ADVANCES IN MICROWAVES

VOLUME 2

This serial publication contains critical reviews covering a wide range of subjects of interest to the microwave engineer and physicist. The articles reflect the most significant contemporary progress in microwaves, and are written by experts who have played an important role in the development of the specific topic being covered. Each subject is thoroughly treated, and purely review material is kept to a minimum. *Advances in Microwaves* provides a permanent record of major developments in the field. It will be of value to all those working directly with microwaves, as well as to scientists and engineers in related areas who need a good general review of the particular advance described.

Volume 2 contains sections on:

- Tunnel Diode Devices
- Recent Advances in Solid State Microwave Generators
- Cooled Varactor Parametric Amplifiers
- Analysis of Varactor Harmonic Generators
- Theory and Design of Diplexers and Multiplexers
- The Numerical Solution of Transmission Line Problems

ADVANCES IN ASTRONOMY AND ASTROPHYSICS

Edited by Zdeněk Kopal

VOLUME 5

1967, about 350 pp., approx. $14.00

REVIEWS OF PREVIOUSLY PUBLISHED VOLUMES:

"Kopal and his associates have done a good service to astronomy by producing this excellent volume."

—SCIENCE

"Advances in Astronomy and Astrophysics will be valuable to a wide range of readers."

—SKY AND TELESCOPE

"The articles are solidly written and will demand frequent reference for many years to come."

—SCIENCE PROGRESS

CONTENTS:

John Hadjidemetriou, Secular Variation of Mass and the Evolution of Binary Systems
F. Zwicky, Compact and Dispersed Cosmic Matter
Miloslav Kopecký, The Periodicity of the Sun Spot Groups
D. E. Blackwell, D. W. Dewhirst, and M. F. Ingham, The Zodiacal Light
T. A. Bray and C. L. Goudas, Three-Dimensional Periodic Oscillations About L1, L2, and L3.

AUTHOR INDEX — SUBJECT INDEX.

YOUR CONTINUATION ORDER AUTHORIZES US TO SHIP AND BILL EACH VOLUME AUTOMATICALLY, IMMEDIATELY UPON PUBLICATION. THIS ORDER WILL REMAIN IN EFFECT UNTIL CANCELLED. PLEASE SPECIFY VOLUME NUMBER WITH WHICH YOUR ORDER IS TO BEGIN.
Editorial Board

D. J. Angelakos  
UNIVERSITY OF CALIFORNIA  
BERKELEY, CALIFORNIA

S. B. Cohn  
RANTEC CORPORATION  
CALABASAS, CALIFORNIA

H. M. Barlow  
UNIVERSITY COLLEGE  
LONDON, ENGLAND

P. D. Coleman  
UNIVERSITY OF ILLINOIS  
URBANA, ILLINOIS

R. W. Beatty  
NATIONAL BUREAU OF STANDARDS  
BOULDER, COLORADO

A. A. Oliner  
POLYTECHNIC INSTITUTE OF BROOKLYN  
FARMINGDALE, NEW YORK

P. J. B. Clarricoats  
UNIVERSITY OF LEEDS  
LEEDS, ENGLAND

G. E. Schafer  
NATIONAL BUREAU OF STANDARDS  
BOULDER, COLORADO
Advances in MICROWAVES

Edited by LEO YOUNG

STANFORD RESEARCH INSTITUTE
MENLO PARK, CALIFORNIA

VOLUME 2

1967

ACADEMIC PRESS New York and London
Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

E. G. Cristal, Stanford Research Institute, Menlo Park, California (237)

Harry E. Green, Weapons Research Establishment, Salisbury, South Australia (327)

B. C. De Loach, Jr., Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey (43)

G. L. Matthaei,* Stanford Research Institute, Menlo Park, California (237)

J. O. Scanlan, Department of Electrical and Electronic Engineering, University of Leeds, Leeds, England (165)

Fred Sterzer, RCA Electronic Components and Devices, Microwave Applied Research Laboratory, David Sarnoff Research Center, Princeton, New Jersey (1)

Michiyuki Uenohara, Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey (89)

* Present address: Department of Electrical Engineering, University of California, Santa Barbara, California.
Preface

The second volume of *Advances in Microwaves* contains six chapters, of which four are concerned with recent developments in microwave solid-state devices and circuits. There has been much activity in this area in recent years, and these chapters are particularly timely. They give comprehensive descriptions of tunnel diode devices, microwave solid-state sources, and cooled varactor parametric amplifiers, and discuss the analysis of varactor harmonic generators.

In general, it is not intended to select a special topic for each volume, and the unusual concentration on solid-state devices in this volume is merely coincidental, though perhaps symptomatic. The last two chapters are concerned with diplexers and multiplexers, and with TEM transmission line problems, both of which are important not only for microwave solid-state circuits, but for many other microwave applications as well.

New microwave techniques and devices are being developed in many laboratories at such a rate that it is hard for most of us to keep up with so much new knowledge. It is hoped that this serial will document at least the most important advances, and will help to keep you, the reader, abreast of new developments by means of these comprehensive and authoritative articles. You will find a list of planned future contributions on a following page.

We again wish to thank Stanford Research Institute for providing secretarial assistance. In particular, the editor is grateful to Mrs. Mary Lou Cahill and Miss Diane Bremer for lightening some of his editorial tasks.

*March, 1967*  

LEO YOUNG
Contents

Contributors ......................................................... v
Preface ................................................................. vii
Contents of Volume I ................................................. xi
Articles Planned for Future Volumes ................................ xiii

Tunnel Diode Devices
Fred Sterzer

I. Introduction ...................................................... 1
II. Tunnel-Diode Parameters ..................................... 2
III. Oscillators ...................................................... 12
IV. Amplifiers ...................................................... 21
V. Converters ...................................................... 30
List of Symbols .................................................... 36
References .......................................................... 38

Recent Advances in Solid State Microwave Generators
B. C. De Loach, Jr.

I. Introduction ...................................................... 44
II. Transistors ....................................................... 45
III. Varactor Harmonic Generators ............................... 56
IV. Tunnel Diodes ................................................ 62
V. The Gunn Oscillator ......................................... 67
VI. Avalanche Transit Time (IMPATT) Oscillators .......... 71
VII. Summary and Conclusion ................................... 80
Appendix .................................................................. 81
References .......................................................... 86

Cooled Varactor Parametric Amplifiers
Michiyuki Uenohara

I. Introduction ...................................................... 89
II. Characteristics of Cooled Amplifier ......................... 91
III. Device Considerations ...................................... 114
IV. Cooled Parametric Amplifiers .............................. 145
CONTENTS

V. Conclusion ........................................... 158
    List of Symbols ..................................... 161
    References ......................................... 161

Analysis of Varactor Harmonic Generators
J. O. Scanlan
I. Introduction ....................................... 165
II. The Varactor Model ............................... 167
III. Varactor Harmonic Generators without Idlers ................................. 172
IV. Varactor Harmonic Generators with Idlers ................................. 227
    List of Principal Symbols ............................. 234
    References ......................................... 235

Theory and Design of Diplexers and Multiplexers
G. L. Matthaei and E. G. Cristal
I. Introduction ....................................... 237
II. Concepts of Minimum Reactance and Minimum Susceptance Networks .......... 241
III. Theory of Complementary and Partly Complementary Diplexers ............... 246
IV. Design of Complementary and Partly Complementary Diplexers ................ 254
V. Theory of Multiplexers Having Noncontiguous Channels .......................... 293
VI. Theory of Multiplexers Having Contiguous Channels ............................ 305
    References ......................................... 324

The Numerical Solution of Transmission Line Problems
Harry E. Green
I. Introduction ....................................... 327
II. Numerical Solution of Laplace's Equation .................................... 328
III. Calculations on Transmission Lines ........................................ 349
IV. Equivalent Circuits .................................... 369
V. Concluding Remarks ..................................... 389
    References ......................................... 390

Author Index .......................................... 395

Subject Index ......................................... 400
Contents of Volume 1

Design and Fabrication of the Accelerating Structure for the Stanford Two-Mile Accelerator

Optical Waveguides
   A. E. Karbowiak

Directional Couplers
   R. Levy

The Use of Singular Integral Equations in the Solution of Waveguide Problems
   L. Lewin

Application of Lie Algebraic Theory to Microwave Networks
   M. Pease

Partially Filled Waveguides and Surface Waveguides of Rectangular Cross Section
   Wolfgang Schlosser and Hans-Georg Unger

Author Index—Subject Index
Articles Planned for Future Volumes

GROUND STATION ANTENNAS FOR SPACE COMMUNICATIONS
   D. C. Hogg

GUIDED TRANSMISSION WAVEBEAMS
   G. Goubau

MICROWAVE COMPONENTS FOR PHASED ARRAYS
   D. H. Temme and W. Ince

THE RAYLEIGH DISTANCE IN ANTENNA THEORY AND BEAM TRANSMISSION
   J. F. Ramsey

HIGH-POWER MILLIMETER WAVE SOURCES
   D. C. Forster

CROSSED-FIELD AMPLIFIERS
   H. McDowell and G. K. Farney

MICROWAVES AND PLASMAS WITH APPLICATION TO THE BALLISTIC MISSILE PROGRAM
   T. Morita and W. E. Scharfman

NONLINEAR INTERACTION OF SPIN WAVES AND ACOUSTIC WAVES
   B. A. Auld

MILLIMETER WAVE ASTRONOMY
   F. J. Low

H-PLANE Y-CIRCULATORS
   H. Bosma

MICROWAVE ULTRASONIC DELAY LINES
   F. A. Olson

THE GUNN EFFECT IN SEMICONDUCTORS
   J. B. Bott and W. Fawcett

ULTRALOW-NOISE PREAMPLIFIERS IN SATELLITE EARTH TERMINAL INSTALLATIONS
   C. L. Cuccia

PRECISION COAXIAL CONNECTORS
   R. C. Powell

MILLIMETER WAVEGUIDES WITH APPLICATIONS TO RAILWAY COMMUNICATIONS
   T. Nakahara and N. Kurauchi

THE MICROWAVE DEMODULATION OF LIGHT
   B. J. McMurtry and M. Di Domenico
ARTICLES PLANNED FOR FUTURE VOLUMES

MODE CONVERSION IN NONUNIFORM MULTIMODE WAVEGUIDES
C. C. H. Tang

HYBRID EH WAVES AND THEIR APPLICATION TO RF SEPARATORS OF VERY HIGH ENERGY PARTICLES
Y. Garault
Tunnel Diode Devices

Fred Sterzer
RCA ELECTRONIC COMPONENTS AND DEVICES
MICROWAVE APPLIED RESEARCH LABORATORY
DAVID SARNOFF RESEARCH CENTER
PRINCETON, NEW JERSEY

I. Introduction ........................................... 1
II. Tunnel-Diode Parameters ............................... 2
   A. Equivalent Circuit .................................. 2
   B. Cutoff and Self-Resonant Frequencies ............. 10
   C. The Noise Constant $I_0|r_j|$ ....................... 11
III. Oscillators ........................................... 12
    A. Theory ............................................... 12
    B. Experimental Results .............................. 17
    C. Higher Power Oscillators ......................... 20
IV. Amplifiers ............................................. 21
    A. Theory ............................................... 21
    B. Experimental Amplifiers ......................... 27
    C. Improved Amplifiers .............................. 30
V. Converters ............................................. 30
    A. Theory ............................................... 30
    B. Experiments ....................................... 33
    C. Improved Converters .............................. 36
List of Symbols ......................................... 36
References ............................................. 38

I. INTRODUCTION

Following the publication of Esaki’s classic paper on tunnel diodes in January 1958 [1], the potential of tunnel diodes for microwave applications was quickly established [2–3c]. A steady stream of papers on microwave tunnel-diode devices started appearing in the literature, and in the early 1960’s practical tunnel-diode amplifiers, frequency converters, and oscillators became commercially available, and were incorporated in industrial and military microwave systems.
Although there is still much significant work on microwave tunnel-diode devices going on today, most advances are relatively small, and the whole field seems to be approaching maturity. It is therefore an appropriate time to review the entire field, to take stock of what has been accomplished, and to point out areas where further work might be profitable.

The chapter is divided into four parts. The first part deals with parameters of the tunnel diode itself, the second deals with oscillators, the third with amplifiers and the fourth with frequency converters.

II. TUNNEL-DIODE PARAMETERS

Superficially tunnel diodes are very simple circuit elements. They are merely diodes that exhibit a negative resistance over part of their current–voltage (I–V) characteristics. This negative resistance is caused by processes that are so fast that there are no transit-time effects even at millimeter-wave frequencies. Closer inspection, however, reveals that the negative resistance of the diode has a capacitance in parallel with it; that the diode structure has resistance and inductance; that for most applications the diode must be packaged and the package invariably adds inductance and capacitance; and that there are thermal, shot, and flicker noise sources in the diode. These parasitic elements play an important part in determining the performance of a tunnel diode in a microwave circuit. It is therefore important to understand the factors that determine the magnitude of all tunnel-diode parameters, and to learn how to control and optimize them for a particular application.

A. EQUIVALENT CIRCUIT

It is generally agreed that the behavior of an encapsulated tunnel diode at microwave frequencies can usually be adequately described by the ac equivalent circuit of Fig. 1. [The notation used in this figure conforms to the IEEE standards on tunnel diodes [4].] The circuit consists of the following elements: a voltage-dependent junction resistance $r_j$, a voltage-dependent junction capacitance $C_j$, a series resistance $r_s$, a series inductance $L$, a package capacitance $C_p$,
a noise-voltage generator \( V_s \) associated with \( r_s \), and a noise-current generator \( i_j \) associated with \( r_j \). For microwave applications it is usually desirable that \( L, r_s, C_p \), and the two noise sources be as small as possible. The optimum values of \( r_j \) and \( C_j \) depend on the particular application.

The equivalent circuit of a given packaged diode is in general not unique at microwave frequencies, but depends to a small, but sometimes important, extent on the properties of the circuit into which the diode is inserted [4a]. The diode elements most likely to vary significantly in different circuits are the series inductance and the case capacitance.

1. The Junction Resistance \( r_j \)

Tunnel diodes have been fabricated from a host of different semiconductor materials including germanium, gallium arsenide, gallium antimonide, silicon, indium antimonide, and indium phosphide. However, only diodes made from the first three materials have found wide applications at microwave frequencies.

Figure 2 shows typical normalized \( I_jV_j \) characteristics (\( I_j \) stands for junction current and \( V_j \), junction voltage) of a Ge and of a GaAs tunnel diode (a similar graph for a GaSb diode is shown in Fig. 9). A plot of junction resistance \( (r_j = dV_j/dI_j) \) as a function of bias voltage \( V_b \) for the two diodes of Fig. 2 and for the GaSb diode of Fig. 9 is given in Fig. 3.

Meyerhoffer et al. [5] and Minton and Glicksman [6] have studied the dependence on doping density of the peak-current density \( J_p (= I_p/A, A = \text{area of the junctions}), \) the peak voltage \( V_p \), the valley-current density \( J_v (= I_v/A) \) and the valley voltage \( V_v \) of Ge diodes. Their experiments show (in agreement with a theory developed by Kane [7]) that \( J_p \) increases exponentially with the reduced doping density \( n^* \), as follows:

\[
J_p \propto n^{1/2} \exp(-\text{const} n^{*1/2})
\]  

(1)

The peak voltage \( V_p \) increases nearly linearly with the sum of the Fermi level penetrations into the conduction and valence bands, and therefore increases monotonically with \( n^* \); \( J_v \) varies exponentially with \( -n^{*1/2} \), and \( V_v \) increases monotonically with \( n^* \). Representative data points are given in Table I.

<table>
<thead>
<tr>
<th>Reduced doping density ( n^* ) (cm(^{-3}))</th>
<th>Peak-current density ( J_p ) (amp/cm(^2))</th>
<th>Peak voltage ( V_p ) (mv)</th>
<th>Valley-current density ( J_v ) (amp/cm(^2))</th>
<th>Valley voltage ( V_v ) (mv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 8 \times 10^{18} )</td>
<td>211</td>
<td>38</td>
<td>40</td>
<td>240</td>
</tr>
<tr>
<td>( 2.56 \times 10^{19} )</td>
<td>11,000</td>
<td>54</td>
<td>469</td>
<td>390</td>
</tr>
<tr>
<td>( 3.24 \times 10^{19} )</td>
<td>17,300</td>
<td>67</td>
<td>1147</td>
<td>420</td>
</tr>
</tbody>
</table>
Fig. 2. Normalized junction current versus junction voltage for a germanium and a gallium arsenide tunnel diode: Ge (---), GaAs (--.--).

Fig. 3. Junction resistance versus junction voltage for the two diodes of Fig. 2 and for the gallium antimonide diode of Fig. 9: GaSb (-----), Ge (----), GaAs (------).

The $I_1-V_1$ characteristics, and therefore also $r_j$, are temperature dependent. For Ge diodes it has been shown [1, 8–10] that (1) the peak voltage is substantially temperature independent, (2) the valley voltage generally decreases with increasing temperature, (3) the peak current may either increase or decrease with temperature depending on doping level, and (4) the valley current generally
decreases with decreasing temperature. Little information has been published on the temperature dependence of the region between the peak and valley points, but it appears that the transition from peak to valley is smooth at all temperatures, and is generally of the shape shown in Fig. 2.

2. The Junction Capacitance $C_j$

For voltages less than the valley voltage, $C_j$ can be calculated from the following well-known equation for the depletion-layer capacitance of an abrupt p-n junction:

$$ C_j = A \left[ \frac{e e n^*}{2(\phi_c - V_j)} \right]^{1/2} \tag{2} $$

where $e$ is the electronic charge, $\varepsilon$ is the dielectric constant, and $\phi_c$ is the contact potential. Reasonable agreement between experimental values of $C_j$ and values predicted from Eq. (2) is obtained by setting $\phi_c$ equal to 0.65 volt for germanium and 1.1 volts for gallium arsenide.

3. The Series Resistance $r_s$

The series resistance $r_s$ is the sum of the package resistance, the resistances of the two contacts, and the resistance of the diode structure. For most

---

**Fig. 4.** Alloyed n on p type of germanium tunnel diode before (a), and after (b), electrolytic etching.
microwave junction diodes, the package resistance and the contact resistances amount to less than 0.01 ohm, and are negligible compared to the resistance of the diode structure itself.

Because most microwave tunnel diodes have cylindrical geometries, the value of \( r_s \) can be estimated for any given diode structure from the equations for the high frequency series resistance of a long cylinder. These equations may be found in Ramo and Whinnery [11].

Varettoni et al. [12] have shown that for etched alloyed germanium diodes by far the greatest part of the series resistance is in the stem of the etched diode. Figure 4a shows a typical tunnel-diode junction before etching. When the diode is etched to reduce its peak current to the desired value (as is done for most microwave diodes) a stem is formed, as shown in Fig. 4b. The length of the stem (and therefore the value of \( r_s \)) depends on the size of the alloy dot and the resistivity of the base material. The smaller the dot size, the less etching is required, and the less "necking down" of the junction occurs. Also, the lower the base resistance, the shorter the stem. The best results obtained by Varettoni and co-workers was a series resistance of 0.8 ohm for a 0.7-mil diam junction. In this diode the alloy dot was 2 mils in diameter, and the resistance of the base material was \( 7.9 \times 10^{-4} \) ohm cm.

Diode structures that require little or no etching have been developed by Lueck et al. [12a] and Gibbons and Davis [12b]. In these structures necking down of the junction is minimized. Both groups use photolythic methods to define the area where alloying is to occur, and both surround the junction with a dielectric film that helps to support the alloyed area and the contact. Diodes of the type developed by Lueck and his co-workers require some etching to reduce the diode peak current to the desired value. Gibbons and Davis do not etch at all, but rely on the uniformity of the crystal and the processing to achieve uniform values of peak current for a given alloyed area.

4. The Series Inductance \( L \)

The series inductance \( L \) is the sum of the package inductance and the inductance of the diode structure. The magnitude of both of these inductances can be estimated from the equations for the high frequency inductance of a long cylinder [11]. These equations show that to minimize inductance it is necessary to minimize the dimensions along the direction of current flow; if skin effect is important (as it usually is for metals at microwave frequencies) one must also maximize the dimensions transverse to the current flow.

The most popular housings for microwave tunnel diodes are ceramic pill packages. The first pill packages were similar to those originally developed for varactor diodes, and used a narrow finger to make contact to the alloy dot (see Fig. 5a). This construction forced all current from the top part of the package to converge to one side of the metal rim and then flow through the
relatively narrow (and therefore high inductance) finger. The next step in the evolution of this package toward lower inductance was to replace the finger by a strip of wire mesh (Fig. 5b); the final step was to cover the entire top of the

![Fig. 5(a-c). Evolution of ceramic pill package.](image)

package with the connection to the alloy dot (Fig. 5c). The inductance of this last type of pill package was measured by Hauer [13] at microwave frequencies. Hauer's package had an over-all height of 0.057 in., and the outer diameter of the ceramic spacer was 0.080 in. The measured inductance was 235 picohenries; this inductance was not the usual self-inductance, but rather the inductance of the coaxial cavity that is formed when the outer diameter of the ceramic spacer is covered with metal.

For minimum inductance, the height of the package must be made as short as possible. Pill packages with ceramic spacers only 10 mils high have been successfully fabricated; reducing the height significantly below 10 mils presents serious technological difficulties.

![Fig. 6(a-b). Low inductance strip-line package.](image)
Packages with dielectric spacers only 2 mils high have been fabricated using copper-clad strip line [14]. In these packages the diodes are mounted directly into a single-ended strip line, as shown in Fig. 6a, and the diode package fits with a minimum of discontinuity into a strip-line circuit (Fig. 6b). The inductance of this package is estimated to be about 50 picothrows.

![Diagram showing lines of current flow in a strip transmission line terminated by a tunnel diode.](image)

Fig. 7. Diagram showing lines of current flow in a strip transmission line terminated by a tunnel diode.

By mounting the diodes into balanced rather than single-ended strip transmission line, further reduction in the inductance can be achieved. For example, Schneider [15] describes a balanced strip-line package with an inductance of only 20 to 30 picothrows. The ground plane spacing in Schneider’s package was 6 mils.

![Tunnel diode mounted in a conical cavity.](image)

Fig. 8. Tunnel diode mounted in a conical cavity.

When a diode is mounted in a circuit, the lines of current flow are altered, and additional inductance is usually introduced. This fact is illustrated in Fig. 7, which shows the current flow in a length of strip line terminated in a diode. Without the diode, the current lines would continue parallel to the end of the strip; with the diode, they converge as shown in the figure, thus increasing the flux linkage and therefore also the inductance of the strip-line section close to the diode. One way to eliminate this “induced” circuit inductance is to insert the diode in the center of a conical cavity, as shown in Fig. 8. In this arrangement the currents, even in the absence of the diode, converge to the center of the cavity, so that placing the diode in the cavity does not significantly change the current distribution. With this type of cavity mounting, Burrus [16] has achieved fundamental oscillations at frequencies exceeding 100 Gc.
5. The Package Capacitance \( C_p \)

The capacitance of most microwave pill packages, measured at 30 Mc, is of the order of 0.3 to 0.4 pf. These packages typically use alumina spacers (dielectric constant ~10) with the following dimensions: OD ~0.056 in., ID ~0.035 in., height ~0.010 in. \( C_p \) can be decreased by increasing the height of the spacers. However, this change increases the inductance of the package, and is therefore applicable only for diodes used in those special applications (certain broad-band amplifiers, for example) where minimizing \( C_p \) is more important than minimizing \( L \). \( C_p \) can also be decreased, without affecting \( L \), by replacing the high dielectric ceramic by a lower dielectric material like quartz. Unfortunately, such packages are presently not commercially available.

The "effective" capacitance of the strip-line package of Fig. 6a depends to an unusually large degree on the type of circuit the package is mounted in. If the circuit is fabricated from strip line of the same height as the package (Fig. 6b), then the package introduces little additional capacitance because it acts merely as a continuation of the transmission line of the circuit. On the other hand, if, for example, the package is clamped between two posts of a waveguide or coaxial circuit, the package behaves like a parallel plate capacitor, and the package capacitance can be quite high.

6. The Noise-Voltage Generator \( V_s \)

The noise voltage associated with the series resistance \( r_s \) is purely thermal in origin. Its mean square value is given by

\[
\overline{V_s^2} = 4kTBr_s
\]

where \( k \) is Boltzmann's constant, \( T \) is the absolute temperature of \( r_s \), and \( B \) is the bandwidth.

7. The Noise-Current Generator \( i_j \)

According to the model proposed by Esaki in his original Physical Review letter [1], there are two opposing tunnel currents flowing in a tunnel-diode junction. The net (measured) current flowing through the diode is the difference between these two tunnel currents. Both of these currents appear to generate full shot noise, and therefore the noise produced by the diode is that caused by full shot noise from a current equal to the sum of the two tunnel currents.

From Esaki's model, the following expression for the equivalent diode noise current \( I_n \) can be derived [17]:

\[
I_n = I_j \coth \frac{eV_i}{2kT}
\]

\( I_n \) is related to the mean square noise current \( \overline{i_j^2} \) by the equation

\[
\overline{i_j^2} = 2eI_n B
\]
Figure 9 shows curves of $I_j$ and $I_n$ for a GaSb diode at room temperature. Both measured values of $I_n$ and values calculated from Eq. (4) are shown. Agreement between measured and calculated values is quite good, except that the measured values are always somewhat higher than the calculated ones. Additional curves of $I_n$ versus $V_j$ for Ga, GaSb, GaAs, and Si at both room and liquid nitrogen temperatures can be found in a paper by King and Sharpe [17a]. Many of these curves show significantly larger discrepancies between measured and calculated values than Fig. 9.

**B. Cutoff and Self-Resonant Frequencies**

The small-signal ac impedance $Z_d$ across the terminals of a tunnel diode (neglecting $C_p$) may be written as follows:

$$Z_d = \left( r_s - \frac{|r_j|}{r_j^2 C_j^2 \omega^2 + 1} \right) + j \left( \omega L - \frac{r_j^2 C_i \omega}{r_j^2 C_j^2 \omega^2 + 1} \right)$$  \hspace{1cm} (6)$$

where it is assumed that the diode is biased into the negative-resistance region. (Unless otherwise noted, it is assumed throughout the rest of this chapter that $r_j$ is negative.) The real part of the impedance, $\text{Re}(Z_d)$, is equal to $(r_s - |r_j|)$ at zero frequency and increases monotonically with frequency. The frequency at
which $\text{Re}(Z_d)$ becomes zero is called the resistive cutoff frequency $f_r$ and is given by

$$f_r = \frac{(|r_j|/r_s - 1)^{1/2}}{2\pi |r_j| C_j} \quad (7)$$

At frequencies above $f_r$, $\text{Re}(Z_d)$ is positive, and the diode ceases to be an active device.

The frequency $f_r$ is a function of bias voltage. If $|r_j| > 2r_s$, as is usually the case for microwave tunnel diodes, then $(f_r)_{\text{max}}$ occurs at the bias voltage corresponding to the minimum negative resistance (neglecting the small effect of the variation of $C_j$ with bias).

Because $r_j$ is inversely proportional to the peak current of the diode, it follows from Eqs. (1) and (2) that the $|r_j| C_j$ product appearing in the expression for cutoff frequency depends exponentially on the reduced doping density of the junction. This exponential dependence is illustrated by the following two data points for germanium epitaxial $n$-on-$p$ tunnel diodes from Minton and Glicksman's [6] paper:

<table>
<thead>
<tr>
<th>$n^*$ (cm$^{-3}$)</th>
<th>$(1/r_j C_j)_{\text{max}}$ (Gc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$8 \times 10^{18}$</td>
<td>1.1</td>
</tr>
<tr>
<td>$3.5 \times 10^{19}$</td>
<td>57.4</td>
</tr>
</tbody>
</table>

Tunnel diodes with cutoff frequencies in the high microwave or even millimeter-wave range have been fabricated by using highly doped semiconductor crystals, and by minimizing the series resistance. Germanium diodes with $f_r$'s as high as 50 Gc are now commercially available from several manufacturers.

The imaginary part of $Z_d$ becomes zero at frequency $f_s$, where

$$f_s = \frac{(r_j^2 C_j L - 1)^{1/2}}{2\pi |r_j| C_j} = \frac{1}{2\pi} \left( \frac{1}{L C_j} - \frac{1}{r_j^2 C_j^2} \right)^{1/2} \quad (8)$$

$f_s$ is referred to as the self-resonant frequency or reactive cutoff of the diode and its value is usually quoted for $r_j = (r_j)_{\text{min}}$. Below self-resonance, the reactance of the diode is capacitive; above self-resonance it is inductive. Commercial diodes, housed in pill packages of the type shown in Fig. 5c are now available with $f_s$'s exceeding 30 Gc.

C. The Noise Constant $I_n|r_j|$

The noise figure of a tunnel-diode amplifier is proportional to the product of the equivalent noise current and the negative resistance of the diode used in the amplifier. This product, $I_n|r_j|$, is often referred to as the noise constant of the diode. Its minimum value usually occurs at junction voltages slightly higher than the voltages corresponding to minimum negative resistance.
The minimum reported values of \((I_n | r_\jmath|)_{\text{min}}\) for GaSb, Ge, and GaAs diodes are as follows:

| Material      | \((I_n | r_\jmath|)_{\text{min}}\) (mv) |
|---------------|----------------------------------------|
| GaSb [17b]    | 43                                     |
| Ge            | 60                                     |
| GaAs [17a]    | 75                                     |

Diodes with similar minimum noise constants are now commercially available. \((I_n | r_\jmath|)_{\text{min}}\) is not only a function of the diode material but also depends strongly on the techniques used in fabricating the diode. Unfortunately, little has been published on how the various parameters of the fabrication process affect the minimum noise constant.

III. OSCILLATORS

Microwave tunnel-diode oscillators are characterized by their simplicity (they consist basically of only a diode, a stabilizing resistor, and a resonator) and by the low voltage required to operate them (0.6 volt or less). They are useful in applications requiring microwatts or at most a few milliwatts of microwave power, and they can be relatively easily tuned over octave frequency ranges.

A. THEORY

A negative resistance is an active circuit element, i.e., a circuit element capable of producing power. If an rms current \(I_{\text{RMS}}\) is passed through an ac negative resistance \(-r\), an rms voltage rise \(I_{\text{RMS}}r\) is produced, and a power output \(I_{\text{RMS}}^2r\) is generated. The theory of negative-resistance oscillators was already well developed by the time the tunnel diode arrived on the scene. The oscillator most commonly treated in pre-tunnel-diode days consisted of a nonlinear negative resistance shunted by a capacitance \(C\), an inductance \(L\), and a load inductance \(G_L\) (see Fig. 10). Following van der Pol’s lead [18], most authors assumed that the \(I-V\) characteristic of the negative resistance is cubic, i.e., that

\[
I = -aV + bV^3
\]
When this assumption is used, analysis of the circuit of Fig. 10 yields the following results:

1. Oscillations will be nearly sinusoidal, provided

\[ \delta \ll 1 \]

where

\[ \delta = (a - G_L)(L/C)^{1/2} \]  \hspace{1cm} (10)

Note that \( \delta \) is defined as the reciprocal of the negative \( Q \) of the oscillator, excluding the nonlinear term \( b \).

2. The frequency of oscillations (provided \( \delta \ll 1 \)) is

\[ \omega = \left(1 - \frac{\delta^2}{16}\right)(LC)^{-1/2} \]  \hspace{1cm} (11)

3. The amplitude of the voltage swing is

\[ V_0 = \left[\frac{4}{3b}(9 - G_L)\right]^{1/2} \]  \hspace{1cm} (12)

4. The power delivered to the load is

\[ P_L = \frac{2}{3b} G_L(a - G_L) \]  \hspace{1cm} (13)

The maximum value of \( P_L \) occurs for \( G_L = a/2 \) and is given by

\[ (P_L)_{\text{max}} = \frac{a^2}{6b} = \frac{3}{16} \Delta I \Delta V \]  \hspace{1cm} (14)

where \( \Delta I \) and \( \Delta V \) are, respectively, the differences between peak and valley currents and voltages.

If the above results are to be applied to tunnel-diode oscillators, the characteristics of the diode must be approximated by a suitable cubic equation. One such approximation is shown in Fig. 11, together with the characteristics of a Ge tunnel diode. Agreement between the two curves is only fair.

Equation (14) shows that the maximum power output of a tunnel-diode oscillator is proportional to \( \Delta V \). For this reason, GaAs diodes are usually preferred in microwave applications requiring maximum power output because they have significantly larger \( \Delta V \) than either Ge or GaSb diodes (see Figs. 2 and 9).

Although van der Pol's solutions provide very useful insights into the behavior of tunnel-diode oscillators, they do not provide good quantitative agreement with actual microwave tunnel-diode oscillators. At microwave frequencies the parasitic elements of the diode equivalent circuit play an important part and cannot be neglected as they are in the equivalent circuit of Fig. 10.
An equivalent oscillator circuit that includes all parasitic elements of the diode is shown in Fig. 12. The question of whether this circuit will oscillate in the first place can be answered by taking its Laplace transform. The characteristic equation will in the general case have a number of eigenfrequencies, $S_n = \sigma_n + j\omega_n$. Oscillations can occur at any frequency $\omega_n$ for which $\sigma_n$ is greater than zero. However, if there exists no $\sigma_n$ greater than zero, the circuit will be stable.

Solutions for the eigenfrequencies of simple oscillator circuits can be found in Sterzer and Nelson [19] and Martin [20]. For example, if the circuit and load consist of a series combination of an inductance $L_c$ and a resistance $r_c$ then...
where

\[ S_{1,2} = \sigma_i \pm j\omega_i \]

\[ S_{1,2} = \frac{-1}{2} \left( r_t - \frac{1}{C_j|r_j|} \right) + \left[ \frac{1}{4} \left( \frac{r_t}{L_t} - \frac{1}{C_j|r_j|} \right)^2 - \frac{1 - r_t|j|}{L_t C_j} \right]^{1/2} \]  \hspace{1cm} (15)

where \[ r_t = r_c + r_s \quad L_t = L_c + L \]

A growing solution is obtained if

\[ L_t > r_t|j| C_j \]  \hspace{1cm} (16)

The initial growth can be either purely exponential (\( \omega_i = 0 \)) or sinusoidal (\( \omega_i \neq 0 \)). The conditions for sinusoidal growth are inequality [16] and

\[ \frac{1}{L_t C_j} > \frac{1}{4} \left( \frac{r_t}{L_t} + \frac{1}{C_j|r_j|} \right)^2 \]  \hspace{1cm} (17)

The steady state oscillations can, of course, be nearly sinusoidal, even though the initial growth is purely exponential.

The analysis of the power output of the oscillator of Fig. 12 is greatly simplified if the two nonlinear elements \( r_j \) and \( C_j \) are replaced by equivalent linear elements. If a sinusoidal voltage is applied across \( r_j \), then the current flowing through \( r_j \) contains a fundamental term at the same frequency as the voltage, plus harmonic terms whose frequencies are integral multiples of the fundamental. The fundamental current is in phase with the voltage and has a magnitude which depends on the voltage. Therefore, in terms of the fundamental frequency, \( r_j \) can be replaced by a linear equivalent resistance \( (r_j)_e \), whose magnitude is given by the voltage-dependent ratio of the fundamental components of RF voltage and current. If the voltage across \( r_j \) is nonsinusoidal, then the fundamental current and voltage are no longer in phase, and \( r_j \) requires a reactance as well as a resistance for its complete representation [21]. However, for nearly sinusoidal voltages, this additional reactance can usually be neglected. The effects of the nonlinearity of \( C_j \) are generally small and \( C_j \) can be assumed to be linear without introducing excessive errors.

In terms of the equivalent resistance \( (r_j)_e \), the power delivered to the load of the oscillator of Fig. 12 can be written as [22]

\[ P_L = P_d - P_s - P_e = \frac{V_0^2}{2|\omega_j|} - \frac{V_0^2}{2(r_j)_e^2} r_s [1 + \omega^2 C_j^2 (r_j)_e^2] - P_{out} \frac{G_e}{G_L} \]

\[ = \frac{V_0^2}{2(1 + G_c/G_L)} \left[ \frac{1}{|\omega_j|} - \frac{r_s}{(r_j)_e^2} - \omega^2 C_j^2 r_s \right] \]  \hspace{1cm} (18)

where \( P_d \) is the power generated by the negative resistance, \( P_s \) is the power lost in the series resistance, \( P_e \) is the power lost in the circuit, \( V_0 \) is the amplitude...
of the RF voltage developed across the junction, and \( \omega \) is the steady state oscillation frequency.

The values of \((r_j)_e\) and \( \omega \) can be estimated from the requirements that in the steady state (assuming sinusoidal oscillations) the impedance presented by the oscillator across the negative resistance must equal \(|(r_j)_e|\). (For analytical expressions for the load impedances presented across a tunnel diode by several types of microwave circuits, see Sterzer and Nelson [19] and Yamashita and Baird [22a]). Therefore

\[
|(r_j)_e| = \text{Re} \frac{(r_s + j\omega L + Z_T)(1/j\omega C_j)}{r_s + j\omega L + Z_T + 1/j\omega C} \tag{19}
\]

where

\[
Z_T = (G_c + G_L + jB_c)^{-1}
\]

\( \omega \) is a root of the equation

\[
\text{Im} \left[ \frac{(r_s + j\omega L + Z_T)(1/j\omega C_j)}{r_s + j\omega L + Z_T + 1/j\omega C} \right] = 0 \tag{20}
\]

If Eq. (20) has more than one root, then the correct root can be found by applying van der Pohl’s principle of minimum dissipation [23]. This principle states that if a self-oscillatory system has more than one possible mode of operation, it will operate in the mode requiring minimum output from the power sources present in the system.

As mentioned earlier, \((r_j)_e\) is a function of \( V_0 \), the amplitude of the RF voltage across the junction. The dependence of \((r_j)_e\) on \( V_0 \) was calculated by Kim and Brändli [24] using a cubic approximation to the \( I_j-V_j \) characteristics, and by the author [22] using a tenth-order power-series approximation. For the cubic approximation

\[
(g_j)_e = \frac{1}{(r_j)_e} = -a + \frac{3b}{4} V_0^2 \tag{21}
\]

where it is assumed that the diode is biased at the inflection point of the approximation. For the tenth-order approximation, which unlike the cubic approximation provides an excellent fit to the \( I_j-V_j \) characteristics, the results of the calculation are best presented in graphical form. Figure 13 shows, for example, a graph of normalized \((r_j)_e\) versus \( V_0 \) for typical GaAs diode. Curves for several values of the dc bias voltage \( V_B \) are shown.

The procedure for calculating the power output is as follows: (1) Calculate \((r_j)_e\) from Eqs. (19) and (20). (2) Find \( V_0 \) corresponding to \((r_j)_e\) and \( V_B \) from either Eq. (21) or a graph of the type shown in Fig. 13. (3) Use Eq. (18) to calculate \( P_L \). Agreement between measured and calculated power outputs is
reasonably good, especially if the graphs based on the tenth-order power-series approximation are used. (See Degen et al. [24a] for a procedure for calculating power output that is based on an approximation using exponential terms.)

![Graph of normalized effective negative resistance \( r_{\text{eff}} \) as a function of RF amplitude for a gallium arsenide tunnel diode. The dc bias voltage \( V_B \) is a parameter.]

**Fig. 13.** Graph of normalized effective negative resistance \( (r)_{\text{eff}} I_B \) as a function of RF amplitude for a gallium arsenide tunnel diode. The dc bias voltage \( V_B \) is a parameter.

**B. EXPERIMENTAL RESULTS**

1. **Fixed-Frequency and Mechanically Tunable Oscillators**

Microwave tunnel-diode oscillators have been built in waveguides, coaxial lines, and strip transmission lines. At frequencies above X band, waveguides are almost always used because losses in coax and strip lines are excessive. At X-band frequencies and below, the Q's of coax and strip-line resonators are high enough for most applications, and they are usually preferred to waveguide resonators because they are more compact.

Figure 14 shows a typical microwave strip-line oscillator [19]. The diode is placed in a re-entrant strip transmission line cavity, and the RF energy is taken from a coaxial connector (in Fig. 14 the coaxial connector is hidden from view by the strip-line board). The function of the stabilizing resistor is to prevent unwanted oscillations, particularly in the dc power supply leads. Such stabilizing resistors are used in most microwave tunnel-diode oscillators. For optimum RF efficiency they must be placed at a point where their loading of the desired oscillations is a minimum.
With circuits of the type shown in Fig. 14, Presser and Roswell [25] have obtained power outputs ranging from 13 to 105 mw in the range from 1.4 to 1.8 Gc. They used GaAs diodes \( (I_p = 200 \text{ ma}, I_v = 10 \text{ ma}, V_p = 230 \text{ mv}, V_v = 600 \text{ mv}, r_s = 0.5 \text{ to 0.8 ohm}, C_j = 8 \text{ to 20 pf}) \) housed in ceramic pill packages. Oscillators using single diodes produced power outputs of about 13 mw, oscillators using two diodes in a single resonator produced power output of about 26 mw and paralleling of 8 oscillators using two diodes each produced a power output of 105 mw.

Nelson et al. [26] described a tunable L-band oscillator using a straight strip transmission line cavity. This oscillator used a 600-ma GaAs diode \( (C_j \sim 30 \text{ pf}, r_s \sim 0.25 \text{ ohm}, f_c \sim 8 \text{ Gc}) \) mounted in a strip-line package of the type shown in Fig. 6, and produced a power output of about 20 mw over a tuning range from 1660 to 1780 Mc. The cavity was built from strip line with a characteristic impedance of only \( \frac{1}{2} \) ohm (thickness of dielectric 0.002 in.) to prevent spurious low frequency oscillations.

The circuits of both the oscillators described above present an inductive reactance across the diode terminals. These circuits will therefore operate only below the self-resonance frequency of the diode. [Note that for oscillators \( (r) \text{e} \) instead of \( r \text{e} \) must be used in the equation for the self-resonance frequency.]

Kim and Brändli [24] describe a 1-mw S-band strip-line oscillator that operates above the self-resonance frequency of the diode. In this oscillator the diode is connected to a transmission line that is shorted at one end to the ground plane, and the stabilizing resistor is connected across a gap in this line (this connection is in contrast to the circuit of Fig. 14, where the stabilizing resistor is connected directly between the strip line and the ground plane). With this circuit a capacitative reactance can be presented across the diode at

**Fig. 14.** Tunnel-diode oscillator using a re-entrant strip transmission line cavity.
the oscillation frequency, and at the same time unwanted lower frequency modes can be heavily loaded and suppressed.

A relatively simple arrangement for using two or more diodes in one circuit has been developed by Hauer [13]. The diodes sit in individual cavities, and are coupled together by \( \lambda/4 \) lines. A 2-diode GaAs oscillator of this type \( (I_p = 100 \text{ ma}) \) has produced a power output of 4 mw at 6 Gc.

Burrus and Trambarulo [16, 27] have operated tunnel-diode oscillators at millimeter-wave frequencies. These oscillators use GaAs, Ge, and Si diodes mounted in conical waveguide resonators of the type shown in Fig. 8. The diodes have no housing, but rather are built directly into the cavity. Power outputs ranging from 25 \( \mu \text{w} \) at 50 Gc to fractions of a microwatt at 103 Gc were obtained. More recently power outputs as high as 200 \( \mu \text{w} \) at 50 Gc have been obtained by Young et al. [27a]. These authors also built the diode directly into the waveguide circuit.

2. Electronically Tunable Oscillators

The frequency of tunnel-diode oscillators can be electronically tuned by several methods: (1) using a low-\( Q \) resonator and varying the bias voltage across the tunnel diode, (2) incorporating a varactor diode in the resonator of the oscillator and tuning the resonant frequency by varying the voltage across the varactor, and (3) using a garnet resonator and changing its resonant frequency by varying the magnetic field in the garnet. Experimental results obtained with these three methods of electronic tuning are shown in Table II [28–31a].

<table>
<thead>
<tr>
<th>Method of tuning</th>
<th>Frequency range (Gc)</th>
<th>Average output power (( \mu \text{w} ))</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias</td>
<td>0.9–1.8</td>
<td>400</td>
<td>28</td>
</tr>
<tr>
<td>Varactor</td>
<td>1.8–4.0</td>
<td>500</td>
<td>29</td>
</tr>
<tr>
<td>Garnet</td>
<td>0.7–2.1</td>
<td>100</td>
<td>30</td>
</tr>
<tr>
<td>Garnet</td>
<td>2.5–4.0</td>
<td>60</td>
<td>31</td>
</tr>
<tr>
<td>Garnet</td>
<td>5.1–10.2</td>
<td>500</td>
<td>31a</td>
</tr>
</tbody>
</table>

3. Frequency Stability

Good frequency stability has been achieved in carefully designed tunnel-diode oscillators. Figure 15 shows, for example, a Rieke diagram for a re-
entrant strip transmission line oscillator of the type shown in Fig. 14 [19]. The maximum fractional frequency deviation is about $3 \times 10^{-3}$ for a change in load voltage and standing wave ratio from 1 to 1.5. Stabilities with temperature variation of about $3 \times 10^{-6}$ per degree centigrade over the range from 0 to 100°C have been obtained with L-band GaAs oscillators, and frequency shifts of less than $2 \times 10^{-6}$ per millivolt of bias voltage over a range of 60 mw have been reported for S-band oscillators [32].

4. **Nonlinear Biasing Resistors**

A significant fraction of the total dc power input of a tunnel-diode oscillator is often dissipated in the stabilizing resistor. Wallmark and Dansky [33] have shown that the use of a nonlinear rather than a linear stabilizing resistor can reduce the dc power dissipation in the resistor by a factor of three for typical germanium tunnel diodes and by a factor of six for typical gallium arsenide tunnel diodes. The nonlinear resistance used by these authors consisted of a reverse-biased tunnel diode.

C. **Higher Power Oscillators**

The maximum power output of a tunnel-diode oscillator is, to a first approximation, proportional to $\Delta V$ and to the peak current $I_p$ [see Eq. (14); note that $\Delta I = I_p - I_c \approx I_p$]. Because $\Delta V$ is a property of the diode material, the only way to increase the power capability of diodes made from a given material is to increase $I_p$. 

---

**Fig. 15.** Rieke diagram of a tunnel-diode oscillator of the type shown in Fig. 14. $f_0 = 3530$ Mc. (---) Power output $\times 10^{-4}$ watt; (----) constant frequency contours.
The maximum peak-current diode that can be used in a microwave oscillator is limited by the parasitic elements of the diode. As the area of the diode junction is increased to increase $I_p$, $C_j$ increases [$C_j A$, see Eq. (2)], and the cutoff frequency usually drops [see Eq. (7); the product $r_j C_j$ is independent of area, but the ratio $|r_j|/r$, generally decreases as the area of the junction is increased]. Because operation above the self-resonance frequency of the diode is extremely difficult with high-current diodes, and because losses in the series resistance of the diode become excessive as $f$ approaches $f_r$, the maximum value of $I_p$ is limited by the requirements that $C_j < 1/(4\pi^2 L)$ and that the cutoff frequency be significantly higher than the operating frequency.

The above discussion shows that high power outputs from microwave tunnel-diode oscillators require diodes that combine large peak currents with high cutoff and self-resonant frequencies. Approaches that can be taken to develop such diodes include: product design of ultra-low-inductance strip-line packages [14, 15], or of diodes that do not require any packaging at all but can be mounted directly into a circuit [12b]; fabrication of diodes from extremely highly doped crystals to maximize the $I_p/C_j$ ratio; and finally fabrication of diodes with line junctions or multiple junctions to lower the resistance and inductance of the diode structure.

IV. AMPLIFIERS

When the potentials of tunnel diodes for microwave applications were first recognized, it was widely believed that the important microwave uses of tunnel diodes would be in oscillators and frequency convertors, and not in amplifiers. The reasoning was that the two-terminal nature of tunnel diodes would be no handicap in oscillators and frequency convertors, but that it would prove to be too cumbersome to adapt a two-terminal (1-port) diode to a two-port device like an amplifier.

The development of compact, high performance, ferrite circulators during the early 1960's has changed the whole outlook for tunnel-diode amplifiers. It is now possible to make compact, unconditionally stable, wideband tunnel diode amplifiers, and as a result amplifiers represent today the most important microwave application of tunnel diodes.

A. THEORY

1. Equivalent Circuit and Small-Signal Gain

   Practically all microwave tunnel-diode amplifiers are reflection-type rather than transmission-type amplifiers. In such reflection amplifiers, a module consisting of a tunnel diode and passive circuitry is placed at the end of a transmission line. The real part of the impedance of the module is negative, so that any incident RF wave is amplified, and the power in the reflected wave
is larger than the power in the incident wave. If the impedance of the module is \(-R+jX\), then the reflection gain \(G\) (i.e., the ratio of the incident power to the reflected power) can be written as follows:

\[
G = \frac{(R_0 + R)^2 + X^2}{(R_0 - R)^2 + X^2}
\]

(22)

where \(R_0\) is the characteristic impedance of the transmission line, which is assumed to be lossless.

Most applications require that the incident and the reflected signal be separated, i.e., they require that the amplifier have separate input and output ports. This separation can be accomplished by connecting two similar amplifiers by means of a hybrid ring [34] or by placing a ferrite circulator at the input of the amplifier. Circulator-coupled amplifiers are much more stable than hybrid-coupled amplifiers [35] and, because compact broad-band circulators are now available for all microwave bands, circulator-coupled amplifiers are used almost exclusively in microwave applications.

Figure 16a is a schematic diagram of a tunnel-diode amplifier using a four-port circulator. The input signal enters at port I of the circulator, leaves the circulator at port II, is amplified by the tunnel-diode module, reflected back to port II, and finally leaves the circulator at port III. An equivalent circuit for the amplifier module is shown in Fig. 17.

The flexibility of tunnel-diode amplifiers is greatly improved if, instead of fixed circulators, switchable latching circulators are used. [Compact (volume < 1/10 cu. in.), switchable, latching, three-port ferrite circulators suitable for use in tunnel-diode amplifiers have been developed recently by W. W. Sickanowicz and his co-workers in the author's laboratory.] Figs. 16b–16d show the front end of a microwave receiver consisting of a tunnel-diode amplifier using two switchable three-port circulators followed by a mixer. There are three useful settings of the circulators:

1. Both circulators are switched clockwise so that the tunnel-diode module and the mixer are switched into the circuit (Fig. 16b).
2. Both circulators are switched counterclockwise so that the module and the mixer are switched out of the circuit and any input power to the receiver is channeled into the load (Fig. 16c).
3. The first circulator is switched counterclockwise and the second clockwise so that any input power bypasses the module and is fed directly into the termination (Fig. 16d).

In pulse radars where it is often necessary to protect the tunnel-diode module and the mixer from TR tube leakage, the circulator can be switched back and forth between the settings of Figs. 16b and 16c. During "receive"
the circulators are set as in Fig. 16b; during “transmit” they are set as in Fig. 16c.

The dynamic range of the receiver can be significantly increased by switching the circulators back and forth between the settings of Figs. 16b and 16d. For small input signals the circulators are set as in Fig. 16b. For input signals that are large enough to drive the tunnel diode into saturation, the circulators are set as in Fig. 16d, and the tunnel module is bypassed. The input signals now feed directly into the mixer, whose saturation level is typically orders of magnitude greater than that of the tunnel diode.

Fig. 16. (a) Schematic diagram of a circulation coupled reflection-type tunnel-diode amplifier; (b–d) circulator settings useful in the front ends of microwave receivers consisting of tunnel-diode amplifiers using two switchable, latching, three-port circulators followed by a mixer.

Fig. 17. Equivalent circuit of tunnel-diode amplifier module.
2. **Bandwidth**

The voltage-gain-bandwidth product of a single-tuned circulator-coupled, tunnel-diode amplifier can be written as follows [3]:

\[ G_v B = \left( \pi |r_j| C_j \right)^{-1} \]  \hspace{1cm} (23)

Equation (23) assumes (1) that \( L \) and \( r_s \) are negligible, (2) that the circuit consists of a lossless inductance, and (3) that the amplifier has high gain.

Several authors have shown that much larger gain-bandwidth products than those of Eq. (23) can be achieved if multielement, broad-band matching networks with Chebyshev response are used to couple the diode to the transmission line feeding the circulator. Procedures for finding the proper values of the elements of the matching networks are given by Getsinger [36]. (Getsinger, like most other authors, assumes that \( L \) and \( r_s \) are negligible.)

Scanlan and Lim [37] give sets of curves which show the maximum bandwidth, and the corresponding gain ripple and phase response, that can be achieved with a given number of circuit elements. They also give similar curves for amplifiers with less than optimum bandwidth, but with lower ripple and better phase response than optimum amplifiers. The improvement in bandwidth over single-tuned amplifiers is considerable: for example, an amplifier with a 2-element matching network and a bandwidth of \( 1/(2\pi |r_j| C_j) \) has a voltage-gain-bandwidth product of about \( 2.8/(\pi |r_j| C_j) \); similar amplifiers with eight and with an infinite number of elements have voltage-gain-bandwidth products of about \( 10/(\pi |r_j| C_j) \) and \( 11.1/(\pi |r_j| C_j) \), respectively. The above results are valid only if \( L \) and \( r_s \) can be neglected. The effect of a nonnegligible series inductance is to restrict the range of gain and bandwidth that may be achieved for a given number of circuit elements [37a].

3. **Stability**

A practical tunnel diode must not oscillate under normal operating conditions. This rather obvious requirement sharply limits the class of tunnel diodes that can be used in practical amplifiers.

Smilen and Youla [38] have shown that a tunnel diode cannot be stabilized unless

\[ r_s/|r_j| < 1 \]  \hspace{1cm} (24)

and

\[ L/|r_j|^2 C_j < F(\theta) \]  \hspace{1cm} (25)

where \( F(\theta) \) is a function of \( r_s/|r_j| \) that varies monotonically from a value of 3 (for \( r_s/|r_j| = 0 \)) to a value of 1 (for \( r_s/|r_j| = 1 \)). Inequalities (24) and (25) are necessary but not necessarily sufficient conditions for stabilizing.
The sufficiency conditions for stability have not yet been fully determined. If

\[ \frac{L}{|r_j|^2 C_j} < 1 \]  

then stability can be achieved with purely resistive load [39]. If loads consisting of a resistance in parallel with a capacitance [40], or filter-type loads [41] are used, the limit on \( \frac{L}{|r_j|^2 C_j} \) can be raised somewhat above unity.

Useful graphical methods for determining whether or not a network containing a tunnel diode will be stable have been given by Henoch and Kvaerna [42] and Bandler [42a].

4. Noise Figure

Nielsen [43], using the equivalent circuit of Fig. 17, derived the following expression for the noise figure of a circulator-coupled tunnel-diode amplifier:

\[ NF = \frac{1 + (eI_n/2KT)|r_j|}{[1 - (r_0/|r_j|)][1 - (f/f_c)^2]} \]  

In the derivation of Eq. (27) it was assumed that the gain of the amplifier approaches infinity, that losses in the circuit and in the circulator are negligible, and that the temperature \( T \) of the amplifier equals the reference temperature used in defining the noise figure. (Nielsen also gives an expression for noise figure for the case where the amplifier noise figure differs from the reference temperature.)

Equation (27) shows that noise figure of a tunnel-diode amplifier depends on three factors: the noise constant \( I_n |r_j| \), the ratio of the parasitic resistance to the negative resistance, and the ratio of operating frequency to cutoff frequency. For lowest noise figure, these factors must be minimized.

Low gain amplifiers can have significantly lower noise figures than high gain amplifiers. However, as Hines and Anderson [44] have shown, a cascade of low gain amplifiers with high over-all gain will have the same noise figure as a single high gain amplifier with the same gain (assuming of course, that all amplifiers use similar diodes).

5. Power Output

The power output of a tunnel-diode amplifier with the equivalent circuit of Fig. 17 can be computed in a manner analogous to that for an oscillator [14] [see Eq. (18)]:

\[ P_{\text{out}} = P_{\text{in}} + P_a - P_s - P_e \]

\[ = V_0^2 \left[ 1 - \frac{1}{(r_j)_{\text{e}}} \left[ r_s + [(r_j)_{\text{min}} - r_s] \left( \frac{f_c (r_j)_{\text{e}}}{(r_j)_{\text{e}}} \right)^2 \right] \right] \]

\[ \frac{2(r_j)_{\text{e}}}{1 - \frac{1}{G} + \frac{2Z}{Z + R_0} \frac{R_0 G_e}{G}} \]

(28)
where \((r_j)_e\) is the large-signal equivalent negative resistance discussed previously in the section on oscillators, and where the cutoff frequency \(f_r\) is defined in terms of the minimum negative resistance. The dependence of \((r_j)_e\) on \(V_0\), the amplitude of the RF voltage across the diode junction, is plotted in Fig. 18 for GaAs and Ge diodes. These graphs assume dc biasing of the diodes at their minimum negative-resistance points.

From Eqs. (22) and (28) and Fig. 18, the saturation characteristics of GaAs and Ge tunnel-diode amplifiers can be computed. (See Hamasaki [44a] for a procedure for computing the saturation characteristics of GaSb tunnel-diode amplifiers.) The procedure is as follows: (1) Assume a value for \(V_0\) and find corresponding value of \((r_j)_e\) from Fig. 18. (2) Use Eq. (22) to find gain corresponding to \((r_j)_e\). (3) Find \(P_{\text{out}}\) and \(P_{\text{in}}=P_{\text{out}}/G\) from Eq. (28).

The maximum power output that can be generated by the junction of the amplifier diode is proportional to the maximum value of the diode peak current and the maximum value of the peak voltage across the diode junction. The maximum peak current is limited by stability considerations [see Eq. (25)] and can be written in terms of the voltage-gain-bandwidth product as follows:

\[
(I_p)_{\text{max}} \sim \frac{\alpha}{\pi G_v B L}
\]

where \(\alpha = I_p(r_j)_{\text{min}}\). The relative values of \(\alpha\) for GaAs, Ge, and GaSb diodes are

\[(\alpha)_{\text{GaAs}}:(\alpha)_{\text{Ge}}:(\alpha)_{\text{GaSb}} \approx 1:0.5:0.27\]

Equations (29) and (30) show that, for the same values of voltage-gain-
bandwidth product and inductance, the maximum usable peak current is considerably greater for GaAs diodes than for either Ge or GaSb diodes. \( (V_0)_{\text{max}} \) is determined by the maximum allowable large-signal gain depression of the amplifier, i.e., the maximum allowable value of \( r_1 \) \( r_{\text{min}} \), and is also approximately proportional to \( \alpha \). Therefore,

\[
(P_d)_{\text{max}} \text{GaAs} : (P_d)_{\text{max}} \text{Ge} : (P_d)_{\text{max}} \text{GaSb} \approx 1 : 0.2 : 0.075
\]

Thus, in amplifiers, just as in oscillators, GaAs diodes provide the highest power output.

6. Cascading

It was shown above that, although GaSb and Ge diodes provide lower noise figures than GaAs diodes, GaAs diodes have higher power-handling capabilities. The low noise properties of Ge and GaSb diodes can be combined with the power capability of GaAs diodes if a first-stage, low noise Ge or GaSb amplifier is cascaded with a second-stage GaAs power amplifier. Gain-power saturation characteristics of such two-stage amplifiers have been calculated by Steinhoff and Sterzer [14]. The same authors also present similar graphs for two-stage amplifiers, where both stages use GaAs diodes. In this case there is, of course, no improvement in noise figure, but the dynamic range of the cascaded amplifier is significantly larger than that of a single amplifier having the same gain as the cascaded amplifier.

B. EXPERIMENTAL AMPLIFIERS

A large number of experimental microwave tunnel-diode amplifiers have been described in the literature, and many amplifiers are now commercially

Fig. 19. Photograph of circulator-coupled C-band tunnel-diode amplifier.
available. Amplifiers have been built in waveguide, coaxial line, and strip transmission line, and operation has been achieved at frequencies as high as 85 Gc.

Figure 19 is a photograph of a typical commercial tunnel-diode amplifier [26]. This amplifier operates at C band, and uses a 1 ma Ge diode with a cutoff frequency of about 20 Gc and a noise factor of about 65 mv. The diode is housed in a low inductance ceramic package (L ~ 100 nh), of the type shown in Fig. 5c. The circulator has four ports and has coaxial input and output connectors. The total weight of the amplifier (including circulator) is approximately 12.5 ounces. Figure 20 shows the gain of the amplifier as a function of frequency, and Fig. 21 shows output power and gain as functions of input power. The change in gain over a temperature range from -20°C to 70°C is less than ± 1 db. The noise figure of the amplifier is typically 4.5 db.

The circulator is well matched (VSWR < 1.5), and has low insertion loss (~ 0.3 db) and high directivity (~ 20 db at input, ~ 40 db at output) over the operating frequency range of the amplifier. Outside this frequency range, however, both match and directivity tend to deteriorate. A stabilizing network is therefore incorporated in the amplifier module. This network presents very little loading at frequencies in the passband of the amplifier, but heavily loads the amplifier at all frequencies where stabilization due to the circulator is inadequate. As a result, the amplifier is stable with any input or output load. (See Hamasaki [44a] and Gallagher [44b] for discussions of stabilizing networks.)

Similar amplifiers are now available at all microwave frequencies up to Ku band. Typical noise figures range from about 3.5 db at L-band frequencies to 5.5–6.0 db at Ku-band, and a few amplifiers have voltage-gain-bandwidth products exceeding 6 Gc. Amplifiers with noise figures of less than about 4 db usually use GaSb diodes. Such amplifiers usually require temperature stabilization and have a smaller dynamic range than Ge amplifiers (see section on power output of amplifiers).
An experimental amplifier that covers the full octave from 2 to 4 Gc with a minimum gain of 10 db and a noise figure of about 4 db has been described by Lepoff and Wheeler [45]. The unusually large bandwidth is achieved by multiplexing a 2 to 3 Gc and a 3 to 4 Gc amplifier, each with its own circulator, and by using filter-type amplifier circuits with Chebyshev response.

Tunnel-diode amplifiers operating at frequencies ranging from 55 to 85 Gc have been built by Burrus and Trambarulo [46]. These amplifiers use formed GaAs diodes mounted in cavities similar to the one shown in Fig. 8. A gain of 20 db, a bandwidth of 40 Mc, and a noise figure of 16 to 18 db were measured at 55 Gc.

Fig. 22. Plot of interdemodulation cross products versus input power for a two-stage, large dynamic range, tunnel-diode amplifier [47].

Amplifiers with unusually large power output have been reported by Steinhoff and Sterzer [14]. They used a 22-ma GaAs diode in a re-entrant strip transmission circuit of the type shown in Fig. 14. The diode was mounted in a low inductance strip-line package (see Fig. 6). The amplifiers operated at L-band frequencies, had a gain of 12 db, and produced a power output of about 0.1 mw at a gain compression of 3 db.

Presser [47] has developed a broad-band, large-dynamic-range, S-band 2-stage amplifier for phased-array applications. The first stage of this amplifier uses a 4-ma Ge diode, and the second stage uses an 8-ma GaAs diode. The amplifier has a gain of 18 db, a bandwidth of 700 Mc, and a power output at
1 db gain compression of about 0.1 mw. The interdemodulation cross products for Presser’s amplifier are shown in Fig. 22.

Push-pull amplifiers for phased array applications have been developed by Lee and his co-workers [47a]. In a stripline push-pull X-band amplifier using two 1.8 ma Ge diodes they achieved a 10 dB increase in dynamic range and a 20 dB decrease in third interdemodulation product over a similar single-ended amplifier using a single 1.8 ma Ge diode. These improvements were achieved without any deterioration in noise figure.

C. IMPROVED AMPLIFIERS

Improvements in the performance of tunnel-diode amplifiers are still likely to occur in several areas: higher gain-bandwidth products, higher power outputs, and lower noise figures. Small, but important, improvements could result from reducing the parasitic elements of the diode by some of the means outlined in Section III,C. Major advances would probably require the development of integrated amplifiers in which the circuit elements of the module are fabricated on the same semiconductor wafer as the diode and the module is built directly into the circulator. A step in this direction already has been taken by Okean [47b] who successfully built a thin film tunnel-diode module using a beam-lead Ge diode [12b].

An important unanswered question about the noise constant would seem to warrant considerable further investigation. Is there a lower limit to \( (n_i |r_j|)_{\text{min}} \) (and therefore to the noise figure of tunnel-diode amplifiers), and, if so, how close to this limit are presently available diodes?

V. CONVERTERS

In microwave frequency converters, tunnel diodes must compete with a much older device, the point-contact crystal diode, and also with a newer device, the Schottky-barrier diode. Tunnel diodes can have lower conversion loss, and often also lower noise figures than either point-contact or Schottky-barrier diodes and, as a result, tunnel diodes now are being used in many microwave frequency converters.

A. THEORY

1. Equivalent Circuit

In a tunnel-diode frequency converter, an input signal of frequency \( f_s \) and a local-oscillator signal of frequency \( f_o \) are impressed across the tunnel diode. In general, a complete set of side-band frequencies is produced, i.e., all positive frequencies \( f_{m,n} = mf_s \pm nf_o \) (\( m \) and \( n \) take on all integral values) are present. However, in most practical converters significant power flow takes place only at frequencies \( f_s \) and \( f_o \), at the intermediate frequency \( f_i = \left| f_s - f_o \right| \), and at the
image frequency \( f_k = |2f_o - f_s| \). In this case, the tunnel-diode converter can be represented by the linear three-port frequency-translation network shown in Fig. 23, provided that (1) the local-oscillator voltage across the negative conductance of the diode is sinusoidal and is much greater than the sum of the amplitudes of the voltages at the signal, intermediate, and image frequencies, and (2) the effects of the variation of the junction capacitance with RF voltage are negligible [48–50]. In Fig. 23, \( Y_s \) is the total admittance at the signal frequency connected across the negative conductance; similarly \( Y_k \) and \( Y_i \) are, respectively, the total image and intermediate frequency admittances;

![Fig. 23. Equivalent linear three-port frequency-translation network for tunnel-diode frequency converter.](image)

g_o is the average value of the conductance as it is driven by the local oscillator, \( g_{c1} \) is the fundamental conversion conductance, and \( g_{c2} \) is the second-harmonic conversion conductance.

Sterzer and Presser [50] have computed the average and the two conversion conductances for typical Ge tunnel diodes, and give normalized curves of \( g_o \), \( g_{c1} \) and \( g_{c2} \) as a function of the amplitude of a sinusoidal local oscillator voltage and of the dc bias voltage. Similar curves are given by Pucel [48] and Christensen [51]. Techniques for measuring the three coefficients of microwave frequencies are described by Pucel [52]. The effects of nonsinusoidal local oscillator voltages are considered in Chisholm [50a] and Gambling and Mallick [50b].

In mixers that use only positive nonlinear conductance elements, like conventional point-contact diodes, the following inequalities must always be satisfied [53]:

\[
g_o > 0 \quad g_o > |g_{c1}| \quad g_o > |g_{c2}|
\]

(32)

However, because tunnel diodes display a negative conductance over part of their characteristics, neither of these restrictions applies to tunnel-diode converters.

2. Conversion Gain

Expressions for the conversion gain (i.e., the ratio of IF output power to the available signal power) can be derived from the equivalent circuit shown in
Fig. 23. If, for simplicity, it is assumed that $Y_s$, $Y_k$, and $Y_l$ are pure conductances equal to the internal conductance of the generator $g_g$, the image conductance $g_k$, and the load conductance $g_l$, respectively, then the conversion gain $K$ can be expressed as follows [50]:

$$K = \frac{4g_1g_g M^2}{(g_g + g_{in})^2}$$  \hspace{1cm} (33)

where

$$M = \frac{g_{cl}(g_o + g_k - g_{cl})}{(g_o + g_k)(g_o + g_l) - g_{cl}^2}$$  \hspace{1cm} (34)

and where $g_{in}$, the input conductance, is given by

$$g_{in} = \frac{g_{cl}^2(2g_{cl} - 2g_o - g_k) - (g_o + g_l)(g_{cl}^2 - g_o^2 - g_o g_k)}{(g_o + g_k)(g_o + g_l) - g_{cl}^2}$$  \hspace{1cm} (35)

Curves of $K$ and $G_{in}$ as a function of bias voltage and peak local oscillator voltage are given by Barber [53a] for both Ge and GaSb diodes. He evaluated these curves using a ninth-order power series approximation to the conductance versus bias characteristics of these diodes.

Passive converters must have conversion loss, as can be shown by inserting inequality (32) into Eq. (33). Tunnel-diode converters, on the other hand, may have conversion gain when any one of the inequalities is reversed. This conversion gain can be made arbitrarily large, and will approach infinity as $-g_{in}$ approaches $g_g$.

3. Bandwidth

The bandwidth of a tunnel-diode frequency converter can be calculated in a straightforward fashion by incorporating the parasitic elements of the diode into the admittances $Y_s$, $Y_k$, and $Y_l$ of the equivalent circuit of Fig. 23, and then using this equivalent circuit to calculate gain as a function of frequency [48]. Graphs for estimating the bandwidth of a converter without going through detailed calculations are given in Sterzer and Presser [50].

4. Noise Figure

The theory of noise in passive nonlinear-conductance frequency converters as developed by Strutt [54] also applies to tunnel-diode converters [48, 49, 55, 56]. Strutt shows that the noise figure, NF, of a nonlinear-conductance converter with short-circuited image impedance ($Y_k = \infty$) and with real input impedance can be written as follows.

$$NF = 1 + \frac{T}{T_0} \left( \frac{G_o(g_o + g_g)^2}{g_g^2 g_{cl}^2} - \frac{2(g_o + g_g)G_{cl}}{g_g g_{cl}} - \frac{G_o}{g_g} \right)$$  \hspace{1cm} (36)
where $T$ is the ambient temperature, $T_0$ is the standard reference temperature, and $G_o$, $G_{c1}$, and $G_{c2}$ are the first three Fourier coefficients of the time-varying noise conductance of the diode as it is driven by the local oscillator. (For the case $Y_k \neq 0$, see Eq. (62) of Kim [49]). In the derivation of Eq. (36) it has been assumed that the noise currents at the signal and intermediate frequencies generated by the pumped nonlinear conductance are fully correlated. Also, in accordance with the IEEE definition, the noise figure is written in terms of available noise output power, and does not take into account any noise contributions from the IF load. [See Breitzer [55] for a noise figure expression that does take the noise contribution from the load into account.] Graphs of minimum noise factor as a function of bias voltage and peak local oscillator voltage are given by Barber [53a] for Ge and GaSb diodes.

At frequencies above about 10 kc, the noise current in tunnel diodes is almost pure shot noise, and the equivalent noise conductance $G_N$ of the diode is given by

$$G_N = \frac{e}{2kT} I_n = G_o + 2G_{c1} \cos \omega_o t + 2G_{c2} \cos 2\omega_o t + \cdots$$

(37)

Graphs of $G_o$, $G_{c1}$ and $G_{c2}$ as functions of $V_o$ for several values of bias voltage are given in ref. 50 for Ge tunnel diodes.

B. Experiments

1. Active Converters (Conversion Gain > 1)

Klapper et al. [57] describe a converter with approximately 1 db gain operating with a signal frequency of 1 Gc and an intermediate frequency of 30 Mc. Input and output impedances are 30 ohms, and the required local-oscillator drive is about 1 mw. A 10-ohm resistor is connected in series with the diode, and when the diode is driven by the local oscillator it self-biases itself into the negative-resistance region. A double-side-band system noise figure of 4.3 db was obtained when the converter was followed by an IF amplifier with a noise figure of 5 db. This system noise figure is significantly lower than could have been obtained with best crystal mixers. (Even an ideal passive mixer would have produced a system noise figure of 5 db.) The saturation characteristics of the converter were quite good; at an input power level of $-10$ dbm, the power gain was still 0.5 db.

A self-oscillating or “autodyne” converter with a gain of 1 db has been described by Sterzer et al. [58]. With an oscillation frequency of 1550 Mc and an intermediate frequency of 30 Mc, a system noise figure of 6.5 db was measured [$N_{IF} = 3$ db]. Autodyne converters have the advantage of simplicity (a single tunnel diode acts as the active element for both the converter and the local-oscillator), have modest power-supply requirements (a few
milliwatts of dc at 0.1 to 0.3 volt), and, unlike nonoscillating converters, can use unstable tunnel diodes (i.e., diodes for which \( L/r_f^2 C_i > 3 \)). A severe drawback of these converters is that at high signal-power levels the self-oscillation frequency is pulled by the input signal, and the intermediate frequency is reduced. For example, in the autodyne converter described above, an input signal of \(-40 \text{ dbm}\) pulled the self-oscillating frequency by 1 Mc, and reduced the intermediate frequency from 30 to 29 Mc.

Conversion gains as high as 22 db at UHF and 12 db at S-band frequencies have been reported respectively by Chang and his co-workers [59, 60] and by Gambling and Mallick [50b]. Although such high gains are desirable in some applications, the sensitivity of the gain of a tunnel-diode frequency converter to variations in input impedance increases sharply as its gain is increased. As a result, high gain converters tend to be unstable unless they are protected against variations in input impedance by a ferrite isolator. This increased complexity has so far discouraged the use of high gain converters. If ferrite device must be used, it is usually preferable to use a tunnel-diode preamplifier, because the noise figure of a tunnel-diode amplifier is generally lower than that of a high gain converter using the same diode.

2. **Passive Converters (Conversion Gain \(< 1\)**)

Sterzer and Presser [50] describe balanced and unbalanced UHF and L-band tunnel-diode frequency converters with unity conversion gain. A typical balanced converter had a bandwidth of 250 Mc centered at 1270 Mc, and a double-side-band system noise figure \([f_i=30 \text{ Mc}, \text{ NF}_{\text{IF}}=1.7 \text{ db}]\) of less than 5 db for a local-oscillator drive anywhere in the range from 2.3 to 7 mw. The 3-db gain compression occurred at a signal input level of about 1 mw. Burnout (i.e., detectable deterioration of performance) occurred for CW powers of 2 watts, and for pulse energies of 170 ergs.

A broad-band converter operating with a local-oscillator frequency of
780 Mc had a double-side-band noise figure of 2.5 db and unity conversion gain. (Passive circuit losses in front of the converter, amounting to 0.5 db, are not included in the quoted gain and noise figure. These losses were caused by the local-oscillator input circuit and an IF rejection filter and can, in principle, at least, be eliminated by more careful design.) It is of interest to note that both the minimum radiometer noise figure and the minimum conversion loss of a broad-band positive-resistance frequency converter always exceed 3 db. Also, the radiometer noise figure of a high gain amplifier using this diode would exceed 3.5 db, as compared to a noise figure of only 2.9 db for a high gain system using the converter together with the best commercially available 30 Mc IF amplifier (NF = 0.7 db).

In converters with image rejection (see Fig. 24), the lowest measured single-side-band (radar) converter noise figure was 3 db. In this case the signal frequency was 512 Mc, and rejection at the image frequency (f<sub>k</sub> = 452 Mc) was 11 db.

Tunnel diodes with peak currents of a few hundred microamperes or less (often referred to as backward diodes because, in contrast to conventional crystal diodes, "easy" current flow occurs for voltages in the back direction) have found wide applications in passive frequency converters with intermediate frequencies below the megacycle range. At low intermediate frequencies, the noise output of crystal-mixer diodes increases rapidly with decreasing frequency because of the presence of 1/f noise. In backward diodes, on the other hand, 1/f noise is negligible down to at least audio frequencies [61, 62]. As a result, large improvement in noise figures can be achieved by using backward diodes rather than crystal diodes in frequency converters with low intermediate frequencies. The local-oscillator power required to drive backward diodes is usually about one order of magnitude less than that required for crystal diodes.

Eng [62] has compared the performance of a Ge backward diode with that of a 1N1838 crystal-mixer diode. The signal frequency was 13.5 Mc and the intermediate frequency 990 cps. With 30 µw of local-oscillator power, the backward diode had a noise figure of 21.5 db versus 45.5 db for the crystal mixer diode. With 250 µw of local-oscillator power, the noise figures were respectively 27.5 db (backward diode) and 37.5 db (crystal diode). Similar improvements in sensitivity have been measured in millimeter converters by Burrus [63] and by Chase and Chang [64].

Backward diodes with noise figures even lower than those reported by Eng [62] have recently been developed by Wright and Goldman [65]. These diodes have aluminum-germanium alloy junctions, and provide noise figures as low as 10.9 db for a signal frequency of 9.375 Gc and an intermediate frequency of 1 kc. For the same signal frequency, but for an intermediate frequency of 30 Mc, over-all receiver noise figures as low as 7 db were measured. Backward diodes also are widely used in video detectors [66]. In this
application they give greater sensitivity than crystal diodes if either the video bandwidth is very large, or the video frequency is so low that the $1/f$ noise in crystal diodes becomes significant.

C. IMPROVED CONVERTERS

Several basic questions about tunnel-diode frequency converters have not yet been resolved:

What is the lowest noise figure that can be achieved in a converter driven by a nonsinusoidal local oscillator, and what is the best semiconductor material to use? It is by now well established that nonsinusoidal local oscillators yield significantly lower noise figures than sinusoidal oscillators [50–50b], but as yet no complete theory for nonsinusoidal oscillators has been published.

