo II, Ph. D.  
Ford Motor Company  
16800 Executive Plaza Drive  
Regent Court Building  
Suite 321  
Dearborn, Michigan 48126

**Abstract**

A new technique for designing directional couplers has been developed. This technique produces directional couplers that are realized entirely from inductances, capacitances, and mutual inductances (LCM), and closely approximate the quarter-wavelength coupled-transmission-line directional coupler in the sinusoidal steady state. The new design technique is based on the concept of even- and odd-mode impedances as applied to the design of coupled-transmission-line directional couplers. The crucial design condition is forcing the propagation functions for the even and odd modes to be equal to each other.

Examples of the new couplers are optimized using frequency-domain gradient-search techniques. The optimization process results in couplers with significantly improved input-port reflection coefficient and isolated-port transmission coefficient.

A prototype has been built and tested in the laboratory. Measurement and theoretical results are found to be in close agreement with each other.

**Background**

The lumped-element LCM directional coupler is designed to approximate the coupled-transmission-line directional coupler over a specific frequency range. This frequency range corresponds to an electrical length in the vicinity of one quarter wavelength for the path from the input port to the through port. This is accomplished by extending well-known coupled-transmission-line directional coupler design methods to lumped-element equivalent circuits.

The frequency response of the coupled-transmission-line directional coupler from the input port to the coupled port is described by Oliver<sup>1</sup>. Using the numbering convention of port 1 for the input port and port 2 for the coupled port, the following function of electrical length  $\theta$  and midband coupling factor  $k$  is found:

$$\frac{V_2}{V_1} = \frac{jk \sin \theta}{\sqrt{1 - k^2 \cos \theta + j \sin \theta}} \quad (1)$$

This function is periodic in  $\theta$ , and maximum coupling occurs at frequencies where the electrical length of the lines is given by:

$$\theta = \frac{(2n+1)\pi}{2}, n = 0, 1, 2, \dots \quad (2)$$

The first maximum of Equation 1 occurs where the electrical length of the transmission lines is equal to one quarter wavelength, hence the name *quarter-wavelength* directional coupler.

An example frequency response from the input port to the coupled port is shown in Figure 1. The

---

<sup>1</sup>Oliver, B.M. "Directional Electromagnetic Couplers." *Proceedings of the IRE* Vol. 42 no. 11 (November 1954): pp. 235-40

frequency range shown is near the first maximum, where the lumped-element LCM directional coupler approximates the coupled-transmission-line directional coupler. The maximum value reached by  $|V_2/V_1|$  is defined as the midband coupling factor  $k$ , and the points where  $|V_2/V_1|$  drops to  $0.8913k$  define the 1 dB bandwidth of the directional coupler.

The coupled-transmission-line directional coupler relies on electric and magnetic coupling for operation. This is also true for the lumped-element LCM directional coupler. Another property that the lumped-element LCM directional coupler and the coupled-transmission-line directional coupler share is that the propagation velocities for the normal modes are equal to each other. Oliver<sup>2</sup> points out that this condition can only be satisfied when the electric and magnetic coupling coefficients are made equal to each other.

### Design

A convenient design method for realizing coupled-transmission-line directional couplers is shown by Shimizu and Jones<sup>3</sup>. Their technique is based on the concept of even- and odd-mode impedances. The even-mode impedance of a coupled-transmission-line directional coupler is determined by driving a coupled pair of infinite transmission lines with equal currents, and calculating the impedance seen looking into one of the lines. The odd-mode impedance is determined in the same way, but with one of the current sources reversed.

The coupled-transmission-line directional coupler is described by the midband coupling factor and characteristic impedance. These are functions of the odd- and even-mode impedances:

<sup>2</sup>ibid.

<sup>3</sup>Shimizu, J.K. and Jones, E.M.T. "Coupled-Transmission-Line Directional Couplers." *IEEE Transactions on Microwave Theory and Techniques* Vol. MTT-6 no. 4 (October 1958): pp. 403-10

$$k = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}}, \quad (3)$$

$$Z_o = \sqrt{Z_{oe}Z_{oo}}. \quad (4)$$

The methods of Oliver and Shimizu and Jones can be extended to the lumped-element case with the application of the well-known concepts of bisection, as described by Temes and Lapatra<sup>4</sup>, and image impedance, as described by Lepage and Seeley<sup>5</sup>. The so-called pi- and T-type sections shown in Figure 2 are used as basic building blocks for the lumped-element LCM directional coupler. The equivalent odd- and even-mode circuits for the basic pi-type section are shown in Figure 3. Two or more of these sections can be connected end-to-end to create a multi-section lumped-element LCM directional coupler, as shown in Figure 4. In general, the frequency response of a lumped-element LCM directional coupler designed using the methods described in this paper more closely approximates the frequency response of a coupled-transmission-line directional coupler as the number of sections is increased. The design process will now be shown.

The image impedances and propagation functions for the pi-type even- and odd-mode circuits shown in Figure 3 are given by:

$$Z_{oe} = \sqrt{\frac{L+M}{C}} \times \sqrt{\frac{4}{4+s^2(L+M)C}}, \quad (5)$$

<sup>4</sup>Temes, Gabor C. and Lapatra, Jack W. *Introduction to Circuit Synthesis and Design* (New York: McGraw Hill Book Company, 1977), pp. 235-40

<sup>5</sup>Lepage, Wilbur R. and Seeley, Samuel. *General Network Analysis* (New York: McGraw-Hill Book Company, 1950), pp. 139-66

$$Z_{oo} = \sqrt{\frac{L-M}{C+2C_1}} \times \sqrt{\frac{4}{4+s^2(L-M)(C+2C_1)}}, \quad (6)$$

$$\gamma_{\text{even}} = \cosh^{-1}\left(1 + \frac{s^2(L+M)C}{2}\right), \quad (7)$$

and

$$\gamma_{\text{odd}} = \cosh^{-1}\left(1 + \frac{s^2(L-M)(C+2C_1)}{2}\right). \quad (8)$$

The corresponding equations for the T-type section are found to be:

$$Z_{oe} = \sqrt{\frac{L+M}{C}} \times \sqrt{\frac{4+s^2(L+M)C}{4}}, \quad (9)$$

$$Z_{oo} = \sqrt{\frac{L-M}{C+2C_1}} \times \sqrt{\frac{4+s^2(L-M)(C+2C_1)}{4}}, \quad (10)$$

$$\gamma_{\text{even}} = \cosh^{-1}\left(1 + \frac{s^2(L+M)C}{2}\right), \quad (11)$$

and

$$\gamma_{\text{odd}} = \cosh^{-1}\left(1 + \frac{s^2(L-M)(C+2C_1)}{2}\right). \quad (12)$$

Notice that the corresponding propagation functions are identical for pi- and T-type sections.

The propagation function is a complex number, representing the phase delay through the circuit.

The value of the propagation function is imaginary when the assumptions that the circuit is

lossless and perfectly matched are made. Since these assumptions are true for the lumped-element LCM directional coupler in the vicinity of maximum coupling, a hyperbolic trigonometric identity will be applied to Equations 7 and 8 and a real number  $\theta$  will be used to represent phase delay.

To approximate the performance of the coupled-transmission-line directional coupler, the phase delay through the lumped-element LCM directional coupler from the input port to the through port must be  $\frac{\pi}{2}$  radians at the intended center frequency. This implies that the phase delay  $\theta$  through each section must be  $\frac{\pi}{2n}$  where  $n$  is the number of sections connected end-to-end in the multi-section lumped-element LCM directional coupler.

When the propagation functions for the even and odd modes are equal to each other the electric and magnetic coupling coefficients are forced to be equal to the coupling factor of the lumped-element LCM directional coupler. Under these conditions, the equations relating  $L$  to  $M$  as well as  $C$  to  $C_1$  are obtained:

$$k = \frac{M}{L} \quad (13)$$

and

$$k = \frac{C_1}{C_1 + C}. \quad (14)$$

As shown in Equation 4, the characteristic impedance of a coupler,  $Z_O$ , is a function of  $Z_{OE}$  and  $Z_{OO}$ . Thus, manipulation of Equations 5 through 8 and 13 and 14 along with the application of  $s = j\omega$  produce equations for the characteristic impedance of the lumped-element LCM directional coupler. For the pi-type section one obtains:

$$Z_o = \sqrt{\frac{L(1-k)}{C}} \times \sqrt{\frac{4}{4 - \omega^2 LC(1+k)}}. \quad (15)$$

The characteristic impedance equation for the T-type section is:

$$Z_o = \sqrt{\frac{L(1-k)}{C}} \times \sqrt{\frac{4 - \omega^2 LC(1+k)}{4}}. \quad (16)$$

The phase delay equations for the pi- and T-type sections can be obtained by manipulating Equations 9 through 14. For the pi-type section the result is:

$$\theta = \cos^{-1} \left( 1 - \frac{\omega^2 LC(1+k)}{2} \right) \quad (17)$$

For the T-type section the result is identical:

$$\theta = \cos^{-1} \left( 1 - \frac{\omega^2 LC(1+k)}{2} \right) \quad (18)$$

Since  $M$  and  $C_I$  are obtained from  $L$  and  $C$  when the value of  $k$  is known, the design of a lumped-element LCM directional coupler can be realized when solutions for  $L$  and  $C$  are found. The values of  $L$  and  $C$  are determined by manipulating the equations for characteristic impedance and phase delay for the pi- and T-type sections. The pi-type section design equations are:

$$L = \frac{2Z_o \sin(\theta/2) \cos(\theta/2)}{\omega \sqrt{1-k^2}} \quad (19)$$

and

$$C = 2 \sqrt{\frac{1-k}{1+k}} \times \frac{\tan(\theta/2)}{\omega Z_o}. \quad (20)$$

The corresponding equations for the T-type section are:

$$L = \frac{2Z_o \tan(\theta/2)}{\omega \sqrt{1-k^2}} \quad (21)$$

and

$$C = 2 \sqrt{\frac{1-k}{1+k}} \times \frac{\sin(\theta/2) \cos(\theta/2)}{\omega Z_o}. \quad (22)$$

This concludes the design procedure.

Designs for pi-type and T-type 10 dB three-section lumped-element LCM directional couplers are shown in Figure 5. These couplers are normalized to a center frequency of one radian per second, and a characteristic impedance of one ohm. The associated frequency responses are shown in Figure 6. Notice the equivalent performance of the pi- and T-type couplers. Designs for pi-type one-, two-, and five-section lumped-element LCM directional couplers were also carried out, and the frequency responses for these circuits are shown in Figure 7. The improvement in performance obtained by increasing the number of sections is evident.

Lumped-element LCM directional couplers can be cascaded for increased coupled-port bandwidth, as done for coupled-transmission-line directional couplers by Crystal and Young<sup>6</sup>. The frequency response of an equal-ripple directional coupler built from a cascade of three pi-type three-section lumped-element LCM directional couplers designed following Crystal and Young is

<sup>6</sup>Crystal, E.G. and Young, L. "Theory and Tables of Optimum Symmetrical TEM-Mode Coupled-Transmission-Line Directional Couplers." *IEEE Transactions on Microwave Theory and Techniques* Vol. MTT-3 no. 4 (October 1955): pp. 29-38



shown in Figure 8. Notice that the bandwidth of this coupler is wider than that of the three-section lumped-element LCM directional coupler.

### Optimization

There is some discrepancy between the theoretical frequency response of the coupled-transmission-line directional coupler and the frequency response of the pi-type three-section lumped-element LCM directional coupler. In particular, the input port reflection coefficient and isolated port transmission coefficients are higher than expected, and the coupled port response is not symmetric about the intended center frequency. Computer optimization has been employed to reduce these discrepancies.

The frequency responses of lumped-element LCM directional couplers have been optimized relative to the frequency responses of coupled-transmission-line directional couplers using methods described by Director and Rohrer<sup>7</sup>. The optimization program is written to preserve the symmetry of the network, and thus adjusts eight unique parameter values in a pi-type three-section lumped-element LCM directional coupler, as shown in Figure 9. A listing of the optimization program used was published in a recent dissertation<sup>8</sup>.

As an example, the input, coupled, and isolated port frequency responses for a pi-type 10 dB three-section lumped-element LCM directional coupler before and after optimization are shown in Figure 10. Note the improvement in input and isolated port performance, as well as the increased

<sup>7</sup>Director, S.W. and Rohrer, R.A. "Automated Network Design-The Frequency-Domain Case." *IEEE Transactions on Circuit Theory* Vol. CT-16 no. 3 (August 1969): pp. 330-6

<sup>8</sup>Tillo, E.J., *Design and Optimization of Lumped-Element LCM Directional Couplers*. Ph. D. dissertation, The University of Michigan 1991

symmetry of the coupled port response about the intended center frequency.

Several pi-type three-section lumped-element LCM directional couplers with coupling factors from 3 to 30 dB were designed and optimized. The element values for the initial design and optimized couplers are plotted parametrically in Figures 11 and 12. The linear relationships between the various element values before and after optimization make it possible to design "optimized" couplers without a computer. The linear equations relating the optimized parameter values to the initial component values for the pi-type three-section lumped-element LCM directional coupler are:

$$P1 = 5.3693e - 2 + 0.7885L \quad (23)$$

$$P2 = -0.10835 + 1.4897L \quad (24)$$

$$1/P3 = 1.2694e - 3 + 0.84955(C1/2) \quad (25)$$

$$1/P4 = -9.8100e - 4 + 1.0320(C1) \quad (26)$$

$$1/P5 = -1.4843e - 2 + 0.69221(C/2) \quad (27)$$

$$1/P6 = 1.2671e - 2 + 1.0904(C) \quad (28)$$

$$P7 = -2.2238e - 3 + 0.82056(M) \quad (29)$$

$$P8 = 4.3173e - 3 + 1.4326(M) \quad (30)$$

Table 1 shows the electric and magnetic coupling for pi- and T-type three-section couplers after optimization. The duality of the electric and magnetic coupling among the pi- and T-type lumped-element LCM directional couplers is encouraging in that it indicates that the optimizer program may be finding a global minimum.

An attempt to broaden the coupled-port frequency response of the lumped-element LCM

directional coupler with computer optimization was made. This attempt resulted in couplers that exhibited virtually the same bandwidth as the non-optimized couplers<sup>9</sup>, and hence was not successful.

### Laboratory Experiments

An optimized pi-type 8.34 dB three-section lumped-element LCM directional coupler was built and tested in the laboratory. The coupling factor of 8.34 dB and a center frequency of 864 KHz were selected to obtain convenient element values. The frequency responses obtained from theory, simulation with measured component values, and measurements on the assembled coupler in the laboratory are shown in Figure 13. Note that all three sets of results are in close agreement with each other.

### Concluding Remarks

Techniques for design and optimization of lumped-element LCM directional couplers have been shown. The performance of circuits obtained using these techniques has been tested by computer and verified in the laboratory. Two structures, the pi and T type, have been proposed. Since the performance of one is equivalent to that of the other, the pi-type section may be more useful. This is because inductors are generally more expensive than capacitors and the pi-type lumped-element LCM directional coupler uses one less pair of coupled coils than the equivalent T-type lumped-element LCM directional coupler.

### Acknowledgment

Guidance for this project was provided by Professor A.B. Macnee of The University of Michigan.

<sup>9</sup>ibid.

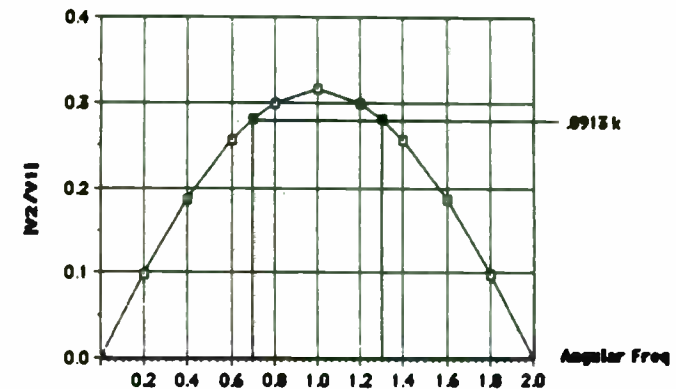


Figure 1. Coupled-Port Frequency Response of a 10 dB Coupled-Transmission-Line Directional Coupler

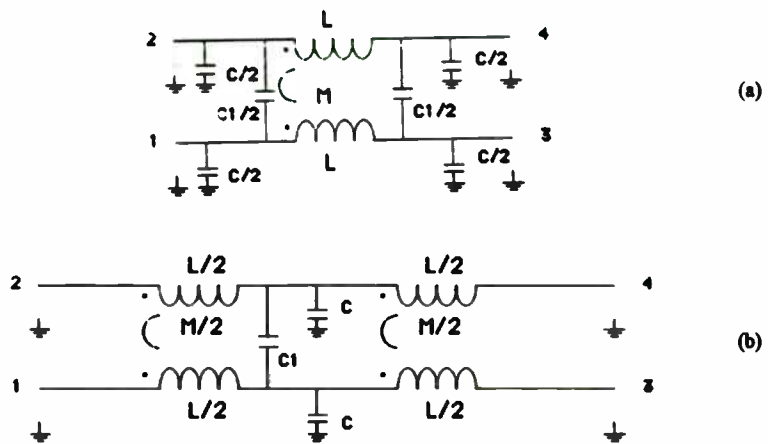


Figure 2. Lumped-Element LCM Directional Coupler Prototypes (a) Pi Type (b) T Type

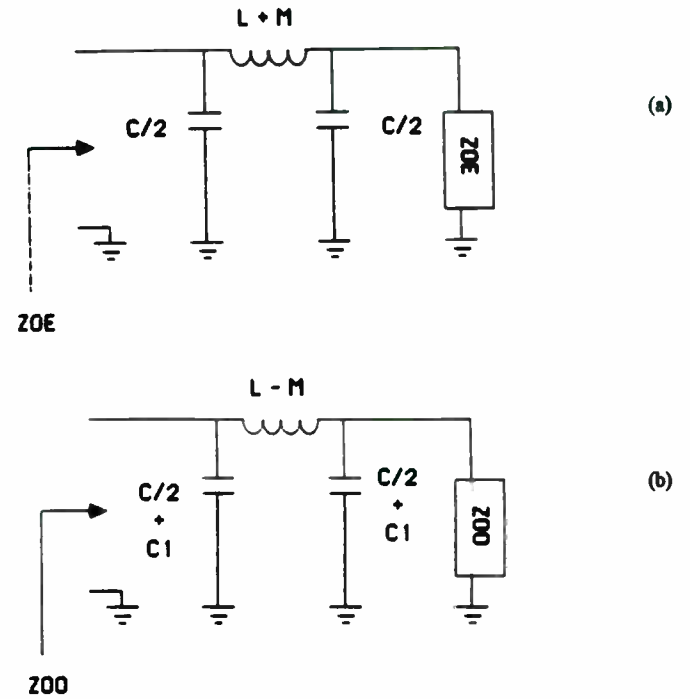


Figure 3. Equivalent Circuits for the Pi-Type Lumped-Element LCM Directional Coupler (a) Even Mode (b) Odd Mode

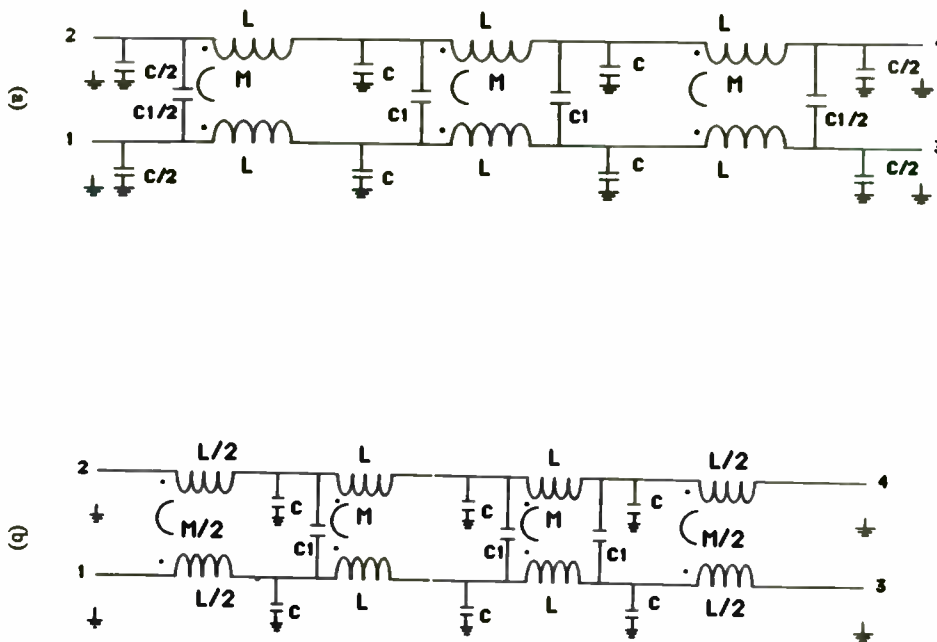
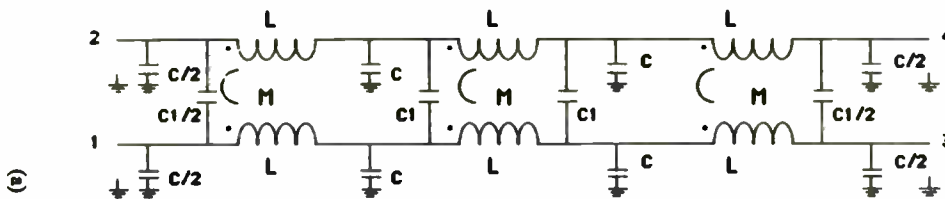
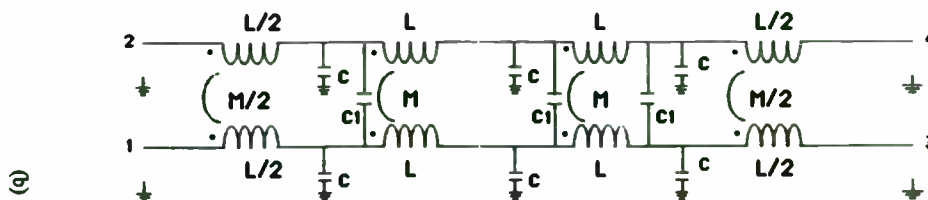


Figure 4. Three-Section Lumped-Element LCM Directional Couplers (a) Pi Type (b) T Type



$L = 52704628 \text{ H}$   
 $M = 16666667 \text{ H}$   
 $C/2 = 19312685 \text{ F}$   
 $C1/2 = 0893164 \text{ F}$



$L/2 = 2864 \text{ H}$   
 $M/2 = 09057 \text{ H}$   
 $C = 3604 \text{ F}$   
 $C1 = 1665 \text{ F}$

Figure 5. 10 dB Three-Section Lumped-Element LCM Directional Couplers (a) Pi Type (b) T Type

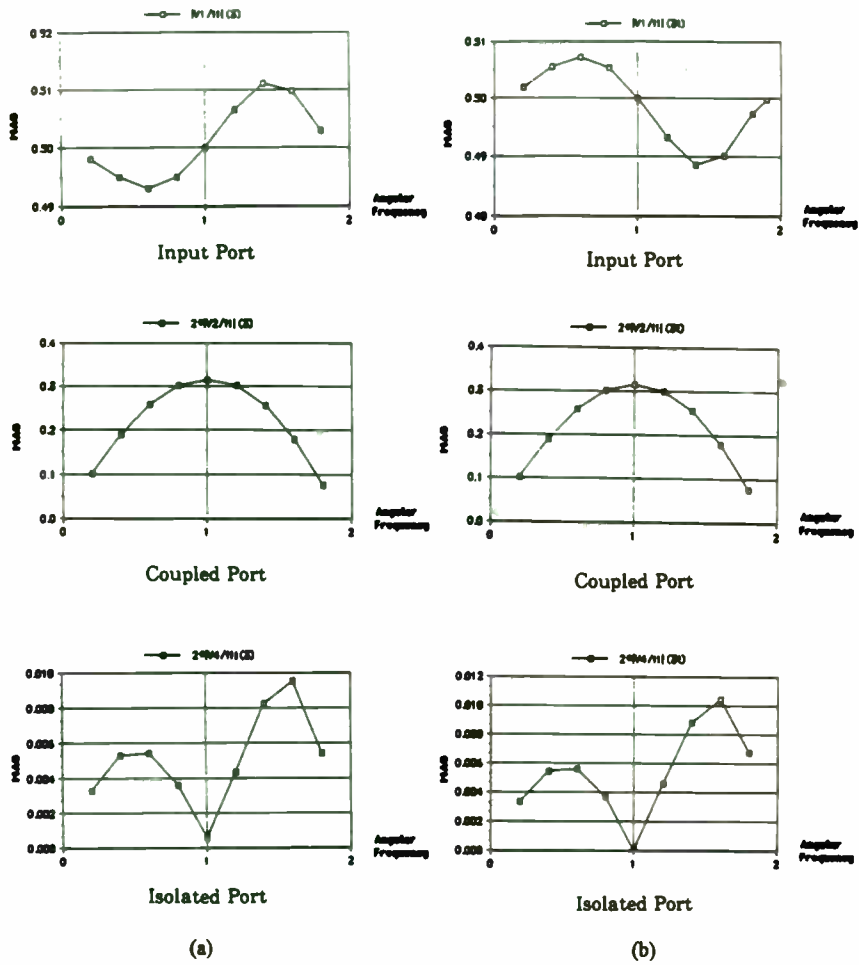


Figure 6. 10 dB Three-Section Lumped-Element LCM Directional Coupler Frequency Response  
 (a) Pi Type (b) T Type

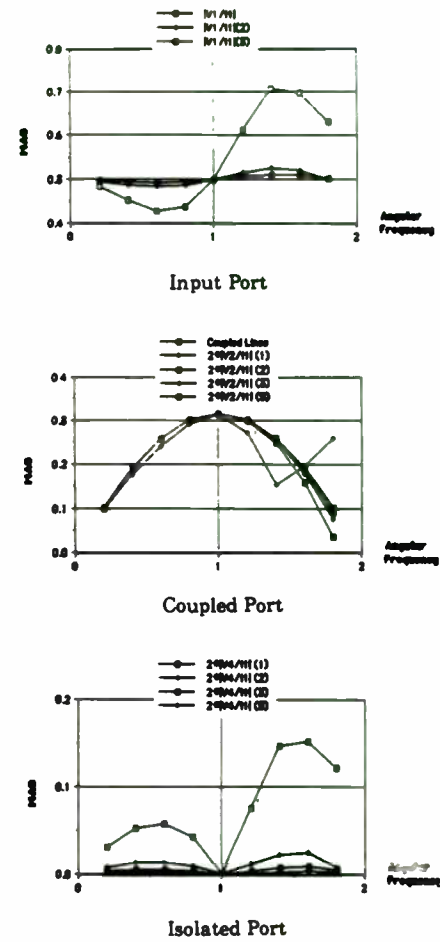


Figure 7. Frequency Response of One- to Five-Section Lumped-Element LCM Directional Couplers

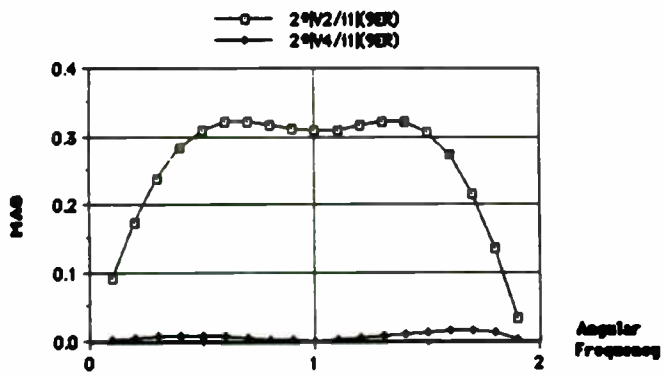


Figure 8. Coupled and Isolated Port Frequency Response of a Cascade of Three Three-Section Lumped-Element LCM Directional Couplers

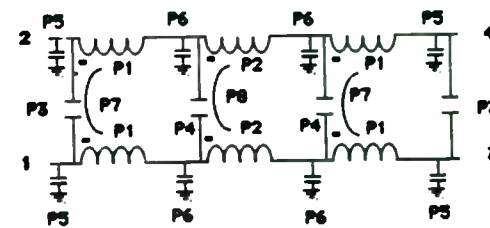


Figure 9. Pi-Type Three-Section Lumped-Element LCM Directional Coupler Labeled to Show Symmetry

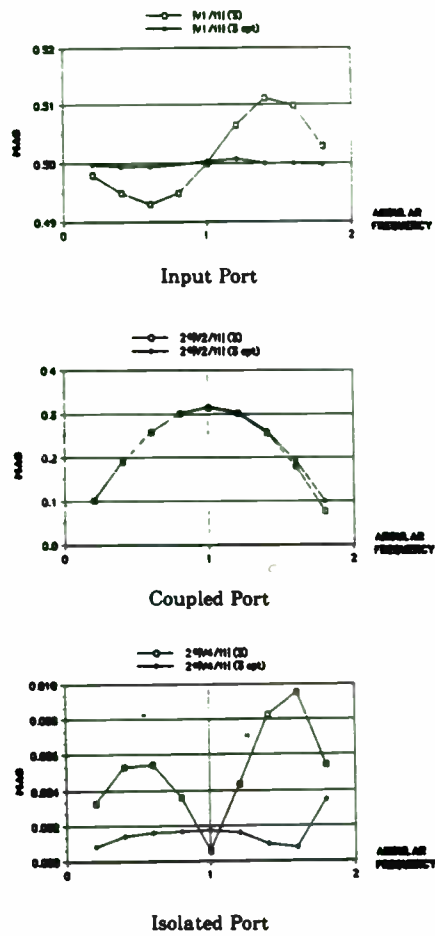


Figure 10. Initial and Optimized Frequency Response of the Pi-Type 10 dB Three-Section Lumped-Element LCM Directional Coupler

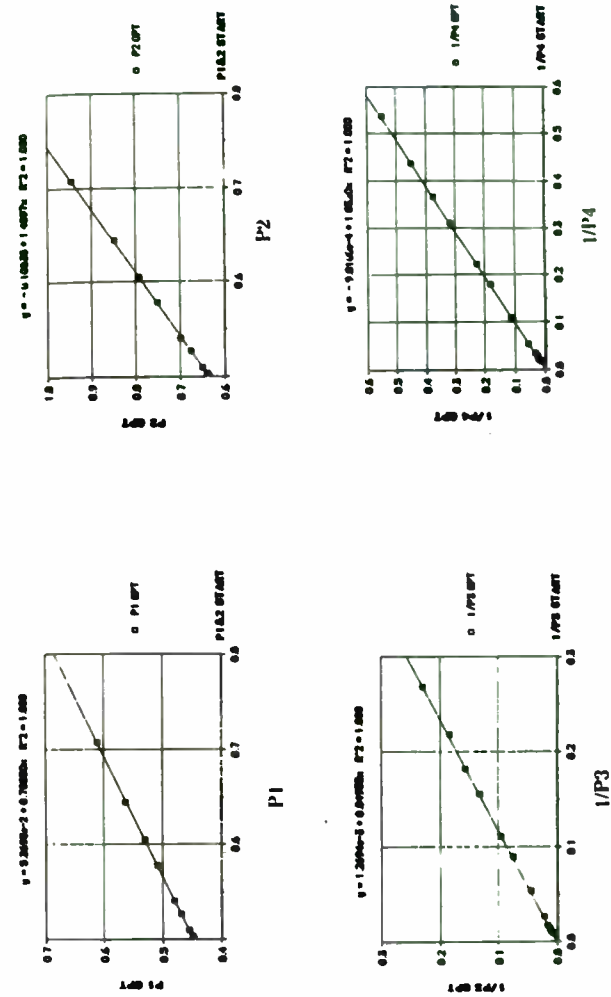


Figure 11. Parametric Plots of Design versus Optimized Element Values -  $P1$  Through  $P4$

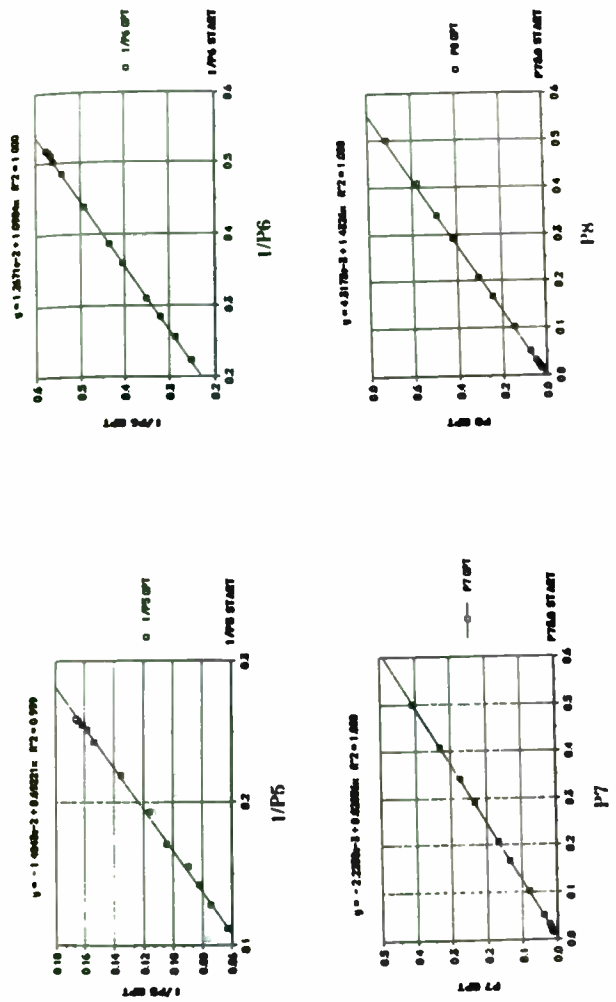


Figure 12. Parametric Plots of Design versus Optimized Element Values - P5 Through P8

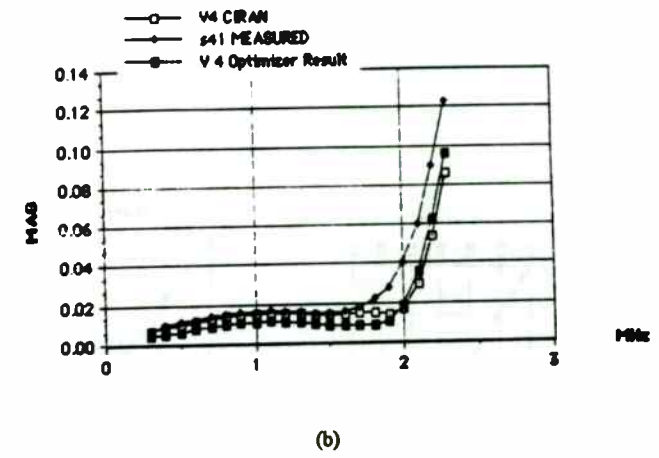
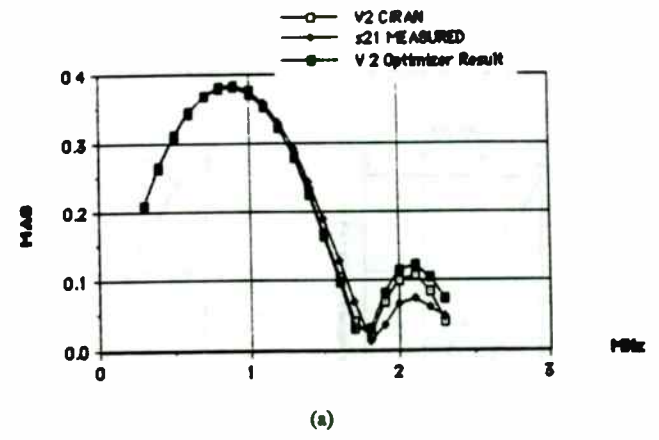


Figure 13. Comparison of Theoretical, Measured, and Simulated-with-Component-Values Frequency Responses of the Prototype Pi-Type Three-Section Lumped-Element LCM Directional Coupler (a) Coupled Port (b) Isolated Port



Coupler Type	$\kappa$ outside	$\kappa$ center
$\pi$	.2847	.3614
T	.3960	.2521

(a)

Coupler Type	$\frac{C_1}{C_1 + C}$ outside	$\frac{C_1}{C_1 + C}$ center
$\pi$	.3950	.2530
T	.2843	.3618

(b)

Table 1. Coupling in the Three-Section Lumped-Element LCM Directional Coupler (a) Magnetic

(b) Electric



# RF and Microwave Circuit Design by Addition and Removal of Components

Kiomars Anvari  
Teknekron Communications Systems  
Berkeley, California

## Abstract:

A new technique for the optimization of electrical networks has been developed whereby the computer program removes components or add new components and nodes to a circuit in a systematic but apparently random manner in order to achieve specified performance targets. It is used to design radio frequency amplifiers, filters, matching circuits, and etc. The program minimizes the difference between the calculated design parameters of the circuit and the specified target values of the parameters. The minimization is achieved by adding and deleting components to the circuit each time the difference is more than an specified value. The process continues until the target design is approached.