What is the optimum shape of the $I-V$ characteristic for a given converter application? In amplifiers, for example, it is known that for lowest noise figure the slope in the negative-resistance region must be as steep as possible, and that for highest power output the voltage difference between peak and valley must be as large as possible. The answers to similar questions about frequency converters are not known.

Answers to the above questions would undoubtedly point the way to converters with lower noise figures.

Compared to little work has thus far been done on converters designed specifically for large dynamic range and high resistance to burnout. Promising approaches include the use of high current diodes housed in ultra-low-inductance packages, and the use of multiple-junction or line-junction diodes.

ACKNOWLEDGMENT

The author thanks D. E. Nelson and A. Presser for their discussions and assistance.

LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>area of tunnel diode (td) junction</td>
</tr>
<tr>
<td>$a, b$</td>
<td>constants in cubic approximation of td current-voltage characteristics [Eq. (9)]</td>
</tr>
<tr>
<td>$B$</td>
<td>bandwidth</td>
</tr>
<tr>
<td>$B_c$</td>
<td>circuit susceptance</td>
</tr>
<tr>
<td>$C_1$</td>
<td>junction capacitance of td</td>
</tr>
<tr>
<td>$C_p$</td>
<td>capacitance of td package</td>
</tr>
<tr>
<td>$e$</td>
<td>charge of the electron</td>
</tr>
<tr>
<td>NF</td>
<td>noise figure</td>
</tr>
<tr>
<td>NF$_{IF}$</td>
<td>IF noise figure</td>
</tr>
<tr>
<td>$f$</td>
<td>frequency</td>
</tr>
<tr>
<td>$f_i$</td>
<td>intermediate frequency</td>
</tr>
<tr>
<td>$f_k$</td>
<td>image frequency</td>
</tr>
<tr>
<td>$f_o$</td>
<td>local oscillator frequency</td>
</tr>
<tr>
<td>$f_c$</td>
<td>cutoff frequency of td</td>
</tr>
<tr>
<td>$f_s$</td>
<td>self-resonant or reactive cutoff frequency of td</td>
</tr>
<tr>
<td>$G$</td>
<td>gain</td>
</tr>
<tr>
<td>$G_c$</td>
<td>circuit conductance</td>
</tr>
<tr>
<td>$G_{c1}, G_{c2}$</td>
<td>coefficients in the Fourier expansion of the noise conductance of a td driven by a sinusoidal local oscillator</td>
</tr>
<tr>
<td>$G_l$</td>
<td>load conductance of td amplifier</td>
</tr>
</tbody>
</table>
$G_N$ equivalent noise conductance of td
$G_o$ first coefficient of the Fourier expansion of the noise conductance of a td driven by a sinusoidal local oscillator
$G_v$ voltage gain of td amplifier
$g_{c1}$ fundamental conversion conductance of td frequency converter
$g_{c2}$ second-harmonic conversion conductance of td frequency converter
$g_i$ load conductance of td frequency converter
$g_{ln}$ input conductance of td frequency converter
$(g_l)_e$ equivalent conductance of the junction of an oscillating td
$g_k$ image conductance of td frequency converters
$g_o$ average value of conductance of td driven by sinusoidal local oscillator
$I_i$ current flowing through junction of td
$I_n$ equivalent noise current associated with junction of td
$I_p$ peak current of td
$I_v$ valley current of td
$\Delta I$ $I_p - I_v$
$i_l$ noise current associated with junction resistance of td
$I_p$ peak-current density of td
$I_v$ valley-current density of td
$K$ Boltzmann’s constant
$L$ series inductance of td
$L_c$ circuit inductance
$L_i$ $L + L_c$
$n^*$ reduced doping density
$P_c$ power dissipated in circuit
$P_d$ power generated by the negative resistance of a td

$P_{in}$ power input to a td amplifier
$P_L$ power delivered to the load of a td amplifier or oscillator
$P_{out}$ output power of a td amplifier or oscillator
$P_s$ power dissipated in the series resistance of a td
$R$ real part of the impedance of td module
$R_0$ characteristic impedance of a transmission line
$r_c$ circuit resistance
$r_j$ junction resistance of td
$(r_j)_e$ equivalent $r_j$ if a sinusoidal voltage is placed across $r_j$
$r_s$ series resistance of td
$r_t$ sum of circuit and diode resistance
$S_n$ $n$th eigenfrequency of td oscillator
$T$ absolute temperature
$T_0$ standard reference temperature
$V_R$ dc bias voltage
$V_j$ voltage across junction of td
$V_0$ amplitude of RF voltage across the junction of a td
$V_p$ peak voltage of td
$V_v$ noise voltage associated with series resistance of td
$V_v$ valley voltage of td
$\Delta V$ $V_p - V_v$
$Y_{n}$, $Y_{k}$, $Y_s$ admittances at the intermediate, image, and signal frequencies of a td frequency converter
$Z$ impedance of td amplifier module
$Z_d$ impedance of td
$\delta$ parameter in van der Pol’s equation
$\epsilon$ dielectric constant
$\phi_c$ contact potential
$\sigma_n$ real part of $n$th eigenfrequency of td oscillator
$\omega$ angular frequency
$\omega_o$ angular frequency of local oscillator
REFERENCES


47. Presser, A. Private communication, RCA Electronic Components and Devices (1965).


Recent Advances in Solid State Microwave Generators

B. C. De Loach, Jr.

BELL TELEPHONE LABORATORIES, INCORPORATED
MURRAY HILL, NEW JERSEY

I. Introduction ........................................... 44
II. Transistors ........................................... 45
   A. Introduction ....................................... 45
   B. Physical Description of Devices ................. 46
   C. State of the Art .................................... 47
   D. Operating Principles ................................ 47
   E. Power-Frequency Limitations ....................... 48
   F. A Summary of Johnson's Study of Power-Frequency Limitations on Transistors ......................... 49
   G. A More Detailed Approach to the Power-Frequency Limitations of Microwave Transistors ............... 50
III. Varactor Harmonic Generators ......................... 56
    A. Introduction ....................................... 56
    B. Physical Description of Devices ................. 56
    C. State of the Art .................................... 56
    D. General Principles of Operation ................... 58
    E. Power-Frequency Limitations ....................... 59
IV. Tunnel Diodes .......................................... 62
    A. Introduction ....................................... 62
    B. Device Description .................................. 63
    C. State of the Art .................................... 63
    D. General Principles of Operation ................... 64
    E. Power-Frequency Limitations ....................... 66
V. The Gunn Oscillator ..................................... 67
    A. Introduction ....................................... 67
    B. Physical Description of Devices ................. 67
    C. State of the Art .................................... 68
    D. Operating Principles ................................ 69
    E. Power-Frequency Limitations ....................... 71
VI. Avalanche Transit Time (IMPATT) Oscillators .......... 71
    A. Introduction ....................................... 71
    B. Physical Description of Devices ................. 72
    C. State of the Art .................................... 73
    D. General Principles of Operation ................... 74
    E. Power-Frequency Limitations ....................... 78
VII. Summary and Conclusion ................................ 80

References .................................................. 86
I. INTRODUCTION

This paper has a threefold purpose:

1. To present a simple physical description and the basic operating principles of the solid state devices currently capable of generating coherent microwave power, including transistors, harmonic generators, tunnel diodes, avalanche transit time (IMPATT) diodes, and Gunn diodes.
2. To summarize the present state of the art for each device.
3. To ascertain the limitations imposed on these devices by device physics.

Recent advances in solid state microwave generators include the maturing of transistor-harmonic generator technology and the appearance on the scene of exciting new sources of microwave power such as the Gunn oscillator and IMPATT oscillators. The relatively sudden transition from no semiconductor microwave power sources to several, has led to considerable confusion in attempts at cross comparisons of device capabilities. Thus, in addition to describing the individual devices and their present performance, a major effort has been made herein to develop a framework for comparison based on inherent limitations imposed by device physics. These limitations are due to attainable velocities of carriers in solids and dielectric breakdown.

The existence of such limitations is a matter of considerable importance. Indeed the costly effort to develop the high frequency technology of ever shrinking dimensions has to be carefully weighed against the potentialities of possible resulting devices.

Throughout the paper "a device" is intended to imply any parallel, but not series or hybrid or traveling wave, combination of individual units whose interconnections are small fractions of a wavelength at the frequency of operation. This enables a meaningful comparison to be made of different devices and is discussed more completely in the transistor section (Section II).

The carrier velocity–dielectric breakdown limitations predict a decrease in attainable power with frequency for all devices considered. The exact dependence of attainable power on frequency is shown to depend on the precise nature of dielectric breakdown within the device and upon the maximum attainable velocity of carriers.

The paper is organized so that the casual reader can separately obtain a device description, general principles of operation, the most recently reported performance data, and fundamental device limitations for each device discussed without perusing the entire paper. A brief discussion of the phenomena of avalanche multiplication and carrier velocity saturation in semiconductors is included in the Appendix. The extent of the limitations imposed by these phenomena is so sweeping that the well-read physicist or semiconductor-microwave specialist can ill afford to remain in ignorance of their nature.
II. TRANSISTORS

A. INTRODUCTION

The advent of planar epitaxial technology has established the transistor, in concert with varactor harmonic generators, as the principal solid state source of microwave power. Although other devices appear promising for the future, the proven reliability and well understood operation of transistor harmonic generator combinations make their choice a "natural" for present day systems.

The precise dimensional control of planar technology allows the fabrication
and easy interconnection of multiple transistors, and the question of just what to call "a transistor" naturally arises. Throughout this paper we call "a transistor" any parallel (but not series or hybrid or traveling wave) combination of individual units whose interconnections are negligible fractions of a wavelength at the frequency being considered. The intent of this classification is to exclude distributed or semidistributed structures such as quarter-wavelength spacing along a transmission line, but to allow the inclusion of multiple devices that are iterated in close proximity, usually for reasons of thermal dissipation or current crowding. This classification permits an orderly discussion of the power impedance tradeoffs within individual classes of devices and allows cross comparisons of different types of devices on a realistic basis.

B. PHYSICAL DESCRIPTION OF DEVICES

A modern overlay transistor is shown in Fig. 1. The reader searching for the conventional emitter, base, and collector contacts will find instead that this one transistor has 156 separate "emitters" all connected by a metal "overlay" which in turn is supported by an oxide layer in regions where contact is not desired. Transistors of this type have greatly enhanced power capabilities over large area noniterated units owing to their higher periphery-to-area ratio, which minimizes current crowding, and to their superior thermal design.

![Fig. 2. Microwave power transistors—state of the art.](image)
The actual fabrication of such a device usually involves many operations in which the epitaxial layer on a heavily doped substrate is repeatedly oxidized, coated with a light-sensitive emulsion, exposed to radiation through a template or mask, developed with unexposed portions removed, then etched to remove the uncoated oxide. The resulting bare portions of the epitaxial layer can then be diffused, removed, or contacted with various metallic combinations.

C. State of the Art

A plot of the powers attained at various frequencies from individual transistors (in the sense previously described) is presented in Fig. 2. The data are also tabulated in Table I. A power proportional to reciprocal frequency squared plot is included for reference and is observed to be a reasonable predictor of power-frequency performance. Its significance will be discussed in another section. The highest reported "normal" frequency operation for a transistor [7] is 20 mw at 5 Gc. The use of the collector depletion layer capacity as a self-contained harmonic generator has produced useful amounts of power [8] at higher frequencies than the transit time cutoff frequency of the transistor employed. This type of operation will be considered a special case of a transistor harmonic generator combination, however, and will not be considered further.

D. Operating Principles

This section presents a simplified model of the operation of a modern power transistor. The particular model chosen illuminates the essentials of the operation of such a transistor, but more important serves as a base for the development of the power-frequency limitations of transistors. There are, of course, many volumes dedicated to the details of the operation of such transis-

---

Table I
Microwave Power Transistors—State of the Art

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Power (watts)</th>
<th>Efficiency (%)</th>
<th>Material</th>
<th>Operation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.25 Gc</td>
<td>2</td>
<td>—</td>
<td>Si</td>
<td>Pulse</td>
<td>[1]</td>
</tr>
<tr>
<td>1.0 Gc</td>
<td>1</td>
<td>30</td>
<td>Si</td>
<td>CW</td>
<td>[2]</td>
</tr>
<tr>
<td>500 Mc</td>
<td>5</td>
<td>—</td>
<td>Si</td>
<td>CW</td>
<td>[3]</td>
</tr>
<tr>
<td>800 Mc</td>
<td>1.5</td>
<td>30</td>
<td>Si</td>
<td>CW</td>
<td>[3]</td>
</tr>
<tr>
<td>400 Mc</td>
<td>20</td>
<td>—</td>
<td>Si</td>
<td>CW</td>
<td>[4]</td>
</tr>
<tr>
<td>2 Gc</td>
<td>0.5</td>
<td>—</td>
<td>Si</td>
<td>CW</td>
<td>[5]</td>
</tr>
<tr>
<td>175 Mc</td>
<td>40</td>
<td>75</td>
<td>Si</td>
<td>CW</td>
<td>[6]</td>
</tr>
</tbody>
</table>
tors and our question is one of function rather than of particulars. A one-dimensional current flow model is employed as shown in Fig. 3. This obviously differs in form from the transistor of Fig. 1, but can be employed to illustrate the workings of such a transistor as well as to develop its limitations. The $pn$ junction whose contacts are labeled emitter and base serve the function of periodically injecting charge into the lightly doped $\pi$-type "drift region" between the base and collector contacts. These drifting carriers then absorb energy from the dc field applied between the base and collector and owing to their variations with time convert it into RF energy. The low field emitter-base region thus functions as an electron (or hole) source, or gun, feeding the high field drift space between the base and collector and this drift space is thus the region in which the RF power is developed.

E. Power-Frequency Limitations

This section develops absolute limitations on the output power of transistors based on inherent limitations in device physics and attainable impedance levels in microwave circuitry. Several authors have treated this topic, with Early's paper [9] first introducing limitations inherent in (a) the limiting velocity of carriers in semiconductors, and (b) the maximum fields attainable in semiconductors without the onset of avalanche multiplication. These basic ideas were recently employed by Johnson [10] in a more general approach, which is of note in that it derives the limitations of transistor structures in terms of "materials constants." At least one family of devices that is treated herein, avalanche transit time diodes, violates one of Johnson's basic assumptions in his study of power-frequency limitations for transistors. Nevertheless, these diodes are limited by the same basic mechanisms as the transistors and a slightly more general development allows their inclusion in the framework of the Early–Johnson approach.
The following section summarizes Johnson's approach [10] to the maximum power obtainable at a given frequency from transistors, including the equation representing his principal result. A general development will then be presented which is more directly related to the underlying phenomena and consequently more readily adapted to other classes of devices. The approximations employed in Johnson's approach then become obvious.

F. A SUMMARY OF JOHNSON'S STUDY OF POWER-FREQUENCY LIMITATIONS ON TRANSISTORS

Johnson's approach was to find an absolute upper limit for the power obtainable from transistors which could not be surpassed under any conditions whatever. Further, he attempted to obtain a limit that was close enough to practicality to be a realistic guide as to what might be attained.

In order to achieve the above, Johnson made three explicit assumptions that are approximations of the actual physical situation. These are:

1. There is a maximum possible velocity of carriers in a semiconductor. This is the "saturated drift velocity" \( v_s \) which is a materials constant and is on the order of \( 6 \times 10^6 \) cm/sec for holes and electrons in silicon and germanium.

2. There is a maximum electric field \( E_m \) that can be sustained in a semiconductor without having dielectric breakdown (which is understood to be avoided in transistor design). This is approximately \( 10^5 \) volts/cm in germanium and about \( 2 \times 10^5 \) volts/cm in silicon.

3. The maximum current that a transistor can carry is limited by base widening.

With these three postulates and the simplified transistor model of Fig. 3 Johnson was able to obtain the following equation:

\[
(P_m X_c)^{1/2} f_t = E_m v_s (2\pi)^{-1}. \tag{1}
\]

Early [9, 11] had previously obtained the general form of this equation but Johnson has evaluated the constant of proportionality in terms of "material constants" \( E_m \) and \( v_s \). In Eq. (1), \( P_m \) is the maximum power that can be delivered to the carriers traversing the transistor, \( f_t \) is a device cutoff frequency (which can be thought to loosely approximate the highest frequency at which the device could be operated before the output power begins to decrease drastically owing to transit time limitations), \( X_c \) is the reactance at \( f_t \) of a capacity approximating that of the collector depletion layer, and \( E_m \) and \( v_s \) are the "materials parameters" previously defined.

The importance of Eq. (1) can be seen from the following. For a given \( E_m v_s \) product, that is, a given material, the maximum power that can be delivered to the carriers traversing the transistor is infinite. This obviously must be the case since we can in principle make our transistor as large (in cross...
section) as we please. Equation (1) also gives the obvious result of our infinite area device: $X_c$ must go to zero. It then follows that when the "optimum" transistor has been fabricated, whether it is a single unit or a parallel combination of individual units, there will always be a trade between device impedance level and power output capability. Thus Eq. (1) allows a prediction of the results attainable from higher-frequency transistors when and if they can be fabricated, provided information is available on levels of $X_c$ that can be accommodated in microwave circuitry. Figure 4 presents Johnson's equation (1) along with the best experimental results reported to date. The experimental $P_m$ values are taken from manufacturers' quoted maximum bias limits. The experimental $f_t$'s are similarly reported device data. As expected, the experimental results fall below the theoretical limit. Also included for comparison is another theoretical curve, which is derived in the next section.

It is suggested in the next section that the attainable impedance level in microwave circuitry falls off inversely with the square root of the frequency. This assumption in combination with either theoretical curve allows a prediction of the frequency dependence of attainable power levels in microwave circuitry. This is discussed more fully in the next section.

G. A MORE DETAILED APPROACH TO THE POWER-FREQUENCY LIMITATIONS OF MICROWAVE TRANSISTORS

As previously indicated, Johnson's assumptions [10] are approximations that simplify the analysis of the complicated physical phenomena involved in

![Fig. 4. $(P_m X_c)^{1/2}$ versus $f_t$ ($f_t$) for existing transistors, including theoretical limit curves of Johnson and De Loach.](image-url)
transistor limitations. In order to understand more clearly the nature of these approximations, and also to lay the ground work for their application to other devices, carrier velocity saturation and avalanche multiplication mechanisms in solids are reviewed in the Appendix. These two mechanisms will now be combined with appropriate assumptions to obtain an equation quite similar to that of Johnson but differing in minor respects. Familiarity with the Appendix is assumed in the following development.

We now pose the general question, What is the maximum rate at which power can be delivered to the carriers within a length $L$ of semiconductor of arbitrary cross section? The answer to this question would certainly represent an upper limit to the power extractable from this device as an RF source.

We first direct our attention to the maximum voltage that can be placed across a semiconductor of given length. This "maximum voltage" will require considerable qualification.

In structures of conventional sizes, the onset of avalanche multiplication limits the voltage that can be applied. We can obtain values of maximum allowable dc voltages under avalanche limitations by employing McKay's equation (A-3) and by assuming $\alpha = \alpha_0 E^m$ as a reasonable approximation to McKay's avalanche data of Fig. A-3 for fields in the range of interest. Then in the breakdown limit where $M$ approaches infinity we have

$$\int_0^L \alpha(E) \, dx = \int_0^L \alpha_0 E^m \, dx = 1 \quad (2)$$

where $E$ is in general a function of $x$. It can be shown, however, that $E =$ constant maximizes the integral

$$\int_0^L E \, dx = V \quad (3)$$

subject to the constraint given by Eq. (2). Since we are seeking the maximum voltage across our piece of semiconductor under avalanche limitations, we replace $E(x)$ by $E_c$, a constant, in the above and have

$$V_m = E_c L \quad \text{with} \quad E_c = \left( \frac{1}{\alpha_0 L} \right)^{1/m} \quad (4)$$

Thus, under avalanche limitations the maximum voltage that can be impressed across the semiconductor sample is given by

$$V_m = E_c L = \left( \frac{1}{\alpha_0} \right)^{1/m} L^{(m-1)/m} \quad \text{m}$$

and the critical field is observed to be a function of the length $L$ of the sample. Since $m$ is usually large, approximately 6 from McKay's data for silicon, $E_c$ varies slowly with $L$. However, a sixth-root dependence changes $E_c$ by approximately 1.5 per order of magnitude variation in $L$, and care must be taken in extending predictions by orders of magnitude. Thus Johnson's "materials
constant” is actually a slowly varying function of sample length for silicon. For other materials its variation may be more or less rapid, depending upon the sensitivity of $\alpha$ and $\beta$ to electric fields.

For very thin samples the phenomenon of quantum mechanical junction barrier tunneling can cause an even lower limitation on the voltage that can be impressed on a semiconductor of length $L$. Also, for extremely high electric fields, electrons can be torn from their normal bound valence states (a tunneling process) to give rise to “internal field emission current.” This concept was first introduced by Zener [12]. These effects must be considered only for extremely thin (high frequency) devices and provide a lower limit for these devices than that expressed by the avalanche multiplication limit. Thus our $v_m$ of Eq. (4) still represents an upper limit but perhaps a more optimistic upper limit than for devices of more conventional lengths. Barrier tunneling will be discussed in some detail in the tunnel diode section (Section IV).

Equation (5) was developed for the maximum dc voltage applicable to our sample of length $L$. It is an assumption then that this dc limitation is also appropriate for peak fields in RF transistor operation. If the word conventional is attached to “RF transistor operation” the appropriateness is fairly obvious. It appears unlikely that this limit will be appreciably exceeded in any conceivable RF transistor operation.

In order to obtain a result specifically appropriate for conventional transistor operation, we assume first, then, that avalanche multiplication represents the principal limitation, and second that this limitation is met by requiring that the multiplication be finite during all parts of the RF cycle. Then in place of Johnson’s assumption of a critical field that is a materials constant, we have a critical field given by Eq. (4) which (at least for silicon) is a slowly varying function of the distance over which a voltage is applied.

We now turn to the more difficult question of the maximum current that can be carried by a semiconductor sample. For both ac and dc currents the general answer is obviously that no maximum current exists in that one can, at least in principle, arbitrarily increase the area of the device without bound. Enhanced doping levels also provide increased current-carrying capability. We are forced then to ask a more restricted question, and we first discuss carrier velocities in solids. The nature of the dependence of the carrier velocities on electric field enables us to place restrictions on current density that provide useful and pertinent information. The maximum current through a “velocity-saturated” sample will still be infinite but will be shown to necessitate a zero impedance level. This in turn will provide finite limits on attainable power from individual units owing to the limited impedance ranges available in microwave circuits.

Although the shape of the $v-E$ curves of Fig. A-1 indicates no precise “saturated” velocity, it is apparent that for fields in excess of some 15,000
volts/cm for n-type silicon, 5000 volts/cm for n-type Ge, or 10,000 volts/cm for p Ge (and an uncertainty about p-type silicon reserved) that carrier velocities are little affected by field magnitudes. These particular values of field are labeled $E_s$ and (as did Johnson) carrier velocities are approximated by a constant labeled $v_s$ for fields in excess of $E_s$. One additional assumption (implicitly assumed by Johnson) must then be made: We limit our attention to samples in which $E$ is always above $E_s$. The above assumptions are now sufficient to obtain reasonable answers to our question of maximum current density.

Initially the assumption is made that current variations with time are slow enough (that is, carrier transit times are much less than a period of the current variation) that the carrier space charge density may be considered independent of distance and given by $\rho_c = J/v_s$ where $J$ is the current density in amperes per square centimeter. In addition the assumption is made that any fixed (net of swept donors and acceptors) charge density $\rho_f$ is also uniformly distributed. These restrictions greatly simplify the approach and, as will be shown, cause our result to be an upper limit.

Owing to the lack of dependence of $\rho_c$ and $\rho_f$ on distance the problem resolves itself into the question of how large a net charge density $\rho_{net}$ can be maintained across a region $L$. The net charge density $\rho_{net}$ necessitates a linear change in field across the sample and $\rho_{net}$ can in general increase until it causes the field to rise (Poisson’s equation) from our minimum value of $E_s$ at one end of the sample to $E_p$ at the other end of the sample, where $E_p$ is determined from the avalanche limitations of Eq. (A-3) as

$$E_p \approx E_c(m+1)^{1/2}$$

(The integral is simple when the region from $E_s$ to 0 field is also included, and this introduces little error since this region contributes very little to the total charge multiplication because of the high dependence of $\alpha$ on $E$.) For $m = 6$, $E_p \approx 1.4E_c$.

Poisson’s equation $\nabla \cdot D = \rho$ yields upon integration,

$$E_p - E_s \approx E_p - \frac{Q}{\epsilon A} \quad \text{where} \quad \frac{Q}{A} = \int_0^L (\rho_{net}) \, dL = (\rho_{net})L$$

$\epsilon$ is the dielectric constant of the semiconductor and $A$ is the cross-sectional area. Since we have equated hole and electron ionization coefficients, we need not concern ourselves with the sign of $\rho_{net}$ and hence the sign of $E_p$. Thus for uniform current and fixed charge distributions the net charge density allowable without exceeding avalanche breakdown is given by

$$\rho_{net} = \frac{E_p \epsilon}{L} = \frac{E_c(m+1)^{1/2} \epsilon}{L}$$
If \( \rho_t \) is made arbitrarily large, it follows from (8) that arbitrarily large dc currents can be obtained. It is thus necessary that, in order to have a meaningful limitation on current density, we must restrict ourselves to ac current considerations. It will be shown that requiring \( E \geq E_s \) limits the difference between the maxima and minima of current variations but does not limit the average or dc current. Since we are primarily interested in RF limitations this is quite appropriate.

It has already been shown in Eq. (8) that the net allowable charge density is limited by the onset of avalanche multiplication. This equation will now be employed to obtain the maximum and minimum currents allowed under our assumptions.

A section of semiconductor of length \( L \) and net fixed charge density \( \rho_f = N_d - N_a \) is considered. From Eq. (8) it is obvious that the normal majority carrier density, which is equal to \( \rho_f \) (that is, electron current for \( n \)-type material), can only be exceeded by

\[
\rho_{net} = \frac{E_p}{L} \epsilon
\]

and for the case in which \( L \) is “unswept” \( \rho_v \) cannot be reduced below the value \( \rho_v = \rho_f \) without having \( E \) fall below \( E_s \). Thus for an “unswept” condition a maximum current variation of \( \Delta I = v_s \rho_{net} A = v_s (E_p / A) A / \epsilon \) is allowed, where \( A \) is the cross-sectional area. Note that this is independent of \( \rho_f \). If the supply of majority carriers at one end of the sample can be limited, as would be the case for the collector depletion region of a transistor, current densities less than the minimum value allowed above can be accommodated while \( E \) is kept greater than \( E_s \). In this case the current can obviously go to zero. When this occurs, however, there is an unavoidable change in \( E \) across \( L \) due to the uncompensated fixed charge density \( \rho_f \). Thus the \( \rho_f \) that is allowable is limited through Eq. (8) and this in turn limits the absolute maximum current variation to

\[
\Delta I_{max} = \frac{2E_p v_s \epsilon A}{L} = \frac{2v_s E_c m (m + 1)^{1/2} \epsilon A}{L}
\]

For \( \rho_f = 0 \), \( \Delta I_{max} \) reduces to one-half the above.

It has been implicitly assumed in the calculation of \( \Delta I \) above that the ac current variations were slow enough compared to the transit time of carriers through the sample that \( \rho_v \) could be considered essentially independent of distance. For frequencies for which this is not the case, restrictions are even more severe and \( \Delta I \) will be even less than \( \Delta I_{max} \) owing to the nonuniformity of charge distribution. In the limit where pulses of charge are considered whose spatial extent in the direction of transport is small compared to \( L \), the maximum \( \Delta I \) is obtained when \( \rho_f = 0 \). This is obviously given by

\[
\Delta I_{max \text{ pulse}} = E_c v_s \frac{\epsilon A}{L}
\]
The $\Delta I_{\text{max}}$ of Eq. (10) thus represents an absolute upper limit which is not likely to be attained in practice for high-frequency operation. It is also obvious that $V_m$ cannot be applied when the current swing of Eq. (10) is at its peak without exceeding our avalanche conditions. We thus obtain an optimistic upper limit on the ac power deliverable to the carriers within our length of semiconductor $L$ as

$$P_m = \frac{V_m \Delta I_{\text{max}}}{2} \quad \text{or} \quad P_m = \frac{v_s E_c^m (m+1)^{1/2} \epsilon A}{(\alpha L)^{1/m}} \quad (12)$$

Since $\Delta I_{\text{max}}$ refers to a peak-to-peak variation, $\Delta I_{\text{max}}/2$ has been employed in Eq. (12).

If our piece of semiconductor were swept free of mobile charges and had metallic or ohmic regions attached at both ends, it would have a capacity given by

$$C = \frac{\epsilon A}{L}$$

If now we define $\tau$, the transit time of carriers across $L$, as $\tau = L/v_s$ and further label $1/2\pi \tau = f_\tau$, a characteristic cutoff frequency of the device, the above can be rewritten as

$$P_m X_{cr} f_\tau^2 = \frac{v_s^2 E_c^2 (m+1)^{1/m}}{(2\pi)^2} \quad (13)$$

where $X_{cr}$ is the reactance $1/\omega_\tau C$.

This is seen to be only slightly different from Johnson’s equation (1) through the presence of the $(m+1)^{1/m}$ factor, but in addition it should be remembered that $E_c \approx (1/\alpha_0 L)^{1/m}$ rather than a materials constant as assumed by Johnson. A plot of Eq. (13) is included in Fig. 4, where $P_m$ is to be interpreted as the maximum ac power deliverable to the carriers of the transistor operating at $f_\tau$ ($f$).

The development just presented had the major intent of illuminating the physics behind Johnson’s simplifying assumptions and as will be seen in later sections, provides a basis for its extension to other types of devices. Equation (13) shows a falloff of the $P_m X_{cr}$ product of a transistor to be less than $1/f_\tau^2$ owing to the inherent limitations of avalanches and carrier transit times. Other limitations such as thermal design, material uniformity, and so on can provide still lower limitations, but when these are overcome Eq. (13) represents an absolute upper limit which cannot be surpassed for normal transistor operation. The question of the maximum power attainable from a transistor at a given frequency is thus a question of the minimum impedance level attainable in practical microwave circuitry. If we assume (from skin effect considerations; see Moreno [13]) that this impedance level varies with the square root of frequency, that is, the higher the frequency the higher the impedance level at which we must work, then we observe from Eq. (13) that
power varies inversely as $f^{(2-2/m+1/2)}$ and that for $m=6$, as $f^{2.17}$. This is quite close to $f^2$ and due to the inaccuracies in the determination of $m$ we have chosen to plot simply a $1/f^2$ relation in the state of the art plot of Fig. 2.

It is likely that materials exist in which the simple McKay approach to avalanche breakdown is not a reasonable predictor of critical fields. (Factors of 2 or so are all that we require as “reasonable” in this general approach.) In such cases a more precise avalanche theory must be employed. The development of this section has hopefully pointed up the underlying physics and indicates the significant materials properties to be sought for the ultimate microwave power transistor.

### III. VARACTOR HARMONIC GENERATORS

#### A. INTRODUCTION

Varactors offer an efficient means of converting power at one frequency to a chosen one of its harmonics. In the face of such a statement the question naturally arises as to the wisdom of seeking higher and higher “prime” frequency sources, that is, a 5-Gc transistor oscillator as opposed to a 50-Mc transistor oscillator followed by a $\times 100$ varactor chain. The answer is to be found in several problems associated with harmonic generation. These include the degradation of signal quality with many stages of multiplication, the criticality of multiplier stage tuning, the large number of components required for many stages of power multiplication, and finally the inherent limitations imposed by device physics on the power-frequency capabilities of varactors. The latter problem is the main concern of this section. The limitations on the power-frequency capabilities of varactors imposed by the finite carrier velocities attainable in solids and the avoidance of avalanche breakdown are derived, and the implications of these results are discussed.

#### B. PHYSICAL DESCRIPTION OF DEVICES

Modern photoresist techniques and epitaxial technology have had a considerable impact on the fabrication of varactor harmonic generators. The combination of precise area and impurity profile control has changed these devices from laboratory curiosities into practical, efficient, high-power microwave devices.

A typical microwave power varactor and its encapsulation are shown in Fig. 5. The power capabilities of such devices have in the past exceeded the output powers available from solid state generators.

#### C. STATE OF THE ART

The highest-power, highest-frequency single-diode doubler and tripler operations reported in the literature are compiled in Fig. 6 and Table II. The
highest-frequency operation reported to date is a doubler operating from 54 to 108 Gc [20] at a reduced power level and is not included. Thus the frequency range up to and including 100 Gc has been covered by operating varactors.
A \(1/f^2\) plot through the best point is shown in Fig. 6 and will be discussed in Section III, E. It is observed to be a good predictor of output power-frequency performance.

### Table II

**Microwave Harmonic Generators—Doublers and Triplers—State of the Art**

<table>
<thead>
<tr>
<th>Output Frequency (Gc)</th>
<th>Output Power</th>
<th>Input Frequency (Gc)</th>
<th>Input Power</th>
<th>Multiplication</th>
<th>Efficiency (%)</th>
<th>Material</th>
<th>Operation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>2 watts</td>
<td>3</td>
<td>4.4 watts</td>
<td>(\times 2)</td>
<td>46</td>
<td>Si</td>
<td>CW</td>
<td>[14]</td>
</tr>
<tr>
<td>9.8</td>
<td>1.05 watts</td>
<td>4.9</td>
<td>3.5 watts</td>
<td>(\times 2)</td>
<td>30</td>
<td>Si</td>
<td>CW</td>
<td>[15]</td>
</tr>
<tr>
<td>12</td>
<td>300 mw</td>
<td>4</td>
<td>600 mw</td>
<td>(\times 3)</td>
<td>50</td>
<td>Si</td>
<td>CW</td>
<td>[16]</td>
</tr>
<tr>
<td>23</td>
<td>200 mw</td>
<td>11.5</td>
<td>470 mw</td>
<td>(\times 2)</td>
<td>43</td>
<td>Si</td>
<td>CW</td>
<td>[17]</td>
</tr>
<tr>
<td>36</td>
<td>97 mw</td>
<td>12</td>
<td>190 mw</td>
<td>(\times 3)</td>
<td>51</td>
<td>GaAs</td>
<td>CW</td>
<td>[18]</td>
</tr>
<tr>
<td>90</td>
<td>15 mw</td>
<td>45</td>
<td>190 mw</td>
<td>(\times 2)</td>
<td>8</td>
<td>GaAs</td>
<td>CW</td>
<td>[19]</td>
</tr>
</tbody>
</table>

D. **General Principles of Operation**

A schematic presentation of a high-frequency depletion layer varactor (we do not treat charge storage varactors in which both holes and electrons are present in comparable densities) is shown in Fig. 7. For peak forward voltages, Fig. 7a, the electrons of the moderately doped \(n\)-type region extend over a

![Figure 7](image)

**Fig. 7.** Idealized model of a "dielectric" varactor assuming one-dimensional current flow: (a) forward bias, (b) reverse bias.
distance $\Delta L$ and occupy almost all of the epitaxial layer in ordinary operation. The capacity of the diode in this state is given by

$$C_{\text{max}} = \frac{eA}{L - \Delta L}$$

(14)

and can be quite large before appreciable amounts of charge penetrate the remaining "barrier" and introduce loss. For peak reverse voltage (Fig. 7b) the carriers are swept out of the entire n-type region, which for optimum design usually occurs at reverse voltages just below avalanche breakdown. Here the capacity of the junction can be written

$$C_{\text{min}} = \frac{eA}{L}$$

(15)

Thus the depletion layer varactor exhibits a capacitance that varies with the applied voltage and can thus be employed as a harmonic generator.

E. POWER-FREQUENCY LIMITATIONS

We have assumed, Fig. 7, a simplified varactor model. These assumptions simplify the argument to follow but do not limit it. The maximum extent of the depletion layer width is labeled $L$ and the extent of its change during a cycle is labeled $\Delta L$. We call $\Delta L/L = \delta$ and note that for reasonably large capacity variations $\delta$ approaches unity. We may also interpret $\delta$ as

$$\delta = \frac{C_{\text{max}} - C_{\text{min}}}{C_{\text{max}}}$$

(16)

in capacitance terminology or

$$\delta = \frac{S_{\text{max}} - S_{\text{min}}}{S_{\text{max}}}$$

(17)

in elastance terminology.

For a varactor of depletion layer width $L$ there is a maximum voltage that can be applied without avalanche breakdown, which is given in Eq. (4) as $V_m = E_c L$. Since carrier velocities are limited, and since the charge $Q$ within the extent of $\Delta L$ must be removed in half a period, a transit time cutoff frequency $f_\tau$ can be defined as the maximum frequency of operation of such a varactor due to carrier velocity limitations,

$$f_\tau = \frac{v_s}{2\Delta L}$$

(18)

Note that at $f_\tau$ carriers could be withdrawn or inserted in a half-cycle only if a field of magnitude equal to or in excess of the saturation field were applied for the full half-cycle. Thus $f_\tau$ represents an upper limit of operation without degradation from transit time effects. This is of course in addition to the
degradation arising from the usual losses associated with conduction mechanisms in semiconductors and usually included in the derivation of a resistive cutoff frequency. This follows simply from setting the transit time \( \tau \) of a carrier across \( \Delta L \) equal to half a period of the maximum frequency, that is,

\[
\tau = T = \frac{1}{2f_r} \quad \text{with} \quad \tau = \frac{\Delta L}{v_s}
\]  

We note that combining the expressions for \( V_m \) and \( f_r \) gives

\[
V_m f_r \delta = \frac{E_c v_s}{2}
\]  

and that for a given \( \delta \) the product of the maximum voltage applicable to a varactor and its transit time cutoff frequency is a materials property represented by the \( E_c v_s \) product.

In order to obtain a limitation on power we need an expression involving the current. Here we distinguish displacement current and particle or transport current. Varactor energy transfer is associated with the motion of charges in electromagnetic fields. Displacement current, although present (and indeed necessary for varactor operation), cannot transfer energy from one frequency to another. The maximum particle or transport current that can flow in a depletion layer varactor is given by the total quantity of charge \( Q \) within \( \Delta L \) in motion at velocity \( v_s \) as

\[
I_m = \frac{Q}{\tau} = \frac{Qv_s}{L}
\]  

We obtain \( Q_{\text{max}} \) from Poisson's equation as

\[
Q_{\text{max}} = \epsilon A E_p \frac{\Delta L}{L}
\]  

where we have again assumed a uniform doping distribution. This assumption is not critical and the reader can quickly verify that other reasonable assumptions of impurity distribution do not yield significantly larger values of \( E_p \). Thus

\[
I_m = \frac{\epsilon A E_p v_s}{L} = C_{\text{min}} E_p v_s
\]  

Then combining Eqs. (6), (20), and (23) we have

\[
P_m X_{c_r} f_r ^2 \delta = \frac{E_c^2 v_s^2}{4\pi} (m + 1)^{1/m}
\]  

where \( P_m = V_m I_m \) is the maximum power convertible to \( f_r \) and where \( X_{c_r} \) is the reactance \( 1/\omega C_{\text{min}} \).
Equation (24) then represents an absolute limitation of a dielectric varactor harmonic generator due to avalanche and transit time effects. As with transistors, other effects such as resistive losses or thermal dissipation can prescribe lower limits than the above. Its similarity to Eq. (13) is to be noted, and should not be surprising owing to the identical limiting factors experienced by transistors.

Figure 8 displays Eq. (24) with the $E_c$ obtained from McKay, $v_t$ assumed equal to $10^7$ cm/sec, and $m$ assumed equal to 6. The best experimental point obtainable from the literature [17] is included for comparison. Equation (24) thus indicates that if dielectric varactors are operated at their $f_r$ (transit time cutoff frequency) the power attainable at a given impedance (reactance) level varies as $f_r^{-2(1-1/m)}$ owing to the dependence of $E_c$ on length. Employing McKay's value $m=6$ for silicon yields $f_r^{1.68}$. Again the impedance level of microwave circuitry is the critical parameter in determining the maximum
transferable power. If this impedance level is assumed to vary directly with the square root of the operating frequency (see Section II, E), that is, the higher the frequency the higher the impedance level at which we must work, then $P_m$ falls off as $f^{-(2-2/m+1/2)}$ or as $f^{2-m}$ for $m = 6$. Again owing to the inaccuracies in determining $m$ we have simply plotted $1/f^2$ in Fig. 6 and this is seen to give a reasonable fit to the data.

From Fig. 8 we conclude that silicon varactors are already operating within a factor of 4 of their theoretical limitations on power reactance product. It is possible that the theoretical "absolute" limit curve of Fig. 8 will actually be surpassed owing to minority carrier charge storage effects that were specifically neglected in the theoretical derivation.

The transit time cutoff frequency for a well-designed power varactor will in general be considerably below the resistive cutoff frequency, at least at the higher microwave frequencies. Thus although the resistive cutoff frequency is a useful figure of merit, $f_r$ is a more faithful predictor of the maximum frequency of useful operation.

The introduction broached the question of a low-frequency source and varactor chain versus a prime power source at the output frequency. The answer to this question is obvious when no prime source exists at the desired frequency but the general case is quite complicated. If the prime source is operating near its absolute limits and experiencing an $f^{-2}$ falloff with frequency which is to be expected with transistors and IMPATT oscillators and probably Gunn oscillators, more power can be obtained at a given frequency by the prime source-varactor chain method provided that the power levels involved fall below the power frequency limits for varactors (Figs. 6 and 8), and provided of course that varactor resistive losses are sufficiently low to allow efficient harmonic generation.

One has but to examine Figs. 2 and 6 to see that in terms of the present state of the art, considerably more power would be expected at 2 Gc by doubling from a 1-Gc transistor prime source rather than from a 2-Gc transistor. The reader is reminded that this argument ignores signal quality, that is, noise, spurious frequencies, and so on.

**IV. TUNNEL DIODES**

A. INTRODUCTION

The first tunnel diode was reported by Esaki [21] in 1958. Since that time the device has undergone intensive investigation as a microwave power source. Its ability to operate at high frequencies, in excess of 100 Gc, has been demonstrated [22] but its power output in the high-frequency range has been disappointingly low, and this section will attempt to elucidate the physical reasons.

---

1 See also p. 1.
behind this fact. The principal culprit will be shown to be the limitation of applied voltages to fractional bandgap values, that is, voltages less than 1 volt for conventional semiconductors. This in turn forces large currents for large power which unavoidably result in low impedance devices.

B. Device Description

Several techniques have been employed in constructing tunnel diodes including ball alloying and, more recently, planar beam leaded construction [23]. Examples of the latter are depicted in Fig. 9.

This process shows promise of good control, batch fabrication capability, and a lead geometry particularly appropriate for incorporation into microwave printed circuitry. It is, to the author's knowledge, the first demonstration that tunnel diodes can be made in such a fashion as to have unit costs comparable with other batch fabricated microwave diodes.

C. State of the Art

A presentation of the highest frequency-power points reported is shown in Fig. 10 and Table III. As indicated these span the microwave range up to 100 Gc. The $1/f^{2.5}$ plot is observed to be a reasonable criterion of performance for the diodes employing $p$-type GaAs. The "best" point was obtained with
n-type GaAs, which forms superior diodes in terms of high-frequency, high-
power performance. The justification of the $1/f^{2.5}$ plot is discussed in Section
IV, E.

<table>
<thead>
<tr>
<th>Frequency (Gc)</th>
<th>Power (mw)</th>
<th>Efficiency (%)</th>
<th>Material</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>9.1</td>
<td>—</td>
<td>GaAs</td>
<td>[24]</td>
</tr>
<tr>
<td>7.9</td>
<td>2.0</td>
<td>—</td>
<td>GaAs</td>
<td>[25]</td>
</tr>
<tr>
<td>9.0</td>
<td>2.5</td>
<td>2</td>
<td>GaAs</td>
<td>[26]</td>
</tr>
<tr>
<td>10.5</td>
<td>1.5</td>
<td>2</td>
<td>GaAs</td>
<td>[26]</td>
</tr>
<tr>
<td>49</td>
<td>0.2</td>
<td>—</td>
<td>GaAs</td>
<td>[27]</td>
</tr>
</tbody>
</table>

D. GENERAL PRINCIPLES OF OPERATION

The tunnel diode employs the quantum mechanical phenomenon of the
tunneling of electrons through potential barriers in order to obtain a negative
resistance. These devices are usually formed by alloying a contact onto a
semiconductor that is very heavily doped with a conductivity type opposite to
that of the alloyed material. At zero volts bias the reverse and forward tunneling
currents are equal so that no net current flows but the slope of the $I-V$ curve at
this point is very large and indeed there is no “reverse breakdown voltage” for
an Esaki diode in the usual sense of the word. Instead one draws increasingly
larger tunneling currents with increasing reverse bias. The negative resistance arises in the "forward bias direction" at a voltage intermediate between 0 volts and the diode's "built-in potential," and does not exist when this "built-in potential" is exceeded, owing to the usual forward bias injection current which then is present.

The operation will be described in more detail with the aid of Figs. 11 and 12. Figure 11a presents the energy band system of two degenerately doped

![Energy band picture of tunnel diode operation.](image)

**Fig. 11.** Energy band picture of tunnel diode operation.

![Tunnel diode current-voltage relation.](image)

**Fig. 12.** Tunnel diode current-voltage relation.
semiconductors of opposite conductivity type at zero bias. The horizontal straight line labeled $E_F$ represents the Fermi level which must be the same throughout the system at equilibrium. (Although this description of tunnel diode operation is strictly valid only at absolute zero it provides a good qualitative picture of conventional room temperature operation.) The dotted region represents the forbidden energy states of the band gap. The slashed region represents the energy states occupied by electrons. It is seen that in the zero bias state no electrons see “allowed” unfilled energy states on the other side of the barrier and thus no tunneling occurs (at absolute zero). This corresponds to point $a$ in Fig. 12. Figure 11b indicates the situation under small forward bias. Now conduction band electrons on the $n$ side of the barrier are opposite empty allowed states (holes) of corresponding energies on the $p$ side of the barrier region and tunneling begins. This corresponds to point $b$ in Fig. 12. As forward bias voltage increases and more states become available, more tunneling current is drawn. As the voltage is increased still further, however, the available states decrease with a corresponding decrease in current (negative resistance) until as shown in Fig. 11c no allowed states are available and tunneling ceases. This corresponds to point $c$ in Fig. 12. As the forward bias voltage is increased even further the usual forward bias diffusion current dominates. The negative resistance shown in Fig. 12 is then the origin of the microwave generation capabilities of the tunnel diode. A particularly illuminating exposition of this operation has been presented by Pucel [28].

E. Power-Frequency Limitations

From the above it is apparent that the maximum voltage $V_m$ that can be applied to a tunnel diode in a negative resistance state is less than the bandgap voltage $\varphi_b$, that is $V_m < \varphi_b$. For Ge, Si, GaAs, and GaSb, $\varphi_b$ is respectively 0.67, 1.1, 1.4, and 0.68 volt.

We now consider $I_m$, the maximum current that can flow in a tunnel diode. The tunneling process is very fast and the transit time limitations previously invoked from optical phonon collisions are inapplicable. This seems to be an advantage to the tunnel diode until one considers that, as $I$ is increased, for a fixed voltage on the order of $\varphi_b$, the impedance level of the device unavoidably goes down. Although absolute limits on $I$ certainly exist for finite areas, it seems inappropriate to develop them, in that existing devices are always limited first by the attainable impedance levels in microwave circuitry, that is, extremely large peak current diodes have been fabricated but have been totally inoperable as microwave sources owing to impedance problems. Thus the severe limitation of $V_m$ makes the tunnel diode a poor choice for a microwave power source unless one is interested in relatively low power applications.

Were it not for the unavoidable junction capacity one might expect no falloff of the output power with frequency. Unfortunately, however, these devices
have a very large capacity per unit area. This capacity shunts the negative resistance and the resulting impedance of the parallel combination

\[ Z = \frac{-R}{1 + R^2 \omega^2 C^2} - \frac{R^2 j\omega C}{1 + R^2 \omega^2 C^2} \]

has a real part that for \( 1/\omega C < R \) goes as \( 1/f^2 \). Thus for a given circuit impedance level, the value of \( R \) (and of \( C \) due to the fact that a given material has a given \( RC \) product) must be reduced as \( f \) increases and consequently reduces the available power. If we assume as in the previous sections that the lowest attainable real part of \( Z \) in microwave circuits is proportional to \( f^{1/2} \), then we expect an \( f^{-2.5} \) dependence of power with frequency. Lines of this slope have been drawn through the best \( p \)-type and \( n \)-type GaAs points in Fig. 10. \( n \)-Type GaAs has the better \( RC \) product and thus as expected has a superiority in power-frequency product. It must be remarked that we have ignored the resistive losses in the diode. When these are included the attainable power will of course fall off much more rapidly than \( f^{-2.5} \) as the diode’s resistive cutoff frequency is approached.

The conclusion is then drawn that tunnel diodes are poor choices for high-power microwave sources owing primarily to the low voltages at which negative resistance exists. This is an unavoidable consequence of the device physics and is intimately related to the bandgap of the semiconductor employed.

V. THE GUNN OSCILLATOR

A. INTRODUCTION

The Gunn effect, first reported by Gunn [29] in 1963, was manifested as oscillations in the current through a gallium arsenide sample when the applied voltage (field) reached a critical value. The oscillations were at microwave frequencies with 0.5 watt obtained at 1 Gc with an efficiency of approximately 2%.

The interest and activity stirred by the report was and is intense and the considerable effort to understand and exploit the underlying phenomena has resulted in a large number of talks and papers. This section presents a brief summary of the work to date and some preliminary comments on the power-frequency capabilities of such devices.

B. PHYSICAL DESCRIPTION OF DEVICES

Phenomenologically, the Gunn effect is quite simple to describe. A piece of bulk semiconductor of moderate but uniform doping concentration has “ohmic” contacts applied at both ends (Fig. 13a). This sample is then mounted in a microwave cavity in such manner as to couple power from any current oscillations into a microwave load (Fig. 13b). Either a low duty cycle pulse, to
minimize heating of the sample, or a dc bias is applied to the sample and power is detected in the microwave load.

From the start some things were fairly obvious. For one, the transit time of carriers through the sample was involved. Second, a relatively low critical electric field was required in the vicinity of from 1000 to 5000 volts/cm, and externally applied magnetic fields were not necessary for the effect to exist. The cryogenic requirements were also nominal and 300°K and higher operation was established. The principal problem involved was merely to remove enough heat from the sample to keep it (or its contacts) from melting.

There the simplicity ends. The interpretation of the phenomena underlying the experiments described above proved illusive.

C. State of the Art

The highest reported Gunn oscillator power-frequency performance for both pulsed and CW operation is plotted in Fig. 14 and given in Table IV. A $1/f^2$ plot is included and is again seen to be a reasonable match to experimental performance. Efficiencies of 5% have been attained in CW operation [31] and 7% [32] in pulse operation.
ADVANCES IN SOLID STATE MICROWAVE GENERATORS

Fig. 14. Gunn oscillator—state of the art.

Table IV

Gunn Oscillator—State of the Art

<table>
<thead>
<tr>
<th>Frequency (Gc)</th>
<th>Power</th>
<th>Efficiency (%)</th>
<th>Material</th>
<th>Operation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.16</td>
<td>100 watts</td>
<td>6</td>
<td>GaAs</td>
<td>Pulsed</td>
<td>[30]a</td>
</tr>
<tr>
<td>3.0</td>
<td>60 mw</td>
<td>5</td>
<td>GaAs</td>
<td>CW</td>
<td>[31]</td>
</tr>
<tr>
<td>3.2</td>
<td>2.5 watts</td>
<td>7</td>
<td>GaAs</td>
<td>Pulsed</td>
<td>[32]</td>
</tr>
<tr>
<td>5.0</td>
<td>2 watts</td>
<td>—</td>
<td>GaAs</td>
<td>Pulsed</td>
<td>[33]</td>
</tr>
<tr>
<td>7.0</td>
<td>35 mw</td>
<td>3</td>
<td>GaAs</td>
<td>CW</td>
<td>[34]</td>
</tr>
<tr>
<td>9.0</td>
<td>20 mw</td>
<td>3</td>
<td>GaAs</td>
<td>CW</td>
<td>[34]</td>
</tr>
</tbody>
</table>

a The authors report 200 watts by paralleling two wafers but it is not clear that their scheme is equivalent to our definition of a “single” device.

D. Operating Principles

Gunn reported [29] “fundamental oscillations” proportional to the length of the sample and harmonically related frequencies. Other investigators [31] have reported frequencies of oscillation representing almost a continuum and certainly not simply harmonically related to the length of the sample. Indeed, the existence of broad bands of negative resistance has now been demonstrated [35] in such samples. Thus different modes of operation seem to be operative under different experimental conditions.
It is now generally held that the phenomenon peculiar to the Gunn effect is the electron dynamics introduced by the multivalley nature of the conduction processes in certain semiconductor compounds [36, 37]. A quantum mechanical study of this process reveals that in a solid only certain values of momentum are allowed for conduction electrons and these vary with the electron's energy and direction of motion through the crystal. Part of the energy-momentum conduction band picture [38] for GaAs is shown in Fig. 15. For electrons having energies just above $E_0$, the energy minima of the lowest energy band, only the lower band is attainable. For electrons in this band a certain effective mass is assignable and consequently an effective mobility. Under an applied electric field, carriers in this band can "heat up," and at a particular value of applied field have sufficient energy ($E_0 + \Delta E$) to be allowed to exist in the momentum energy states of the upper "valley." The effective mass in this valley is very large or equivalently the effective mobility is very low. Thus carriers that transfer from the lower valley to the upper valley will slow down markedly for the same value of applied electric field, which gives rise to a negative resistance. Ridley and Watkins [39], and Hilsum [40] had independently predicted oscillations from this effect prior to Gunn's observations. Experiments have been performed which offer convincing evidence that it is indeed the source of the Gunn effect. Hutson et al. [36] applied hydrostatic pressure to a Gunn sample. This affects the relative energy separation of the conduction valleys in a known way and the Gunn effect was observed to disappear when the minimum of the energy level of the low-mobility valley was reduced to that of the high-mobility valley. Allen et al. [37] were similarly able to theoretically derive the experimentally observed critical fields for Gunn oscillations in GaAs-GaP mixtures in which the energy separation of the valleys is continuously tailorable through control of the percentage of admixture.
Whereas these experiments offered convincing proof of the mechanism involved, the dynamics of the situation have proved more difficult to elucidate. McCumber and Chynoweth [41] have recently obtained computer "movies" of the charge-field dynamics of a Gunn sample. Their work analytically predicts most of the experimental phenomena observed. This field is still in an active state of flux both theoretically and experimentally.

E. POWER-FREQUENCY LIMITATIONS

The electron dynamics of Gunn effect devices is complicated enough that their ultimate power-frequency performance remains an open question at present. The indications are, however, that their peak voltage will be limited to values considerably less than \( V_m = E_c L \) where \( L \) is the length of the sample, at least in the most efficient mode of operation. The maximum allowable current swing \( \Delta I_{\text{max}} \) is also undetermined at present but will not exceed the values previously developed from space charge limitations. The comparison of the ultimate power-frequency capabilities of these devices with other solid state sources must then await a more definitive theoretical treatment of their operation, but it appears from the above considerations that the power impedance product of the Gunn oscillator will be poorer than that of the IMPATT diodes (Section VI) and thus that they will be poorer sources of high-frequency power. On the other hand, Gunn oscillators have already produced exceptional amounts of pulsed low-frequency microwave power. These power levels are well within the avalanche-transit time limitations of IMPATT diodes, but it remains to be demonstrated that the large-area high-voltage IMPATT devices required for this low-frequency operation can actually be fabricated.

VI. AVALANCHE TRANSIT TIME (IMPATT) OSCILLATORS

A. INTRODUCTION

A new microwave power source has recently been reported by Johnston et al. [42] which employs the properties of avalanche and carrier velocity peculiar to semiconductors to obtain a microwave negative resistance. Read had predicted [43] in 1958 that these two phenomena could be combined in a certain manner to produce a microwave power source that was both efficient (30\%\;) and powerful (1.6 watt at 50 Gc). The Read diode, however, proved difficult to construct and the first operating Avalanche Transit Time oscillator [42] employed the depletion region of a simple silicon \( p n \) junction to produce 80 mw of pulsed 12-Gc power. This work was closely followed by Lee et al. [44] in reporting the first "Read" diode, which employed the same basic mechanisms but a considerably different field structure to obtain CW 1-\( \mu \)w, 200-Mc oscillations. \( p n \)-Junction IMPATT (\textit{IMPact Avalanche and Transit Time})
oscillators have since been reported in Ge [45] and GaAs [46] and microwave oscillations have been obtained [45] from a Read structure.

The "Read" structure was analyzed and indeed predicted by Read [43] in 1958. The pn-junction structures as well as a pin structure were recently analyzed by Misawa [47] for small-signal operation. Misawa's calculations also allow the inclusion of arbitrary field-distance profiles and are thus capable of calculating variations on these simple structures. Experimental work [48] on pin structures has confirmed the accuracy of Misawa's theoretical model. This section, in addition to summarizing results in this new field, develops the limitations for IMPATT oscillator power-frequency performance.

B. PHYSICAL DESCRIPTION OF DEVICES

The first IMPATT oscillator was a conventional pn-junction mesa-type computer diode, and was operable only on a pulse basis. By employing an

---

**Fig. 16.** Epitaxial pn junction IMPATT oscillator.

**Fig. 17.** A 200-Mc Read diode oscillator structure. (Courtesy of R. L. Batdorf. See Lee et al. [44].)
epitaxial layer to reduce loss (Fig. 16), continuous operation [42] was achieved. The first "Read" diode reported [44] operated at 200 Mc and was constructed as shown in Fig. 17. A liquid gallium-indium contact was required to avoid microplasmas, and approximately 1 μw of CW 200-Mc power was obtained. The low-frequency version avoided some of the problems of high-frequency circuitry at the expense of relatively high operating voltages, 700 to 900 volts, and modest drift fields on the order of 20,000 volts/cm.

Fig. 18. (a) A microwave Read diode oscillator structure and (b) its associated electric field. (See De Loach and Johnston [45]).
The first microwave version [45] of the Read diode operated in the 5-Gc region and was constructed as shown in Fig. 18. This diode employed conventional double diffusion techniques to approximate the doping profile suggested by Read. Although microplasmas were present, indicating non-uniform current flow at low currents, it produced 19 mw of CW 5-Gc power.

C. State of the Art

The experimental performance of IMPATT oscillators is in a high state of flux. Burrus has recently reported [49] 350 mw at 50 Gc for pulsed operation of a silicon pn junction. De Loach and Johnston [45] reported 19 mw of 5-Gc CW power from a microwave Read diode with 1.5% efficiency, and

<table>
<thead>
<tr>
<th>Frequency (Gc)</th>
<th>Power (mw)</th>
<th>Efficiency (%)</th>
<th>Material</th>
<th>Type</th>
<th>Operation</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3</td>
<td>130</td>
<td>4.9</td>
<td>Si</td>
<td>pnpn</td>
<td>CW</td>
<td>[51]</td>
</tr>
<tr>
<td>7.7</td>
<td>80</td>
<td>2.8</td>
<td>Si</td>
<td>pin</td>
<td>(CW cooled)</td>
<td>[52]</td>
</tr>
<tr>
<td>11.0</td>
<td>30</td>
<td>1.6</td>
<td>Si</td>
<td>pin</td>
<td>CW</td>
<td>[52]</td>
</tr>
<tr>
<td>12.5</td>
<td>13</td>
<td>0.5</td>
<td>Si</td>
<td>pn</td>
<td>CW</td>
<td>[42]</td>
</tr>
<tr>
<td>13.6</td>
<td>24</td>
<td>4.25</td>
<td>GaAs</td>
<td>pn</td>
<td>CW</td>
<td>[50]</td>
</tr>
<tr>
<td>50</td>
<td>350</td>
<td>0.5</td>
<td>Si</td>
<td>pn</td>
<td>Pulse</td>
<td>[49]</td>
</tr>
</tbody>
</table>
Higgins et al. [50] have reported 20 mw of CW 11-Gc power with 4.5% efficiency from a GaAs varactor structure. Still more recently Johnston [51] has reported 130 mw of CW 5.3-Gc power with approximately 5% efficiency from a silicon Read diode. Although sufficient data have not been accumulated to substantiate the expected power-frequency performance, a $1/f^2$ plot is shown in Fig. 19 through the best pulsed power point and the best CW point. The data points are summarized in Table V.

D. GENERAL PRINCIPLES OF OPERATION

We now present the elements of the theory of the “Read” avalanche-transit time diode. The other IMPATT oscillators are simpler to construct but are

![Figure 20](https://via.placeholder.com/150)

Fig. 20. (a) The NPIP Read diode structure and (b) its associated electric field. (Courtesy of W. T. Read. [43].)

more difficult to discuss in that the avalanche and “drift” regions are no longer distinct.

Read made two basic assumptions. (1) He assumed that for electric fields greater than some critical field (in the vicinity of 5000 volts/cm for silicon) carrier velocities “saturate” and become constants, independent of the magnitude of the field. (2) He assumed that high-field charge multiplication in solids is predictable by the approach of McKay (see Appendix) and in particular that $\alpha$ equals $\beta$, and that the current grows without bound when

$$\int_0^L \alpha dX \geq 1$$

[Both of these assumptions are questionable (see Appendix) but nevertheless lead to a reasonably accurate picture of “Read” diode operation.] Read then examined a structure shown in Fig. 20 in which a “hyperabrupt” impurity profile is employed to obtain a thin very high-field avalanche region which serves to inject charge, and a relatively long moderately high-field region which
serves as a drift space. The analogies with vacuum tube operation are obvious with the two major differences being the inherent delay in the avalanching cathode and the “limited” velocity of the carriers in the drift region, which is much lower than velocities employed in vacuum tube oscillators.

The reader who wishes a detailed quantitative picture of the physics involved is referred to Read’s paper [43]. Herein we give a simple qualitative picture that is intended to illuminate the two principal effects that are employed to obtain a microwave negative resistance. We turn first to the delay inherent in the avalanche process.

We consider first a thin (compared to the distance traveled by a carrier moving at $v_s$ in a half-cycle at the operating frequency) section of semiconductor. At some critical field $E_c$ where across this region

$$\int_0^L \alpha(E_c) \, dX = 1$$

the particle current through this region will begin to build up exponentially. For lower fields there is a stable value of current and for higher fields the exponential rate of buildup will increase even more. One can observe the delay inherent in such a “cathode” by applying a sine wave modulation in $E$ to such a region biased at $E_c$. In order to depict this delay, we neglect space charge effects and assume that the voltage across the region is related to the field by

$$V = \int_0^L E \, dX \equiv EL$$
That is, we approximate the high-field region by assuming $E$ independent of distance over the effective region of multiplication, and in phase with the applied voltage. Read treated more complicated cases and showed that they behave similarly.

These simplifying assumptions are incorporated in Fig. 21 to show the delay inherent in the avalanche process. This figure shows time snapshots of the voltage applied to the multiplying region and the simultaneous particle current through the region, that is, we ignore displacement current and concentrate on the transport current due to moving electrons and holes within the multiplying region. At time $t_0$ a step in voltage equal to $V_c$ is applied to the region and the particle current begins to flow. A sine wave of voltage is assumed superimposed on $V_c$, that is, $V = V_a \sin \omega t + V_c$. At time $t_1$, a quarter-cycle later the second snapshot shows that the particle current has continually grown from time $t_0$. Note that the applied voltage has peaked at $t_1$, but that at this time the transport current is still growing. At time $t_2$, after a half cycle, the applied voltage is again reduced to $V_c$ but the time snapshot shows that the particle current has continued to grow during this entire half cycle, since the applied voltage was always greater than $V_c$. A snapshot at $t_3$ after three quarters of a cycle shows a decrease in particle current since between $t_2$ and $t_3$ the applied voltage is less than $V_c$. Thus the particle current is a maximum at $t_2$ which occurs one quarter-cycle later than the peak in the applied voltage which occurred at time $t_1$. Whereas for illustrative purposes this delay was developed on a transient picture, that is, we assumed no current at time $t_0$ at which time a step function in voltage was applied, the same delay can be obtained analytically on a continuous basis. The effect of space charge, which was neglected, can be shown to reduce the delay in large signal operation and indeed provide one of the saturation mechanisms of the "Read" diode.

Having thus obtained a cathode which injects a particle current delayed $90^\circ$ from an applied voltage we now develop the delay inherent in a "drift" region. Following Read we assume that the field in this region always exceeds $E_s$, the velocity saturation field, and consequently that all carriers move with the same average velocity $V_s$. If a thin avalanche cathode is placed at one edge of this relatively long drift region we can describe, for small enough signals, the particle injection rate as a sine wave that is delayed $90^\circ$ from the applied voltage as has just been shown. This particle injection rate is shown as a function of time in the first column of Fig. 22.

Again employing a transient condition for simplicity we proceed to develop the delay inherent in the drifting of carriers. Column 2 contains a time snapshot of the charge density within a drift region, whose length $L$ is adjusted to be the distance traveled by the carriers in one half-period of the applied voltage wave, that is, $L/V_s = T/2$ where $T$ is the period of the applied voltage wave. Column 3 contains the total current per unit area through the region of length $L$ (measured
at its terminals) resulting from the carriers in transit within it. This current at any instant is simply obtained by summing all the charges in transit and dividing them by \( \tau \), the transit time, that is, \( I = Q_{\text{total}} / \tau \). As can be seen \( I \) peaks 90° later than the particle injection rate which was in turn delayed 90° from the applied voltage. This results in a total delay of 180° in \( I \) from the applied voltage \( V \) and is the source of the negative resistance of the Read diode.

### E. Power-Frequency Limitations

We consider first the Read diode. Here avalanche is assumed at one edge of the drift region. Thus we require that the field in the drift region of length \( L \) be limited by

\[
\int_0^L \alpha dX = 1
\]

or that \( V_m = E_c L \) is the maximum voltage that can be applied across this region. (This will in practice consist of a dc and an RF term.) The maximum charge that can be injected into the drift region is approximately that which drops \( E_c \) to \( E_s \) since in the drift region \( E \) does not exceed \( E_c \) in operation. Thus

\[
I_m \approx C_m E_c v_s
\]

where \( C_m = \varepsilon A / L \) = capacitance of the device at high reverse bias. Thus

\[
P_m X_c f_{\text{op}}^2 = \frac{E_c^2 v_s^2}{4\pi}
\]
where \( f_{op} \) is the optimum operating frequency (for which \( r = T/2 \)) and \( X_c = 1/C_m \omega_{op} \) or \( f_{op} = v_g/2L \).

These absolute limitations are of course above Read's suggested operating points and a comparison is shown in Fig. 23. A dashed line indicated the best experimental results to date.

Fig. 23. Absolute power-frequency limitations for Read diodes, Read's suggested operating points, and performance obtained to date.

Absolute limits on power again depend on the attainable impedance levels in microwave circuitry, and as for silicon transistors (see Section II) the attainable power is expected to show an \( f^{-2.17} \) dependence. As before, owing to the uncertainty of \( m \) we have plotted \( f^{-2} \) in the state-of-the-art plot of Fig. 19.

The other versions (\( pn \) and \( pin \)) of IMPATT oscillators are more difficult to analyze. Here the peak field considerably exceeds \( E_c \) over a reasonable extent of the device. It appears unlikely, however, that the peak voltage will exceed \( 2V_c \) in any of these devices. One can also envision currents perhaps a factor of 2 larger than those of the Read diode owing to the two types of carriers involved. It thus appears that the \( P_mX_c \) product for some variations of these devices can be as much as a factor of 4 higher than that of the Read diode. The practicality of such devices is currently being studied.

It appears then that the entire class of IMPATT oscillators will have an output power capability falling off approximately as \( 1/f^2 \), with the "approximately" depending upon the exact nature of the avalanche process. It is hard
VII. SUMMARY AND CONCLUSION

Pertinent data has been collated summarizing the state of the art of microwave solid state power generation. An attempt has been made to rationalize this data in terms of attainable impedance levels in microwave circuitry and the limitations imposed by device physics. These limitations were shown to follow from dielectric breakdown (avalanche multiplication) which determined

Fig. 24. Summary of the state of the art of existing solid state microwave sources.
the maximum voltage that could be impressed on a device, and upon the "limited" velocity of carriers in solids, except in the case of the tunnel diode. For it the principal limitation was shown to be due to the fractional bandgap voltage that could be applied. In order to avoid detail, the simplified theory of McKay [53] was employed in the study of avalanche limitations. While more precise treatments can be carried out, it is felt that this approach gives reasonable accuracy for the generalized approach taken herein.

A summary of the state of the art of solid state microwave power sources is presented in Fig. 24. The solid portions of the curves indicate regions in which operation has actually been obtained while the dashed portions indicate operating ranges allowed by device physics but not yet achieved. The curve plotted for tunnel diodes is for p-type GaAs rather than the n-type GaAs since the latter, although slightly higher, would be all dashed in Fig. 24. On the basis of high-power, high-frequency capability the IMPATT devices already exceed all other solid state sources. It is hoped that the techniques developed in this paper for interdevice comparison have provided a framework suitable for the microwave or device engineer to decide whether or not a "hot" new result represents a real improvement in the light of the capabilities of existing sources of microwave power.

The answer to the question of absolute power limitations of individual semiconductor devices resides in the attainable impedance levels in microwave circuitry. This question must be studied separately for each device owing to the different boundary conditions imposed by the individual device physics on circuit requirements.

No survey can hope to be accurate on the date of its appearance in the literature owing to the unavoidable time lag in the publishing process. This is particularly true in the field on which this chapter is based. The state-of-the-art curves are, to the best of the author's knowledge, inclusive of the best reported experimental results to Nov. 10, 1965.

APPENDIX

CARRIER VELOCITY SATURATION

The concept of saturation of carrier velocities in solids is fairly simple, at least when the transfer of carriers from a high-mobility valley to a low-mobility valley is not involved. (This mechanism has been invoked in explaining Gunn oscillations.) It is evidenced as a marked decrease in the dependence of carrier velocities on field magnitudes for fields in excess of from 5000 to 15,000 volts/cm for the curves of Fig. A-1. As is well known, the motion of individual carriers in solids is not limited to the direction of an applied external field. Rather carriers are continually scattered by electron-electron collisions, electron-lattice collisions, and electron-impurity atom collisions. A small externally applied electric field can then be regarded as supplying a net drift velocity to a
group of carriers whose individual (thermal) velocities are of the order of $10^6$ to $10^7$ cm/sec at room temperature. For steady state conditions of current flow the carriers on the average must lose as much energy in the collisions as they gain in the accelerating process between collisions. In the low-field, linear mobility region, the carrier energies are insufficient to excite the high-energy, optical mode, lattice vibrations and instead interact with the lower-energy acoustical vibrations. By invoking conservation of momentum considerations it can be shown [54] that the acoustical interactions are good randomizers of the velocities of the electron population but are relatively ineffective in removing energy from the electrons. Thus as the applied external electric field is increased, the average velocity and energy of the carriers (and consequently the current through the sample) increase. The optical phonons, on the other hand, are quite effective in the energy transfer process and when the electrons “heat up” sufficiently to excite optical mode interactions, these interactions act to “limit” their energy and hence their velocity. This “scattering-limited velocity” is thus a characteristic of the crystal involved, and its magnitude is a function of the effective mass of the carriers and the energies of the optical modes of lattice vibrations. Impurity scattering can dominate lattice scattering at low temperatures, or in some cases at room temperature, but this mechanism is not dominant in device limitations considered herein and will not be discussed further.

Fig. A-1. Carrier velocity saturation at high electric fields. After Gunn [56] and Prior [57].
Although the qualitative arguments presented above (and a large body of supporting theory not presented) provide a firm theoretical basis for scattering-limited velocities, experimental confirmations have proved exceedingly difficult to obtain [55]. Indeed the "Gunn effect" [29] was discovered in 1963 in an attempt by J. B. Gunn to measure scattering-limited velocities in GaAs. On the basis of our discussion one might expect to simply place a voltage across a semiconductor sample, measure the current through the sample, and from $J = qnv$ with $q$ the charge of the electron, $n$ the density of carriers of the material, and $v$ the velocity of the carriers, infer the velocity as a function of electric field $E$. Unfortunately things are not so simple [54]. Under an applied dc voltage the crystal lattice heats as well as the electron population. This can change the carrier density $n$ as well as scattering cross sections, and particularly for high fields forces the use of short (nanosecond) pulses. With such short pulses the presence of displacement current then becomes a complicating problem. Also, contacts that are ohmic at low fields can become nonohmic at high fields and minority carrier injection can becloud the picture. Charge multiplication effects further complicate matters at extremely high fields or at moderate fields in the case of impact ionization of impurities. Some of the data of Gunn [56] and Prior [57] on the velocities of holes and electrons in Ge and electrons in Si as a function of electric field are presented in Fig. A-1. These curves have been obtained from their published $J$-$E$ data by simply assuming that the carrier density does not vary with field. The curve for hole velocity in Si is conspicuous by its absence. If one attempts to obtain a $v$-$E$ curve for holes in Si from Prior's data by assuming $p$ constant, one obtains degrees of saturation varying with the background doping of the material. The variations range from no indicated velocity saturation in intrinsic material to modest saturation in 3 $\Omega$-cm material. Thus the saturation of hole velocities in Si is still an open question. Nevertheless, it is apparent that with the possible exception of holes in silicon, we can regard it as a reasonable simplification of experimental data to assume that for fields in excess of some value labeled $E_s$ (5000 volts/cm for electrons and 10,000 volts/cm for holes in Ge; 15,000 volts/cm for electrons in Si) carrier velocities are "saturated" and can be labeled $v_s$, a constant, independent of field. This is the approximation employed by Johnson (Section I, F). Whereas only the symbol $v_s$ appears in the text, $v_s = 10^7$ cm/sec is assumed for electrons in silicon for the calculations of Figs. 4, 8, and 23.

Avalanche Multiplication in Solids

The previous section has described the basic mechanisms involved in carrier velocity saturation in solids due to lattice scattering. For applied electric fields considerably in excess of those necessary for the onset of velocity limiting, some carriers can gain enough energy to create a hole electron pair before being "scattered" by an optical phonon. This in general requires that carrier
energies be greater than the energy gap of the material. This phenomenon drastically affects device operation and will now be discussed in some detail.

The study of avalanche multiplication in solids [52] has been pursued by many authors both theoretically and experimentally. We present here the simplified approach of McKay [53]. It is fairly straightforward and gives physical insight into the nature of avalanche multiplication. It also serves as an introduction to the theory of the "Read" avalanche transit time diode and was employed by Read in his theoretical study. Although recent work [58] has produced improvements in McKay's model, his approach gives reasonable answers in our range of interest and will be employed throughout this paper. The extension to a more general model is straightforward but not worth the extra complication for our approximation.

![Diagram](image)

**Fig. A-2.** The geometry assumed for the calculation of avalanche multiplication.

McKay's approach follows Townsend's [59] theories of avalanche breakdown in gases. The basic mechanism is that of charge carriers (either holes or electrons) being accelerated by an applied electric field until they gain enough energy to create hole-electron pairs. This is then a cumulative process in that both holes and electrons can create new hole-electron pairs and thus provide positive feedback no matter which carrier initiates the process. The ionization rate for electrons is \( \alpha \), defined as the average number of hole-electron pairs produced by an electron per centimeter traveled against the direction of the electric field \( E \); \( \beta \), the ionization rate for holes, is similarly defined as the average number of hole-electron pairs produced by a hole per centimeter traveled in the direction of the electric field \( E \). For simplification McKay assumed that \( \alpha = \beta \), but the results do not critically depend on this assumption. He further assumed that \( \alpha \) is solely a function of the electric field; that recombination can be neglected in the multiplying region; and that the mutual interaction between conduction electrons can be ignored. The space charge of the carriers is also neglected. A thin high-field multiplying region of thickness \( W \) is then considered as shown in Fig. A-2. \( E \) is assumed to be a function of \( x \) only and we consider a number of electrons \( n_0 \) injected into the region at \( x = 0 \) per unit area. Let \( n_1 \) be the number of electrons produced by electrons or holes between 0 and \( x \) per unit area, and \( n_2 \) the number produced between \( x \) and \( W \).
per unit area. Then between \( x \) and \( x + dx \), the number produced per unit area is given by

\[
dn_1 = (n_0 + n_1) \alpha dx + n_2 \alpha dx = n \alpha dx
\]  

where \( n = n_0 + n_1 + n_2 \). Integrating with the boundary conditions \( n_0 = 0 \) at \( x = 0 \) and \( n = n_1 + n_0 \) at \( x = W \) yields

\[
1 - \frac{1}{M} = \int_0^W \alpha dx
\]

where \( M = n/n_0 \) is the multiplication factor. When the integral above becomes unity, \( M \) becomes infinite and “breakdown” occurs. That is,

\[
M = \infty \quad \text{for} \quad \int_0^W \alpha dx = 1
\]

For values of \( E(x) \) such that the integral is less than unity a steady state value of current exists. For values of \( E(x) \) such that the integral is greater than or equal to unity, the current through the region grows exponentially until our starting assumptions are violated. These considerations are discussed further in Section VI.

McKay [53] obtained the field dependence of \( \alpha \) in silicon as shown in Fig. A-3. More recent work [58] has indicated a slightly higher dependence of \( \alpha \) on

---

**Fig. A-3.** Ionization rate versus electric field. (Courtesy of K. G. McKay.)
E but in addition has demonstrated that $\alpha \neq \beta$. Nevertheless, McKay's approach will be assumed here for simplicity and the nature of the argument will be seen to be little affected by the exact field dependence of $\alpha$ or $\beta$ or the fact that they are not equal.

ACKNOWLEDGMENTS

Many people assisted in the compilation of these results. In particular I would like to thank A. G. Chynoweth, R. Edwards, J. C. Irvin, R. M. Ryder, C. B. Swan, and A. R. Uhlir. I am further indebted to R. M. Ryder for a critical review of the manuscript and for many illuminating discussions.

REFERENCES

5. Texas Instruments Advertisement.
15. RCA advertisement.
33. Hakki, B. W., private communication.


I. INTRODUCTION

Much of the recent progress in our understanding of the world around us was achieved through the ability to detect very weak signals, for example, from outer space, from nerve fibers, or from the ocean floor. To detect these weak signals, one requires very sensitive receivers, and hence the ability to amplify these signals without, at the same time, introducing any appreciable amount of noise.

In the microwave frequency range, this need was especially acute, in particular for radio astronomy, satellite communications, and long range radars. This need was met by the improvement of well-established amplifiers such as the traveling-wave tube, the invention of new ones such as the maser, and the realization of low noise parametric amplifiers.

The most exciting developments in low noise amplifier technology in the recent years have been the maser and the parametric amplifier. Although the principle of the maser [1] is relatively new, the principle of parametric amplification [2] has been known for a long time. However, the low noise

1 The idea for the maser was conceived by Dr. C. H. Townes in 1951. Similar ideas occurred independently to Professors A. Prokhorou, N. Basou, and J. Weber.
characteristic of the parametric amplifier had not been recognized until the late 1940's [3] and the realization of low noise parametric amplifiers [4–7] was further delayed to 1957 when the high quality semiconductor pn junction diode (varactor) became available.

The parametric amplifier has made great progress in the past seven years. Among nonrefrigerated amplifiers, the parametric amplifier is the most sensitive amplifier in existence. The effective input noise temperature of varactor parametric amplifiers can be greatly reduced by refrigerating the amplifier down to liquid nitrogen temperature [8] and further to liquid helium temperature [9, 10].

The liquid helium temperature parametric amplifier and the maser are comparable in noise performance and are practically noiseless. Since these amplifiers add so little noise to desirable signals, the antenna and the input transmission line often contribute more noise to the receiver than these amplifiers. Therefore, for many high sensitivity receiver applications, the ultimate low noise temperature such as that of the maser or liquid helium temperature parametric amplifier may not be needed.

The varactor parametric amplifier can be operated at convenient temperatures, from room temperature down to liquid helium temperature, and offers versatility for system design. One of the convenient temperatures for open dewar systems may be liquid nitrogen temperature. The liquid nitrogen temperature parametric amplifier is economical to operate and yet offers high sensitivity; therefore, it is suited for most high sensitivity receiver applications. For closed-cycle refrigerator systems ambient temperature can be lowered down to about 10°K without losing the advantages of medium temperature refrigerators such as compactness, reliability, and lower cost. This is a very important factor for most practical system applications which require reliable and economical operation but no ultimate low noise performance. This is one of the advantages of cooled parametric amplifiers over the maser.

The percentage bandwidth of nondegenerate amplifiers with 20-db gain can be 20% or even higher in the low microwave frequency range. The bandwidth of high microwave frequency amplifiers is still limited to around 5–10% by the low self-resonant frequency of varactors in existing encapsulations; however, a bandwidth similar to that of low microwave frequency amplifiers may be obtained in the future. This is another advantage of parametric amplifiers over the maser.

Most parametric amplifiers in operation are of the circulator type. The gain and transmission characteristics are susceptible to pump source and circuit impedance variations. However, advanced technologies have reduced these problems to such a level that the amplifiers have been used satisfactorily [11, 12] in radio observatory radiometric systems which require extremely high stability.
COOLED VARACTOR PARAMETRIC AMPLIFIERS

With these advantages and advanced technologies the future of cooled parametric amplifiers in high sensitivity receiver systems such as satellite ground stations, radio astronomy, and long range radars appears very optimistic. With the advent of a simple solid state pump source and compact, reliable refrigerator, the cooled parametric amplifier may further be explored in new space microwave systems applications.

This chapter will discuss the characteristics of cooled parametric amplifiers, including effective input noise temperature, gain-bandwidth product, and gain stability. Special attention is given to shot noise and varactor heating due to pump power dissipation, which are detrimental to the noise performance of cooled parametric amplifiers. Careful selection of the varactor, circulator, and refrigerator is extremely important to achieve a reliable, stable low noise amplifier. These device considerations will be briefly discussed in Section III. Several cooled amplifiers will be described and the characteristics of these amplifiers will be presented in Section IV.

II. CHARACTERISTICS OF COOLED AMPLIFIER

A. GENERAL THEORY

Manley and Rowe [13] have derived a very general set of equations relating power flowing into and out of an ideal nonlinear reactance element. These relations are a powerful tool in understanding the principle of parametric amplification and predicting the ultimate maximum power gain that can be obtained. In the parametric amplifier, power is supplied by an RF source, whose frequency \( f_p \) is usually higher than that \( f_1 \) of a signal to be amplified; it is usually named the pump source. The pump power is much larger than the signal power. Besides these frequencies at least a third frequency \( f_2 \), which is either the sum of (upper sideband) or the difference of (lower sideband) \( f_p \) and \( f_1 \), is needed to achieve the amplification. The Manley–Rowe power relations predict that if the third frequency is the upper sideband, the amplification is only possible by converting the frequency from \( f_1 \) to \( f_2 \) and its power gain does not exceed a frequency ratio \( f_2/f_1 \); and that if the third frequency is the lower sideband, arbitrary gains can be obtained at both \( f_1 \) and \( f_2 \) with and without frequency conversion and the available power at \( f_2 \) at the nonlinear capacitance is \( f_2/f_1 \) times greater than that at \( f_1 \). When the output frequency \( f_{out} \) is \( f_2 \), the amplifier is a two-port device; when \( f_{out} = f_1 \), it is usually a one-port device.

Parametric amplifiers are usually subdivided into three basic types:

1. upper sideband up-converter \( (f_{out} = f_p + f_1) \);
2. lower sideband up-converter \( (f_{out} = f_p - f_1) \);
3. one-port amplifier \( (f_{out} = f_1) \).
In the one-port amplifier, since the output and input signals are of the same frequency and are at the same port, a nonreciprocal device such as a circulator must be used to separate the output from the input. Therefore, it is usually called a circulator-type amplifier. When $f_1$ and $f_2$ are widely separated and the input circuit of the amplifier does not pass $f_2$, the lower sideband power is completely dissipated in the amplifier. In this case the amplifier is named a circulator-type nondegenerate amplifier and its lower sideband frequency is usually called an idler frequency. As the signal frequency $f_1$ moves closer to half the pump frequency $f_p$, the idler frequency $f_2$ moves closer to the signal, and eventually both the signal and idler channels overlap partly or in whole. Under such conditions, the amplifier must have enough bandwidth to encompass both the signal and its idler. When the input circuit passes both the signal and idler bands and the input termination is common to both, the amplifier is called a degenerate amplifier.

The theoretical minimum noise figures of the three basic types of the amplifier are identical [14]. However, a limited gain obtainable from the upper sideband up-converter and a narrow band characteristic of the lower sideband up-converter under the minimum noise figure condition dictate the one-port amplifier as the best candidate for extremely high sensitivity system applications. Since the degenerate amplifier is not suited for communications applications except for some radiometer applications, this section will discuss only the circulator-type nondegenerate amplifier.

A simplified amplifier circuit is shown in Fig. 1. In the figure a varactor is represented by the elastance $S(t)$ and the series resistance $R_s$ and other parasitic impedance elements are absorbed in the external signal and idler circuits. $\omega_1$ and $\omega_2$ represent the ideal filters which pass only the signal frequency.
COOLED VARACTOR PARAMETRIC AMPLIFIERS

band and idler frequency band, respectively. If the elastance $S(t)$ of the varactor pumped at the frequency $f_p$ (angular frequency $\omega_p$) is expressed by the following equation

$$S(t) = S_0 + S_1 \cos \omega_p t + \cdots$$

the junction is characterized by the equation [14]

$$
\begin{bmatrix}
  e_1 \\
  e_2^*
\end{bmatrix} =
\begin{bmatrix}
  S_0 / j\omega_1 & -S_1 / j2\omega_2 \\
  S_1 / j2\omega_1 & -S_0 / j\omega_2
\end{bmatrix}
\begin{bmatrix}
  i_1 \\
  i_2^*
\end{bmatrix}
$$

where the star indicates the complex conjugate and other symbols are shown

in Fig. 1. (A factor 2 in the denominators of the off-diagonal elements of the impedance matrix in Fig. 2 disappears if $S(t)$ is defined by

$$S(t) = S_0 + 2S_1 \cos \omega_p t + \cdots$$

If we assume that the series resistance $R_s$ is the only dissipative element of the varactor, which is a realistic assumption for most of the low noise amplifiers, and define the dynamic quality factor [15] $\bar{Q}_n$ by the following equation,

$$
\bar{Q}_n = \frac{|S_1|}{2\omega_n R_s} = \gamma \frac{\bar{Q}_n}{2}
$$
where
\[ \gamma = \frac{|S_1|}{S_0} \]
and
\[ Q_n = \frac{S_0}{\omega_n R_s} \]

Eq. (2) is rewritten as follows:
\[ \begin{bmatrix} e_1 \\ e_2^* \end{bmatrix} = \begin{bmatrix} \frac{S_0}{j\omega_1} & jQ_2 R_s \\ -jQ_1 R_s & -\frac{S_0}{j\omega_2} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2^* \end{bmatrix} \quad (4) \]

By solving the above equation an equivalent circuit of the amplifier can be determined; the equivalent circuit is shown in Fig. 2. Since thermal noise generated in the series resistance \( R_s \) of the varactor and in the idler circuit is the dominant noise source, these noise generators are included in the circuit to aid the calculation of the effective input noise temperature. In the figure \( Z_{11} \) and \( Z_{22} \) include most passive circuit elements including \( S_0 \), but \( R_s \) is left independent since it plays an important role in determining amplifier characteristics.

From the equivalent circuit the reflection power gain is derived as
\[ G = \frac{|Z_{11}^* - 1 + \frac{Q_1 Q_2}{1 + Z_{22}^*/R_s}|^2}{1 + \frac{Q_1 Q_2}{1 + Z_{22}^*/R_s}} \quad (5) \]

In the above calculation the circulator was assumed to be ideal. The term
\[ -\frac{Q_1 Q_2}{1 + Z_{22}^*/R_s} \quad (6) \]
is a dynamic impedance presented by the pumped varactor at the signal port.

1. **Effective Input Noise Temperature (Thermal Noise)**

The effective input noise temperature of the amplifier at the center band frequency due to thermal noise sources, excluding the circulator insertion loss and assuming \( R_s \) and \( R_d \) are at the same ambient temperature, is derived from Fig. 2 as the following equation [16],
\[ T_e = \left(1 - \frac{1}{G}\right) \frac{1 + \left(\frac{f_1}{f_2}\right) \frac{Q_1 Q_2}{1 + R_d/R_s}}{1 + \frac{Q_1 Q_2}{1 + R_d/R_s}} T_d \quad (7) \]
where \( T_d \) is the ambient temperature of the amplifier, and \( R_1 \) and \( R_2 \) are assumed to be negligibly small. The effective input noise temperature is directly proportional to ambient temperature. This suggests that cooling the amplifier is the best method of improving the noise temperature of the amplifier. By rewriting Eq. (7), the noise measure [17] is obtained as

\[
M = \frac{1 + \frac{f_2}{f_1} \frac{Q_1 Q_2}{1 + R_1/R_s} T_d}{\frac{Q_1 Q_2}{1 + R_1/R_s} - 1} \frac{T_0}{T_d}
\]  

(8)

where \( T_0 \) is the standard temperature (290\(^\circ\)K). Therefore, the noise measure is independent of the gain and is determined by the varactor, circuit design, and ambient temperature. The noise measure is minimized when the idler circuit is loaded by the series resistance of the varactor alone \((R_t = 0)\).

Since \( Q_2 = Q_1(f_1/f_2) \), the optimum idler frequencies for low noise operation are obtained from the equation [16]

\[
\frac{f_2}{f_1} = \left(1 + \frac{Q_1^2}{1 + R_t/R_s}\right)^{1/2} - 1 \quad \text{for } R_t \neq 0
\]  

(9)

and

\[
\frac{f_2}{f_1} = (1 + Q_1^2)^{1/2} - 1 \quad \text{for } R_t = 0
\]  

(10)

For a given varactor and idler loading factor \( R_t/R_s \), to achieve a minimum noise temperature, the passive input resistance of the amplifier must be adjusted to satisfy the condition

\[
\frac{R_g}{R_s} = \frac{[\sqrt{(G)} + 1]^2}{G - 1} \left(\frac{Q_1 Q_2}{1 + R_t/R_s} - 1\right)
\]  

(11)

2. Refrigeration

According to Eq. (7) the effective input noise temperature is directly proportional to the ambient temperature of the amplifier. This is based on the assumption that the series resistance \( R_s \) and capacitance modulation factor \( \gamma \) or the dynamic quality factor \( Q \) of the varactor remain constant from room temperature to cryogenic temperature. The experimental investigations [9, 18–20] indicate that such varactors do certainly exist. The most commonly used varactor for refrigerated operation is the varactor made from heavily doped gallium arsenide. Most commercially available silicon varactors cannot be used for refrigeration below 100\(^\circ\)K since the series resistance increases very rapidly and below 20\(^\circ\)K they cease to show any variable capacitance. However, if the silicon is heavily doped and its doping profile is carefully controlled, the varactor can be refrigerated to at least 4\(^\circ\)K without any degradation of its noise
temperature from the estimated value [20]. More detailed discussion on the varactors for cooled parametric amplifiers will be given in Section III.

In the effective input noise temperature derivation we have assumed thermal noise as the dominant noise source and neglected other noises. However, if average forward and reverse currents are allowed to flow, noise generated in these currents cannot be neglected. Microplasma noise can be easily detected for most amplifiers; therefore, the reverse current has been carefully avoided in the past. However, shot noise due to the forward current is usually insignificant for amplifiers whose noise temperature is higher than 100°K. As the ambient temperature of the amplifier is lowered and thermal noise is decreased, shot noise becomes an important noise source. The excess noise temperature due to shot noise is about 3°K/μA for a 4-Gc amplifier with a gallium arsenide varactor, and about 25°K/μA with a silicon varactor. Therefore, cooled amplifiers should be carefully designed in order to avoid these currents.

At extremely low ambient temperatures, the system noise temperature is mainly determined by the insertion loss in the input transmission line and the increase in noise temperature due to the use of poor quality diodes seems

![Graph](image-url)

**Fig. 3.** Theoretical minimum noise temperatures of parametric amplifiers and experimental results of seven nondegenerate amplifiers ranging from 1–8 Gc. The points × show the measured over-all noise temperature at 300°K ambient temperature; ◊ at 77°K; ⊶ at 4°K. The points • indicate the measured noise temperature contribution of varactor and microwave circuits. (Bracketed numbers indicate references.)
to be practically insignificant. However, the temperature rise in a poor varactor due to large pump power dissipation increases the achievable noise temperature proportional to the junction temperature. For this reason, a high quality varactor must be used even in the cooled amplifier. Effects of shot noise and varactor heating on the effective input noise temperature will be further discussed later in this section.

Another important noise source in the cooled amplifier is the high temperature load resistances in spurious sideband frequency networks. The multiport parametric amplifier theory [21] indicates that the noise temperature of the amplifier is higher if the number of ports is larger. This means that any unnecessary sideband responses should be eliminated in order to achieve the minimum noise temperature. One of the sidebands which must be carefully investigated is the upper-sideband frequency. As long as this is constrained within the amplifier, there is not much harm to noise performance of the cooled amplifiers. However, if the upper sideband circuit couples to an external room temperature load through, for example, the pump circuit, this results in a considerably higher effective input noise temperature than the estimated value.

Figure 3 shows the theoretical minimum noise temperature [Eq. (7)] of parametric amplifiers at ambient temperatures of 300°, 77°, 15°, and 4.2°K. The dynamic quality factors of the varactors including the input circuit losses are assumed to be 20, 16, 12, 8, and 6 at 1 Gc, 2 Gc, 4 Gc, 6 Gc, and 8 Gc, respectively. The experimental results of seven amplifiers are plotted in the figure. The differences between the theoretical minima and the experimental results are mainly due to the choice, for practical reasons, of pump frequencies that are much lower than the theoretical optima.

An accurate value for the theoretical effective input noise temperature of the amplifier is difficult to estimate, since $\bar{Q}_1 \bar{Q}_2 / (1 + R_2 / R_s)$ cannot be accurately estimated from the passive measurements of $\bar{Q}_1$ and $\bar{Q}_2$. This is because the actual pump voltage and wave shape under operating conditions cannot be measured at microwave frequencies. In order to solve this difficulty, the value of $\bar{Q}_1 \bar{Q}_2 / (1 + R_2 / R_s)$ under operating condition is determined from a measured gain $G$, $R_2 / R_s$, and Eq. (11) as

$$\bar{Q}_1 \bar{Q}_2 / (1 + R_2 / R_s) = \frac{R_s}{G - 1} \left( \sqrt{\frac{G - 1}{R_s}} \right)^2 + 1$$

(12)

If the amplifier is carefully designed and operated to avoid the above mentioned spurious noise, the measured noise temperatures are usually within the measurement errors of the values estimated from Eqs. (7) and (12).

3. Bandwidth

The gain-bandwidth limitation of parametric amplifiers has been discussed by several authors [22, 25–29]. Among these papers only that of De Jager [22]...
Michiyuki Uenohara considers the effects on the gain-bandwidth product of the amplifier of both the parasitic elements in the varactor package and the isolating filters between the signal and idler circuits. The gain-bandwidth product of the amplifier with single-tuned signal and idler circuits for large gain is given [22] by

$$\frac{BQ_m}{f_{10}}g = \frac{2}{(2+q^2+1/q^2)^{1/2}}$$  \hspace{1cm} (13)$$

where $B$ is the 3-db bandwidth, $f_{10}$ is the center signal frequency, $g$ is the voltage gain,

$$Q_m = \frac{2}{\gamma} \left( \frac{d_1 d_2 f_{10}}{f_{20}} \right)^{1/2}$$  \hspace{1cm} (14)$$

and

$$q = \frac{Q_1 R_s}{R_r + R_s} \left[ \frac{d_2 (1/q^2)^3}{\sqrt{2} f_{20}} \right]^{1/2}$$  \hspace{1cm} (15)$$

where $d_1$ and $d_2$ are the slope factors [22] of the signal and idler circuits respectively, and $f_{20}$ is the center idler frequency. The gain-bandwidth product can be improved by compensating the input and idler circuits. The maximum limiting flat bandwidth for signal circuit compensation only is given [22] by

$$\frac{BQ_m}{f_{10}}\ln g = \frac{\pi ((1+3q^2)^{1/3} - 1)}{2}$$  \hspace{1cm} (16)$$

In order to achieve the maximum possible gain-bandwidth product, $Q_m$ should be as small as possible and $q$ should be the optimum value.

The slope factors $d_1$ and $d_2$ are determined by the signal and idler circuit configurations as well as the parasitic elements of the varactor. The minimum value of $(d_1 d_2)$ is four for the single-varactor configuration, and it can approach unity for the balanced-varactor configurations [30]. As long as $(d_1 d_2)$ remains constant, the higher the idler frequency, the larger will be the resulting gain-bandwidth product. However, beyond the self-resonant frequency $f_d$ of the varactor, $(d_1 d_2)$ increases and the gain-bandwidth product no longer increases monotonically. The maximum gain-bandwidth product is achieved when the idler frequency is approximately $\sqrt{3} f_d$ for the single-varactor configuration [22], and for the balanced-varactor configuration when the idler frequency is at, or slightly lower than, the self-resonant frequency [26].

The self-resonant frequency of practical microwave varactors ranges from 6 Gc to 12 Gc. Therefore, the bandwidth of the amplifier in the higher microwave frequency range is seriously degraded by poor varactor package design. Figure 4 shows the percentage bandwidths of single-varactor and balanced-varactor amplifiers with 20-db gains for single-tuned, double-tuned, and maximally compensated signal circuits. The self-resonant frequencies of the varactors are fixed at 10 Gc, and the dynamic quality factors of the varactors are assumed to be the same as those in Fig. 3.
COOLED VARACTOR PARAMETRIC AMPLIFIERS

Fig. 4. Theoretical percentage bandwidth of single-varactor and balanced amplifiers with 20-db gain for single-tuned, double-tuned, and maximally compensated signal circuits.

The gain-bandwidth product can further be improved by stagger-tuning several circulator-type amplifiers with a small degradation in noise performance.

Some experimental results are tabulated in Table I.

Table I
PERCENTAGE BANDWIDTH OF SINGLE-VARACTOR AND BALANCED-VARACTOR AMPLIFIERS

<table>
<thead>
<tr>
<th>$f_1$ (Gc)</th>
<th>$G$ (db)</th>
<th>Bandwidth (3-db points, %)</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-Varactor Amplifier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>15&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Single stage, double-tuned signal circuit [31]</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>20&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Two stages, double-tuned signal circuits [32]</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>15</td>
<td>Single stage, double-tuned signal circuit [33]</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>4.6</td>
<td>Single stage, double-tuned idler circuit [34]</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>4.6</td>
<td>Single stage, double-tuned idler circuit [35]</td>
</tr>
<tr>
<td>Balanced-Varactor Amplifier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>21&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Single stage, double-tuned signal circuit [31]</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>40</td>
<td>Four stages, compensated signal circuits [36]</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>6.5</td>
<td>Single stage, double-tuned signal circuit [30]</td>
</tr>
</tbody>
</table>

<sup>a</sup>The bandwidths quoted are for maximally flat characteristics ($\pm 0.2$ db).
4. **Gain Stability**

The reflection power gain of the circulator-type amplifier is given by Eq. (5). For the large gain condition the gain is dominantly determined by the denominator of Eq. (5). Therefore, in order to maintain constant gain, \( Z_{11} \) and \( Q_1 Q_2 R_s^2/(R_s + Z_{22}) \) must be maintained constant. Variation in \( Z_{11} \) is primarily caused by circulator impedance and varactor impedance variations and secondarily by antenna and passive amplifier circuit impedance variations. The active amplifier impedance \( -Q_1 Q_2 R_s^2/(R_s + Z_{22}) \) is mainly affected by pump power level and frequency and varactor capacitance variations and secondarily by passive idler and pump circuit impedance variations. Fluctuations in ambient temperature and mechanical vibration cause changes in all circuit impedances and hence fluctuations in the amplifier gain. Therefore, the amplifier circuits should be made rigid, the circulator should be connected as close as possible to the varactor in order to minimize impedance variations in the interconnecting transmission line due to mechanical vibrations, and large variations in ambient temperature avoided.

In the cooled amplifier ambient temperature is practically constant; therefore, if the circulator is also refrigerated, the pump source variation is the main source of gain variation. If the circulator remains at room ambient, not only does its insertion loss contribute considerable thermal noise to the amplifier but also impedance variations in the interconnecting transmission line and circulator itself cause gain variation. In the liquid nitrogen temperature amplifiers the difference in thermal noise contribution between the room temperature and cooled circulators is insignificant, considering the fact that the characteristics of cooled circulators are often inferior to those of room temperature circulators. However, the gain stability requirement often dictates the need of cooling the circulator.

For a low noise amplifier with about 20-db gain, the gain changes by about 0.1 db for a 0.01-db change in pump power. This gain sensitivity can be reduced, as seen from Eq. (5), by reducing the gain per stage and using several cascaded stages in order to achieve the desired gain. As discussed in Section II,B the self-bias technique cannot be used for the first-stage amplifier of an extremely low noise receiver to improve the gain stability, since the forward current contributes shot noise. However, this technique can be applied to later stages with minor deterioration in over-all noise performance.

The pump frequency variation has less effect than the pump power variation on the gain stability. However, if the pump circuit is not broad enough, which is often the case, the pump frequency must be well stabilized to maintain the effective power level constant. For broad band amplifiers, which have highly compensated circuits, it is extremely important to stabilize both the pump power and frequency to maintain a constant transmission characteristic of the amplifier.
B. Shot Noise

In Section II.A.2 a brief discussion was given on shot noise, which cannot be overlooked for cooled parametric amplifiers. In this section the effective input noise temperature including shot noise will be discussed on the basis of the paper by Josenhans [37].

In order to improve the gain-bandwidth product and gain stability of the amplifier operated at room temperature, the pump voltage excursion is usually extended into the forward conduction region of the diode. This forward conduction results in shot noise which is additive to the thermal noise previously discussed.

1. Effective Input Noise Temperature

An approximate equivalent circuit of the noisy varactor is shown in Fig. 5. Here we assumed that all the injected minority carriers make one pass across the barrier layers and are recombined before the pump voltage changes its polarity; therefore, the dc component $I_o$ of conduction current fully represents carrier fluctuation. It is also assumed that the effective shunt resistance $R_j=kT_d/eI_o$ is much larger than $S_0/\omega$, and the varactor is a linear-time-variant capacitor. The equivalent circuit for calculating the effective noise temperature is shown in Fig. 6. All the symbols except shot noise terms are identical to those in Fig. 2.

The high-gain approximation for the effective input noise temperature is found to be

$$T_e = T_d \left[ \frac{R_g}{R_j} + \frac{(S_0/\omega_1)^2}{2R_jR_g} + \frac{R_j + R_g + \frac{1}{2} (S_0/\omega_2)^2}{R_j} \frac{1}{f_1} \left( 1 \right) \frac{1}{f_2} \left( 1 \right) \frac{R_j}{R_g} \frac{(S_0/\omega_1)^2}{R_j} \right]$$

(17)
For low current condition, \( R_e + R_s \gg (S_0/\omega) L_2/R_f \), Eq. (17) is simplified as
\[
T_e \approx \left[ \frac{R_s + f_1(1 + \frac{R_e}{R_s})}{f_2} \right] T_d + \left( 1 + 2 \frac{f_1}{f_2} \right) \frac{(S_0/\omega L_2)^2 eI_0}{2R_s} \frac{1}{k} \tag{18}
\]

\( \omega_1 \) CIRCUIT \quad \omega_2 \) CIRCUIT

\[
\frac{\dot{G}_s R^2}{R_s + Z_{22}^*} \quad \frac{\dot{G}_s R^2}{R_s + Z_{22}^*}
\]

The first bracket in Eq. (18) is the effective input noise temperature of the amplifier with \( I_0 = 0 \). If we define the excess shot noise temperature as
\[
T_s = T_e(I_0) - T_e(I_0 = 0) \tag{19}
\]
then
\[
T_s \approx \left( 1 + 2 \frac{f_1}{f_2} \right) \frac{(S_0/\omega L_2)^2 eI_0}{2R_s} \frac{1}{k} \tag{20}
\]

When the assumptions used in this analysis are not valid, a much more sophisticated analysis is needed. However, the excess shot noise temperatures estimated from Eq. (20) are always optimistic values, and it is valuable to know how detrimental shot noise is to the sensitivity of cooled parametric amplifiers.

2. Experimental Results

The excess shot noise temperatures were measured at room temperature by Josenhans [37] for various varactors in a 4-Gc parametric amplifier pumped
at 23 Gc. Figures 7–9 show the excess shot noise temperatures as a function of rectified diode current, $I_0$, for epitaxial GaAs surface barrier varactors, epitaxial GaAs diffused varactors, and epitaxial Si diffused varactors, respectively. The capacitances and dynamic quality factors of the varactors, the

experimental shot noise slope $T_{\text{ex}}/I_0$, and the theoretical slope derived from Eq. (20) are tabulated in Table II. The noise measurement had a maximum error of $\pm 7^\circ\text{K}$, while the over-all experiment was repeatable within 10%. For the gallium arsenide varactors the difference between the theoretical and experimental results is insignificant. However, for the silicon varactors the experimental results are an order of magnitude larger than the theoretical results. The simple theory presented here does not adequately represent the excess noise present in silicon varactors. The reason for this is probably tied

Fig. 7. Measured shot noise temperature of a room temperature 4-Gc amplifier with a surface barrier GaAs varactor pumped at 23 Gc as a function of rectified current, $I_0$. 

The number 103
up with the long lifetime of carriers inherent in silicon. A minority carrier lifetime which is longer than the pump cycle would allow for minority carrier storage. The situation as described by Uhlir [39] may exist; that is, the injected carriers are stored for a finite length of time and then allowed to traverse the space charge layer again during the reverse portion of the pump cycle and possibly the following cycle. The net dc current contributed by multiple transit carriers is much smaller than that of single transit carriers. The net noise power contribution from multiple transit carriers depends on the degree of correlation between each transit. This kind of storage effect would negate the single-pass assumption used in this calculation and may account for the anomalous noise behavior of silicon varactors.

The results presented here indicate that even a small rectified diode current will degrade the noise performance of the cooled amplifiers considerably.
Fig. 9. Measured shot noise temperature of a room temperature 4-Gc amplifier with a graded $pn$-junction Si varactor pumped at 23 Gc as a function of rectified current, $I_0$.

Table II

<table>
<thead>
<tr>
<th>Varactor</th>
<th>Capacitance (pf)</th>
<th>$Q$ at 4 Gc</th>
<th>$T_s/I_0$ ($^\circ$K/μa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Operating bias</td>
<td>Zero bias</td>
<td></td>
</tr>
<tr>
<td>Surface barrier</td>
<td>0.464</td>
<td>0.555</td>
<td>17.2</td>
</tr>
<tr>
<td>No. 1</td>
<td>0.487</td>
<td>0.609</td>
<td>14.3</td>
</tr>
<tr>
<td>No. 2</td>
<td>0.523</td>
<td>0.664</td>
<td>19.5</td>
</tr>
<tr>
<td>No. 3</td>
<td>0.564</td>
<td>0.684</td>
<td>—</td>
</tr>
<tr>
<td>No. 4</td>
<td>0.595</td>
<td>0.790</td>
<td>17.5</td>
</tr>
<tr>
<td>Diffused GaAs</td>
<td>0.391</td>
<td>0.446</td>
<td>13.0</td>
</tr>
<tr>
<td>No. 1</td>
<td>0.525</td>
<td>0.620</td>
<td>19.4</td>
</tr>
<tr>
<td>No. 2</td>
<td>0.580</td>
<td>0.655</td>
<td>14.7</td>
</tr>
<tr>
<td>Diffused Si</td>
<td>0.45</td>
<td>0.51</td>
<td>14.2</td>
</tr>
<tr>
<td>No. 1</td>
<td>0.43</td>
<td>0.44</td>
<td>12.9</td>
</tr>
<tr>
<td>No. 2</td>
<td>0.59</td>
<td>0.66</td>
<td>13.5</td>
</tr>
<tr>
<td>No. 3</td>
<td>0.645</td>
<td>0.88</td>
<td>11.8</td>
</tr>
</tbody>
</table>
Special attention must be exercised for the amplifiers with silicon varactors in order to prevent any rectified current from flowing.

C. VARACTOR HEATING DUE TO PUMP POWER DISSIPATION

A part of the pump power supplied to the varactor is dissipated in the series resistance by the pumping current. This power is converted to heat, and the series resistance is heated up. If the dissipated pump power is large and the thermal resistance of the package is also high, temperature of the series resistance could be much higher than that of the helium cooled amplifier mount. This will increase the effective input noise temperature of the amplifier almost in proportion to the diode temperature. This section will discuss, on the basis of the paper by Garbrecht [40], the effect of varactor heating due to pump power dissipation on the effective input noise temperature. The optimum pump frequency and external idler loading factor for the minimum effective input noise temperature will also be derived.

1. Equivalent Thermal Resistance

Figure 10 shows a cross-sectional view of a varactor. The semiconductor substrate is usually a heavily doped $n^+$-type semiconductor with a thin layer of $n$-type epitaxial film. The space charge barrier is formed either by doping $p$-type impurity in the $n$-type epitaxial film ($pn$ junction) or by putting a proper metal to semiconductor contact (Schottky barrier). The series resistance is

![Fig. 10. Cross-sectional view of a typical mesa varactor.](image-url)
located primarily in the vicinity of the barrier and in the ohmic contacts. Therefore most heat is generated in these areas and flows out through the substrate, contacts, and package to the mount. To minimize a temperature rise, the semiconductor wafer should be encapsulated in a package with a low thermal resistance, and the package should be tightly mounted in the heat sink to minimize the contact thermal resistance.

Since the series resistance and its temperature are not uniformly distributed throughout the varactor, the available thermal noise power from the series resistance is given by

$$ P_{av} = kT_{eq}B $$

(21)

where $T_{eq}$ is the equivalent noise temperature derived from the following equation:

$$ T_{eq} = \frac{1}{R_s} \int dR / dx Tdx = T_d + \Delta T $$

(22)

In the above equation $R_s$ is the series resistance, $dR/dx$ is the resistance gradient along the shortest current path to the contact, $T$ is the physical temperature of the resistance at $x$, and $T_d$ is the ambient temperature of the amplifier. The integration is made from one end of the package to another.

The equivalent temperature rise $\Delta T$ is related by the pump power $P$ and the equivalent thermal resistance $R_{th}$ as

$$ \Delta T = PR_{th} $$

(23)

The equivalent thermal resistance $R_{th}$ is about two thirds of the thermal resistance of the varactor.

2. Effective Input Noise Temperature (Thermal Noise)

The effective input noise temperature of the amplifier with a pump heated varactor is simply derived by replacing $T_d$ of Fig. 7 by $T_{eq}$ when the idler circuit is loaded by the series resistance $R_s$ alone ($R_i/R_s = 0$). However, if the external idler load $R_i$ is not zero and its temperature is at $T_d$, $\tilde{v}_r^2$ in Fig. 2 must be replaced by $4kT_{eq}R_sB$ and the effective input noise temperature is derived as

$$ T_e = \left(1 - \frac{1}{G}\right) \left(1 + \frac{f_1}{f_2} \frac{\bar{\tilde{Q}}_1 \bar{\tilde{Q}}_2}{1 + R_i/R_s} \right) T_d + \left(1 + \frac{f_1}{f_2} \frac{\bar{\tilde{Q}}_1 \bar{\tilde{Q}}_2}{(1 + R_i/R_s)^2} \right) \Delta T $$

(24)

The excess noise temperature due to pump heating is

$$ \Delta T_e = \left(1 - \frac{1}{G}\right) \left(1 + \frac{f_1}{f_2} \frac{\bar{\tilde{Q}}_1 \bar{\tilde{Q}}_2}{1 + R_i/R_s} \right) \Delta T $$

(25)

In order to reduce $\Delta T_e$, $\Delta T$ must be reduced.
3. Optimum Pump Frequency

In Section II,A we have derived the optimum pump frequency for the minimum effective input noise temperature assuming $\Delta T=0$. However, if $\Delta T \neq 0$, Eqs. (9) and (10) can no longer be used to determine the optimum pump frequency, since $\Delta T$ increases in proportion to the square of the pump frequency. To determine a new optimum pump frequency we shall substitute a relation

$$\Delta T = \Delta T_0 \left(1 + \frac{f_2}{f_1}\right)^2$$

(26)

into Eq. (24). $\Delta T_0$ is the value of temperature rise when the varactor is pumped at the signal frequency to achieve the same capacitance variation (or the same $\gamma$) as that at the pump frequency. Equation (24) is rewritten as

$$T_e = \left(1 - \frac{1}{G}\right) \left[1 + \left(\frac{f_1}{f_2}\right)^2 \frac{Q_1^2}{1 + R/\epsilon R_a}\right] T_d + \left[1 + \left(\frac{f_1}{f_2}\right)^2 \frac{Q_1^2}{(1 + R/\epsilon R_a)^2}\right] \Delta T_0 \left(1 + \frac{f_2}{f_1}\right)^2$$

$$\frac{f_1}{f_2} \frac{Q_1^2}{1 + R/\epsilon R_a} - 1$$

(27)

Since this equation is too complicated to determine the optimum pump frequency analytically, $T_e/T_d$ is calculated numerically as a function of $f_2/f_1$

![Fig. 11. Normalized effective input noise temperature of nondegenerate parametric amplifier as a function of $f_2/f_1$ for various values of $\Delta T_0/T_d$. The dynamic quality factor $Q_1$ is assumed to be 10 and $R_\epsilon$ is negligibly small.](image-url)
COOLED VARACTOR PARAMETRIC AMPLIFIERS

for various values of $\Delta T_0/T_d$. The results for an amplifier with $\tilde{Q}_1 = 10$, $G = 20$ db, and $R_e = 0$ are shown in Fig. 11. If $\tilde{Q}_1$ is very large and $\Delta T_0/T_d$ is small, $f_2/f_1$ can be estimated from

$$\frac{f_2}{f_1} \approx \left( 1 + \frac{1}{\tilde{Q}_1^2} + \frac{1}{1 + \frac{R_e}{R_s}} \frac{\Delta T_0}{T_d} \right)^{-1/2}$$

(28)

It is clear from Fig. 11 that the optimum pump frequency for the pump heated amplifier is much lower than that of the amplifier with no pump heating.

![Diagram](image)

**Fig. 12.** Optimum idler signal frequency ratio $f_2/f_1$ as a function of $\tilde{Q}_1$ for various values of $\Delta T_0/T_d$. For the varactors with the same thermal resistance $R_{th}$, $f_2/f_1$ curves are shown by dotted lines.

Figure 12 shows the optimum idler to signal frequency ratio as a function of $\tilde{Q}_1$ for various $\Delta T_0/T_d$. The optimum pump frequency is rather a noncritical function of the dynamic quality factor $\tilde{Q}_1$. Figure 13 shows the minimum $T_e/T_d$ for various $\Delta T_0/T_d$.

The pump power needed to drive the varactor is also a function of varactor quality and is related as

$$P \propto \frac{1}{\tilde{Q}_1^2} \left( \frac{f_p}{f_1} \right)^2$$

(29)

The pump powers supplied to achieve the same capacitance modulation factor $\gamma$ for various varactors with different dynamic quality factors are related by the following equation:

$$\frac{P(\tilde{Q}_1)}{P'(\tilde{Q}_1')} = \left( \frac{\tilde{Q}_1'}{\tilde{Q}_1} \right)^2$$

(30)
Therefore, if the thermal resistances are the same for all varactors, $\Delta T_0$ and $\Delta T_0'$ are related by

$$\frac{\Delta T_0}{\Delta T_0'} = \left(\frac{\bar{Q}_1}{\bar{Q}_1'}\right)^2$$

(31)

Assuming the above relation holds for all varactors with different dynamic quality factors and fixing $\Delta T_0/T_d$ for the varactor with $\bar{Q}_1 = 15$, the optimum

$$f_2/f_1$$

and the minimum $T_e/T_d$ are calculated and shown by the dotted lines in Figs. 12 and 13.

4. Optimum External Idler Loading Factor

In Section II,A we discussed the effect of the external idler load resistance $R_{\epsilon}$ on the effective input noise temperature and concluded that the minimum effective input noise temperature can be obtained when $R_{\epsilon} = 0$. However, when the pump heating cannot be neglected the above conclusion is no longer valid. The optimum external idler loading factor $R_{\epsilon}/R_s$ for the minimum effective input noise temperature, assuming $R_{\epsilon}$ is at ambient temperature $T_d$ and

$$f_1 \frac{\bar{Q}_1^2}{f_2 1 + R_{\epsilon}/R_s} \gg 1$$
is obtained from the following equation:

\[
\frac{R_e}{R_s} \approx \left( \frac{\Delta T_0/T_d}{1 + (\Delta T_0/T_d)(1 + f_2/f_1)^3} \right)^{1/2} \left( 1 + \frac{f_1}{f_2} \right) \tilde{Q}_1 - 1
\]  

(32)

The optimum external idler loading factor \( R_e/R_s \) is calculated from Eq. (32) as a function of \( \tilde{Q}_1 \) for various values of \( \Delta T_0/T_d \), and the results are shown in Fig. 14. The optimum values of \( f_2/f_1 \) are determined from Fig. 12. Substituting the values \( f_2/f_1 \) and \( R_e/R_s \) obtained from Figs. 12 and 14 in Eq. (27), the minimum normalized effective input noise temperature \( (T_e/T_d)_{\text{min}} \) is obtained and is shown in Fig. 15 as a function of \( \tilde{Q}_1 \) for various values of \( \Delta T_0/T_d \). In comparison with Fig. 13, a considerable improvement in the noise temperature has been accomplished by optimizing the external idler loading.

5. Experimental Results

The equivalent thermal resistance \( R_{\text{th}} \) and heating parameter \( \Delta T_0/T_d \) of the varactor are difficult to determine. One interesting, but still difficult, method is the noise power method used by Garbrecht [40]. A block diagram of his measurement setup is shown in Fig. 16. The varactor mount is so tuned that the series resistance is matched to the signal load resistance as well as the pump generator resistance, and other frequency circuits, especially the idler circuit, are completely detuned so that no parametric amplification takes place when the pump power is applied. First, without pump power, the series resistance is matched to the low noise amplifier, and the noise output power \( P_{\text{out}} \) is recorded. Next the pump power is applied to the varactor and the output impedance is rematched to the amplifier by slightly adjusting the bias voltage of the varactor. The noise output power \( (P_{\text{out}} + \Delta P_{\text{out}}) \) is then recorded. The equivalent pump
Fig. 15. Minimum normalized effective input noise temperature as a function of $\bar{Q}_1$ for various values of $\Delta T_0/T_d$.

Fig. 16. Circuit representation of noise output measurement setup for the equivalent thermal resistance determination. (After Garbrecht [40].)
heating $\Delta T$ and the equivalent thermal resistance $R_{th}$ are determined from the equation

$$\Delta T = P R_{th} = \frac{\Delta P_{\text{out}}}{P_{\text{out}}} (T_d + T_e)$$ (33)

where $P$ is the pump power dissipated in the varactor, $T_d$ is the ambient temperature of the varactor mount, and $T_e$ is the effective input noise temperature of the low noise amplifier measured at the output port of the varactor mount.

The pump heating $\Delta T$ of a varactor (MS 4107) was measured by Garbrecht [40]. At room temperature he dissipated about 70 mw of 36-Gc pump power and measured $\Delta T = 35^\circ\text{K}$. This corresponded to the equivalent thermal resistance $R_{th}$ of 500°K/watt. He also obtained essentially the same values at liquid nitrogen temperature.

Garbrecht constructed a 4-Gc amplifier with the MS 4107 GaAs varactor and pumped at 36 Gc. The input normalized generator impedance $R_g/R_s$ was 11, which corresponds to

$$\bar{Q}_1 \bar{Q}_2 \left|1 + R_e/R_s\right| = 10$$

and the pump power was about 70 mw. The measured effective input noise temperatures are compared with the theoretically estimated values in Table III.

### Table III

**Measured and Estimated Effective Input Noise Temperatures of a 4-Gc Nondegenerate Amplifier Pumped at 36 Gc**

<table>
<thead>
<tr>
<th>Ambient temp. (°K)</th>
<th>300</th>
<th>77</th>
<th>4.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured noise temperature, $f_2/f_1 = 8$ (°K)</td>
<td>85</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>Theoretical noise temperature, $f_2/f_1 = 8$ (°K)</td>
<td>83</td>
<td>28</td>
<td>10</td>
</tr>
<tr>
<td>Optimum $f_2/f_1$ ($\bar{Q}_1 = 9$)</td>
<td>7.5</td>
<td>6.5</td>
<td>3</td>
</tr>
<tr>
<td>Optimum $R_e/R_s$ ($\bar{Q} = 9$)</td>
<td>0</td>
<td>0</td>
<td>1.3</td>
</tr>
<tr>
<td>Theoretical minimum noise temperature, $\bar{Q}_1 = 9$ (°K)</td>
<td>82</td>
<td>26</td>
<td>3.7</td>
</tr>
</tbody>
</table>

*a $f_1 = 4$ Gc, $f_0 = 36$ Gc, $P = 70$ mw, $R_e = 0$, $\Delta T_0 = 0.45^\circ\text{K}$, $R_g/R_s = 11$, $\bar{Q}_1 \bar{Q}_2 = 10$.

The effects of varactor heating due to pump power dissipation have been discussed. The effective input noise temperature of cooled parametric amplifiers can be greatly degraded if the pump heating effect is not carefully considered. The equivalent thermal resistance of most varactors can be improved...
by at least a factor 2 by properly designing the varactor encapsulation. Reducing the thermal resistance of the stud and that of the contacting wire by using a mesh or cohn structure, and filling exchange gas such as helium in the package will reduce the equivalent thermal resistance of the varactor. With a carefully designed varactor and amplifier an effective input noise temperature of less than 2°K should be obtained. The device design considerations will be discussed in the next section.

III. DEVICE CONSIDERATIONS

In the previous section we have discussed the effective input noise temperature and bandwidth of cooled parametric amplifiers. Several factors, which are detrimental to the noise performance of cooled parametric amplifiers, were considered and the important design and operational problems were discussed. However, to simplify the analysis, we have made two assumptions: the varactor characteristic remains essentially constant from room temperature down to liquid helium temperature and the circulator and interconnecting transmission line are ideal.

A parametric amplifier with satisfactory performance at room temperature may have a greatly different performance or may not work at all at liquid helium temperature. This is mainly due to the change in varactor characteristics. If the junction capacitance and series resistance of the varactor at various temperatures are known and the dynamic quality factor $Q_1$ is in an acceptable range, the amplifier can be designed to operate properly at lower temperatures. However, the dynamic quality factor $Q_1$ decreases considerably at lower temperatures owing either to the increase in the series resistance or to the decrease in the capacitance modulation factor, $\gamma$, so more pump power has to be dissipated in the varactor to maintain the same gain-bandwidth product. This results in a much higher effective input noise temperature, and loses the advantage of cooling the amplifier down to such a low temperature.

The cooled circulator is difficult to make and expensive. The characteristics of cooled circulators are often inferior to those of room temperature circulators. The advantages of cooling the circulator are, as discussed in Section II,A,4, the lower thermal noise contribution and the better gain stability. These advantages can be fully appreciated for the liquid helium temperature amplifier; however, considering the poorer characteristics of cooled circulators, the advantage in thermal noise contribution of liquid nitrogen temperature circulators is often insignificant.

This section will briefly describe varactor and circulator characteristics at lower temperatures and some important device considerations for cooled amplifier applications will be discussed. A brief discussion of the refrigerators will also be given.
A. Varactors

1. Temperature Dependence of Barrier Capacitance

Modern practical varactors for cooled parametric amplifiers are semiconductor \textit{pn} junction diodes [41] and semiconductor-metal surface barrier (Schottky barrier) diodes [42]. The varactor exhibits nonlinear barrier capacitance of the form

\[ C = C_0 \left(1 + \frac{V}{V_d}\right)^{-1/n} \tag{34} \]

or elastance

\[ S = S_0 \left(1 + \frac{V}{V_d}\right)^{1/n} \tag{34'} \]

where \(C_0\) and \(S_0\) are the capacitance and elastance at the zero bias voltage, \(V_d\) is the diffusion potential, \(V\) the external bias voltage, and \(n\) a positive number.

![Fig. 17. Temperature dependence of barrier capacitance of diffused GaAs varactor No. 1.](image)

The constant \(n\) is a function of impurity doping profile, and is 2 for an abrupt-junction diode and 3 for a linearly graded-junction diode. For most practical varactors for microwave applications \(n\) ranges between 2 and 3. For the abrupt-junction diode, Eqs. (34) and (34') are also shown as

\[ C = A \left(\frac{\varepsilon q N_a N_d}{2(V_d + V)(N_a + N_d)}\right)^{1/2} = \frac{1}{S} \tag{35} \]

where \(\varepsilon\) is the permittivity, \(N_d\) the ionized donor concentration, \(N_a\) the ionized acceptor concentration, \(A\) the junction area, and \(q\) the electronic charge. In
Eq. (35) $V_d$, $N_d$, and $N_a$ are temperature dependent; $V_d$ increases, and $N_d$ and $N_a$ decrease at lower temperatures. Figure 17 shows the barrier capacitance of a diffused GaAs varactor as a function of bias voltage at three ambient temperatures, $300^\circ$, $77^\circ$, and $4.2^\circ$K. Figure 18 shows the similar curves for Si varactor No. 1, which is made of boron acceptors and phosphorus donors.
Fig. 20. Mesa structure of typical high-button mesa varactor.

Fig. 21. Spreading resistance in a 0.003 \( \Omega \)-cm wafer of 0.020-in. diameter as a function of wafer thickness. (After Irvin [44].)

and has a breakdown voltage of about 6 volts, and Fig. 19 for Si varactor No. 2, which is made of boron acceptors and arsenic donors and has a breakdown voltage of about 15 volts. Most GaAs varactors and heavily doped Si
varactors with low breakdown voltages behaved similarly in $C$ versus $T$
characteristic. However, the B-As doped varactors with high breakdown
voltage behaved differently; the junction capacitance decreased very rapidly
below 77°K, there was a small capacitance variation left at 20.4°K, and at
4.2°K the capacitance was 0.139 pf and did not vary with bias voltage. This
bias-independent capacity decreased between 4.2° and 2°K. This is a sign of
carrier freeze-out. More detailed information on Si varactors can be found
in Ref. [43].

2. Temperature Dependence of Series Resistance

The series resistance of the $pn$ junction varactor can be expressed as the
sum of three distinct resistances [44]:

$$R_s = R_d + R_m + R_b + R_c$$  (36)

Fig. 22. Temperature dependence of electrical conductivity for arsenic-doped silicon
samples. (After Morin and Maita [46].)
where $R_d$ is the resistance of the diffused region, $R_m$ the resistance of the remaining portion of the mesa, $R_b$ the spreading resistance in the base, and $R_c$ the contact resistance. These various regions are illustrated in Fig. 20. For a nonepitaxial structure, by minimizing junction depths and careful control of etching, the sum of $R_d$ and $R_m$ may be kept less than 0.1 $\Omega$. The contact resistance of the nonepitaxial structure can be made less than 0.1 $\Omega$ if the contact is carefully made. Therefore, the main source of series resistance is the spreading resistance of the base. This resistance was calculated by Irvin [44], using the published solutions [45] for a wafer diameter of 0.020 in. and 0.003 $\Omega$-cm resistivity, and the results are shown in Fig. 21 as a function of wafer thickness for 1-mil and 2-mil mesas. It is apparent from the results that no substantial reduction in resistance is possible unless the wafer thickness is less than a mil, which is about a present practical limit. Since epitaxial films on low resistivity substrates ($\approx$0.001 $\Omega$-cm) are readily available today, the low series resistance can be achieved through the use of an epitaxial structure.
without using the ultrathin structure. However, for high microwave frequency applications a skin-depth effect should be considered. At 16 Gc the skin depths are 0.85 mil and 0.5 mil for the 0.003 and 0.001 $\Omega$-cm materials. Therefore, for the thick wafer, the high microwave current would take a detour along the wafer surface and encounter a much higher spreading resistance than that predicted from Fig. 21. In this case the wafer size should be minimized to

![Impedance loci of gallium arsenide varactor No. 1 as a function of bias voltage at 300°, 77°, and 4.2°K and measured at 4.17 Gc.](image)

reduce the current path or the wafer surface be plated with a high conductivity metal.

At lower temperatures $R_d$ and $R_m$ play important roles in the series resistance, since the mesa is made of relatively high resistivity materials (0.04–1 $\Omega$-cm) and its resistance increases at low temperatures. The temperature dependence of the electrical conductivity for a set of silicon crystals is shown in Fig. 22 [46] and the similar curves for a set of gallium arsenide crystals are shown in Fig. 23 [47]. These results indicate that considerable increases in $R_d$ and $R_m$ are expected at extremely low temperatures except for heavily doped materials. Therefore, if the mesa is unnecessarily thick or the doping profile is improper, a thick high resistivity region is left near the junction and the total series resistance increases significantly at low temperatures.
Figure 24 shows the impedance loci of GaAs varactor No. 1 in a test cavity as a function of bias voltage at 300°, 77°, and 4.2°K. Figures 25 and 26 show the similar results of Si varactors No. 1 and No. 2. If we assume that the cavity loss is negligibly small compared with the series resistance of the varactors and the impedance transformation ratio between the cavity and the input transmission line is constant over the entire temperature range, which are the reasonable assumptions for a well-designed test circuit, the normalized resistances of the loci represent the effective series resistances at corresponding ambient temperatures. Figures 24 and 25 show very small changes in the series resistances down to liquid helium temperatures. However, Fig. 26 shows a great change in the series resistance.

At the time of this writing, most commercially available silicon varactors could not be used for the cooled amplifier below 100°K since the series resistance increased very rapidly and below 20°K they ceased to show any variable capacitance. On the other hand, most gallium arsenide varactors and selected silicon varactors such as one shown in Fig. 26 were useful down to liquid helium temperature. However, considerable variations in the resistance at liquid...
helium temperature were observed from one sample to another. This is considered to be due to the difference in the impurity gradient at the junction and the mesa thickness.

Fig. 26. Impedance loci of epitaxial silicon varactor No. 2 as a function of bias voltage at 300°, 77°, and 4.2°K and measured at 5.85 Gc.

3. Thermal Resistance

Another important parameter to be considered for cooled varactors is their thermal resistance. The equivalent thermal resistance of varactors was discussed in Section II,C,1.

Figures 27 and 28 show the temperature dependence of thermal conductivity for a set of gallium arsenide [48] and silicon crystals [49]. The thermal resistance can be calculated in the same way as was the dc spreading resistance [45].

4. Figure of Merit

An equivalent circuit of the varactor is shown in Fig. 29. In the figure $R_s$ is the series resistance, $L_s$ the series inductance, $C_s$ the stray capacitance in the vicinity of the junction, and $C_p$ the package capacitance. These elements are
unnecessary for parametric amplification and have undesirable effects, and are therefore called parasitic elements. The series resistance limits the ultimate noise performance and the maximum operating frequency of the amplifier as discussed in Section II. Other reactive parasitic elements degrade the gain-bandwidth product.

The figure of merit of varactors is usually defined by several parameters. These are the conventional quality factor $Q$, cutoff frequency $f_c$, and dynamic quality factor $\tilde{Q}$, and are respectively defined by

$$\tilde{Q} = \frac{1}{\omega C R_s}$$  \hspace{1cm} (37)

where $C$ is the capacitance at a specific bias voltage

$$f_c = \frac{1}{2\pi CR_s}$$  \hspace{1cm} (38)
or

\[ f' = \frac{S_{\text{max}} - S_{\text{min}}}{2\pi R_s} \]  

(39)

Fig. 28. Temperature dependence of thermal conductivity for single crystal Si with indicated carrier concentrations and dopants. (After Slack [49].)

Fig. 29. Equivalent circuit of a varactor.
[50] where \( S_{\text{max}} \) is the elastance at the maximum negative bias and \( S_{\text{min}} \) is that at the maximum forward bias, and

\[
\bar{\eta} = \frac{|S_1|}{2\omega R_s}
\]  

Equation (40) can easily be rewritten in terms of measurable circuit parameters [51] as

\[
\bar{\eta} = \frac{\text{total reactance variation over bias range}}{4R_s} = \frac{\text{normalized total reactance variation}}{4\frac{n^2 R_s}{R_s}}
\]

and \( R_s/n^2 R_s \) is the VSWR of passive resonant circuit at an operating bias of the amplifier if it is overcoupled to the generator. Then it is apparent that \( \bar{\eta} \) and \( f_c' \) are related by

\[
\bar{\eta} = \frac{1}{4} \frac{f_c'}{f}
\]

The dynamic quality factors of various varactors over the bias range from the maximum forward bias voltage (rectified current \( I_0 = 1 \mu A \)) to \(-5.0 \) volts measured at \( 4 \) Gc are tabulated in Table IV. These measurements were made

<table>
<thead>
<tr>
<th>Type\textsuperscript{a}</th>
<th>Zero-bias capacitance ( C_0 ) (pf)</th>
<th>Temp. ((\degree K))</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs Schottky barrier (epitaxial)</td>
<td>0.72</td>
<td>20</td>
</tr>
<tr>
<td>GaAs \textit{pn} (epitaxial)</td>
<td>0.5</td>
<td>14</td>
</tr>
<tr>
<td>Si \textit{pn} (epitaxial)</td>
<td>0.7</td>
<td>10</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The GaAs varactors were supplied by J. C. Irvin and the Si varactors were supplied by R. L. Rulison, both of Bell Telephone Laboratories.

in 1963 at Bell Telephone Laboratories, and the recent results\textsuperscript{1} of diffused epitaxial GaAs varactors are considerably better than those in the table.

\textsuperscript{1} J. G. Josenhans of Bell Telephone Laboratories measured \( \bar{\eta} \approx 18 \) at \( 4 \) Gc. The varactor was made by J. C. Irvin of Bell Telephone Laboratories.
B. CIRCULATORS

1. Effect of Insertion Loss on Effective Input Noise Temperature

The need for low temperature circulators to be used in conjunction with cooled parametric amplifiers was briefly discussed in Section II,A,4. In order to supplement the previous discussion we shall calculate the effective input noise temperatures of the receivers with a cooled parametric amplifier as a preamplifier.

The receiver arrangements are shown in Fig. 30. In Fig. 30a-1 both the amplifier and circulator are cooled at 4.2°K and in Fig. 30a-2 only the amplifier is cooled at 4.2°K. Figures 30b-1 and 30b-2 show similar arrangements for liquid nitrogen temperature amplifiers.

Assuming lossless transmission lines, the effective input noise temperature of the receiver is calculated from the following equation:

$$T_e = (L_1 - 1) T_1 + L_1 T_{el} + \frac{L_1 (L_2 - 1) T_1 + L_1 L_2 T_{e2}}{G_1}$$ (43)

All symbols in the equation are shown in Fig. 30; $L$ is the insertion loss, $T_1$ the ambient temperature of the circulator, $T_{el}$ the effective input noise tem-
perature of the parametric amplifier, $T_{e2}$ the effective input noise temperature of the second stage amplifier, and $G_1$ the available power gain of the parametric amplifier.

The effective input noise temperatures of the receivers are shown in Fig. 31 as a function of $L_1$. In the calculation $L_2$ is assumed to be equal to $L_1$ and

**Fig. 31.** Effective input noise temperatures of the receivers shown in Fig. 30 as a function of circulator insertion loss.

$G_1$ is 20 db. The figure indicates that, in order to achieve a maserlike noise temperature, a cryogenic circulator is an essential for the liquid helium temperature amplifier. On the other hand, for the liquid nitrogen temperature, degradation in noise temperature by use of a room temperature circulator is not substantial, considering the difficulty of reducing the insertion loss of the liquid nitrogen temperature circulator below 0.3 db at present, while the insertion loss as low as 0.1 db has been accomplished for the room temperature circulator. However, a gain stability requirement imposed upon many systems often necessitates, as discussed before, the use of a liquid nitrogen temperature circulator.
2. Temperature Dependence of Operating Frequency

The most critical part of the circulator for low temperature operations is the ferrimagnetic material, and the ferromagnetic resonance loss is the most critical property. Since the ferrimagnetic garnets have been observed to have lower linewidths than spinels or other ferrimagnets, the ferrimagnetic garnets are mainly considered for low temperature circulator applications. The detailed considerations on the ferrimagnetic garnets are discussed in the paper by Comstock and Fay [52].

Fig. 32. Schematic diagram of three-port Y-junction circulator. (After Fay and Comstock [53].)

The most commonly used circulator for low temperature operation is the ferrimagnetically loaded symmetrical Y junction with a strip line configuration since it is compact and suited for broadbanding. A typical three-port strip line circulator is shown schematically in Fig. 32. A four-port circulator is usually constructed by joining two three-port circulators together.

In many respects the junction circulator is similar to a symmetrical transmission cavity; for instance, the ferrimagnetic disk acts as a resonant cavity and the strip line circuits as the coupling networks. The lowest frequency resonance of the circular disk structure of Fig. 32 is the dipolar mode in which the electric field vectors are perpendicular to the plane of the disk and the RF magnetic field vectors lie parallel to the plane of the disk. This mode, as excited at port 1, is illustrated in the unmagnetized case by the standing wave pattern in Fig. 33a. Ports 2 and 3, if open circuited, will see voltages which are $180^\circ$ out of phase with the input voltage and of value about half of the input voltage. If the standing wave pattern is rotated about $30^\circ$ as in Fig. 33b, then port 3
is situated at the voltage null of the disk and the voltages at ports 1 and 2 are equal. This standing wave pattern results from the combination of two counter rotating modes (fields varying as $e^{\pm i\phi}$). If a dc magnetic biasing field is applied in the direction of the axis of the disk, the two counter rotating modes are no longer resonant at the same frequency, but their resonant frequencies are split by the internal magnetic field ($H_{\text{int}}$): the resonant angular frequency ($\omega^+$) of the mode rotating in the same sense as the electron spins of the ferrite disk is higher than that of another mode ($\omega^-$). If the degree of splitting is adjusted so that the phase angles of the impedances of the two rotating modes are $\pm 30^\circ$ at the operating frequency ($\omega_{\text{op}} = (\omega^+ + \omega^-)/2$), then the standing wave pattern will be rotated $30^\circ$ from that of Fig. 33a as is in Fig. 33b. Under this condition

![Diagram](image-url)
the impedance of the $\omega^+$ mode is capacitive and that of the $\omega^-$ mode is inductive and they are in equal magnitude. The total impedance will be real at the operating frequency. At other operating frequencies, since the impedances are complex quantities, the bandwidth of the circulator is determined by the loaded $Q$ of the circulator mode and the fractional splitting $\delta = (\omega^+ - \omega^-)/2\omega$. A detailed theoretical discussion of symmetrical junction circulators is given in the paper by Fay and Comstock [53]. In order to understand problems involved with low temperature circulators a few important equations will be quoted here from their paper and their discussions will be summarized.

The loaded $Q$ of the circulator mode is given by

$$Q_L = 1.5 \frac{\omega R^2 \epsilon}{G_R d}$$

(44)

where $d$ is the ferrite disk thickness, $R$ the disk radius, $G_R$ the load conductance referred to the edge of the disk, $\epsilon$ the dielectric constant of the disk, and $\omega$ the angular frequency of the signal wave. For proper circulation of the circulator mode from one coupling port to another it is also required that

$$Q_L = 1/[2\sqrt{(3\delta)]}$$

(45)

In order to keep the circulator characteristic constant when temperature is lowered, the splitting $\delta$ and $Q_L$ must be kept constant. The loaded $Q$ is not expected to vary appreciably with temperature. However, in order to keep $\delta$ constant with increasing magnetization $M$, which results when temperature is lowered, it is necessary to decrease the internal magnetic field $H_{int}$ for "below resonance mode" ($H_{int} < \omega/\gamma$) and increase it for "above resonance mode" ($H_{int} > \omega/\gamma$). $\gamma$ is defined by $M = \gamma J$, where $J$ is the angular momentum density. Since the center frequency of the circulator ($\omega_r$) is given by

$$\omega_r = 1.84/\sqrt{\epsilon \mu_{eff}}$$

(46)

where $\mu_{eff}$ is the scalar microwave permeability of the disk, the change in dc magnetic field required to maintain a constant value of $\delta$ also changes $\mu_{eff}$ and hence the operating frequency. The resulting operating field and frequency can be found if the temperature characteristic of $M$ is known, by using curves of $K/\mu$ and $\mu_{eff}$ versus $H_{int}$ and $M$. ($K$ and $\mu$ are the Polder tensor components.)

In general, for the "below resonance" operation the required magnetic field decreases and the center frequency increases with decreasing temperature, and for the "above resonance" operation the opposite is true. For "below resonance" operation the ferrite disk can be operated at magnetic saturation and the required value of $4\pi M$ for a 4-Gc circulator with $Q_L = 2$ is 500 gauss. It is of course acceptable to use a lower value of $4\pi M$ at the expense of higher biasing field. If too low a value of $4\pi M$ is used, a biasing field which is too close to the resonance value must be used and ferromagnetic resonance losses may
become excessive. On the other hand, if higher values of $4\pi M$ are used, the ferrite must be operated below saturation and low field losses may become excessive. For "above resonance" operation the required $4\pi M$ is much larger than that for "below saturation" operation and the disk diameter is reduced.

![Fig. 34. Measured and theoretical linewidth of polycrystalline YIG at 5.5 Gc. (After Comstock and Fay [52].)](image)

3. Temperature Dependence of Insertion Loss

The insertion loss of the circulator is mainly determined by the magnetic losses in ferrites. The unloaded $Q$ of the disk is approximately given by

$$Q_0 = \frac{2\omega^2}{\gamma^2} 4\pi M \Delta H$$  \hspace{1cm} (47)

where $\Delta H$ is the magnetic linewidth of the disk. The insertion loss in decibels is determined by

$$L = 20 \log \left(1 - \frac{Q_L}{Q_0}\right)$$  \hspace{1cm} (48)

The loaded $Q$ is specified by the reverse isolation bandwidth and is typically in the range $1 < Q_L < 4$. In order to keep the insertion loss below 0.4 db it is necessary to have $Q_0$ higher than $20 Q_L$. From Eq. (47) with $Q_0 = 20, 4\pi M = 500$ gauss, at 4 Gc, the linewidth is restricted to $\Delta H < 400$ oe.

The polycrystalline ferrimagnetic garnets such as YIG (yttrium iron garnet)
and YIG-Al have linewidths smaller than 400 oe over the wide temperature range of interest. Figure 34 shows the measured results of linewidths $\Delta H$ as a function of temperature for a dense polycrystalline sample of YIG fired at a temperature which produced a minimum of divalent iron. Similar considerations for the "above resonance" operation lead to similar restriction on ferromagnetic resonance linewidth.

4. Low Temperature Circulators

a. Liquid Nitrogen Temperature Circulators. Liquid nitrogen temperature circulators in 1-, 4-, and 6-Gc frequency ranges have been developed by Raytheon Company. These circulators utilize the "below resonance" mode of operation. The characteristics of these circulators are temperature sensitive, and they change markedly above 90°K and below 77°K. The minimum insertion loss per path of narrow band circulators at 77°K is 0.2 db and varies from
Cooled Varactor Parametric Amplifiers

0.2 to 0.5 db from unit to unit. The measured insertion losses were consistently low (0.2–0.3 db) at around 80°C for all units tested and largely varied at 77°C as mentioned above.

Permanent magnets are used in these circulators to supply the magnetic biasing field.

![Graphs showing insertion loss and isolation at different temperatures](image)

**Fig. 36.** Characteristic of circulator at various operating temperatures. (After Heinz and Okwit [54].)

**b. Liquid Helium Temperature Circulator.** A three-port circulator [52] and four-port circulator were developed at Bell Telephone Laboratories for use in the liquid helium temperature parametric amplifiers. These circulators utilize the "above resonance" mode of operation, and the biasing fields are applied by the superconducting magnets (niobium-zirconium alloy wire)
mounted on the circulator bodies. The ferrimagnetic material used was polycrystalline YIG (see Fig. 34). The measured results of insertion loss and isolation of the three-port circulator are shown in Fig. 35. Since, in the above circulator no attempt was made to match it over a broadband, the bandwidth of the circulator isolation at 30-db points was less than 100 Mc. This bandwidth could be improved by using more sophisticated impedance matching circuits.

A temperature insensitive circulator [54] has been developed at Airborne Instruments Laboratories. This circulator uses a ferrimagnetic material and operates at the magnetic biasing field of about 775 gauss, which is supplied by a permanent magnet having small reversible variations with temperature. The measured characteristics at 290°, 77°, and 4.2°K are shown in Fig. 36.

Hughes Aircraft Company, Raytheon Company, TRG, Inc. [55], and Western Microwave Company are also developing the cryogenic temperature circulators.

C. Refrigerators

Reliability of cooled parametric amplifiers is mainly determined by that of refrigerators and of pump supplies. Since reliability of the pump supply has been improved beyond 10,000 hours by using solid state sources or a combination of varactor harmonic generators and reliable vacuum tubes, the design of refrigerator is the most important factor of determining reliability of cooled parametric amplifiers.

The open dewar system is the most reliable refrigeration system at present; however, frequent refilling operations may be considered inconvenient and the limited steerable angle may exclude a fixed temperature bath from some system applications. The closed-cycle refrigerator, on the other hand, requires frequent periodical maintenance, which for refrigerators operating below 15°K can become costly.

Since the cooled parametric amplifier can be operated at an arbitrary ambient temperature without losing its gain, if it is operated at 20°K, for which temperature a reliable refrigeration system is available, one still can achieve the high sensitivity requirements for most system applications. This is one of the advantages of the cooled parametric amplifier over the maser.

This section will briefly describe the principles of closed-cycle refrigerators and review the present status of several refrigerators available commercially. Design considerations for open-dewar systems, which are mostly applicable to closed-cycle refrigerator systems, will also be discussed.

Almost all refrigeration is effected by gas expansion in some thermodynamic cycle. Two kinds of basic gas expansion, isenthalpic and isentropic, and three

---

1 References [56–58]
refrigeration cycles based on them, Joule–Thompson, Sterling, and Gifford–McMahon cycles, will be discussed here. At extremely low temperatures most substances are solids and at 4°K all but helium are solids; therefore, an obvious choice of gas is helium.

1. Isenthalpic and Isentropic Expansions

a. Isenthalpic Expansion. Free expansion of a gas through a valve, capillary tube, or porous plug, during which the gas does not work either on itself or on any external body, is isenthalpic and is called a Joule–Thompson expansion.

![Diagram of Temperature versus Entropy](image)

**Fig. 37.** Temperature versus entropy diagram of a gas under isenthalpic expansion.

The commonly employed refrigerants for moderate refrigeration such as ammonia and sulfur dioxide will undergo a drop in temperature when they expand from a high pressure at room temperature. However, helium under the same conditions will become hotter. Only when helium is precooled below a certain temperature before it goes through the Joule–Thompson expansion, does its temperature decrease as it expands. The temperature at which no change in temperature takes places upon expansion is called the inversion temperature. Therefore, the Joule–Thompson expansion can be used only for the refrigerator whose temperature is lower than the inversion temperature of the refrigerant.

The temperature versus entropy diagram of a gas when it expands isenthalpically is shown in Fig. 37. One can see the change in the sign of slope of constant enthalpy lines above and below the inversion temperature. When
helium is compressed or precooled below its inversion temperature, about 50°K, in the compressed state, it experiences a drop in temperature on isenthalpic expansion. The temperature decreases, for example, from $T_1$ at pressure $P_1$ to $T_2$ at $P_2$ along the enthalpy curve $h_1$. Refrigeration is made available by restoring helium to its original temperature, and it is proportional to $(h_2 - h_1)$. 
A schematic diagram of a single Joule–Thompson expansion loop is shown in Fig. 38. With the use of a counterflow heat exchanger, if thermal loading of the refrigerator is low enough, the gas is cooled down to a temperature at which the saturated vapor pressure of the liquefied gas is equal to the pressure of the gas after expansion.

As seen from Fig. 37 refrigeration at constant temperature can only occur when a part of the gas is liquefied.

![Schematic diagram of a single Joule–Thompson expansion loop](image)

**Fig. 40.** Schematic diagram of single stage expansion-engine refrigerator.

**b. Isentropic Expansion.** Expansion of a gas in an expansion engine in which the gas does work on itself and (or) some external body is called an isentropic expansion.

A typical temperature versus entropy diagram is shown in Fig. 39. A gas entering at an initial pressure $P_1$ is compressed isothermally along a constant temperature line $T_1$ to pressure $P_2$. The isothermal compression is followed by an isentropic expansion to $P_1$. This results in a temperature drop to $T_2$. Refrigeration is made available by restoring the gas to its original temperature, and it is proportional to the difference in enthalpy ($h_1 - h_2$). The isenthalpic expansion cycle is also shown in the figure. It is clear from the figure that the isentropic expansion is always more efficient than the isenthalpic expansion; $h_1 - h_2 > h_1 - h_3$. At temperatures above the inversion temperature of the gas, as discussed before, refrigeration cannot be obtained by isenthalpic expansion; therefore, isentropic expansion is the only way to obtain refrigeration.

Figure 40 shows a schematic diagram of single stage expansion-engine refrigerator.
2. **Closed-Cycle Refrigerators**

   a. **Claude Cycle Refrigerator.** The Claude cycle refrigerator consists of an expansion engine or engines and a Joule–Thompson valve. Figure 41 shows a simple system using one expansion engine and a Joule–Thompson expansion valve. A large part of the compressed gas is tapped off the heat exchanger at some intermediate point and passed through the expansion engine. The expanded gas which is much colder than the incoming gas passes through the heat exchanger and helps to cool the incoming gas. If the remaining part of the gas is cooled below its inversion temperature, by quenching it through a counterflow heat exchanger, it will, when passing the expansion valve, provide some liquid which can be used for constant temperature refrigeration. Claude’s expansion engine was of the reciprocating type, but a turbine expander was used by Kapitza. The turbine expander is more reliable than a metallic bellows engine and a diaphragm engine. Commercial versions of such turbines for small refrigerators have passed the experimental stage and may soon be available.

   The refrigerators using conventional expansion engines, either in cascade or with Joule–Thompson expansion valve, require very high precision expansion engines and valves that can be operated at low temperatures. This further requires very high purity gas, since any impurities will freeze out in the valves.
or heat exchangers and may hinder the operation of such high precision components. Furthermore, the necessary high engine speed and the vibration and high force values resulting from the extracted work make a reliable design of compact refrigerator difficult.

**b. Sterling Cycle Refrigerator.** Figure 42 shows a schematic diagram of a typical Stirling cycle refrigerator. It consists of a compressor and expander combined into a single cylinder with a piston and a displacer. This refrigerator utilizes periodic-flow heat exchangers (regenerators) instead of the counterflow heat exchanger used in the previous refrigerators. A regenerator is a heat reservoir which readily exchanges heat with gas flowing through it.

The gas is compressed between the piston and the closed end of the cylinder. The mass of gas remains constant in the system and is simply pushed back and forth between the expansion space and the compression space through the regenerator.

The regenerator adjusts the temperature of the gas so that the heat of compression is always generated at the room temperature end. This heat is removed by air or water cooling. The expansion is performed at the low-temperature end, thus supplying refrigeration at this low temperature.

The Stirling cycle has four major phases. Assume the cycle starts with the displacer as far up as possible and the piston as far down as possible, thus allowing most of the gas to be in the compression space.

Phase 1. Since the displacer and piston are 90° out of phase, the gas is compressed by moving the piston up, raising the temperature of gas by compression.

Phase 2. As the displacer moves down to meet the piston the gas is displaced from the compression space, it passes through the cooler, and then the re-
generator, where it is further cooled, assuming previous cycles have already produced refrigeration in the refrigerator.

Phase 3. The gas expands into the expansion space as both the piston and displacer move downward, and its temperature drops.

Phase 4. In its expanded and cooled state, the gas passes through the heat station where it is restored to its preexpanded temperature, thus providing refrigeration outside the system. As the displacer and the piston move back toward their original position, the gas flows back into the compression space through the regenerator, where it absorbs the heat originally given up.

This refrigerator is mechanically simple; therefore, it is suited for compact system applications. Compression and expansion of the gas are done without any valves, and the moving parts do not need close fitting in the low temperature end. Therefore, high reliability can be expected. The small thermal regenerator is remarkably efficient and can be over 99% efficient. The theoretical and the reliable efficiency of the refrigerator are very high.

A disadvantage of this refrigerator is its high mechanical vibrations. Since compression and expansion are done in the same cylinder and the refrigeration cycle must be repeated about 1800 times a minute for a single stage engine, it is very difficult to avoid mechanical vibrations. This is not desirable for parametric amplifier applications, and special mechanical design is needed to minimize the effect of this vibration on the amplifier performances.

C. Gifford–McMahon Cycle Refrigerator. This refrigerator consists of a compressor, an inlet and an exhaust valve, a displacer (or displacers), and a regenerator (or regenerators). Figure 43 shows a schematic diagram of single stage Gifford–McMahon cycle refrigerator. The cycle is completed by four steps.

Step 1. With displacer at the bottom of the cylinder, high pressure gas at room temperature is admitted by opening the inlet valve. The gas fills the space above the displacer and the regenerator. If the gas is helium, it becomes slightly warmer in this process as discussed before, but the regenerator absorbs this heat.

Step 2. As the displacer is moved upward the gas fills the lower part of the cylinder through the regenerator where it gives up additional heat, contracts, and allows additional gas to be admitted.

Step 3. The inlet valve is closed and the exhaust valve opened. The high pressure gas expands and forces the gas at the exhaust valve out. Its temperature drops. This cool gas passes through the regenerator and absorbs the heat originally given up.

Step 4. As the displacer is moved downward to its original position, the remaining cool gas is forced out through the regenerator which is cooled below the initial temperature.

Successive cycles continue to reduce the temperature of the lower end of the
regenerator and cylinder until the thermal load is equal to the refrigeration capacity.