Although the assessment of a large number of possible alternative circuits consumes a large amount of computer time, two examples show that acceptable circuits are produced with very little operator interaction and reasonable computer time.

The program is suitable for integrated circuit design since constraints may be imposed, limiting the number, type, and size of elements in the circuit. The program will optimize the values of the elements in a specified circuit as well as developing new topologies.

## Introduction:

The trend in electronics manufacturing today is to design products that are lighter, smaller, more efficient, more reliable, easy to manufacture, and easy to maintain-not to mention more economical to produce. The first generation of any technological innovation is designed from the start to become a museum relic: equipment built to spec, but not for mass production or consumption. However, with rapid advances in technology manufacturers overcome the obstacles gradually and produce equipments which are suitable for mass productions.

In order to satisfy all or some of the product requirements which mentioned above for future products considerable integration is unavoidable. Any integration is preferable to be in a form of chip set which can result in minimum development time, easy and speedy productions process and generic applications. Monolithic integrated circuits, HIC, VLSI and LSI are perhaps the most significant force behind the evolution of next generation electronic products. These technologies will allow manufacturers to design units that are smaller, better capable of handling a myriad of functions and more economical.

The advent of monolithic RF and Microwave integrated circuits and the desire for minimum chip size has encouraged the investigation of new design techniques. Despite the fact that application of the conventional circuit design techniques simplifies the process, a good design requires considerable network synthesis, or trial and error using tables, charts, even though computer is used. In a desire to automate the effects of the designer's work on various circuits, and to eliminate the effects of the designer's preconceptions, a free-ranging program has been developed for two port network design.

This paper deals with a new technique for the optimization of electrical networks whereby a computer program removes components or adds new components and nodes to a circuit, in an apparently random manner, in order to achieve specified performance targets. It is capable of handling most of the usual configurations. The program is designed to operate with the usual RF and microwave active (transistor, diode) and passive lumped and distributed components. The program is given an initial circuit which specifies the complexity and the characteristic of the components to be used. It then removes or adds elements to the circuit in a systematic way in order to approach the target design performance. It optimizes for gain, input and output VSWR, bandwidth, noise figure, stability and etc. The program is suitable for integrated circuit design since constraints may be imposed, limiting the number and type and size of elements in the circuit. The program will optimize the values of the elements in a specified circuit as well as developing new topologies. This paper describes the different features of the design program and gives two applications.

## 2. Design procedure

There are various methods of RF and microwave circuit analysis suitable for CAD techniques. In our work the overall circuit performance parameters are evaluated from the characterization of individual components. Each element is considered as a two-port network and is represented by its T-matrix. S-parameters are used for the characterization of all circuits, so that the S-parameters are obtained from the overall T-matrix. The parameters of the circuit must be optimized in order to achieve the best performance. There are two different ways of optimization, known as gradient methods and direct search methods. For this particular problem it was found that the gradient technique is more appropriate. There are a few commonly used gradient methods from which the quasi-Newton technique has been adopted. The optimization process involves the minimization of an error function composed of the sum of squares of the differences between the circuit design targets and the calculated performance.

The program will analyze the performance of the circuit given to it and optimize the values of the components in that circuit in order to approach specified performance targets. If the optimized initial configuration does not meet the design targets, the program searches for a new circuit which either meets the specified targets or gives improved performance compared with the previous circuit. The program adds or removes elements from the circuit based on a specific criterion. Components can be added either in parallel or in series with existing circuit elements. After each removal or addition, the new topology will be optimized. If the new topology provides improved performance, it will be retained and the old configuration will be replaced with the new one. If the target design is reached, the program stops with the final circuit as the result. Otherwise the process continues.

The program keeps adding elements indefinitely if the target design is not met, so that a criterion for stopping the program needs to be defined. One possible criterion is a specified value of the error function, but this might result in a topology having a large number of elements. Therefore, the maximum acceptable number of components in each part of the circuit is another reasonable criterion. If the required performance is not achieved, the designer will have to modify the design constraints. There are times when a circuit with only a distributed element network, a lump element network, a combination of both, or a lossy equalizer is required. Some designers might not wish to have inductors or another component type in their circuit. Any constraints on the type of component used, or the number of elements in a network, may be included in the program, so that the designer may achieve the target design with his desired circuit configuration. Constraints on the values of the components can be easily defined

## 3. Specimen amplifier and filter design

The program was tested for different amplifier and filter circuits, two of which are presented below. The transistor which was used is a standard X3 FET, whose S-parameters are shown in Table 1. All the results were obtained on a personal computer.



**Table 1: Standard X3 FET S parameters**

Frequency GHz	S11		S21		S12		S22	
	MAG	ANG deg	MAG	ANG deg	MAG	ANG deg	MAG	ANG deg
8	0.895	-66	1.41	115	0.068	52	0.839	-27
9	0.879	-72	1.35	108	0.075	49	0.833	-30
10	0.863	-78	1.3	101	0.082	46	0.828	-33
11	0.849	-83	1.23	96	0.088	44	0.823	-36
12	0.835	-88	1.18	91	0.094	41	0.819	-39

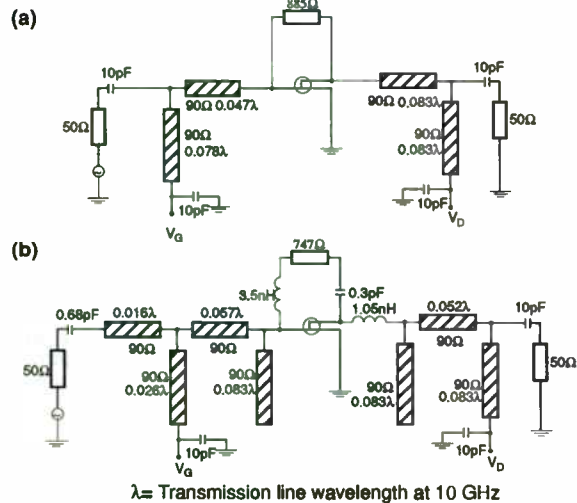
**3.1 Broadband single-stage feedback amplifier**

This example illustrates the performance improvement of a single stage feedback amplifier. Design specifications are:

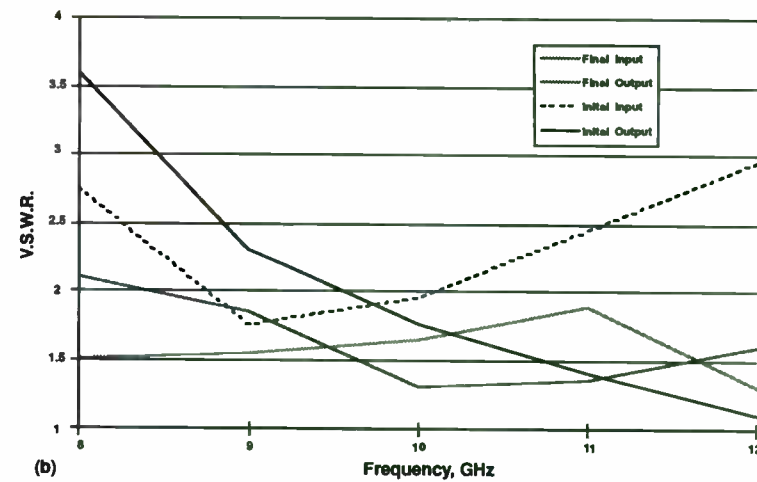
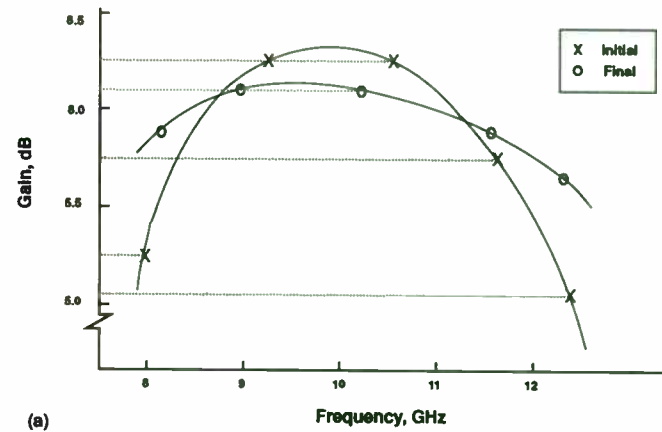
Frequency 8-12 GHz  
 Gain 6 +/- 0.5 dB  
 VSWR less than 2:1

The available components are:

Resistor 50-5000  
 Inductor 0-5 nH  
 Capacitor 0-10 pF  
 Trans line Zo = 90



**Figure 1: Circuit Diagrams of the Single-Stage Feedback Amplifier**



**Figure 2: Performance of the Circuits shown in Figure 1.**

The initial circuit schematic diagram of the amplifier stage is shown in Fig. 1a. This circuit is now optimized with the error function composed of the sum of squares of the individual error functions of gain, input and output VSWR and stability factor. After optimization, the error function (EF) was reduced to 0.95, which corresponds to Fig. 1a. As the desired value of EF was set to 0.05, the adding and removal stage was entered. The maximum allowable number of components in the input and output equalizers was chosen to be five and in the feedback network three. After examination of 34 topologies, the computer produced the final circuit shown in Fig. 1b. Despite the fact that error function did not reach the target value, the program stopped because of the constraints on the number of elements in each network. In Fig. 2, curves of gain and input and output VSWR for the initial and final circuits are shown. The improvement in performance of the circuit can be seen. The values of the components and characteristics of the transmission lines are appropriate for monolithic integrated circuits.

### 3.2 Broadband Band Pass Filter

In a communication channel, filters play one of the most important roles. Primarily, to bandlimit the channel noise or reject the interference and spurious signals. In order to avoid any signal distortion, it is essential that these filters have flat group delay and amplitude response across the passband.

In the case of a bandpass filter conventional design techniques result in amplitude asymmetry and group delay distortion. In this example we have compared the transfer function of a bandpass filter using conventional design method and the approach presented in this paper. The filter design specifications are:

Center frequency	70 MHz
Input/Output impedances	75 $\Omega$
-3 dB Bandwidth	22 MHz
-11 dB Bandwidth	40 MHz
Group delay ripple	1 nsec across 22 MHz

The circuit schematic diagram generated from the program is shown in figure 3. In figure 4a and 4b the amplitude response and group delay response of the filters are shown. In the conventional method a low pass Bessel filter is simply transformed to bandpass. The program generated filter has used the above parameters as its target design.

It can be seen that the filter transfer function is significantly improved. The component values shown in figure 3 also show that the filter can be easily realized.

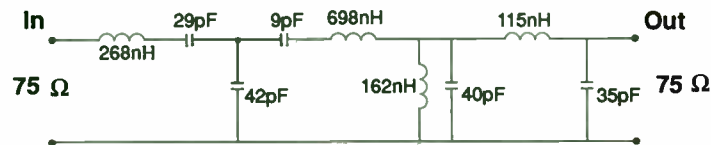


Figure 3: Circuit diagram of the filter generated by the Program

### 4. Conclusion

In most integrated circuits, chip area is the dominant feature limiting the value of lumped elements and transmission lines. Microstrip transmission lines in RF and Microwave networks use the largest amount of chip area. It is obvious that long lengths of transmission lines are not practical, but low characteristic impedances have a more significant effect, and can cause a coupling problem. Therefore, in the two examples which are given above, the characteristic impedance of the transmission line was specified to be 90 and 100, respectively, and the bounds on the line length of the microstrips were 0.0005 m and 0.003 m (free space). The lumped element limits which were used were those recommended as practical by the IC manufacturer. The examples given illustrate the efficiency and versatility of the program for RF and microwave circuit design, monolithic integrated circuits, in particular.

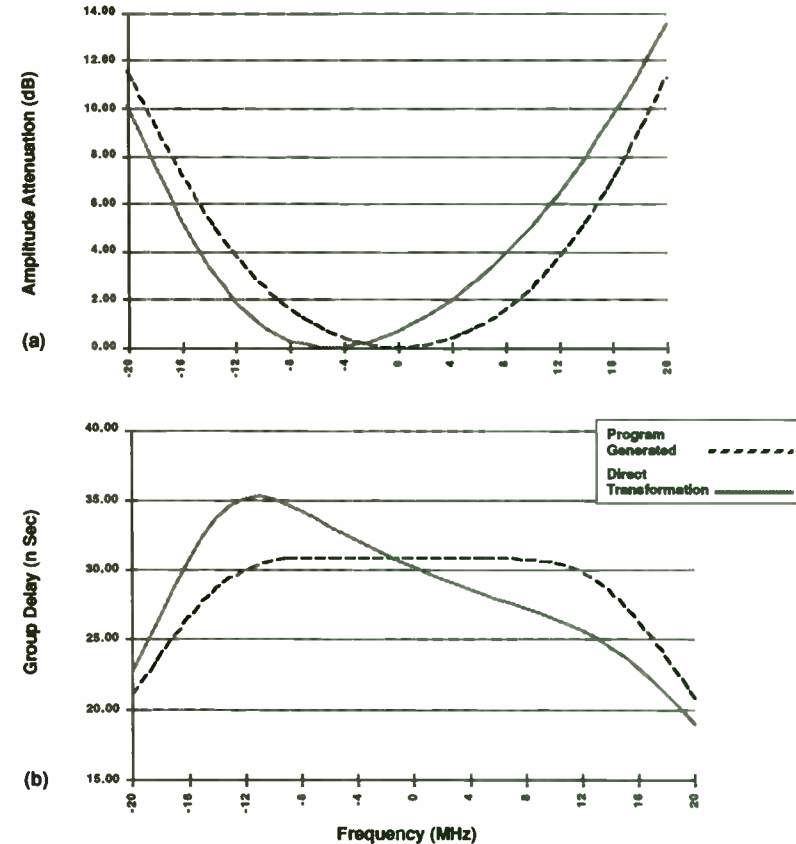


Figure 4: Filter's Amplitude and Group Delay Responses.



## FIELD STRENGTH MEASUREMENTS FOR FCC COMPLIANCE

by  
Earl McCune  
Proxim Inc.

**ABSTRACT:** The FCC has issued its regulations on maximum radiated signal strength in units of microvolts (or millivolts) per meter. These are the units of electric field, rather than RF power. This paper describes a process used to convert the regulatory field strength numbers into the more conventional power measurement units of watts and dBm. From this conversion process a method is described that allows any antenna type to be calibrated for the measurement of field strength. The accuracy of this simple field strength measurement is good to a few dB, which is adequate for quick checks of compliance with regulatory limits. THIS TECHNIQUE IS AN APPROXIMATION, AND CANNOT BE SUBSTITUTED FOR A CALIBRATED LABORATORY FOR ACTUAL CERTIFICATION.

Consider the problem facing the Federal Communications Commission (FCC) when it has to write a regulation. The new regulation must be specific enough to meet the purpose of minimizing interference from any system developed to and operating under the regulation. However, the regulation must not dictate how any system must be built, since the FCC also must encourage creativity from the RF communications industry.

These requirements appear to directly conflict with each other. What the FCC decided to do was to regulate the radiated signal itself, and not the equipment generating the signal. Physics tells us that the primary parameter of a radiating signal is its field - therefore field strength is the way the FCC guarantees control of a radiated signal.

An important side effect of choosing field strength as the parameter to be regulated, from the FCC's point of view, is the inherent simplicity of such regulations. Transmitter DC input power, output power, antenna type, and other hardware specific parameters do not necessarily need to be addressed.

### ISOTROPIC RADIATED FIELD STRENGTH

The most straightforward way to address field strength is to start from an isotropic (equal in all directions) radiated field. A search through the literature for isotropic radiated field equations came up surprisingly empty. Fortunately the isotropic case can be addressed very simply, and one derivation is presented here. This derivation is based on Ohm's law.

Ohm's Law is the basic relationship between voltage (E) and power (P):

$$P = E^2 / R \quad (1)$$

Expanding this basic relationship to radiated fields is best done using units analysis. In a radiated field the voltage does not exist at a point, as in a circuit, but must be measured over a distance. The units for E are therefore Volts per meter (V/m). The resistance encountered by a radiated field is the free space impedance ( $\eta$ ), which has a value of  $120\pi = 377\Omega$ . Applying these units to the right hand side of (1) gives

$$\begin{aligned} P[\text{units}] &= (V/m)^2 / \Omega \\ &= (V^2 / \Omega) / m^2 \\ &= W / m^2 \end{aligned} \quad (2)$$

So in a radiated field we do not consider absolute watts, but rather a power density of Watts per square meter.

However, we can only measure power, not power density. To relate these two, we use the straightforward concept that all radiated power must flow away from the antenna. Mathematically, this means that an integral of the radiated power density taken over a closed surface surrounding the antenna must equal the total radiated power from the antenna. If this were not true, then conservation of energy would be violated.

Notice that this is effectively the same as Gauss' Law, which states that the electric field emanating from a closed surface is governed solely by the total charge enclosed by that surface. This comes from conservation of charge rather than conservation of energy, but they both express the same idea.

Closed surface integrals are usually nasty mathematical problems. We can simplify the mathematics and avoid the integral completely if we choose the surface and power density carefully. The isotropic case naturally brings both of these into their simplest form.

Isotropic is, by definition, uniform in all directions. By defining the radiated field as isotropic, we are spreading the total radiated power uniformly in all directions. This makes the power density constant at any given distance. The closed surface of choice is one that has uniform distance in all directions: the sphere. These choices allow the difficult integrals to be completely avoided.