This refrigerator is very similar to the Stirling cycle refrigerator and thus retains most of the advantages of the Stirling cycle refrigerator. The refrigerator requires a separate compressor and two valves. These add some complexity; however, the advantage of being able to choose the optimum compressor design and compressor ratios is of importance for low-temperature refrigeration when the size of the compressor becomes dominant. Since the compressor

Fig. 44. Photograph of Cryodyne 340L, manufactured by Arthur D. Little, Inc.
is a separate part of the main refrigeration body and the displacer can be operated at low speed (100 rpm or less) mechanical vibrations are greatly reduced in comparison with the expansion engine and Stirling cycle refrigerators. The pressure differences across the displacers are only a few pounds per square inch, rather than a few hundred pounds per square inch for others, thereby minimizing seal problems. In addition, the engines are self-purging, that is, tolerant to gas contamination. These factors, low engine speed, low pressure difference, and self-purging, favor reliability and long life.

Use of a refrigerator with a large number of moving parts at low temperatures would reduce the reliability of the system considerably. Lubricants cannot be used in such refrigerators at low temperatures as they would solidify, and elastomers, which ordinarily might be considered for piston seals, become brittle. An additional difficulty in extremely low temperature operation is a purity required for the input high pressure gas. Any contaminants condense and solidify at low temperature, leading to refrigerator failure by plugging the gas passages. Therefore, as the operating temperature of the refrigerator goes down the design of a reliable refrigerator becomes very difficult.

3. Cryogenic System Design

It is necessary for simple maintenance and satisfactory amplifier performance to design an efficient cryogenic system. Noise considerations dictate that not only the varactor but also the total structure of the first-stage amplifier, including the circulator, the idler load, and the input transmission line, should be refrigerated. Immersing the whole amplifier in the refrigerant is a convenient way of refrigerating. However, it is very difficult to seal all parts of such a complex geometry in order to avoid refrigerant leakage, which produces erratic electrical performance. Therefore, it is recommended to enclose the whole unit in a tight cylindrical copper vessel which is then immersed in the refrigerant. This is not necessary for the dry closed-cycle refrigerator system, when heat is extracted by thermal conduction.

In order to achieve a long operating time with each filling of the dewar (or to minimize the load for the refrigerator in closed-cycle refrigerator systems) the heat input must be minimized through the proper use of materials of low thermal conductivity. The design must make full use of the cryogenic value of the refrigerant. One liter of liquid nitrogen can absorb 38,000 cal as heat of vaporization, and in addition the nitrogen gas which is produced can absorb 39,000 cal when brought from 77° to 273°K. If an arrangement can be made so that the nitrogen gas leaves the dewar with a temperature equal to that of ice, for instance, then the cryogenic value of 1 liter of liquid nitrogen is 77,000 cal. Furthermore, if the vaporized gas escapes at low temperature there will be condensation or freeze-out of moisture at the lid of the dewar, where many electrical connections are located. A long operating time between refills can
be obtained by immersing the amplifier deeply into the coolant, thereby reducing the heat input. On the other hand, the input transmission line to the amplifier must be short to keep its insertion loss and the corresponding noise contribution down. This imposes some restrictions on the cryogenic design.

For most metals the electrical conductivity is proportional to the thermal conductivity in the temperature range of interest. Therefore, the minimum electrical insertion loss requirement for low noise systems and the minimum heat leakage for efficient cryogenic system are mutually exclusive. The transmission line should have enough mechanical strength to support the amplifier and withstand mechanical stresses and vibrations. This requires a minimum thickness of the conductors. The transmission line is commonly made of thin walled stainless steel tubing that is copper plated or copper clad at the relevant surfaces, where electromagnetic waves propagate. Stainless steel provides enough mechanical strength but is of relatively low thermal conductivity. The thickness of copper-plated or copper-clad layer is minimized within a practical limit for low insertion loss (about 10 skin depths).

An increase with time in the surface resistivity of plated transmission lines has been observed by many investigators. All of the test samples inhibited corrosion products on their conducting surfaces. The generally accepted procedure to reduce corrosion has been to electroplate a stable, highly conductive coating over the base metal. This method is not always successful because the porous deposits allow the plating solution to produce, at the junction of dissimilar metals, electrolytic action that may result in accelerated corrosion. A finish of high-quality surface on tubing can be obtained by having a precision die-drawn tubing.

The largest heat input in any good dewar occurs through the neck tube, which therefore is usually long and narrow. However, for the parametric amplifier applications this cannot be realized because of the amplifier dimensions (including a circulator) and for another very important reason: the input transmission line must be kept short to minimize the insertion loss. In order to reduce the heat input the neck tube should be made of poor thermal conductor material such as fiberglass reinforced phenolic or a corrugated structure, which increases the thermal path. But this alone does not suffice to obtain a long operating time. One has to shield against thermal radiation and increase the path of flow for the cold gas with several layers of shields. These shields are thermally connected to both the neck tube and the interconnecting transmission lines and surfaces. Another important factor is that all interconnecting transmission lines to the amplifier must be hermetically sealed at the points where temperature is above the dew point. This is extremely important to achieve a reliable amplifier operation since any moisture or ice condensed in the transmission lines increases the insertion losses of the lines and thus increases the system noise temperature and (or) decreases the amplifier gain.
For the closed-cycle refrigerator systems, most of the design considerations discussed above are applicable. The design of amplifier package or the selection of a refrigerator should be made carefully to optimize the amplifier performance and reliability. An important design consideration is to minimize the effect of mechanical vibrations which cannot be avoided with the closed-cycle refrigerator. The cooled parametric amplifiers, being negative resistance amplifiers, are sensitive to any impedance variations. This has been discussed in Sections II,A,4 and III,B,1.

4. Commercial Refrigerators

Several refrigerators that can be used for cooled parametric amplifiers are commercially available today. They are Cryodyne refrigerators by Arthur D. Little, Inc., Cryogem refrigerators by Cryogenators, Inc., Cryomite refrigerators by Malaker Laboratories, and Joule-Thompson and Claude cycle devices by Air Products and Chemicals, Inc. [59]. Hughes Aircraft Company [60] and General Electric Company [61] are also developing their own refrigerators. Promising developments have been shown by the AiResearch Manufacturing Company in closed-cycle refrigerators that use conventional means as well as sophisticated turboexpanders [62].

Fig. 45. Photograph of Cryogem, manufactured by Cryogenators, Inc.
Figure 44 shows a photograph of Cryodyne 340L, which is based on the Gifford–McMahon cycle and provides simultaneous 77° and 20°K cooling stations. Figure 45 shows a photograph of a Cryogem which uses the modified Stirling cycle developed by the Philips Company. It provides refrigeration between 20° and 40°K. The Cryomite also uses a modified Stirling cycle.

For liquid helium temperature operation there can be used a refrigerator [63] (which was developed for Bell Laboratories by Arthur D. Little, Inc.) consisting of three series connected heat engines (the Gifford–McMahon cycle) and a coupled Joule–Thompson circuit.

IV. COOLED PARAMETRIC AMPLIFIERS

The cooled parametric amplifier has passed the purely laboratory development stage, and many amplifiers have been manufactured by several companies.
for high sensitivity system applications such as the satellite ground station receiver, the radio observatory radiometer, and the radar receiver.

Since detailed information of most of the amplifiers is not yet available, the amplifiers developed at Bell Laboratories will be mainly described. As examples, a brief description will be given of a liquid nitrogen cooled 4-Gc amplifier for a satellite ground station [24]; a 1.4-Gc liquid nitrogen cooled amplifier for a radio observatory [12]; a narrow band 4-Gc liquid helium cooled amplifier [23]; and a broad band 4-Gc liquid helium cooled amplifier [64]. An X-band amplifier with a closed-cycle refrigerator [65] developed by Sperry Microwave Electronics Company will also be described. The performances of several amplifiers will be summarized.

1. 4-Gc Liquid Nitrogen Cooled Amplifier

The 4-Gc liquid nitrogen cooled amplifier consists of two cascaded stages of similar design: the first of these is operated at liquid nitrogen temperature and the second at room temperature. The first stage circulator is also operated at

![Diagram of 4-Gc liquid nitrogen cooled parametric amplifier](image-url)
liquid nitrogen temperature. One 23-Gc pump source is used for both amplifier stages. This pump source consists of an X-band klystron of proven reliability which has thermal (vapor phase) frequency stabilization and it is followed by a varactor frequency doubler and an automatic level control (A.L.C.). A block diagram of the receiver is shown in Fig. 46 and a photograph of the amplifier is shown in Fig. 47. The idler and pump tuners are tunable from the control panel by means of two stepping motors which are located inside the dewar. The design parameters can be found in Uenohara et al. [24]. Figure 48 shows a drawing of the cryostat. A commercial Linde Dewar featuring "Super Insulation" was chosen. It has a neck tube of 5.5-in. diameter and holds 13.5 liters of liquid nitrogen. Part of this volume is taken up by the amplifier chamber,
which encloses the varactor mount, circulator, and two stepping motors, so that about 10 liters of liquid nitrogen is available for refrigeration. After immersion of the copper vessel the neck tube is closed off by a lid consisting of a grooved Teflon plug, to the bottom of which is glued a styrofoam cylinder 4 in. deep and \( \frac{1}{2} \) in. smaller in diameter than the neck tube. This arrangement forces the nitrogen gas to rise along the wall of the neck tube and thereby cools it. The pump waveguide, two coaxial lines, a Teflon filling tube, and two polyester support rods (which add mechanical strength to the structure) pass through this lid down to the top of the amplifier chamber. The 11-in. input and output coaxial lines are made of 5-mil wall stainless steel tubes with 0.5-mil copper clad interior. A thin-wall stainless steel waveguide with silver plated
COOLED VARACTOR PARAMETRIC AMPLIFIERS

interior is used for the pump waveguide. These transmission lines are hermetically sealed at the neck of the dewar (where the temperature is higher than the dew point) to prevent moisture condensation. The top of the copper chamber is sealed vacuum tight by an indium O ring. The inside pressure of the chamber is left at atmospheric pressure to improve the speed of cooling.

The amplifier performance is listed in Table V. The measured noise temperatures are very close to the theoretically estimated ones.

Figures 49 and 50 show the front and back view of the receiver.

2. 1.4-Gc Liquid Nitrogen Cooled Amplifier

The 1.4-Gc amplifier was designed to achieve the very low noise performance and high stability that were required for the radiometer in a radio astronomy
<table>
<thead>
<tr>
<th>Performance Data of 4-Gc Liquid Nitrogen Cooled Parametric Amplifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input and output frequency</td>
</tr>
<tr>
<td>Pump frequency</td>
</tr>
<tr>
<td>Varactor diode</td>
</tr>
<tr>
<td>Dynamic quality factor $\tilde{Q}$</td>
</tr>
<tr>
<td>Gain</td>
</tr>
<tr>
<td>Bandwidth</td>
</tr>
<tr>
<td>Noise temperature</td>
</tr>
<tr>
<td>First stage</td>
</tr>
<tr>
<td>Contribution from second</td>
</tr>
<tr>
<td>stage + mixer</td>
</tr>
<tr>
<td>Gain Stability</td>
</tr>
<tr>
<td>Short term</td>
</tr>
<tr>
<td>Long term</td>
</tr>
<tr>
<td>Gain compression, 23-Gc pump</td>
</tr>
<tr>
<td>power required</td>
</tr>
<tr>
<td>Pump stability</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Power</td>
</tr>
<tr>
<td>Dewar nitrogen service</td>
</tr>
<tr>
<td>Evaporation rate</td>
</tr>
</tbody>
</table>

**Fig. 51.** Second stage L-band amplifier mount and its equivalent circuit.
The general design of amplifier is almost identical to the previous amplifier. The first stage, which is operated at liquid nitrogen temperature, is pumped at 23 Gc and the second stage is pumped at 11.5 Gc. One WE 457A Klystron provides these pump powers. A part of 11.5 Gc output power is tapped off to drive the second stage amplifier by a 6-db coupler at somewhere between the variolosser and the doubler.

The idler circuit of the first stage is reactively terminated and is mechanically tunable by means of a motor driven Teflon tuner. This allows the amplifier to be tuned from 1.37 Gc to 1.43 Gc. The varactor used is a gallium arsenide pn-junction diode. The cryostat design is identical to that of the previous amplifier. Since the circulator used is bulky, the dewar has a neck tube of 7.2-in. diameter and holds about 7 liters of liquid nitrogen for refrigeration. This reduced volume of liquid nitrogen, most of which is held above the copper chamber, and the large neck tube reduced the liquid nitrogen holding time to slightly over two days.

The bandwidth of the second stage covers the entire tunable band of the receiver. Figure 51 shows the design of the second stage amplifier and the equivalent circuit of passive signal and idler circuits. The varactor used is an

<table>
<thead>
<tr>
<th>Table VI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance Data of 1.4 Gc Liquid Nitrogen Cooled Parametric Amplifier</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>First stage</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input frequency</strong></td>
<td>1.41 Gc ±20 Mc</td>
<td>1.41 Gc ±20 Mc</td>
</tr>
<tr>
<td><strong>Pump frequency</strong></td>
<td>23 Gc</td>
<td>11.5 Gc</td>
</tr>
<tr>
<td><strong>Idler frequency</strong></td>
<td>21.59 Gc ±20 Mc</td>
<td>10.09 Gc ±20 Mc</td>
</tr>
<tr>
<td><strong>Ambient temperature</strong></td>
<td>77°K</td>
<td>300°K</td>
</tr>
<tr>
<td><strong>Dynamic quality factor ( \tilde{Q} ) (including the circuit loss)</strong></td>
<td>20 at 1.4 Gc, 3 at 21 Gc</td>
<td>20 at 1.4 Gc, 4 at 10 Gc</td>
</tr>
<tr>
<td><strong>Diode capacitance (at zero bias)</strong></td>
<td>0.7 pf</td>
<td>0.7 pf</td>
</tr>
<tr>
<td><strong>Gain</strong></td>
<td>16 db</td>
<td>16 db</td>
</tr>
<tr>
<td><strong>Bandwidth</strong></td>
<td>15 Mc</td>
<td>60 Mc</td>
</tr>
<tr>
<td><strong>Frequency range</strong></td>
<td></td>
<td>1370-1430 Mc</td>
</tr>
<tr>
<td><strong>Noise temperature</strong></td>
<td></td>
<td>&lt;40°K</td>
</tr>
<tr>
<td><strong>First stage</strong></td>
<td></td>
<td>&lt;34°K</td>
</tr>
<tr>
<td><strong>Contribution from second stage + mixer</strong></td>
<td></td>
<td>&lt;6°K</td>
</tr>
<tr>
<td><strong>Gain stability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Short term</strong></td>
<td></td>
<td>&lt;0.01 db</td>
</tr>
<tr>
<td><strong>Long term</strong></td>
<td></td>
<td>&lt;0.02 db</td>
</tr>
<tr>
<td><strong>Pump stability (long term)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td></td>
<td>&lt;1 Mc</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td></td>
<td>&lt;0.02 db</td>
</tr>
<tr>
<td><strong>Pump power</strong></td>
<td>60 mw at 23 Gc</td>
<td>20 mw at 11.5 Gc</td>
</tr>
</tbody>
</table>
epitaxial silicon pn-junction diode. Table VI lists the measured performance of the receiver. The design parameters can be found in Uenohara and Elward [12]. The estimated noise temperature of the first stage amplifier [excluding the circulator (0.5 db) and input transmission line (0.2 db) losses] is 10°C and that of the second stage is 58°C. The measured noise temperature of the second stage was 100°C. This increase in noise temperature was due to the forward current of the silicon varactor (about 20°C/μA) which was intentionally allowed to flow to increase the gain stability. Taking the insertion loss and the second stage noise contribution into account, the first stage noise temperature comes very close to the estimated value.

With this receiver, the radiometer at the Ohio State University Radio Observatory has an operating system noise temperature of about 120°C. With a 5-Mc bandwidth and 12 sec integration time the resulting rms noise temperature variation is 0.025°C for a single record.

3. 4-Gc Liquid Helium Cooled Amplifiers
In order to demonstrate a maserlike noise temperature of the cooled parametric amplifier, the 4-Gc liquid helium cooled amplifiers were developed. Over-all system noise temperatures of less than 10°C have been demonstrated;

![Photograph of 4-Gc liquid helium cooled narrow band amplifier.](image)

however, to achieve such low noise temperatures extremely careful development of each component is necessary as discussed in the previous sections.

a. A Narrow Band Amplifier (1%). For a narrow band amplifier experiment, a 4-Gc liquid nitrogen cooled amplifier mount was slightly modified to operate
it at liquid helium temperature. The idler circuit is reactively terminated externally. A three-port circulator which was developed by Fay (Fay and Comstock [53]) was used in the initial test. Figure 52 shows a photograph of the amplifier. Diffused silicon and gallium arsenide varactors and surface barrier gallium arsenide varactors were tested in the amplifier. The average $\bar{Q}$ at 4 Ge was about 15. The amplifier was pumped at 23 Ge.

In order to avoid shot noise and to minimize pump heating, the amplifier
Michiyuki Uenohara was designed conservatively. The normalized generator impedance $\frac{R_g}{R_s}$ averaged 12 instead of the theoretical maximum value of 43. The average pump power required to obtain a 20-db gain was 10 mw, of which less than half was dissipated in the varactors. Assuming a thermal resistance of 400°K/watt for the varactors, a temperature rise less than $2^\circ$K is expected in the varactor. This results in the effective input noise temperature of 1.5°K for the amplifier (excluding the circulator loss). The measured effective input noise temperature of the receiver with a liquid helium cooled amplifier as the first stage and a liquid nitrogen cooled amplifier as the second stage was $6^{\pm}2^\circ$K. The theoretically estimated value was about $4.5^\circ$K.

b. A Broad Band Amplifier (10%). To build a broad band 4-Gc amplifier with a similar noise temperature to that for the narrow band amplifier, many factors have to be carefully considered: The varactor has to be fully pumped to obtain a large $\gamma$ ($C_1/C_0$, a capacitance modulation factor). This may result in a considerable increase in noise temperature because of the increase in varactor temperature due to pump heating and of shot noise if the pump voltage is swept into the forward conduction. The pump circuit has to be carefully investigated to eliminate any significant amount of idler coupling to the room temperature load through the pump circuit over the idler frequency range of interest.

Figure 53 shows a photograph of the amplifier assembly. The varactor mount and four-port circulator are enclosed in a sealed copper can. The
circulator was developed by Fay, and has a similar characteristic to the three-port circulator [53]. Figure 54 shows a cutaway view of the amplifier, which is based on the design philosophy by De Jager [22]. The varactor used is a gallium arsenide \textit{pn}-junction diode developed by J. C. Irvin and has the zero bias capacitance of 0.3 pf, the dynamic quality factor of 15 at 4 Gc, and a self-resonant frequency of about 16 Gc. The amplifier is pumped at 26 Gc. The amplifier with 16-db gain provides about 450 Mc bandwidth at 3-db down points. The average number of measured effective input noise temperatures is 10°K at the top of the dewar.

4. \textit{X-Band Amplifier with Closed-Cycle Refrigerator}\footnote{Reference [65].}

The entire amplifier assembly, excluding the room temperature circulator, is shown in Fig. 55. It consists of a vacuum-tight amplifier container and a Norelco Cryogem refrigerator. The amplifier mount with the input transmission lines and the circulator is shown in Fig. 56. The mount is constructed of tellurium copper because of its excellent machinability and good thermal and electrical conductivity at low temperatures. The amplifier uses a com-

![Diagram of X-band amplifier with closed-cycle refrigerator]

\textbf{Fig. 55.} Photograph of X-band amplifier with closed-cycle refrigerator. The room temperature circulator is not shown. (After Rucker \textit{et al.} [65].)
commercially available gallium arsenide varactor having dynamic quality factor of about 7.5 at 8 Ge. The varactor is pumped at 23.8 Ge.

The amplifier mount is attached to the cold "finger" of the refrigerator by means of flexible copper straps. In order to minimize effects of mechanical vibrations on amplifier performances, a bellows assembly was used at the joint between the refrigerator and the container, and a spider-type mount support was employed to hold the mount rigidly in place. The spider mount consists simply of three 0.010-in. diameter stainless steel wires which connect to the mount at 120° intervals and anchor to the container body.

The amplifier body temperature reaches a stable value of approximately 40°K 2 hours after the refrigerator is started. On the basis of a thermal load versus temperature diagram supplied by the manufacturer, the thermal load under operating condition was determined to be about 4 watts.

The amplifier with 20-db gain is tunable from 7.0 Ge to 8.0 Ge by changing the pump frequency, and a typical bandwidth is about 15 Mc. The effective input noise temperature of the amplifier (including all input losses) is about 80°K, of which 35 deg are attributed to the input losses. Since $R_g/R_s$ was adjusted for 7.5, the theoretically estimated noise temperature of the amplifier (excluding the circulator and input losses) is 30°K. Therefore, the measured result is about 15°K higher than the theoretical value.
Performances of ten cooled parametric amplifiers are summarized in Table VII.

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-4.3</td>
<td>23 Gc</td>
<td>60</td>
<td>77</td>
<td>Dewar</td>
<td>38</td>
<td>77</td>
<td>45</td>
<td>Idler (mechanical)</td>
<td>BTL</td>
</tr>
<tr>
<td></td>
<td>1.37-1.43</td>
<td></td>
<td>77</td>
<td>Dewar</td>
<td>32</td>
<td>77</td>
<td>40</td>
<td>Idler (mechanical)</td>
<td>BTL</td>
</tr>
<tr>
<td></td>
<td>9.1-9.5</td>
<td></td>
<td>77</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Pump</td>
<td>MP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.7</td>
<td>23</td>
<td>50</td>
<td>20</td>
<td>Cryodyne</td>
<td>20</td>
<td>30</td>
<td>50</td>
<td>Idler (mechanical)</td>
<td>BTL</td>
</tr>
<tr>
<td>3.6-4.3</td>
<td>11.0-14.4</td>
<td></td>
<td>25-30</td>
<td>Cryomite</td>
<td>17</td>
<td>300</td>
<td>30</td>
<td>Idler (mechanical)</td>
<td>BTL</td>
</tr>
<tr>
<td></td>
<td>7.25-7.75</td>
<td></td>
<td>30</td>
<td>Cryogem</td>
<td></td>
<td></td>
<td></td>
<td>Pump</td>
<td>MP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.17</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>AIL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>ITT</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Cryodyne</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>TRG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7-4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency range (Gc)</th>
<th>Pump frequency (Gc)</th>
<th>Bandwidth (Mc)</th>
<th>Cryogenic temperature (°K)</th>
<th>Refrigerator</th>
<th>Gain (db)</th>
<th>Circulator temperature (°K)</th>
<th>Noise temperature (°K)</th>
<th>Tuning</th>
<th>Companya</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>23</td>
<td>40</td>
<td>4.2</td>
<td>Dewar</td>
<td>45</td>
<td>4.2</td>
<td>6</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>26</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>16</td>
<td>4.2</td>
<td>10</td>
<td>Fixed</td>
<td>BTL</td>
</tr>
<tr>
<td>3.7-4.2</td>
<td>24</td>
<td>500</td>
<td>4.2</td>
<td>Dewar</td>
<td>24</td>
<td>4.2</td>
<td>20</td>
<td>Fixed</td>
<td>MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.5-10</td>
<td>Dewar</td>
<td></td>
<td></td>
<td></td>
<td>Fixed</td>
<td>TI</td>
</tr>
</tbody>
</table>

5. Survey of the Cooled Amplifiers

Performances of several amplifiers developed by various companies are summarized and shown in Table VII.

a BTL, Bell Telephone Laboratories; AIL, Airborne Instrument Laboratories; TI, Texas Instruments, Inc.; MP, Microwave Physics Corp.; ITT, International Telephone and Telegraph Company.
Figure 57 shows a photograph of the integrated two-stage parametric amplifier developed by TRG, Inc. [55] and its block diagram.

V. CONCLUSION

The cooled parametric amplifier and its associated devices have been discussed and some practical amplifiers have been described. With available technologies it is not difficult to build the amplifier with noise temperature higher than 20°K and high reliability. However, it is still difficult to achieve...
the broadband receiver with the effective input noise temperature lower than 20°K (especially below 10°K). This is mainly due to the increase in series resistance of most varactors commercially available below 20°K and thus increasing the varactor temperature due to pump heating. To achieve the extremely low noise amplifier with noise temperature lower than 10°K, a further development of better varactors with low thermal resistance packages is necessary. The Schottky barrier diode may be a good solution to minimize the pump heating problem. It is worth comparing the maser and the parametric amplifier, since they provide comparable noise performance at similar ambient temperatures. Since the traveling wave maser (TW maser) is the most commonly used type of maser in practical systems only the TW maser is considered in the comparison. There are many minor differences between the two amplifiers; only the major differences will be discussed here.

Some characteristics which are different for the maser and the parametric amplifier are listed in Table VIII.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Maser</th>
<th>Parametric amplifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>Narrow</td>
<td>Moderately broad</td>
</tr>
<tr>
<td>Saturation power</td>
<td>-50 to -40 dbm</td>
<td>-10 to 0 dbm</td>
</tr>
<tr>
<td>Pump frequency</td>
<td>Fixed by paramagnetic crystal</td>
<td>Arbitrary, optimum choice for specific characteristic</td>
</tr>
<tr>
<td>Ambient temperature</td>
<td>Cryogenic temperature</td>
<td>Arbitrary 3°-400°K</td>
</tr>
<tr>
<td>Gain sensitivity for pump for magnetic field</td>
<td>Insensitive</td>
<td>Sensitive</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Insensitive</td>
</tr>
</tbody>
</table>

The 3-db bandwidth for an operating frequency of say 4 Gc of a single stage maser is about 25 Mc compared with 400 Mc for parametric amplifiers. The bandwidth of a maser can be improved by cascading several stages, but only at a considerable increase in complexity and cost, since each stage requires a different pump frequency.

The saturation power level is usually not of much concern for low noise receiver applications, though one does have to bear in mind that the parametric amplifier after being saturated recovers almost instantaneously, while the recovery time of a maser is quite long, typically of the order of a millisecond; this is because the population inversion has to be reestablished. On the other hand, intermodulation noise due to large interference signals is negligible for masers, while it can be detrimental to the signal-to-noise ratio of parametric amplifiers.

For the maser the pump frequency and magnetic field are fixed by the
operating frequency and the crystal used. On the other hand, the pump frequency of a parametric amplifier can be arbitrarily chosen, keeping in mind only the noise performance desired and the availability of a suitable microwave power source. To provide a magnetic field for the maser poses no great problem, since a superconducting magnet can be used, and this is small and does not require any large excitation power.

The varactor parametric amplifier can be operated at an arbitrary ambient temperature with only minor circuit adjustments. For most low noise receiver applications, the noise performance is sufficiently good even if the parametric amplifier is operated at, say, 20°K instead of at liquid helium temperatures. Operating at such a higher temperature enables one to use a much more reliable closed-cycle refrigerator than would be needed to attain 4°K. This reduces the maintenance cost of the system.

The characteristics of a cooled circulator are a function of its ambient temperature. It is therefore difficult to maintain the performance of a cooled parametric amplifier of which the circulator is also refrigerated if its cryogenic system fails. However, the receiving system would still function, though with a degradation of the noise performance to that of its second stage. It may even be possible to build an amplifier that can be operated continuously from cryogenic temperatures to room temperature. This is not so for the maser, where a failure of the cryogenic system would make the amplifier completely inoperative.

The gain of the circulator-type parametric amplifier is extremely sensitive to the pump power and frequency and to the circulator impedance. However, with a careful design the gain fluctuation can be reduced sufficiently so that the amplifier can even be used effectively in radiometer systems.

It is not an easy task to build a broad band, reliable amplifier with extremely low noise characteristics. However, all devices which are needed to build such an amplifier have been demonstrated.

The maser may still remain as a narrow band preamplifier with the highest sensitivity; however, the parametric amplifier will undoubtedly be used in most high sensitivity receivers, especially of broad band characteristics.

ACKNOWLEDGMENTS

The materials covered in this chapter are results of cooperative studies with many colleagues at Bell Telephone Laboratories. They include Messrs. A. E. Bakanowski, W. J. Bertram, K. D. Bowers, R. L. Comstock, M. J. Cowan, J. T. de Jager, K. M. Eisele, J. P. Elward, Jr., R. S. Engelbrecht, C. E. Fay, H. J. Fink, J. H. Forster, D. C. Hanson, J. C. Irvin, J. G. Josenhans, K. Kurokawa, R. L. Rulison, R. M. Ryder, H. Seidel, W. M. Sharpless, K. M. Poole, and N. C. Vanderwal. Many other people assisted these members during the course of studies. The author would like to express his sincere appreciation for their cooperation and contributions. Special thanks are also due Messrs. L. K. Anderson, W. J. Bertram, K. M. Eisele, R. S. Engelbrecht, C. E. Fay, and J. G. Josenhans for their valuable comments on the manuscript.
### LIST OF SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>junction capacitance</td>
</tr>
<tr>
<td>$C_0$</td>
<td>junction capacitance at zero bias</td>
</tr>
<tr>
<td>$d_1$</td>
<td>slope factor of the signal circuit</td>
</tr>
<tr>
<td>$d_2$</td>
<td>slope factor of the idler circuit</td>
</tr>
<tr>
<td>$e_1$</td>
<td>varactor junction voltage at signal frequency</td>
</tr>
<tr>
<td>$e_2$</td>
<td>varactor junction voltage at idler frequency</td>
</tr>
<tr>
<td>$f_1$</td>
<td>signal frequency</td>
</tr>
<tr>
<td>$f_2$</td>
<td>idler frequency</td>
</tr>
<tr>
<td>$f_{10}$</td>
<td>center signal frequency</td>
</tr>
<tr>
<td>$f_{20}$</td>
<td>center idler frequency</td>
</tr>
<tr>
<td>$f_c$</td>
<td>cutoff frequency of the varactor</td>
</tr>
<tr>
<td>$f_d$</td>
<td>self resonant frequency of the varactor (diode)</td>
</tr>
<tr>
<td>$f_p$</td>
<td>pump frequency</td>
</tr>
<tr>
<td>$g$</td>
<td>voltage gain</td>
</tr>
<tr>
<td>$G$</td>
<td>power gain</td>
</tr>
<tr>
<td>$i_1$</td>
<td>varactor junction current at $f_1$</td>
</tr>
<tr>
<td>$i_2$</td>
<td>varactor junction current at $f_2$</td>
</tr>
<tr>
<td>$i_0$</td>
<td>dc component of conduction current</td>
</tr>
<tr>
<td>$M$</td>
<td>noise measure</td>
</tr>
<tr>
<td>$N_a$</td>
<td>ionized acceptor density</td>
</tr>
<tr>
<td>$N_d$</td>
<td>ionized donor density</td>
</tr>
<tr>
<td>$Q_1$</td>
<td>dynamic quality factor of the varactor at $f_1$</td>
</tr>
<tr>
<td>$Q_2$</td>
<td>dynamic quality factor of the varactor at $f_2$</td>
</tr>
<tr>
<td>$Q$</td>
<td>quality factor of the varactor</td>
</tr>
<tr>
<td>$Q_m$</td>
<td>gain-bandwidth quality factor</td>
</tr>
<tr>
<td>$R_1$</td>
<td>spurious resistance in the signal circuit</td>
</tr>
<tr>
<td>$R_2$</td>
<td>spurious resistance in the idler circuit</td>
</tr>
<tr>
<td>$R_b$</td>
<td>spreading resistance in the base</td>
</tr>
<tr>
<td>$R_c$</td>
<td>contact resistance</td>
</tr>
<tr>
<td>$R_d$</td>
<td>resistance of the diffusion region</td>
</tr>
<tr>
<td>$R_e$</td>
<td>generator resistance</td>
</tr>
<tr>
<td>$R_i$</td>
<td>effective shunt resistance of the varactor</td>
</tr>
<tr>
<td>$R_f$</td>
<td>external idler load resistance</td>
</tr>
<tr>
<td>$R_s$</td>
<td>series resistance of the varactor</td>
</tr>
<tr>
<td>$R_{th}$</td>
<td>equivalent thermal resistance</td>
</tr>
<tr>
<td>$S(t)$</td>
<td>time variable elastance</td>
</tr>
<tr>
<td>$S_0$</td>
<td>average elastance</td>
</tr>
<tr>
<td>$S_1$</td>
<td>first harmonic Fourier coefficient of $S(t)$</td>
</tr>
<tr>
<td>$T_0$</td>
<td>standard temperature (290°K)</td>
</tr>
<tr>
<td>$T_d$</td>
<td>varactor ambient temperature</td>
</tr>
<tr>
<td>$T_e$</td>
<td>effective input noise temperature</td>
</tr>
<tr>
<td>$T_{eq}$</td>
<td>equivalent noise temperature of the varactor</td>
</tr>
<tr>
<td>$T_s$</td>
<td>excess shot noise temperature</td>
</tr>
<tr>
<td>$\Delta T$</td>
<td>equivalent temperature rise due to pump heating</td>
</tr>
<tr>
<td>$\Delta T_e$</td>
<td>excess noise temperature due to pump heating</td>
</tr>
<tr>
<td>$\Delta T_0$</td>
<td>equivalent temperature rise of the varactor pumped at $f_1$ to achieve the same $\gamma$</td>
</tr>
<tr>
<td>$X_{11}$</td>
<td>reactance in the signal circuit</td>
</tr>
<tr>
<td>$X_{22}$</td>
<td>reactance in the idler circuit</td>
</tr>
<tr>
<td>$Z_{11}$</td>
<td>passive signal circuit impedance excluding $R_s$</td>
</tr>
<tr>
<td>$Z_{22}$</td>
<td>passive idler circuit impedance excluding $R_s$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>elastance (capacitance) modulation factor</td>
</tr>
<tr>
<td>$\omega$</td>
<td>angular frequency</td>
</tr>
<tr>
<td>ALC</td>
<td>automatic level control</td>
</tr>
<tr>
<td>PA</td>
<td>parametric amplifier</td>
</tr>
<tr>
<td>BPF</td>
<td>bandpass filter</td>
</tr>
</tbody>
</table>

### REFERENCES

COOLED VARACTOR PARAMETRIC AMPLIFIERS 163


Analysis of Varactor Harmonic Generators

J. O. Scanlan

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
UNIVERSITY OF LEEDS, LEEDS, ENGLAND

I. Introduction

II. The Varactor Model
   A. Charge-Capacitance-Voltage Relationships
   B. Sinusoidal Charges or Voltages

III. Varactor Harmonic Generators without Idlers
   A. Resonant Solution
   B. Limitation on Charge Excursion
   C. Bias Voltage
   D. Optimization of Efficiency and Output Power
   E. The Forward-Driven Varactor Multiplier
   F. The Charge-Storage Diode Multiplier
   G. Bandwidth in Multipliers without Idlers
   H. Variations in Multiplier Circuit Configuration

IV. Varactor Harmonic Generators with Idlers
   A. Resonant Solution
   B. Solution with a Specified Input Current
   C. Small Signal Approximations
   List of Principal Symbols
   References

I. INTRODUCTION

In recent years the demand for a solid state source of microwave power has been increased considerably by the advent of situations where space and weight have become of prime importance in the selection of components and systems for highly mobile military and civil installations. Important additional benefits conferred by the use of solid state sources are increased reliability and lifetime. The most obvious choices for solid state microwave sources are the tunnel diode and (if available at the particular frequency) the microwave transistor. The tunnel diode suffers from the disadvantage of inherently low power output. Although this may be increased by increasing the peak current, a limit is reached where the ever decreasing circuit impedances demanded may not be realized in practice. Frequencies of oscillation as high as 25 Gc may be obtained using commercially available diodes, and about 1 mw appears to represent an upper limit to the power output. The microwave transistor is a comparatively
recent addition to the list of possibilities. The maximum frequency of oscillation currently available from commercial transistors lies in the S band, and forecasts of ultimate performance vary, while output power may be of the order of 1 watt. A very recent device which shows promise of both high-power and high-frequency operation is the "Gunn effect" oscillator. This device is still, in general, in the research stage, and since it is difficult at this time to make a meaningful assessment of its potentiality no discussion is possible.¹

The arrangement whereby a varactor harmonic generator is used to provide a solid state source of microwave power is shown in Fig. 1. A high-power transistor oscillator is followed by a varactor harmonic generator of order \( N \) to provide an output at \( N \) times the frequency of oscillation of the transistor. Varactors are currently available which have useful performance at output frequencies of 40 Gc or even higher, and are capable of delivering several watts of output power. The order of multiplication required from the varactor is evidently dependent on the maximum frequency at which transistor oscillators of the required power output are available, so that any advance in the development of microwave transistors, as well as the varactor itself, will undoubtedly affect the situation as far as the design of varactor harmonic generators is concerned. The section of Fig. 1 labeled "varactor multiplier" may well consist of a chain of multipliers rather than a single multiplier if the required order of multiplication is high, and may or may not have idler circuits associated with it. Thus it will be necessary to examine various configurations in order to arrive at an optimum design. The factors that ultimately limit performance in a particular situation vary according to the circumstances. For example, when the frequency involved is low it may well be that more than adequate power is available from the transistor oscillator and the limiting factor is the power handling capabilities of the varactor, while at high frequencies the amount of power available may be limited and multiplier efficiency is then of vital importance. A number of other possibilities will often occur, and in general each requires a somewhat different design.

The emphasis throughout this chapter is on principles of design and analysis rather than on providing detailed designs to meet particular specifications. Lumped component equivalent circuits are used throughout and little attention is given to translating these into practical microwave realizations. The latter process is of critical importance and is generally the point at which ultimate

¹ See pp. 67–71.
success or failure in achieving a particular aim occurs. The analysis presented here should enable the designer to assess the feasibility of a proposed course of action, or to make a choice of several possibilities.

The main part of the chapter is concerned with "spot-frequency" analysis and design, although coverage is given to the problems of transient response and bandwidth. The analyses given are not restricted to abrupt-junction varactors but are applicable to any varactor with a single power-law dependence of capacitance on voltage. In general the analysis is an exact large-signal one, but at various points the use of small-signal and truncated power-series approximations is demonstrated. The varactor harmonic generator is essentially a nonlinear device, and as is widely known the analysis of circuits including nonlinear elements is generally cumbersome, with few meaningful results emerging. The varactor multiplier is one such circuit for which many useful and quite general results are known, although even in the theoretical analysis on a large-signal basis it is sometimes expedient to introduce slight approximations. Throughout the text it is clearly shown where such approximations are made, and in most cases their effect has been evaluated by numerical methods. As a general philosophy approximations, when introduced, are such as not to favor any particular mode of operation (high-power, low-power, high-frequency, low-frequency, and so on) and generally yield a theory with accuracies within a few per cent of the actual performance. The varactor model used is the simple series type, with a frequency and power-invariant series resistance being the same as that employed by Penfield and Rafuse [1] in their classic text on the subject. It is recognized that a more sophisticated model is required, particularly at high frequencies, in order to derive more severe limits on performance. Furthermore, it is assumed, in general, that the varactor is not driven into forward conduction, where charge-storage effects arise, but the method of analysis and results for this situation are also given. The use of the charge-storage or "step-recovery" diode is described, and related to the varactor. Finally recent work on harmonic generators with idlers is described.

An attempt is made to avoid duplication of material available in Penfield and Rafuse [1], to which reference is made for all background knowledge.

II. THE VARACTOR MODEL

The "series" model of a varactor is shown in Fig. 2. The series resistance is assumed to be frequency and voltage invariant although this is not strictly true. As the frequency of operation is increased high-frequency effects will inevitably cause some variation in $R_s$, while a dependence on drive level is basically inherent in the mechanism of operation. As the voltage across the semiconductor junction varies, the width of the depletion layer also varies, causing a variation in the value of $R_s$. The smaller the total value the more significant
is this varying component, and in a rigorous analysis of multipliers employing extremely high-quality varactors the frequency-converting effects of this varying component of resistance would have to be considered.

The model chosen also fails to incorporate parasitic reactances that are always present and that affect the performance at high frequencies particularly with respect to bandwidth. A rather complete model is shown in Fig. 3. The network of $C_1$, $C_2$, and $L_s$ represents the effect of the diode encapsulation, although at all but the very highest frequencies $C_1$ may be lumped together with $C_2$. However, for the analysis presented here the simple series circuit is used.

An important parameter of the diode equivalent circuit is the cutoff frequency, $f_c$. This is defined as

$$f_c = \frac{1}{2\pi R_s C(v)}$$

The definition is obviously dependent on the value that is chosen for $C(v)$. Various alternatives may be appropriate depending on the application, but provided the maximum value of $C$ is large compared to the minimum value, the latter is normally used and is generally meaningful. Thus as proposed by Uhlir [2] $f_c$ is defined in this chapter as

$$f_c = \frac{1}{2\pi R_s C_0}$$  \hspace{1cm} (1)$$

where $C_0$ is the minimum value, which occurs at the reverse breakdown voltage, $V_R$. This voltage is such that any greater reverse bias causes the diode
to break down, owing to avalanche multiplication, and therefore is one of the factors that limit the power-handling capabilities.

The $Q$ factor of the varactor at a particular frequency $f$ is defined as

$$Q_F = \frac{f_c}{f} = \frac{1}{\omega R_s C_0}$$

Since the series resistance inevitably includes a varying component due to the modulation of the depletion layer width, there is obviously a lower limit on its value, which occurs when all resistance other than the varying component has been removed. The total resistance then varies from zero to a maximum value, and Penfield [3] has defined an "effective" cutoff frequency by taking half the maximum value as the average value for $R_s$. Since this cannot be reduced further, for a particular reverse breakdown voltage, it corresponds to the maximum possible cutoff frequency. Figure 4 shows the results obtained for different materials. High values of reverse breakdown voltage are associated with varactors suitable for harmonic generation and it is evident from Fig. 4 that varactors of this type with cutoff frequencies approaching the theoretical maximum are currently available.

A. CHARGE-CAPACITANCE-VOLTAGE RELATIONSHIPS

The type of relationship between capacitance and voltage considered in this chapter is

$$C(v) = \frac{C_0'}{(1-v/\phi)^m} \quad \left( = \frac{dq}{dv} \right)$$

where $\phi$ is known as the contact potential or diffusion voltage of the junction, and is the voltage which appears across the junction in the absence of any
applied bias. The polarities of voltage are chosen such that $\phi$ is positive. It is evident that when $v = \phi$ $C$ is infinite and in fact as $v$ approaches $\phi$ the effect of the junction disappears and the forward current is virtually limited only by the series resistance. $C_0'$ is the capacitance when $v = 0$, and the index $m$ is $1/2$ for the abrupt junction diode and $1/3$ for a junction with linearly graded doping. Figure 5 shows a sketch of the general type of $C-v$ characteristic that results from Eq. (3). The minimum capacitance, $C_0$, occurs at the reverse breakdown voltage, $v = -V_R$, so that

$$C_0 = C(-V_R) = \frac{C_0'}{(1 + V_R/\phi)^m}$$  \hspace{1cm} (4)

If we write

$$V_0 = V_R + \phi \quad C(v) = C_0 \left(\frac{V_0}{\phi - v}\right)^m$$ \hspace{1cm} (5)

now the charge on the varactor is given by

$$Q = \int C(v) \, dv$$

$$= -C_0 \frac{V_0^m}{1 - m} (\phi - v)^{1-m} + K$$

When $v = \phi$ the voltage across the junction is zero and therefore the charge is also zero, which gives $K = 0$, so that

$$Q(v) = -C_0 \frac{V_0^m}{1 - m} (\phi - v)^{1-m}$$ \hspace{1cm} (6)

At the reverse breakdown voltage the charge $Q_R$ is given by

$$Q_R = -\frac{C_0 V_0}{1 - m}$$

so that

$$\frac{Q(v)}{Q_R} = (\frac{\phi - v}{V_0})^{1-m}$$ \hspace{1cm} (7)

The general shape of the charge-voltage curve is shown in Fig. 6. If the elastance $S(v) = 1/C(v)$,

$$\frac{S(v)}{S_0} = \left(\frac{\phi - v}{V_0}\right)^m$$

and

$$\frac{S(v)}{S_0} = \left(\frac{Q(v)}{Q_R}\right)^{m(1-m)}$$ \hspace{1cm} (8)
ANALYSIS OF VARACTOR HARMONIC GENERATORS

In the case of the abrupt-junction varactor where $m = 1/2$, Eq. (8) indicates that the elastance is directly proportional to the charge. Instead of expressing $Q$ as a function of $v$, Eq. (7) may be rearranged to give

$$v(Q) = \phi - V_0 \left( \frac{Q}{Q_R} \right)^{1/(1-m)}$$

or

$$v + V_R = V_0 \left\{ 1 - \left( \frac{Q}{Q_R} \right)^\gamma \right\}$$

where $\gamma = 1/(1 - m)$.

B. Sinusoidal Charges or Voltages

If either the charge or the voltage appearing across the nonlinear capacitance is constrained to be composed of a finite number of sinusoidal components, then it is evident from Eqs. (7) and (9) that the resulting voltage or charge will not be sinusoidal. For example, if the charge is constrained to be

$$Q = Q_B + Q_0 \cos \omega t$$

where $Q_B$ is the bias charge, then

$$\frac{v + V_R}{V_0} = 1 - \left( \frac{Q_B + Q_0 \cos \omega t}{Q_R} \right)^\gamma$$

$$= 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma (1 + q_0 \cos \omega t)^\gamma$$

where

$$q_0 = \frac{Q_0}{Q_B}$$

and expanding in a binomial series yields

$$\frac{v + V_R}{V_0} = 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma \left\{ 1 + \gamma q_0 \cos \omega t + \frac{\gamma(\gamma - 1)}{2!} q_0^2 \cos^2 \omega t + \frac{\gamma(\gamma - 1)(\gamma - 2)}{3!} q_0^3 \cos^3 \omega t + \cdots \right\}$$

$$= 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma \{a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \cdots \}$$

Thus voltages at all harmonic frequencies appear even though the charge is constrained to be a single sinusoid. Similarly, if the charge is constrained to two
harmonically related sinusoids, as may well be the case with a harmonic generator, so that

$$Q = Q_0 + Q_1 \cos \omega t + Q_N \cos (N\omega t + \varphi)$$

(10)

then

$$\frac{v + V_R}{V_0} = 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma \left\{ (1 + q_1 \cos \omega t)^\gamma + \gamma (1 + q_1 \cos \omega t)^{\gamma-1} q_N \cos (N\omega t + \varphi) \right\}$$

$$+ \frac{\gamma(\gamma-1)}{2!} (1 + q_1 \cos \omega t)^{\gamma-2} q_N^2 \cos^2 (N\omega t + \varphi) + \cdots$$

$$= 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma \left\{ a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \cdots \right.$$  

$$+ b_1 \sin \omega t + b_2 \sin 2\omega t + \cdots \right\}$$

(11)

so that again voltage components appear at all harmonic frequencies.

The voltage waveform appearing on the application of two harmonically related voltages may be analyzed by the standard methods of Fourier series as follows:

$$\frac{v + V_R}{V_0} = 1 - \left( \frac{Q_B}{Q_R} \right)^\gamma \left\{ 1 + q_1 \cos \omega t + q_N \cos (N\omega t + \varphi) \right\}^\gamma$$

$$= 1 + \left( \frac{Q_B}{Q_R} \right)^\gamma \left\{ a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \right\}$$

(12)

where

$$a_n = -\frac{1}{\pi} \int_{-\pi}^{\pi} \left\{ 1 + q_1 \cos \omega t + q_N \cos (N\omega t + \varphi) \right\}^\gamma \cos n\omega t \, d(\omega t)$$

$$b_n = -\frac{1}{\pi} \int_{-\pi}^{\pi} \left\{ 1 + q_1 \cos \omega t + q_N \cos (N\omega t + \varphi) \right\}^\gamma \sin n\omega t \, d(\omega t)$$

Alternatively the power series expansion preceding Eq. (11) may be used to obtain the coefficients. In a large signal analysis every term of this infinite series must be used in the calculation.

When only a single sinusoid of charge (or voltage) is used then in the resulting voltage (or charge), the Fourier coefficients may be expressed in terms of a hypergeometric series [4], and a related publication [5] also contains many useful results on integration, differentiation, and use of these expressions. Further discussion of the evaluation of these coefficients is also given in Penfield [6].

III. VARACTOR HARMONIC GENERATORS WITHOUT IDLERS

When no idlers are used the equivalent circuit of the harmonic generator may take either of the forms shown in Fig. 7. The circuit of Fig. 7a is known as the
series model, and the filters $F_1$ and $F_2$ are such that they permit current to flow only at the fundamental and $N$th harmonic, respectively, being open circuited at all frequencies other than that desired, and short circuited at that frequency. The circuit of Fig. 7b is a shunt model: the filters $F_1$ and $F_2$ allow voltages to exist at the fundamental and $N$th harmonic only, being short circuited at all but the desired frequency. Although the two circuits are in appearance duals,

![Circuit Diagram](diagram)

**Fig. 7.** Equivalent circuits of a varactor multiplier without idlers: (a) shunt diode, (b) series diode.

the nonlinear capacitance renders their behavior different. In the shunt circuit the series $R_s-C$ representation is sometimes replaced by a parallel equivalent in which the components are assumed to be frequency independent, which introduces negligible error at high values of $Q$ factor, but the resulting parallel conductance should really be considered as voltage dependent, giving rise to frequency conversion. These effects can be taken into account in an approximate manner which is quite accurate for high $Q$ factors [7] and this circuit will be discussed later. The representation of Fig. 7a is mainly used throughout the chapter.

If the reference in time is taken to be such that the charge on the variable capacitance at the fundamental frequency is

$$Q_1 = q_1 \cos \omega t$$

and at the harmonic frequency

$$Q_N = q_N \cos (N\omega t + \varphi)$$
we have

\[ I_1 = \frac{dQ_1}{dt} = -\omega q_1 \sin \omega t = i_1 \sin \omega t \]

and

\[ I_N = \frac{dQ_N}{dt} = -N\omega q_N \sin (N\omega t + \varphi) = i_N \sin (N\omega t + \varphi) \]

Here \( q_1 \) and \( q_N \) are chosen to be negative for later convenience. The total charge on the variable capacitance is then

\[ Q = Q_B + q_1 \cos \omega t - q_N \cos (N\omega t + \varphi) \tag{13} \]

and the voltage appearing across it is given by Eq. (12), where

\[
\begin{align*}
a_n &= -\frac{1}{\pi} \int_{-\pi}^{\pi} \left( 1 + \frac{q_1}{Q_B} \cos \omega t - \frac{q_N}{Q_B} \cos (N\omega t + \varphi) \right)^\gamma \cos n\omega t \, d(\omega t) \tag{14} \\
b_n &= -\frac{1}{\pi} \int_{-\pi}^{\pi} \left( 1 + \frac{q_1}{Q_B} \cos \omega t - \frac{q_N}{Q_B} \cos (N\omega t + \varphi) \right)^\gamma \sin n\omega t \, d(\omega t) \tag{15}
\end{align*}
\]

The ac components of the voltage are then given by

\[ V_0 \left( \frac{Q_B}{Q_R} \right)^\gamma (a_n \cos n\omega t + b_n \sin n\omega t) \tag{16} \]

The circuit equations of Fig. 7a are

\[
E_g \sin (\omega t + \alpha) = -\omega q_1 (R_g + R_\ell) \sin \omega t - L_1 \omega^2 q_1 \cos \omega t + V_0(Q_B/Q_R)^\gamma (a_1 \cos \omega t + b_1 \sin \omega t) \tag{17}
\]

and

\[
V_0(Q_B/Q_R)^\gamma (a_N \cos N\omega t + b_N \sin N\omega t)
= -N\omega q_N (R_L + R_g) \sin (N\omega t + \varphi) - L_N N^2 \omega^2 q_N \cos (N\omega t + \varphi) \tag{18}
\]

Equations (14) and (15) together with (17) and (18) are sufficient to determine the performance of the harmonic generator. However, even in the most simple case of \( \gamma = 2 \) \((m = 1/2)\) a direct computation involves the solution of a cubic equation for \( q_N \) in terms of \( E_g \) and the circuit parameters.

A. Resonant Solution

Some simplification is obtained by treating only the resonant case. Resonance in the input circuit is defined (at a particular drive level) as the condition in which the (nonlinear) input impedance is purely real. Eq. (17) then gives

\[
\alpha = 0 \quad E_g = -\omega q_1 (R_g + R_\ell) + V_0(Q_B/Q_R)^\gamma b_1 \tag{19}
\]

\[
L_1 \omega^2 q_1 = V_0(Q_B/Q_R)^\gamma a_1 \tag{20}
\]
The meaning of resonance in the output circuit is not by any means as clear. The desirable condition is that the maximum possible power should be developed in the load resistance, \( R_L \), for a given value of the peak voltage appearing across the varactor. The value of the peak voltage for a given value of \( E_q \) is evidently a function of, among other things, the phase angle of the output charge. Thus in order to find the true optimum situation the output power should be found from Eqs. (14), (15), (18), and (19) as a function of \( \phi \), and then the value of \( \phi \) which minimizes the peak voltage across the nonlinear capacitor should be found. In practice this is not analytically possible, and a "linearized" resonance condition is used by all authors in the field. This consists in assuming that the inductance in the output circuit is chosen to resonate with some "effective" capacitance, which is a function of drive level, at the output frequency. The difference between this condition and the true optimum in the sense described above is generally very small. Except in the case where \( m = 1/2 \) and the various quantities can be evaluated analytically, the problem of finding the linearized resonant condition is not straightforward. A useful approach has been found to be the following: The problem is essentially that the voltage across the variable capacitance at the output frequency is

\[
v_N = V_0 Q_B/Q_R \gamma (a_N \cos N\omega t + b_N \sin N\omega t)
\]

but as the phase of the output current is contained in \( a_N \) and \( b_N \) it is not easy to see how a resonance condition is arrived at. If, however, we consider \( v_N \) to be composed of two parts, one independent of \( q_N \) (and of \( \phi \)) and the remainder a function of both \( q_1 \) and \( q_N \) (also \( \phi \)), the output circuit may be redrawn as shown in Fig. 8. The voltage \( v_2' \) is a function of both \( q_1 \) and \( q_2 \) (and \( \phi \)), and \( v_2 = v_2' + v_2'' \). Now, a "linearized" resonance condition, in the sense previously mentioned, occurs when the current, \( I_N \), is in phase with \( v_2' \), since as far as the output circuit is concerned this represents a constant voltage generator (that is, a voltage which is independent of the current). The remaining voltage \( v_2'' \) represents the voltage across a nonlinear impedance which may have both resistive and reactive components which are dependent on the drive level. Thus resonance is achieved when, at particular values of \( q_1 \) and \( q_N \), \( L_N \) resonates with a capacitance whose reactance is the ratio of the reactive

![Fig. 8. Equivalent circuit of a varactor multiplier at the output frequency.](image-url)
components of $v_2'$ to the current $I_N$, that is, the "effective" elastance. In the case of the abrupt-junction doubler $v_2'$ consists only of a reactive component proportional to $q_N$, namely the voltage across the average elastance. Essentially the process of "linearized" resonance consists of performing the same steps, at particular values of $q_1$ and $q_N$, as if the circuit were linear, and the steps described above are just those that would be taken in the linear case. The method is, of course, open to the same objections as stated previously, namely that the truly optimum condition may require a somewhat different phase for the secondary current.

The voltage $v_2'$ is given by

$$v_2' = V_0(Q_B/Q_R)^\gamma (k_N \cos N\omega t + I_N \sin N\omega t)$$

where

$$k_N = -\frac{1}{\pi}\int_{-\pi}^{\pi} \left(1 + \frac{q_1}{Q_B} \cos \omega t\right)^\gamma \cos N\omega t d(\omega t)$$

$$I_N = -\frac{1}{\pi}\int_{-\pi}^{\pi} \left(1 + \frac{q_1}{Q_B} \cos \omega t\right)^\gamma \sin N\omega t d(\omega t)$$

In this case $I_N=0$ so that at resonance the current in the output circuit always has the phase $k_N \cos N\omega t$. When the time origin is chosen such that $Q_1 = q_1 \sin \omega t$, $I_N$ is zero when $N$ is even, and $k_N$ is zero when $N$ is odd. As a check on the usefulness of this method one can construct the appropriate charge waveforms and see that, apart from slight variations of $\phi$ about the values found, the results yield the smallest peak values for given values of $q_1$ and $q_N$ (other than results appropriate only to reactive circuits). This method is applied throughout this article and where similar cases are considered yields the same results as those found by Penfield and Rafuse [1].

Proceeding now from Eq. (18) the phase of the current in the output circuit should be that of $k_N \cos N\omega t$, so that the positive sign is taken if $k_N$ is positive and vice versa. Then

$$V_0(Q_B/Q_R)^\gamma b_N = \pm L_N N^2 \omega^2 q_N$$

and

$$V_0(Q_B/Q_R)^\gamma a_N = \pm N\omega q_N(R_L + R_s)$$

The power developed in the load resistance $R_L$ is

$$P_{out} = I_N^2 R_L/2$$

$$= \frac{V_0^2(Q_B/Q_R)^{2\gamma} a_N^2 R_L}{2(R_L + R_s)^2}$$

and defining efficiency, $\eta$, as the ratio of the power in the load to the available power from the source, we have

$$\eta = \frac{4R_s R_L V_0^2(Q_B/Q_R)^{2\gamma} a_N^2}{E_0^2(R_L + R_s)^2}$$
ANALYSIS OF VARACTOR HARMONIC GENERATORS

Equations (25) and (26) are true, in general, for any order of multiplication by a varactor with any single power law for its capacitance-voltage characteristic, subject to the assumption made regarding the resonant condition. The two expressions are, however, difficult to manipulate in general, since $a_N$ is a function of both $q_N$ and $q_1$ and thus implicitly of $E_g$. Equation (19) can be used to eliminate $E_g$ from Eq. (26) to give

$$\eta = \frac{4R_g R_L V_0^2 (Q_B/Q_R)^{2^\gamma} a_N^2}{(R_L + R_g)^2 (-\omega q_1 (R_g + R_s) + V_0 (Q_B/Q_R)^{\gamma b_1})^2}$$

(27)

B. LIMITATION ON CHARGE EXCURSION

The charge appearing on the variable capacitor at any instant is, from Eq. (13) with $\varphi = \pm \pi/2$,

$$Q(t) = Q_B + q_1 \cos \omega t \pm q_N \sin N\omega t$$

the negative sign occurring when $N$ is even and the positive sign when $N$ is odd. If it is desired to prevent forward conduction, and if the total (negative) charge must never exceed the charge at the reverse breakdown voltage,

$$Q_R \leq Q(t) \leq 0$$

(28)

Thus for a given bias charge the magnitudes of $q_1$ and $q_N$ are restricted so that Eq. (28) is satisfied for all values of $t$. Differentiating the charge to find the maximum and minimum values gives

$$-q_1 \omega \sin \omega t \pm q_N N\omega \cos N\omega t = 0$$

(29)

and the solution of this equation gives the times at which maximum and minimum charges occur. These are then substituted into the equation for the charge and the relationship among $Q_B$, $q_1$, and $q_2$ found such that the limits of Eq. (28) are satisfied. Finding the times of maximum and minimum charge, in general, involves the solution of an $N$th-order polynomial equation so that a general result cannot be stated. The solution for $N=2$ is readily found as shown below. Equation (29) in this case becomes

$$q_1 \sin \omega t + 2q_N \cos 2\omega t = 0$$

or

$$\sin \omega t = \frac{1 \pm \sqrt{[1 + 32(q_N/q_1)^2]^{1/2}}}{8(q_N/q_1)}$$

Hence

$$Q = Q_B + q_1 \{16(q_N/q_1)^2 - 1 \mp [1 + 32(q_N/q_1)^2]^{1/2}\}^{1/2}$$

$$\times \left\{\frac{3 \mp [1 + 32(q_N/q_1)^2]^{1/2}}{16 \sqrt{2}(q_N/q_1)}\right\}$$
where \( \bar{Q} \) indicates the maximum and minimum values of \( Q \), which occur when the negative and positive signs, respectively, are taken. Thus

\[
Q_b + q_1 \left\{ 16 \left( \frac{q_N}{q_1} \right)^2 - 1 \left[ 1 + 32 \left( \frac{q_N}{q_1} \right)^2 \right]^{1/2} \right\}^{1/2} \left\{ 3 - \frac{[1 + 32(q_N/q_1)^2]^{1/2}}{16\sqrt{2}(q_N/q_1)} \right\} \leq 0
\]  

(30)

and

\[
q_1 \left\{ 16 \left( \frac{q_N}{q_1} \right)^2 - 1 + \left[ 1 + 32 \left( \frac{q_N}{q_1} \right)^2 \right]^{1/2} \right\}^{1/2} \left\{ 3 + \frac{[1 + 32(q_N/q_1)^2]^{1/2}}{16\sqrt{2}(q_N/q_1)} \right\} \geq Q_R - Q_B
\]  

(31)

(remember that both \( Q_R \) and \( Q_B \) are negative). In order that the largest value of \( q_1 \) for a given ratio of \( q_1 \) to \( q_N \) be possible both the above inequalities should be simultaneously satisfied with the equality signs so that

\[
Q_b = \frac{Q_R}{2}
\]

In this case if we choose normalized charge variables

\[
p_1 = 2q_1/Q_R
\]

(32)

\[
p_N = 2q_N/Q_R
\]

(33)

both the above conditions become

\[
\{16p_N^2 - p_1^2 + p_1[p_1^2 + 32p_N^2]^{1/2}\}^{1/2} \{3p_1 + [p_1^2 + 32p_N^2]^{1/2}\} = 16\sqrt{2}p_N
\]

(34)
This relationship is displayed in Fig. 9 and defines a region of allowable values for $p_1$ and $p_N$. Any other bias charge will reduce the area of allowed values of $q_1$ and $q_N$. Obviously a different curve relating the permissible values of $q_1$ and $q_N$ exists for each value of $N$. An output current with a phase angle other than $\pm \pi/2$ will relax the curve given in Fig. 9 slightly and thereby lead to slightly higher efficiency and output power, but the difference is small and is not considered here.

If $Q_B < \frac{1}{2}Q_R$ the second inequality, Eq. (31), predominates and if $Q_B > \frac{1}{2}Q_R$ the first, Eq. (30), limits the magnitudes of $q_1$ and $q_N$. In the former case the allowed value of $p_1$ for a given value of $p_N/p_1$ is reduced by a factor

$$2[(Q_R - Q_B)/Q_R]$$

and in the latter case by a factor $2Q_B/Q_R$, so that it is again evident that the bias point should be at $Q_B = Q_R/2$, in order to obtain the maximum charge swing.

C. Bias Voltage

The dc voltage appearing across the variable capacitance for a particular value of $Q_B$ is a function of the values of $q_1$ and $q_N$ for any particular diode, and is given by Eq. (12) as

$$(V_{dc} + V_R)/V_0 = 1 - (Q_B/Q_R)^\gamma a_0$$

where $a_0$ is the zero-frequency Fourier component of the voltage waveform. Again $a_0$ is not found analytically for arbitrary values of $N$ and $\gamma$, but for the abrupt junction varactor where $N=2$ and $\gamma=2$ expansion of the voltage-charge relationship gives

$$a_0 = 1 + (q_1/Q_B)^2 + (q_N/Q_B)^2$$

and

$$V_{dc} = \phi - V_0 \left \{ Q_B^2 + (q_1^2/2) + (q_N^2/2) \right \} Q_R^2$$

and when $Q_B = \frac{1}{2}Q_R$

$$V_{dc} = \phi - \frac{V_0}{4} \left \{ 1 + \frac{p_1^2}{2} + \frac{p_N^2}{2} \right \}$$

D. Optimization of Efficiency and Output Power

It is now necessary to consider optimization of efficiency and output power under various constraints, while in each case maintaining the previously discussed resonance condition. The fundamental limitation in all these processes is the requirement that the charge components fall within the region specified in Subsection B. There are many further constraints which can be added as demanded in a particular situation. For example, the value of the
input drive level, \( p_1 \), may be specified, or the available power from the source may be fixed, or the output power may be predetermined, each leading to a different optimization procedure. The most general case where only the breakdown limit of Fig. 9 is specified will, of course, lead to the absolute maximum values of both efficiency and power output. This case is extremely difficult to treat analytically, even for the abrupt junction doubler. Penfield and Rafuse [1] have given the numerical data corresponding to this case, and in this contribution some discussion is given of simpler cases which in fact yield results differing by less than a few per cent from the true optima, and in addition provide useful design data in a simple form.

1. Optimization with Input Drive Level Specified

Returning to Eq. (27) and setting \( Q_B = \frac{1}{2} Q_R \), we have

\[
\eta = \frac{4 R_g R_L V_0^2 a_N^2}{2^{3\gamma}(R_L + R_s)^2 (-\omega q_1(R_g + R_s) + (V_0/2\gamma) b_1)^2}
\]

(37)

If \( q_1 \) is specified, then \( a_N \) and \( b_1 \) are independent of \( R_g \) and we may immediately differentiate with respect to this quantity to obtain a maximum. This gives

\[
R_g = R_s - \frac{V_0 b_1}{2\gamma \omega q_1}
\]

(38)

This is the condition where, at the specified drive level, the generator resistance is equal to the input resistance, as can be seen from Eq. (19), which gives

\[
\frac{E_g}{I_1} = R_g + R_s - \frac{V_0 b_1}{2\gamma \omega q_1}
\]

This is not a surprising result since for maximum efficiency we evidently require that all the available power from the generator be transferred to the multiplier. Substitution of this value for \( R_g \) into Eq. (37) yields

\[
\eta = \frac{V_0^2 a_N^2}{2^{3\gamma} \omega^2 q_1^2} \frac{R_L}{(R_L + R_s)^2} \frac{1}{R_s - V_0 b_1/2\gamma \omega q_1}
\]

(39)

One cannot now immediately differentiate with respect to \( R_L \) since both \( a_N \) and \( b_1 \), being functions of \( q_N \) are therefore functions of \( R_L \). In order to do this it is necessary to know the explicit forms of \( a_N \) and \( b_1 \), which are only available for the abrupt junction doubler. In that case

\[
a_N = -\frac{p_1^2}{2} = -\frac{2 q_1^2}{Q_R^2}
\]

(40)

\[
b_1 = p_1 p_N = 4 q_1 q_N
\]

(41)
Equation (39) then becomes

\[
\eta = \frac{1}{4} \frac{V_0 q_1^2}{Q_R^2} \frac{R_L}{(R_s + R_L)^2 R_s - V_0 q_N/\omega Q_R^2}
\] (42)

or

\[
\eta = \frac{1}{16} \frac{V_0^2 p_1^2}{Q_R^2} \frac{R_L}{(R_s + R_L)^2 R_s - V_0 p_N/2\omega Q_R}
\] (43)

also, \( q_1 \) and \( q_N \) are related by Eq. (24) so that

\[
\frac{V_0 p_1^2}{8} = -\omega Q_R p_N (R_L + R_s)
\] (44)

Substituting for \( p_N \) from Eq. (44) in Eq. (43) gives

\[
\eta = \frac{1}{16} \frac{V_0^2 p_1^2}{Q_R^2} \frac{R_L}{(R_s + R_L)^2 R_s + (V_0^2 p_1^2/16\omega^2 Q_R^2) [1/(R_L + R_s)]}
\]

Using, from Eq. (2), the definition of \( Q_F \), the diode \( Q \) factor at the reverse breakdown voltage

\[ Q_F = \frac{1}{\omega R_s C_0} \]

and

\[ Q_R = -\gamma C_0 V_0 \]

We find

\[ \eta = \frac{R_L}{R_s} \frac{1}{\{1 + (64/p_1^2 Q_F^2) (1 + R_L/R_s)\} (1 + R_L/R_s)} \]

and since \( p_1 \) is specified we set \( \partial \eta / \partial (R_L/R_s) = 0 \) for a maximum, which gives

\[ R_L = 1 + \frac{Q_F^2 p_1^2}{64(1 + R_L/R_s)} \]

But substituting for \( b_1 \) gives

\[ \frac{R_L}{R_s} = 1 + \frac{Q_F^2 p_1^2}{64(1 + R_L/R_s)} \]

Thus this optimization procedure leads to equal source and load resistances and

\[ \eta = \frac{1}{64} \frac{p_1^2 Q_F^2}{\{1 + [1 + p_1^2 Q_F^2/64]^{1/2}\}^2} \] (45)

also

\[ \frac{R_L}{R_s} = \frac{R_g}{R_s} = \left\{1 + \frac{Q_F^2 p_1^2}{64}\right\}^{1/2} \] (46)
Fig. 10. Efficiency as a function of $p_j Q_F$ for an abrupt-junction doubler with $R_L = R_g$ (note: breakdown point depends on $Q_F$).

Fig. 11. Optimum load resistance as a function of $p_j Q_F$ for an abrupt-junction doubler with $R_L = R_g$ (note: breakdown point depends on $Q_F$).
ANALYSIS OF VARACTOR HARMONIC GENERATORS

183

\begin{align}
\gamma &= 2.0 \\
\gamma &= 1.5
\end{align}

\begin{align}
\eta_{\text{max}} \% &= \gamma \\
\eta_{\text{max}} \% &= 1.5
\end{align}

Fig. 12. Maximum efficiency as a function of diode \( Q \) factor for abrupt and graded junction doublers with \( R_L = R_s \).

\begin{align}
\gamma &= 2.0 \\
\gamma &= 1.5
\end{align}

\begin{align}
\eta_{\text{max}} \% &= \gamma \\
\eta_{\text{max}} \% &= 1.5
\end{align}

Fig. 13. Efficiency as a function of load resistance [28]. Curve a: \( N=2, \gamma=1.5, p_1=0.6, Q_F=50 \); curve b: \( N=3, \gamma=1.5, p_1=0.8, Q_F=200 \); curve c: \( N=4, \gamma=1.6, p_1=0.9, Q_F=200 \).

Equation (47) may be substituted into Eq. (34), which yields the maximum value of \( p_1 \) for each value of \( Q_F \), and Eq. (45) then gives the maximum efficiency. This is shown in Fig. 12 as a function of \( Q_F \). As \( Q_F \to 0 \), \( p_1 \to 1 \) and

\begin{align}
\eta \to \frac{1}{256} Q_F^2 = \frac{1}{256} \left( \frac{f_c}{f} \right)^2
\end{align}
Fig. 14. Efficiency as a function of $p_1$ [28]. (a) $Q_F = 400$; (b) $Q_F = 200$; (c) $Q_F = 100$

(i) $N = 2 \quad \gamma = 2$
(ii) $N = 2 \quad \gamma = 1.5$
(iii) $N = 3 \quad \gamma = 1.5$
(iv) $N = 3 \quad \gamma = 1.7$
(v) $N = 3 \quad \gamma = 1.9$
(vi) $N = 4 \quad \gamma = 1.5$
(vii) $N = 4 \quad \gamma = 1.7$
As \( Q_F \to \infty, P_N \to \frac{1}{4} P_1 \) and

\[
\eta \to 1 - 20.8 \left( \frac{f}{f_c} \right)
\]  

(49)

The expression found by Penfield and Rafuse [1] has a factor 19.9 instead of 20.8 since their formula corresponds to absolute maximum efficiency whereas here \( p_1 \) is assumed given. The maximum power output is found when \( R_L = R_s \), from Eq. (25), and is

\[
P_{\text{out}} = \frac{1}{512} P_n p_1^2
\]

where the normalization power \( P_n = V_o^2/R_s \). At high frequencies \( p_1 \to 1 \) and

\[
P_{\text{out}} \to 0.00195 P_n
\]

(50)

and as the frequency tends to zero

\[
P_{\text{out}} \to 0.0015 P_n
\]

(51)

The high-frequency approximation for the input and output resistance is given by

\[
R_g/R_s = R_1/R_s = 1 + 0.0078(\omega_c/\omega)^2
\]

and at low frequencies

\[
R_g/R_s = R_1/R_s \to 0.096(\omega_c/\omega)
\]
It is interesting to note that the condition of maximum efficiency for a given value of $p_1$, which implies $R_L = R_s$, also gives the absolute maximum power output in both the limiting cases, although this has not been established analytically but can be seen from table 8.5 of Penfield and Rafuse [1].

In the case of varactors other than the abrupt junction type, the various Fourier coefficients cannot be found explicitly, and numerical techniques must be used. From Eq. (38) we have seen that, with $p_1$ given, maximum efficiency is obtained when the source resistance is equal to the input resistance at the particular drive level, irrespective of the values of $\gamma$ and $N$. In order to find
ANALYSIS OF VARACTOR HARMONIC GENERATORS

Fig. 15—continued.
(b) \( N = 3 \).

(i) \( \gamma = 1.5 \quad Q_F = 400 \quad (iv) \quad \gamma = 1.7 \quad Q_F = 200 \)
(ii) \( \gamma = 1.7 \quad Q_F = 400 \quad (v) \quad \gamma = 1.5 \quad Q_F = 100 \)
(iii) \( \gamma = 1.5 \quad Q_F = 200 \quad (vi) \quad \gamma = 1.7 \quad Q_F = 100 \)

the optimum value of \( R_L \) efficiency has been calculated as a function of this parameter for a representative selection of cases, and the results are shown in Fig. 13, from which it is seen that \( R_L = R_g \) in all cases gives maximum efficiency, so that we shall take

\[
\begin{align*}
\frac{R_L}{R_s} &= \frac{R_g}{R_s} = 1 - \frac{V_0 b_1}{2\gamma \omega q_1 R_s} \\
&= 1 + \frac{Q_F b_1}{\gamma^2 p_1} 
\end{align*}
\]

Then the efficiency is

\[
\eta = \frac{a_N^2 Q_F^2}{\gamma^2 2^{2\gamma-2} p_1^2 (2 + Q_F b_1/\gamma 2^{\gamma-1} p_1)^2} 
\]

\( a_N \) and \( b_1 \) must be evaluated numerically for a given \( p_1 \). Figure 14 shows efficiency as a function of \( p_1 \), with \( R_L = R_g \) for three different values of \( Q_F \),
while Fig. 15 shows corresponding values of load and source resistance, normalized to the diode series resistance, $R_s$, for $N=2$ and $N=3$. The points at which the curves are stopped correspond to the breakdown limit in each case. This was inherent in the computer program where the peak value of the charge was monitored to prevent the limits from being exceeded.

Figure 16 shows efficiency as a function of normalized input power, with contours of constant normalized output power superimposed. The power normalization factor used in these curves was $(V_0^2/R_s 2^{2y+1}) \times 10^{-5}$.

The optimization and analysis carried out in this section on the assumption that $p_1$ is specified do not yield the absolute maximum of either efficiency or output power, but the actual results obtained do not differ by more than a
few per cent from true optima. This approach has the advantage of providing convenient design criteria for practical use, since the input or output power corresponding to a given $p_1$ can readily be found, as well as the load and generator resistances.

2. Exact Analysis of Graded-Junction Doubler

Greenspan [8] has analyzed the graded-junction doubler ($\gamma = 1.5, N = 2$) in some detail, and Fig. 17 shows the results obtained for the absolute maximum efficiency (subject to the usual tuning condition) as a function of $Q_F$, together with the corresponding results for the abrupt-junction case, from Penfield and Rafuse [1], from which it is seen that the abrupt-junction diode is somewhat more efficient. Also, comparison of Greenspan's results with the results when $p_1$ is assumed to be given shows that the latter assumption leads to slightly lower efficiencies, the error being of the order of a few per cent. Greenspan's
asymptotic expressions for maximum efficiency when \( Q_B = Q_R / 2 \) are
\[
\eta = 0.00235 (f_c / f)^2 \quad \text{as} \quad Q_F \to 0
\]
and
\[
\eta = 1 - 26 (f / f_c) \quad \text{as} \quad Q_F \to \infty
\]
which are both lower than the corresponding expressions in the abrupt junction case.

Figure 18 shows the results obtained by Greenspan for the various powers, while Fig. 19 shows the corresponding resistance values. The comparative results for the abrupt junction case are given in figures 8.7 and 8.8 of Penfield and Rafuse \[1\].

3. Small Signal Approximations

One can obtain an approximate solution in the "small signal" case, as shown by Leeson and Weinreb \[9\]. This corresponds to a fixed bias being applied (for example, so that \( Q_B = Q_R / 2 \)) and \( p_1 \) and consequently \( p_N \) being small. Since \( a_N \) and \( b_1 \), the quantities of interest in calculating the efficiency, are found from the expansion of
\[
(1 + p_1 \cos \omega t \pm p_N \sin N \omega t)^\gamma
\]
we write this as
\[
(1 + p_1 \cos \omega t)^\gamma \pm \gamma (1 + p_1 \cos \omega t)^{\gamma-1} p_N \sin N \omega t
\]
\[
\pm \frac{\gamma (\gamma - 1)}{2!} (1 \pm p_1 \cos \omega t)^{\gamma-2} p_N^2 \sin^2 N \omega t + \cdots
\]
\( a_N \) is given by the terms involving \( \cos N \omega t \) in this expansion, which in the first term arise from \( \cos^N \omega t \) and higher-powered even terms; none arise from the second term; those in the third term arise from \( \cos^N \omega t \) and higher-powered terms, and so on for the other terms of the expansion. The contribution arising from \( \cos^N \omega t \) has a multiplier \( p_1^N \) in the first term, a multiplier \( p_1^N p_N^2 \) in the third term, and so on. Thus if \( p_1 \) and \( p_N \) are small compared to unity a reasonable approximation is to take account only of the term involving \( p_1^N \).