The surface area of a sphere of radius (r) is well known to be  $4\pi r^2$ . Spreading the total radiated power (W) over this area provides the isotropic power density ( $D_p$ )

$$D_p = \frac{W}{4\pi r^2} \quad (3)$$

Equating this with the radiated power density relation of (1, 2) gives

$$\frac{E^2}{\eta} = \frac{W}{4\pi r^2} \quad (4)$$

This can be solved for the radiated field strength E:

$$E = \sqrt{\frac{W\eta}{4\pi r^2}} \quad (5)$$

The units of E are Volts per meter at the distance r from the antenna.

Later we will also use the inverse relationship, in which (4) is solved for the total radiated power W

$$W = \frac{4\pi r^2 E^2}{\eta} \quad (6)$$

where E is specified in units of Volts per meter at the distance r from the antenna.

It turns out that the constants in equations (5) and (6) combine into a simple number.  $377/4\pi = 30.00$ , so it may be easier to remember and use these equations in the following forms

$$E = \sqrt{\frac{30W}{r^2}} \quad (5a)$$

$$W = \frac{r^2 E^2}{30} \quad (6a)$$

EXAMPLE 1: Calculate the isotropic field strength of a 5 Watt radiator at a distance of 5 meters.

Using (5), these conditions calculate to

$$E = \sqrt{[(5 * 377) / (4\pi * 5^2)]} = 2.45 \text{ V/m.}$$

EXAMPLE 2: How much isotropic radiated power is allowed for a field strength

specification of 50 mV/m at a distance of 3 meters?

This requires using (6)

$$W = (4\pi * 3^2 * .050^2) / 377 = 0.00075 \text{ W} \\ = -1.2 \text{ dBm.}$$

## REAL ANTENNAS - GAIN AND MATCHING EFFECTS

Isotropic antennas are convenient mathematically, as we have seen. However, a true isotropic antenna has yet to be built. Any real antenna has some gain, and this gain will make the field strength larger in some direction compared to the isotropic. Units of this antenna gain are dBi, which means the gain is quoted in dB related to the isotropic radiator. Even though a truly isotropic radiator is not built, it is still a useful reference for performance comparison and measurements.

By definition, antenna gain is specified in the direction of maximum radiated field. This direction is often referred to as the boresight. Though the gain is specified in dB, the field equations are linear so the gain must be converted. Since we are working with power,

$$g_A = 10^{G_A/10} \quad (7)$$

where:  $G_A$  is the antenna gain specified in dBi, and  $g_A$  is the linear antenna gain factor.

This gain makes the transmitted power appear in the boresight direction to be larger than it really is. This can be incorporated into equations (5) and (6) as

$$E = \sqrt{\frac{g_A W \eta}{4\pi r^2}} \quad (8)$$

$$W = \frac{4\pi r^2 E^2}{g_A \eta} \quad (9)$$

These equations are now valid only at boresight. Since the FCC regulations govern the maximum radiated field, boresight is the direction of interest and equations (8) and (9) are appropriate.



Matching of the antenna feed to the antenna affects the actual power radiated. A poor match reflects some of the applied power back to the transmitter. This makes the antenna gain appear to be less than it actually is. Antenna match is a difficult measurement to make, unless sophisticated equipment is available. For purposes of this paper the measurement methods are defined so that the gains measured will include effects of feed matching.

If the precision of matching effects is required, and the equipment is available to measure return loss of both the transmitter final amplifier ( $s_{22}$ ) and antenna ( $s_{11}$ ) to determine the quality of the match, then these equations can be readily expanded to include the power transfer efficiency as separate from antenna gain. Another factor is required,  $m_A$ , which is the linear power transfer:  $0 \leq m_A \leq 1$ . This antenna matching factor must be a linear conversion from the actual matching measurement, whether it is return loss, VSWR, etc.

The expanded equations then become

$$E = \sqrt{\frac{g_A m_A W \eta}{4\pi r^2}} \quad (10)$$

$$W = \frac{4\pi r^2 E^2}{g_A m_A \eta} \quad (11)$$

#### CALIBRATION OF ANTENNA ISOTROPIC GAIN

To accurately measure antenna gains, it is essential that fields radiated from the antennas of interest are all that is measured. This is not as easy as it sounds. In order to achieve reasonable accuracies without extensive and expensive equipment, the following baseline assumptions are established for this method:

- 1) The operating wavelength is small compared to the size of the test apparatus,
- 2) RF power is the only available measurement, and
- 3) a matched pair of antennas are available for calibration as references.

It should also be noted that the accuracy of this calibration method improves with increasing antenna gain. This is due primarily to the reduced sensitivity of high gain antennas to reflections from the test area.

The basic idea for this method is to determine the performance expected from ideal isotropic radiators, and compare it to the measured performance of the actual radiators. All differences will be accounted as gain in the antennas. These

measurements will be converted into dBi performance calibrations for the reference antennas.

The starting point is the isotropic path loss of RF radiation over a distance. Path loss ( $L$ ) is well known as

$$L = \frac{\lambda}{4\pi r} \quad (12)$$

where:  $\lambda$  is the wavelength of the test signal, and  
 $r$  is the distance of the measurement.

The test wavelength must be small so that ground effects, range reflections, and other error sources are minimized. Practical frequencies are therefore at UHF and higher.

The test distance  $r$  separating the antennas should be greater than five wavelengths, and preferably ten wavelengths or more. This will assure that far field measurements are being taken from the antennas.

To calibrate the reference antennas, mount them facing each other separated by the reference distance  $r$ . Carefully align them along their boresight directions so that the maximum signal strengths will be radiated by the source antenna and measured by the receiving antenna. Connect a source of known RF power to the source antenna. Measure the power received by the receiving antenna.

A simple form of the radio link equation is used to determine the reference antenna gains in dBi:

$$P_R = G_T + G_R + P_T + L_p \quad (13)$$

where:  $P_R$  is the measured receive power in dBm,  
 $P_T$  is the known transmit power in dBm,  
 $L_p$  is the path loss in dB =  $20\log(L)$ , and  
 $G_T$  and  $G_R$  are the dBi gains of the transmit and receive antennas respectively.

When solved for the antenna gains we get

$$G_T + G_R = (P_R - P_T) - L_p \quad (14)$$

Notice that the gains are not yet independently separable. However since identical antennas were used in the calibration we can split the measured path gain equally between both antennas

$$G_T = G_R = G_A \quad (15)$$

so that

$$G_A = \frac{(P_R - P_T) - L_P}{2} \quad \text{dBi} \quad (16)$$

EXAMPLE 3: Calibrate the dBi gain of a pair of antennas at 915 MHz.

First check: the test wavelength is  $300/915 = 0.328\text{m}$ . This qualifies as small.

A standard test distance for UHF and higher frequencies is 3m. At this test frequency this corresponds to just over 9 wavelengths, so far field measurements will be made at this test distance.

Path Loss: The isotropic path loss is found from (12) to be

$$L = 0.328/(4\pi \cdot 3) = 0.0087$$

$$L_P = 20\log(0.0087) = -41 \text{ dB}$$

Transmit power: It is always a good idea to use small source powers to minimize any possible interference. -20dBm is often a good choice for these small test distances.

Measure receive power: This is usually done with a spectrum analyzer. For this example the received power is found to be -45 dBm.

Reference antenna boresight gain can now be found from (16):

$$G_A = [(-45 - (-20)) - (-41)]/2$$

$$= 8 \text{ dBi}$$

## MEASURING FIELD STRENGTH

The first step in measuring the actual field strength of any system under test is to convert the regulatory limit to an equivalent isotropic radiator power at the specified distance. This is done with the technique given in example 2. For the purposes of this discussion the results of that example will be used here: 50mV/m at 3m corresponds to a 0.75 mW isotropic radiator.

We now need to determine what received power from our test setup corresponds to this field strength at the receive antenna. This is done using equation (13). Remember that the equivalent transmitter power was defined from a perfectly isotropic radiator,  $G_T = 0\text{dBi}$ . At the test distance of 3m and using 915 MHz as the test frequency, Example 3 showed the isotropic path loss to be -41dB. Knowing that the reference antenna used at the receiver has a gain of 8dBi we can now find the maximum allowed receive power from our system under test:

$$P_R = G_T + G_R + P_T + L_P$$

$$= 0 + 8 + (-1.2) + (-41)$$

$$= -34 \text{ dBm}$$

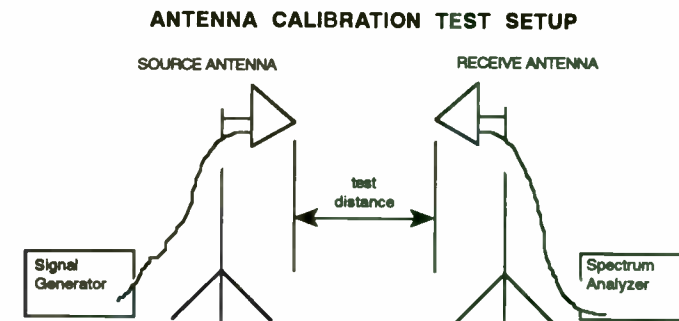
This tells us that whatever the combination of transmitter power and antenna radiating efficiency used by the system under test, at a distance of 3m with **this test system** the received power must never exceed -34dBm. This will assure compliance with the regulations and clear the way for FCC approval of the product.

## CONCLUSION

A simple method has been presented to relate the field strength specifications from FCC regulations to more conventional RF power units. From this relationship two important results have been derived

- 1) how to measure the isotropic (dBi) gain of any antenna, and
- 2) how to use these calibrated antennas to measure field strength from a particular unit or system under test.

These field strength measurements are accurate to within a few dB if care is taken in the calibration of the reference antennas. This accuracy is sufficient to demonstrate basic compliance with the FCC regulations. This accuracy is NOT sufficient to get compliance certification from the FCC. For actual FCC certification there is no substitute for a calibrated laboratory.



Choosing the Optimum  
Synthesizer Architecture  
for Your  
Receiver Application

Larry Kimbrough  
Harris GCSD  
Palm Bay, Fl  
(407) 729-7824



# Synthesizer Requirements - The 6 Basic Design Parameters

- Output Frequency ( $F_{out}$ )
- Step Size at  $F_{out}$  ( $F_{step}$ )
- Output Spurious (dBc)
- Phase Noise
  - Close-in integrated (residual FM or degrees RMS)
  - Far-out (SSB dBc/Hz)
- Settling Time
  - Freq1 to Freq2 switching time
  - Freq1 to fixed phase at Freq2 time
- Size and Power (Complexity)

ARCHITECTURE	ADVANTAGES	DISADVANTAGES	PERFORMANCE
DIRECT ANALOG	<ul style="list-style-type: none"> <li>• LOW PHASE NOISE (REF. MULTIPLIED TO OUTPUT FREQUENCY)</li> <li>• HIGH SWITCHING SPEED</li> <li>• NOT TECHNOLOGY DEPENDENT</li> <li>• MODULAR CONSTRUCTION</li> <li>• FINE FREQUENCY RESOLUTION CAN BE OBTAINED BY REPETITION OF MODULES</li> </ul>	<ul style="list-style-type: none"> <li>• HARDWARE INTENSIVE</li> <li>• HIGH POWER DISSIPATION</li> <li>• LARGE SIZE</li> <li>• PHASE NOISE AND SPURIOUS SUPPRESSION DEGRADE AS OUTPUT FREQUENCY INCREASES</li> </ul>	<ul style="list-style-type: none"> <li>• SPURIOUS PRODUCTS APPROX. -60 TO -70 dBc</li> <li>• SETTLING TIME ~ 1<math>\mu</math>s to 10<math>\mu</math>s</li> </ul>
INDIRECT ANLOG	<ul style="list-style-type: none"> <li>• WIDE TUNING RANGE</li> <li>• LOW SPURIOUS PRODUCTS</li> <li>• SMALL SIZE</li> <li>• CAN BE LOW POWER</li> <li>• LOW COST</li> <li>• CAN HAVE FINE FREQ. RESOLUTION</li> </ul>	<ul style="list-style-type: none"> <li>• SLOW SETTLING TIME</li> <li>• FREQUENCY RESOLUTION VS SETTLING TIME TRADEOFF</li> </ul>	<ul style="list-style-type: none"> <li>• PROGRAMMABLE DIVIDERS LIMITED TO 2 GHz</li> <li>• SETTLING TIME 1ms OR MORE</li> <li>• SPURIOUS PRODUCTS -70 TO -100 dBc</li> </ul>
DIRECT DIGITAL	<ul style="list-style-type: none"> <li>• FAST SETTLING TIME</li> <li>• CONTINOUS PHASE</li> <li>• FINE FREQ. RESOLUTION</li> <li>• MMIC COMPATIBLE</li> <li>• SMALL SIZE</li> <li>• LOW COST</li> <li>• REDUCED SHIELDING REQUIREMENTS</li> <li>• INHERENT PROGRAMMABLY</li> <li>• NO ALIGNMENT EXCEPT FOR D/A CONVERTER AND LOWPASS FILTER</li> </ul>	<ul style="list-style-type: none"> <li>• LIMITED OUTPUT FREQUENCY</li> <li>• LIMITED SPURIOUS REJECTION</li> </ul>	<ul style="list-style-type: none"> <li>• OUTPUT FREQUENCY GENERALLY LESS THAN 10 MHz</li> <li>• SETTLING TIME ~ 0.1<math>\mu</math>s to 20<math>\mu</math>s</li> <li>• FREQUENCY RESOLUTION TO 0.001 Hz</li> <li>• SPURIOUS PRODUCTS -60 TO -70 dBc</li> </ul>
HYBRID	<ul style="list-style-type: none"> <li>• FAST SETTLING</li> <li>• FINE RESOLUTION</li> </ul>	<ul style="list-style-type: none"> <li>• REQUIRES TWO OR MORE SYNTHESIZERS</li> </ul>	<ul style="list-style-type: none"> <li>• SPURIOUS PRODUCTS -60 dBc</li> <li>• SETTLING TIME 1<math>\mu</math>s -100ms</li> <li>• FREQUENCY RESOLUTION 0.1Hz</li> </ul>

Advantages and Disadvantages of Frequency Synthesizer Architectures

## Example of 3 different Synthesizer Requirements for Rx Applications

	Fout	Fstep	Spurious	Phase Noise	Settling Time	Size/Power
Synth A (Voice Com Rx from 20-500 MHz 1st IF = 700 MHz))	720-1300MHz	5 kHz	-40 dBc	300 Hz FM	< 250 msec for F1 to F2	< 1 watt 10 sq in
Synth B (HF Comm Rx)	83-110 MHz	1 kHz	-80 dBc	3 deg RMS 100Hz-3kHz  -90 dBc/Hz SSB at 10 kHz	< 5 msec for F1 to F2	<5 watts 50 sq in
Synth C (Freq Hop Rx for Data)	100-200 MHz	10 Hz	-60 dBc	2 deg RMS 10Hz-10kHz  -75 dBc/Hz SSB at 100 kHz	<50 usec to within 45 deg of final phase	< 10 watts 1/4 rack

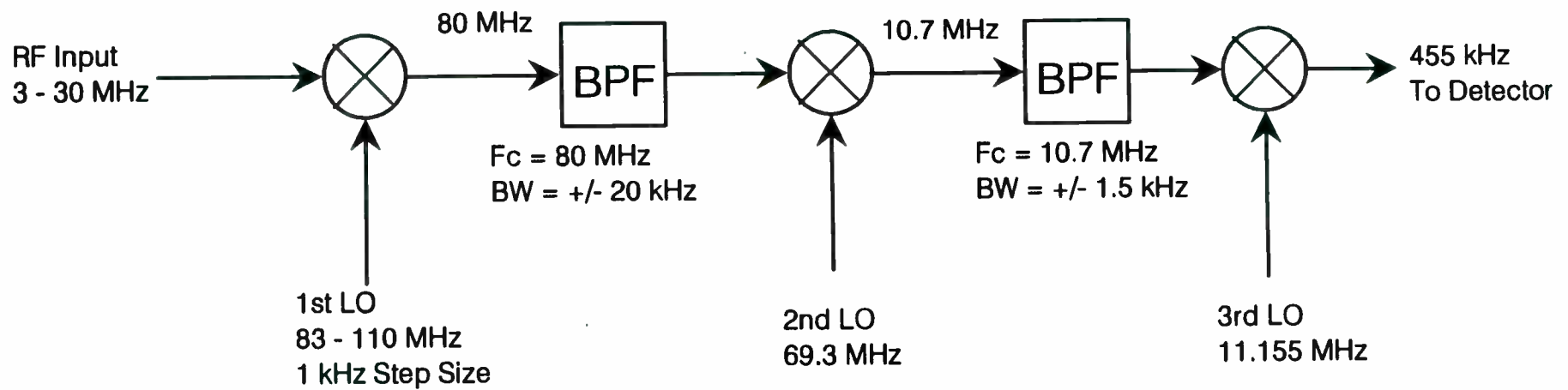


Figure 1 - HF Receiver Frequency Plan



# HF Receiver Requirements

for 1st LO

- $F_{out} = 83 - 110 \text{ MHz}$
- $F_{step} = 1 \text{ kHz}$
- Phase Noise:
  - 100 Hz-3 kHz,  $< 3 \text{ deg}$
  - 10 kHz offset,  $-90 \text{ dBc/Hz SSB}$
- Spurious  $< -80 \text{ dBc}$
- Settling Time  $< 5 \text{ msec}$  from F1 to F2
- Size/Power = TBD

## Rules of Thumb for Determining Loop Performance

### o Phase Noise

#### 1. Outside LoopBW use VCO noise

- o VCO Noise determined from Leeson's model

Given:  $F_{out}$ , NF, Loaded Q of Resonator, and noise corner of device - Determine  $L(f)$ .

#### 2. Inside LoopBW use Cat 1 Noise increased by $20 \log N_{max}$

- o Cat 1 noise is Logic noise and Fref noise and PD noise.

#### 3. Other noise sources will appear at $F_{out}$ modified by transfer functions as shown in Figure 4.

### o Spurious

#### 1. Fref spurs = $20 \log((V_m * K_{vco}) / (2 * f_m))$

$V_m$  = PD Vout component \*  $F(s)$  \* LPF(s)

$K_{vco}$  = VCO tuning constant (Hz/volt)

$f_m$  = offset frequency from VCO carrier

#### 2. With a mixer in the Loop, The mixer spurs that are input into the programmable divider will appear at $F_{out}$ at the same level (dBc), modified by transfer function of Figure 4.