Thus
\[
a_N = - \frac{\gamma (\gamma - 1) (\gamma - 2) \cdots (\gamma - N + 1) p_1^N}{2^{N-1} N!}
\]  (54)

Similarly \( b_1 \) can be found making the same assumptions, as
\[
b_1 = \pm \frac{\gamma (\gamma - 1) (\gamma - 2) \cdots (\gamma - N + 1)}{2^{N-1} (N-1)!} p_1^{N-1} p_N
\]  (55)

The positive sign is taken when \( N \) is odd and the negative sign when \( N \) is even. Substituting in Eq. (52) gives
\[
\frac{R_s}{R_i} = \frac{R_s}{R_b} = 1 \pm \frac{(\gamma - 1) (\gamma - 2) \cdots (\gamma - N + 1) p_1^{N-2} p_N Q_F}{(N-1)! 2^{2N-2}}
\]
Fig. 18. Normalized input, output, and dissipated powers for graded junction doubler [8]. $P_1 =$ input power, $P_2 =$ output power, $P_D =$ dissipated power.

Fig. 19 Normalized source and load resistances for maximum power and efficiency in a graded junction varactor doubler [8].
and it can readily be shown that in general in this small signal case, with $p_1$ specified, $R_L = R_s$ gives maximum efficiency. $p_1$ and $p_N$ are related by Eq. (24), which gives

$$\pm \frac{Q_F a_N}{\gamma N 2^{\gamma - 1}} = p_N \left( \frac{R_L}{R_s} + 1 \right)$$

and using Eq. (56)

$$\frac{R_L}{R_s} = \frac{R_s}{R_s} = \left[ 1 + \left( \frac{Q_F (\gamma - 1)(\gamma - 2) \cdots (\gamma - N + 1)p_1^{N-1}}{N! 2^{\gamma + N - 2}} \right) \right]^{1/2}$$

From Eq. (53) the efficiency is

$$\eta = \left\{ \frac{Q_F p_1^{N-1}(\gamma - 1)(\gamma - 2) \cdots (\gamma - N + 1)^2}{N! 2^{\gamma + N - 2}} \right\} \times \left[ 1 + \left( \frac{Q_F p_1^{N-1}(\gamma - 1)(\gamma - 2) \cdots (\gamma - N + 1)}{N! 2^{\gamma + N - 2}} \right) \right]^{1/2}$$

Similarly conditions of maximum power output and so on can be evaluated.

These small signal expressions will inevitably yield pessimistic results, but are none the less convenient in a preliminary trial design, for estimation of the order of magnitude of the quantities involved. It is interesting to note that

![Graph](image_url)

**Fig. 20.** Comparison of approximate and exact calculation of efficiency as a function of $p_1$ for graded junction doubler with $Q_F = 200$. 
when $\gamma = 2$, $N = 2$, the abrupt junction doubler, the approximate expressions of Eqs. (57) and (58) agree with the exact expressions of Eqs. (45) and (46). This is, of course, because the neglected terms are all zero in that case. For the doubler where $\gamma \neq 2$ we have

$$
\eta = \frac{Q_F^2 p_1^2 (\gamma - 1)^2}{2^{2\gamma+2}} \left[ 1 + \left( 1 + \frac{Q_F^2 p_1^2 (\gamma - 1)^2}{2^{2\gamma+2}} \right)^{1/2} \right]^{-2}
$$

when $\gamma = 1.5$, $Q_F = 200$

$$
\frac{10^4 p_1^2}{32} \left[ 1 + \left( 1 + \frac{10^4 p_1^2}{32} \right)^{1/2} \right]^{-2}
$$

This can then be compared with the exact curve of Fig. 14b, as shown in Fig. 20, from which it is seen that the approximation is indeed quite useful, and Leeson and Weinreb [8] have found experimental verification of this fact.

4. Optimization for a Specified Generator Power

A situation which may arise in practice is that the available power from the fundamental frequency generator is specified. This can occur when the power handling capabilities of the varactor are greater than the power available from the generator, the latter being coupled to the varactor by a transformer. The available power is

$$
P_{AV} = \frac{E_g^2}{8R_g}
$$

and since efficiency is defined as the ratio of the output power to the available power from the generator, optimization of efficiency corresponds to optimization of output power. Equation (25) gives the output power as

$$
P_{out} = \frac{V_0^2 (Q_B/Q_R)^{2\gamma} a_N^2 R_L}{2(R_L + R_s)^2}
$$

also, from Eq. (19),

$$
E_g = -\omega q_1 (R_g + R_s) + V_0 \left( \frac{Q_B}{Q_R} \right)^{\gamma} b_1
$$

or

$$
\frac{E_g}{V_0} = \gamma \frac{p_1}{2Q_F} \left( \frac{R_s}{R_g} + 1 \right) + \left( \frac{Q_B}{Q_R} \right)^{\gamma} b_1
$$

and

$$
\pm Q_F a_N \left( \frac{Q_B}{Q_R} \right)^{\gamma} = \frac{N \gamma p_N}{2} \left( \frac{R_L}{R_s} + 1 \right)
$$

In the case of the abrupt junction doubler

$$
a_N = -\frac{p_1^2}{2}
$$

$$
b_1 = p_1 p_N = \frac{Q_F p_1^3 (Q_B/Q_R)^2}{4[(R_L/R_s) + 1]}
$$
Hence

\[
\frac{E_0}{V_0} = \frac{P_1}{Q_F} \left( R_0 + 1 \right) + \frac{Q_F P_1 Q_B Q_R}{4[(R_L/R_s) + 1]}
\] (62)

\[
\frac{P_{\text{out}}}{P_n} = \frac{P_1^4 Q_B Q_R}{8[(R_L/R_s) + 1]^2}
\] (63)

where \( P_n = V_0^2 / R_s \) is the normalization power. Equation (62) may be rewritten as

\[
2 \sqrt{2} \left( \frac{R_g}{R_s} \right)^{1/2} \left( \frac{P_{\text{AV}}}{P_n} \right)^{1/2} = \frac{P_1}{Q_F} \left( R_0 + 1 \right) + \frac{Q_F P_1^3 Q_B Q_R}{4[(R_L/R_s) + 1]}
\]

Now, assuming that for maximum efficiency the generator should be matched at the input, we have

\[
R_g = R_{\text{in}}
\]

or

\[
\frac{R_g}{R_s} = 1 + \frac{Q_F^2 P_1^2 Q_B Q_R}{4[(R_L/R_s) + 1]}
\] (64)

so that

\[
2 \frac{P_{\text{AV}}}{P_n} = \frac{P_1^2}{Q_F^2} \left[ 1 + \frac{Q_F^2 P_1^2 Q_B Q_R}{4[(R_L/R_s) + 1]} \right]
\] (65)

Differentiating Eq. (63) for a maximum gives

\[
\left( R_L - 1 \right) \frac{R_L}{R_s} \left( R_0 + 1 \right) \dot{P}_1 = 4 \frac{R_L}{R_s} \left( R_0 + 1 \right) \dot{P}_1
\] (66)

where \( \dot{P}_1 \) is the derivative of \( P_1 \) with respect to \( R_L / R_s \). Differentiating Eq. (65) leads to

\[
0 = 2 P_1 \ddot{P}_1 \left[ 1 + \frac{Q_F^2 P_1^2 Q_B Q_R}{4[(R_L/R_s) + 1]} \right]
\]

\[
+ \frac{P_1^2 Q_F^2 Q_B Q_R}{4[(R_L/R_s) + 1]^2} \left[ 2 P_1 \dot{P}_1 \left( R_0 + 1 \right) - P_1^2 \right]
\]

and substitution for \( \ddot{P}_1 \) from Eq. (66) gives

\[
R_L \left( R_L - 1 \right) \left( R_0 + 1 \right) = \left[ \frac{1 + \frac{Q_F^2 P_1^2 Q_B Q_R}{2}}{2} \right]^{1/2}
\] (67)

Substituting this value in Eq. (64) leads to

\[
2 \left( \frac{R_g}{R_s} - 1 \right) \left( \frac{R_0}{R_s} + 1 \right) = \left( \frac{R_0}{R_s} \right)^2 - 1
\]

or

\[
R_g = \frac{1}{2} (R_L + 1)
\] (68)
The maximum output power is then
\[
P_{\text{out}} = \frac{P_{1}^4(Q_{B}/Q_{R})^{4(1 + \frac{1}{2}[Q_{F}^{2}p_{1}^{2}(Q_{B}/Q_{R})^{4}]^{1/2})}}{8(1 + [1 + \frac{1}{2}[Q_{F}^{2}p_{1}^{2}(Q_{B}/Q_{R})^{4}]^{1/2}]^{2})} \tag{69}
\]
From Eq. (65) and using Eq. (68) we find
\[
p_{1}^{2} = \frac{4Q_{F}^{2}P_{AV}}{P_{n}(1 + R_{L}/R_{s})} \tag{70}
\]
so that
\[
P_{\text{out}} = \frac{2Q_{F}^{4}P_{AV}^{2}(Q_{B}/Q_{R})^{4}R_{L}/R_{s}}{P_{n}^{2}[(R_{L}/R_{s}) + 1]^{4}} \tag{71}
\]
\[
\eta = \frac{2Q_{F}^{4}P_{AV}(Q_{B}/Q_{R})^{4}R_{L}/R_{s}}{P_{n}[(R_{L}/R_{s}) + 1]^{4}} \tag{72}
\]
\[
\left(\frac{R_{L} + 1}{R_{s}}\right)^{2} - 1 = 2Q_{F}^{4}\left(\frac{Q_{B}}{Q_{R}}\right)^{4}\frac{P_{AV}}{P_{n}} \tag{73}
\]
If the available power is such that under these optimum conditions \(Q_{B} = \frac{1}{4}Q_{R}\), that is, the varactor is driven from the point of forward conduction to the reverse breakdown voltage, the above expressions may be computed. If on the other hand, the operation is such that the charge swing is between \(Q=0\) and some charge less than the reverse breakdown, so that \(Q_{R^{'}}\) is the maximum negative charge, then \(Q_{B} = Q_{R^{'}}/2\) and \(V_{0}\) is replaced by \(V_{0^{'}}\), where
\[
V_{0^{'}}/V_{0} = (Q_{R^{'}}/Q_{R})^{y}
\]
so that \(P_{n}\) is changed to \(P_{n^{'}}\), where
\[
P_{n^{'}}/P_{n} = (Q_{R^{'}}/Q_{R})^{2y}
\]
Similarly \(C_{0}\) is changed to \(C_{0^{'}}\), where
\[
C_{0^{'}}/C_{0} = (Q_{R}/Q_{R^{'}})^{y-1}
\]
and \(Q_{F}\) is changed by the inverse of this ratio.

Thus for the case where the varactor is underdriven in the manner described (which should be optimum) the equations may be appropriately scaled. However, in the underdriven case, the design of a multiplier is quite difficult since essentially Eqs. (73) and (67) provide a relationship between \(P_{AV}, p_{1},\) and \(Q_{B}\). Eq. (61) provides a relationship between \(p_{N}, p_{1}, Q_{B},\) and \(P_{AV}\), and Eq. (31) enables one to determine \(Q_{B}\) as a function of \(p_{1}\) and \(p_{N}\). These equations must then be solved simultaneously to obtain the efficiency, \(R_{s}\), and \(R_{L}\), which is quite a tedious process and must be performed numerically. If results are computed for the situation where \(P_{AV}\) is such as to fully drive the varactor and these results are normalized to \(P_{n}\) and \(Q_{F}\), then scaling may be accomplished for any other case by using \(P_{n^{'}}\) and \(Q_{F^{'}}\) found from the transformations given above.
When $P_{AV}$ is specified the low-frequency efficiency is
\[ 1 - 22(f/f_c) \]
Since the available power is specified, the output power and efficiency will tend to zero as $f/f_c$ tends to infinity. In other cases the input power must be continuously increased as $f$ tends to infinity in order to provide the asymptotic output power.

5. Optimization of Efficiency for a Given Output Power

A problem sometimes encountered is that the output power of the multiplier is specified and it is desired to optimize the efficiency in order to be able to provide the minimum possible input power. Equation (25) gives the output power as
\[ \frac{P_{out}}{P_n} = \frac{1}{2^{2^7+1}} \frac{a_n^2(R_L/R_s)}{[(R_L/R_s) + 1]^2} \]
when $Q_B = \frac{1}{2}Q_R$. In the abrupt junction doubler $a_n^2 = -P_1^2/2$ so that
\[ \frac{P_{out}}{P_n} = \frac{p_1^4(R_L/R_s)}{128[(R_L/R_s) + 1]^2} \]
and, assuming that the input is matched for maximum efficiency,
\[ \eta = \frac{R_L/R_s}{(1 + R_L/R_s) + (64/p_1^2 Q_F^2)(1 + R_L/R_s)^2} \]
or
\[ \eta = \frac{R_L/R_s}{(1 + R_L/R_s)(1 + (64/Q_F^2) K (R_L/R_s)^{1/2})} \]
where $K = (P_n/128P_{out})^{1/2}$. Setting
\[ \partial \eta/\partial (R_L/R_s) = 0 \]
gives
\[ (R_L/R_s)^{3/2} - (R_L/R_s)^{1/2} - (Q_F^2/2\sqrt{2})(P_{out}/P_n)^{1/2} = 0 \]
If we write this equation as
\[ x^3 - x - k = 0 \]
Cardan's solution gives
\[ x = \left( \frac{R_L}{R_s} \right)^{1/2} = \left\{ \frac{1}{2} + \left[ \frac{1}{4} + \frac{k^3}{27} \right]^{1/2} \right\}^{1/3} + \left\{ \frac{1}{2} - \left[ \frac{1}{4} + \frac{k^3}{27} \right]^{1/2} \right\}^{1/3} \]
so that $R_L/R_s$ can readily be found if $Q_F$ and $P_{out}/P_n$ are known, from which the efficiency and input power are found.
E. The Forward-Driven Varactor Multiplier

In all the analyses discussed so far the charge on the variable capacitance has been constrained to lie between the charge at the reverse breakdown voltage, \( Q_R \), and zero, at all times. This implied that no forward conduction takes place; however, if forward conduction does take place it is evident that the power-handling capabilities of the varactor are improved because of the larger permissible voltage swing. Forward conduction occurs if the voltage across the variable capacitance exceeds the contact potential, \( \phi \), and since rectification will generally take place a resistive component of current will flow through the junction and the incremental conductance is given by the usual equation

\[
g(v) = g_0 \exp \frac{eV}{kT}
\]

where \( k \) is the Boltzmann constant, \( T \) is the absolute temperature, and \( e \) is the electronic charge. The voltage \( \phi \) is such that the conductance given above is negligible when the forward voltage is less than \( \phi \). Various expressions for the incremental capacitance of the junction with forward applied bias are discussed by Davis [10], with the general conclusion that an expression of the form

\[
C = C_0 \exp (KeV/kT)
\]

describes the behavior.

If the frequency of operation of the varactor is sufficiently high, current rectification fails and the conductance term may be regarded as approximately zero. For ease of manipulation it is assumed that when the forward voltage is greater than \( \phi \) the incremental capacitance is infinite, and consequently the voltage across this capacitance is zero, even though charge is stored. Thus the model proposed by Davis [10] for the forward-driven varactor is that the voltage across the variable capacitance is constrained, at all times, to lie between the reverse breakdown voltage and zero, while the charge can vary between the value at the reverse breakdown voltage (which is negative) and some positive value. Figure 21 shows a sketch of the voltage-charge relationship to be used. In the ensuing analysis it has been assumed, as for all previous cases that \( Q=0 \) when \( V=\phi \).

Now if we assume

\[
Q = Q_B + q_1 \cos \omega t + q_N \cos (N\omega t + \varphi)
\]

or

\[
\frac{Q}{Q_B} = 1 + \frac{q_1}{Q_B} \cos \omega t + \frac{q_N}{Q_B} \cos (N\omega t + \varphi)
\]

and if the charge swing is such that it varies between \( Q_R \) (\( Q_R < 0 \)) and some positive charge \( Q_D \), and if we take \( \varphi = \pm \pi/2 \) as was the case for the tuned
condition previously assumed, then the charge waveform is symmetrical with respect to the average charge \( Q_B \). In this case the peak-to-peak charge swing is given by

\[
\frac{Q_{pp}}{Q_R} = 2\left(\frac{Q_B}{Q_R} - 1\right)
\]

or

\[
\frac{Q_{pp}}{|Q_R|} = 2\left(\frac{|Q_B|}{Q_R} + 1\right)
\]

and \( Q_B = \frac{1}{2}(Q_D + Q_R) \).

In order to determine the allowable limits in this case, when \( Q_B \) is given, the time at which the peak positive or negative charge occurs is found and

![Diagram](image)

**Fig. 21.** Voltage-charge model for the overdriven varactor.

the charge at this time is set equal to \( Q_D \) or \( Q_R \). For the case where \( N=2 \), \( \varphi = +\pi/2 \), and

\[
Q = Q_b + q_1 \cos \omega t - q_N \sin 2\omega t
\]

\( q_1 \) and \( q_N \) are assumed to be negative as before and Eq. (31) then gives

\[
q_1 \left[ 16 \left( \frac{q_N}{q_1} \right)^2 - 1 + \left[ 1 + 32 \left( \frac{q_N}{q_1} \right)^2 \right]^{1/2} \right]^{1/2} \left[ 3 + \frac{1 + 32(q_N/q_1)^2}{16\sqrt{2}(q_N/q_1)} \right]
\]

\[
= Q_R - Q_B = \frac{1}{2}(Q_R - Q_D)
\]

Thus \( Q_D \) obviously increases the allowed values of \( q_1 \) and \( q_N \) and for a given ratio of \( q_N \) to \( q_1 \) the largest value of \( q_1 \) is scaled by a factor \((Q_R - Q_D)/Q_R\).

The analysis then proceeds in exactly the same way as before, except that the factor \( Q_B/Q_R \) must not be set equal to 1/2 as was done for the varactor driven between the point of forward conduction and the reverse breakdown point. Equation (27) is used to calculate the efficiency and Eq. (25) the output
Fig. 22. Efficiency as a function of normalized average charge with $f_{in} = f_{in}/f_c$ as a parameter [10]: (a) abrupt-junction doubler; (b) graded-junction doubler.
Fig. 23. Normalized output power as a function of normalized input frequency with normalized average charge as a parameter [10]: (a) abrupt-junction doubler; (b) graded-junction doubler.
power. The Fourier coefficients are the same if \( p_1 \) and \( p_N \) are each multiplied by \( (Q_R/2Q_B) \). The \( Q \) factor, \( Q_F \), remains the same and so if \( Q_B \) is known all the previous results may be rescaled to give the performance in the overdriven case. Davis [10] has carried out extensive computations for both abrupt junction \((\gamma = 2)\) and graded junction \((\gamma = 1.5)\) doublers, and all of the following results are taken from his thesis.

First, with the previously assumed resonant condition, and with given values of \( Q_B/Q_R \) and \( Q_F \), the values of \( q_1 \) and \( q_N \) were numerically varied to find the absolute maximum efficiency, and curves obtained for both the abrupt junction and graded junction case, showing efficiency as a function of \( \dot{q}_0 = Q_B/Q_R \) with \( f^2 = f/f_c = 1/Q_F \) as a parameter. These curves are shown in Figs. 22a and 22b. Again, under the same conditions, \( q_1 \) and \( q_N \) were varied to find the absolute maximum power output in both cases, and these results are shown in Figs. 23a and 23b. The values of \( R_q (= R_m) \) resulting from these operations are shown in Figs. 24a and 24b, while the load resistance required to obtain optimum efficiency or optimum power output is shown in Figs. 25a and 25b. In the case where the varactor is not overdriven \( \dot{q}_0 = 0.5 \), while \( \dot{q}_0 = 0 \) is the case where the stored charge in the forward direction is equal in magnitude to \( Q_R \).

It is evident from Fig. 22 that the improvement in efficiency obtained by overdriving is considerably greater in the case of the graded junction than in the case of the abrupt junction. At each frequency there is a maximum of efficiency at a particular drive level, for the model used. The curves of power output in Fig. 23 show that the output power in the case of the graded junction doubler is increased about 10 times by overdriving to \( \dot{q}_0 = 0 \), but that the abrupt junction type still delivers the greater maximum power. When \( \dot{q}_0 = 0.5 \) the asymptotic expression for the output power in the abrupt junction doubler is 0.00195\( P_n \) and in the graded junction diode, 0.000662\( P_n \), while the corresponding values when \( \dot{q}_0 = 0 \) are 0.00781\( P_n \) and 0.00710\( P_n \). The differences between the resistance values for maximum efficiency and maximum output power are small, and at high frequencies these differences tend to disappear.

Davis has also investigated the effect of using a condition other than the resonance situation previously assumed. In this investigation the phase angle of the output circuit was varied to maximize efficiency and output power. The main differences occurred in the most heavily driven cases, being at most about 10% (of the “resonant value”) for the maximum efficiency and power output for the abrupt junction case, with corresponding figures of about 4% in the graded junction case. It is also of interest that these largest deviations were found when \( Q_F = 10 \), and that for the case where the varactor was not overdriven the difference between the “resonant” condition and the optimum is quite negligible, so that for convenience in analysis “resonance” in the sense previously discussed may be assumed with negligible error.

Thus, on the basis of the model studied by Davis [10] it can be concluded
Fig. 24. Normalized input resistance as a function of normalized average charge, with normalized input frequency as a parameter [10]: (a) abrupt-junction doubler; (b) graded-junction doubler.
Fig. 25. Optimum normalized load resistance as a function of normalized average charge with normalized input frequency as a parameter [10]: (a) abrupt-junction doubler; (b) graded-junction doubler.
that there is considerable advantage, particularly in terms of output power, to be obtained from overdriving the varactor. Whether the model would be valid if the drive level were increased indefinitely is doubtful, but it is clear that at least a moderate overdriving is beneficial.

Grayzel [11] has investigated the overdriven varactor doubler for various values of $m$ in the $C-V$ equation. The results are shown in Figs. 26 and 27.

![Figure 26](image)

**Fig. 26.** Maximum efficiency as a function of the diode $C-V$ index, $m$ [11].

![Figure 27](image)

**Fig. 27.** Maximum output power as a function of the diode $C-V$ index, $m$ [11].

These were obtained by varying the fundamental and second harmonic charge coefficients as well as the phase of the second harmonic component to find the optimum efficiency and output power, as functions of $m$ and the $Q$ factor, $Q_F$. The results show an improved efficiency as $m$ is reduced, and for $m$ less than about 0.1 the efficiency is constant. The maximum power output however, decreases as $m$ is reduced and becomes constant for $m$ less than about 0.1. Thus, the behavior of any varactor where $m$ is less than about 0.1 will be approximately the same.

**F. The Charge-Storage Diode Multiplier**

As the index $m$ is made to tend to zero, the charge-voltage characteristic approaches that of the ideal charge-storage, or step-recovery, diode shunted by a constant capacitance, $C_0$, biased by the voltage $\phi$. Figure 28 shows a sketch of the $Q-V$ characteristic as $m \rightarrow 0$ with the limiting case also shown. The usual assumption that in the forward-driven region charge is stored without increase in voltage ($C = \infty$) is implied. Thus Grayzel’s results [11] imply that this idealized model should have about the same performance as that obtained when $m < 0.1$ in his analysis. Leenov and Uhlir [12] have considered the generation of harmonics by the idealized charge-storage diode whose $Q-V$ characteristic is shown in Fig. 29, when the embedding circuit is purely resistive; that is, no frequency separation or resonating circuits were employed. Thus the analysis consisted in calculating the Fourier components of power developed in the load resistor of the circuit shown in Fig. 30. Table I summarizes the results obtained and provides a comparison with the ideal non-linear resistive diode as analyzed by Page [13]. The considerable advantage of
the nonlinear capacitance diode, particularly at high orders of multiplication, is apparent from the table.

Roulston [14] has analyzed the performance of the charge-storage diode with tuned output circuits and compared the results with the untuned case. The main result of this investigation was to show that there is a fundamental limit on efficiency as the ratio of generator to load resistance is varied. In the case of

![Figure 28](image1.png) Charge-voltage characteristic of a varactor with \( m = 0 \).

![Figure 29](image2.png) Charge-voltage characteristic for an idealized charge-storage diode.

| Table I |

<table>
<thead>
<tr>
<th>Harmonic</th>
<th>Ideal resistive (NR) diode, ( d )</th>
<th>Ideal capac. (NC) diode, ( d )</th>
<th>Advantage of NC over NR diode (( d ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>13.5</td>
<td>9.2</td>
<td>4.3</td>
</tr>
<tr>
<td>3</td>
<td>23.2</td>
<td>15.1</td>
<td>8.1</td>
</tr>
<tr>
<td>4</td>
<td>27.5</td>
<td>18.5</td>
<td>9.0</td>
</tr>
<tr>
<td>5</td>
<td>32.0</td>
<td>20.4</td>
<td>11.6</td>
</tr>
<tr>
<td>6</td>
<td>34.7</td>
<td>22.6</td>
<td>12.1</td>
</tr>
<tr>
<td>7</td>
<td>37.7</td>
<td>24.2</td>
<td>13.5</td>
</tr>
<tr>
<td>8</td>
<td>39.9</td>
<td>25.7</td>
<td>14.2</td>
</tr>
<tr>
<td>9</td>
<td>41.9</td>
<td>26.4</td>
<td>15.5</td>
</tr>
<tr>
<td>10</td>
<td>43.8</td>
<td>27.3</td>
<td>16.5</td>
</tr>
<tr>
<td>100</td>
<td>83.9</td>
<td>47.9</td>
<td>36.0</td>
</tr>
</tbody>
</table>
a × 2 and a × 10 multiplier, improvements of about 3 db in conversion loss were obtained, as compared to the untuned case. When the generator circuit was resonated rather than the load, improvements of about 5 db were found. The same type of circuit, but with the load circuit tuned, has also been analyzed by Hedderly [15]. His circuit is shown in Fig. 31 and the analysis assumes that when the voltage across the diode becomes positive conduction takes place and continues until the charge passed through the diode becomes zero, at which time conduction ceases leaving energy of, say, \( \frac{1}{2}Li_n^2 \) stored in the inductance. While the diode is nonconducting this energy decays (preferably) in a periodic manner dissipating harmonic power in the load. The following approximations are made in the ensuing analysis.

1. The natural resonant frequency of the tuned circuit is sufficiently high for the energy stored in the capacitor to be negligible as compared to that in the inductor at the end of the conduction period, and also for the capacitance to be ignored during conduction. Both these approximations are quite good when

![Fig. 30. Resistive circuit for generation of harmonics using ideal charge-storage diode.](image)

![Fig. 31. Circuit for generation of harmonics using ideal charge-storage diode, where output circuit is resonant.](image)

![Fig. 32. Current and voltage waveforms for the circuit of Fig. 31 with \( Q_s = 1.75 \), \( k = -0.427 \) [15].](image)
the order of the desired harmonic is large and the accuracy improves as the order is increased.

(2) The energy stored in the inductor at the end of one conduction period is all dissipated before the start of the next conduction period.

Typical voltage and current waveforms for this circuit are shown in Fig. 32, where

\[ v = E\{\sin(\omega t - \theta) + k\} \]

\[ k = \sin \theta \]

and

\[ Q_g = L\omega/R_g = \tan \varphi \]

Hence

\[ i = \frac{E}{R_g}\left[ -\sin \theta + \frac{\sin(\theta + \varphi)}{1 + Q_g^2} \right] \exp\left( -\frac{R_g t}{L} \right) + \sin \theta + \frac{\sin(\omega t - \theta - \varphi)}{(1 + Q_g^2)^{1/2}} \]

On the assumption that conduction ceases when \( q = \int i dt = 0 \), one then calculates the current at the end of the conduction period and hence the energy stored in \( L \). This energy is stored (and removed) \( \omega/2\pi \) times per second, and thus the available power for dissipation at the harmonic output frequency is determined. The nonconducting state is then considered as this energy decays from the inductor and is dissipated in the load resistor, the output circuit being tuned to the desired harmonic \((n\omega)\). The total loss in the circuit is then the difference between the power dissipated in the load and the available power from the generator at the fundamental frequency. This is composed of a loss \( L_g \) in converting the generator power to power stored in the inductance at the end of conduction, and a loss \( L_L \) in converting this stored power into ohmic power in the load. Figure 33a and 33b show the results of this analysis in terms of \( k \), the ratio of bias voltage to peak fundamental frequency voltage, \( Q_g \) and \( Q_L = n\omega C/g_L \). The total loss is \( L_g + L_L \). It is seen that the lowest loss for a generator of this type is about 8.5 db and is independent of the order of multiplication. This illustrates the great attraction of this form of capacitive non-linearity for high-order harmonic generation as compared to the more conventional C-V characteristic found in varactors.

Uhlir [16] has analyzed a bridge-type doubler using diodes whose charge-voltage relationship has the form shown in Fig. 34. \( S_{\text{max}} \) is the maximum elastance and \( S_{\text{min}} \) the minimum elastance. Thus the charge-storage effect is accounted for by a capacitance \( S_{\text{min}}^{-1} \) which in the analysis was made large enough so that its value was immaterial but which was finite for mathematical convenience. The analysis consisted in solving the response of a bridge-doubler circuit as a function of time on the application of a sinusoidal fundamental frequency voltage. It was found that when the load and source resistances (assumed equal) were small the harmonic current waveform in some conditions...
of bias did not settle down to a steady value and hysteresis effects were found in the efficiency-bias curve as shown in Fig. 35. $V_H$ is the bias voltage, with the fundamental frequency voltage having a peak value of 100 volts. $R_g$ and $R_L$ are normalized to the diode series resistances. The efficiency falls to zero when the peak voltage does not reach the point of discontinuity in the $Q$-$V$ curve. When $R_L$ and $R_g$ are increased a more satisfactory situation with respect to bias
voltage is achieved. Figures 36a and 36b show the results obtained for efficiency as a function of normalized frequency in the high $Q$-factor and low $Q$-factor cases, respectively, from which it is seen that considerable danger exists in the use of a very high $Q$ circuit since the frequency of maximum efficiency is very close to the sharp dropoff on one side of the curve. Again the use of a lower $Q$ factor is shown to lead to a more desirable practical situation.

Fig. 36. Frequency response of bridge doubler [16]: (a) high-$Q$ ($R_L = R_s = 0.1904R_o$); (b) moderate-$Q$ ($R_L = R_s = 0.5R_o$).
G. Bandwidth in Multipliers without Idlers

All the analyses described so far have been concerned with "spot-frequency" behavior (except for that just discussed). It is evidently of considerable interest to know the variation of efficiency with frequency, particularly when the design of variable-frequency sources of microwave power is contemplated.

1. Bandwidth of Doubler with Single Tuned Load and Source

The simplest case of the single tuned input and output circuit abrupt junction doubler has been analyzed by Wolfson [17]. Using the expressions derived by Penfield and Rafuse [1] and the same circuit as Fig. 7a except that the current $I_N$ is reversed, we find that the following equations hold:

\[
Z_{in} = \frac{V_1}{I_1} = R_s + \frac{S_0}{j\omega} + R_s m_2 \frac{\omega_c}{\omega} e^{-j\theta} \quad (78)
\]
\[
Z_L = \frac{V_2}{-I_2} = -R_s - \frac{S_0}{2j\omega} + R_s \frac{m_1^2 \omega_c}{4m_2 \omega} e^{j\theta} \quad (79)
\]
\[
R_{in} = R_s \left( 1 + m_2 \frac{\omega_c}{\omega} \cos \theta \right) \quad (80)
\]
\[
R_L = R_s \left\{ \frac{m_1^2 \omega_c}{4m_2 \omega} \cos \theta - 1 \right\} \quad (81)
\]
\[
|I_1|^2 = 4 \left( \frac{\omega}{\omega_c} \right)^2 \frac{P_n}{R_s} m_1^2 \quad (82)
\]
\[
|I_2|^2 = 16 \left( \frac{\omega}{\omega_c} \right)^2 \frac{P_n}{R_s} m_2^2 \quad (83)
\]
\[
P_{in} = 8P_n m_1^2 \left( \frac{\omega}{\omega_c} \right)^2 \left( 1 + m_2 \frac{\omega_c}{\omega} \cos \theta \right) \quad (84)
\]
\[
P_{out} = 32P_n m_2^2 \left( \frac{\omega}{\omega_c} \right)^2 \left\{ \frac{m_1^2 \omega_c}{4m_2 \omega} \cos \theta - 1 \right\} \quad (85)
\]

where

\[
m_k = \frac{|S_k|}{S_{max}} \quad R_s \omega_c = S_{max} \quad (86)
\]

\[
I_1 = |I_1| \cos \omega t \quad I_2 = |I_2| \cos (2\omega t - \theta)
\]

and the other symbols have the same meaning as before.

When $\theta = 0$, the previously designated resonance condition, and the input circuit is tuned, the load and source inductances are chosen so that

\[
L_1 = \frac{S_{0,\text{res}}}{\omega_0^2} \quad L_2 = \frac{S_{0,\text{res}}}{4\omega_0^2}
\]
Fig. 37. Efficiency as a function of fractional detuning with fixed bias in an abrupt-junction doubler [17]: (a) maximum frequency with relative input frequency $f_{in}$ as a parameter.

and the bandwidth is defined as

$$\Delta = (\omega_2 - \omega_1)/\omega_0$$

where $\omega_2$ and $\omega_1$ are the frequencies at which the output power is 3 dB less than the maximum. The frequency at which maximum output power occurs may not be $\omega_0$.

Equations (78) and (79) yield

$$Z_{in} = R_s + \frac{S_0}{j\omega} + u \left[ Z_L^* + R_s - \frac{S_0}{2j\omega} \right]$$
ANALYSIS OF VARACTOR HARMONIC GENERATORS

Fig. 37—continued. (b) Efficiency with load resistance relative to the optimum value as a parameter.

where

\[ u = (2m_2/m_1)^2 \] (89)

and * denotes complex conjugate.

Considering the case of the single tuned source and load impedances, where in the condition of maximum output power the varactor is fully driven so that \( m_{0,\text{res}} = 1/2 \), \( m_0 \) remains at this value off resonance if the bias is so adjusted, while if self-bias or fixed bias is used \( m_0 \) changes as \( m_1 \) and \( m_2 \) change. Then if \( \delta = (\omega - \omega_0)/\omega_0 \), the source resistance \( Z_g \) may be written as

\[ Z_g = R_g + j \frac{S_{0,\text{res}}}{\omega_0} \left[ 1 + \delta + \frac{S_0}{S_{0,\text{res}}} \frac{1}{1 + \delta} \right] \] (90)
$Z_0$ includes $S_0$. Similarly,

$$Z_L = R_L + j \frac{S_{0,\text{res}}}{2\omega_0} (1 + \delta)$$

and does not include $S_0$. Now if, at $\omega_0$, $R_\pi = R_{\text{in}}$, the operation of the doubler is described by the following equations:

$$A^2 E m_1^2 (uB + 1) (1 + \delta)^2$$

$$= \frac{64 D (uB + 1)}{16 (D + uB + 1)^2 + (1/A^2) (2 - u)^2 [1 + \delta - (2m_0)/(1 + \delta)]}$$

(91)

$$m_1^2 = u(1 + \delta)^2 \left[ 4 A^2 B^2 + \frac{1}{4} \left( 1 + \delta - \frac{2m_0^2}{1 + \delta} \right) \right]$$

(92)

$$m_0^2 = C - \frac{1}{4} m_1^2 (u + 4)$$

(93)

$$\eta = A E m_1^2 u (B - 1) (1 + \delta)^2$$

(94)

in the case of fixed bias, while in the case of self-bias Eqs. (91) and (92) together with (94) also hold, and in addition, equations to determine the value of the average elastance at each frequency are needed. These are

$$\cos \omega t_0 + \sqrt{(u)} \cos (2\omega t_0 - \theta) = 0.$$  

(95)

$$m_0 + 2m_1 \sin \omega t_0 + m_1 \sqrt{(u)} \sin (2\omega t_0 - \theta) = 0$$  

(96)

$$\cos \theta = \frac{2}{m_1} \sqrt{(u)} AB(1 + \delta)$$  

(97)

where $t_0$ is the time at which minimum elastance occurs and

$$A = \frac{\omega_0}{\omega_c}, \quad B = \frac{R_L}{R_\pi} + 1, \quad C = \left( \frac{V_R + \phi}{V_B + \phi} \right)$$

$$D = \frac{R_\pi}{R_\pi} = u_0 B + 1, \quad E = \frac{8P_n}{P_{AV}} = \frac{1}{A^2 Dm_1^2_{\text{res}}}$$

(98)

If in the maximum efficiency condition the efficiency is the largest possible (where the breakdown limit is not exceeded), then $R_L$ is known, and also the other quantities required to determine $A$, $C$, $D$, and $E$. The value of $R_L$ when the maximum efficiency is optimized is called $R_{L,\text{opt}}$. The numerical results obtained from this study are shown in Figs. 37a and 37b for both $R_{L,\text{opt}}$ and other values of $R_L$, with fixed bias, and in Figs. 38a and 38b for self-bias. The dotted lines indicate that the computer solution converged to the value shown but failed to meet the required error criterion. This may indicate hysteresis effects of the type revealed by Uhlir [16]. It is seen that for a high-$Q$-factor varactor with optimum efficiency the maximum bandwidth is about 30% for fixed bias and about 38% for self-bias. Details of the analysis indicate that the
main cause of falling efficiency with detuning is the circuit resonances and not the conversion efficiency \( P_{\text{out}}/P_{\text{in}} \) of the varactor itself. The results of Figs. 37 and 38 obviously indicate a possibility of tradeoff between efficiency and bandwidth by varying \( R_L \), but no specific results are available on this point.

Käch [18] has derived results for the bandwidth of an abrupt junction doubler where the diode series resistance is ignored. These correspond to the case where the frequency of operation is very small compared to the diode cutoff frequency. When the diode operates with maximum possible efficiency it is found that the bandwidth is 24% and at maximum output power the bandwidth is 25.7%. Although these are limiting results they are of the same order as those found by Wolfson [17] and give an indication of the performance to be expected. In practice, however, the stray reactances associated with both the diode and the physical realization of the required circuit will reduce the bandwidth somewhat, so that the results given above can be taken as upper bounds. The stray reactances mean, in effect, that the rate of charge of the external circuit reactance is greater than that of the simple inductance assumed in the analysis.

2. General Limitations on Bandwidth

By using generator and load impedances which are more complicated than the simple \( R-L \) series circuits considered so far it is possible to obtain an improved bandwidth for a given efficiency. Grayzel [19] has considered this problem and his results are discussed below. The notation is the same as that used in connection with the case of the single tuned circuits, except that two additional variables

\[
y = m_2 \left( \frac{\omega_c}{\omega} \right) \cos \theta \tag{98}
\]

\[
F = \frac{32P_n}{\cos^2 \theta} \left( \frac{\omega}{\omega_c} \right)^4 \tag{99}
\]

are introduced. When these are substituted

\[
\frac{Z_2}{R_s} = \frac{Z_L}{R_s} + 1 + \frac{S_0}{2j\omega R_s} = \frac{ye^{j\theta}}{u \cos \theta} \tag{100}
\]

\[
\frac{R_2}{R_s} = \frac{y}{u} \tag{101}
\]

\[
\frac{Z_1}{R_s} = \frac{Z_{\text{in}}}{R_s} - 1 - \frac{S_0}{j\omega R_s} = \frac{ye^{-j\theta}}{\cos \theta} \tag{102}
\]

\[
\frac{R_1}{R_s} = y \tag{103}
\]
Fig. 38. Efficiency as a function of fractional detuning with self-bias in an abrupt-junction doubler [17]; (a) maximum efficiency with relative input frequency $f_{in}$ as a parameter.

\[
P_{in} = \frac{P y^2 (y+1)}{u}
\]

(104)

\[
P_{out} = \frac{P y^2 (y-u)}{u}
\]

(105)

and the conversion efficiency $\epsilon (= P_{out}/P_{in})$ is

\[
\epsilon = \frac{(y-u)}{(y+1)}
\]

(106)

Thus $P_{in}$, $P_{out}$, $R_1$, and $R_2$ are functions only of $u$ and $y$ and may be plotted on a $(u-y)$ plane. The breakdown curve on the other hand is a function of $\theta$.
as well as of $u$ and $y$ but it is found that the variation in this curve as $\theta$ varies from 0° to 90° is small. From Eq. (88)

$$Z_1 = uZ_2^*$$

(107)

Thus Fig. 39 shows three equivalent representations of the doubler.

The efficiency is

$$\eta = T\epsilon = T(\omega) \frac{1-u(\omega)/y(\omega)}{1+1/y(\omega)}$$

(108)
where $T = P_{in}/P_{AV}$ and is the transmission coefficient through the lossless matching network of Fig. 39. Now

$$T = 1 - \left| \frac{Z_g(\omega) - (R_s + uZ_2(2\omega))}{Z_g(\omega) + (R_s + uZ_2^*(2\omega))} \right|^2$$  \hspace{1cm} (109)$$

![Diagram](https://via.placeholder.com/150)

**Fig. 39.** Equivalent circuits of an abrupt-junction varactor doubler with matching network [19].

But by definition of $T$

$$P_{AV} T(\omega) = P_{in}$$  \hspace{1cm} (110)$$

and using Eqs. (101), (103), and (104)

$$T = A \left( \frac{\omega}{\omega_c} \right)^4 u(\omega)|Z_2(2\omega)|^2 [u(\omega)R_2(\omega) + R_s]$$  \hspace{1cm} (111)$$

where now

$$A = \frac{2P_n}{P_{AV} R_s^3}$$

Thus $A$, and $P_{AV}$ are assumed to be known quantities and if $Z_g(\omega)$ and $Z_2(2\omega)$ are known Eqs. (109) and (111) may be solved to obtain $T(\omega)$ and $u(\omega)$, and hence the efficiency may be found. The problem is then to choose $Z_g(\omega)$ and $Z_L(\omega)$ in order to maximize the bandwidth for a given efficiency. It is assumed throughout the analysis that $S_0$, the average elastance, remains constant as the frequency varies.
ANALYSIS OF VARACTOR HARMONIC GENERATORS

Considering first the case where \( R_s \ll R_L \), it is shown [20] that the impedance \( Z(s) \), which is the input impedance of a lossless network terminated in a resistance \( R \) and having a series elastance \( S_0 \) at the input, is restricted by

\[
\int_{\omega_1}^{\omega_2} \frac{\ln |Y(\omega)| \, d\omega}{\omega((\omega^2-\omega_1^2)(\omega^2-\omega_2^2))^{1/2}} \leq \frac{\pi}{2\omega_1 \omega_2} \ln \left( \frac{2\omega_1 \omega_2}{(\omega_2-\omega_1) S_0} \right) \quad (112)
\]

Helgesson [20] has shown that one impedance satisfying Eq. (112) with the equality sign has a constant magnitude in the passband and is

\[
Z(s) = \frac{(\omega_2-\omega_1) S_0}{2\omega_1 \omega_2} \left[ \left( 1 + \left( \frac{s^2 + \omega_1 \omega_2}{(\omega_2-\omega_1) s} \right)^{2 \frac{1}{2}} \right) \frac{s^2 + \omega_1 \omega_2}{(\omega_2-\omega_1) s} + \frac{k(\omega_2-\omega_1)}{2s} \right] \quad (113)
\]

and

\[
|Z(j\omega)| = \frac{(\omega_2-\omega_1) S_0}{2\omega_1 \omega_2} \quad (\omega_1 \leq \omega \leq \omega_2) \quad (114)
\]

\[
|Z(j\omega)| = \frac{W-1}{2\omega_1 W} \quad (115)
\]

where

\[
W = \omega_2/\omega_1 \quad (116)
\]

\( Z(s) \) as given by Eq. (113) is one of the general class

\[
Z_k(s) = A_k \left[ \left( 1 + \left( \frac{s^2 + \omega_1 \omega_2}{(\omega_2-\omega_1) s} \right)^{2 \frac{1}{2}} \right) \frac{s^2 + \omega_1 \omega_2}{(\omega_2-\omega_1) s} + \frac{k(\omega_2-\omega_1)}{2s} \right] \quad (117)
\]

where

\[
A_k = \frac{2(W-1) S_0}{\omega_2(k(W^2+1)-2W(k-2))} \quad (118)
\]

which all fulfill Eq. (112) with the equality sign. The impedances \( Z_k(s) \) are all merely impedance scaled replicas of each other with the addition of some extra series capacitance at the input. For maximum bandwidth both \( |Y_g(\omega)| \) and \( |Y_2(2\omega)| \) should satisfy Eq. (112) with the equality sign and for perfect match \( |Y_g| = |Y_2| \), which leads to

\[
\int_{\omega_1}^{\omega_2} \frac{\ln (2|u(\omega)|) d(2\omega)}{2\omega((4\omega^2-\omega_1^2)(\omega^2-\omega_2^2))^{1/2}} = 0 \quad (119)
\]

where \( \omega_{12} \) and \( \omega_{22} \) are the lower and upper limits of the band at the harmonic frequency. Obviously \( u(\omega) \equiv 2 \) satisfies this equation and if \( u(\omega) \) is not constant it should vary about the value 2. Operation along the breakdown curve maximizes the values of \( |Y_2| \) and \( |Y_g| \), leading to maximum bandwidth. The breakdown curve is approximated by

\[
10m_2 = u^{0.268}
\]
With good agreement in the vicinity of \( u=2 \), and using this relation and the fact that

\[
Z_2 = \frac{2S_0m_2}{u\omega_1}
\]

(120)
together with Eqs. (119) and (112) for \( Y_2(\omega) \), one finds

\[
16.67 = \frac{4(W+1)}{W-1} \quad \text{or} \quad W = 1.63
\]

This indicates the maximum bandwidth for perfect match which can be achieved by working on the (approximate) breakdown curve. A slightly higher value (1.65) can be achieved by making \( u/m_2 \) constant, which is the absolute maximum for a perfect match. This, however, cannot be achieved in practice since \( Z_2(s) \) is not a realizable impedance.

If some mismatch is allowed so that

\[
|Y_g(\omega)| = \frac{r}{u} |Y_2(2\omega)|
\]

then

\[
\int_{\omega_1}^{\omega_2} \ln \left[ \frac{2r}{u(\omega)} \right] d(2\omega) = 0
\]

and the integral restriction of \( Y_2(2\omega) \) is

\[
\int_{\omega_1}^{\omega_2} \ln \left( \frac{2r}{m_2} \right) d(2\omega) = \frac{\pi}{2}\omega_{12}\omega_{22} \ln \left( \frac{4(W+1)}{W-1} \right)
\]

The maximum bandwidth is then given by

\[
\frac{u}{m_2} \bigg|_{u=2r} = \frac{4(W+1)}{W-1}
\]

and occurs when \( u=2r \) on the breakdown curve. In this case

\[
T = \frac{2r(1+\cos 2\theta)}{1+2r\cos 2\theta+r^2}
\]

where it is assumed that \( \theta = \theta_g \) where \( Z_g = |Z_g| < \theta_g \). One can then find \( T(\theta) \) and \( W \), having chosen a value of \( r \), and for each value of \( \theta \) plot \( T \) as a function of \( W \). These results are shown in Fig. 40 and represent an absolute upper bound on the efficiency. This bound cannot be achieved in practice as the phase condition \( \theta = \theta_g \) is certain to be violated over some part of the bandwidth, so that one can define a usable portion of \( W \) such that \( T \) remains constant over this band.

The next step is to calculate \( Z_2 \) for the desired operating conditions using Eqs. (117) and (120).
When $R_s$ is not negligible the performance of the lossless doubler is first calculated and then the modified values of efficiency are calculated. When the frequency is very high the reactance of $S_0$ becomes negligible and so $Z_2$ and $Z_g$ are made purely resistive.

If the required impedances $Z_2$ and $Z_g$ are approximated by an infinite ladder the transmission coefficient will not remain constant over the operating band, but a bandwidth can be defined such that over this range $T$ exceeds some

![Graph](https://via.placeholder.com/150)

**Fig. 40.** An upper bound on the transmission coefficient of a broadband varactor doubler as a function of the fractional bandwidth [19].

value $T_{\text{min}}$. The impedances $Z_g$ and $Z_2$ to be approximated are chosen to give the highest efficiency for a given bandwidth. When the (approximated) functional dependence of $Z_g$ and $Z_2$ are known $u$ is found as the solution of a cubic equation in the lossless case, and of a quartic equation when $R_s$ is included. The results of these computations are shown in Fig. 41, where $\eta$ represents the minimum efficiency in the corresponding bandwidth, and the approximating impedances have been chosen so as to maximize the minimum value.

The impedance $Z_0(s)/A_0 (k=0)$ can be approximated by the bandpass form of the finite networks shown in Figs. 42a, and 42b, which were derived
Fig. 41. Maximum efficiency as a function of bandwidth for an abrupt-junction doubler with broadband matching network, consisting of an infinite ladder [19]. $\omega_{L1}$ is the lowest input frequency at which the multiplier operates and the curves are for (a) lossless; (b) $\omega_{L1}/\omega_C=10^{-3}$; (c) $\omega_{L1}/\omega_C=2 \times 10^{-3}$; (d) $\omega_{L1}/\omega_C=4 \times 10^{-3}$; (e) $\omega_{L1}/\omega_C=8 \times 10^{-3}$; (f) $\omega_{L1}/\omega_C=16 \times 10^{-3}$.

Fig. 42. Low-pass prototype circuits for matching in broadband varactor doublers [19].

from those proposed by Bode [21]. For other values of $k$ these networks are scaled in impedance and a series capacity is added to the input. Again the solution can be found in the same way as for the infinite ladder, by solving
ANALYSIS OF VARACTOR HARMONIC GENERATORS

a cubic or quartic equation for \( u \). It was found that results comparable to the infinite ladder could be obtained with the finite networks used.

Thus varactor doublers can be designed with extremely broad bandwidths over which the output power remains relatively constant as the input frequency is tuned. Upper bounds on the bandwidth as a function of efficiency have been found and it is established that comparatively uncomplicated networks provide a very useful performance.

H. VARIATIONS IN MULTIPLIER CIRCUIT CONFIGURATION

Apart from the circuit considered exclusively so far, several variations are both possible and useful. These may occur for a variety of reasons, such as increased bandwidth, separation of harmonics, or generation of high-order harmonics. The two variations principally discussed here are the four-diode bridge doubler, and the series-diode circuit.

1. The Bridge Doubler

The circuit of the four-diode bridge doubler is shown in Fig. 43. If the diodes are identical then \( e_k = f(i_k) \) for all \( k \) and

\[
\begin{align*}
  i_1 &= i_a - i_d = i_c - i_b \\
  i_2 &= i_d + i_c \\
  e_1 &= e_c - e_d = e_a - e_b \\
  e_2 &= e_c + e_b \\
  e_2 &= f(i_c) - f(i_d) = f(i_a) - f(i_b)
\end{align*}
\]

Fig. 43. Bridge doubler circuit.
If one then reverses each diode and the polarities of voltages and currents, the symmetry of the circuit guarantees that the same currents with opposite directions flow through corresponding diodes so that

\[ i_c = i_a \quad i_d = i_b \quad e_c = e_a \quad e_d = e_b \]

The current and voltage relationship for each varactor (abrupt-junction) is then

\[ E_k = \left( R_s + \frac{S_0}{jk\omega_0} \right) I_k + \sum_{r=\infty}^{r=-\infty, r\neq k} \frac{S_{k-r} I_r}{jk\omega_0} \]

as before. In the analysis [19] one need only deal with the diodes \( a \) and \( b \) since \( c \) and \( d \) have identical variables to these, via the relations expressed above.

If \( I_1 \) and \( I_2 \) are restricted to be pure sinusoids at the input and output frequencies, respectively, then

\[ E_{ak} = \left( R_s + \frac{S_0}{jk\omega_0} \right) I_{ak} - \frac{S_{\text{max}}}{Q_R} \sum_{r=\infty}^{r=-\infty, r\neq k} \frac{I_{a(k-r)} I_{ar}}{2r k \omega_0^2} \]
\[ E_{bk} = \left( R_s + \frac{S_0}{jk\omega_0} \right) I_{bk} - \frac{S_{\text{max}}}{Q_R} \sum_{r=\infty}^{r=-\infty, r\neq k} \frac{I_{b(k-r)} I_{br}}{2r k \omega_0^2} \]

But

\[ i_1 = i_a - i_b \quad e_1 = e_a - e_b \]
\[ i_2 = i_a + i_b \quad e_2 = e_a + e_b \]

which gives

\[ E_{1k} = \left( R_s + \frac{S_0}{jk\omega_0} \right) I_{1k} - \frac{S_{\text{max}}}{Q_R} \sum_{r=\infty}^{r=-\infty} \frac{I_{1(k-r)} I_{2r} + I_{2(k-r)} I_{1r}}{2r k \omega_0^2} \]
\[ E_{2k} = \left( R_s + \frac{S_0}{jk\omega_0} \right) I_{2k} - \frac{S_{\text{max}}}{Q_R} \sum_{r=\infty}^{r=-\infty} \frac{I_{1(k-r)} I_{1r} + I_{2(k-r)} I_{2r}}{2r k \omega_0^2} \]

But when \( i_1 \) is constrained to be sinusoidal only \( I_{1(1)} \) and \( I_{1(-1)} \) exist and with \( i_2 \) also purely sinusoidal only \( I_{2(2)} \) and \( I_{2(-2)} \) exist, so that the nonzero components of \( E_1 \) and \( E_2 \) are

\[ E_{1(1)} = \left( R_s + \frac{S_0}{j\omega_0} \right) I_{1(1)} + \frac{S_{\text{max}} I_{2(2)} I_{1(-1)}}{Q_R 4 \omega_0^2} \quad \text{(at the fundamental frequency)} \]
\[ E_{1(3)} = - \frac{S_{\text{max}} I_{2(2)} I_{1(1)}}{6 \omega_0^2} \quad \text{(at the third harmonic)} \]
\[ E_{2(2)} = \left( R_s + \frac{S_0}{j\omega_0} \right) I_{2(2)} - \frac{S_{\text{max}} I_{1(1)}^2}{Q_R 4 \omega_0^2} \quad \text{(at the second harmonic)} \]
\[ E_{2(4)} = - \frac{S_{\text{max}} I_{2(2)}^2}{Q_R 16 \omega_0^2} \quad \text{(at the fourth harmonic)} \]

(121)
Thus a filter in the input is needed to reject the third harmonic while a filter in the output is needed to reject the fourth. The current components can be related to the elastance components of an individual diode to show that $E_{1(1)}$ and $E_{2(2)}$ are identical to the equations for a single diode operating with the same current. Thus the bridge operates in the same way as the straightforward doubler, except that filters at fundamental and second harmonic frequencies are not required.

![Diagram](https://via.placeholder.com/150)

**Fig. 44.** Bridge doubler with extra diode to eliminate fourth harmonic [19].

The analysis described above was given by Grayzel [19], who notes that the output filter at $4\omega_0$ limits the input tuning range to 2:1; however it is shown that the addition of another diode with the same cutoff frequency as the fourth but half the series resistance, as shown in Fig. 44, will remove the need for a fourth harmonic filter in the output. The voltage across this diode is

$$E_k = \frac{1}{2} \left( R_s + \frac{S_0}{jk\omega_0} \right) I_{2k} + \frac{S_{\max}}{Q_R} \sum_{r=-\infty}^{\infty} \frac{I_{2(k-r)} I_{2r}}{2k r \omega_0^2}$$

This produces a fourth harmonic voltage component equal but opposite in sign to that of Eq. (121), thus enabling the input frequency to be tuned to $3\omega_0$ and giving greater tuning bandwidth.

The bridge doubler is therefore a versatile circuit which removes the necessity for filters at the fundamental and second harmonic frequency, and permits tuning over a 2:1 frequency range with the same performance as the conventional circuit, which is restricted to 1.5:1 in tuning range. The analyses of Uhlir [16] and Wolfson [17] previously discussed were based on the bridge circuit.
2. The Series Diode

This type of circuit is shown in Fig. 7b. The fundamental and output harmonic voltages across the diode are constrained to pure sinusoids so that

\[ V = V_B + V_1 \cos \omega t - V_N \cos (N\omega t + \varphi) \]

The corresponding charges are then found using Eq. (7).

\[ Q = Q_R \left( \frac{\phi - V_B}{V_0} - \frac{V_1}{V_0} \cos \omega t + \frac{V_N}{V_0} \cos (N\omega t + \varphi) \right)^{1-n} \]

so that in the abrupt junction case the index is 1/2 rather than the corresponding value 2 in the parallel-diode circuit, thus enabling harmonics other than the

Fig. 45. Power-loss curves for an abrupt-junction multiplier with series diode [7]: (a) \( N = 2 \), (b) \( N = 3 \), (c) \( N = 10 \). The symbols are defined in the text.

Fig. 46. Conversion loss as a function of harmonic number for a typical Hughes Aircraft Company diode. Capacity at the dc bias point = 1 pf, \( V_{RB}/2V_R = 40 \), \( R_s = 4 \Omega \) where \( V_{RB} \) is the reverse breakdown voltage and \( V_R \) is the voltage at which forward conduction losses reduce the diode \( Q \) factor to one half of its original value [22].
second to be generated without idlers. If the diode series resistance is replaced by its parallel equivalent, which introduces only a small error when \( Q_F \) is large [7], the analysis proceeds exactly as before on a dual basis. The tuning condition is found in the way previously described for the parallel diode. Roulston [7] has carried out this analysis and some of his results are shown in Figs. 45a, 45b, and 45c.

These results show that a minimum conversion loss occurs at a particular value of the parameter \( G_{m}/\omega_n C_0 \) where \( Q_{dn} \) is the diode \( Q \) factor at the output frequency, and \( G_i \) is a conductance composed of the load and a term representing the diode losses. It is assumed that input circuit is matched. These results indicate that an abrupt junction diode may be used in the series mode to generate high harmonics with reasonable efficiency, while in the parallel mode only the second harmonic can be generated.

Johnson [22] has also analyzed the series-diode circuit, using a parallel conductance to represent the diode losses. Figure 46 shows the results obtained for maximum efficiency as a function of the order of the output harmonic.

### IV. VARACTOR HARMONIC GENERATORS WITH IDLERS

The introduction of so-called "idler" circuits into harmonic generators as described by Diamond [23] can materially increase the available efficiencies.

The idea is that currents (or voltages) at other than the fundamental frequency and output harmonic are allowed to exist in virtually lossless circuits. These currents are mixed with the fundamental frequency and with each other to produce additional components of current at the desired harmonic. In the case of the abrupt junction diode connected in the parallel configuration this is the only way to produce output at other than the second harmonic, while for other types of diode the efficiency is considerably increased. This requires some additional circuit complexity, but should be compared with the possibility of having to use a chain of low-order multipliers without idlers to achieve the same result.

![Fig. 47. Circuit of varactor harmonic generator with idlers.](image-url)
The circuit of a general multiplier with idlers is shown in Fig. 47. The filters $F_1$, $F_m$, and $F_N$ constrain the currents at the fundamental, idler, and output frequencies, respectively, to be purely sinusoidal. Again one writes

$$Q_1 = q_1 \cos \omega t \quad Q_m = q_m \cos (m\omega t + \varphi_m) \quad Q_N = q_N \cos (N\omega t + \varphi_N)$$

(122)

and, if $Q_B = Q_R/2$,

$$v + V_R = V_0 \left[ 1 - \left( \frac{1}{2}\right)^\gamma \left( 1 + p_1 \cos \omega t - \sum_p p_m \cos (m\omega t + \varphi_m) - p_N \cos (N\omega t + \varphi_N) \right) \right]$$

(123)

where the summation is taken over all the idlers. One can also write

$$v + V_R = V_0 \left[ 1 + \left( \frac{1}{2}\right)^\gamma \left( a_o + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t) \right) \right]$$

(124)

The Fourier coefficients $a_n$ and $b_n$ are evaluated as before. The circuit equations are

$$E_g \sin (\omega t + \alpha) = i_1 (R_g + R_s) \sin \omega t + L_1 i_1 \cos \omega t + \frac{V_0 a_1}{2^\gamma} \cos \omega t + \frac{V_0 b_1}{2^\gamma} \sin \omega t$$

(125)

for the input circuit,

$$\frac{V_0}{2^\gamma} (a_m \cos m\omega t + b_m \sin m\omega t) = R_s i_m \sin (m\omega t + \varphi_m) + L_m m\omega i_m \cos (m\omega t + \varphi_m)$$

(126)

for each idler circuit, and

$$\frac{V_0}{2^\gamma} (a_N \cos N\omega t + b_N \sin N\omega t) = (R_s + R_L) i_N \sin (N\omega t + \varphi_N) + L_N N\omega i_N \cos (N\omega t + \varphi_N)$$

(127)

for the output circuit. The resonant condition in the input circuit as before gives

$$L_1 = \frac{-V_0 a_1}{2^\gamma \omega i_1}$$

and

$$E_g = i_1 (R_g + R_s) + \frac{V_0 b_1}{2^\gamma}$$

(128)

and if the input circuit is matched

$$R_g = R_s + \frac{V_0 b_1}{2^\gamma i_1}$$

(129)
so that

\[ E_s = 2i_1 \left( R_s + \frac{V_0 b_1}{2 \gamma i_1} \right) \]  \hspace{1cm} (130)

Solving Eq. (127) for the output circuit yields

\[ i_N(R_L + R_s) = (V_0/2\gamma)(a_N \sin \varphi_N + b_N \cos \varphi_N) \]  \hspace{1cm} (131)

\[ L_N N\omega i_N = (V_0/2\gamma)(a_N \cos \varphi_N - b_N \sin \varphi_N) \]  \hspace{1cm} (132)

Similarly for each idler circuit

\[ i_m R_s = (V_0/2\gamma)(a_m \sin \varphi_m + b_m \cos \varphi_m) \]  \hspace{1cm} (133)

\[ L_m m\omega i_m = (V_0/2\gamma)(a_m \cos \varphi_m - b_m \sin \varphi_m) \]  \hspace{1cm} (134)

In principle one can now solve Eqs. (131)–(134) for specified values of \( i_1, R_L, \) \( L_m, \) and \( L_N \) or specified values of \( i_1, R_L, \) \( \varphi_m, \) and \( \varphi_N. \) The latter is an easier approach since Eqs. (132) and (134) then merely give the necessary values of \( L_N \) and \( L_m, \) which are not generally of prime interest. In this case Eqs. (131) and (133) are solved simultaneously for \( i_m \) and \( i_N. \) These may be rewritten as

\[ p_N \frac{N\gamma 2^{\gamma-1}}{Q_F} \left( \frac{R_L}{R_s} + 1 \right) = a_N \sin \varphi_N + b_N \cos \varphi_N \]  \hspace{1cm} (135)

\[ p_m \frac{m\gamma 2^{\gamma-1}}{Q_F} = a_m \sin \varphi_m + b_m \cos \varphi_m \]  \hspace{1cm} (136)

The usual expressions for efficiency and output power in terms of \( p_N, p_1, \) and \( b_1 \) apply also in this case.

If operation in the forward driven region is not allowed then the charge on the varactor must remain within the range

\[ 0 > Q > Q_R \]

or

\[ 0 \leq 1 + p_1 \cos \omega t - \sum_m p_m \cos (m\omega t + \varphi_m) - p_N \cos (N\omega t + \varphi_N) \leq 1 \]  \hspace{1cm} (137)

assuming that the average change is held at \( Q_R/2. \) The plotting of this curve is not straightforward as in the nonidler case, but once \( \varphi_N \) and \( \varphi_m \) are known it expresses a limiting relationship between the various charges. In analyzing the circuit one knows that maximum efficiency and output power are obtained by operating along the breakdown curve, so that instead of choosing a value of \( R_L, \) the specification that operation on this curve is required is inserted instead, thus enabling a solution to be found. The absolute maxima of efficiency and output power can be found by exploring all possible combinations of values of \( p_1, \varphi_N, \) and \( \varphi_m \) with operation on the breakdown curve.
A. Resonant Solution

While the procedure described above can be used to provide an optimum solution for any particular configuration, numerical methods are, in general, necessary and little insight is gained regarding the behavior of the multiplier. It is therefore advantageous, as in the generator without idlers, to consider some simplified cases. The first simplification is to assume that maximum efficiency and output power are obtained for particular values of \( \varphi_N \) and \( \varphi_m \). The rule given for this “linearized” resonance condition in the nonidlers case applies equally well here, so that if \( k_m, l_m, k_N, l_N \) are the portions of \((a_m, b_m)\) and \((a_N, b_N)\) that are independent of \( \varphi_m \) and \( \varphi_N \), respectively, the angles are given by

\[
\varphi_m = \tan^{-1}(k_m/l_m) \tag{138}
\]

\[
\varphi_N = \tan^{-1}(k_N/l_N) \tag{139}
\]

Here \( k_N \) and \( l_N \) are integrals which in general must be evaluated by numerical techniques, and for chosen values of \( p_1 \) and \( R_L \) a simultaneous solution of Eqs. (138), (139), (135), and (136) is necessary; \( k_m \) and \( l_m \) arise from the expression

\[
(1 + p_1 \cos \omega t - \bar{p}_N \cos (N \omega t + \bar{\varphi}_N)
\]

where \( p_m \) is set equal to zero, \( \bar{p}_N \) is the resulting value of \( p_N \), and \( \bar{\varphi}_N \) the resulting value of \( \varphi_N \). Similarly \( k_N \) and \( l_N \) are found. It is evident that in general this procedure is just as cumbersome as the general procedure suggested previously. However, in the abrupt junction doubler, one is forced to choose one of the values of \( m = 2 \), since otherwise no output can be obtained. In applying the rule the second harmonic idler is treated first and \( k_2, l_2 \) are given by the expression

\[
(1 + p_1 \cos \omega t)^2
\]

since removal of the second harmonic idler reduces all other outputs to zero. Thus \( l_2 = 0 \) and so \( \varphi_2 = \pm \pi/2 \). The next highest idler is then removed and the phase (determined by \( k_m \) and \( l_m \)) depends only on \( p_2 \), whose phase is already known; similarly, proceeding in ascending order we can find \( k_r, l_r \). It should, of course, be realized that the phase angles found in this way will, in general, be slightly nonoptimum, and in fact when there are many idlers the process of finding the phases is quite complicated. The difficulty arises when the removal of a particular idler in order to calculate \( k_m \) and \( l_m \) does not render all higher harmonics zero.

Penfield and Rafuse [1] have extensively treated many cases of the abrupt-junction multiplier with idlers, operating at maximum efficiency or maximum power, with the “linearized” tuning condition described above. Most of the original work on the subject is due to Diamond [23], and complete results for
triplers, quadruplers, quintuplers, and octuplers are given together with asymptotic formulas, at high and low frequencies.

Additional results have been given by Rafuse [24], concerning the improvement obtainable through use of idlers additional to the minimum number required to obtain an output. It was found that a considerable improvement in efficiency is possible. Defining the following quantities, which are the low-frequency asymptotic values,

\[
\eta \approx \exp \{-\alpha (\omega_{\text{out}}/\omega_c)\} \\
P_{\text{out}}/P_n = \beta (\omega_0/\omega_c) \\
R_{\text{in}}/R_s = A (\omega_c/\omega_0) \\
R_l/R_s = B (\omega_c/\omega_0)
\]  

(140)  
(141)  
(142)  
(143)

where \(\omega_0\) is the fundamental frequency and \(\omega_{\text{out}}\) is the output frequency, Rafuse [24] has compared the performances of a quadrupler with a second harmonic idler and of a quadrupler with both second harmonic and third harmonic idlers. The superior performance of the circuit with an extra idler is apparent from the results in the tabulation.

<table>
<thead>
<tr>
<th></th>
<th>(\alpha)</th>
<th>(\beta)</th>
<th>(A)</th>
<th>(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd Harmonic only</td>
<td>15.6</td>
<td>0.0196</td>
<td>0.15</td>
<td>0.0513</td>
</tr>
<tr>
<td>2nd and 3rd Harmonics</td>
<td>11.4</td>
<td>0.0226</td>
<td>0.096</td>
<td>0.0625</td>
</tr>
</tbody>
</table>

Burckhardt [25] has analyzed a further selection of circuits for both the fully driven and overdrive cases. The same notation as given in Eqs. (140)–(143) is used together with the drive level defined as

\[
\text{Drive} = \frac{Q_{\text{max}} - Q_{\text{min}}}{Q_R}
\]

(Drive = 1 corresponds to the fully driven varactor, and drive greater than unity defines the overdriven case.) It was found that both the power-handling capabilities and the efficiency can be increased by overdriving the diode. The idealized case \((m=0)\) is included in the analysis, and is found to yield the highest efficiency which occurs when the drive = 2.0. For example, in the case of the quadrupler with second harmonic idler \(m=0\) gives \(\alpha = 10.3\), \(\beta = 0.0298\), with drive = 2.0, while \(m=1/3\) gives \(\alpha = 12.2\), \(\beta = 0.0351\), with drive = 1.6, and \(m=1/2\) gives \(\alpha = 14.1\), \(\beta = 0.53\), with drive = 1.6. Again the considerable improvement to be obtained through use of additional idlers is evident also when the varactor is overdriven. For example, the quadrupler with a third
harmonic idler has $\alpha = 8.1$, $\beta = 0.0438$ at drive = 1.6 with $m = 1/3$, as compared to the values quoted above for the case with second harmonic idler only.

It is difficult, however, to draw any more specific conclusion than that extra idlers improve the efficiency, but to what extent, in general, it is difficult to say. Both Rafuse [24] and Burckhardt [25] have shown that by using a number of idlers the abrupt-junction multiplier can be made to have virtually the same efficiency (when fully driven) in generating a specific harmonic as for the ideal ($m = 0$) case without idlers and with drive = 2, and that the power handling of the circuit with idlers is somewhat better.

Kaylie [26] has developed equivalent circuits for the abrupt-junction multiplier with idlers. A circuit is given for each harmonic frequency in terms of the current amplitudes at the various harmonics, and their relative phase angles. The transfer of power between the various harmonics is expressed in terms of a voltage generator whose voltage is independent of the particular harmonic, and a nonlinear resistance whose value is a function of the particular harmonic. In order to make use of these circuits one must specify the operating conditions, in order to determine the current amplitudes. One might, for example, specify the "linearized" resonant condition, together with either the load resistance or the condition that operation should be confined to the breakdown curve. The circuits are then useful in writing down the circuit equations for the particular situation chosen.

B. Solution with a Specified Input Current

In a way analogous to the solution for a specified value of $p_1$, in the case without idlers, the solution with specified idlers and specified $p_1$ can be found. Again this procedure is not optimum but does provide some insight into the operation. As an example let us consider the abrupt-junction tripler. The frequencies involved are $\omega$, $2\omega$, $3\omega$, the normalized charge is

$$1 + p_1 \cos \omega t - p_2 \cos (2\omega t + \phi_2) - p_3 \cos (3\omega t + \phi_3)$$

and the circuit equations are (130), (135), and (136). In this case the linearized resonance condition gives $\phi_2 = -\pi/2$, $\phi_3 = 0$ so that

$$\frac{12p_3}{Q_F} \left( \frac{R_L}{R_s} + 1 \right) = b_3 \quad \frac{8p_2}{Q_F} = -a_2$$

$$\frac{E_g}{V_0} = \frac{2p_1}{Q_F} \pm \frac{b_1}{4}$$

and

$$b_1 = p_1 p_2 + p_2 p_3$$  \hspace{1cm} (144)

$$a_2 = -\frac{p_1^2}{2} + p_1 p_3$$  \hspace{1cm} (145)

$$b_3 = p_1 p_2$$  \hspace{1cm} (146)
which gives

\[ p_3 = \frac{Q_f^2 p_1^3}{192(1 + R_L/R_s) + 2Q_f^2 p_1^2} \]  \hspace{1cm} (147)

and

\[ \frac{R_s}{R_e} = 1 + \frac{Q_f^2}{32} \left( \frac{p_1}{2} - p_3 \right) \left( p_1 + p_3 \right) \]  \hspace{1cm} (148)

The efficiency is then

\[ \eta = \left\{ \frac{3Q_f^2 p_1^2}{192[(R_L/R_s) + 1] + 2Q_f^2 p_1^2} \right\}^2 \frac{R_L}{R_s} \]  \hspace{1cm} (149)

where \( R_s \) is given in Eq. (148). Differentiating \( \eta \) with respect to \( R_L \) to obtain a maximum gives

\[ \frac{R_L}{R_s} = \frac{1}{24} \left\{ \frac{13Q_f^2 p_1^4 + 1344Q_f^2 p_1^2 + (192)^2}{64 + Q_f^2 p_1^2} \right\}^{1/2} \]  \hspace{1cm} (150)

Thus one can compute the input and output resistances and the efficiency for a particular value of \( p_1 \), the maximum output power and efficiency occurring when \( p_1 \) is such that the diode operates on the breakdown curve. In the case of the graded junction multiplier the expressions cannot be obtained in closed form, and numerical methods must be used.

In a similar manner one may compute optimum conditions for the case of a specified available power, or a specified power output. Reasonable solutions can be obtained in this manner which yield results which do not differ greatly from the true optima.

C. SMALL SIGNAL APPROXIMATIONS

Utsunomiya and Yuan [27] have developed a small signal approximate solution for multipliers with idlers. Again, as in the case of the circuit without idlers, only a limited number of terms are taken in computing the voltage across the variable capacitance. In fact the voltage is taken to be a quadratic function of the charge, and the series resistance is neglected. The output power and efficiency are then calculated as functions of the fundamental frequency charge.

Thus it is assumed that

\[ v = \beta_1 q + \beta_2 q^2 \]  \hspace{1cm} (151)

where \( \beta_1 \) and \( \beta_2 \) are found by expanding the true charge-voltage relationship in a Taylor series and taking only the first two terms. Thus the ac voltage is given by

\[ \frac{v}{V_0} = \frac{1}{2\gamma} \left\{ \gamma p + \frac{\gamma(\gamma - 1)p^2}{2} \right\} \]  \hspace{1cm} (152)
so that

\[ \beta_1 = \frac{\gamma}{2^\gamma} \]  
\[ \beta_2 = \frac{\gamma(\gamma-1)}{2^\gamma+1} \]

and \( p = p_1 \cos \omega t + p_2 \sin 2\omega t - p_3 \cos 3\omega t \) for the tripler with second harmonic idler and the same phase condition as before. The various harmonic components may be calculated from Eq. (152). When \( \gamma = 2 \) (the abrupt-junction case), Eq. (152) is, of course, exact, and is a relatively good approximation for other diodes.

In the case of the graded-junction diode, \( \gamma = 1.5, \beta_1 = 0.53, \) and \( \beta_2 = 0.133 \) as compared to \( \gamma = 2, \beta_1 = 0.5, \) and \( \beta_2 = 0.25 \) for the abrupt-junction case. The quantities \( b_1, a_2, b_3 \) appearing in Eqs. (144)–(146) are proportional to \( \beta_2 \) and are therefore reduced for a given value of \( p_1 \). Using this small signal approximation a graded-junction multiplier may be analyzed as readily as one employing an abrupt junction, and a first-order comparison of the results obtained. Utsunomiya and Yuan calculated the efficiency and output power of a 1-2-4-5 multiplier using the small signal approximation, and described experimental results which show that the approximation is quite useful.

ACKNOWLEDGMENT

The author is indebted to Mr. P. J. R. Laybourn, Department of Electrical and Electronic Engineering, University of Leeds, for his considerable assistance in the preparation of this chapter.