### o Settling Time

F1 to F2 settling =  $10 * 1 / \text{LoopBW}(\text{rad})$ , (conservative)

### o Phase Margin should be > 45 degrees

### o Loop Bandwidth should be as high as possible to reduce microphonics. The Loop will track off all microphonics that fall below the LoopBW frequency.

Loop Bandwidth should be low enough to allow adequate low pass filtering to be added for Fref attenuation.

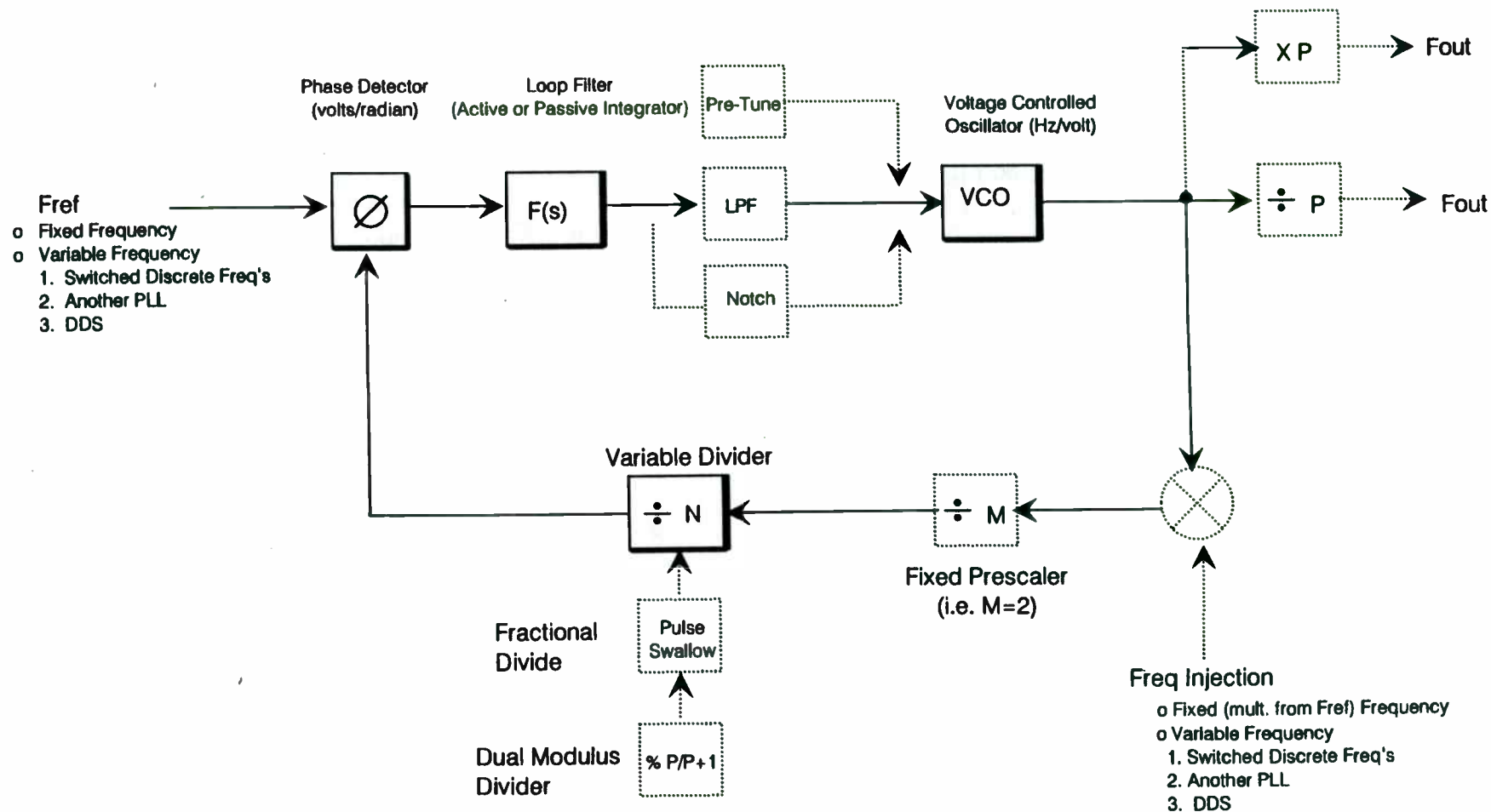
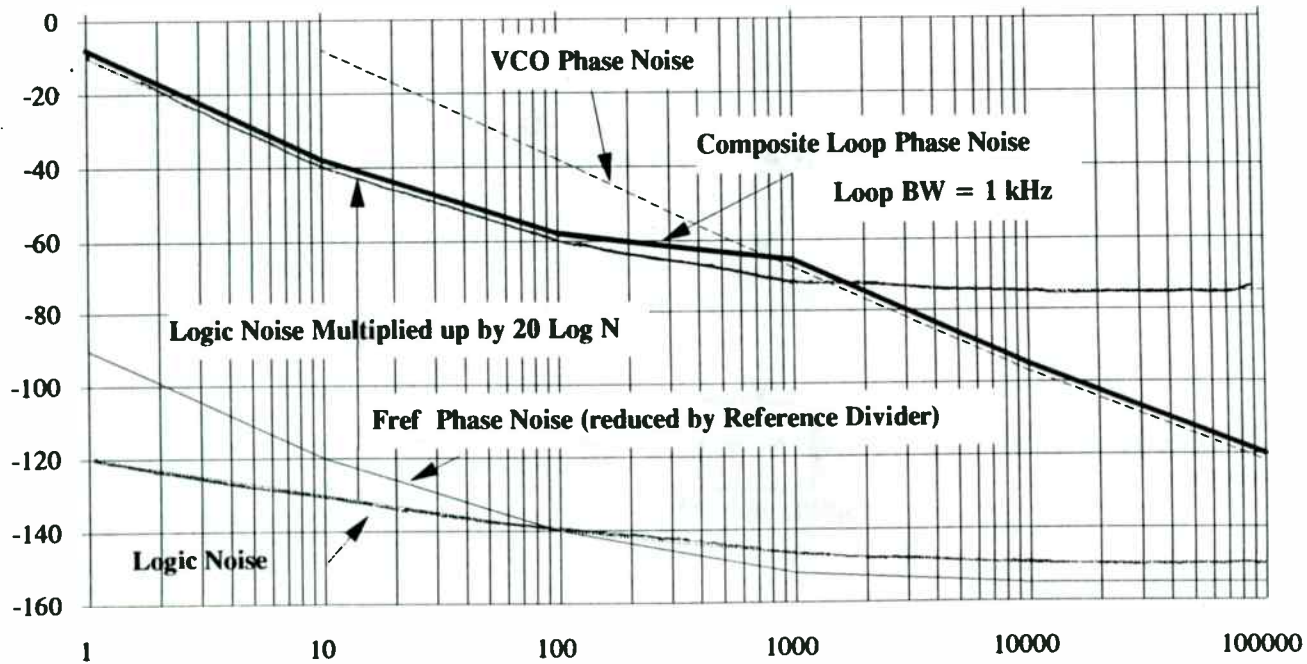


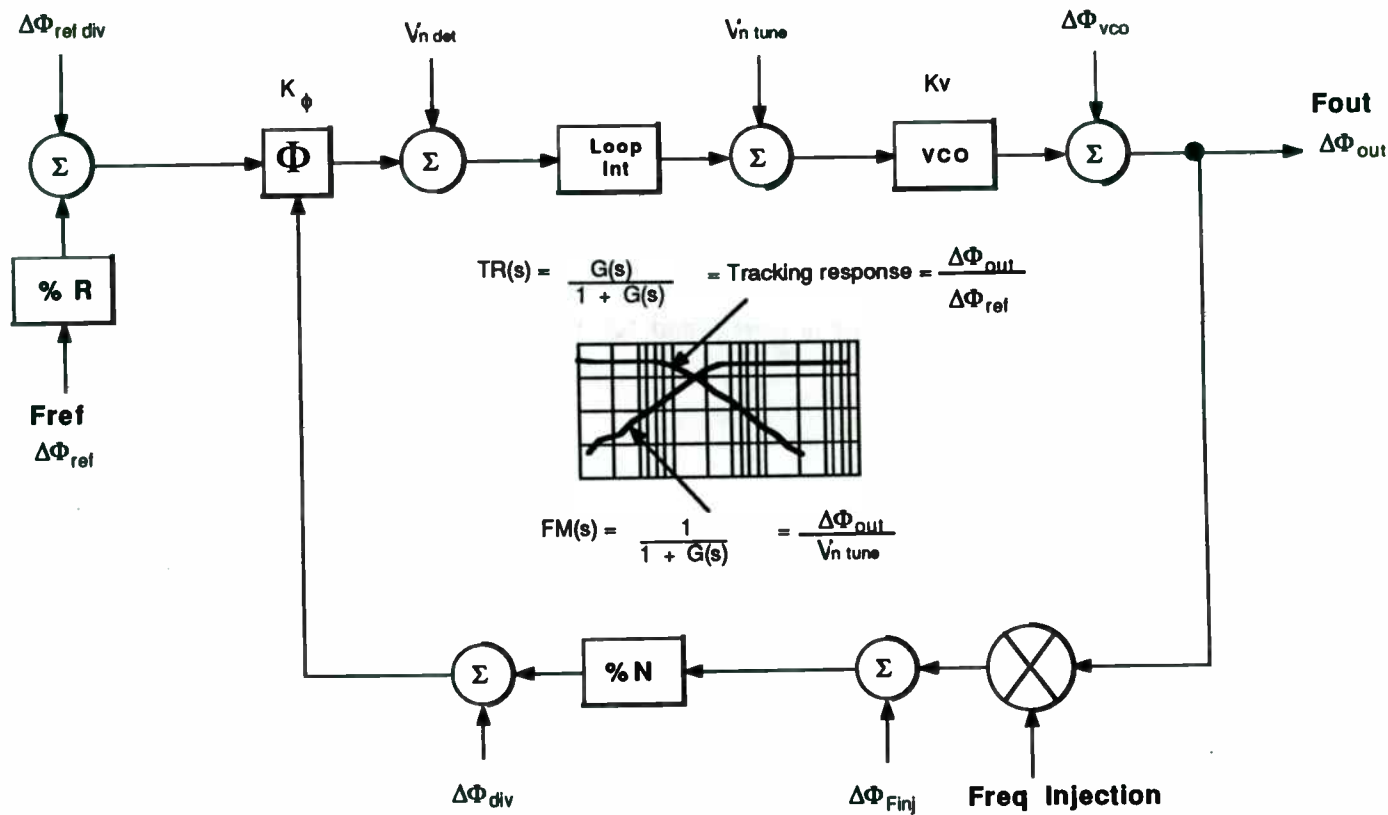
Figure 2 - Basic Loop with Optional (dashed lines) Functions, that will change Synthesizer Architecture and Output Performance

### Selection of Loop BW for Optimum Noise Response



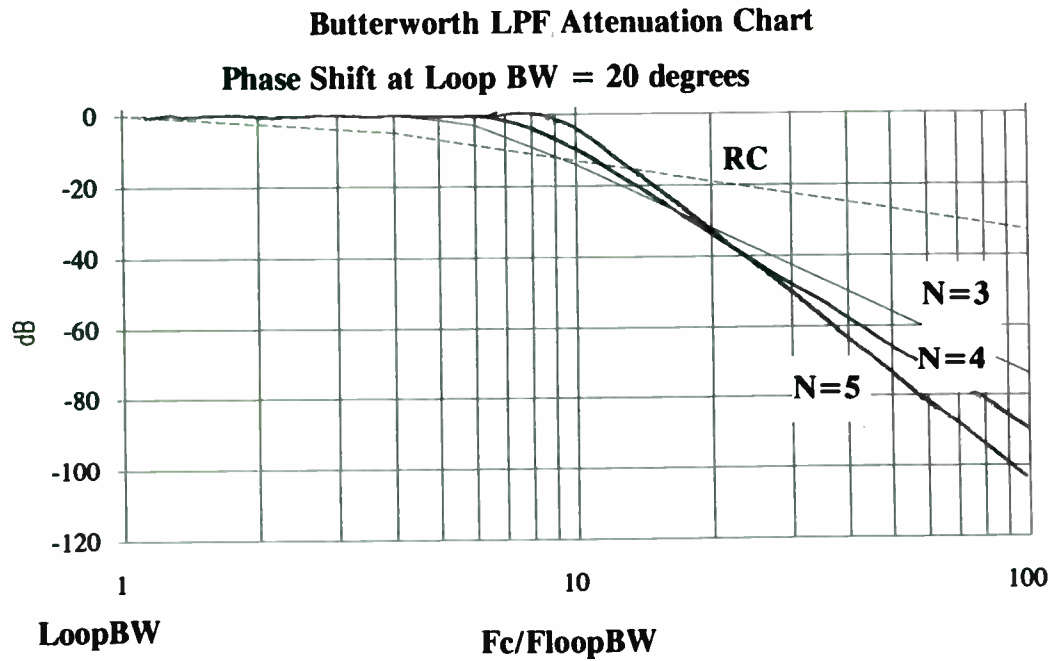
- o Loop BW is set to the frequency at which the crossover between VCO noise and Multiplied up Category 1 Noise intersects.
- o Category 1 Noise is the sum of all noise sources that is degraded by  $20 \log N$ , at the output. This includes Fref Phase Noise, Logic Phase noise, and Phase Detector Noise.

Figure 3 - Setting the Loop Bandwidth for Optimum Phase Noise Response at the Output



$$\Delta\Phi_{out} = TR(s) \left[ \frac{N^2}{R^2} \Delta\Phi_{ref}^2 + N^2 \Delta\Phi_{ref\ div}^2 + N^2 \Delta\Phi_{div}^2 + \frac{N^2}{K_\phi^2} (V_{n\ det}^2) + \Delta\Phi_{Finj}^2 \right] \\
 + FM(s) \left[ \frac{K_v^2}{s} V_{n\ tune}^2 + \Delta\Phi_{vco}^2 \right]$$

Figure 4 - Contribution of Noise Sources in a PLL to the Output Phase Noise



- o All filters are normalized so that the phase shift at FloopBW = 20 degrees
- o If the Fref = 1 kHz and FloopBW = 100 Hz, then  $F_c/\text{FloopBW} = 10$
- o For a N=3 LPF the attenuation of the Fref component would be -18 dB