LIST OF PRINCIPAL SYMBOLS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_m )</td>
<td>Fourier coefficient of ( \cos m\omega t )</td>
</tr>
<tr>
<td>( b_m )</td>
<td>Fourier coefficient of ( \sin m\omega t )</td>
</tr>
<tr>
<td>( C )</td>
<td>Capacitance</td>
</tr>
<tr>
<td>( C_0 )</td>
<td>Minimum diode capacitance</td>
</tr>
<tr>
<td>( C_0' )</td>
<td>Diode capacitance with zero applied voltage</td>
</tr>
<tr>
<td>( E )</td>
<td>A voltage</td>
</tr>
<tr>
<td>( E_g )</td>
<td>Generator e.m.f.</td>
</tr>
<tr>
<td>( e )</td>
<td>Electronic charge</td>
</tr>
<tr>
<td>( f_c )</td>
<td>Diode cutoff frequency</td>
</tr>
<tr>
<td>( f )</td>
<td>Frequency normalized to ( f_c )</td>
</tr>
<tr>
<td>( i_m )</td>
<td>Magnitude of a current at the ( m )th harmonic frequency</td>
</tr>
<tr>
<td>( I_m )</td>
<td>Instantaneous value of a current at the ( m )th harmonic frequency</td>
</tr>
<tr>
<td>( j )</td>
<td>( \sqrt{-1} )</td>
</tr>
<tr>
<td>( k_m )</td>
<td>Fourier coefficient of ( \cos m\omega t )</td>
</tr>
<tr>
<td>( L_m )</td>
<td>Inductance in the ( m )th harmonic circuit</td>
</tr>
<tr>
<td>( m )</td>
<td>Exponent of the varactor ( C-V ) characteristic</td>
</tr>
<tr>
<td>( m_n )</td>
<td>Ratio of the ( R )th harmonic component of elastance to the maximum elastance</td>
</tr>
<tr>
<td>( n )</td>
<td>Designation of a general harmonic number</td>
</tr>
<tr>
<td>( N )</td>
<td>Output harmonic number</td>
</tr>
<tr>
<td>( p_m )</td>
<td>Ratio of the ( m )th harmonic charge amplitude to half the maximum charge</td>
</tr>
<tr>
<td>( p^0 )</td>
<td>Derivative of ( p ) with respect to ( (R_L/R_S) ).</td>
</tr>
</tbody>
</table>
ANALYSIS OF VARACTOR HARMONIC GENERATORS

\[ P_n \] normalization power
\[ P_{AV} \] available power from a generator
\[ P_{out} \] output power
\[ \dot{P} \] power normalized to \( P_n \)
\[ q_m \] magnitude of a charge at the \( m \)th harmonic frequency
\[ Q_m \] instantaneous value of a charge at the \( m \)th harmonic frequency
\[ Q_B \] charge at the bias voltage
\[ \dot{Q} \] charge relative to \( Q_B \)
\[ Q_R \] charge at reverse breakdown
\[ Q_F \] diode \( Q \) factor
\[ R \] resistance
\[ R_s \] internal resistance of a generator
\[ R_L \] load resistance
\[ R_S \] diode series resistance
\[ S \] elastance
\[ s \] complex frequency
\[ S_0 \] maximum elastance
\[ T \] transmission coefficient
\[ V \] voltage
\[ v \] ac component of voltage
\[ V_{d.c.} \] dc component of voltage
\[ V_R \] reverse breakdown voltage
\[ V_0 \] sum of \( V_R \) and the contact potential
\[ Y \] admittance
\[ Z_L \] load impedance
\[ Z_{in} \] input impedance
\[ \alpha \] phase angle of generator voltage
\[ \gamma \] (exponent) = \( 1/1 - m \)
\[ \Delta \] fractional bandwidth
\[ \epsilon \] conversion efficiency, ratio of output to input power
\[ \phi \] contact potential of a semiconductor junction
\[ \psi_m \] phase angle of the \( m \)th harmonic charge
\[ \eta \] efficiency
\[ \omega \] angular frequency
\[ \omega_{out} \] output angular frequency
\[ \omega_c \] angular cutoff frequency

REFERENCES


The design of multiplexers, which separate out frequencies in certain ranges from a spectrum of signals covering a large range of frequencies, is a problem frequently encountered in microwave engineering. The separation of the desired frequency bands can be accomplished using bandpass filters, or by

1 The work described herein was supported by the U.S. Army Electronics Laboratories, Fort Monmouth, New Jersey, under Contracts DA-36-039 SC-74862, DA-36-039 SC87398, and DA-36-039-AMC-00084(E).

2 Present address: Department of Electrical Engineering, University of California, Santa Barbara, California.
using bandpass and bandstop filters. However, the design problem is much more difficult than it might at first seem. If a number of ordinary bandpass filters are simply connected together, interaction effects will usually disrupt the performance of the system, unless the filters and their interconnections are very carefully designed.

One way of getting around this difficulty is to use directional filters [1, 2]. In theory, directional filters are ideal for this application because they have unity input VSWR at all frequencies, so that any number of them can be cascaded without undesirable interaction effects. However, in practice, the coupling apertures required in cavity-type directional filters are quite large, and the structures typically have a rather large off-channel VSWR. A number of such multiplexer units connected together, therefore, may have quite large interaction effects. Another disadvantage of directional filters is that if more than one or two cavity resonators are required, they become quite difficult to tune [3]. In this chapter, analyses are made of the potentialities of other forms of channel separation units, which are relatively easy to tune, which need not be physically complicated, and which (in theory) can be designed to have the constant resistance properties (or nearly constant resistance properties) of directional filters.

Another problem often occurring in complex microwave systems is that of designing diplexers to separate signals ranging from dc to a given cutoff frequency from signals having frequencies above the given cutoff frequency. Diplexers having precisely controlled frequency response characteristics are usually quite difficult to design and construct. Thus, new design techniques and diplexer configurations that provide simplification of both the precision design process and the fabrication process are needed. Work in this area is also described in this chapter.

The presentation here covers work on diplexers and multiplexers carried out at Stanford Research Institute under contracts sponsored by the U.S. Army Electronics Laboratory, Fort Monmouth, New Jersey. The work presented is not intended to cover all design techniques of possible interest. It will be noted in the presentation that certain narrow-band techniques (such as those of Fox cited in [4]), by the use of directional filters [1, 2], and the use of filters and hybrids [4, 5] are not discussed. The techniques described, however, are believed to have several advantages over these other techniques for many applications.

In the description of the theory and also the description of the microwave design techniques, reference is frequently made to either the singly terminated or the doubly terminated low pass prototype filter. It is convenient to describe these filter prototypes (with accompanying definitions) at this point in order to avoid repetition later on in the discussions.

The singly terminated, low pass, prototype filter is shown in Fig. 1. This
prototype filter has a resistive termination at the load end only, and is designed to be driven by an ideal source. In Fig. 1a, the first element is a shunt element and the source is therefore a zero-internal-admittance current generator. In Fig. 1b, the first element is a series element and the source is therefore a zero-internal-impedance voltage generator.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Singly terminated, low pass, prototype filter. (a) Driven by a zero-admittance current generator; (b) driven by a zero-impedance voltage generator.}
\end{figure}

The singly terminated filter has the property that its transmission attenuation response has the same behavior as the real part of its input impedance (for Fig. 1a) or admittance (for Fig. 1b). Thus, for the filter of Fig. 1a

\[ 20 \log \frac{I_s}{I_L} = 10 \log \frac{\text{Re}[Z_{in}]}{R_L} \]  

(1)

and for the filter of Fig. 1b

\[ 20 \log \frac{E_s}{E_L} = 10 \log \frac{\text{Re}[Y_{in}]}{G_L} \]  

(2)

Equations (1) and (2) represent properties of the singly terminated filter that will be used later in the design of multiplexers.

A doubly terminated, low pass, prototype filter and its dual are shown in Figs. 2a and 2b. A doubly terminated filter has resistive loading at both the source and load ends so that, in effect, the internal impedance of the source has been incorporated into the design.
Note in both Figs. 1 and 2 that the $g_i$ of the prototype filters have been designated. Tables of numerical values for the $g_i$ are available that give to the low pass, prototype filters of Figs. 1 and 2 prescribed Chebyshev or maximally flat responses [6, 7]. Also note that the parameters of the prototype filters are primed. The frequency variable of the prototype filters will also be primed, and this convention will be used consistently throughout this chapter whenever the prototype filter is referred to. The frequency variable and other parameters of derived filters will be unprimed.

The theory of diplexers and multiplexers presented in this chapter is general. However, the design techniques described are most suitable for devices operating in the microwave band, which is where our attention is focused. At lower frequencies, where lumped inductors and capacitors can be realized, the theory may be applied directly without the use of special design techniques. At frequencies above the microwave band, other techniques will probably be more appropriate.

The description in this chapter, although general, is often illustrated in terms of lumped-element filter circuits in order to explain more easily the concepts involved. For the most part, it is presented in nonmathematical terms by relying on existing designs of low pass, lumped-element prototype filters (that is, Chebyshev and maximally flat filters). However, in order to understand fully the concepts involved, it is necessary at this point to introduce the

\[ R'_0 = q_0 \]
\[ C'_n = q_n \]
\[ G'_n = q_{n+1} \]

\[ R'_1 = q_1 \]
\[ C'_n = q_n \]
\[ G'_n = q_{n+1} \]

Fig. 2. (a) The doubly terminated, low pass, prototype filter; (b) its dual. (From [7].)
definitions, and examine in some detail the properties, of minimum reactance and minimum susceptance networks. The presentation is mainly descriptive rather than rigorous, and is without proofs. The interested reader will find proofs and more complete discussion elsewhere [8, 9].

II. CONCEPTS OF MINIMUM REACTANCE AND MINIMUM SUSCEPTANCE NETWORKS

It is shown in books on network theory that the input impedance of a network composed of inductors, capacitors, and resistors may be expressed as a quotient of two polynomials [8, 9]:

$$Z_{in}(s) = \frac{a_0 + a_1 s + a_2 s^2 + \cdots + a_n s^n}{b_0 + b_1 s + b_2 s^2 + \cdots + b_m s^m}$$  (3)

where \( m-n = \pm 1 \) or 0. The variable \( s \) in Eq. (3) is denoted as the complex frequency and is equal to

$$s = \sigma + j\omega$$  (4)

where \( \omega \) is the familiar radian frequency and \( j \) is \( \sqrt{-1} \). In order to determine the input impedance as a function of the frequency variable \( j\omega \), one sets \( \sigma \) equal to 0 in Eq. (4) and evaluates Eq. (3) as a function of \( j\omega \).

The numerator and denominator polynomials of Eq. (3) may also be written in terms of their factors. The expression for \( Z_{in}(s) \) then becomes

$$Z_{in}(s) = H \prod_{i=1}^{n} \frac{(s-z_i)}{(s-p_i)}$$  (5)

where \( H \) is a constant equal to \( a_n/b_m \).

The factors of the numerator of \( Z_{in}(s) \) in Eq. (5) have been expressed in terms of the quantities \((s-z_i)\). The parameter \( z_i \) is generally denoted in the literature as a "zero" of \( Z_{in}(s) \), because for \( s=z_i \), \( Z_{in}(s)=0 \). The factors of the denominator of \( Z_{in}(s) \) have been expressed in terms of the quantities \((s-p_i)\), where \( p_i \) is generally denoted in the literature as a "pole" of \( Z_{in}(s) \). For \( s=p_i \), the denominator of \( Z_{in}(s) \) is zero and the magnitude of the function \( Z_{in}(s) \) goes to infinity.

**Definition:** An impedance function, such as is denoted in Eq. (5), in which none of the \( p_i \) are pure imaginary and for which \( n\leq m \), is referred to as a *minimum reactance network*. An impedance function, such as is denoted in Eq. (5), in which none of the \( z_i \) are pure imaginary and for which \( m\leq n \), is referred to as a *minimum susceptance network* [10].

Stated in another way, the input impedance of a minimum reactance network is finite for all real frequencies \( \omega \), including infinity. Similarly, the
input admittance of a minimum susceptance network is finite for all real frequencies $\omega$, including infinity.

An example of a minimum reactance network is the singly terminated low pass prototype filter shown in Fig. 1a. The input impedance of this network is finite for all frequencies including infinity. Note that the first element of the filter is a shunt capacitor. If the first element were a series inductor, the network would not be minimum reactance, because the input impedance at very high frequencies would be essentially that of the inductor and would be infinite at infinite frequency.

An example of a network that is not minimum reactance is shown in Fig. 3 [11]. Note that at a frequency of 1 rad/sec, the middle shunt branch resonates, grounding the series inductors. The equivalent network at this frequency is simply an $L-C$ parallel circuit that is resonant. Therefore, for the given element values the input impedance is infinite when $\omega = 1$ and therefore does not fulfill the conditions for being minimum reactance. (However, if the input capacitor were changed from $C = 1$ to $C \neq 1$, the network would be a minimum reactance network.) The network of Fig. 3 will not be encountered in the work described in this chapter. It is presented only to clarify the concept of minimum reactance networks.

An example of a minimum susceptance network is the singly terminated, low pass prototype filter shown in Fig. 1b. The input admittance of this network is finite for all frequencies, including infinity. Note here that the first element of the filter is a series inductor. In this case if the first element were a shunt capacitor, the network would not be minimum susceptance because the admittance would become infinite at infinite frequency.

For the work described in this chapter we may simplify somewhat the characterization of minimum reactance and minimum susceptance networks in the manner shown in Figs. 4a and 4b. The important points to note in Figs. 4a and b are that the minimum reactance network starts out with a
shunt susceptance, while the minimum susceptance network starts out with a series reactance. There are exceptions to the description given by Fig. 4, as was noted by the example in Fig. 3. However, for the purposes of the discussion in this chapter, no exceptions will be encountered, and the recognition of minimum reactance and minimum susceptance networks may be easily accomplished by comparison with Figs. 4a and 4b.

Fig. 4. (a) Illustration of a frequently occurring minimum-reactance network (see exception in text). (b) Illustration of frequently occurring minimum-susceptance network.

It is also necessary to point out that if two (or more) minimum susceptance networks are connected in parallel, the resulting network is also minimum susceptance. Similarly, if two (or more) minimum reactance networks are connected in series, the resulting network is also minimum reactance.

An important property of minimum reactance networks is that the real and imaginary parts of the input impedance are mathematically related. Thus, if either the real or imaginary component of the input impedance is known, the remaining component is unique and is known implicitly. The same holds true for the real and imaginary parts of the input admittance of minimum susceptance networks.

Bode [12] has given several mathematical formulas relating the real and imaginary parts of minimum reactance and minimum susceptance networks. One of these is particularly useful in explaining the concepts of the diplexer and multiplexer design theory. The formula (adapted to our terminology and nomenclature) is [13]

\[
X(\omega) = \frac{\omega}{\pi} \int_{0}^{\infty} \frac{R(\omega) - R(\omega)}{u^2 - \omega^2} \, du
\]  

(6)
where $Z(\omega)$ is the complex impedance (admittance) of the minimum reactance (minimum susceptance) network; $R(\omega)$ is the real part of the minimum reactance (minimum susceptance) network; and $X(\omega)$ is the imaginary part of the minimum reactance (minimum susceptance) network.

This formula shows how to compute the imaginary component of the input impedance (admittance) of a minimum reactance (minimum susceptance) network, if the real component is known.

A second mathematical relationship that is useful to us is given by [14]

$$\int_0^\infty [R(\omega) - R(\infty)]^2 \, d\omega = \int_0^\infty [X(\omega)]^2 \, d\omega$$

where $R(\infty)$ means the real part evaluated at infinite frequency.

A very important consequence of Eq. (7) is that if $R(\omega)$ is a constant for all frequencies, then the integrand (and therefore the integral also) of the left side of Eq. (7) is identically zero. Since the right side of Eq. (7) is the integral of a quantity squared (which is therefore always positive), it can be concluded that $X(\omega)$ is identically zero. It is emphasized that this result does not imply that the network contains only resistors. It states only that if the real part of the impedance measured at a given pair of terminals is constant at all frequencies, the imaginary part must be zero at all frequencies, even though the network contains reactive elements. Equation (7) also implies that if $R(\omega)$ is nearly constant as a function of $\omega$ (that is, it has only a small variation about a constant value), then $X(\omega)$ is very nearly zero, but not identically zero at all frequencies.

Equation (6) may be used to show that if the real part of the input impedance is approximately constant over a band of frequencies (referred to as the passband) and then falls to zero and remains zero outside the passband (referred to as the stop band), then the imaginary part has

1. on the average, a negative slope in the passband;
2. an extremum in the vicinity of the transition from the passband to stop band; and
3. on the average, a positive slope in the stop band.

This result is illustrated qualitatively in Fig. 5, in which are sketched (not to scale) typical input impedances (admittances) of minimum reactance (susceptance) networks of four, singly loaded filter types.

These relationships between real and imaginary components of the input impedance (admittance) of minimum reactance (susceptance) networks underlie the theory of the design of diplexers and multiplexers as described in this chapter. It is for this reason that we have described in some detail the definitions and some of the properties of minimum reactance and minimum susceptance networks.
Fig. 5. Illustrations (not to scale) of input impedance (admittance) of minimum-reactance (susceptance) networks of singly loaded filter types. (a) Low pass network characteristics; (b) bandpass network characteristics; (c) high pass network characteristics; (d) bandstop network characteristics. (The foreshortened doubly loaded filters discussed in Section III will have similar characteristics.)
III. THEORY OF COMPLEMENTARY AND PARTLY COMPLEMENTARY DIPLEXERS

A. THEORY OF COMPLEMENTARY DIPLEXERS

Three slightly different design approaches for diplexers will be described in order to cover three somewhat different situations. These ideas will later be carried over to the design of multiplexers. The first case discussed will be referred to as the complementary diplexer. This design is based on constructing low pass and high pass filters to have complementary input impedances and connecting the filters in series. Alternatively, in the dual formulation, one constructs low pass and high pass filters to have complementary input admittances and connects the filters in parallel.

The phrase "complementary input admittances" means that the sum of the input admittance of the low pass and high pass filter is real and constant for all frequencies [15]; the phrase "complementary input impedance" has the same definition, except the sum of the input impedances is real and constant for all frequencies. Guillemin has shown that only minimum reactance and minimum susceptance networks may be made complementary [15]. These ideas will be developed in more detail in the following paragraphs.

Figure 6a shows a parallel connected complementary diplexer. It consists of a minimum susceptance low pass filter (the first element is a series inductor) connected in parallel with a minimum susceptance high pass filter. The diplexer is excited by a source having internal conductance $G_g$. Since the networks are minimum susceptance, the input admittance of the low pass filter approximates the general form given in Fig. 5a, while the input admittance of the high pass filter approximates the general form given in Fig. 5c. The input admittance of the diplexer, denoted by $Y_T$, is given by

$$Y_T = Y_L + Y_H$$

because the filters are parallel connected. The admittance $Y_T$ may be estimated by superimposing Figs. 5a and 5c and performing the summation. This has been done in Fig. 7, wherein also it has been assumed that the admittance levels of the filters and their cutoff frequencies have been properly chosen to achieve the results given in the figure.

Notice in Fig. 7 that by adjusting the real parts of the input admittances of the low pass and high pass filters to cross over at approximately the value $G_g/2$, the real part of $Y_T$ is made nearly constant for all frequencies, and the imaginary part of $Y_T$ is virtually zero. Thus, the filters are approximately complementary. Since the input admittance of the diplexer approximately equals the source conductance, the diplexer is well matched to the generator, and the available power of the source is delivered to the diplexer almost without reflection.
It is very important to the understanding of the theory to recognize in Fig. 7 that the imaginary parts of the admittances of the low pass and high pass filters have not been arbitrarily constructed so that their sum is approximately zero. Rather, this result is a direct consequence of Bode's relationship given by Eq. (7), and results from the fact that the real part of the input admittance of the diplexer is approximately constant for all frequencies.

In the example just given, it was shown in general how to design an (approximately) complementary diplexer using singly terminated low pass and high pass filters. It can be shown that if maximally flat, singly terminated filters are used as prototypes, the low pass and high pass filters may be designed to be exactly complementary, in theory. However, requiring two filters to be exactly complementary is an unnecessarily severe requirement. As will be shown, by using specially modified doubly terminated Chebyshev prototype filters, faster rates of cutoff (better selectivity) can be obtained than with maximally

---

**Fig. 6.** (a) Parallel- and (b) series-connected diplexers using complementary filters.
flat filters, and a predictable Chebyshev transmission attenuation and VSWR response can be obtained.

Figure 8 shows a doubly terminated, low pass prototype filter (which may be assumed Chebyshev for the present discussion) and suggests a foreshortening procedure (to be described) that makes it suitable for use in diplexer applications. Notice that two input admittances have been designated:

\[ Y'_L, \text{ the input admittance of the doubly terminated filter, and } Y'_{L_2}, \text{ the input admittance of the same filter with the first reactive element removed.} \]

In the present context, a doubly terminated filter with the first reactive element removed will be referred to as a foreshortened, doubly terminated filter, or simply as a foreshortened filter. Notice that foreshortening the filter in Fig. 8a causes the input admittance \( Y'_{L_2} \) to be minimum susceptance. The imaginary part of \( Y'_{L_2} \) (and also \( Z'_{L_2} \) for Fig. 8b) will have a shape similar to the imaginary part in Fig. 5a.

The input admittance of the doubly terminated filter and also the foreshortened filter are sketched in Fig. 9. The real part of the input admittance is not affected by the foreshortening procedure because it is a shunt capacitor that is removed. Thus \( \text{Re } Y'_{L_2} = \text{Re } Y'_{L_1} \). The \( \text{Im } Y'_{L_1} \), however, which was
THEORY AND DESIGN OF DIPLEXERS AND MULTIPLEXERS

R = 9»

R = 9»

C = g,

C = $C'

R = 9»

R = 9»

Fig. 8. (a) Doubly terminated and foreshortened, doubly terminated low pass prototype filter; (b) its dual.

Fig. 9. Typical admittance characteristics of the doubly terminated and foreshortened filter of Fig. 8a ($Y_L$ is the admittance before foreshortening, and $Y'_L$ is the admittance after foreshortening).

typically small throughout the passband, is altered considerably. The form of the input admittance of the foreshortened filter is similar to that shown in Fig. 5a. This result is expected for the reason that the input admittance is minimum susceptance and the network is designed to be a low pass filter.
In designing a complementary diplexer, the low pass filter is based on a foreshortened, doubly terminated Chebyshev filter. The high pass filter is also designed as a foreshortened, doubly terminated Chebyshev filter. Consequently, its input admittance takes the general form shown in Fig. 5c.

Let us assume that the respective cutoff frequencies of the two filters have been chosen so that the real parts cross over at \( G_g/2 \), and that the real part of the admittances of the filters in their passbands have been adjusted to equal the source conductance. The filters are next connected in parallel, as in Fig. 6a. Here, in this example, as in the previous case, there results a nearly constant input admittance equal to the source conductance. By Bode’s relationship, the imaginary part is approximately zero. The diplexer is therefore approximately matched to the source for all frequencies.

As the theory has been described to this point, it might appear that using either the foreshortened, doubly terminated Chebyshev filter as a prototype, or the singly terminated Chebyshev filter as a prototype would give the same results. However, this is not the case. It only appears thus because we have been describing the theory of diplexers in general terms. A closer examination shows that by using the foreshortened, doubly terminated prototype filter, a predictable Chebyshev response can be obtained, whereas if one uses the singly terminated Chebyshev filter, the response will be much more difficult to estimate. (However, one should still use the singly terminated maximally flat prototype filter to obtain exactly complementary filters.)

The reasons for this conclusion rest on the manner in which the prototype filters themselves were designed. The singly terminated filter is designed to be driven by a zero-internal-conductance current generator (or a zero-internal-resistance voltage generator in the dual case). Therefore, if the diplexer is designed using the singly terminated Chebyshev prototype filter, its response when driven by a zero-internal-conductance current generator would be Chebyshev. But, when it is driven by a source having finite internal conductance, the response is altered by the current drain through the internal conductance and the Chebyshev characteristic is considerably altered.

On the other hand, the doubly terminated filter was designed to be driven by a source having finite conductance, so that this fact is accounted for in advance. This situation manifests itself, to a large degree, by altering the real part of the input admittance from that of the singly terminated case. Therefore, when the diplexer (consisting of foreshortened low pass and high pass filters) is driven by a source with internal conductance, the real part of the input admittance is that of the doubly terminated filter, while the imaginary part is approximately zero (owing to the approximately complementary nature of the filters); therefore, the Chebyshev transmission attenuation and VSWR responses are retained.
B. Theory of Partly Complementary Diplexers

An idealized diplexer would consist of lumped element, low pass and high pass filters. If the elements in these filters were truly lumped at all frequencies, the low pass filter would have a stopband of unlimited extent, and the high pass filter would have a passband of unlimited extent. However, in practice, a microwave low pass filter is often constructed from semilumped elements, and its stopband will deteriorate at frequencies at which the size of the circuit elements becomes appreciable as compared to a wavelength. Likewise, a microwave high pass filter is also often constructed from semilumped elements, and its passband will deteriorate at frequencies at which the size of the circuit elements becomes appreciable as compared to a wavelength. Microwave low pass filter structures are, in principle, not particularly difficult to design, but if a diplexer with a very accurately defined crossover is required, it may be difficult to design the low pass filter with the amount of accuracy required to get the specified crossover frequency. Microwave high pass filters with precisely defined characteristics are considerably more difficult to design than microwave low pass filters; for this reason, wide-band bandpass filters are very often used in place of semilumped-element high pass filters. Since any practical microwave high pass filter will have a limited passband anyway, a wide-band bandpass filter may serve the purpose equally well, and be much more convenient to design and construct.

A partly complementary diplexer may consist of a low pass filter for the low frequency channel and a wide-band bandpass filter for the high frequency channel. Such a parallel connected, partly complementary diplexer is shown in Fig. 10a and its dual, a series connected partly complementary diplexer, is shown in Fig. 10b. The susceptance-annulling network in Fig. 10a is required in the partly complementary diplexer theory for reasons that will be discussed in the succeeding paragraphs.

The theory of partly complementary diplexers calls for the low pass filter preferably to be designed from a foreshortened, doubly terminated Chebyshev, low pass, prototype filter. The reasons are the same as those given in the discussion for complementary diplexers. (Briefly, use of the foreshortened, doubly terminated, Chebyshev, prototype filters permits the realization of a predictable Chebyshev transmission and VSWR response.) For the same reasons, the design of the bandpass filter is also based on the foreshortened, doubly terminated, Chebyshev, prototype filter.

The input admittance of the bandpass filter, in Fig. 10a, prior to foreshortening and after, is given in Fig. 11. Analogous to the description in Fig. 8,

1 By partly complementary impedances, it is meant that two or more impedances are nearly complementary over a finite band of frequencies. By analogy, a partly complementary diplexer is one composed of filters having partly complementary impedances or admittances.
$Y_{B_1}$ is the admittance prior to foreshortening, and $Y_{B_2}$ is the admittance after foreshortening. In the example afforded by Fig. 10a, the bandpass filter is developed from a prototype filter having as its first element a shunt capacitor.

![Diagram of partly complementary diplexers]

When this element is removed, the real part of the input admittance is unaffected, and also the network is made minimum susceptance. Note that the foreshortened bandpass filter has an input admittance, such as that given in Fig. 5b.

When the foreshortened low pass and bandpass filters are connected in parallel, then the respective admittances add so that

$$Y_T = Y_{L_2} + Y_{B_2}$$

(9)
Re \(Y_T = \Re Y_{B_1} = \Re Y_{B_2}\)

**Fig. 11.** Typical input admittance characteristic for a Chebyshev bandpass filter \((Y_{B_1} \text{ is the admittance before foreshortening, and } Y_{B_2} \text{ is the admittance after foreshortening}).

\[
Y_T = Y_{L2} + Y_{B2}
\]

\(\Im Y_{B_1} \quad \Im Y_{B_2} \quad \Im Y_T\)

**Fig. 12.** (a) Input admittance of the parallel connected partly complementary diplexer before the susceptance-annulling network is added. (b) Susceptance of the susceptance-annulling network.

As in the complementary diplexer theory, the cutoff frequencies of the filters are adjusted so that the real parts of their admittances cross over at approximately one-half the generator conductance. After this is done, the input admittance \(Y_T\) will take the form shown in Fig. 12a.
Notice that the input admittance $Y_T$ is similar to that of a minimum susceptance low pass filter, and that (on the average) Im $Y_T$ is negative throughout the passband. It is because Im $Y_T$ is not zero in the passband that the susceptance-annulling network is required. The susceptance of the lossless susceptance-annulling network, shown in Fig. 10a, has a positive slope for all frequencies so that, in theory, its parameters can be adjusted to make its susceptance largely cancel Im $Y_T$ over the passband of the diplexer. This is suggested by the sketch in Fig. 12b, in which the susceptance of the susceptance-annulling network has been scaled so as largely to cancel Im $Y_T$.

Once the parameters of the susceptance-annulling network are adjusted to the proper values, Im $Y_T + jB_a$ is approximately zero throughout the passband of the diplexer, and the diplexer is matched to the generator.

The examples presented in explaining the theories of the complementary and the partly complementary diplexers used parallel connected filters. However, the same theory holds for the dual case wherein the filters are connected in series. For the series-connected case, the impedance, rather than the admittance, is dealt with. This has been indicated in Figs. 6b and 10b. For the partly complementary diplexer, a reactance-annulling network is used in series with the low pass and bandpass filters as indicated in Fig. 10b.

IV. DESIGN OF COMPLEMENTARY AND PARTLY COMPLEMENTARY DIPLEXERS

A. Example Design of a Parallel Connected Partly Complementary Diplexer

In this section, the problem of designing a partly complementary diplexer using a low pass filter and a bandpass filter will be treated. It is desired that the bandpass filter have a band-edge ratio of 3:1 or greater. In this type of application the bandpass filter is being used as a high pass filter, primarily because a bandpass filter is easier to design with precision at microwave frequencies. The possible use of an interdigital filter [17] for the high pass channel was considered, but it was found that if a band-edge ratio as great as 3:1 is desired, the dimensions for an interdigital filter structure become impractical (because of the close spacing between the resonator elements). For this reason it was decided to use a form of stub-resonator bandpass filter. In this type of filter, the wider the bandwidth the easier it becomes to realize the physical dimensions.

1. Proposed Bandpass Filter Structure

Figure 13 shows the proposed form of bandpass filter structure. The stubs are all one-quarter wavelength long at the midband frequency $\omega_0$. The shunt

---

1 This work is reported in greater detail in Ref. [16]. The design equations in Table I were derived using the methods of Refs. [16a, b].
Fig. 13. Bandpass filter structure proposed for use in a wide band partly complementary diplexer. (All of the stubs are $\lambda_0/4$ long, where $\lambda_0$ is the wavelength at the mid-passband frequency $\omega_0$.)

stubs in the interior portions of the filter are separated by connecting lines that are also one-quarter wavelength long at midband. Note that the filter has a short-circuited shunt stub at each end, which is followed by an open-circuited series stub and then by a short-circuited shunt stub. When the bandpass filter is foreshortened, the shunt stub $Y_1$ will be removed.

Table I presents design equations for filters of the form in Fig. 13. Using these equations, the bandpass filter is designed from a doubly terminated low pass filter.

**Table I**

**Design Equations for the Bandpass Filter in Fig. 13**

\[
\begin{align*}
\theta_1 &= \frac{\pi \omega_1}{2 \omega_0} = \frac{\pi}{2} \left(1 - \frac{w}{2}\right) \\
Y_{k,k+1} &= \left(\frac{g_k g_{k+1}}{\tan \theta_1}\right)^{1/2} \quad C_a = g_0 g_3 Y_A \\
M_{k,k+1} &= \left[ (Y_{k,k+1})^2 + \left(\frac{\omega_1' C_a \tan \theta_1}{2}\right)^2 \right]^{1/2} \\
Y_1 &= \omega_1' g_0 g_1 Y_A \tan \theta_1 \\
Z_2 &= \frac{\omega_1' g_2 Z_A}{g_0} \tan \theta_1 \\
Y_3 &= \frac{\omega_1' C_a \tan \theta_1}{2} + M_{34} - Y_{34} \\
Y_k|_{k=4,10} &= M_{k-1,k} + M_{k,k+1} - Y_{k-1,k} - Y_{k,k+1} \\
Y_{n-2} &= M_{n-3,n-2} - Y_{n-3,n-2} + \left(\omega_1' g_{n-2} g_{n+1} Y_A - \frac{C_a}{2}\right) \tan \theta_1 \\
Z_{n-1} &= \frac{\omega_1' g_{n-1} Z_A \tan \theta_1}{g_0} \\
Y_n &= \omega_1' g_n g_0 Y_A \tan \theta_1
\end{align*}
\]
prototype filter, from which the bandpass filter derives its passband characteristics. Figure 2 defines the parameters for the low pass prototype filters to be used. There is one additional parameter, however, that was not defined in the figure, which is the passband-edge frequency \( \omega_1' \). In the case of Chebyshev filters, \( \omega_1' \) is usually defined as the equal-ripple band edge.

The low pass prototype filter band-edge frequency, \( \omega_1' \), maps into two band-edge frequencies, \( \omega_1 \) and \( \omega_2 \), for the bandpass filter. Zero frequency for the low pass prototype filter maps to the midband frequency, \( \omega_0 \), of the bandpass filter. The response of the bandpass filter can be estimated from that of the low pass prototype filter by use of the mapping equation:

\[
\frac{\omega'}{\omega_1'} = \frac{2}{w} \left( \frac{\omega - \omega_0}{\omega_0} \right)
\]

where

\[
w = \frac{\omega_2 - \omega_1}{\omega_0}
\]

and

\[
\omega_0 = \frac{\omega_2 + \omega_1}{2}
\]

In these equations \( \omega' \) is the frequency variable for the low pass prototype, while \( \omega \) is the corresponding frequency variable for the bandpass filter.

In order to check out the design equations in Table I, a design was worked out using a 0.10-db ripple Chebyshev low pass prototype filter having \( n = 15 \) reactive elements. The prototype element values are summarized in Table II. The design was worked out in normalized form so that the terminations were \( Z_A = Z_B = 1 \ \Omega \), and the band-edge parameter \( \omega_1/\omega_0 = 0.450 \). This gives a band-edge ratio \( \omega_2/\omega_1 = 3.44 \). The normalized element values for the filter are summarized in Table III, while Fig. 14 shows the computed attenuation characteristic. Note that the passband attenuation ripple is approximately 0.10 \( \text{db} \) (as specified), while the band-edge ratio at the 0.10-db points is approximately 3.40, as compared with desired 3.45 design value.

Table II

<table>
<thead>
<tr>
<th>ELEMENT VALUES FOR A LOW PASS PROTOTYPE FILTERa</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g_0 = g_{16} = 1 )</td>
</tr>
<tr>
<td>( g_1 = g_{15} = 1.2101 )</td>
</tr>
<tr>
<td>( g_2 = g_{14} = 1.4612 )</td>
</tr>
<tr>
<td>( g_3 = g_{13} = 2.1660 )</td>
</tr>
<tr>
<td>( g_4 = g_{12} = 1.6461 )</td>
</tr>
</tbody>
</table>

a Having \( n = 15 \) reactive elements and 0.10-db Chebyshev ripple with \( \omega_1' = 1 \).
Table III

NORMALIZED ELEMENT VALUES FOR A TRIAL BANDPASS FILTER DESIGN OF THE FORM SHOWN IN FIG 13a

| \( Y_1 = Y_{15} \) = | 1.0335 | \( Y_{3,4} = Y_{12,13} \) = | 1.1471 |
| \( Z_2 = Z_{14} \) = | 1.2480 | \( Y_{4,5} = Y_{11,12} \) = | 1.1230 |
| \( Y_3 = Y_{13} \) = | 1.2514 | \( Y_{5,6} = Y_{10,11} \) = | 1.1124 |
| \( Y_4 = Y_{12} \) = | 0.6583 | \( Y_{6,7} = Y_{9,10} \) = | 1.1074 |
| \( Y_5 = Y_{11} \) = | 0.6619 | \( Y_{7,8} = Y_{8,9} \) = | 1.1053 |
| \( Y_6 = Y_{10} \) = | 0.6698 | \( Z_A = Z_B \) = | 1.0000 |
| \( Y_7 = Y_9 \) = | 0.6714 |
| \( Y_8 \) = | 0.6719 |

a This design is for 0.10-db ripple with \( \omega_1/\omega_0 = 0.450 \).

Fig. 14. Computed response for the trial design in Table III.

2. Low Pass Filter and Interconnection

The low pass filter structure to be used is of the form in Fig. 8a. Notice that this structure begins with the shunt capacitor on the left. When the filter is foreshortened and the diplexer connection made, this capacitor will be removed.
The low pass filter may be designed directly from a low pass prototype filter by simply scaling the element values using the formulas

\[ L_k = \frac{L_k'}{\omega_1 Z_A} \]

(13)

\[ C_j = \frac{C_j'}{\omega_1 Z_A} \]

(14)

where \( \omega_1' \) is the band-edge frequency for the low pass prototype, \( \omega_1 \) is the corresponding band-edge frequency for the scaled filter, \( Z_A' \) is the impedance level of one of the terminations of the prototype filter, \( Z_A \) is the corresponding impedance level of the scaled filter, \( L_k' \) and \( C_j' \) are for the prototype, and \( L_k \) and \( C_j \) are for the scaled filter.

When the diplexer connection is made, both the low pass and bandpass filters are first foreshortened by removing their first element. Thus the shunt capacitor \( C_1 \) is removed from the low pass filter and the shunt stub \( Y_1 \) is removed from the bandpass filter. The resulting connection is as shown in Fig. 15, where it will be noted that the shunt susceptance-annulling network has been added at the junction. (The necessity for the susceptance-annulling network was explained in Section III,B.)

In order to calculate the numerical values of the parameters in the susceptance-annulling network, it is helpful to look at the diplexer problem from another point of view. Before the doubly terminated low pass filter is foreshortened, the susceptance in its passband is very small (see Fig. 9). After foreshortening (by removing the first reactive element) a nonnegligible susceptance component appears in the passband of the low pass filter. When the foreshortened bandpass filter and the susceptance-annulling network are connected in parallel with the low pass filter, however, the susceptance in the passband of the low pass filter is again reduced to negligible values. We see, therefore, that we are
effectively replacing the first reactive element of the low pass filter by the stopband susceptance of the bandpass filter plus the susceptance of the annuling network. A similar argument holds for the bandpass filter, so that for the bandpass filter, we are replacing the susceptance of the shunt stub \( Y_1 \) in Fig. 13 by the stopband susceptance of the low pass filter plus the susceptance of the annuling network. In this manner each filter sees a susceptance at its input that is very much the same as the susceptance of the first element that was removed.

The above point of view gives us a criterion by which the susceptance-annuling network for the diplexer can be designed. Within the passband of the bandpass filter we wish the equation

\[
\frac{1}{\omega L_a - 1/\omega C_a} + \text{Im} \ Y_{L_2} = -Y_1 \cot \left( \frac{\pi \omega}{2\omega_0} \right)
\]  

(15)

to be satisfied. This equation says that within the passband of the bandpass filter we wish to have the susceptance of the annuling network plus the imaginary part of the input admittance of the low pass filter to be equal to the susceptance of the stub \( Y_1 \) in Fig. 13. This equivalence cannot be made to hold exactly at all frequencies, but we can choose \( L_a \) and \( C_a \) so as to force this equivalence to hold exactly at two frequencies. We shall choose the frequencies to be at \( \omega_0 \), the mid-passband frequency of the bandpass filter, and \( \omega_2 \), the upper band-edge frequency of the bandpass filter.\(^1\) The computation of \( \text{Im} \ Y_{L_2} \) exactly could be very tedious, but in the stopband of the low pass filter we may compute the input admittance of the low pass filter with reasonably good accuracy if we use only the first two elements of the filter. Thus we use the approximation

\[
\text{Im} \ Y_{L_2} \approx \frac{1}{\omega L_2 - 1/\omega C_3}
\]  

(16)

Making use of this approximation, and writing Eq. (15) with \( \omega = \omega_0 \), and again with \( \omega = \omega_2 \), we obtain two simultaneous equations from which \( L_a \) and \( C_a \) can be obtained. The computations are facilitated by using the equations

\[
A_0 = \frac{\omega_0 C_3}{1 - \omega_0^2 L_2 C_3}
\]  

(17)

\[
A_2 = \frac{\omega_2 C_3}{1 - \omega_2^2 L_2 C_3} + Y_1 \cot \left( \frac{\pi \omega_2}{2\omega_0} \right)
\]  

(18)

\[
\omega_a = \left[ \frac{A_0 \omega_0 - A_2 \omega_2}{(A_0/\omega_0) - (A_2/\omega_2)} \right]^{1/2}
\]  

(19)

\(^1\) As can be seen from Fig. 12, the upper band-edge frequency is particularly important.
where $\omega_a$ is the radian frequency at which the susceptance-annulling network is series resonant. Then $L_a$ and $C_a$ can be obtained from

$$L_a = \frac{\omega_0}{A_0(\omega_0^2 - \omega_a^2)}$$  \hspace{1cm} (20)$$

$$C_a = \frac{-A_0}{\omega_0} \left[ 1 - \left(\frac{\omega_0}{\omega_a}\right)^2 \right]$$  \hspace{1cm} (21)$$

In most microwave structures it would not be practical to realize the susceptance-annulling network in lumped-element form, but it is quite easily realized by use of an equivalent circuit consisting of an open-circuited stub one-quarter wavelength long at the frequency $\omega_a$. In order to obtain the desired susceptance-annulling effect the characteristic impedance of the stubs could be approximately

$$Z_0 = 4\omega_a L_a/\pi$$  \hspace{1cm} (22)$$

In designing a diplexer it is usually desirable to have the low pass and bandpass filters cross over at their 3-db points. Thus, their band-edge frequencies should be adjusted so as to give this result. The locations of the 3-db points relative to the equal-ripple band-edge frequencies of Chebyshev filters can be determined from charts given in Ref. [7], Sec. 4.03. If Chebyshev filters are used it will usually be desirable for the low pass filter to have an odd number of elements. This is because Chebyshev low pass filters with an even number of elements require unequal terminations. It will usually be desirable also for the rates of cutoff of the low pass filter and the bandpass filter to be approximately the same. The required number of elements in the low pass filter and the bandpass filter can be easily estimated by use of the charts of low pass prototype filter cutoff characteristics given in Ref. [7], Sec. 4.03, along with the low pass to bandpass mapping equations, Eqs. (10)–(12), of this section.

Figure 16 is a sketch of a possible physical configuration for the proposed
wide-band diplexers. The low pass filter is of the common coaxial split-block form [18, 19]. This filter structure consists of sections of high impedance wire that simulate the series inductances, and metal disks with dielectric rims that simulate the shunt capacitors. The bandpass filter is made in strip line form and is supported by the short-circuited stubs. The series stubs at each end of the filter are realized by use of open-circuited stubs that protrude into holes within the input and output transmission lines. The susceptance-annulling network consists of a short, open-circuited stub, mounted on the common input line, and the shunt stub at the output end of the bandpass filter is also connected directly to the output line of the bandpass filter.

3. Computed Diplexer Performance

In order to check out this design approach the bandpass filter design in Table III was paralleled with a low pass filter. It was decided to use a 15-element low pass filter with 0.10-db Chebyshev ripple, with its band edge adjusted so that the 3-db point falls at the normalized frequency $\omega = 1$. The normalized design of Table II has the equal-ripple band edge at $\omega_1' = 1$, while its 3-db point occurs at $\omega' = 1.0145$. Thus the desired normalized design with its 3-db point at $\omega = 1$ was obtained by dividing element values $g_1$ to $g_{15}$ in Table II by 1.0145. In order to set the 3-db point of the bandpass filter at approximately $\omega = 1$, the midband frequency $\omega_0$ of the bandpass filter was set equal to 2.2381, which gave the value for the upper band-edge frequency $\omega_2$ of 3.4691. The scaled

![Fig. 17. Computed performance in low pass channel of diplexer.](image-url)
values $L_2$ and $C_3$ for the low pass filter were 1.44031 and 2.1350, respectively. Using these data in Eqs. (17) to (22), the resonant frequency of the susceptance-annulling network was found to be $\omega_a = 3.984$, and $L_a = 0.641$, while $C_a = 0.0982$, and the characteristic impedance for an equivalent annulling network composed of an open-circuited stub is $Z_0 = 3.26$ (when $Z_A = 1$). The low pass and bandpass filter structures were foreshortened and assumed to be connected as shown in Fig. 15.

Figure 17 shows the attenuation characteristic computed for the low pass channel. Note that the desired 0.10-db Chebyshev passband ripple is maintained quite well. Figure 18 shows the bandpass channel. Note that in this case the size of the Chebyshev ripples is actually reduced somewhat over part of the band from what it was originally, as shown in the response in Fig. 14. A Chebyshev ripple of 0.1 db corresponds to a maximum passband VSWR of approximately 1.36. The VSWR at the common junction of the diplexer was computed, and it was found to be 1.36 or less across the entire range from

![Fig. 18. Computed performance in bandpass channel of diplexer.](image-url)
\( \omega = 0 \) to \( \omega = 3.5 \), except right at crossover where the VSWR jumped up to around 1.8. This rise in VSWR at crossover could have easily been corrected by a small adjustment of center frequency of the bandpass filter.

### B. Example Design and Measured Responses of Two Series Connected Partly Complementary Diplexers

#### 1. Introduction

Figure 10b shows a series-connected partly complementary diplexer using foreshortened doubly terminated filters and a series reactance-annulling network. In the analogous example about to be discussed, the diplexer uses a low pass filter designed from a Chebyshev prototype with 0.10-db passband ripple and \( n = 15 \) reactive elements. At the time this diplexer design was worked out, however, tabulated element values were not available for Chebyshev filters with this many elements, so that it was decided to use a corresponding design with \( n = 9 \) elements and then repeat the two middle elements of the design a number of times in order to increase the number of elements to 15. This gives a design differing slightly from a true \( n = 15 \) Chebyshev design, but the difference is small. Table IV gives element values for the resulting low pass filter with the impedance level normalized so that \( R_0 = R_L = 1 \), and with the frequency scale normalized so that the equal-ripple band-edge frequency is \( \omega_1' = 1 \).

#### Table IV

<table>
<thead>
<tr>
<th>Normalizing Element Values for the Low Pass Filter*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_0 = R_L = Z_A = 1 )</td>
</tr>
<tr>
<td>( L_1 = L_{15} \quad = 1.1956 )</td>
</tr>
<tr>
<td>( C_2 = C_{14} \quad = 1.4425 )</td>
</tr>
<tr>
<td>( L_3 = L_{13} \quad = 2.1345 )</td>
</tr>
<tr>
<td>( C_4 = C_{12} \quad = 1.6167 )</td>
</tr>
</tbody>
</table>

<p>| |</p>
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( L_5 = L_{11} = 2.2053 )</td>
</tr>
<tr>
<td>( C_6 = C_{10} = 1.6167 )</td>
</tr>
<tr>
<td>( L_7 = L_9 = 2.2053 )</td>
</tr>
<tr>
<td>( C_8 = 1.6167 )</td>
</tr>
</tbody>
</table>

* These element values were obtained from an \( n = 9 \) reactive element, Chebyshev low pass filter with 0.1-db ripple and with the equal-ripple band edge at \( \omega_1' = 1 \). The \( n = 9 \) filter was augmented to become an \( n = 15 \) reactive element design as described in the text.

It was also decided to use a bandpass filter of the type shown in Fig. 19 for the high pass channel. Using the design equations in Table V, with \( d = 1 \),

---

1 This work is reported in greater detail in Ref. [18], Chapters 20–21.

2 The validity of this procedure can be seen more easily from the image impedance design viewpoint. Working from that viewpoint it is common practice to use certain matching end sections with any number of identical interior sections.

3 The design equations in Table V were obtained by the methods of Refs. [16a, 16b].
Design Equations for Filters with $\lambda_0/4$ Stubs and $\lambda_0/4$ Connecting Lines as Shown in Fig. 19**

Compute:

$$
\theta_1 = \frac{\pi \omega_1}{2 \omega_0} = \frac{\pi}{2} \left(1 - \frac{w}{2}\right)
$$

$$
\frac{J_{23}}{Y_A} = \frac{1}{g_0} \left(\frac{C_n}{g_3}\right)^{1/2} \frac{J_{k,k+1}^{+} Y_A}{(k-3 \text{ to } n-2)} = \frac{C_n}{g_0 (g_k g_{k+1})^{1/2}}
$$

$$
J_{n-1,n}^{+} Y_A = \frac{1}{g_0} \left(\frac{C_n g_{n+1} g_0}{g_{n-1}}\right)^{1/2}
$$

where $C_n = 2d g_2$

and $d$ is a dimensionless constant (typically chosen to be 1.0) which can be chosen so as to give a convenient admittance level in the interior of the filter.

$$
N_{k,k+1}^{+} Y_A = \left(\left(\frac{J_{k,k+1}^{+} Y_A}{Y_A}\right)^2 + \left(\frac{\omega_1' C_n \tan \theta_1}{2 g_0}\right)^2\right)^{1/2}
$$

The characteristic impedance of the series stub is

$$
Z_1 = \frac{\omega_1' g_1 g_0 \tan \theta_1}{Y_A}
$$

The characteristic admittances of the shunt stubs are

$$
Y_2 = \frac{Y_A \omega_1' (1-d) g_2 \tan \theta_1 + Y_A \left(N_{23} - \frac{J_{23}}{Y_A}\right)}{g_0}
$$

$$
Y_k^{+} Y_A \bigg|_{k-3 \text{ to } n-1} = Y_A \left(N_{k-1,k} + N_{k,k+1} - \frac{J_{k-1,k}}{Y_A} - \frac{J_{k,k+1}}{Y_A}\right)
$$

$$
Y_n = Y_A \omega_1' \left(g_n g_{n+1} - \frac{d g_1}{g_0}\right) \tan \theta_1 + Y_A \left(N_{n-1,n} - \frac{J_{n-1,n}}{Y_A}\right)
$$

The characteristic admittances of the connecting lines are

$$
Y_{k,k+1}^{+} Y_A \bigg|_{k+1 \text{ to } n-1} = Y_A \left(\frac{J_{k,k+1}^{+} Y_A}{Y_A}\right)
$$

All stubs and connecting lines are $\lambda_0/4$ long, where $\lambda_0$ is the wavelength in the medium of propagation at the midband frequency $\omega_0$.  

* Use mapping Eqs. (10)-(12) to select a low pass prototype with the required value of $n$. The filter will have one series stub and $n-1$ shunt stubs for an $n$-reactive-element prototype.

$\omega_1/\omega_0 = 0.630$ (which calls for $f_2/f_1 = 2.175$), along with an $n = 10$ reactive-element low pass prototype with 0.10-db Chebyshev ripple, the transmission line impedances can be calculated. As was done in the case of Table IV, the
n = 10 design was augmented by additional sections in order to raise the value of n. In this case, five additional sections were added to make n = 15 (since in the final design one stub will be removed).

![Diagram of Bandpass Filter with Series Stub](image)

**Fig. 19.** Bandpass filter with series stub at one end.

The design of the series reactance-annulling network follows from the previous discussion of the shunt susceptance-annulling network. Within the passband of the bandpass filter we wish to have

\[ \frac{1}{\omega C_a - 1/\omega L_a} + \text{Im} Z_{L_2} = -Z_1 \cot \left( \frac{\pi \omega}{2 \omega_0} \right) \]  

(23)

To a good approximation in the stopband of the low pass filter,

\[ \text{Im} Z_{L_2} \approx \frac{1}{\omega C_2 - 1/\omega L_3} \]  

(24)

Next define

\[ A_0 = \frac{\omega_0 L_3}{1 - \omega_0^2 C_2 L_3} \]  

(25)

\[ A_2 = \frac{\omega_2 L_3}{1 - \omega_2^2 C_2 L_3} + Z_1 \cot \left( \frac{\pi \omega_2}{2 \omega_0} \right) \]  

(26)

\[ \omega_a = \left[ \frac{A_0 \omega_0 - A_2 \omega_2}{(A_0/\omega_0) - (A_2/\omega_2)} \right]^{1/2} \]  

(27)

where \( \omega_a \) is the radian frequency at which the reactance-annulling network is parallel resonant. Then \( L_a \) and \( C_a \) are given by

\[ C_a = \frac{\omega_0}{A(\omega_0^2 - \omega_a^2)} \]  

(28)

\[ L_a = -\frac{A_0}{\omega_0} \left[ 1 - \left( \frac{\omega_0}{\omega_a} \right)^2 \right] \]  

(29)

In most microwave structures it is not practical to realize the reactance-annulling network in lumped-element form, but it can be realized by using an equivalent circuit of a short-circuited series stub one-quarter wavelength long.
Fig. 20. Possible means for realizing reactance-annulling network in coaxial line for series-connected diplexer.

at the frequency \( \omega_a \), such as is shown in Fig. 20. The characteristic admittance of the stub should be approximately

\[
Y_0 = 4\omega_a C_0 / \pi
\]

(30)

2. Construction Details and Measured Response of a Series Connected Partly Complementary Diplexer for the 1- to 4-Gc Range

The series connected diplexer previously described was constructed to operate in the 1- to 4-Gc range with a 2-Gc cutoff frequency for the low

Fig. 21. Photograph of 1- to 4-Gc band diplexer with cover plates removed. (From [3], Chapter 16.)
frequency channel. A photograph of the completed diplexer is shown in Fig. 21. The low frequency channel is composed of a low pass coaxial filter, whose constructional details are shown in Fig. 22. The high frequency channel of the diplexer is composed of a bandpass filter as shown in Fig. 23.\(^1\) Details of the diplexer junction are shown in Figs. 24 and 25. The reactance-annulling

\[\text{REXOLITE 1422 RINGS, EACH HAVING A WIDTH EQUAL TO THE THICKNESS OF THE DISK IT SURROUNDS} \]

![Diagram of diplexer junction](image)

**Fig. 22.** Dimensions of low pass filter.

**Fig. 23.** Bandpass filter structure for 1- to 4-Gc diplexer. (Most of the shunt stubs have been realized as double stubs in order to give more practical dimensions.) For details of stub No. 2 and junction, see Figs. 24 and 25. (Stub No. 1 is removed.)

\(^1\) Since most of the stub admittances called for in this design were quite large, all stubs except the stubs at the ends were fabricated as double stubs in parallel. This permitted the use of more reasonable dimensions.
network in this example consists of a radial-line disk, also shown in Figs. 24 and 25.

Figure 26 shows the measured response for transmission through the lower-frequency channel while Fig. 27 shows the corresponding response for transmission through the higher-frequency channel. (In both cases 50-Ω terminating or source impedances were maintained at all three ports.) In order to detect the details of the passband responses, extensive VSWR measurements were
made in the passbands and the reflection loss was computed from the VSWR values. Insertion loss measurements were made at various spot check points in the passbands so that the total loss (that is, reflection loss plus dissipation loss) would also be evident. These measurements were made using tuned pads,

![Diagram of diplexer junction](image)

**Fig. 25.** Further details of diplexer junction.

the source and load circuits being tuned to a very low standing wave ratio with a slotted line prior to making each insertion loss measurement. The cutoff characteristics of the filters were also determined by insertion loss measurements. Figure 28 shows the cutoff characteristics of the two channels superimposed and also the VSWR measured at the junction port of the diplexer.
**Fig. 26.** Measured transmission characteristic of low pass channel of diplexer of Fig. 21.

**Fig. 27.** Measured transmission characteristics of bandpass channel of diplexer of Fig. 21.
3. Construction Details and Measured Response of a Series Connected Partly Complementary Diplexer for the 4- to 12.4-Gc Range

A drawing of this diplexer is shown in Fig. 29 and a photograph is shown in Fig. 30. In most respects it is a scaled-down version of the diplexer for the 1- to 4-Gc range. However, since the filter structure becomes quite small when scaled down for use in the 4- to 12.4-Gc frequency range, it was necessary to design a special transition in order to connect type N connectors to the miniaturized filter structures with a minimum of reflection. The type N connectors themselves were modified to give low reflection through X band, and to hold the center pin rigidly within the connector. Details of the transitions and modified type N connectors are given in Ref. [20].

When the 4- to 12.4-Gc diplexer was first tested, its response appeared to be quite close to what was expected. However, at around 11 Gc there were some points where there was anomalous behavior. At these points the VSWR went up sharply to around 3.5, while the insertion loss went as high as 9 db. After some experimentation it was determined that this phenomenon was due to a resonance of a ground plane mode within the bandpass filter structure. This mode was apparently excited by some dissymmetry of the center conductor structure with respect to the ground planes. It was found that this behavior could be largely eliminated by insertion of small pieces of polyiron within the
structure. This, however, also tended to introduce extra loss. Finally, some 0-80 screws were inserted between the stubs to prevent propagation of a ground plane mode. This stopped the anomalous behavior, but probably introduced some mistuning. After some adjustment of the lengths of the stubs in the bandpass filter, the responses shown in Figs. 31-33 were obtained. In all probability the VSWR at crossover (see Fig. 33) could be reduced to about that of the adjacent VSWR peaks if further adjustments were made with respect to the lengths of the stubs in the bandpass filter. As was done for the measurements on the lower-frequency diplexer, the source and load impedances were tuned to nearly unity VSWR for each insertion loss measurement, except for some measurements where high attenuation was involved.

C. EXAMPLE DESIGN OF A PARALLEL CONNECTED COMPLEMENTARY DIPLEXER USING A STUB-TYPE BANDSTOP FILTER AND AN INTERDIGITAL BANDPASS FILTER

1. Introduction

One form of bandpass filter that is particularly attractive in the high pass channel of diplexers is the interdigital bandpass filter [21]. This type of filter can be designed with good accuracy, gives a compact structure, and is quite economical to build (especially if the resonators are constructed from round rods [22]). Interdigital filters are suitable for bandwidths up to somewhat over an octave. When the bandwidths become sufficiently large (say, 3 to 1 or greater) the rod spacings become so close as to become impractical. For such cases, stub forms of bandpass filters are preferable. Figure 34 shows a diplexer with an interdigital bandpass filter consisting of an array of round rods between parallel ground planes.

Though a low pass filter of moderate accuracy is quite easy to design, a microwave low pass filter with a very rigidly prescribed response is very tedious to design. For this reason in many cases it is preferable to replace the low pass filter by a bandstop filter having a wide stopband. Such a filter can be designed with good accuracy, and can be tuned so as to place the cutoff frequency exactly where it is desired. Figure 34 shows a diplexer using such a bandstop filter consisting of an array of stubs open-circuited at one end and connected to the main transmission line at the other end [23, 24]. Each stub is a quarter-wavelength long at the center of the stopband; the stubs are spaced a quarter-wavelength apart at the stopband center frequency. Using this type of bandstop filter structure, the bandstop filter and the bandpass interdigital filter can be designed so that their characteristics will complement each other. Thus, if the frequency scale is normalized so that \( \omega_0 = 1 \) is the stopband center

---

\[1\] Wenzel [20a] has discussed complementary diplexers or multiplexers of different physical form than the example discussed herein.
Fig. 29. Drawing of the 4- to 12.4-Gc band diplexer.

Fig. 30. Photograph of 4- to 12.4-band diplexer with cover plate removed.
Fig. 31. Measured transmission characteristics of low pass channel of diplexer of Fig. 30.

Fig. 32. Measured transmission characteristics of bandpass channel of diplexer of Fig. 30.
Fig. 33. VSWR at junction port of diplexer of Fig. 30 and cutoff characteristics of the low pass and bandpass channel superimposed.

Fig. 34. Approximate form of a complementary diplexer using a stub-type bandstop filter and an interdigital bandpass filter (cover plate removed).
for the bandstop filter, the bandstop filter in Fig. 34 has a passband from 
\( \omega = 0 \) to \( \omega_1 = 0.625 \) and then its attenuation rises abruptly. When the bandstop 
filter cuts off at 0.625, the bandpass interdigital filter begins to transmit and it 
has a passband from \( \omega_1 = 0.625 \) to \( \omega_2 = 1.375 \). Though in most cases the higher 
order passbands are not of any practical interest, this type of structure can be 
designed so that the width of the stopband of the bandstop filter corresponds 
to the width of the passband of the bandpass filter, so that at the upper cutoff 
frequency of the bandpass filter (that is, \( \omega_2 = 1.375 \)) the bandstop filter will 
begin to transmit once again. Though it is not essential that the filter character-
istics be designed so as to complement each other in this way, doing so 
permits the obtaining of good performance without the addition of a sus-
ceptance-annulling network.\(^1\) When the two filters are designed to complement 
each other, the electrical length of the stubs of the bandstop filter should be 
the same as the electrical length of the resonator lines of the bandpass inter-
digital filter.

This discussion is addressed to the problem of designing diplexers of the 
form in Fig. 34 consisting of a stub-type bandstop filter and an interdigital 
bandpass filter, so designed that their attenuation characteristics complement 
each other. The design procedure is based on use of lumped-element, low pass 
prototype filters in the design of the bandstop and bandpass filters.

2. **Design of the Bandstop Filter**

Figure 35 shows a more detailed view of the type of bandstop filter to be 
considered herein. As previously mentioned, when this filter is connected to 
the interdigital bandpass filter, the stub \( Y_1 \) is removed. Assuming that the 

![Diagram](image)

**Fig. 35.** Form of bandstop filter for use in diplexers. (The stub \( Y_1 \) is removed when the 
diplexer connection is made. All stubs and the connecting lines are \( \lambda_0/4 \) long at the center 
of the stop band.)

same impedance level is desired at both ends of this filter, and if a Chebyshev 
response is desired, only an odd number of stubs should be used in the filter 
design (that is, \( m \) should be odd). This is because, for equal terminations at 
both ends of the filter, a perfect match must be obtained at dc, and Chebyshev 
filters with an odd number of elements will have theoretically zero attenuation

\(^1\) Thus, if the filters are designed this way, the diplexer would correspond to the com-
plementary-diplexer type, which does not require a susceptibility-annulling network.
at dc, while Chebyshev filters with an even number of elements will have a Chebyshev attenuation-ripple maximum at dc.

The filter in Fig. 35 is the general type discussed in Refs. [23, 24]; however, for this application the design procedures described in Refs. [23, 24] must be altered slightly. Techniques for the design of filters of the form in Fig. 35 are presented at the end of this section (that is, at the end of Section IV,C) along with tables of normalized designs. These tables give normalized values $h_k$ for the stub admittances, and the actual stub admittances can be computed by use of

$$Y_k = h_k Y_A$$  \hspace{1cm} (31)$$

where $Y_A$ is the admittance of the terminations. The tables also include normalized values $h_{k,k+1}$ for the connecting line admittances, from which the actual line admittances $Y_{k,k+1}$ can be computed by use of

$$Y_{k,k+1} = h_{k,k+1} Y_A$$  \hspace{1cm} (32)$$

It will be noticed that even though all of the bandstop filter designs tabulated at the end of this section were computed from symmetrical, lumped-element, low pass filter designs, the various transformations used resulted in bandstop filters that are quite unsymmetrical. As indicated in Fig. 35, the stubs near the number one end of the filter are of relatively large characteristic admittance, while those at the other end of the filter are of quite small characteristic admittance.

3. Design of the Bandpass Filter

Figure 36 shows a modified form of interdigital filter having special properties suitable for diplexers. This filter differs from the interdigital filters described in Ref. [21] in that in Fig. 36 the first resonator is not an interdigital element, but a short-circuited shunt stub. When the diplexer connection is made, the shunt stub is removed and is replaced by the bandstop filter. If the interdigital filter is designed to have an attenuation characteristic complementary to that of the bandstop filter, the length of the resonator elements in the interdigital filter is the same as the length of the stubs in the bandstop filter.

Table VI summarizes the design equations required for the design of interdigital filters of the form in Fig. 36. In most respects the design procedure is the same as is discussed in Ref. [21], so herein we will only treat the points at which the present design procedure differs from that previously presented. Of course, one way in which the design equations and Table VI differ from those in Table V of Ref. [21] is that Table VI provides for the presence of the stub $Y_{p1}$. Also in Table VI, the parameter $k$ of Ref. [21] has been replaced by $1/N_A$. The significance of the parameter $N_A$ can be better understood by considering the equivalent circuit in Fig. 37. This circuit is basic to the derivation of the design equations for interdigital filters of the form in Fig. 36, and
Design Equations for Interdigital Filters of the Form in Fig. 36

\[ \theta_1 = \frac{\pi \omega_1}{2 \omega_0} = \frac{\pi}{2} \left( 1 - \frac{w}{2} \right) \]

\[ u = g_0 g_3 \]

\[ D_{k,k+1} = \begin{cases} \frac{u}{(g_k g_{k+1})^{1/2}} & \text{if } k = 3 \text{ to } n-2 \\ \left[ \left( D_{k,k+1} \right)^2 + \left( \frac{\omega_1' u \tan \theta_1}{2} \right)^2 \right]^{1/2} & \text{if } k = n-1 \end{cases} \]

\[ M_{k,k+1} = \begin{cases} \frac{\omega_1' u \tan \theta_1}{2} + M_{34} - D_{34} & \text{if } k = 3 \text{ to } n-2 \\ \omega_1' g_k g_{k+1} \tan \theta_1 & \text{if } k = n-1 \end{cases} \]

\[ \frac{Y_{p1}}{Y_A} = \omega_1' g_0 g_1 \tan \theta_1 \]

\[ Z_2 Y_A = \frac{\omega_1' g_2 \tan \theta_1}{g_0} \]

\[ S_3 = \frac{\omega_1' u \tan \theta_1}{2} + M_{34} - D_{34} \]

\[ S_k = \frac{\omega_1' u \tan \theta_1}{2} + M_{n-2,n-1} - D_{n-2,n-1} \]

\[ N_B^2 = \frac{g_{n-1} Y_B N_A^2}{g_{n+1} Y_A u} \]

\[ Z_n Y_B = \omega_1' g_n g_{n+1} \tan \theta_1 \]

The normalized self-capacitances, \( C_k / \epsilon \), per unit length for the interdigital line elements are

\[ C_2 = \frac{376.7 Y_A N_A Z_2 Y_A}{\sqrt{\epsilon} N_A} \]

\[ C_3 = \frac{376.7 Y_A S_3}{\sqrt{\epsilon} N_A^2} - \frac{1}{N_A} \frac{C_2}{\epsilon} \]

\[ C_k = \frac{376.7 Y_A}{\sqrt{\epsilon} N_A^2} S_k \]

\[ C_{n-1} = \frac{376.7 Y_A}{\sqrt{\epsilon} N_A^2} S_{n-1} - \frac{1}{N_B} \frac{C_n}{\epsilon} \]

\[ C_n = \frac{376.7 Y_B N_B Z_n Y_B}{\sqrt{\epsilon} N_B} \]

where \( \epsilon \) is the dielectric constant, \( \epsilon_r \) is the relative dielectric constant in the medium of propagation, and \( N_A \) is a parameter which is chosen so as to give a practical admittance level in the interior of the filter.
The normalized mutual capacitances \( \frac{C_{k,k+1}}{\varepsilon} \) per unit length between adjacent line elements are

\[
\frac{C_{k,k+1}}{\varepsilon} = \frac{376.7}{\sqrt{(\varepsilon_r)} N_A^2 D_{k,k+1}}
\]

\[
\frac{C_{n-1,n}}{\varepsilon} = \frac{376.7}{\sqrt{(\varepsilon_r)} N_B^2 \left( Z_n Y_B \right)}
\]

Parameter \( Y_{p1} \) of the preceding equations is the characteristic admittance of a short-circuited shunt stub which in a diplexer is replaced by the bandstop filter circuit.

\[ Y_{p1} \]

---

**Fig. 36.** Form of interdigital filter for use in diplexers. (The shunt stub of admittance \( Y_{p1} \) is replaced by the bandstop filter in the diplexer connection.)

---

**Fig. 37.** Open-wire equivalent circuit for the filter of Fig. 36.
the circuit in Fig. 37 would have identical response to that in Fig. 36 except for an approximation introduced by the folding process (discussed in Part VI of Ref. [21]). The parameter $N_A$ is seen to be the turns ratio of an ideal transformer, and since the circuit in Fig. 36 will not be symmetrical even though the low pass prototype used in its design may have been symmetrical, a different turns ratio $N_B$ is generally required at the other end of the filter. The low pass prototype circuit, $Y_A$, $Y_B$, and $N_A$ having been specified, then $N_B$ is automatically fixed. The parameter $N_A$ can be chosen so as to give a convenient impedance level in the interior of the filter circuit; but another perhaps more important criterion for the choice of this value is that the distributed capacitance values $C_k/e$ and $C_{k,k+1}/e$ be of more or less uniform size throughout the filter. By improper choice of $N_A$ it is possible that some of the distributed capacitance values toward the ends of the filter might turn out to be negative. If computer services are available, it is convenient to program the design equations in Table VI and run off a series of designs for different values of $N_A$ in order to obtain a value of $N_A$ which gives a most practical value for the distributed capacitances.

The interdigital filter in Fig. 36 is designed from a doubly terminated, lumped-element low pass Chebyshev prototype filter. The number of elements

Table VII

<table>
<thead>
<tr>
<th>$k$</th>
<th>$Y_k$</th>
<th>$Y_{k,k+1}$</th>
<th>$k$</th>
<th>$Y_k$</th>
<th>$Y_{k,k+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.807</td>
<td>0.688</td>
<td>8</td>
<td>1.248</td>
<td>0.574</td>
</tr>
<tr>
<td>2</td>
<td>1.395</td>
<td>0.644</td>
<td>9</td>
<td>1.008</td>
<td>0.660</td>
</tr>
<tr>
<td>3</td>
<td>1.487</td>
<td>0.634</td>
<td>10</td>
<td>0.705</td>
<td>0.754</td>
</tr>
<tr>
<td>4</td>
<td>1.509</td>
<td>0.628</td>
<td>11</td>
<td>0.451</td>
<td>0.843</td>
</tr>
<tr>
<td>5</td>
<td>1.488</td>
<td>0.598</td>
<td>12</td>
<td>0.246</td>
<td>0.924</td>
</tr>
<tr>
<td>6</td>
<td>1.350</td>
<td>0.534</td>
<td>13</td>
<td>0.076</td>
<td>—</td>
</tr>
<tr>
<td>7</td>
<td>1.220</td>
<td>0.555</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Design is for 0.1-db Chebyshev ripple, $m=13$ resonators, with a stopband fractional bandwidth of $w=0.75$ at the 0.1-db band-edge points. The element values are normalized so that $Y_A=1$.

$n$ in the low pass prototype is chosen to give the desired rate of cutoff, and for this filter the number of elements may be either odd or even. In most cases it will be desirable to have the terminating lines $Y_A$ and $Y_B$ of equal characteristic admittance; but if desired, they can be made unequal without affecting the power transmission characteristics of the filter.
Table VII summarizes the normalized element values for a trial bandstop filter of the form in Fig. 35. This filter was designed from a doubly terminated prototype having 0.10-db Chebyshev ripple and \( m = 13 \) reactive elements. The filter was designed so as to have a stopband fractional bandwidth of \( w = 0.75 \) as measured at the 0.10-db band-edge point. Table VIII presents the normalized distributed capacitance values for a corresponding interdigital bandpass filter structure, which was designed from the same low pass prototype.

Table VIII

| Normalized Distributed Capacitances for a Trial Interdigital Structure of the Type in Fig. 36a |
| --- | --- | --- |
| \( k \) | \( C_k/\varepsilon \) | \( C_{k,k+1}/\varepsilon \) | \( k \) | \( C_k/\varepsilon \) | \( C_{k,k+1}/\varepsilon \) |
| 2 | 2.220 | 1.233 | 8 | 1.634 | 1.070 |
| 3 | 1.563 | 1.103 | 9 | 1.628 | 1.080 |
| 4 | 1.615 | 1.080 | 10 | 1.615 | 1.103 |
| 5 | 1.628 | 1.070 | 11 | 1.578 | 1.170 |
| 6 | 1.634 | 1.067 | 12 | 1.304 | 1.813 |
| 7 | 1.636 | 1.067 | 13 | 2.357 | — |

*This design is for 0.1-db Chebyshev ripple, \( n = 13 \), with a fractional bandwidth of \( w = 0.750 \), \( N_A = 2.80 \), and \( Y_A = Y_B = 0.020 \). \( Y_{p1} = 0.036 \, \Omega^{-1} \).*

filter, and was also designed for a fractional bandwidth of \( w = 0.750 \). In order to check out the design procedure described herein, the attenuation characteristics in the bandstop and bandpass channels of the diplexer were computed. For computational purposes the interdigital filter was represented by the circuit shown in Fig. 37, and the normalized element values for this circuit are presented in Table IX.

Table IX

| Parameters for a Circuit of the Form in Fig. 37 Corresponding to the Interdigital Filter Design in Table VIIIa |
| --- | --- | --- | --- | --- |
| \( k \) | \( Y_k \) | \( Y_{k,k+1} \) | \( k \) | \( Y_k \) | \( Y_{k,k+1} \) |
| 1 | 1.807 | — | 7 | 0.2171 | 0.1416 |
| 2 | — | — | 8 | 0.2169 | 0.1421 |
| 3 | 0.3127 | 0.1463 | 9 | 0.2161 | 0.1433 |
| 4 | 0.2143 | 0.1433 | 10 | 0.2143 | 0.1463 |
| 5 | 0.2161 | 0.1421 | 11 | 0.2094 | 0.1553 |
| 6 | 0.2169 | 0.1416 | 12 | 0.3091 | — |

*The admittances and impedances are normalized so that \( Y_A = Y_B = 1 \). \( Z_2 = 2.182 \); \( Z_{13} = 1.807 \). \( N_A = 2.800 \); \( N_B = 2.300 \).*
The computed performance of the diplexer (including all interaction effects between the two filters) is shown in Fig. 38. In this figure the attenuation characteristics of the bandstop channel of the diplexer are indicated by channel 1, and the attenuation characteristics of the bandpass channel of the diplexer are indicated by channel 2. The stubs of the bandstop filter were exactly one-quarter wavelength long at the normalized frequency \( \omega = 1 \). However, in order to make the crossover at \( \omega = 0.625 \) occur at the 3-db points as desired, the electrical lengths of the resonators in the interdigital filter were shortened slightly so as to raise the passband of the interdigital filter a small amount. This disrupted the crossover at the upper band edge of the bandpass filter slightly, but since this crossover is assumed to be of no interest for diplexer applications, the degradation of the crossover characteristics at the upper edge of the bandpass filter is not considered to be of importance. Note that both the bandpass and bandstop filters are close to the 0.10-db Chebyshev ripple that was prescribed, and any ripples that are not 0.10 db are less than 0.10 db, a situation that is almost always acceptable. It will also be noticed that there are not quite as many ripples as would normally be expected from 13-resonator Chebyshev filters, but this is also of little practical importance.

**Fig. 38.** Computed attenuation versus normalized frequency through channels 1 and 2 of the trial diplexer design. (Channel 1 is bandstop filter channel; channel 2 is bandpass filter channel.)
Fig. 39. Stages in the transformation of a low pass prototype filter into a bandstop transmission-line filter. (a) Singly loaded prototype; (b) mapped prototype; (c) after applying Kuroda's identity to $Y_3$ and $Z_{12}$ in (b); (d) after applying Kuroda's identity to $Z_2$ and $Z_{12}'$ and to $Z_3'$ and $Z_{23}'$ in (c).
The ripples that are missing are most likely ones that would have been crowded very close to cutoff of the filters. All in all the computed performance of the diplexer is judged to be quite satisfactory.

4. Design Procedures and Tables of Bandstop Filter Designs for Use in Diplexers

References [23, 24] describe procedures for exact design of bandstop filters from low pass lumped-element prototype filters. In the procedure, the mapping

\[ \omega' = \frac{\omega}{A \tan \left( \frac{\pi \omega}{2 \omega_0} \right)} \]  

(33)

is used in order to transform the lumped-element, low pass filter into a transmission-line bandstop filter. Use of this mapping technique results in a bandstop filter structure having a ladder configuration with open-circuited stubs connected in shunt and short-circuited stubs connected in series.

**Table X**

<table>
<thead>
<tr>
<th>( w )</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.0812</td>
<td>0.1634</td>
<td>0.2474</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.9850</td>
<td>0.9603</td>
<td>0.9285</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.0962</td>
<td>0.2030</td>
<td>0.3191</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.9302</td>
<td>0.8769</td>
<td>0.8344</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.0698</td>
<td>0.1231</td>
<td>0.1656</td>
</tr>
</tbody>
</table>

**Table XI**

<table>
<thead>
<tr>
<th>( w )</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.1816</td>
<td>0.2753</td>
<td>0.3726</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.9851</td>
<td>0.9453</td>
<td>0.8968</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.2681</td>
<td>0.4271</td>
<td>0.5967</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.8562</td>
<td>0.8208</td>
<td>0.7878</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.2519</td>
<td>0.3787</td>
<td>0.5196</td>
</tr>
<tr>
<td>( h_{34} )</td>
<td>0.8821</td>
<td>0.8120</td>
<td>0.7515</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.2512</td>
<td>0.3655</td>
<td>0.4820</td>
</tr>
<tr>
<td>( h_{45} )</td>
<td>0.8948</td>
<td>0.8690</td>
<td>0.8504</td>
</tr>
<tr>
<td>( h_5 )</td>
<td>0.1052</td>
<td>0.1310</td>
<td>0.1496</td>
</tr>
</tbody>
</table>
Table XII

Designs Having \( m = 7 \) and \( L_{Ar} = 0.10 \) dB

<table>
<thead>
<tr>
<th>( w )</th>
<th>( 0.30 )</th>
<th>( 0.40 )</th>
<th>( 0.50 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.2836</td>
<td>0.3838</td>
<td>0.4892</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.9549</td>
<td>0.8990</td>
<td>0.8407</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.4623</td>
<td>0.6393</td>
<td>0.8255</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.8612</td>
<td>0.8276</td>
<td>0.7834</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.4333</td>
<td>0.6185</td>
<td>0.8215</td>
</tr>
<tr>
<td>( h_{34} )</td>
<td>0.7757</td>
<td>0.7302</td>
<td>0.6947</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.4274</td>
<td>0.5591</td>
<td>0.7073</td>
</tr>
<tr>
<td>( h_{45} )</td>
<td>0.8230</td>
<td>0.7551</td>
<td>0.6908</td>
</tr>
<tr>
<td>( h_5 )</td>
<td>0.4195</td>
<td>0.5660</td>
<td>0.6966</td>
</tr>
<tr>
<td>( h_{36} )</td>
<td>0.8122</td>
<td>0.7752</td>
<td>0.7488</td>
</tr>
<tr>
<td>( h_6 )</td>
<td>0.3206</td>
<td>0.3762</td>
<td>0.4167</td>
</tr>
<tr>
<td>( h_{47} )</td>
<td>0.8950</td>
<td>0.8838</td>
<td>0.8757</td>
</tr>
<tr>
<td>( h_7 )</td>
<td>0.1050</td>
<td>0.1162</td>
<td>0.1243</td>
</tr>
</tbody>
</table>

Table XIII

Designs Having \( m = 9 \) and \( L_{Ar} = 0.10 \) dB

<table>
<thead>
<tr>
<th>( w )</th>
<th>( 0.40 )</th>
<th>( 0.50 )</th>
<th>( 0.68 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.3885</td>
<td>0.4952</td>
<td>0.7008</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.8987</td>
<td>0.8383</td>
<td>0.7347</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.6541</td>
<td>0.8414</td>
<td>1.2045</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.8505</td>
<td>0.7953</td>
<td>0.6916</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.6739</td>
<td>0.8833</td>
<td>1.2761</td>
</tr>
<tr>
<td>( h_{34} )</td>
<td>0.7706</td>
<td>0.7412</td>
<td>0.6681</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.5984</td>
<td>0.8004</td>
<td>1.2130</td>
</tr>
<tr>
<td>( h_{45} )</td>
<td>0.7124</td>
<td>0.6623</td>
<td>0.6009</td>
</tr>
<tr>
<td>( h_5 )</td>
<td>0.6082</td>
<td>0.7447</td>
<td>1.0351</td>
</tr>
<tr>
<td>( h_{56} )</td>
<td>0.7598</td>
<td>0.6974</td>
<td>0.5917</td>
</tr>
<tr>
<td>( h_6 )</td>
<td>0.5871</td>
<td>0.7540</td>
<td>1.0117</td>
</tr>
<tr>
<td>( h_{67} )</td>
<td>0.7449</td>
<td>0.7034</td>
<td>0.6568</td>
</tr>
<tr>
<td>( h_7 )</td>
<td>0.5169</td>
<td>0.5905</td>
<td>0.6765</td>
</tr>
<tr>
<td>( h_{78} )</td>
<td>0.8097</td>
<td>0.7944</td>
<td>0.7770</td>
</tr>
<tr>
<td>( h_8 )</td>
<td>0.3069</td>
<td>0.3303</td>
<td>0.3576</td>
</tr>
<tr>
<td>( h_{89} )</td>
<td>0.9054</td>
<td>0.9002</td>
<td>0.8939</td>
</tr>
<tr>
<td>( h_9 )</td>
<td>0.0946</td>
<td>0.0998</td>
<td>0.1061</td>
</tr>
</tbody>
</table>
The bandstop filter configuration described above has the desired input admittance properties for diplexing (that is, after the first stub is removed from the filter), but since the series stubs in such a structure are very difficult to fabricate, such a ladder configuration is not very practical. In Ref. [23] this difficulty is avoided by inserting extra lengths of transmission line into the structure by use of Kuroda’s identity [24a,b]. These extra lengths of line are exactly one-quarter wavelength long at the mid-stopband frequency \( \omega_0 \); we shall refer to these line lengths as “unit lines.” Kuroda’s identity makes it possible to work unit lines into the structure so as to convert a circuit consisting of a ladder configuration of series and shunt stubs into a structure such as that in Fig. 35, which uses only shunt stubs [24a, b]. The final configuration, then, is much more practical to build than is the original ladder configuration. This procedure is shown schematically in Fig. 39. However, if any unit lines are fed through the structure from the end of the filter structure which is to be connected to the common junction of the diplexer, the input admittance properties of the filter at that end will be considerably changed, and the

<table>
<thead>
<tr>
<th>Table XIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESIGNS HAVING ( m = 11 ) AND ( L_{AR} = 0.10 ) DB</td>
</tr>
<tr>
<td>( 0.50 )</td>
</tr>
<tr>
<td>( h_1 )</td>
</tr>
<tr>
<td>( h_{12} )</td>
</tr>
<tr>
<td>( h_2 )</td>
</tr>
<tr>
<td>( h_{23} )</td>
</tr>
<tr>
<td>( h_3 )</td>
</tr>
<tr>
<td>( h_{34} )</td>
</tr>
<tr>
<td>( h_4 )</td>
</tr>
<tr>
<td>( h_{45} )</td>
</tr>
<tr>
<td>( h_5 )</td>
</tr>
<tr>
<td>( h_{56} )</td>
</tr>
<tr>
<td>( h_6 )</td>
</tr>
<tr>
<td>( h_{67} )</td>
</tr>
<tr>
<td>( h_7 )</td>
</tr>
<tr>
<td>( h_{78} )</td>
</tr>
<tr>
<td>( h_8 )</td>
</tr>
<tr>
<td>( h_{89} )</td>
</tr>
<tr>
<td>( h_9 )</td>
</tr>
<tr>
<td>( h_{910} )</td>
</tr>
<tr>
<td>( h_{10} )</td>
</tr>
<tr>
<td>( h_{1011} )</td>
</tr>
<tr>
<td>( h_{11} )</td>
</tr>
</tbody>
</table>
performance of the diplexer will be greatly disrupted. The procedure described in Refs. [23, 24] provides for introducing unit lines through both ends of the filter structure, and for that reason the designs obtained by the information in Refs. [23, 24] are not suitable for use in diplexers, at least not unless the design procedures are modified somewhat.

The designs, which are tabulated in Tables X through XXIII, were obtained by feeding unit lines into the filter structure from the load end only, so that the input admittance properties of the structure at the other end would not be disturbed. For example, when the structure in Fig. 35 is designed from these tables, the input admittance $Y_s$ at the right end of the structure can be obtained directly from the corresponding input admittance of the lumped-element, low pass prototype filter by use of the mapping function in Eq. (33). Thus, the

<table>
<thead>
<tr>
<th>Table XV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Designs Having $m=13$ and $L_{Ar}=0.10$ dB</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.68</th>
<th>0.75</th>
<th>1.00</th>
<th>1.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.7077</td>
<td>0.8067</td>
<td>1.2074</td>
<td>1.6618</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.7314</td>
<td>0.6887</td>
<td>0.5453</td>
<td>0.4341</td>
</tr>
<tr>
<td>$h_2$</td>
<td>1.2185</td>
<td>1.3950</td>
<td>2.1138</td>
<td>2.9323</td>
</tr>
<tr>
<td>$h_{23}$</td>
<td>0.6885</td>
<td>0.6435</td>
<td>0.5013</td>
<td>0.3943</td>
</tr>
<tr>
<td>$h_3$</td>
<td>1.3020</td>
<td>1.4872</td>
<td>2.2366</td>
<td>3.0869</td>
</tr>
<tr>
<td>$h_{34}$</td>
<td>0.6784</td>
<td>0.6339</td>
<td>0.4930</td>
<td>0.3875</td>
</tr>
<tr>
<td>$h_4$</td>
<td>1.3201</td>
<td>1.5086</td>
<td>2.2643</td>
<td>3.1191</td>
</tr>
<tr>
<td>$h_{45}$</td>
<td>0.6674</td>
<td>0.6276</td>
<td>0.4911</td>
<td>0.3864</td>
</tr>
<tr>
<td>$h_5$</td>
<td>1.2863</td>
<td>1.4883</td>
<td>2.2618</td>
<td>3.1140</td>
</tr>
<tr>
<td>$h_{56}$</td>
<td>0.6223</td>
<td>0.5979</td>
<td>0.4878</td>
<td>0.3879</td>
</tr>
<tr>
<td>$h_6$</td>
<td>1.1442</td>
<td>1.3496</td>
<td>2.1758</td>
<td>3.0432</td>
</tr>
<tr>
<td>$h_{67}$</td>
<td>0.5629</td>
<td>0.5344</td>
<td>0.4565</td>
<td>0.3831</td>
</tr>
<tr>
<td>$h_7$</td>
<td>1.0920</td>
<td>1.2202</td>
<td>1.8434</td>
<td>2.6125</td>
</tr>
<tr>
<td>$h_{78}$</td>
<td>0.5991</td>
<td>0.5552</td>
<td>0.4322</td>
<td>0.3684</td>
</tr>
<tr>
<td>$h_8$</td>
<td>1.1016</td>
<td>1.2476</td>
<td>1.6433</td>
<td>1.8649</td>
</tr>
<tr>
<td>$h_{89}$</td>
<td>0.5981</td>
<td>0.5743</td>
<td>0.5254</td>
<td>0.5030</td>
</tr>
<tr>
<td>$h_9$</td>
<td>0.9527</td>
<td>1.0083</td>
<td>1.1308</td>
<td>1.1925</td>
</tr>
<tr>
<td>$h_{9,10}$</td>
<td>0.6699</td>
<td>0.6603</td>
<td>0.6394</td>
<td>0.6288</td>
</tr>
<tr>
<td>$h_{10}$</td>
<td>0.6843</td>
<td>0.7045</td>
<td>0.7505</td>
<td>0.7751</td>
</tr>
<tr>
<td>$h_{10,11}$</td>
<td>0.7588</td>
<td>0.7542</td>
<td>0.7436</td>
<td>0.7378</td>
</tr>
<tr>
<td>$h_{11}$</td>
<td>0.4429</td>
<td>0.4515</td>
<td>0.4720</td>
<td>0.4835</td>
</tr>
<tr>
<td>$h_{11,12}$</td>
<td>0.8451</td>
<td>0.8428</td>
<td>0.8372</td>
<td>0.8340</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.2425</td>
<td>0.2462</td>
<td>0.2552</td>
<td>0.2604</td>
</tr>
<tr>
<td>$h_{12,13}$</td>
<td>0.9254</td>
<td>0.9245</td>
<td>0.9220</td>
<td>0.9206</td>
</tr>
<tr>
<td>$h_{13}$</td>
<td>0.0746</td>
<td>0.0755</td>
<td>0.0780</td>
<td>0.0794</td>
</tr>
</tbody>
</table>
Table XVI

**Designs Having** $m = 15$ AND $L_{Ar} = 0.10$ **dB**

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.68</th>
<th>0.75</th>
<th>1.00</th>
<th>1.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.7093</td>
<td>0.8086</td>
<td>1.2101</td>
<td>1.6656</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.7306</td>
<td>0.6869</td>
<td>0.5445</td>
<td>0.4333</td>
</tr>
<tr>
<td>$h_2$</td>
<td>1.2216</td>
<td>1.3985</td>
<td>2.1191</td>
<td>2.9398</td>
</tr>
<tr>
<td>$h_{33}$</td>
<td>0.6877</td>
<td>0.6426</td>
<td>0.5003</td>
<td>0.3934</td>
</tr>
<tr>
<td>$h_3$</td>
<td>1.3059</td>
<td>1.4916</td>
<td>2.2435</td>
<td>3.0967</td>
</tr>
<tr>
<td>$h_{34}$</td>
<td>0.6778</td>
<td>0.6330</td>
<td>0.4918</td>
<td>0.3864</td>
</tr>
<tr>
<td>$h_4$</td>
<td>1.3282</td>
<td>1.5161</td>
<td>2.2743</td>
<td>3.1340</td>
</tr>
<tr>
<td>$h_{45}$</td>
<td>0.6728</td>
<td>0.6293</td>
<td>0.4894</td>
<td>0.3845</td>
</tr>
<tr>
<td>$h_5$</td>
<td>1.3261</td>
<td>1.5192</td>
<td>2.2820</td>
<td>3.1413</td>
</tr>
<tr>
<td>$h_{56}$</td>
<td>0.6557</td>
<td>0.6205</td>
<td>0.4887</td>
<td>0.3847</td>
</tr>
<tr>
<td>$h_6$</td>
<td>1.2600</td>
<td>1.4731</td>
<td>2.2701</td>
<td>3.1290</td>
</tr>
<tr>
<td>$h_{67}$</td>
<td>0.5982</td>
<td>0.5790</td>
<td>0.4827</td>
<td>0.3859</td>
</tr>
<tr>
<td>$h_7$</td>
<td>1.1133</td>
<td>1.3101</td>
<td>2.1563</td>
<td>3.0486</td>
</tr>
<tr>
<td>$h_{78}$</td>
<td>0.5601</td>
<td>0.5244</td>
<td>0.4438</td>
<td>0.3762</td>
</tr>
<tr>
<td>$h_8$</td>
<td>1.1329</td>
<td>1.2532</td>
<td>1.8364</td>
<td>2.6109</td>
</tr>
<tr>
<td>$h_{89}$</td>
<td>0.5994</td>
<td>0.5588</td>
<td>0.4276</td>
<td>0.3566</td>
</tr>
<tr>
<td>$h_9$</td>
<td>1.1025</td>
<td>1.2654</td>
<td>1.7428</td>
<td>2.0125</td>
</tr>
<tr>
<td>$h_{9,10}$</td>
<td>0.5877</td>
<td>0.5577</td>
<td>0.4980</td>
<td>0.4727</td>
</tr>
<tr>
<td>$h_{10}$</td>
<td>1.0359</td>
<td>1.1177</td>
<td>1.2726</td>
<td>1.3488</td>
</tr>
<tr>
<td>$h_{10,11}$</td>
<td>0.6369</td>
<td>0.6245</td>
<td>0.5993</td>
<td>0.5874</td>
</tr>
<tr>
<td>$h_{11}$</td>
<td>0.8065</td>
<td>0.8345</td>
<td>0.8950</td>
<td>0.9256</td>
</tr>
<tr>
<td>$h_{11,12}$</td>
<td>0.7135</td>
<td>0.7076</td>
<td>0.6947</td>
<td>0.6879</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.5746</td>
<td>0.5867</td>
<td>0.6145</td>
<td>0.6295</td>
</tr>
<tr>
<td>$h_{12,13}$</td>
<td>0.7917</td>
<td>0.7885</td>
<td>0.7812</td>
<td>0.7771</td>
</tr>
<tr>
<td>$h_{13}$</td>
<td>0.3764</td>
<td>0.3822</td>
<td>0.3961</td>
<td>0.4039</td>
</tr>
<tr>
<td>$h_{13,14}$</td>
<td>0.8658</td>
<td>0.8641</td>
<td>0.8600</td>
<td>0.8577</td>
</tr>
<tr>
<td>$h_{14}$</td>
<td>0.2088</td>
<td>0.2115</td>
<td>0.2180</td>
<td>0.2217</td>
</tr>
<tr>
<td>$h_{14,15}$</td>
<td>0.9351</td>
<td>0.9344</td>
<td>0.9326</td>
<td>0.9315</td>
</tr>
<tr>
<td>$h_{15}$</td>
<td>0.0649</td>
<td>0.0656</td>
<td>0.0674</td>
<td>0.0685</td>
</tr>
</tbody>
</table>

Table XVII

**Designs Having** $m = 3$ AND $L_{Ar} = 0.01$ **dB**

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.10</th>
<th>0.20</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.0495</td>
<td>0.0996</td>
<td>0.1510</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.9716</td>
<td>0.9407</td>
<td>0.9077</td>
</tr>
<tr>
<td>$h_2$</td>
<td>0.0779</td>
<td>0.1590</td>
<td>0.2434</td>
</tr>
<tr>
<td>$h_{23}$</td>
<td>0.9549</td>
<td>0.9169</td>
<td>0.8840</td>
</tr>
<tr>
<td>$h_3$</td>
<td>0.0451</td>
<td>0.0831</td>
<td>0.1160</td>
</tr>
</tbody>
</table>
### Table XVIII

**Designs Having $m=5$ and $L_{Ar}=0.01$ dB**

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.20</th>
<th>0.30</th>
<th>0.40</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.1198</td>
<td>0.1816</td>
<td>0.2457</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.9351</td>
<td>0.8936</td>
<td>0.8495</td>
</tr>
<tr>
<td>$h_2$</td>
<td>0.2240</td>
<td>0.3474</td>
<td>0.4782</td>
</tr>
<tr>
<td>$h_{23}$</td>
<td>0.8812</td>
<td>0.8354</td>
<td>0.7914</td>
</tr>
<tr>
<td>$h_3$</td>
<td>0.2326</td>
<td>0.3506</td>
<td>0.4744</td>
</tr>
<tr>
<td>$h_{34}$</td>
<td>0.8920</td>
<td>0.8382</td>
<td>0.7900</td>
</tr>
<tr>
<td>$h_4$</td>
<td>0.2046</td>
<td>0.2950</td>
<td>0.3747</td>
</tr>
<tr>
<td>$h_{45}$</td>
<td>0.9190</td>
<td>0.8948</td>
<td>0.8761</td>
</tr>
<tr>
<td>$h_5$</td>
<td>0.0810</td>
<td>0.1052</td>
<td>0.1239</td>
</tr>
</tbody>
</table>

### Table XIX

**Designs Having $m=7$ and $L_{Ar}=0.01$ dB**

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.40</th>
<th>0.50</th>
<th>0.68</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.2646</td>
<td>0.3373</td>
<td>0.4773</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.8391</td>
<td>0.7887</td>
<td>0.7006</td>
</tr>
<tr>
<td>$h_2$</td>
<td>0.5344</td>
<td>0.6888</td>
<td>0.9915</td>
</tr>
<tr>
<td>$h_{23}$</td>
<td>0.7911</td>
<td>0.7366</td>
<td>0.6413</td>
</tr>
<tr>
<td>$h_3$</td>
<td>0.5941</td>
<td>0.7673</td>
<td>1.0995</td>
</tr>
<tr>
<td>$h_{34}$</td>
<td>0.7582</td>
<td>0.7110</td>
<td>0.6249</td>
</tr>
<tr>
<td>$h_4$</td>
<td>0.5847</td>
<td>0.7570</td>
<td>1.0938</td>
</tr>
<tr>
<td>$h_{45}$</td>
<td>0.7390</td>
<td>0.6865</td>
<td>0.6072</td>
</tr>
<tr>
<td>$h_5$</td>
<td>0.5818</td>
<td>0.7280</td>
<td>1.0111</td>
</tr>
<tr>
<td>$h_{56}$</td>
<td>0.7592</td>
<td>0.7051</td>
<td>0.6181</td>
</tr>
<tr>
<td>$h_6$</td>
<td>0.5492</td>
<td>0.6855</td>
<td>0.9004</td>
</tr>
<tr>
<td>$h_{67}$</td>
<td>0.7701</td>
<td>0.7326</td>
<td>0.6859</td>
</tr>
<tr>
<td>$h_7$</td>
<td>0.4533</td>
<td>0.5227</td>
<td>0.6116</td>
</tr>
<tr>
<td>$h_{78}$</td>
<td>0.8308</td>
<td>0.8140</td>
<td>0.7935</td>
</tr>
<tr>
<td>$h_8$</td>
<td>0.2715</td>
<td>0.2976</td>
<td>0.3301</td>
</tr>
<tr>
<td>$h_{89}$</td>
<td>0.9151</td>
<td>0.9088</td>
<td>0.9009</td>
</tr>
<tr>
<td>$h_9$</td>
<td>0.0849</td>
<td>0.0912</td>
<td>0.0991</td>
</tr>
</tbody>
</table>
Table XX

**Designs Having \( m = 9 \) and \( L_{Ar} = 0.01 \, \text{db} \)**

<table>
<thead>
<tr>
<th>( w )</th>
<th>0.30</th>
<th>0.40</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.1913</td>
<td>0.2589</td>
<td>0.3301</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.8907</td>
<td>0.8425</td>
<td>0.7934</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.3776</td>
<td>0.5189</td>
<td>0.6694</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.8371</td>
<td>0.7893</td>
<td>0.7391</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.4057</td>
<td>0.5601</td>
<td>0.7252</td>
</tr>
<tr>
<td>( h_{34} )</td>
<td>0.8054</td>
<td>0.7539</td>
<td>0.7073</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.4048</td>
<td>0.5394</td>
<td>0.6819</td>
</tr>
<tr>
<td>( h_{45} )</td>
<td>0.8246</td>
<td>0.7688</td>
<td>0.7169</td>
</tr>
<tr>
<td>( h_5 )</td>
<td>0.3764</td>
<td>0.4948</td>
<td>0.6027</td>
</tr>
<tr>
<td>( h_{56} )</td>
<td>0.8395</td>
<td>0.8055</td>
<td>0.7791</td>
</tr>
<tr>
<td>( h_6 )</td>
<td>0.2689</td>
<td>0.3217</td>
<td>0.3631</td>
</tr>
<tr>
<td>( h_{67} )</td>
<td>0.9109</td>
<td>0.8986</td>
<td>0.8893</td>
</tr>
<tr>
<td>( h_7 )</td>
<td>0.0891</td>
<td>0.1014</td>
<td>0.1107</td>
</tr>
</tbody>
</table>

Table XXI

**Designs Having \( m = 11 \) and \( L_{Ar} = 0.01 \, \text{db} \)**

<table>
<thead>
<tr>
<th>( w )</th>
<th>0.50</th>
<th>0.68</th>
<th>0.75</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_1 )</td>
<td>0.3411</td>
<td>0.4826</td>
<td>0.5502</td>
<td>0.8234</td>
</tr>
<tr>
<td>( h_{12} )</td>
<td>0.7862</td>
<td>0.6974</td>
<td>0.6593</td>
<td>0.5318</td>
</tr>
<tr>
<td>( h_2 )</td>
<td>0.6981</td>
<td>1.0048</td>
<td>1.1533</td>
<td>1.7612</td>
</tr>
<tr>
<td>( h_{23} )</td>
<td>0.7345</td>
<td>0.6372</td>
<td>0.5968</td>
<td>0.4675</td>
</tr>
<tr>
<td>( h_3 )</td>
<td>0.7844</td>
<td>1.1210</td>
<td>1.2820</td>
<td>1.9342</td>
</tr>
<tr>
<td>( h_{34} )</td>
<td>0.7158</td>
<td>0.6223</td>
<td>0.5826</td>
<td>0.4556</td>
</tr>
<tr>
<td>( h_4 )</td>
<td>0.7958</td>
<td>1.1444</td>
<td>1.3086</td>
<td>1.9649</td>
</tr>
<tr>
<td>( h_{45} )</td>
<td>0.6916</td>
<td>0.6110</td>
<td>0.5752</td>
<td>0.4549</td>
</tr>
<tr>
<td>( h_5 )</td>
<td>0.7677</td>
<td>1.1109</td>
<td>1.2759</td>
<td>1.9254</td>
</tr>
<tr>
<td>( h_{56} )</td>
<td>0.6781</td>
<td>0.5926</td>
<td>0.5604</td>
<td>0.4571</td>
</tr>
<tr>
<td>( h_6 )</td>
<td>0.7630</td>
<td>1.0477</td>
<td>1.1850</td>
<td>1.7368</td>
</tr>
<tr>
<td>( h_{67} )</td>
<td>0.6977</td>
<td>0.6080</td>
<td>0.5726</td>
<td>0.4759</td>
</tr>
<tr>
<td>( h_7 )</td>
<td>0.7262</td>
<td>0.9844</td>
<td>1.0854</td>
<td>1.3706</td>
</tr>
<tr>
<td>( h_{78} )</td>
<td>0.7079</td>
<td>0.6479</td>
<td>0.6291</td>
<td>0.5845</td>
</tr>
<tr>
<td>( h_8 )</td>
<td>0.6312</td>
<td>0.7593</td>
<td>0.8297</td>
<td>0.9042</td>
</tr>
<tr>
<td>( h_{89} )</td>
<td>0.7627</td>
<td>0.7340</td>
<td>0.7252</td>
<td>0.7037</td>
</tr>
<tr>
<td>( h_9 )</td>
<td>0.4446</td>
<td>0.4977</td>
<td>0.5146</td>
<td>0.5568</td>
</tr>
<tr>
<td>( h_{9,10} )</td>
<td>0.8414</td>
<td>0.8278</td>
<td>0.8235</td>
<td>0.8127</td>
</tr>
<tr>
<td>( h_{10} )</td>
<td>0.2501</td>
<td>0.2715</td>
<td>0.2784</td>
<td>0.2958</td>
</tr>
<tr>
<td>( h_{10,11} )</td>
<td>0.9227</td>
<td>0.9172</td>
<td>0.9154</td>
<td>0.9108</td>
</tr>
<tr>
<td>( h_{11} )</td>
<td>0.0773</td>
<td>0.0828</td>
<td>0.0846</td>
<td>0.0892</td>
</tr>
</tbody>
</table>
input admittance at the right-hand side of the structure is well behaved, and that end is suitable for use at the common junction of a diplexer.

It should be noted that Tables X to XVI are for $L_{Ar}=0.10$-db Chebyshev ripple in the passbands, while Tables XVII to XXIII are for $L_{Ar}=0.01$-db

Table XXII

<table>
<thead>
<tr>
<th>Designs Having $m=13$ and $L_{Ar}=0.01$ db</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$</td>
</tr>
<tr>
<td>$h_1$</td>
</tr>
<tr>
<td>$h_{12}$</td>
</tr>
<tr>
<td>$h_2$</td>
</tr>
<tr>
<td>$h_{23}$</td>
</tr>
<tr>
<td>$h_3$</td>
</tr>
<tr>
<td>$h_{34}$</td>
</tr>
<tr>
<td>$h_4$</td>
</tr>
<tr>
<td>$h_{45}$</td>
</tr>
<tr>
<td>$h_5$</td>
</tr>
<tr>
<td>$h_{56}$</td>
</tr>
<tr>
<td>$h_6$</td>
</tr>
<tr>
<td>$h_{67}$</td>
</tr>
<tr>
<td>$h_7$</td>
</tr>
<tr>
<td>$h_{78}$</td>
</tr>
<tr>
<td>$h_8$</td>
</tr>
<tr>
<td>$h_{89}$</td>
</tr>
<tr>
<td>$h_9$</td>
</tr>
<tr>
<td>$h_{9,10}$</td>
</tr>
<tr>
<td>$h_{10}$</td>
</tr>
<tr>
<td>$h_{10,11}$</td>
</tr>
<tr>
<td>$h_{11}$</td>
</tr>
<tr>
<td>$h_{11,12}$</td>
</tr>
<tr>
<td>$h_{12}$</td>
</tr>
<tr>
<td>$h_{12,13}$</td>
</tr>
<tr>
<td>$h_{13}$</td>
</tr>
</tbody>
</table>

Chebyshev ripple in the passbands. The parameter $w$ is the fractional stopband width as measured to the edges of the equal-ripple passbands. The parameter $m$ is the number of resonators (that is, stubs) in the filter structure. The interpretation of the parameters $h_k$ and $h_{k, k-1}$ was explained previously.

Tables X to XXIII were prepared by using a modification of Cristal’s algorithm, which is discussed in Ref. [24].
Table XXIII

| Designs Having $m=15$ and $L_{A1}=0.01$ dB |

<table>
<thead>
<tr>
<th>$w$</th>
<th>0.68</th>
<th>0.75</th>
<th>1.00</th>
<th>1.20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>0.4876</td>
<td>0.5559</td>
<td>0.8320</td>
<td>1.1451</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.6944</td>
<td>0.6562</td>
<td>0.5283</td>
<td>0.4250</td>
</tr>
<tr>
<td>$h_2$</td>
<td>1.0168</td>
<td>1.1672</td>
<td>1.7827</td>
<td>2.4867</td>
</tr>
<tr>
<td>$h_{23}$</td>
<td>0.6336</td>
<td>0.5930</td>
<td>0.4634</td>
<td>0.3652</td>
</tr>
<tr>
<td>$h_3$</td>
<td>1.1374</td>
<td>1.3009</td>
<td>1.9640</td>
<td>2.7170</td>
</tr>
<tr>
<td>$h_{34}$</td>
<td>0.6185</td>
<td>0.5780</td>
<td>0.4500</td>
<td>0.3540</td>
</tr>
<tr>
<td>$h_4$</td>
<td>1.1724</td>
<td>1.3388</td>
<td>2.0118</td>
<td>2.7748</td>
</tr>
<tr>
<td>$h_{45}$</td>
<td>0.6129</td>
<td>0.5729</td>
<td>0.4461</td>
<td>0.3510</td>
</tr>
<tr>
<td>$h_5$</td>
<td>1.1816</td>
<td>1.3499</td>
<td>2.0244</td>
<td>2.7866</td>
</tr>
<tr>
<td>$h_{56}$</td>
<td>0.6067</td>
<td>0.5693</td>
<td>0.4456</td>
<td>0.3513</td>
</tr>
<tr>
<td>$h_6$</td>
<td>1.1643</td>
<td>1.3375</td>
<td>2.0161</td>
<td>2.7691</td>
</tr>
<tr>
<td>$h_{67}$</td>
<td>0.5900</td>
<td>0.5580</td>
<td>0.4459</td>
<td>0.3546</td>
</tr>
<tr>
<td>$h_7$</td>
<td>1.1149</td>
<td>1.2818</td>
<td>1.9609</td>
<td>2.6951</td>
</tr>
<tr>
<td>$h_{78}$</td>
<td>0.5799</td>
<td>0.5437</td>
<td>0.4409</td>
<td>0.3616</td>
</tr>
<tr>
<td>$h_8$</td>
<td>1.1048</td>
<td>1.2411</td>
<td>1.8017</td>
<td>2.4215</td>
</tr>
<tr>
<td>$h_{89}$</td>
<td>0.5983</td>
<td>0.5617</td>
<td>0.4515</td>
<td>0.3858</td>
</tr>
<tr>
<td>$h_9$</td>
<td>1.0624</td>
<td>1.1953</td>
<td>1.5971</td>
<td>1.8550</td>
</tr>
<tr>
<td>$h_{9,10}$</td>
<td>0.6085</td>
<td>0.5815</td>
<td>0.5209</td>
<td>0.4912</td>
</tr>
<tr>
<td>$h_{10}$</td>
<td>0.9549</td>
<td>1.0258</td>
<td>1.1942</td>
<td>1.2845</td>
</tr>
<tr>
<td>$h_{10,11}$</td>
<td>0.6581</td>
<td>0.6444</td>
<td>0.6141</td>
<td>0.5987</td>
</tr>
<tr>
<td>$h_{11}$</td>
<td>0.7502</td>
<td>0.7817</td>
<td>0.8548</td>
<td>0.8945</td>
</tr>
<tr>
<td>$h_{11,12}$</td>
<td>0.7283</td>
<td>0.7211</td>
<td>0.7043</td>
<td>0.6952</td>
</tr>
<tr>
<td>$h_{12}$</td>
<td>0.5416</td>
<td>0.5566</td>
<td>0.5926</td>
<td>0.6128</td>
</tr>
<tr>
<td>$h_{12,13}$</td>
<td>0.8012</td>
<td>0.7972</td>
<td>0.7873</td>
<td>0.7818</td>
</tr>
<tr>
<td>$h_{13}$</td>
<td>0.3580</td>
<td>0.3655</td>
<td>0.3841</td>
<td>0.3948</td>
</tr>
<tr>
<td>$h_{13,14}$</td>
<td>0.8714</td>
<td>0.8692</td>
<td>0.8636</td>
<td>0.8604</td>
</tr>
<tr>
<td>$h_{14}$</td>
<td>0.1997</td>
<td>0.2033</td>
<td>0.2121</td>
<td>0.2173</td>
</tr>
<tr>
<td>$h_{14,15}$</td>
<td>0.9377</td>
<td>0.9367</td>
<td>0.9342</td>
<td>0.9328</td>
</tr>
<tr>
<td>$h_{15}$</td>
<td>0.0623</td>
<td>0.0633</td>
<td>0.0658</td>
<td>0.0672</td>
</tr>
</tbody>
</table>

The modifications are, first, 

$$P = m - 1 \tag{34}$$

where $m$ is the number of resonators in the filter structure, and second, 

$$h_1 = g_{1'} \tag{35}$$

The reader is referred to Ref. [24] for details of the algorithm and definitions of the parameters, $P$ and $g_{1'}$. 

V. THEORY OF MULTIPLEXERS HAVING NONCONTIGUOUS CHANNELS

A. INTRODUCTION

The theory of multiplexers having noncontiguous channels (that is, frequency channels that are separated from one another by frequency guard bands) is described in this part. A representative transmission attenuation response of a multiplexer of this type is shown in Fig. 40. Note that adjacent channels are typically separated by a band of frequencies, and that there is good isolation between adjacent channels at the crossover frequency. A design realizing the characteristics given in Fig. 40 is shown schematically in Fig. 41. Here a tandem arrangement of \( n \) bandpass and \( n \) bandstop filters is illustrated. The cascade of filters is divided into groups of channel-separating units as indicated in the figure. Each channel-separating unit consists of a bandpass filter and a bandstop filter arranged so that the bandpass filter of each unit is closer to the generator. Each channel-separating unit separates a band of frequencies about the center frequency of the unit from the remaining frequencies of the total spectrum. In this design, each pair of bandpass and bandstop filters is designed to produce a constant-resistance input impedance at all frequencies. Each channel-separating unit is matched to the previous channel-
separating unit or to the generator; therefore, any number of channel-separating units may be cascaded without harmful interaction effects.

The theory of operation of the multiplexer design shown in Fig. 41 can be understood by considering the design of a single channel-separating unit.

**Fig. 41.** Schematic representation of a multiplexer.

In Fig. 42 is shown a lumped-element realization of a series connected channel-separating unit of the type under consideration. The unit consists of a minimum reactance three-resonator bandstop filter and a minimum reactance three-resonator bandpass filter. Each filter is tuned to the same center frequency.

**Fig. 42.** Lumped-element, series connected, channel-separating unit.

Note that the first resonator of each filter is in shunt, as required for the filter to be minimum reactance. The input impedances $Z_{AC}$ and $Z_{CB}$ of the bandpass and bandstop filters, respectively, are sketched in Fig. 43. Note that the general appearance of the input impedance is that of Fig. 5b for the bandpass filter, and that of Fig. 5d for that of the bandstop filter. This results, of course, from
the fact that the filters are minimum reactance. However, also note in this example that the real parts of the input impedances have been drawn so that they add up to a constant for all frequencies; that is (the filters are exactly complementary), the real part of the input impedance seen at port 1 in Fig. 42 is a constant. Hence, by Bode's relationship for minimum reactance networks [Eq. (7)], the imaginary part of the input impedance is identically zero. Therefore, in Fig. 43 the imaginary components of the input impedance of the bandpass and bandstop filter have been drawn so that they are conjugates of each other, in order that they may sum to zero at port 1 at all frequencies. Note next that because each channel-separating unit provides a constant resistance input impedance, the terminating resistor, \( R_L \), at port 3 of a given channel-separating unit may be replaced by another channel-separating unit also having a constant-resistance input impedance, without disrupting the performance of the first. The bandwidth and center frequency of the second channel-separating unit may be arbitrarily chosen without affecting the first unit. In this way any number of channel-separating units may be cascaded to realize a desired multiplexer response.

Let us rephrase the basic ideas expressed above. The bandpass and bandstop filters of each channel-separating unit are designed to be complementary filters. If the filters are to be electrically connected in series, the sum of their input impedances should be a constant equal to the generator resistance or equal to the load resistance of the preceding channel-separating unit. The filters must be minimum reactance for the series connected case; this is accomplished by having the first resonator of the two filters be in shunt. If the (complementary) bandpass and bandstop filters are to be electrically connected in parallel, the sum of their input admittances should be a constant equal to the generator conductance or the conductance of the preceding channel-separating unit. The filters must be minimum susceptance for the parallel connected case; this
is accomplished by having the first resonator of the two filters be in series. The multiplexer is synthesized by designing the specified number of channel-separating units, each having their prescribed bandwidth and center frequency, and connecting the units in tandem.

As has been pointed out previously, it is not essential that the bandpass and bandstop filters be exactly complementary. In fact, this can seldom be accomplished. The important point is that the sum of the real parts of the impedances should be approximately constant. If this is done there will result only a small residual reactance in the out-of-band region of the channel-separating unit, which may be characterized by stating that each channel-separating unit has a small residual VSWR. If the residual VSWR of each channel-separating unit is kept small, many channel-separating units may be cascaded without harmful interaction effects. Eventually, however, there is a buildup in VSWR which finally limits the number of channel-separating units that may be cascaded. In theory, however, there is no limitation.

When the channel to be separated has a wide or moderate bandwidth, a bandpass and bandstop filter combination as in Fig. 34, designed for a small Chebyshev ripple, can be used. However, in most cases of noncontiguous channels, the fractional bandwidth of the channels is small and various other filter structures (such as the strip line structure about to be discussed) are preferable.

B. EXAMPLE DESIGN AND MEASURED RESPONSE OF A SERIES CONNECTED MULTIPLEXER CHANNEL-SEPARATING UNIT IN STRIP LINE

A sketch of a strip line multiplexer channel-separating unit is given in Fig. 44. It consists of an interdigital bandpass filter and a parallel-coupled-resonator bandstop filter electrically connected in series at point B in Fig. 44. In the figure, \( Z_A \) represents the generator and terminating resistance. The input impedance \( Z_T \) is, in theory, exactly equal to \( Z_A \) at all frequencies. Thus the load impedance

\[
Z_Z = Z_A + R_B
\]

![Fig. 44. Strip line multiplexer channel unit.](image)

\(^1\) This section is largely taken from Ref. [25], with the permission of the IEEE.
Z_A may be replaced by another channel-separating unit, and a number of channel-separating units may be cascaded without harmful interaction effects.

The design of this example is based on using singly terminated maximally flat prototype filters, since this prototype offers the greatest possibility for minimizing the residual VSWR. If a large number of channel-separating units are to be cascaded, this singly terminated maximally flat prototype is recommended in order to reduce interchannel interference to a minimum.

1. Design of the Bandstop Filter

The bandstop filter of the multiplexer channel-separating unit is realized by applying the exact synthesis theory of Schiffman and Matthaei [23], and a modification of the computation algorithm of Cristal [24]. This design method was discussed previously in connection with the design of the bandstop filter in Section IV.C, and will not be further described.

![Diagram of bandstop filter](image)

Fig. 45. (a) Open-wire-line bandstop filter. (b) Its parallel coupled strip line equivalent.

In the present example, the bandstop filter was designed from a three-resonator, singly terminated, maximally flat, low pass prototype filter. The stopband fractional design bandwidth was 0.05. The h parameters of the bandstop filters are given in Table XXIV. (Recall that herein the h parameters are the relative characteristic admittances or impedance of the stubs and connecting lines normalized with respect to the load conductance or impedance.)

The significance of the h parameters is further clarified by Fig. 45, which shows the open-wire line equivalent circuit for the bandstop filter and also the parallel-coupled-resonator realization of the filter.
Design equations relating the admittance of the lines and stubs in Fig. 45a to the normalized self- and mutual-distributed capacitances of the coupled resonators in Fig. 45b are given in Ref. [23]. These equations, together with the design procedure given in Ref. [22], are sufficient to realize the bandstop filter shown in Fig. 45b.

Table XXIV

<table>
<thead>
<tr>
<th>$h_1$</th>
<th>$h_{12}$</th>
<th>$h_2$</th>
<th>$h_{23}$</th>
<th>$h_3$</th>
<th>$h_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0589</td>
<td>0.9679</td>
<td>0.0517</td>
<td>0.9811</td>
<td>0.0189</td>
<td>1</td>
</tr>
</tbody>
</table>

2. Design of the Bandpass Filter and the Interconnection to the Bandstop Filter

Figure 46a shows an interdigital bandpass filter such as is used in the multiplexer channel unit in Fig. 44. Equations for the design of filters of this type

\[
T = \text{Turns Ratio} = \frac{v\alpha_{01}}{v\alpha_{54}}
\]

Fig. 46. An interdigital filter and its open-wire-line equivalent. (The capacitance $C_i$ is the capacitance per unit length between strip $i$ and ground, and capacitance $C_{i,i+1}$ is the capacitance between strips $i$ and $i+1$. The transformers shown are ideal transformers, and $v$ is the velocity of propagation.)
from low pass prototype filters are presented in Ref. [17]. The interdigital filter structure in Fig. 46a is of the type that uses short-circuited input and output lines, which act as impedance transforming sections (not as resonators) [17]. Thus, the structure shown is a three-resonator filter (any number of resonators may be used, of course), the resonators being formed from lines 1, 2, and 3. This form of interdigital bandpass filter, which uses short-circuited input and output lines, is most practical for narrow to moderate bandwidths [17].

Figure 46b shows an equivalent circuit for the filter in Fig. 46a. Though this equivalent circuit involves some approximations [17], it has been found to give very good accuracy in representing the performance of interdigital filters of the form shown in Fig. 46a. Two parts of this equivalent circuit are not evident in Ref. [17]. These are the lengths of transmission line of length $\theta$ and characteristic admittance $Y_A$ at each end of the filter. The exact equivalent circuit in Fig. 21 of Ref. [17] should have included such a length of line between the ideal transformer and the load conductance $Y_A$ at the left in the open-wire circuit, in order for the open-wire equivalent circuit to be exactly equivalent to the parallel-coupled-strip circuit in all respects. However, the discussion in Ref. [17] was only concerned with the input admittance seen looking in from the right, so the presence of this section of matched transmission line was of no importance for that analysis. But for the present situation, the presence of this length of line is important to the proper operation of the multiplexer channel unit. (Note that points $T$ and $D$ in Fig. 46a correspond to points $T$ and $D$ indicated in Fig. 46b.)

In order to use a filter of the form in Fig. 46a in a bandpass-plus-bandstop filter connection such as that in Fig. 44, it is necessary to break the ground connection on line 0 at point $B$ indicated in Fig. 46a. By study of various parallel-coupled-line configurations and their known, exact, open-wire equivalent circuits [26], it was determined that at least to an excellent approximation (and probably as an exact equivalence), breaking the circuit in Fig. 46a at point $B$ and inserting an added impedance between the end of line 0 and ground, has the effect of breaking the circuit in Fig. 46b at the indicated point $B$ and inserting the added impedance in series with the ideal transformer.

Making use of the equivalent circuits in Figs. 45a and 46b, along with the above observation, we see that the multiplexer channel-separating unit in Fig. 44 has the open-wire equivalent circuit shown in Fig. 47. Note that in this circuit the connection at point $B$ in Fig. 46b has been broken, and the input terminals of the bandstop filter circuit have been connected so as to reclose the circuit. Note that looking left from the left half of the bisected point $B$, a constant-resistance $Z_A$ will be seen, since the transmission line of characteristic impedance $Z_A$ is terminated in matching generator impedance $Z_A$. The ideal transformer at the input of the bandpass filter has an impedance-level-trans-
forming effect, but otherwise has no effect on the character of the input impedance to the bandpass filter. Thus, the impedance $Z_p$ associated with the bandpass filter is of the form typical of a bandpass filter that starts out with a shunt bandpass resonator. The bandstop filter starts out with a shunt bandstop resonator, which makes the two filters of the proper form for series interconnection. (Note that if the bandstop filter started out with a series bandstop resonator, it would present an open circuit at resonance, which would disrupt the performance of the bandpass filter.)

Since the bandstop filter is designed from a singly terminated, maximally flat, low pass prototype filter using the methods previously described, and since the bandpass filter is also designed using the same low pass prototype and for the same 3-db bandwidth as the bandstop filter, the input impedances $Z_p$ and $Z_s$ will be very nearly complementary. (If the mapping for the bandstop and bandpass filters were identical, the impedances $Z_p$ and $Z_s$ would be exactly...
complementary.) Thus, as has been previously mentioned, it is possible for the input impedance of multiplexer units of the form in Fig. 44 to have a very good impedance match with respect to the line impedance $Z_A$ at all frequencies.

3. A Trial Multiplexer Channel-Separating Unit in Strip Line

On the basis of the theory and design techniques just presented, a multiplexer channel-separating unit was constructed in strip line. The filters of the channel unit were based on a three-resonator, singly terminated, maximally flat low pass prototype. The design center frequency was 1.5 Gc, and the design fractional bandwidth was 0.05. The interdigital bandpass filter and the parallel-coupled-resonator bandstop filter were both constructed using round rods between parallel ground planes [22].

A drawing of the channel-separating unit, showing important dimensions of the filters, is given in Fig. 48, and Table XXV gives further data. A photograph of the constructed channel unit with its top ground plane removed is shown in Fig. 49.

The bandstop filter was tuned by individually resonating each of the bandstop filter resonators while the remaining two were detuned. Resonance was determined by tuning for minimum transmission from port 1 to port 3. Initial tuning of the bandpass filter was accomplished using the alternating short-circuit and open-circuit procedure [27, 28]. Fine tuning adjustments were made by observing the reflected wave of the channel unit using an electronically swept frequency source and reflectometer.

![Fig. 48. Drawing of the trial multiplexer channel-separating unit in strip line.](image)
Table XXV

**DIMENSIONING OF FILTERS IN CHANNEL-SEPARATING UNIT**

(a) Bandpass Filter Rod Diameters and Center-to-Center Spacing (in.)

<table>
<thead>
<tr>
<th>k</th>
<th>Diameter of rod k (in.)</th>
<th>Center-to-center spacing (c_{k,k+1}) (in.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.331</td>
<td>0.486</td>
</tr>
<tr>
<td>1</td>
<td>0.217</td>
<td>0.792</td>
</tr>
<tr>
<td>2</td>
<td>0.224</td>
<td>0.674</td>
</tr>
<tr>
<td>3</td>
<td>0.203</td>
<td>0.384</td>
</tr>
<tr>
<td>4</td>
<td>0.318</td>
<td></td>
</tr>
</tbody>
</table>

(b) Bandstop Filter Rod Diameters and Center-to-Center Spacing (in.)

<table>
<thead>
<tr>
<th>k</th>
<th>Diameter of rod (k_a) (in.)</th>
<th>Center-to-center spacing (c_{k,k}) of rod (k_a) and (k_b) (in.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.332</td>
<td>0.215, 0.663</td>
</tr>
<tr>
<td>2</td>
<td>0.336</td>
<td>0.219, 0.686</td>
</tr>
<tr>
<td>3</td>
<td>0.344</td>
<td>0.237, 0.846</td>
</tr>
</tbody>
</table>

* Ground plane spacing 0.625.

---

Fig. 49. Photograph of the trial multiplexer channel-separating unit (cover plate removed.)
The VSWR of the channel unit (at port 1) from 1 to 4 Gc is given in Fig. 50. An expanded frequency scale is also shown in this figure, giving the VSWR in the vicinity of the passband of the interdigital filter. It is seen to be better than 1.25 everywhere in the 1- to 4-Gc interval. The sinusoidal variation of VSWR seen in Fig. 50 is due primarily to the junction effect of the transitions associated with the type N connector at ports 1 and 3, although there is some contribution to the VSWR from the connectors themselves as well as the termination at port 3. The maximum magnitude of the sinusoidal variation of VSWR suggests that the VSWR of the transition from coaxial line to the round conductor between parallel ground planes is approximately 1.1 per transition. This could be reduced by further refining the transitions. However, the measured results adequately demonstrate the design principles that were to be shown.

The measured attenuations from port 1 to port 3 and from port 1 to port 2 in the vicinity of the passband of the bandpass filter are given in Fig. 51. The attenuation at the band-center frequency, 1.5 Gc, is 0.38 db. This value compares favorably with that given by a first-order formula for the attenuation of the bandpass filter of a channel unit at center frequency:

\[ (L_A)_0 \approx \frac{4.34 \omega_1}{w} \left[ \left( \sum_{j=1}^{n} \frac{g_i}{Q_{u_j}} \right) + \frac{1}{Q_{u_1} \hat{g}_1} \right] \text{ db} \]  

(36)
where \( w \) is the fractional bandwidth; \( \omega_1 \) is the cutoff frequency of the low pass prototype filter; \( Q_{ui} \) is the unloaded \( Q \) of the \( i \)th resonator of the bandpass filter; \( Q_{u1} \) is the unloaded \( Q \) of the first resonator of the band-stop filter;

\[
\begin{align*}
\text{Fig. 51. Measured attenuations of trial channel-separating unit in strip line.} & \quad \bullet \text{Measured values.} \\
g_i \text{ is the } i\text{th element value of the low pass prototype filter used for designing the bandpass filter; and } \hat{g}_1 \text{ is the } g_1 \text{ element value of the low pass prototype filter used for designing the bandstop filter. Using the same unloaded } Q \text{ of 1000 for each of the resonators and Eq. (36) gives a calculated value of 0.347 db for } L_{A0}. 
\end{align*}
\]
From Fig. 51, it is seen that the crossover frequencies occur virtually at the 3-db points, and that the 3-db fractional bandwidth is

$$2 \left( \frac{1.536 - 1.463}{1.536 + 1.463} \right) = 0.0488$$

which is in good agreement with the design value of 0.05.

VI. THEORY OF MULTIPLEXERS HAVING CONTIGUOUS CHANNELS

A. INTRODUCTION

Multiplexers having contiguous channels typically have attenuation characteristics that cross over at the 3-db points; there is no frequency guard band between adjacent channels. The theory that will be presented for the design of such multiplexers comes under the category of partly complementary multiplexers, since the theory closely follows the ideas discussed in the section on partly complementary diplexers.

Figure 52a shows a parallel-connected multiplexer of three channels. Additional channels are added by adding additional bandpass filters in parallel. Note that the filters are minimum susceptance for the parallel-connected multiplexer. Figure 52b shows the dual of Fig. 52a, a series connected multiplexer. For this configuration, additional channels are added by connecting additional bandpass filters in series. Also note that the filters are minimum reactance for the series connected multiplexer. A shunt, susceptance-annulling network is required for the parallel connected multiplexer and a series, reactance-annulling network is required for the series connected multiplexer. This situation is analogous to the theory of parallel and series connected partly complementary diplexers previously presented.

From here on we shall focus our attention on the parallel connected multiplexer, but the same reasoning applies to the series connected multiplexer on a dual basis.

Consider the input admittance of each of the three bandpass filters comprising channels A, B, and C of the multiplexer of Fig. 52a; since each filter is minimum susceptance, the input admittance of each will take the general form shown in Fig. 5b.

Let us concentrate our attention on the real part of the input admittance of the bandpass filters. Since the channels of the multiplexers are to be contiguous, with 3-db crossover points, the real part of the input admittance of adjacent channels must cross over at approximately half of their maximum values. This is shown in Fig. 53a. The real part of the multiplexer input admittances, $\text{Re } Y_T$, is the sum of the real parts of the admittances of the bandpass filters. This sum is given in Fig. 53b. Notice that as a result of the multiplexers having
contiguous channels, $\text{Re} Y_T$ also has the general form of the real part of the input admittance of a bandpass filter. Since $Y_T$ is minimum susceptance, and its real part has the form shown in Fig. 53b, its imaginary part must be of the general form shown in Fig. 5b. Therefore, the input-admittance $Y_T$ appears as shown in Fig. 54a. (Of course, this conclusion can also be reached by superimposing both the real and imaginary parts of the admittances of each filter.) Since the imaginary component of $Y_T$ has, on the average, a negative slope.
throughout the operating band of the multiplexer, it can be largely canceled by a susceptance-annulling network. This is suggested by the curve of susceptance shown in Fig. 54b.

![Graph showing real part of admittance](image)

**Fig. 53.** (a) Typical real part of the input admittance of the bandpass filters of Fig. 52a. (b) Their sum.

With the addition of the susceptance of the susceptance-annulling network to the admittance $Y_T$, the generator is approximately matched to the multiplexer throughout its operating band.

**B. Example Design and Measured Response of a Parallel Connected Three-Channel Multiplexer**

The design of the three-channel multiplexer, which will be described, was carried out before the design viewpoints in Section III were fully developed.

---

1 This work is reported on in greater detail in Refs. [30, 31]. The presentation here is taken from Ref. [32], with the permission of the IEEE.
This example is based on using singly terminated prototype filters. The experimental results show that both the design technique and the use of the singly terminated prototype give satisfactory results. However, after the completion of the work on this three-channel multiplexer, subsequent refinements in the general theory have led the authors to the conclusion that a more satisfactory response can be obtained if the design of the multiplexer is based on the doubly terminated prototype filter, with the first reactive element of the prototype removed. The reason for this is that a predictable Chebyshev transmission and VSWR response can be obtained by the latter technique, but not with the former. (A discussion of this point was given previously in Section III,B.) Nevertheless, the present design, based on the singly terminated prototype filter, is given here for illustrative purposes, as well as to show one possible microwave realization of the theory.

It was pointed out in Section I [Eq. (2)] that the singly terminated filter having a minimum susceptance input admittance has the property that the
real part of its input admittance has the same characteristic as its transmission attenuation response. Thus, a singly terminated low pass prototype filter that is designed to have a Chebyshev transmission response has the typical characteristic shown in Fig. 55 for its real part. Notice that in Fig. 55 a low pass

![Diagram](image)

**Fig. 55.** Re $Y_1$ response of a low pass prototype filter and definition of the quantity $g_{n+1}''$ with $L_{A\nu} =$ Chebyshev attenuation ripple of singly terminated prototype in db. (a) Case of $n$ even; (b) $n$ odd.

prototype filter parameter $g_{n+1}''$ is defined. This parameter will be referred to in the discussion of the details of multiplexer design. However, it should be noted at this point that either $g_{n+1}''$ or its reciprocal corresponds to the geometric mean between the value of Re $Y_k'$ at the top of the ripples and Re $Y_k'$ at the bottom of the ripples. This admittance level for the low pass prototype is analogous to the driving source admittance $G_g$ in Fig. 52. That is, the multiplexer described herein is designed to have a Chebyshev real part characteristic with $G_g$ equal to the mean value of the ripples.
The multiplexer design techniques described in this section are applicable using many different types of filter structures. To illustrate the design theory of this section it was decided to use comb-line filters [33] for the bandpass filters of the multiplexer.

Figure 56 shows a comb-line bandpass filter in strip line form. The resonators in this type of filter consist of TEM-mode transmission line elements that are short circuited at one end and have a lumped capacitance $C_k$ between the other end of each resonator line element and ground. In Fig. 56, lines 1 to $n$ and their associated capacitances, $C_1^s$ to $C_n^s$, constitute resonators, while lines 0 and $n+1$ are not resonators, but simply part of impedance-transforming sections at the ends of the filter. Coupling between resonators is achieved in this type of filter by way of fringing fields between resonator lines. With the capacitors $C_k^s$ present, the resonator lines will be less than $\lambda_0/4$ long at resonance (where $\lambda_0$ is the wavelength in the medium of propagation at midband), and the coupling between resonators is predominantly magnetic. Interestingly enough, if the capacitors $C_k^s$ were not present, the resonator lines would be a full $\lambda_0/4$ long at resonance, and the structure would have no passband [34]. Without some kind of reactive loading at the ends of the resonator line elements, the magnetic and electric coupling effects would cancel each other, and the comb-line structure would become an all-stop structure.

For reasons described above, it is usually desirable to make the capacitances $C_k^s$ in this type of filter sufficiently large so that the resonator lines will be $\lambda_0/8$ or less long at resonance. Besides permitting efficient coupling between resonators (with sizable spacings between adjacent resonator lines), the resulting filter will be relatively small. In this type of filter, the second passband occurs when the resonator line elements are somewhat over one-half wavelength long, so that, if the resonator lines are $\lambda_0/8$ long at the primary passband,
the second passband will have its center at slightly over four times the frequency of the center of the first passband. If the resonator line elements are made to be less than \( \lambda_0/8 \) long at the primary passband, the second passband will be even further removed from the primary passband.

To adapt the comb-line filter design [33] for use in multiplexers, it was necessary to modify the input end of the filter in order to provide for coupling several filters to a common input junction. The modification used is shown schematically in Fig. 57. The comb-line filter in this figure is seen to include a fine-wire, high-impedance line at the input end.

Figure 58 shows a high-impedance wire and its equivalent circuit when used as an admittance inverter \( J_{n,n+1} \). (An ideal admittance inverter is defined as a device which operates like a quarter-wavelength line of characteristic admittance \( J_{n,n+1} \) mho at all frequencies [33].)\(^1\) The comb-line filter design equations of Ref. [33], which utilize the admittance inverter concept, were modified to use this high-impedance wire type of inverter at one end of the filter. The susceptance at one side of the inverter in Fig. 58 was compensated for by absorbing it into resonator \( n \), while the susceptance \( B \) on the other side of the inverter was effectively absorbed into the susceptance-annulling network. In this case, the high-impedance wire type of coupling was suggested because it would help prevent the common junction from being crowded. However, it is desirable to keep the high-impedance wire short in order to avoid unwanted resonances, while not having it so small in diameter that it would increase the losses. The comb-line filter equations of Ref. [33] were modified for designing comb-line filters for use in multiplexers and are presented in Table XXVI.

\(^1\) Cohn [35] describes impedance inverters. An admittance inverter is an admittance representation of an impedance inverter.
Table XXVI is typically applied in the following way:

(1) Choose a lumped-element, singly terminated Chebyshev low pass filter prototype (or preferably a doubly terminated prototype foreshortened as discussed in connection with Fig. 8) of specified ripple and number of reactive elements. The normalized values of the prototype filter's elements, 

$$g_0, g_1, g_2, \ldots, g_n,$$

are given in Refs. [6, 7].

(2) Choose the electrical length, $\theta_0$, of the resonator bars at resonance, and the electrical length of the coupling wire, $\theta_c$. It is essential to the proper performance of the multiplexer that the sum of these two electrical lengths be less than $\pi$ for all frequencies in the multiplexer operating band.

(3) Choose the multiplexer junction and channel output terminating admittances $Y_B = 1/R_s$ and $Y_A = G_L$ respectively, and choose the resonator normalized slope parameters, $b_l/Y_A$ [33]. The normalized slope parameters are chosen generally to realize a maximum $Q$. To the best knowledge of the authors, a slope parameter which realizes an equivalent line impedance of 70 $\Omega$ for the resonator is nearly optimum. The last quantity to be chosen is the number $g_{n+1}^*$. This quantity is defined in Fig. 55. Ideally it represents the mean value of $\text{Re} Y_T$ in the operating band and is a prototype parameter corresponding to the multiplexer termination $Y_B = 1/R_s$ at its common junction.

(4) The self- and mutual capacitances per unit length and the normalized admittance of the high-impedance wire are then given by the formulas of Table XXVI.

(5) The physical dimensions of the resonator and their relative spacings may then be determined from Ref. [36] or [22].

Using the equations in Table XXVI for the modified comb-line filter, a four-resonator, 10% bandwidth filter design was worked out using an $n = 4$, singly terminated, Chebyshev, prototype filter having 1-db ripple. The filter design was expressed analytically in an approximate form and its input admittance $Y_k$ was computed using a digital computer. The results are shown
Table XXVI

**Design Equations for Comb-Line Filters for Multiplexer Applications**

General:

\[ Y_a = \sqrt{\varepsilon_r} \left( \frac{C_0}{376.7} \right) \]

Normalized mutual capacitances per unit length:

\[ \frac{C_{01}}{\varepsilon} = Y_a \frac{376.7}{\sqrt{\varepsilon_r}} \left[ \frac{w (b_1/Y_A)}{\omega_l g_0 b_1} \right]^{1/2} \]

\[ \frac{C_{i,i+1}}{\varepsilon} = Y_a \frac{376.7}{\sqrt{\varepsilon_r}} \left[ \frac{(b_i/Y_A)(b_{i+1}/Y_A)}{\delta_i \delta_{i+1}} \right]^{1/2} \tan \theta_0 \quad (i = 1, 2, \ldots, n-1) \]

Normalized admittance of high-\( Z_0 \) coupling wire:

\[ \frac{Y_s}{Y_A} = \left[ \frac{w (b_n/Y_A)(Y_B/Y_A)}{\omega_l \delta n \delta n+1} \right]^{1/2} \sin \theta_s \]

Normalized self-capacitances per unit length:

\[ \frac{C_0}{\varepsilon} = \frac{376.7}{\sqrt{\varepsilon_r}} \left( Y_A - \frac{C_{01} \sqrt{\varepsilon_r}}{376.7} \right) \]

\[ \frac{C_1}{\varepsilon} = Y_A \left[ \frac{376.7 2b_1}{\sqrt{\varepsilon_r}} Y_A \frac{\csc \theta_0 \theta_0 \csc \theta_0}{\csc \theta_0 \theta_0 \csc \theta_0} \right] - \left( \frac{(C_{12}/\varepsilon_r)(C_{01}/\varepsilon_r)}{376.7} \right) \]

\[ \frac{(C_i)}{\varepsilon} = Y_A \frac{376.7}{\sqrt{\varepsilon_r}} \left[ 2 \frac{b_i}{Y_A} \frac{\csc \theta_0 \theta_0 \csc \theta_0}{\csc \theta_0 \theta_0 \csc \theta_0} \right] \]

\[ \frac{\varepsilon}{\delta_{n-1,n}} = - \frac{C_{n-1,n}}{\varepsilon} + \frac{376.7 Y_A}{\sqrt{\varepsilon_r}} \left[ \frac{2(b_n/Y_A) - (Y_B/Y_A) \left[ \cot \theta_s + \theta_s \csc \theta_s \right]}{\cot \theta_0 + \theta_0 \csc \theta_0} \right] \]

Normalized lumped loading susceptance:

\[ \frac{\omega_0 C_i}{\varepsilon Y_A} \]

\[ = \left[ \frac{b_i}{Y_A} \right] \left[ \cot \theta_0 + \theta_0 \csc \theta_0 \right] \cot \theta_0 \quad (i = 1, 2, \ldots, n-1) \]

\[ \frac{\omega_0 C_n}{\varepsilon Y_A} \]

\[ = \frac{Y_s}{Y_A} \cot \theta_s + \frac{2(b_n/Y_A) - (Y_B/Y_A) \left[ \cot \theta_s + \theta_s \csc \theta_s \right]}{\cot \theta_0 + \theta_0 \csc \theta_0} \cot \theta_0 \]

in Fig. 59, normalized with respect to \( G_A \). Note that \( \text{Re} \ Y_k/G_A \) is very nearly perfectly Chebyshev, while in the passband, the slope of \( \text{Im} \ Y_k/G_A \) is negative on the average.

Next, two additional bandpass filters with input admittance characteristics such as that in Fig. 59 were designed to cover contiguous bands. Their center
Fig. 59. Normalized $Y_k$ of a modified comb-line filter. (The filter corresponds to a 1.0-db Chebyshev, low pass filter.)

frequencies were determined so that the Re $Y_k/G_A$ of adjacent filters overlapped at approximately their Re $Y_k/G_A = 0.5$ points.\(^1\)

The input admittance of the multiplexers, $Y_T$, was then calculated on the basis $Y_T = Y_1 + Y_2 + Y_3$. The result is shown in Figs. 60 and 61. Note that on the average, the slope of Im $Y_T/G_A$ is negative throughout the operating band of the multiplexer. This susceptance can be largely cancelled by adding the susceptance of a shunt susceptance-annulling network, as has been described in the theory.

\(^1\) By Eqs. (1) and (2) it can be shown that Re $Y_k/G_A$ characteristic for a singly terminated Chebyshev filter can be computed using standard Chebyshev transfer functions. Equations for doing this are given in Ref. [30].
In this case, a susceptance-annulling branch could consist of a short-circuited stub of such a length as to give resonance at the normalized frequency \( \omega/\omega_0 = 1.02 \), where the \( \text{Im} Y_T/G_A \) curve in Fig. 61 is approximately zero.

![Fig. 60. Computed normalized Re \( Y_T \) of three-channel multiplexer.](image1)

![Fig. 61. Computed normalized Im \( Y_T \) of three-channel multiplexer.](image2)

Estimates indicate that, if the stub had a normalized characteristic admittance of \( Y_s/G_A = 3.85 \), the susceptance slope of the annulling network should be about right. Figure 62 shows the normalized susceptance after the susceptance-annulling network has been added. Note that although the susceptance has
not been completely eliminated, it has been greatly reduced. (Note the change of scale in Fig. 62.)

The driving generator conductance $G_B$ for the trial multiplexer design was given a normalized value of $G_B/G_A = 1.15$, which makes the generator con-
ductance equal to the mean value of the ripples in Fig. 60. Figure 63 shows the computed response of the multiplexer, while Fig. 64 shows the details of the passband response in enlarged scale. Note that the attenuation characteristics cross over at about the 3-db points, and that, although the filters were designed to have 1-db Chebyshev ripple when driven by a zero-impedance generator, the passband attenuation is much less than that in the completed multiplexer. (This is the reason that foreshortened doubly terminated prototype filters are recommended for use in future designs. Their use will facilitate obtaining a prescribed Chebyshev response.) This is largely due to the fact that adding the generator internal conductance $G_B$ tends to mask out the variations in the real part of the input admittance. Also, choosing $G_B$ to be equal to the mean value of the real part of the input admittance in the operating band tends to reduce the amount of mismatch that will occur.

Using the design theory and equations previously discussed and referenced, a microwave multiplexer based on the preceding trial design was constructed. A sketch of the multiplexer is given in Fig. 65. The view in Fig. 65 is that seen when looking at the top of the multiplexer. The susceptance-annulling network of the multiplexer cannot be seen in this particular view. It consists of a specially constructed, low impedance, coaxial line that is connected to the common junction and extends through the lower ground plane opposite the common input of the triplexer. The electrical length of the coaxial line was varied by

![Figure 64. Magnified scale version of Fig. 63.](image-url)
sliding a short-circuit block.\(^1\) Details of the annulling stub are shown in Figs. 66 and 67.

The multiplexer attenuation, after final tuning and adjusting of the annulling network for an optimum response, is given in Fig. 68. Figure 69 gives the multiplexer attenuation in the passband using an expanded ordinate. The attenuation in the passband of the lower- and middle-frequency channels is typically 0.5 db. The attenuation in the passband of the highest frequency channel is approximately 0.4 db. The VSWR of the multiplexer (measured at the common input) is shown in Fig. 70 and is less than 1.6 throughout the passband except at the multiplexer band edges.

\(^1\) For a more detailed description of the construction of the triplexer, see Ref. [31].
The 3-db percentage bandwidths of each channel were determined. These were defined as the difference of the frequencies at which the attenuation is 3 db, divided by the arithmetic mean of those frequencies. The bandwidths were found to be 17, 16, and 13% for the lowest, middle, and highest frequency channels, respectively. These percentages are larger than the corresponding calculated values for the triplexer example previously discussed, whose channel 3-db percentage bandwidths were 11% each. The increase in bandwidth in the trial design over the design value is denoted as "bandwidth spreading." It has been noticed in comb-line filters before [33]. However, in the case of a
Fig. 67. Detailed drawing of the annulling stub outer conductor.
Fig. 68. Measured attenuation of the three-channel comb-line multiplexer.

Fig. 69. Measured attenuation in the passband of the three-channel multiplexer.
previous experimental comb-line filter having a design bandwidth of 10% [33], the bandwidth spreading was less than in the present cases. Bandwidth spreading is believed to be due to coupling effects beyond nearest neighbor line elements, which were neglected in the derivation of the design equations for comb-line filters. It can be compensated for by increasing the spacings between the resonators slightly.

![Graph](image)

**Fig. 70.** Measured VSWR in the passband of the three-channel multiplexer. • Measured values.

**C. DESIGN OF MULTIPLEXER FILTERS AND ANNULLING NETWORKS FROM DOUBLY TERMINATED LOW PASS PROTOTYPES**

As has been previously emphasized, it would have been preferable in the preceding triplexer example to have used doubly terminated low pass filter prototypes rather than singly terminated prototypes. This would have helped with regard to obtaining bandpass filter responses with Chebyshev ripples more nearly as prescribed.

In order to clarify the use of doubly terminated prototypes for multiplexers, let us suppose that essentially the same multiplexer design is desired as in the previous example, but that the filters are to be designed to have \( n = 4 \) resonators with 0.1-db Chebyshev ripple. To do this a 0.1-db ripple, low pass prototype with \( m = 5 \) reactive elements is used, and then the prototype filter is fore-shortened by removing one reactive element. This process was illustrated in
Fig. 8, though using the numbering system as for Fig. 57, element \( g_{m}=g_{5} \) should be removed, rather than \( g_{1} \). Using the prototype element values \( g_{0} \) to \( g_{4} \) obtained in this way, the design of the bandpass filters is the same as for the preceding example, except for the fact that the information in Fig. 55 is not needed using this procedure. The element value \( g_{n+1}^{*} \) in Fig. 55 and in the fourth equation of Table XXVI would in this case be

\[
g_{n+1}^{*} = \frac{1}{g_{m+1}}
\]  

where \( n=m-1 \) and \( g_{m+1} \) is the resistive termination element \( g_{6} \) from the original \( m=5 \) reactive element, unforeshortened, doubly terminated prototype. Using \( g_{0}, g_{1}, g_{2}, g_{3}, g_{4}, \) and \( g_{5}^{*} \) obtained in this way, the design process would be the same as described in the preceding example.

In the preceding example, the susceptance-annulling network design was determined after first computing the input admittance characteristic for the triplexer. Using doubly terminated prototypes, it should be relatively simple to obtain a useful approximate design for the annulling network. The basis for the approximation is that the Re \( Y_{P} \) characteristic for a triplexer having three four-resonator filters is similar to the real part of the input admittance for a three \( 3 \times 4 (=12) \)-resonator foreshortened filter having a bandwidth equal to the total operating band of the triplexer. (Compare Figs. 5b and 54a.)

If we have a foreshortened bandpass filter with an input admittance characteristic similar to that in Fig. 5b, the sizable passband susceptance present can be canceled out by simply replacing the resonator that was removed (or omitted) when designing the foreshortened filter. Let us suppose that the resonator to be added is to be designed as a short-circuited stub, which is a quarter-wavelength long at midband, and that the element values for the unforeshortened doubly loaded low pass prototype are \( g_{0}, g_{1}, \ldots, g_{m}, g_{m+1} \), where \( g_{m} \) is the element omitted in the foreshortening process. It is easily shown that in order for the stub to have a susceptance at the band edges corresponding to \( B=\omega'_{1} g_{m} \) for the prototype

\[
\frac{Y_{s}}{G_{B}} = \tan \theta_{1} \omega'_{1} g_{m} g_{m+1}
\]

where

\[
\theta_{1} = \frac{\pi \omega_{1}}{\omega_{2} + \omega_{1}}
\]

is the electrical length of the stub at the lower band-edge frequency \( \omega_{1} \), and \( \omega_{2} \) is the upper band-edge frequency. \( Y_{s} \) is the stub admittance and \( G_{B} \) is the conductance of the adjacent termination.

For the triplexer example under consideration, using three four-resonator filters, the total input admittance would be somewhat like that of a 12-resonator foreshortened filter designed from an \( m=13 \) reactive element prototype. For a 0.1-db ripple prototype with 13 reactive elements, \( g_{13}=1.2074 \) if the adjacent termination is \( g_{14}=1 \), and \( \omega'_{1}=1 \). From the plot in Fig. 60, \( \omega'_{1}/\omega_{0} \) is about 0.95, while \( \omega_{2}/\omega_{0} \) is about 1.3. From Eqs. (38) and (39), the normalized
stub admittance is computed to be $Y_s/G_B=4.76$, as compared to the value of 3.85, which was chosen in the computed design example, after studying Fig. 61. However, it should be recalled that the use of a 0.1-db ripple prototype does not correspond exactly to the case of Fig. 61, and the susceptance characteristic in Fig. 61 includes the effect of the susceptance $jB$ on the right end of the inverter network in Fig. 58. (The coupling wire on each filter contributes such a component.) The effect of these coupling wire susceptances accounts for the skewedness of the susceptance characteristic in Fig. 61. To correct for this effect the stub length $l$ and admittance $Y_s$ should be chosen so that

$$-Y_s \cot (\omega_k/v) + B_1(\omega_k) + B_2(\omega_k) + B_3(\omega_k) = (-1)^k G_B w_1 g_m g_{m+1}$$

(40)

where $k=1$ is for the lower band-edge frequency $\omega_1$, $k=2$ is for the upper band edge $\omega_2$, and $v$ is the velocity of propagation. The $B_1(\omega_k)$, $B_2(\omega_k)$, and $B_3(\omega_k)$ are the coupling-wire susceptances (as in Fig. 58b) for the three channels, evaluated at $\omega_1$ and $\omega_2$. By simultaneous solution using the equations for $k=1$ and $k=2$, appropriate values for $Y_s$ and $l$ can be obtained.

ACKNOWLEDGMENTS

W. J. Getsinger and D. B. Weller worked out several of the trial designs discussed in this chapter. The laboratory testing of the trial designs was performed by M. Di Domenico, Y. Sato, P. R. Reznik, and R. B. Larrick, while the trial devices were fabricated by R. Pierce and C. A. Knight. The computer programs for computing the responses of trial designs were prepared by P. H. Omlor, V. H. Sagherian, and J. Herndon.

REFERENCES


I. INTRODUCTION

This chapter is concerned with the numerical solution of two broad classes of problems that arise in the use of TEM-mode transmission lines. The first is the determination of the basic constants of these lines, and the second the determination of the elements of the equivalent circuits which may be used to represent certain simple obstacles that interrupt their longitudinal uniformity. Both are united in the common requirement of initially requiring the solution of Laplace's equation, namely,

\[ \nabla^2 V = 0 \]  

(1)
within some region of interest subject to boundary conditions imposed by that region.

Laplace's equation, or related forms, appears frequently in many branches of physics and there has been great effort for nearly two centuries to find methods of solution. In a few cases where the boundaries can be represented simply in an appropriate coordinate system or where particularly fortunate symmetry is in evidence in the problem a direct mathematical solution is possible.

More commonly other analytic means of less direct application must be used. It may, for example, be possible to conformally transform the given boundary by one or more successive transformations into some other shape more amenable to direct solution [1]. This is based on the well-known property that a solution of Laplace's equation remains a solution under conformal transformation. While this can be a powerful mathematical tool it is limited to two-dimensional problems and, not uncommonly, its application is frustrated by an inability to integrate the transformation. This can sometimes be avoided by the introduction of approximations which are sufficiently ingenious to allow a solution with only small error.

In yet other cases, where bounds on the answer rather than an explicit solution are acceptable, a successful assault on the problem can be made by the application of variational techniques [1]. These allow a first-order description of the field within the region to yield answers for the constants sought correct to second order. It is possible by this technique to obtain upper and lower bounds on the solution which, if the initial assumptions are sufficiently detailed, may be close and so give considerable confidence in the results. Unfortunately the analytic effort required greatly increases as the precision of the initial approximations is improved.

Nevertheless despite all these analytic efforts there still remain many problems of engineering importance which are at best only approximately solved with an undetermined accuracy or which have no solutions at all. This is particularly true where the dielectric properties of the region within the conducting boundaries are not homogeneous, giving rise to additional boundary conditions. An essential purpose of this chapter is to show how these problems may be solved very economically by numerical means.

II. NUMERICAL SOLUTION OF LAPLACE'S EQUATION

A. GENERAL REMARKS

Interest in the numerical solution of partial differential equations is not of recent origin. None of the numerical methods contained in this monograph are new although some of the results obtained with them are believed to be. Some parts of the theory were formalized at least 60 years ago when the practical
importance of radio was rather slight. The more modern developments seem to have been stimulated by work in the atomic energy and civil engineering [2] fields where many of the problems are analytically intractable but the capital investment so high that a solution is necessary.

New developments in TEM-mode transmission lines have presented the microwave engineer with problems which are also not easily solved but, judging from the literature, the predominant tendency appears to have been to attempt to twist the older techniques of conformal transformation and variational solution to meet the new demands. There does not appear to have been much interest in purely numerical techniques in which a specific underlying theory peculiar to the problem is not even sought, although some small amount of work relating to cavity resonators in microwave tubes has been published [3]. The real contribution of this work is thought to be to draw attention to the use of these techniques and to indicate how, with the aid of a digital computer, a program of considerable generality may be written to solve not one specific problem but whole classes of problems [4].

The basic handicap of numerical methods is the large amount of arithmetic frequently involved to obtain an accurate solution to even simple problems. Some of the methods involved which have been designed for hand calculation have made use of the intuitive abilities of the human computer [5] in speeding a solution but it still remains true that really significant successes have only been obtained in the past decade or so since the development of the modern high-speed digital computing machine. This has led to great interest in numerical methods which have been lifted from an often rather theoretical plane to a thoroughly practical level and there is now a large and increasing volume of literature relating to these problems.

Laplace’s equation is the simplest example of an elliptic partial differential equation. A common numerical method for solving this class of equations relies upon the use of finite difference techniques which aim to approximate the original problem by a set of linear algebraic equations sufficient in number to give an adequately accurate representation of the original problem. The main emphasis in the solution is then directed away from purely partial differential equation theory toward finding means for the economic solution of large arrays of simultaneous linear equations, often numbered in thousands when highly accurate results are sought.

B. BASIC FINITE DIFFERENCE THEORY

Although a more general approach would be possible it must be made clear from the outset that this chapter will be concerned only with two-dimensional problems which can be represented in either Cartesian coordinates or cylindrical coordinates subject to the special but important restriction of rotational symmetry (no variation with \(\phi\)). In each case, while the dielectric need not be
homogeneous throughout the region in which the solution of Laplace's equation is sought, the restriction of piecewise homogeneity will be imposed, that is, the dielectric is supposed to be made up by the juxtaposition of discrete pieces of homogeneous dielectrics. The practical utility of the work suffers little if it is further demanded that no more than two dielectrics are present,

one being free space. Thus, unless the context clearly demands otherwise, the word "dielectric" will always be taken to mean that part of the region which contains nonconducting material substance.

For Cartesian problems Eq. (1) assumes the special form

\[
\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0
\]

(2)

while for cylindrical problems the equation to be solved is

\[
\frac{\partial^2 V}{\partial r^2} + \frac{1}{r} \frac{\partial V}{\partial r} + \frac{\partial^2 V}{\partial z^2} = 0
\]

(3)
The numerical solutions of both these equations have so much in common that it will suffice to develop the theory in detail for the Cartesian case only. Points of difference in the solution of cylindrical problems can then be illustrated without needless duplication.