Figure 5 - Low Pass Filter Attenuation for Phase Shift = 20 degrees at Loop BW

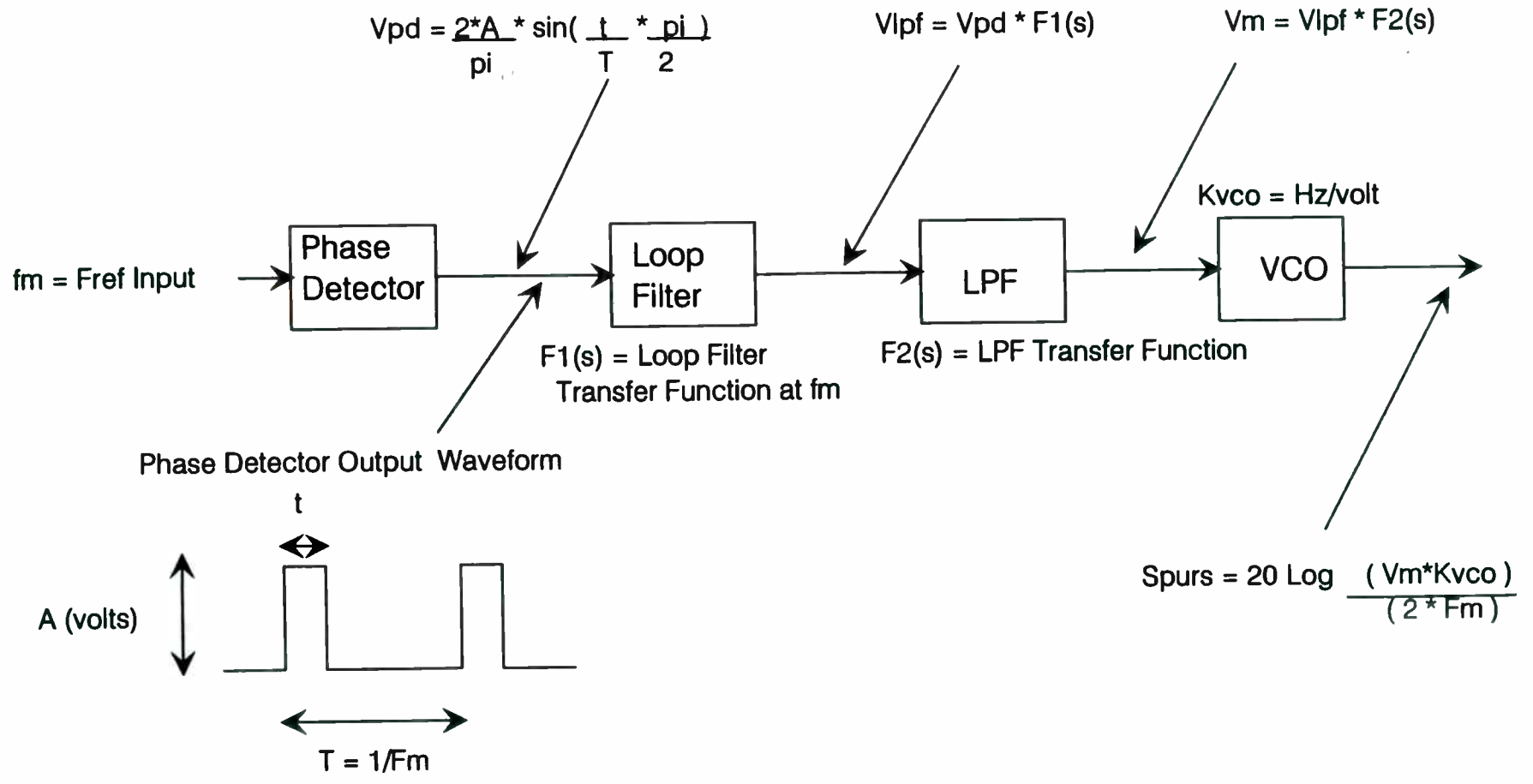


Figure 6 - Spurs caused by Fref leakage in a PLL

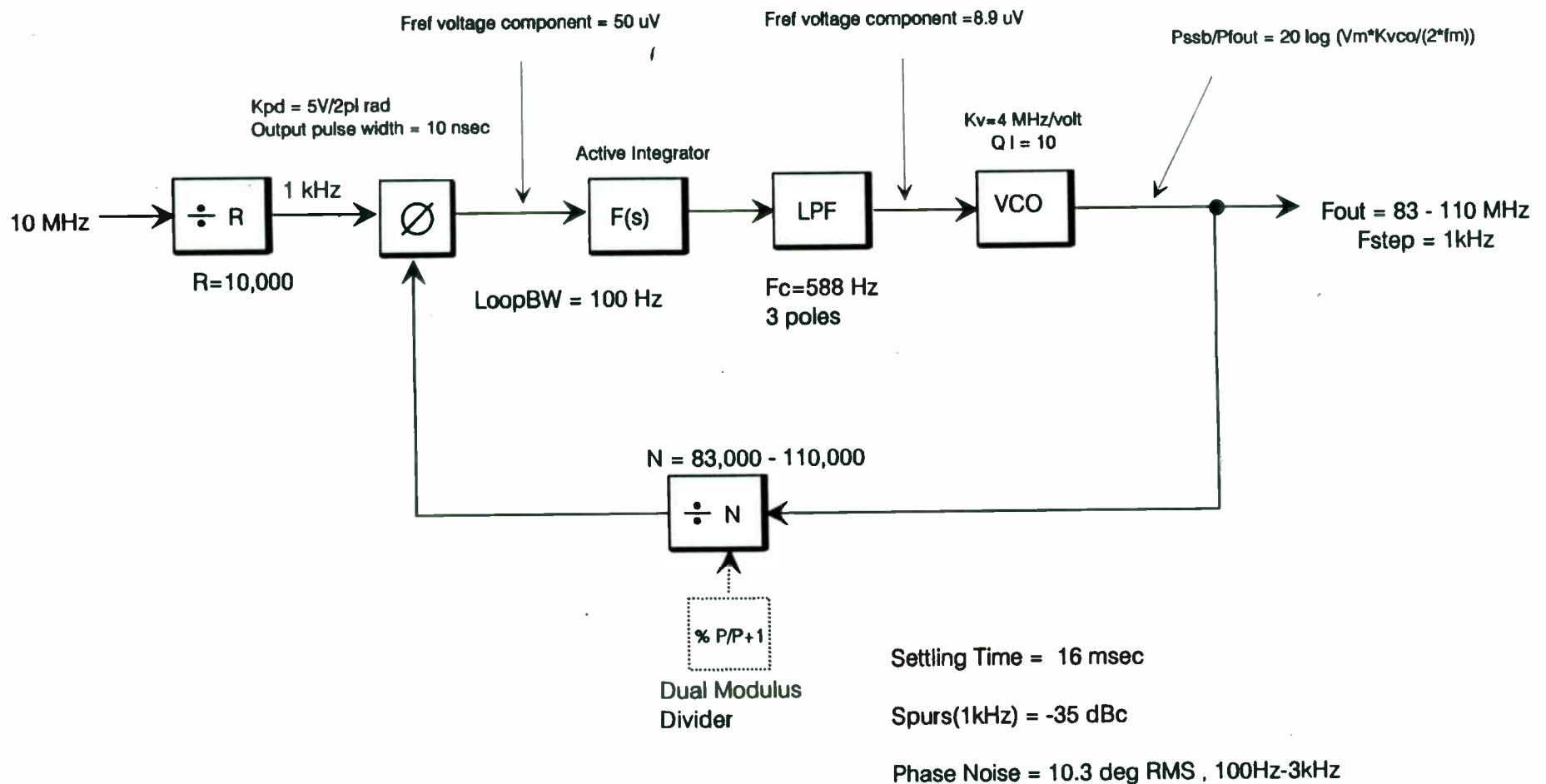
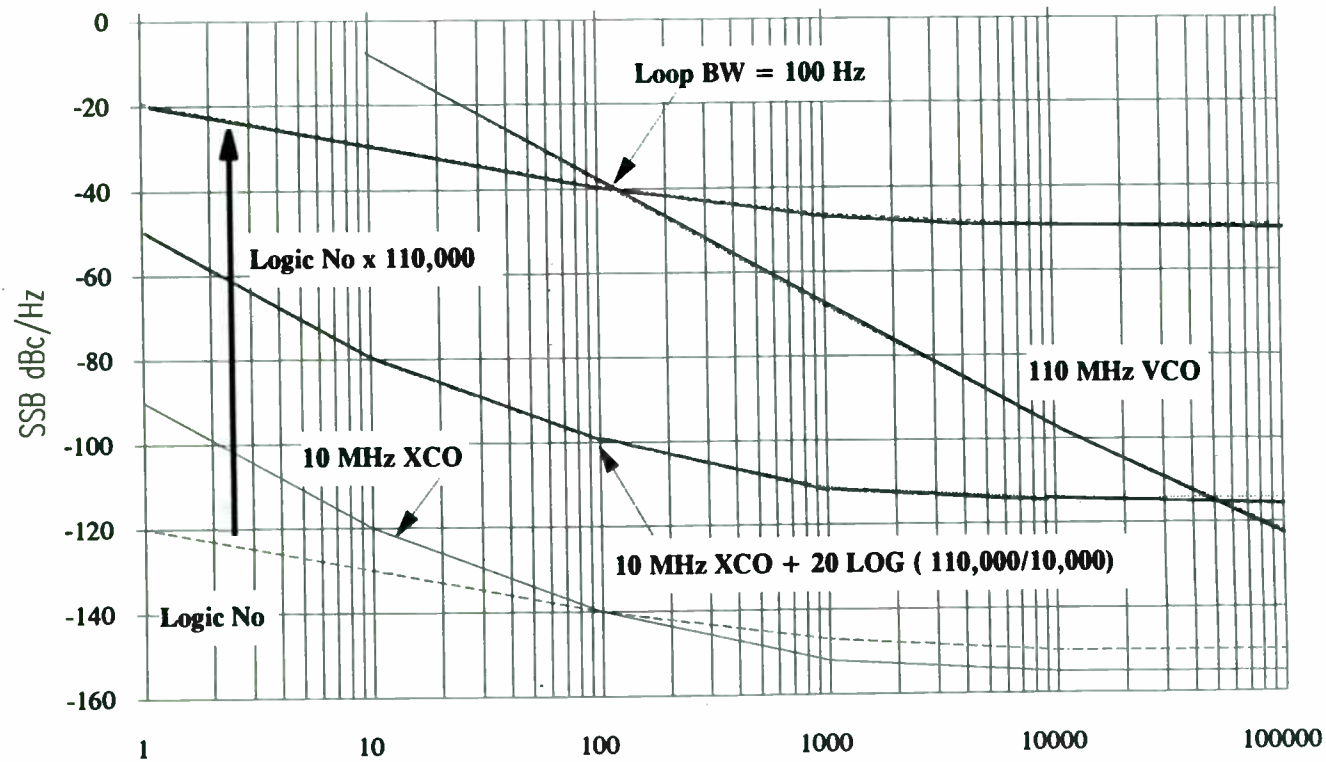


Figure 7 - Single Loop Synthesizer for 1st LO of HF Receiver





- o Loop BW is set to the frequency at which the crossover between VCO noise and Multiplied up Logic Noise
- o  $F_{out\ max} = 110\ \text{MHz}$  ,  $N = 110,000$ , Logic Noise multiplied by 100.8 dB at  $F_{out}$
- o  $F_{ref} = 1\ \text{kHz}$ , Loop BW = 100 Hz

Figure 8 - Output Phase Noise of a Single Loop Synthesizer

# Analysis for Single Loop Synthesizer

- Phase Noise

- 10 MHz noise is reduced by  $20 \text{ Log}(10,000)$  making it fall below the Logic noise
- Inside LoopBW : Cat 1 noise +  $20 \text{ Log } N_{\text{max}}$
- Logic Noise = -150 dBc/Hz with 1/f corner at 1 kHz
- Total Noise = Logic noise + 100.8 dB
- Outside LoopBW : VCO with  $Q_l = 10$
- Integrated Noise 100Hz-3kHz = 10.3 deg RMS
- SSB phase noise at 10 kHz = -90 dBc/Hz

- Spurs

- 1 kHz spurs =  $20 \text{ Log}(V_m \cdot K_{\text{vco}})/(2 \cdot 10^3)$
- $= 20 \text{ Log}((8.9 \cdot 10^{-6} \cdot 4 \cdot 10^6)/(2 \cdot 10^3)) = -35 \text{ dBc}$

- Settling Time =  $10 \cdot 1/\text{LoopBW}(\text{rad})$

- Time =  $10 \cdot 1/(2\pi \cdot 100) = 16 \text{ msec}$

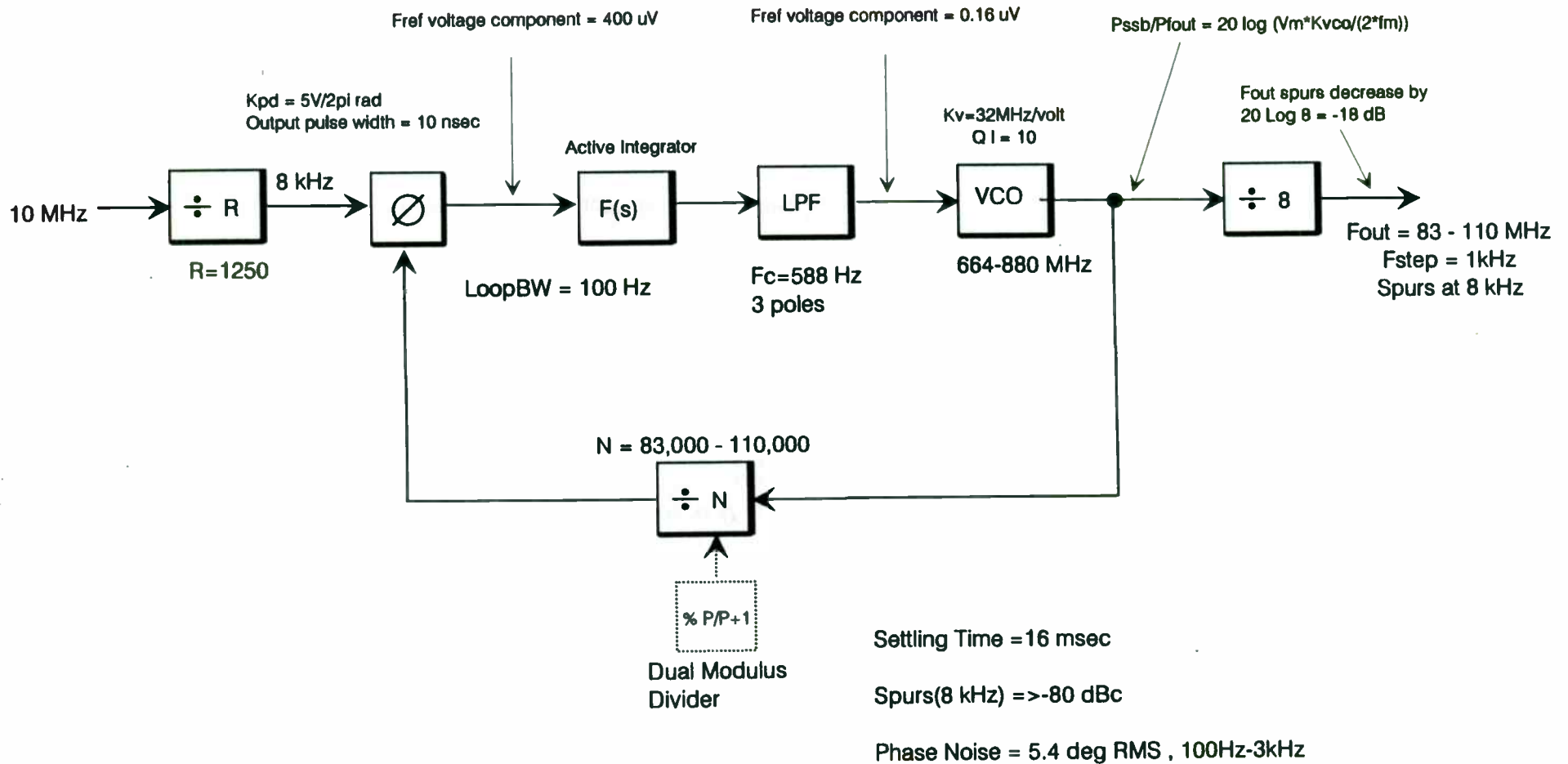
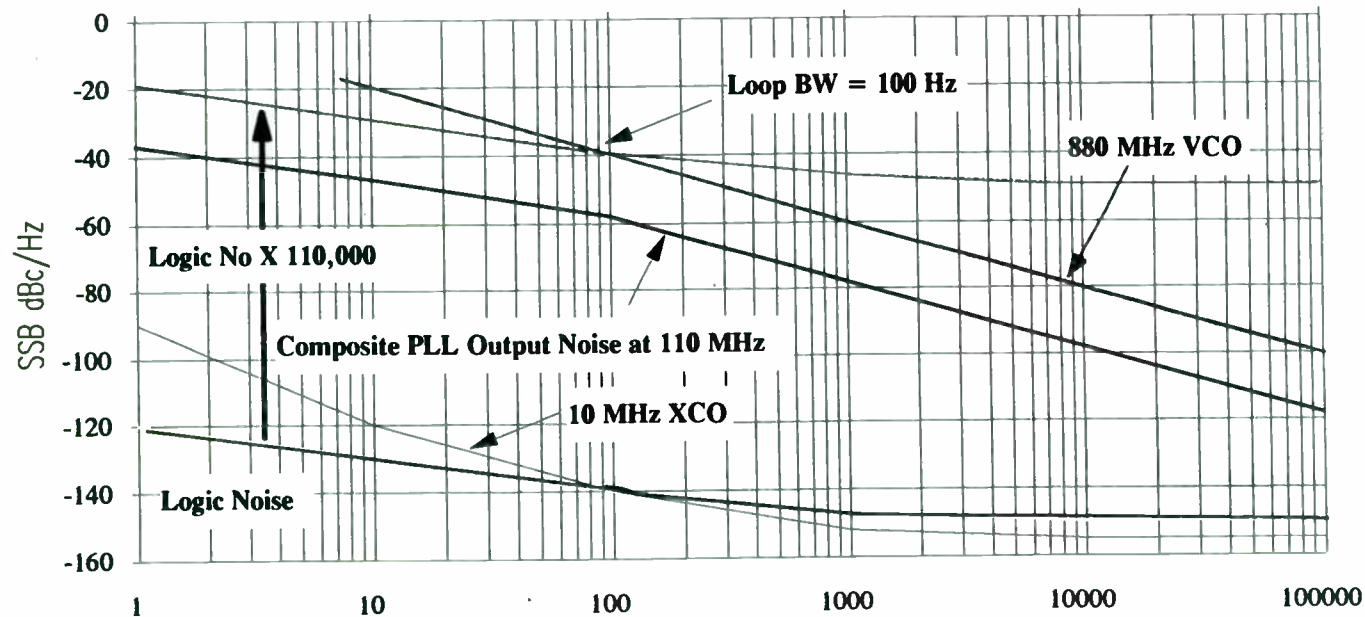


Figure 9 - Single Loop Synthesizer with Output Divider for 1st LO of HF Receiver



- o Divide by 8 reduces output phase noise and spurs by 18 dB
- o  $F_{out}$  of VCO = 880 MHz,  $F_{out}$  of Loop = 110 MHz, Noise multiplied by 82.8 dB at  $F_{out}$
- o  $F_{ref}$  = 8 kHz,  $F_{step}$  = 8 kHz / 8 = 1 kHz, Spurs at 110 MHz  $F_{out}$  are at 8 kHz rate

Figure 10 - Output Phase Noise of a Single Loop Synthesizer with Divide by 8 on Output

## Analysis for Single Loop Synthesizer with Divider on Output

### ● Phase Noise

● 10 MHz noise reduced by  $20 \log(125)$ , making it lower than Logic noise

● Inside LoopBW : Cat 1 noise +  $20 \log N_{max}$  -  $20 \log(8)$

● Logic Noise = -150 dBc/Hz with  $1/f$  corner at 1 kHz

● Total Noise = Logic noise + 100.8 dB - 18 dB

● Outside LoopBW : VCO with  $Q_l = 10$

● Integrated Noise 100Hz-3kHz = 5.4 deg RMS

● SSB phase noise at 10 kHz = -90 dBc/Hz

### ● Spurs - Reduced by 3 pole LPF at 588 Hz and by divide by 8 on Output

● 8 kHz spurs =  $20 \log((v_m * K_{vco}) / (2 * 1e3))$  -  $20 \log(8)$

● =  $20 \log((0.16e-6 * 32e6) / (2 * 8e3))$  - 18 dB = >-80 dBc

### ● Settling Time = $10 * 1 / \text{LoopBW}(\text{rad})$

● Time =  $10 * 1 / (2\pi * 100) = 16 \text{ msec}$

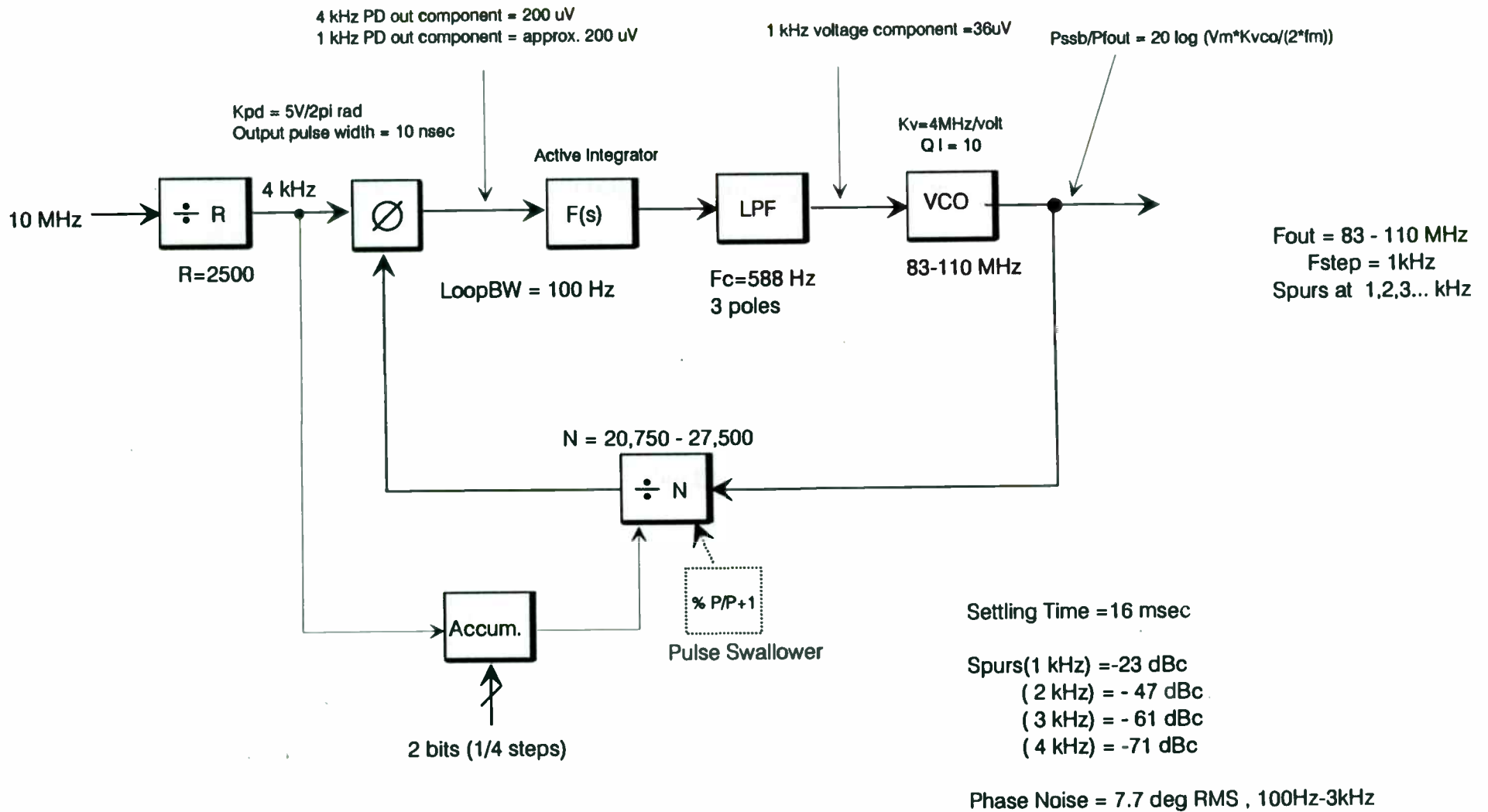
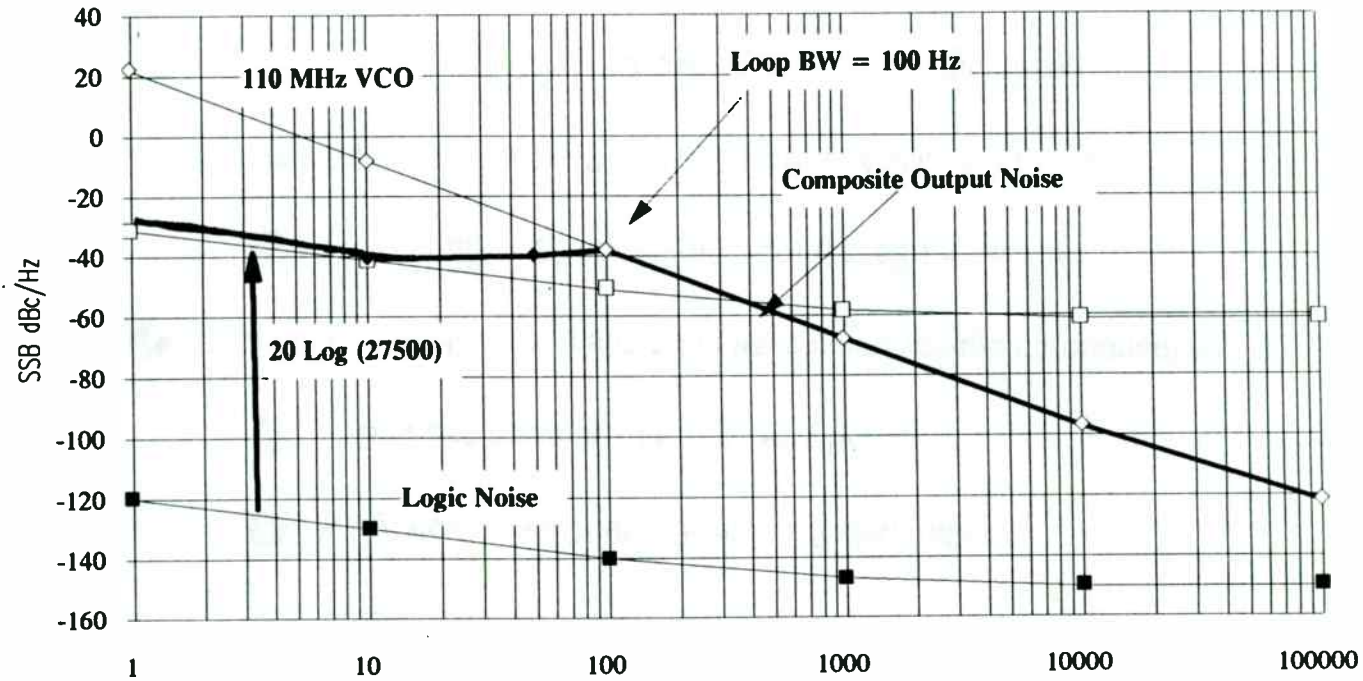


Figure 11 - Single Loop Fractional N Synthesizer for 1st LO of HF Receiver



- o Loop BW = 100 Hz
- o  $F_{out} = 110 \text{ MHz}$ ,  $F_{ref} = 4 \text{ kHz}$ , Noise multiplied by 88.8 dB at  $F_{out}$
- o  $F_{ref} = 4 \text{ kHz}$ ,  $F_{step} = 4\text{kHz} / 4 = 1 \text{ kHz}$ , Spurs at 1 kHz, 2 kHz ... rate

Figure 12 - Output Phase Noise of a Fractional N Loop Synthesizer

# Analysis for Fractional N Single Loop Synthesizer

- Phase Noise
  - Inside LoopBW : Cat 1 noise + 20 Log Nmax
  - Logic Noise = -150 dBc/Hz with 1/f corner at 1 kHz
  - Total Noise = Logic noise + 88.8 dB
  - Outside LoopBW : VCO with QI = 10
  - Integrated Noise 100Hz-3kHz = 7.7 deg RMS
  - SSB phase noise at 10 kHz = -90 dBc/Hz
- Spurs - Fractional spurs assumed to have same voltage component as Fref
  - 1 kHz spurs =  $20 \text{ Log} ( (v_m * K_{vco}) / (2 * 1e3) ) = -23 \text{ dBc}$
  - 2 kHz = -47 dBc, 3 kHz = -61 dBc, 4 kHz = -71 dBc
  - LPF reduces spurs as the spur freq offset increases
- Settling Time =  $10 * 1 / \text{LoopBW}(\text{rad})$ 
  - Time =  $10 * 1 / (2\pi * 100) = 16 \text{ msec}$



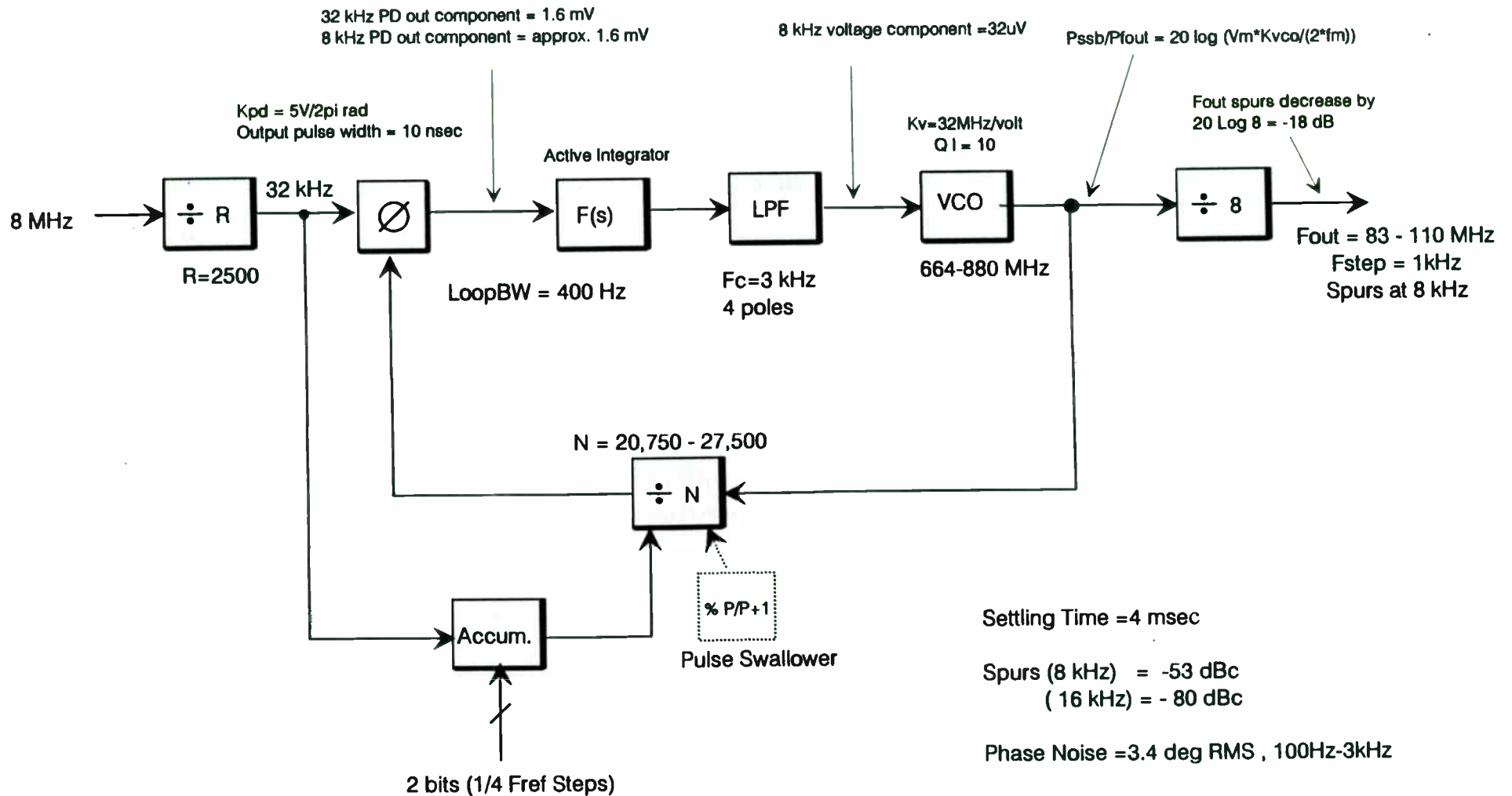
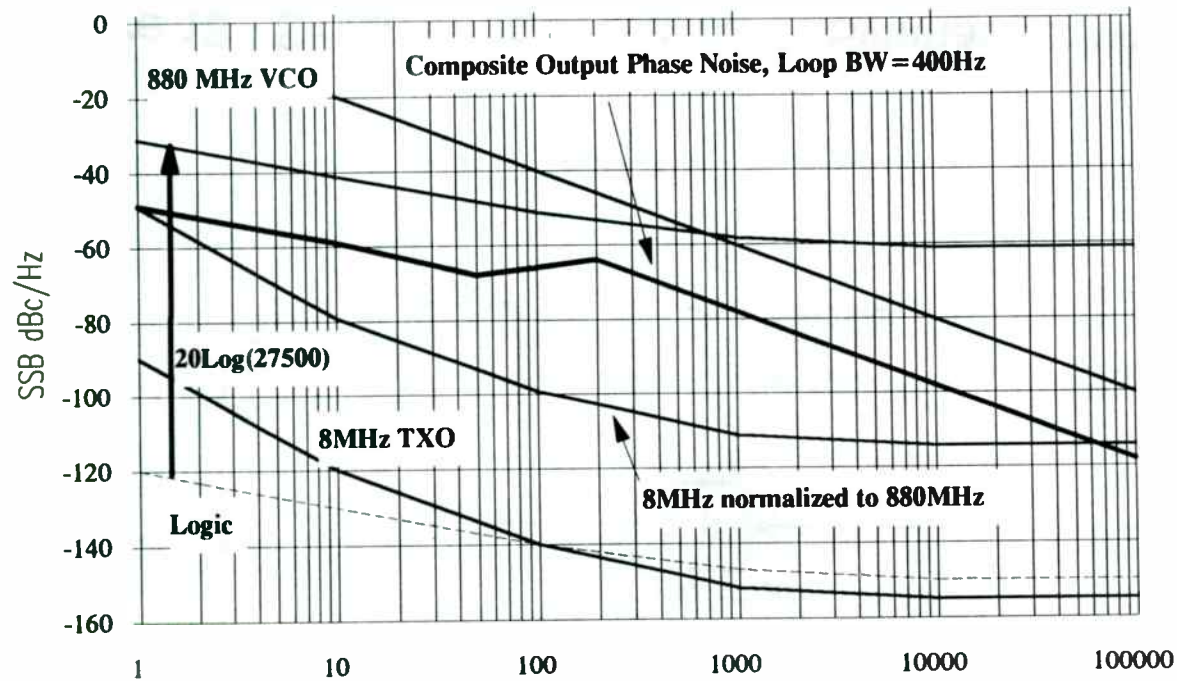


Figure 13 - Single Loop Fractional N Synthesizer with Output Divider for 1st LO of HF Receiver



- o Loop BW = 400 Hz
- o Output Phase Noise and Spurs are reduced by 18 dB,
- o  $F_{out} = 880 \text{ MHz} / 8 = 110 \text{ MHz}$ ,  $F_{ref} = 32 \text{ kHz}$ , Fractional Steps of  $1/4 F_{ref}$   
 $F_{step} \text{ at output} = 8 \text{ kHz} / 8 = 1 \text{ kHz}$ , Spurs at 8 kHz, 16 kHz ...

Figure 14 - Output Phase Noise of a Fractional N Loop with Divide by 8 on Output

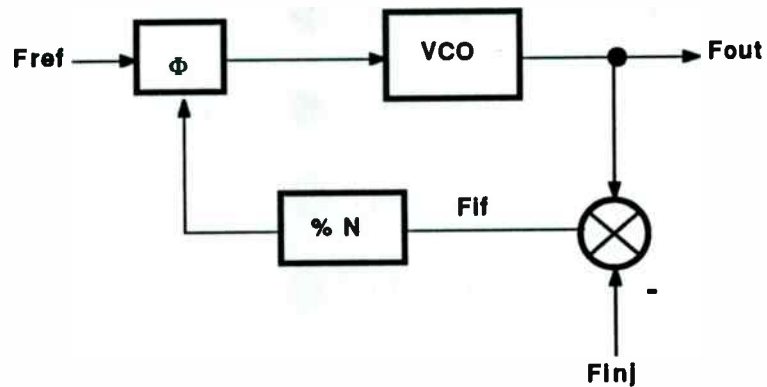
# Analysis for Single Loop Fractional N Synthesizer

## with Divider on Output

- Phase Noise
  - Inside LoopBW : Cat 1 noise + 20 Log Nmax-20 Log (8)
  - Logic Noise = -150 dBc/Hz with 1/f corner at 1 kHz
  - Total Noise = Logic noise + 100.8 dB - 18 dB
  - Outside LoopBW : VCO with Ql = 10
  - Integrated Noise 100Hz-3kHz = 3.4 deg RMS
  - SSB phase noise at 10 kHz = -90 dBc/Hz
- Spurs - Fractional spurs assumed to have same voltage component as Fref Spur
  - Fractional Spurs at 8 kHz = -35 dBc - 20 Log 8 = -53 dBc, 16 kHz = - 80 dBc ( due to additional LPF attenuation)
- Settling Time =  $10 * 1/LoopBW(rad)$ 
  - Time =  $10 * 1/(2\pi * 400) = 4 \text{ msec}$

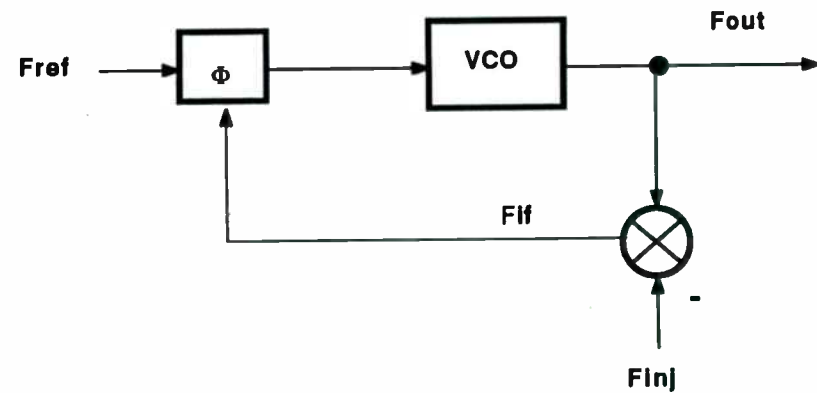
## 2 Options of Output Loop for a Multi-Loop Synthesizer

Programmable Divider in Output Loop



- o  $F_{out} = F_{inj} + F_{ref} (N)$
- o Fine steps are done in  $F_{inj}$  ( $F_{inj}$  could be composed of many loop)
- o Typically the BW of  $F_{inj}$  is  $\ll$   $F_{out}$  BW,  $F_{inj}$  is close to  $F_{out}$
- o Typically  $N$  is a low value
- o Typically the BW of  $F_{out}$  is small % BW

Summing Loop



- o  $F_{out} = F_{inj} + F_{ref}$
- o  $F_{ref}$  will move in fine step sizes  
 $F_{inj}$  will move in coarse step sizes
- o The BW of  $F_{inj}$  must be equal to  $F_{out}$  BW
- o Typically used for large  $F_{out}$  % BW

Figure 14a.

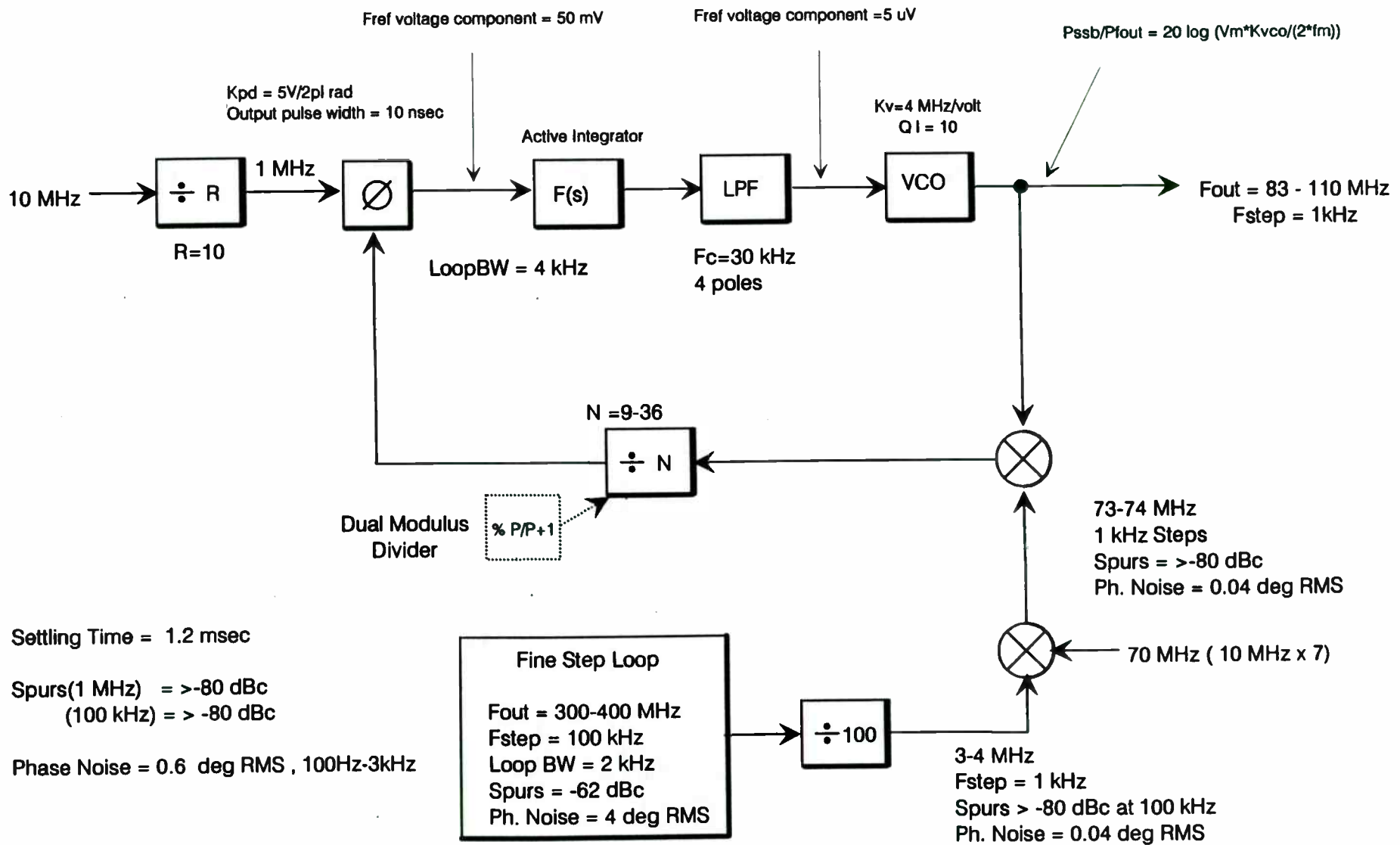
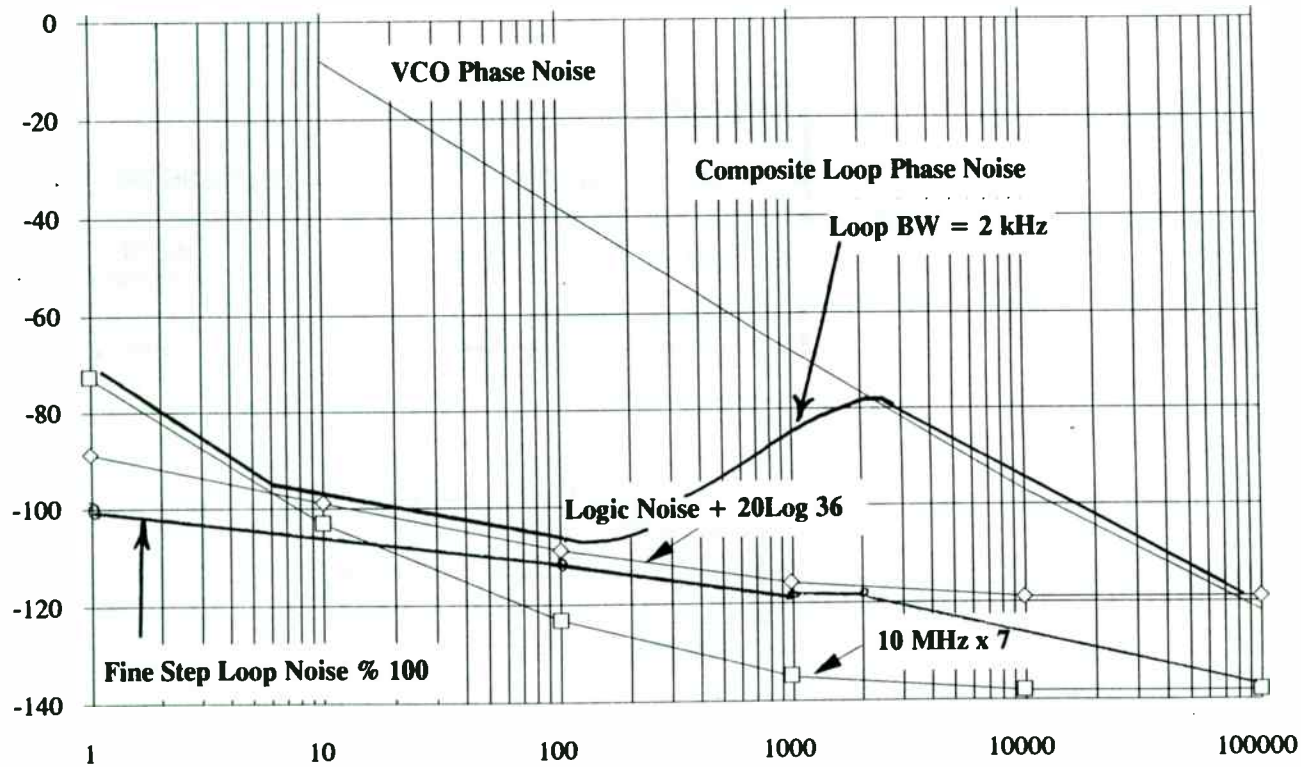


Figure 15 - Dual Loop Synthesizer for 1st LO of HF Receiver



- o Fine Step Loop = 300-400 MHz Loop divided by 100 and summed with 70 MHz source
- o The contribution of the Fine Step Loop Noise and 70 MHz noise is negligible
- o The Output Loop's Loop BW = 2 kHz. Logic Noise of Output Loop is multiplied up by 31.1 dB

Figure 16 - Output Phase Noise of Dual Loop Synthesizer

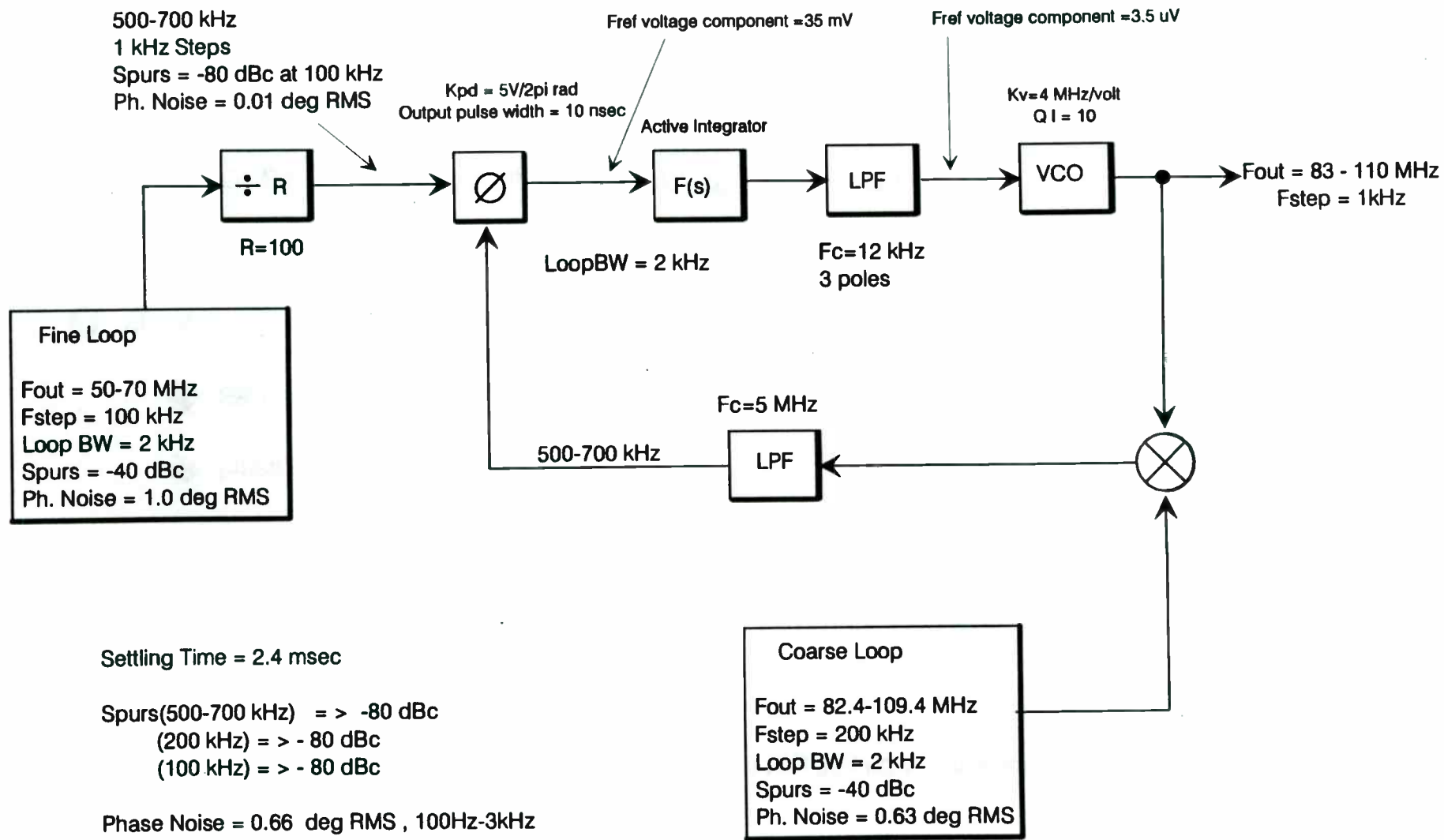
## Analysis for Dual Loop Synthesizer

### ● Phase Noise

- Inside LoopBW : Comprised of Cat 1 noise + 20 Log 36 and contribution of Fine Step Loop and 70 MHz
- Logic Noise = -150 dBc/Hz with 1/f corner at 1 kHz
- Total Noise = Logic noise + 31.1 dB + 70 MHz noise + Fine Loop Noise
- Outside LoopBW : VCO with QI = 10
- Integrated Noise 100Hz-3kHz = 0.6 deg RMS
- SSB phase noise at 10 kHz = -90 dBc/Hz

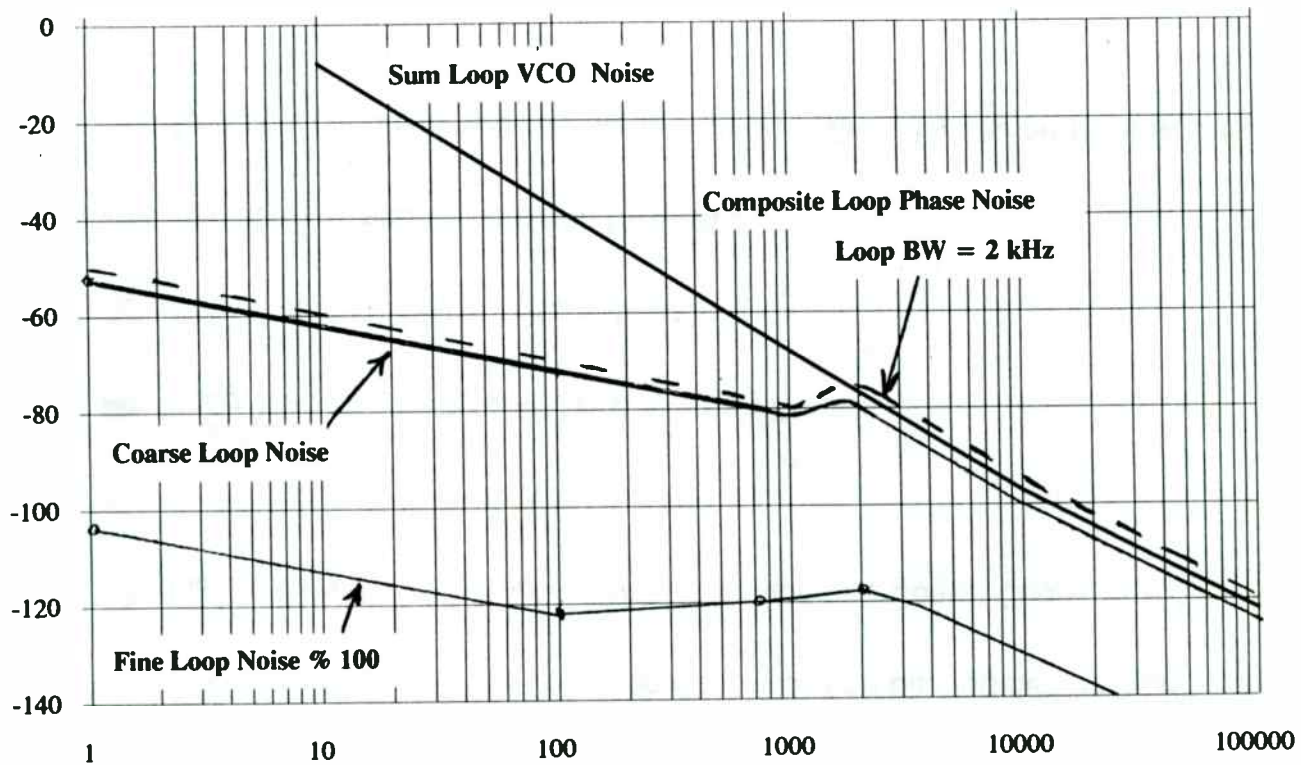
### ● Spurs

- Fine Step Loop (100 kHz) are > -80 dBc (divide by 100 reduces -40 dBc spurs by 40 more dB)
  - Output Loop (1 MHz) > -80 dBc (30 kHz LPF reduces Fref)
  - Mixer spurious will appear at the output at the same level, modified by Tracking response of Output Loop. (Tracking response can be modeled as LPF with  $F_c = \text{Loop BW}$ )
- Settling Time = Sum of both Loops settling =  $2 * 0.6 \text{ msec} = 1.2 \text{ msec}$



**Figure 17 - Triple Loop Synthesizer  
 for 1st LO of HF Receiver**





- o Fine Step Loop = 50-70 MHz Loop divided by 100 , its noise is reduced by 40 dB
- o Coarse Loop's noise appears dominates the Sum Loop's noise inside the Loop BW
- o Sum Loop's Loop BW = 2 kHz. N=1 so there is no multiplication of the Fine Loop's noise

Figure 18 - Output Phase Noise of Triple Loop Synthesizer

# Analysis for Triple Loop Synthesizer

- Phase Noise
  - Sum Loop (Output Noise ) includes contributions from Fine and Coarse Loops
  - Fine Loop's Noise is reduced by 40 dB so that it has no effect
  - Coarse Loop's Noise will be translated directly to Fout of the Sum Loop, modified by the Tracking Response of Sum Loop. (Tracking Response can be modeled as a LPF with  $F_c = \text{Loop BW}$ )
  - Inside Loop BW , the Fout noise is dominated by Coarse Loop
  - Outside LoopBW : VCO with  $QI = 10$
  - Integrated Noise 100Hz-3kHz = 0.7 deg RMS
  - SSB phase noise at 10 kHz = -90 dBc/Hz
- Spurs - All Fref spurs should be down by >-80 dBc
  - Fine Loop Spurs at 100 kHz are reduced by divide by 100 and by low pass in the Sum Loop ( $F_c = 12 \text{ kHz}$ )
  - Coarse Loop Spurs are reduced by Sum Loop's LPF ( $F_c = 12 \text{ kHz}$ )
- Settling Time = Sum of all 3 Loop's settling =  $3 * 0.8 \text{ msec} = 2.4 \text{ msec}$

Summary of Synthesizer Architecture Trades						
Architecture	Fout	Fstep	Phase Noise	Spurs	Settling	Complexity
	(MHz)		(100Hz-3kHz)		Time	
Single Loop	83-110	1 kHz	10.3 deg RMS	1 kHz = -35 dBc	16 msec	simple
Single with %8	83-110	1 kHz	5.4 deg RMS	8 kHz = -80 dBc	16 msec	High speed divider & VCO
Fract N	83-110	1 kHz	7.7 deg RMS	1kHz = -23 dBc, 2kHz = -47 dBc	16 msec	Fractional circuitry
Fract N %8	83-110	1 kHz	3.4 deg RMS	8kHz = -53 dBc, 16kHz = -80 dB	4 msec	High speed Div., VCO, Fract circuit
Dual Loop	83-110	1 kHz	0.6 deg RMS	all spurs > -80 dBc	1.2 msec	Two Loops, multiplier, narrow BPF
Triple Loop	83-110	1 kHz	0.7 deg RMS	all spurs > -80 dBc	2.4 msec	Three Loops