Consider the two-dimensional problem shown in Fig. 1. For the present, some generality may be sacrificed without essential loss to the outline of the theory by considering the medium between the conductors to be homogeneous. Details of the treatment of inhomogeneous media require only a simple extension of the theory and will be relegated to a later stage. Imagine the region between the conductors to be divided into equal rectangles by a net of interlaced rows and columns. Although the conductors are shown falling along these rows and columns this is not an essential element in the theory. However, for the purpose of organizing the problem on a computing machine it is much easier if this is the case; this restriction has been voluntarily imposed throughout this work. In practice the inconvenience is slight since the machinery used to generate most transmission line components tends to produce surfaces which can be so represented.

Consider a typical nodal point \( P \) in the medium between the conductors. The potential at \( P \) must satisfy Eq. (2). For numerical computation, this equation can be reduced by considering it in terms of the potential at \( P \) and its four immediate neighbors distant by mesh widths \( a \) and \( b \) in the \( X \)- and \( Y \)-coordinate directions, respectively. Then, applying Taylor’s theorem in the \( X \) direction,

\[
V_B - V_P = \frac{\partial V}{\partial x} + \frac{a^2 \partial^2 V}{2! \partial x^2} + \frac{a^3 \partial^3 V}{3! \partial x^3} + \frac{a^4 \partial^4 V}{4! \partial x^4} + \ldots
\]

\[
V_A - V_P = -\frac{\partial V}{\partial x} + \frac{a^2 \partial^2 V}{2! \partial x^2} - \frac{a^3 \partial^3 V}{3! \partial x^3} + \frac{a^4 \partial^4 V}{4! \partial x^4} - \ldots
\]

From this it follows that

\[
\frac{\partial^2 V}{\partial x^2} = \frac{V_A + V_B - 2V_P}{a^2} - \frac{a^2 \partial^4 V}{12 \partial x^4} - \ldots
\]

\[
\approx \frac{V_A + V_B - 2V_P}{a^2} \tag{4}
\]

that is, an approximation has been obtained to the second derivative with respect to \( x \) at \( P \), accurate within the order \( a^2 \), and which can be arbitrarily refined by decreasing the mesh width.

The same argument may be applied in the \( Y \) direction. Equation (2) at \( P \) therefore has the equivalent form

\[
V_A + V_B + h^2 V_C + h^2 V_D - 2V_P(1 + h^2) = 0 \tag{5}
\]
where $h = a/b$. In the usual case of a square net ($a = b$) this reduces to the particularly simple form

$$V_A + V_B + V_C + V_D - 4V_P = 0 \quad (6)$$

and this treatment may be extended to every other point in the medium. For nodes on the conducting boundaries by definition

$$V_P = V_0 \quad (7)$$

where $V_0$ is the assigned boundary potential. The original problem has therefore been replaced by one which gives an approximate representation in terms of simultaneous linear equations, that is, a complex analytic problem has been reduced to solving simultaneous linear equations.

Nearly all problems treated in this chapter have been computed on square nets, although there have been a few specialized exceptions (Section III,A,2). Thus in what follows mainly square nets will be considered. In every case, however, uniform rectangular nets, that is, nets in which $h$ stays constant throughout the cross section, have been used. This is somewhat inefficient in regions where the field is of small magnitude, such as in reentrant corners, but has considerable advantages, particularly in the improvement of the numerical solutions by use of Richardson’s method of the “deferred approach to the limit,” which is discussed later (Section II,G).

Equation (6) is a statement of what is commonly known as the “five point operator” and states, in effect, that in a square net the potential at $P$ is the average of the potentials of the nodes surrounding it. The finite difference solution can thus exhibit neither maxima nor minima within the potential field, which is also true of Laplace’s equation itself. It is the simplest of all the various finite difference operators that may be used, but is preferred to the more elaborate forms (such as the nine point operator) as the array of linear finite difference equations which results always possesses or can be reordered to possess Young’s property A [6]. This provides a sound theoretical basis for the convergence of the iterative method of solution which is used.

It is perhaps worth noting in passing that Eq. (6) also occurs in the description of other physical systems. Thus, for example, it is a statement of Kirchhoff’s first circuitual law for a node in a mesh of equal resistors meeting four at a node. It is also true for small lateral displacements in a grid of pin jointed elastic bars. These properties have been used in devising analog computers for the solution of Laplace’s equation [7, 8].

C. Development in Cylindrical Coordinates

It is now appropriate to give a brief outline of the development of a finite difference equation to represent Laplace’s equation in cylindrical coordinates for a general interior nodal point, subject to the restriction of rotational
symmetry. It will be assumed that the medium between the conducting surfaces is homogeneous and that these conducting surfaces can be drawn in by joining lines of nodes parallel to the coordinate axes, that is, the same assumptions that underlie Eqs. (5) and (6).

Consider the point $P$ shown in Fig. 2, a node in a square net of mesh width $a$. Owing to symmetry about the axis only one-half the problem need be treated. The convention will therefore be adopted that the first row in the net lies along the axis of symmetry. Let $P$ lie in the $N$th row (and for the present assume $N > 1$), that is, on a radius $r = (N - 1)a$. The potential at $P$ must satisfy Eq. (3)

\[
\begin{align*}
V_A + V_B + (1 - a/2r) V_C + (1 + a/2r) V_D - 4V_P &= 0 \\
\end{align*}
\]  

which, unlike its Cartesian counterpart, Eq. (2), involves derivatives of the first order.

Nonetheless, a Taylor expansion closely similar to that used in the previous subsection serves to derive approximations to these derivatives in terms of the potential differences between $P$ and its four immediate neighbors. Thus

\[
\begin{align*}
V_A + V_B + (1 - a/2r) V_C + (1 + a/2r) V_D - 4V_P &= 0 \\
\end{align*}
\]  

is an approximation to the Laplace equation at $P$ with a dominant error term in the order $a^2$. Equation (8) may be further simplified by substituting for the radius in terms of the position of $P$ in the net; thus

\[
\begin{align*}
V_A + V_B + \left( \frac{2N-3}{2N-2} \right) V_C + \left( \frac{2N-1}{2N-2} \right) V_D - 4V_P &= 0 \\
\end{align*}
\]  

Special interest attaches to nodal points in row 1 which are not on a conducting surface. As the radius is then zero and since, by symmetry, all odd order
derivatives must be zero, the middle term in Eq. (3) assumes an indeterminate form in this case. This difficulty may be surmounted by reverting to the Cartesian form of the Laplace equation, for which purpose a \( w \) axis is introduced to form the third of an orthogonal set with the \( r \) and \( z \) axes (Fig. 3). The point \( P \) must then satisfy

\[
\frac{\partial^2 V}{\partial r^2} + \frac{\partial^2 V}{\partial z^2} + \frac{\partial^2 V}{\partial w^2} = 0
\]  

This is not of course peculiar to the axis of symmetry; any point in the space between the conductors must satisfy Eq. (10). It has not been generally applied,

since on all but the axis of symmetry it leads to a three-dimensional net. In this one particular case, by symmetry

\[
V_C = V_D = V_E = V_F
\]

and therefore expanded in finite difference form Eq. (10) leads to the two-dimensional form

\[
V_A + V_B + 4V_D - 6V_P = 0
\]  

D. Special Finite Difference Equations

So far, for either coordinate system, the finite difference representation has been constructed only for interior points in continuous dielectric media. The intention to permit stepwise continuous media introduces the possibility of dielectric boundaries and it remains to show how these may be accommodated.
As with conductors the restriction that these boundaries must lie along lines of nodes in the overlying finite difference net will again be made.

In addition, quite apart from any dielectric variations, many problems exhibit surfaces of symmetry across which the normal component of potential gradient is either identically zero or can be assumed to be zero. To limit the number of finite difference equations it is always desirable to make full use of any symmetry possessed by a given problem. It can thus happen that the finite difference net terminates in one or more "open circuit" edges, the nodes along which need special finite difference equations for their representation. An example is shown in Fig. 4.

![Fig. 4. Typical problem simplified by use of open circuit edge.](image)

To take account of all the special cases or "oddities" which may occur within the restrictions imposed on problems treated in this chapter some 28 special equations must be developed. Their large number and the general lack of any particular novelty in the procedures used precludes individual derivation of each of these equations herein. Taylor expansions may again be used but with the introduction of the additional conditions that the normal components of electric displacement and the tangential components of electric intensity are preserved in traversing dielectric boundaries, while for open circuit boundaries the normal electric intensity is zero.

These special equations, written for square nets, are tabulated in Table I, which also gives inset diagrams showing the conditions in which each applies.

E. THE SOLUTION OF LARGE GROUPS OF SIMULTANEOUS EQUATIONS

Having obtained a system of finite difference equations adequate to represent the problem, perhaps numbering several thousand, it is necessary to have an economical method for their solution. The simplest procedure is, of course, direct solution, but, while this may have advantages for specific problems, it suffers from the general disadvantage in machine use of being much too
<table>
<thead>
<tr>
<th>Oddity No.</th>
<th>Description</th>
<th>Figure</th>
<th>Cartesian equation</th>
<th>Cylindrical equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ordinary interior point</td>
<td><img src="image" alt="Figure" /></td>
<td>( V_A + V_B + 2V_D - 4V_P = 0 )</td>
<td>See Eq. (9)</td>
</tr>
</tbody>
</table>

<p>| 2         | Ordinary point         | <img src="image" alt="Figure" /> | ( V_A + V_B + 2V_c - 4V_P = 0 ) | See Eqs. (11, 32)    |
| 3         | Top edge               | <img src="image" alt="Figure" /> | Simple permutation of (2) | ( V_A + V_B + 2V_c - 4V_P = 0 ) |
| 4         | Left-hand edge         | <img src="image" alt="Figure" /> | Simple permutation of (2) | ( 4(N-1)V_B + (2N-3)V_c + (2N-1)V_D - B(N-1)V_P = 0 ) |
| 5         | Right-hand edge        | <img src="image" alt="Figure" /> | Simple permutation of (2) | Simple permutation of (4) |
| 6         | Corner point           | <img src="image" alt="Figure" /> | ( V_B + V_D - 2V_P = 0 ) | ( -V_B + 2V_D - 3V_P = 0 ) |
| 7         | Right-hand side        | <img src="image" alt="Figure" /> | Simple permutation of (6) | Simple permutation of (6) |</p>
<table>
<thead>
<tr>
<th>Page</th>
<th>Top</th>
<th>Left-hand side</th>
<th>Right-hand side</th>
<th>Simple permutation of (6)</th>
<th>Simple permutation of (6)</th>
<th>Simple permutation of (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>$V_B + V_C - 2V_p = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>$(1 + k_x) V_A + (1 + k_x) V_B + 2V_C + 2k_x V_D - 4(1 + k_x) V_p = 0$</td>
<td>$(2N(k_x+1) - (k_x+3)) V_A + (2N(k_x+1) - (k_x+3)) V_B + 2(2N-3) V_C + 2k_x(2N-1) V_D - 4(2N(k_x+1) - (k_x+3)) V_p = 0$</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$(2N(k_x+1) - (3k_x+1)) V_A + (2N(k_x+1) - (3k_x+1)) V_B + 2k_x(2N-1) V_C + 2(2N-1) V_D - 4(2N(k_x+1) - (3k_x+1)) V_p = 0$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$4k_x(N-1) V_A + 4(N-1) V_B + (2N-3)(k_x+1) V_C + (2N-1)(k_x+1) V_D - 8(N-1)(k_x+1) V_p = 0$</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (12)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td>$2V_A + (k_x + 1) V_B + 2V_C + (k_x + 1) V_D - 2(k_x + 3) V_p = 0$</td>
<td>$4(N-1) V_A + (2N(k_x+1) - (k_x+3)) V_B + (2N-3) V_C + (2N-1)(k_x+1) V_D - 2(2N(k_x+3) - (k_x+7)) V_p = 0$</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (14)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Simple permutation of (14)</td>
<td></td>
</tr>
<tr>
<td>Oddity No.</td>
<td>Description</td>
<td>Figure</td>
<td>Cartesian equation</td>
<td>Cylindrical equation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Fourth quadrant</td>
<td><img src="image" alt="Fourth quadrant" /></td>
<td>$2k_e V_A + (k_e + 1) V_B + 2k_e V_C - 2(3k_e + 1) V_p = 0$</td>
<td>Simple permutation of (16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Obtuse</td>
<td><img src="image" alt="First quadrant" /></td>
<td>$2(N-1)V_A + (2N(k_e+1) - (k_e+3)) V_B + 2(N-3) V_C + 2(N-1)(k_e+1) V_D - 4(N-1)(3k_e+1) V_p = 0$</td>
<td>Simple permutation of (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Second quadrant</td>
<td><img src="image" alt="Second quadrant" /></td>
<td>Simple permutation of (18)</td>
<td>Simple permutation of (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Third quadrant</td>
<td><img src="image" alt="Third quadrant" /></td>
<td>$2(N(k_e+1) - (k_e+3)) V_A + 4(N-1) V_B + 2(N-3)(k_e+1) V_C + 2(N-1) V_D - 4(k_e+3)(N-1) V_p = 0$</td>
<td>Simple permutation of (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Fourth quadrant</td>
<td><img src="image" alt="Fourth quadrant" /></td>
<td>Simple permutation of (18)</td>
<td>Simple permutation of (20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Dielectric interface to:</td>
<td><img src="image" alt="Dielectric interface" /></td>
<td>$k_e V_A + V_B + (k_e + 1) V_D - 2(k_e + 1) V_p = 0$</td>
<td>Simple permutation of (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Dielectric to right-hand side</td>
<td><img src="image" alt="Dielectric to right-hand side" /></td>
<td>Simple permutation of (22)</td>
<td>Simple permutation of (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Top edge</td>
<td><img src="image" alt="Top edge" /></td>
<td>Simple permutation of (22)</td>
<td>Simple permutation of (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Page</td>
<td>Description</td>
<td>Equation 1</td>
<td>Equation 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Dielectric to right-hand side</td>
<td>Simple permutation of (22)</td>
<td>Simple permutation of (24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Left-hand edge Dielectric to top</td>
<td>Simple permutation of (22)</td>
<td>[(2N(k_e+1)-(k_e+3))V_B + (2N-3)V_C + (2N-1)k_eV_D - 2(2N(k_e+1)-(k_e+3))V_p=0]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Dielectric to bottom</td>
<td>Simple permutation of (22)</td>
<td>[(2N(k_e+1)-(3k_e+1))V_B + (2N-3)k_eV_C + (2N-1)V_D - 2(2N(k_e+1)-(3k_e+1))V_p=0]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Right-hand edge Dielectric to top</td>
<td>Simple permutation of (22)</td>
<td>Simple permutation of (26)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Dielectric to bottom</td>
<td>Simple permutation of (22)</td>
<td>Simple permutation of (27)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
profligate in its use of the high-speed store and also slow because it makes no use of the sparse nature of the matrix of the coefficients of the finite difference equations, which can have nonzero entries on no more than five diagonals.

As a way out of these difficulties some form of systematic iterative solution may be used [9]. This has distinct advantages in the solution of finite difference arrays in that, since these represent only an approximation to the original problem, it is clearly quite fruitless to force a solution beyond the limits of validity of the original approximations. Iterative methods can be terminated easily whenever desired, while with direct solution it is usually necessary to take the whole solution to finality, using the full "accuracy" permitted by the machine, to obtain any answers at all. They also have the further advantages of requiring the constant repetition of only a small group of machine orders which makes for economy in program storage. The principal disadvantage can be failure to converge, but with proper ordering of the solution [10] this is never a difficulty in any case to be considered in this chapter.

It is not appropriate in this chapter to consider the merits of the various iterative schemes for the solution of the finite difference equations. In much modern work the favored method appears to be successive over relaxation (abbreviated SOR), an overrelaxed version of the common Gauss-Seidel cycle which, under optimum conditions, can be made several tens of times faster. In this process relaxation occurs according to the cycle [11]

$$V_{i(j+1)} = V_{i(j)} - \frac{\Omega}{a_{ii}} \left\{ \sum_{k=1}^{n} a_{ik} V_{i(k)} - b_i \right\}$$  \hspace{1cm} (12)$$

whereby a gradually improving solution is obtained to the simultaneous equations, written in matrix form as

$$A \mathbf{V} = \mathbf{B}$$

where $a_{ik}, V_i, b_i$ are elements of the matrices $A, V, B$, respectively; $j$ measures the number of iteration cycles; $n$ is the number of simultaneous equations; $\Omega$ is the accelerating factor.

The necessary and sufficient conditions for the convergence of SOR do not appear to be fully known but a sufficient condition is that $A$ be symmetric positive definite, possess Young's property $A$, and be ordered consistently with this property [10]. Use of the five point operator under the conditions imposed in this chapter always leads to matrices that satisfy these requirements. In that case SOR is convergent for all $\Omega$ between 0 and 2 and reduces to the common Gauss-Seidel cycle when $\Omega$ is unity. There is an optimum $\Omega_0$, such that $1 < \Omega_0 < 2$, for which the rate of convergence is a maximum. It is obviously crucial, therefore, to have some means for determining $\Omega_0$.

Except in the most trivial cases it is not possible to find $\Omega_0$ before commencement of the iteration process. The optimum accelerating factor depends in an
implicit way upon the largest latent root of an associated matrix derived from \( A \) and there is therefore generally no way of determining it separately. To overcome this difficulty a method whereby it may be approximated with increasing accuracy as the computation proceeds has been devised by Carre [12] and, while obviously not as fast as optimum SOR, it offers a good practical compromise in most cases.

Since this chapter is not intended as a treatise on numerical analysis, rather than give a necessarily brief and thus probably inadequate treatment of the basis of the Carre method, we refer the reader to either the original paper [12] or Fox's summary of it [10]. Carre also gives in his paper a method for estimating an upper bound on the largest remaining error after any complete iteration cycle, and this serves to set up a termination criterion.

F. Use of the Computed Potential Field

A knowledge of the potential field suffices to compute the capacity between the conductors, the charges on them, and the losses under RF excitation associated both with them and the dielectric medium. From these the transmission line constants or obstacle equivalent circuit parameters can then be derived. Since determination of capacity is the most fundamental use for the potential field it will be considered first.

1. Calculation of Capacity

a. Gauss's Theorem Solution. According to Gauss’s theorem [13] the total charge on any conductor is equal to the integral of the electric displacement through any surface which completely encloses it. Thus consider Fig. 5, in which the enclosing surface is drawn around the conductor in the dielectric medium by a series of line segments parallel to the coordinate directions. At any point on this surface the normal component of electric displacement is given by

\[
D_n = \varepsilon E_n = -\varepsilon (\partial V/\partial n)
\]

where \( D_n \) is the normal displacement; \( E_n \) is the normal electric intensity; \( n \) is the normal coordinate.

From the Taylor expansion given earlier in Section II.B it follows that, in terms of the nodal potentials at \( A \) and \( B \) on each side of it, the potential gradient at \( P \) is

\[
\frac{\partial V}{\partial n} = \frac{V_B - V_A}{2a}
\]

with a dominant error in the order of \( a^2 \). Equation (14) is true irrespective of whether a Cartesian or cylindrical coordinate system is being employed but the methods of integration to determine charge are materially different in each case.
Thus, for a Cartesian problem, if the enclosing surface is made up of $s$ chords, the $q$th one of which contains $r_q$ nodes, the charge per unit axial length on the conductor enclosed is given by

$$Q = a \sum_s \sum_{1}^{s} \left[ \epsilon \frac{\partial V}{\partial n} \right]_P$$

(15)

where the symbol $\sum'$ is used to indicate that the first and last terms in the summation are halved. In the above if the dielectric is uniform over the cross section $\epsilon$ may be taken outside the sum.

In cylindrical coordinates it is necessary to distinguish two kinds of surfaces over which integration may be required. These are surfaces of constant $r$, or cylinders about the axis of symmetry, and surfaces of constant $z$, or plane annular surfaces normal to the axis of symmetry. In any given problem the total surface of integration will, in general, be made up by the interconnection of subsurfaces of each kind. If the enclosing surface consists of $n$ subsurfaces, the $q$th one of which is a cylinder in the $i$th row containing $r_q$ nodes, then the contribution to the total charge enclosed found by integration over this subsurface is

$$\Delta Q_q = 2\pi (i-1) a^2 \sum_{P=1}^{r_q} \left[ \epsilon \frac{\partial V}{\partial r} \right]_P$$

(16)
Alternatively if the $q$th subsurface is an annulus in the $j$th column, spanning between the $K$th and $m$th rows, then its contribution to the integral is

$$
\Delta Q_q = 2\pi a^2 \sum_{p=1}^{m-K+1} (K+p-2) [\epsilon(\partial V/\partial z)]_p
$$

(17)

The total charge is then

$$
Q = \sum_{q=1}^{n} \Delta Q_q
$$

(18)

In both the Cartesian and cylindrical cases these integrations will be seen to have been performed by what is equivalent to the application of the trapezoidal rule. This is known to involve a dominant error of the order $a^2$ and may therefore be looked upon as consistent with the remainder of the finite difference process.

From charge, capacity then follows as

$$
C = QV_r^{-1}
$$

(19)

where $V_r$ is the potential difference between the conductors. A numerical procedure for the determination of capacity of the type outlined above appears to have been first used by Gandy and Southwell [14] to compute the capacity between two coaxial dissimilarly orientated square prisms. It must be pointed out that when this method is used the location of the surface of integration is not without influence on the accuracy of the answer obtained although for an accurately determined potential field this should not be so. This is partly due to the use of a relaxation process that is not uniformly accurate at all points to solve for the nodal potentials, but over and above this is the fact that the
adequacy of the finite difference equations to represent the true field is not the same for all regions of the problem.

A possible variation of the surface integral method is to integrate along the conductors themselves. Thus, consider Fig. 6: A "one-sided" approximation to the normal derivative of the field at $P$, accurate also to within the order $a^2$, may be found by considering the potentials at the first two nodes along a line normal to the surface, that is,

$$\frac{\partial V}{\partial n} = \frac{4V_B - V_C - 3V_P}{2a}$$  \hspace{1cm} (20)$$

The remainder of the calculation then proceeds as before.

This method is not as accurate as Eq. (14) since, while the order of the dominant error is the same in each case, the coefficient of the $a^2$ in the error term associated with Eq. (20) is doubled. However, there are certain problems in which a knowledge of the charge distribution along the conducting surfaces is required and Eq. (20) may be used to obtain it. It will also be used later in this chapter in considering attenuation (Sections II,F,2,b).

b. Energy Integral Solution. Determination of capacity by the energy integral approach is based upon the fact that the energy contained in the region between the conductors is given by

$$W = \frac{1}{2} \int \varepsilon E^2 \, dv = \frac{1}{2} CV_0^2$$  \hspace{1cm} (21)$$

where $v$ is the volume [1]. This may be translated into finite difference form by first considering Fig. 7, a single square in the finite difference net. It is desired to find the energy associated with this square. This in turn demands a knowledge of the average value of the square of the electric intensity over the region bounded by the square.

Resolving gradients in each coordinate direction along the sides of the square it is readily shown by using Eq. (14) that, in either the Cartesian or cylindrical coordinate systems, at the centre of the square

$$E^2 = \frac{(V_A - V_C)^2 + (V_B - V_D)^2}{2a^2}$$  \hspace{1cm} (22)$$

If the net contains $m$ rows and $n$ columns, integration over the net serves to show that (1) in the Cartesian case,

$$C = a^2/V_0^2 \sum_{p=1}^{m-1} \sum_{q=1}^{n-1} (\varepsilon E^2)_{pq}$$  \hspace{1cm} (23)$$

and (2) in the cylindrical case,

$$C = \pi a^2/V_0^2 \sum_{p=1}^{m-1} (2p-1) \sum_{q=1}^{n-1} (\varepsilon E^2)_{pq}$$  \hspace{1cm} (24)$$
Again in each case, if the dielectric is homogeneous, \( e \) may be taken outside the summations.

Although the calculation of capacity by Gauss's theorem requires less computing time, this second method is considerably more accurate and, where checks have been possible, has been shown to give errors three to ten times smaller. This commends its use since most of the computer time is expended in relaxing the potential field, and to use this method adds almost negligibly to the total time.

On a purely heuristic basis some increase in accuracy is not unexpected of the energy integral method since full use is made of the whole computed potential field rather than of a sample drawn from it. More specifically, finding capacity by a method which depends on the square of the potential gradient should give more accurate answers as, while the squaring process raises the order of the quantities involved, it does no more than approximately double the residual errors. Equation (21), the integral form, constitutes the basis of the variational method for determining the upper bound on capacity, and it may be shown by the Euler–Lagrange equations [15] to be stationary with respect to first-order variations in the potential field. A good theoretical understanding of the larger errors involved in the Gauss theorem integration is yet to be had.

2. Determination of Transmission Line Constants

a. The Lossless Case. Given capacity, characteristic impedance follows easily, although two cases need to be distinguished. In the simplest, where the medium between the conductors is homogeneous, characteristic impedance is found from

\[
Z_0 = (Cv)^{-1}
\] (25)
where \( v = v_0 / \sqrt{(k_e)} \) is the phase velocity in the medium; \( v_0 \) is the velocity of light in free space \((2.997925 \times 10^8 \text{ meters/sec})\) and \( k_e \) is the dielectric constant of the medium between the conductors.

If the medium is not homogeneous two steps are necessary, capacity first being determined with all dielectrics removed and then with them present. In the usual case of a nonmagnetic dielectric the inductance per unit length is independent of the presence of the dielectric, and characteristic impedance of the assumed TEM mode then follows as

\[
Z_0 = [v_0(C_0C)^{1/2}]^{-1}
\]

where \( C_0 \) is the capacity without dielectrics and \( C \) is the capacity with dielectrics present, with the phase velocity in the line being given by

\[
v = v_0(C_0/C)^{1/2}
\]

It is worth noting that this argument is not absolutely correct as it is not difficult to see that, in general, a line having a dielectric medium which is discontinuous over its cross section cannot support a pure TEM wave [16–18]. The error involved is frequency sensitive but in typical cases, for lines of simple geometry, has been shown analytically to be not more than a per cent or so in typical lines up to frequencies of several gigacycles per second, and can therefore usually be neglected.

b. Lines with Small Losses. On a matched line, that is, a line carrying a traveling wave in one direction only, the attenuation constant is defined as

\[
\alpha = \frac{W_L}{2W_T}
\]

where \( W_L \) is the power lost per unit length and \( W_T \) is the power transmitted.

A line with conductor (series) losses cannot support a pure TEM mode as there must be some component of the electric field in the direction of propagation, although this is not true of dielectric or shunt losses. However, for small losses, mode purity may be assumed and the fields in the cross section taken to be those computed for the lossless case. Actual losses may then be estimated by perturbation, conductor and dielectric losses being considered separately.

The behavior of a lossy dielectric may be expressed as a complex permittivity \( \varepsilon = \varepsilon' - j\varepsilon'' \). If this is substituted in Eq. (21) and the loss term extracted, dielectric losses per unit length are given by, in analogy with Eq. (23),

\[
W_{LD} = \omega d^2 \sum_{p=1}^{m-1} \sum_{q=1}^{n-1} (\varepsilon'' E^2)_{pq}
\]

where \( \omega \) is the angular frequency of RF excitation. In the machine this computation presents no problems and is easily carried out as part of the same operation as calculates the stored energy integral.

To find conductor losses it is first necessary to estimate the surface current
density on the conductors, assuming the skin effect to be well established. The electric intensity at any nodal point \( P \) on a conducting boundary is found from Eq. (29). It then follows that the surface current density at \( P \) is given by

\[
J_P = \frac{E_P}{\eta}
\]

(30)

where \( \eta \) is the intrinsic impedance of the medium. It will be assumed that this surface current flows in a uniform sheet to a depth equal to the skin depth \( \delta \). It then follows that the power lost per unit length in the conductors is given by

\[
W_{LC} = (\rho/\delta) \sum J_P^2
\]

(31)

where \( \rho \) is the resistivity of the conductor material and \( \sum \) denotes the operation of summing over the conductor surfaces.

This estimate of conductor loss will always be either correct or pessimistic [19]. This arises from the assumption of a uniformly deep sheet current which is only true for a few simple line geometries, such as ordinary cylindrical conductor coaxial line.

**G. Methods of Improving the Solution**

Southwell [5] has proposed "advancing to a finer net" to speed computation. In this method the computation is started on a relatively coarse grid and the values obtained for the nodal potentials used to interpolate starting solutions on a finer net which is then further relaxed. This may be continued any number of times. Although originally proposed for hand relaxation this process is easily applied on a digital computer; at the outset a coarse net is used with all interior nodal points set at, say, zero potential. These are iterated to termination when a simple linear interpolation routine serves to set up initial solutions on a more refined net. In this way a significant saving of computer time is effected. However, more than this simple advantage can be obtained.

The Richardson method of the "deferred approach to the limit" [20] has already been mentioned. In this method solutions are worked out for capacity, or any of the other quantities sought, using at least two and preferably three nets in increasing order of refinement. Each solution is said to be worked out at a mesh number \( n \), where the mesh number is defined as the number of meshes (squares) abutting some assigned defining dimension in the problem, such as \( ST \) in Fig. 1. Then if \( C_1, C_2, \) and \( C_3 \) are solutions obtained for mesh numbers \( n_1, n_2, n_3 \) \((n_3 > n_2 > n_1)\) a better solution is [21]

\[
C = b_3 C_3 - b_2 C_2 + b_1 C_1
\]

(32)

where

\[
b_1 = n_1^4(n_3^2-n_2^2)/D \quad b_2 = n_2^4(n_3^2-n_1^2)/D \quad b_3 = n_3^4(n_2^2-n_1^2)/D
\]

\[
D = n_1^2 n_2^2(n_2^2-n_1^2)-n_1^2 n_3^2(n_3^2-n_1^2)+n_2^2 n_3^2(n_3^2-n_2^2)
\]

provided that the approach to the limit is monotonic.
Some idea of whether this process has been reliably applied may be obtained by computing the ratio

\[ R = \frac{a_2 C_2 - a_1 C_1}{a_3 C_3} \]  

(33)

where

\[ a_1 = n_1^2(n_3^2 - n_2^2) \quad a_2 = n_2^2(n_3^2 - n_1^2) \quad a_3 = n_3^2(n_2^2 - n_1^2) \]

For true monotonic convergence \( R \) is unity and in practice its closeness to unity gives a measure of the goodness of the extrapolation.

Richardson's method is based on the supposition that the true solution is related to the finite difference solution by an equation of the form

\[ C = C_f + Aa + Ba^2 + \cdots \]  

(34)

where \( C \) is the exact solution and \( C_f \) is the finite difference solution. Its application can often produce an answer in which the error has been reduced by a factor of 2 or 3. Fox [22] has stated that the conditions under which it can be applied with certainty of producing a useful improvement are not yet known but it can be expected to give its best results with "well-behaved" problems which do not possess reentrant corners or the like. It is the author's experience that these conditions are not too stringent, for, in using it with strip lines having very thin inner conductors, for example, it has always been found to yield improved results whenever a check has been possible.

H. ESTIMATING THE ACCURACY OF THE NUMERICAL SOLUTIONS

There does not appear to be any direct theoretical method for estimating the accuracy of these finite difference solutions although, if the energy integral method is used, there are good theoretical grounds for believing the answer obtained for capacity will be too large. Nonetheless there are ways of establishing with some precision the degree of confidence of the solution. The simplest of these which is at all reliable is to obtain a numerical result for a problem which ideally has an exact analytic solution and which resembles as closely as possible the unknown problem. In each case a comparable number of nodes is used and the assumption made that the error levels in both problems will be of comparable order.

While this has been found to be the only really suitable method for cylindrical coordinate problems, at the time of writing investigations are in progress on the application of the duality principle to Cartesian problems. With this method the open circuit boundaries of the problem are made electric conductors, with the original conductors being treated as open circuits. In this way it is possible to obtain an upper bound on the reciprocal of capacity for the original cross section, that is, a lower bound on that capacity. The region in which the exact answer must lie can then be established with precision without the need for any comparison problem.
Estimates made in these ways have verified that, with a maximum of a few thousand nodal points in the most refined net, capacity can be obtained to an accuracy of between 1 in $10^3$ and 1 in $10^4$ in 2 to 3 minutes' computing time. For characteristic impedance determination this is an order or so better than is commonly allowed by tolerances in construction and on the properties of dielectric materials used.

III. CALCULATIONS ON TRANSMISSION LINES

The main purpose of this section is to present examples of problems on uniform transmission lines which may be solved using the finite difference technique outlined in the first part of this chapter. A second and very important intention is to show the relationship of the purely numerical approach to the more familiar techniques of conformal transformation and the variational method, as well as other numerical methods. It is therefore proposed to adopt as broad an outlook in the selection of these examples as necessary limitations of space permit. This should serve to illustrate the generality of the numerical method and its lack of any need for involved preliminary analysis. It will also serve to show its weakness, the most important of which is the amount of computer time required.

A. POLYGONAL LINES

A class of problem which has attracted attention in the recent literature is the calculation of the characteristic impedance of transmission lines formed by two conductors, which are polygons in section, one inside the other. These lines will be referred to as "polygonal lines." The need to consider them has arisen, for example, in the design of transitions between conventional coaxial lines and the various types of strip line and as mounts for certain microwave solid state components.

For obvious practical reasons much of the work published, though by no means all, has been concerned with what may be called regular polygonal lines, that is, lines in which the conductors are similar, similarly orientated, regular concentric prisms. Another important subgroup which will be considered later is coaxial rectangular conductors.

1. SQUARE COAXIAL LINE

The square coaxial line shown in section in Fig. 8 is an excellent example of a regular polygonal line which has been treated by a number of authors, two types of conformal transformation solution (accurate and approximate), a finite difference solution, and a solution by a process called "orthonormal block analysis" having been presented in the literature [23–26].

The method of orthonormal block analysis consists in series expansion of the
potential function in the region of interest and using the boundary conditions to provide two infinite sets of simultaneous linear equations from which the coefficients of the series may be obtained, giving upper and lower bounds on the solution sought. So brief a description does not do justice to a very interesting analytic method but space does not permit a further explanation and the method will not be considered further in this chapter. The interested reader should consult Cruzan and Garver [27].

The accurate solution is based upon a Schwarz–Christoffel conformal transformation and is rather involved. From Anderson’s analysis [23] it follows that

\[
Z_0 = 5\pi v \times 10^{-8} \frac{K(k')}{K(k)}
\]  

where \(K\) denotes a complete elliptic integral of the first kind

\[
k = (\lambda' - \lambda)^2(\lambda' + \lambda)^2 \quad \lambda' = (1 - \lambda^2)^{1/2}
\]

\(\lambda\) is determined in terms of the problem geometry from

\[
\frac{K(\lambda')}{K(\lambda)} = \frac{b-a}{b+a}
\]

where \(a\) is the flat width of the inner conductor and \(b\) is the flat width of the outer conductor. This formula is not easy to use for hand calculation unless a good table of elliptic integrals is available [28]. A good approximate formula is therefore welcome.
The approximate conformal transformation solution relies on the fact that it is comparatively easy to obtain a Schwarz–Christoffel transformation which will map the region between two equal semiinfinite parallel plate lines meeting at right angles and thus allow the determination of the capacity associated with the corner region so formed [24]. It is then assumed that if the corner regions in a square coaxial line are not too close together they may be reckoned as isolated and the total capacity found by summing the contributions of the four corners and the parallel plate regions joining them. Characteristic impedance then follows as

$$Z_0 = \frac{1}{4\pi(C_{pp} + C_c)}$$

(36)

where $C_{pp}$ is the parallel plate component of capacity and $C_c = 0.558\epsilon$ is the corner component.

It may be expected that this formula is valid only for a side length ratio near unity where the corners are well separated, but a comparison with the accurate formula reveals an error of no more than 1% to a side length ratio of 4. At that stage the corners contribute nearly one-half the total capacity and are separated by just two-thirds the spacing between the inner and outer conductors. The approximate formula is thus seen to be good over most of the range of characteristic impedances that are of engineering interest, an observation which will be generalized in the next subsection.

This problem has also been treated by the finite difference method [25] using both Gauss's theorem and the energy integral methods to obtain capacity from the same computed potential field, and hence characteristic impedance. Consider Fig. 8; owing to symmetry about the two axes $AA'$ and $BB'$ one quarter only of the cross section need be considered with the open ends so formed being considered closed with magnetic conductors (perfect open circuits). In this way a net effectively four times as large may be treated in the

### Table II

<table>
<thead>
<tr>
<th>Flat width ratio</th>
<th>Conformal transformation solution</th>
<th>Numerical analysis solution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exact</td>
<td>Approximate</td>
</tr>
<tr>
<td>2</td>
<td>36.81</td>
<td>36.82</td>
</tr>
<tr>
<td>3</td>
<td>60.61</td>
<td>60.64</td>
</tr>
<tr>
<td>4</td>
<td>77.74</td>
<td>76.89</td>
</tr>
</tbody>
</table>
same computing time. As with the analytic methods it is always important to make use of any simplifications allowed by symmetry in using the finite difference method. Computing time to obtain an accuracy of better than 0.5% is about 1 minute per data point on an IBM 7090 machine.

The results of these various methods are compared in Table II. The usefulness of the approximate formula, already pointed out, is highlighted by this table as is also the closeness to the exact solution of the answers obtained purely numerically. The increase in accuracy which follows from the use of the energy integral method is also apparent [29].

2. A More General Problem

The approximate analytic method may in principle be generalized to any regular polygonal line provided only that the corners can be considered isolated. Consider Fig. 9. If the conductors each have \( n \) sides then it is required to find

![Fig. 9. Regular polygonal line.](image-url)
the capacity of a bend of angle $\phi = 2\pi/n$ at the junction of two equal semiinfinite parallel plate lines.

This problem has been solved analytically by the conformal transformation technique. When $\phi$ takes the values $\pi/2$ or $\pi$ (the right angle bend [24] and bifurcated parallel plate line [30], respectively) a relatively simple application of the Schwarz–Christoffel transformation is sufficient. Though more complex, the problem has been solved for arbitrary $\phi$, for example, in determining the equivalent network of an E plane bend in rectangular waveguide [31], leading to the formula

$$C_c = \frac{2e}{\pi} [\Psi(-\frac{1}{2}(1 - \phi/\pi)) - \Psi(-\frac{1}{2})]$$

(37)

where $\Psi(x) = -\gamma + \sum_{n=1}^{\infty} [(1/n) - 1/(x + n - 1)]$ is the derivative of the logarithm of $\Gamma(x + 1)$ and $\gamma$ is Euler’s constant (0.5772).

![Fig. 10. Capacity versus bend angle for a mitered bend in parallel plate line.](image-url)
Provided that a rectangular rather than a square net is used there is no basic difficulty in applying finite difference theory to this problem. This was done to prepare the data presented in Fig. 10, which shows corner capacity against bend angle. Since, of course, it is impractical to treat infinite parallel plate lines numerically they must be terminated a finite distance on each side of the corner. As the comparison of the accurate and approximate solutions for square coaxial line makes clear, lengths of parallel plate line equal to twice the conductor separation are more than adequate to approximate semiinfinite lines. The ends of these lines are sealed off with magnetic conductors since the field is assumed to have become uniform at the point of truncation.

With these results Eq. (36) may be generalized to

$$Z_0 = \frac{1}{n\pi(C_c + C_{pp})}$$  \hspace{1cm} (38)

where \(C_c\) is the corner capacity found by entering Fig. 10 at \(\phi = 2\pi/n\); \(C_{pp} = 2\varepsilon\tan(\pi/n)/(s-1)\) is the parallel plate capacity; and \(s\) is the ratio of outer to inner conductor flat width.

The data points used to construct Fig. 10 were computed using the surface integral method and are a little on the high side with an average error of about 0.5%.

This and the previous example bring out certain of the features of the numerical method which have been alluded to previously. No preliminary mathematical work is required to set up the problem and the answers obtained are adequately accurate for most engineering purposes. On the other hand, the computing time per data point, about 30 seconds, is typically a hundred times greater than is required to compute a value from the analytic formula once it is available.

**B. SHIELDED STRIP LINE**

The last decade and a half has seen the large scale introduction of various forms of strip line, both shielded and unshielded, in an attempt to simplify the construction of microwave components. This section is concerned with the determination of the properties of shielded strip line, which consists basically of an inner strip conductor sandwiched between and spaced from two conducting ground planes wide compared with the width of the strip. These ground planes are sometimes closed at the ends to form a complete rectangular conducting outer shell.

Various forms of shielded strip line have been devised and discussed in the literature. Some have inner conductors consisting of a thin metal shim totally embedded in solid dielectric [32, 33]. Others have an inner conductor made up of a dielectric card with a thin conducting strip on one or both sides [34, 35] (with the dielectric occasionally not extending beyond the edges of the strips) and air spaced from the ground planes. Yet another has a solid rectangular
bar as inner and may be air spaced or dielectric filled [1, 36, 37]. Replacement of the rectangular bar by a circular rod has also been discussed as this has obvious advantages in construction [38, 39]. These various forms of strip line are illustrated in Fig. 11.

Fig. 11. Various forms of shielded strip line. (a) Single strip; (b) double strip, dielectric supported; (c) single strip, dielectric supported; (d) double strip, partial dielectric support; (e) solid inner bar; (f) circular rod, inner.

Assuming infinitely wide ground planes the cross section shown in Fig. 11a may be solved by the use of standard conformal transformation techniques [32], giving the result

\[ Z_0 = \frac{30\pi}{\sqrt{(k_e)}} \frac{K'(\kappa')}{K(\kappa)} \]  

(39)

where \( \kappa = \tanh (\pi w/2b) \) and \( \kappa' = (1 - \kappa^2)^{1/2} \). The approximation of infinitely wide ground planes is justified by the rapid exponential decay sideways of the field within the line. Hayt [40] has shown that provided \( D > 2.5 \, w \) and \( D > b + 0.5 \, w \) this assumption is effectively met, the error being about 0.25%.

When the width of the inner conductor is sufficiently great compared with its spacing from the ground planes, the fringing fields at each end may be considered as effectively isolated [32] and the line taken as consisting of two
bifurcated parallel plate lines connected back to back. Equation (39) then simplifies to

$$Z_0 = \frac{1}{2\pi(C_p + C_r)}$$

(40)

where $C_p$ is the parallel plate component of capacity and $C_r = 0.882\epsilon$ is the fringing capacity [30]. For common dielectric materials this formula applies to most lines having characteristic impedances in the range of practical interest. Use will be made of this simplification when coupled strip lines are considered later in this chapter.

The configuration shown in Fig. 11b has been solved analytically for the case where the dielectric slab is removed, using either a Schwarz–Christoffel transformation involving an approximation which is designed to remove one vertex from the path of integration [41] or by a variational approach which gives an upper bound [42]. The conformal transformation solution leads to the formula

$$Z_0 = 30\pi \frac{K(\kappa')}{K(\kappa)}$$

(41)

where $\kappa$ is a parameter related to the line geometry by

$$\frac{w}{b} = \frac{2}{\pi} \left[ \tanh^{-1} \left( \frac{\kappa(b/s) - 1}{\kappa^{-1}(b/s) - 1} \right) \right]^{1/2} - \frac{s}{b} \tanh^{-1} \left[ \frac{1}{\kappa} \left( \frac{\kappa(b/s) - 1}{\kappa^{-1}(b/s) - 1} \right) \right]^{1/2}$$

which is valid for $w/s > 0.35$.

The variational solution relies upon finding a sufficiently accurate approximation for the charge distribution on the conducting strips and uses the Green's function technique to solve the resulting boundary value problem for an upper bound on characteristic impedance. The method may be made arbitrarily accurate by taking sufficient terms in a series representation of the charge distribution but only at the expense of rapidly mounting mathematical labor [1]. By using a three-term series Duncan [42] has obtained answers accurate within 2% over a wide range of geometries. His work may be considered as complementary to the approximate conformal transformation method since it is useful in regions where the latter solution is not completely valid. It also covers the case of coupled strip lines (Section III.C).

The variational formula is complicated to write as it involves several infinite series and thus will not be reproduced herein. To obtain numerical results these series are summed on a digital machine and it is reported that 900 data points can be computed in only 2 minutes' machine time. The method, however, can hardly be described as "numerical" since much detailed analysis precedes any computation. Further the variational method requires a special formulation for each individual class of problem and the dielectric cannot be included easily.

An approximate solution which does take account of the dielectric has been
THE NUMERICAL SOLUTION OF TRANSMISSION LINE PROBLEMS 357

published by Foster [43]. This uses a hybrid conformal transformation variational approach to evolve a method from which answers may be obtained with an accuracy of a few per cent.

The analytic difficulties experienced with this problem are common to most which contain a mixed dielectric and it is in treating these problems that numerical methods excel. Extensive data have been obtained for various strip widths by finite difference solution [4, 44] for the line cross section shown in Fig. 11b and having dimension \( b = 5/16 \) in., \( t = 1/16 \) in. and \( D = 1 \) in. and closed ends. These data are summarized in Table III, which shows results for

Table III

<table>
<thead>
<tr>
<th>Dielectric material</th>
<th>Teflon</th>
<th>Rexolite 2200</th>
<th>Air</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strip width (inch)</td>
<td>Characteristic impedance</td>
<td>Velocity ratio</td>
<td>Characteristic impedance</td>
</tr>
<tr>
<td>0.2188</td>
<td>58.75</td>
<td>0.9506</td>
<td>57.38</td>
</tr>
<tr>
<td>0.2500</td>
<td>54.48</td>
<td>0.9540</td>
<td>53.29</td>
</tr>
<tr>
<td>0.2813</td>
<td>50.79</td>
<td>0.9569</td>
<td>49.74</td>
</tr>
<tr>
<td>0.3125</td>
<td>47.57</td>
<td>0.9594</td>
<td>46.43</td>
</tr>
<tr>
<td>0.3438</td>
<td>44.72</td>
<td>0.9617</td>
<td>43.89</td>
</tr>
</tbody>
</table>

*Note: (1) All characteristic impedances are expressed in ohms. (2) Phase velocity is given with respect to light in free space \( (2.997925 \times 10^8 \) meters/sec). (3) These data apply to lines with the cross section shown in Fig. 11b, having leading dimensions \( b = 5/16 \) in., \( t = 1/16 \) in., \( D = 1 \) in.*

Rexolite 2200, PTFE, and air dielectrics. Characteristic impedance and velocity ratio are given. In computing the data full use was made of symmetry to allow treatment by consideration of one quarter of the cross section.

To test the accuracy of these results a check solution was worked out for air dielectric line using Eq. (41) and it agreed within 0.6%. Similar accuracy can be expected when Rexolite 2200 or Teflon are in place. The results for velocity ratio should, of course, be somewhat better as this involves the ratio of quantities which probably have errors of comparable magnitude. As an estimate of error this method will give results which are slightly pessimistic since the effect of the end walls, neglected in the analysis, will be to decrease the line impedance a little.

It is not surprising that no complete analytic solution appears to have been given for the configuration shown in Fig. 11c, although some experimental
results have been published [45]. This problem is easily treated by finite difference methods, although there is now symmetry about only one axis to simplify the work. Figure 12 gives data for a Rexolite 2200 dielectric. It is clear then, that since about one-half the electric flux lines pass for some portion of their length through the dielectric, losses will be higher for this type of line than in the symmetrical construction.

Only limited attention has been given to the line shown in Fig. 11d [4], although it too is easily handled by numerical means with the simplification of symmetry about two axes. The work that has been done suggests that the effect of the dielectric is negligible and therefore that Eq. (41) is adequate for this case.

Table IV

<table>
<thead>
<tr>
<th>Characteristic Impedance of Strip Line with Solid Inner Conductor (Fig. 11c)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bar width (inches)</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>0.1250</td>
</tr>
<tr>
<td>0.1875</td>
</tr>
<tr>
<td>0.2500</td>
</tr>
<tr>
<td>0.3125</td>
</tr>
<tr>
<td>0.3750</td>
</tr>
</tbody>
</table>

* Note: These data apply to a line with leading dimensions $b=5/16$ in., $t=1/16$ in., $D=1\frac{1}{4}$ in.

For the solid inner bar (Fig. 11e) solutions have been obtained analytically by conformal transformation and by the variational method (both bounds) [1]. Two conformal transformation solutions have been published, one by Getsinger [36], whose paper contains a large collection of results in graphical form, and whose solution is direct, and the other by Cohn [37] in the form of thickness corrections to the infinitely thin case [32].

Some results have also been obtained by the finite difference method for bars 1/16 in. thick, symmetrically placed between ground planes spaced 5/16 in. [4]. These are summarized in Table IV. As a check, solutions for a bar 5/16 in. wide have been worked out by the variational method and the results, given in Table IV, have been found to agree with the mean of the upper and lower bounds within 0.04%, an excellent agreement.

To perform experimental work on strip line components without the need of coaxial-to-strip line transitions there is much to be said for having a slotted section available. The probe may be introduced along either of the axes of
Fig. 12. Data for one sided strip line.
symmetry of the line; an apparatus in which it enters from the side has been described by Cohn [46]. Alternatively a slot may be cut in the center of one of the ground planes [47]; this is much less demanding mechanically since the probe is then inserted in a region of the field where the electric intensity vector has little or no component transverse to it. If the wall thickness of the ground plane is adequate, little slot leakage will result, but some compensation to the width of the center strip is obviously necessary, if the characteristic impedance is to be maintained constant.

The slotted cross section shown in Fig. 13 was studied with the aid of the computer [4]. Since the line is no longer closed one of the boundary conditions to be met is that the potential at infinity is zero. While the application of numerical methods to unbounded fields will be considered in greater detail in a subsequent section, it is satisfactory to say in the present instance that this is met with sufficient approximation by considering the slot to couple into a large conducting box which straddles the slotted ground plane. Assuming Rexolite 2200 as the dielectric support material a strip width of 0.286 in. was found necessary to maintain 50-Ω characteristic impedance. This represents an

![Fig. 13. 50-Ω slotted strip line.](image)

<table>
<thead>
<tr>
<th>Type of line</th>
<th>Reference figure</th>
<th>Strip width (inch)</th>
<th>Velocity ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double strip, Teflon support</td>
<td>Fig. 11b</td>
<td>0.289</td>
<td>0.9380</td>
</tr>
<tr>
<td>Double strip, Rexolite 2200 support</td>
<td>Fig. 11b</td>
<td>0.279</td>
<td>0.9369</td>
</tr>
<tr>
<td>Double strip, thin unsupported strips</td>
<td>Fig. 11b</td>
<td>0.309</td>
<td>1.0000</td>
</tr>
<tr>
<td>Single strip, Rexolite 2200 support</td>
<td>Fig. 11c</td>
<td>0.375</td>
<td>0.9035</td>
</tr>
<tr>
<td>Solid inner bar</td>
<td>Fig. 11c</td>
<td>0.298</td>
<td>1.0000</td>
</tr>
<tr>
<td>Slotted strip line</td>
<td>Fig. 13</td>
<td>0.286</td>
<td>0.9360</td>
</tr>
</tbody>
</table>
increase of 0.007 in. over that in the absence of the slot. The corresponding velocity ratio was found to be 0.9360.

Table V gives a summary of data for the construction of 50-Ω lines in certain of the various forms of strip line treated in this section.

C. Coupled Strip Lines

Various types of strip line devices, such as directional couplers, some types of bandpass filters, and delay lines rely for their action upon the coupling between adjacent strip lines. A pair of coupled strip lines, being basically a three-conductor system, may support two distinct TEM modes—the so-called “even” and “odd” modes—and the behavior of any coupled strip line device is completely determinable if the characteristic impedances and phase velocities of these modes are known. In the even mode case the inner conductors are at equal potentials above the ground planes, but in the odd mode these potentials, while still of equal magnitude, are of opposite sign above and below the ground planes.

Where the inner conductors consist of thin strips or solid conducting bars totally immersed in a uniform dielectric, adequate solutions of this problem have been deduced by conformal transformation and variational processes [36, 41, 42].

In this chapter coupled strip lines of the types shown in Figs. 14a and 14b will be considered exclusively. The first has been treated by the finite difference method and the second by a rather different numerical procedure which affords an interesting comparison with the finite difference method, and could perhaps be coupled with it to yield a useful hybrid method [48].

Consider Fig. 14a. For determination of the odd mode parameters it is sufficient to consider a cell in which the plane AA’ is replaced with an electric conductor and the plane BB’ by a magnetic conductor. For the even mode
case a similar cell may be used except that $AA'$ is also now a magnetic partition. Since in neither case is the dielectric medium homogeneous it is necessary to consider each cell twice, as outlined in Section II,F,2,a.

Provided that the width $w$ of each strip is at least comparable with its separation from the ground plane (almost always so in cases of practical interest) then for all cases the capacity may be written in one of two forms: (1) For the even mode,

$$C_e = C_{pp} + C_f + C_e'$$

and (2) for the odd mode,

$$C_o = C_{pp} + C_f + C_o'$$

where $C_{pp}$ is the parallel plate capacity; $C_f$ is the fringing capacity at the outer ends of the strips [30]; $C_e'$ and $C_o'$ are the fringing capacities at the adjacent ends of the strips under even and odd mode excitation. Figure 15 makes the meaning of these capacities clear.

Since it is not necessary for design to know $C_f$ and $C_e'$ or $C_o'$ explicitly (although they are quite easily separated, $C_f$ being found by halving the even mode intercept with the ordinate axis in Fig. 16), Fig. 16, which also shows the basic dimensions for the case considered, presents answers for the sums of
these quantities. It is assumed that \( w \geq \frac{1}{2}(b - t) \) and that the dielectric material is Rexolite 2200. The parallel plate component of capacity is, of course, \( 4\varepsilon w/(b - t) \). With these data Eqs. (26) and (27) serve to obtain the characteristic impedances and velocity ratios of either mode for a variety of strip widths and separations.

![Graph showing even and odd mode fringing capacities in dielectric (Rexolite 2200) supported, coupled strip lines. Note: These data apply for strips on 1/16 in. thick Rexolite 2200 dielectric board between ground planes spaced 5/16 in.](image)

For many purposes such as the construction of interdigital bandpass filters [49] it is desirable to use coupled strip lines in which the inner conductor is a solid bar supported in air. When the filters were first devised it was customary to use a rectangular bar for the inner conductor, as good theoretical data were available for this type of line [36]. However, from a manufacturing point of view it is obviously better to use cylindrical rods.
Recently Cristal [48] has considered an infinite array of parallel coupled rods between infinite ground planes in an attempt to find the even and odd mode characteristic impedances. The even mode is said to occur when all rods are at the same potential with respect to the ground planes and the odd when alternate rods are at equal magnitude but oppositely signed potentials. The analysis in either case can therefore be reduced to considering a rectangular cell of width equal to the rod center line spacing and with the rod at its center. For the odd mode, all four walls are supposed to be conducting and at ground potential, while, in the even mode, only the top and bottom walls, sections of the ground planes, are electric conductors. The remaining two walls, the parallel sides, are magnetic conductors, or surfaces across which the normal component of potential gradient is zero.

The surface charge density may be found by solving the integral equation associated with the boundary value problem [50], which gives the potential at any point on the surfaces in terms of the charge distribution. Cristal uses a numerical method to solve this equation which it is interesting to describe briefly and to contrast with the finite difference method which, if suitably modified to treat circular conductors (theoretically quite possible), could also be used to solve this problem.

The numerical method adopted is to imagine the inner and outer boundary surfaces to be cut into small segments, the surface charge density (or the potential on a magnetic conductor) being assumed constant but unknown over any of these segments. Over each single segment it is therefore possible to take the unknown constant term outside the integral and integrate, numerically if need be, the remaining terms, which involve only the known boundary geometry. In this way the assigned potential on any segment may be written down in terms of the charge or potential distribution on all segments, including itself. This gives rise to a system of simultaneous linear equations, the coefficients of which are the segment integrals.

To obtain an adequate answer, typically some tens of simultaneous equations are necessary. These may be solved to yield the charge distribution, hence the total charge, and finally capacity. The most efficient method of solution is by matrix inversion on a digital machine.

It will be noticed that, unlike the finite difference method, which considers the whole volume of the medium between the conductors, this method works directly with the conducting surfaces themselves. Clearly therefore it involves considerably fewer simultaneous equations, tens instead of hundreds. However, the matrix of the coefficients is not now "sparse" as with the finite difference method, there being in general no zero elements. Thus to yield the same ultimate accuracy in capacity the time taken for matrix inversion is still likely to be comparable with that taken to relax the finite difference equations.
D. Microstrip Transmission Line

An alternative form of strip line known as microstrip has been used in certain applications [51]. This is shown in cross section in Fig. 17. It is clearly much easier to construct than a shielded strip line but it suffers the disadvantages that not only must there always be some small leakage flux but losses must be higher since, for all practical strip widths, most of the electric field is contained in the dielectric. Mathematical analysis of this form of transmission line has been attempted by various writers but a rigorous solution of even the associated electrostatic boundary value problem is difficult [52, 53]. An outline of a solution for the exact electrodynamic case where the TEM-mode assumption is not made has been given by Wu [18].

Until recently most of the best available data on microstrip was derived experimentally [51, 54, 55] but Wheeler [56, 57] has now given an ingenious approximate solution by the conformal transformation technique. This solution is important because it represents one of the few really successful attempts to cope with mixed dielectrics by simple analytic methods, and may lead the way to new successes.
Fig. 19. Data for two representative microstrip transmission lines. (a) Microstrip using Fiberglas G6.

Since Wheeler's papers are lengthy and purely analytic there is not space here for more than a rather inadequate indication of the scope of his work. The basic problem treated is one of a dielectric sheet carrying, one on each side,
two equal width, symmetrically placed, thin metal shims which, by the image principle, may be regarded as two microstrip lines with infinitely wide ground planes placed back to back. The problem is treated in two parts by considering wide strips (that is, wide compared with the thickness of the dielectric sheet).
and narrow strips, leading to formulas for both analysis and synthesis. While approximate, comparison of this work with solutions obtained by the finite difference method suggests an average accuracy of around 1% (after making due allowances for known errors in the latter method). The side strip formulas are found to hold to good accuracy right down to strip widths comparable to the dielectric thickness.

A set of design graphs based on this study is presented in the second of Wheeler's papers [57] and these effectively provide all the data which ever seem likely to be needed for all microstrip geometries, with all present day microwave dielectrics. While this considerably diminishes the importance of any numerical solution, the inclusion of this section is merited because it illustrates the method of handling unbounded fields.

Since one of the boundary conditions to be imposed is that the potential at infinity is zero it is evident from Fig. 17 that failure to meet this condition will occasion maximum error when \( w \ll s \), as there will then be most leakage flux from between the conductors. Clearly a field of infinite extent cannot be handled in a numerical solution. This difficulty may be avoided by "bringing infinity up close" by enclosing the line in a large conducting box, big enough to contain all regions of the field where the energy density is significant.

In Ref. [58] it is pointed out that for \( w \ll s \) the strip behaves as essentially a circular wire of diameter \( w/2 \). Solutions have been published for the characteristic impedance of circular wires eccentrically placed in air filled rectangular enclosures [39]. If the wire is located near one wall of the box that acts as the ground plane, the dimensions of the box being large compared with the spacing of the wire from the wall (Fig. 18), then it may be shown by making appropriate approximations in the formula given in Ref. [39] that the relative change in characteristic impedance (or capacity) is

\[
\frac{\Delta Z_0}{Z_0} = \frac{5(\pi s/h)^2(1+\exp(-\pi g/2h))}{60 \ln(8s/w)}
\]  

From this it follows that for a box of specified cross section, and therefore a specified number of nodes, the error is a minimum when \( g = 2h \).

This may be used as a rough guide in designing a net for numerical analysis in which the enclosure error is to be kept within acceptable limits. For this problem a choice for \( s:h:g \) of 1:10:20 can be expected to lead to an error in impedance of about 1 \( \Omega \) with the narrowest strips investigated. For wider strips, and with a dielectric present to further limit the leakage flux, even better results can be expected.

Figure 19 shows plots of the characteristic impedance and velocity ratio for microstrip lines having Teflon (\( k_e = 2.05 \)) and Fiberglas G6 (\( k_e = 4.20 \)) dielectrics.
IV. EQUIVALENT CIRCUITS

A. THE QUASI-STATIC METHOD

This section is concerned with the use of numerical methods to determine the equivalent circuits of certain simple nondissipative obstacles which interrupt the longitudinal uniformity of TEM-mode transmission lines. The effect of a discontinuity in a wave transmission system is to give rise in its immediate vicinity to an infinite number of higher order modes which represent energy storage at the obstacle. If the physical size of the obstacle is small compared with the wavelength of excitation the effect of these evanescent higher order modes may be treated as though the frequency was substantially zero.

Arguments based on electrostatic or magnetostatic considerations then apply with considerable accuracy and it is not necessary to do battle with complicated modal expansions to obtain useful solutions. This is known as the quasistatic approach and was described by MacFarlane [59] about 25 years ago. However, even with this simplification there are many problems still lacking a useful solution and it is now to be hoped that the availability of fully mechanized numerical analysis procedures will lead to a speedy demolition of these outstanding problem areas. This section is devoted to the combined application of the quasistatic theory and numerical methods to the solution of some of these problems.

It must, of course, be pointed out that while the quasistatic method gives working approximations to many problems there are other analytic methods of greater power such as those of Marcuvitz [60]. These allow examination of the behavior of many obstacles as functions of frequency, the quasistatic solutions falling out as the low frequency and if they are to be useful, usually the dominant term. These methods however, cannot be embraced by a numerical procedure which is designed purely to solve potential problems and the quasistatic results must suffice. In very many cases this involves negligible real loss and in some cases may well afford the starting point on which to construct a more elaborate electrodynamic solution.

B. STEPS IN COAXIAL LINE

1. Single Step in One Conductor Only

One of the most commonly occurring interruptions to the longitudinal continuity of a coaxial line is the abrupt step in either the inner or outer conductor of the line (Fig. 20). It was shown by Whinnery and Jamieson [61] and Whinnery et al. [62] that this type of obstacle can be represented by a shunt discontinuity capacity at the plane of the step, and that this capacity is invariant with frequency if the dimensions of the line cross section remain
small fractions of the wavelength of excitation. In most coaxial line applications this is the case. An analytic determination of this capacity was attempted with success.

For the parallel plate line, a limiting case of coaxial line, a formula has been obtained by conformal transformation [59]. For coaxial lines the conformal transformation procedure is not applicable but the problem can be attacked by the mode-matching method in which the fields on each side of the junction are expanded in an infinite series of modes matched across the boundary to preserve continuity. This was the approach used by Whinnery and his co-workers in 1944 and although theirs has been virtually the only data available in that time their paper suffers from two deficiencies that this new determination removes, namely,

(a) they do not consider diameter ratios beyond 5, inadequate to meet many needs; and
(b) they published their answers as rather small difficult-to-read graphs.

In scaling these up for publication in handbooks, often it seems by redraffing, considerable reproduction errors have occurred, causing values taken from various books to differ by several per cent.

\[ \text{Fig. 20. Step discontinuities in coaxial lines.} \]
Discontinuity capacity can be computed numerically by noting that it is equal to the difference between the total capacity of a line section including a step, and that computed by adding the contributions of two single unperturbed lines with cross-sectional dimensions and lengths equal to the actual lines on each side of the step. In setting up a model on a digital computer the observed fact that the inhomogeneous fields do not extend beyond the step that generates them by more than a diameter of the outer conductor may be used to limit the volume over which capacity must be calculated. The line may therefore be terminated by magnetic conductors one diameter each side of the step, yet fully include its effects.

This problem was run on the machine to give the discontinuity capacities per unit circumferential length shown in Table VI. Diameter ratios to 11 are considered. As a guide to the accuracy of the data the discontinuity capacities obtained numerically for the limiting case of parallel plate lines were compared with the conformal transformation solutions. The agreement was found to be very good, the average correspondence being 0.4% with an error exceeding 1% only for the case where the capacity is smallest.

This method of computation can be attacked as inefficient since it demands the calculation of the total capacity to great accuracy to obtain an adequately small error in the discontinuity capacity (typically only a few per cent of the total). While this is true, the computation time for the data presented in Table VI, averaged no more than 2 minutes per point.

<table>
<thead>
<tr>
<th>Table VI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discontinuity Capacity in Coaxial Line</strong></td>
</tr>
<tr>
<td><strong>a</strong></td>
</tr>
<tr>
<td><strong>Step in inner conductor</strong></td>
</tr>
<tr>
<td>( \alpha )</td>
</tr>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
</tr>
<tr>
<td>0.4</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
</tr>
<tr>
<td>1.0</td>
</tr>
</tbody>
</table>

*Note: Discontinuity capacities are given in picofarads per centimeter of circumference.*
2. Steps in Both Conductors Simultaneously

In some applications, such as a bead support required to restrain the inner conductor against axial displacement, it is necessary to undercut the inner conductor and overcut the outer conductor at the same reference plane to produce a doubly opening-out discontinuity. A circuit representation of this double step by a single shunt capacity is still legitimate but a problem arises in determining its magnitude. Although there is no general case and any individual problem can be treated either analytically or numerically, it is obviously desirable to be able to use the data given in Table VI.

Consider the discontinuity shown in Fig. 21a. In the undisturbed field

---

Fig. 21. Location of continuous equipotential surface through double discontinuity. (a) Double step in coaxial line; (b) double step in parallel plate line.
regions $R$ and $S$, well away from the step, the lines will be radial, and the potential in the space between the inner and outer conductors will vary logarithmically with radius. There will be a certain diameter where the potentials in each of $R$ and $S$ are equal. In terms of the notation of the figure this may be determined readily as

$$ r = \log^{-1}\left\{ \frac{\log a \log (d/c) - \log c \log (b/a)}{\log (d/c) - \log (b/a)} \right\} $$  \hspace{1cm} (45) $$

or in the limiting case where the line has become a pair of parallel plates (Fig. 21b),

$$ r = \frac{bc}{b+c-d} $$ \hspace{1cm} (46) $$

It is assumed that this equipotential surface will continue through the region of inhomogeneity without substantial deviation from cylindrical form. The discontinuity thus splits into two series connected sections, each of which may be estimated from Table VI.

### Table VII

**Comparison of Accurate and Approximate Methods of Computing a Doubly Opening-out Discontinuity**

<table>
<thead>
<tr>
<th>Type of line</th>
<th>Figure</th>
<th>Discontinuity capacity (pf)</th>
<th>Accurate</th>
<th>Approximate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel plate</td>
<td><img src="image" alt="Parallel Plate" /></td>
<td>0.125</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>Coaxial</td>
<td><img src="image" alt="Coaxial" /></td>
<td>0.597</td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td>Coaxial</td>
<td><img src="image" alt="Coaxial" /></td>
<td>0.298</td>
<td>0.280</td>
<td></td>
</tr>
</tbody>
</table>

*Note: (1) Dimensions are given in centimeters. (2) for the parallel plate line, capacity is given per centimeter width normal to the section shown.*
This assumption is obviously rigorous in certain limiting cases and its general validity has been examined by comparing results obtained using it with direct numerical computations. Table VII shows typical cases. It will be seen that agreement is within 10%.

3. Proximity Effects between Neighboring Discontinuities

Often, to anchor a support bead, for example, double step discontinuities such as shown in Fig. 22 are created. If the distance between them is short the fringing fields generated at the steps interact in such a way as to decrease the effective discontinuity capacity at each step from the value calculated in isolation. This may be taken into account by multiplication with a proximity factor $P$.

![Fig. 22. Proximity factor for double step in parallel plate line.](image)

Table VIII

<table>
<thead>
<tr>
<th>Proximity ratio</th>
<th>Parallel plate line</th>
<th>Coaxial line ($T=6$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inner step</td>
<td>Outer step</td>
</tr>
<tr>
<td>0.0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.2</td>
<td>0.604</td>
<td>0.586</td>
</tr>
<tr>
<td>0.4</td>
<td>0.873</td>
<td>0.858</td>
</tr>
<tr>
<td>0.6</td>
<td>0.969</td>
<td>0.960</td>
</tr>
<tr>
<td>0.8</td>
<td>0.996</td>
<td>0.991</td>
</tr>
<tr>
<td>1.0</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*Note: See Fig. 22 for notation.*
Using the notation shown in Fig. 22, a reasonably extensive table of proximity factors has been computed for neighboring discontinuities in parallel plate line and these are shown in that figure. The error involved in applying these to discontinuities in coaxial line—where the radius of curvature of the conductors is no longer infinite—has been investigated. "Corresponding" discontinuities in parallel plate and coaxial line having a diameter ratio of 6 are compared in Table VIII. It is evident that there is negligible difference to the proximity factor whether the step is cut in the inner or outer conductors, and that the divergence from the parallel plate case does not exceed 3%.

C. BEAD SUPPORTS IN COAXIAL LINE

1. Simple Compensation Schemes

The most primitive possible bead support in coaxial line is the unadorned dielectric spacer (Fig. 23a). Since this introduces in cascade with the line a
section of incorrect characteristic impedance, severe reflections can be expected at all but the lowest frequencies unless special spacing layouts, which are, of course, frequency sensitive, are used [63].

An increasing stage of development shown in Fig. 23b is to undercut or overcut the line conductors (or do both if axial anchorage is required) to maintain continuity of characteristic impedance between the line section within the bead and those abutting [64], that is, to fulfill the relationship

$$b_1/a_1 = (b_2/a_2)^{k_s-1/2}$$  \hspace{1cm} \text{(47)}$$

where \(b_1, b_2\) are the outer conductor diameters without and within the bead, respectively, and \(a_1, a_2\) are the inner conductor diameters, similarly subscripted.

This helps considerably but it cannot eliminate the problem as the beads are now shunted at each end by the discontinuity capacities introduced by the under- or overcutting. Worse, since the fringing fields generated in this way occur almost entirely within the bead material, these capacities are multiplied by its dielectric constant.

As a further step in minimizing reflections the whole bead may be considered as a network which, provided that it remains electrically short, may be broad-
banded by increasing the depth of cut into the conductors to obtain enough extra inductance to offset the discontinuity capacity, that is, the characteristic impedance throughout the bead is increased slightly. An exact design may be achieved by trial and error using the data given in Table VI. In practice a better average performance can be obtained over a wider band by using an undercut somewhere between this and that demanded by Eq. (47).

The relative merits of each of these procedures can be seen in Fig. 24, which shows curves of VSWR as a function of frequency resulting from their application to a particular basic bead support structure. It will be agreed that except at low frequencies all leave much to be desired in their attainable performance.

2. Beads with Undercut Faces

In both cases above where some sort of compensation has been tried to limit the effects of the discontinuity capacities, it will be observed that this has been attempted "on the average" over an extended region about the point of reflection rather than at that point itself. It is this that is the basic weakness of both schemes.

Of the better methods which have been developed for bead compensation two of the most effective, both of which use basically the same principle, are shown in Fig. 25. While neither quite achieves compensation exactly at the point of discontinuity it is apparent that the region over which it is attempted is now much smaller. This type of compensation has been studied empirically in great detail by Kraus [65] but it does not yet appear to have had much mathematical analysis. A numerical treatment has been given by the finite difference method and the results compared with Kraus' experimental work. The agreement has been found to be very good [4].
If it can be assumed that the beads are sufficiently long for their end regions to be considered isolated the problem is then one of compensating the junction between two semiinfinite lines of equal characteristic impedance, one of which is dielectric filled. The fact that the problem can be treated in this way at once serves to emphasize one of the basic advantages of this method of compensation: thin, and therefore mechanically weak, beads are not needed to obtain broad banding.

It will be assumed that the transition region may be represented by a single shunt capacitor at the plane of the step followed by a very short length of unmatched transmission line, giving the equivalent circuit shown in Fig. 26.

![Fig. 26. Equivalent circuit for one face of stepped dielectric bead.](image)

Table IX

<table>
<thead>
<tr>
<th>Stepped conductor</th>
<th>Undercut (δ) (inch)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inner</td>
<td>0.036</td>
</tr>
<tr>
<td>Outer</td>
<td>0.056</td>
</tr>
</tbody>
</table>

*Note: These data apply for a 50-Ω line with a maximum outer conductor diameter of 1 in. and a bead material of dielectric constant 2.05.*

The object is to determine the cutout depth δ in the dielectric so that the region between the reference planes AA' and BB' will have a zero frequency image impedance equal to that of the adjoining transmission lines.

By terminating the lines with magnetic planes spaced about a diameter on each side of the reference planes the problem is readily handled numerically [4]. The machine calculates the total capacity of the lines and by deduction of the capacities of the terminal lines (assumed unperturbed) gives the net capacity C of the junction region (transmission line component plus dis-
continuity component) between $AA'$ and $BB'$. The zero frequency image impedance of the junction network is therefore

$$Z_{i0} = \left(\frac{L}{C}\right)^{1/2}$$

(48)

where $L = (\delta/2\pi)\ln(b/c)$ is the inductance of the line between $AA'$ and $BB'$, $b$ and $c$ being the inner and outer conductor diameters, respectively, within the bead. By computing $Z_{i0}$ for several values of $\delta$ it is easy to interpolate the depth which gives impedance continuity. Table IX summarizes the conditions for match in a 50-$\Omega$ line where the dielectric is Teflon ($\varepsilon_r = 2.05$) and Fig. 27 shows curves of junction performance against frequency. Compensation on the zero frequency image impedance basis is essentially what Kraus [65] attempts, since the type of frequency response sought in his experimental work is consistent with the adoption of this criterion.

It is interesting to make a comparison of these results with Kraus' experimentally derived design curve for a compensated stepped inner conductor, which is reproduced in Fig. 28. The agreement is seen to be very good. A rather surprising conclusion of Kraus' study is that the normalized offset ($\delta/D$) is independent of the characteristic impedance over the range 50 to 75 $\Omega$ (it could be greater, but the experimental work was not extended beyond this range), being a function of the dielectric constant of the bead material only. While all numerical results which have been computed have tended to show the same good agreement with experimental work they have not been sufficient in number for it to be claimed that they completely confirm this conclusion. There is, however, a strong suggestion that it is substantially correct.

It may be noted that reasonable success has been obtained in designing compensated junctions, with either a discontinuous inner or outer conductor, with a simple semiempirical method of analysis. The problem in any intended
analysis is to obtain a value for the discontinuity capacity at the face of the bead in the presence of a profiled dielectric which does not contain all the fringing flux. A method of approximation is to use the discontinuity capacities given in Table VI and seek to find an effective dielectric constant by which to multiply them.

It has been found that a dielectric offset depth not in error by more than 10% may be obtained by using an effective dielectric constant of

$$k_e' = \frac{k_e^{3/2}}{k_e + k_e^{1/2} + 1}$$

(49)

where $k_e$ is the dielectric constant of the bead material. This has been tested up to $k_e = 4$ [66]. It will be noted that this is the dielectric constant of a uniform medium which, if placed in the region between the reference planes $AA'$ and $BB'$ in Fig. 25 would give the same capacity as the existing partial dielectric filling.

The region between the reference planes is equivalent to a length of line

$$1 = \varphi(LC)^{1/2}$$

(50)

This formula is rigorous only for a uniform distribution of series inductance and shunt capacitance but it is a good approximation where the equivalent length is electrically small.

D. Butt Transitions in Coaxial Line

A common requirement in coaxial systems is a low VSWR connection between a line of one size and another of the same characteristic impedance. This requirement may arise, for example, in the connection of test equipment brought out to type N sockets to large rigid coaxial cable runs.
Although a taper transition may be used this is usually bulky and is relatively
difficult to machine. A possible alternative is the offset butt joint shown in
Fig. 29. The design requirement is to make the offset between the steps in the
inner and outer conductors give sufficient inductance to compensate the excess
capacity of the inhomogeneous field region.

By electrostatic means it is not possible to compute the individual dis-
continuity capacities occurring at each end of the junction region, but it is
relatively easy to determine the total capacity of the junction. This is sufficient
for design even though it is not now, as was the case with the bead support
problem, possible to compute the behavior of the transition with frequency
without some assumption on the distribution of the reactances in the network.

The effective length of the junction region can be computed from Eq. (50)
using the appropriate value of phase velocity.

The butt transition has been studied experimentally by Kraus [67], who
has produced a series of excellent design curves for lines of 50, 60, and 75 Ω
in which the dielectric medium in the transition region is air. These are re-
produced in Fig. 30. In view of this and the large number of possible combina-
tions of standard lines and common dielectric materials, it has not been
thought practical to attempt to provide generalized data. The program may,
however, be used to advantage in individual problems, an example of which
will be given.

It was desired to construct a matched butt transition in a 50-Ω line under-
going a 2:1 step in diameter, the transition region to be filled throughout with
Fluon dielectric \( k_e = 2.0 \). Leading dimensions are shown in Fig. 29. Offsets
increasing in steps of 0.0075 in. from 0.015 in. to 0.060 in. were tried and the
image impedance of the transition computed in each case. By interpolation
it was found that an offset of 0.045 in. preserves continuity of impedance. This
transition was constructed experimentally by Pyle [68], who obtained an
optimum match also at an offset of 0.045 in.; the agreement between theory
and practice is, therefore, excellent.

It is of interest to note that without the dielectric filling the problem would
have been equivalent to designing a matched transition in a 71-Ω line. Although Kraus’ experimentally derived curves do not include one for this impedance it is within the range of his work, allowing a value to be interpolated. This gives an offset of 0.046 in. This is typical of the excellent agreement that was obtained in a number of cases between Kraus’ work and the numerically computed answers.

Fig. 30. Optimum dimensions for a butt transition in coaxial line.

E. Capcitive Gaps in Coaxial Line

1. Numerical Solution

An obstacle similar to that shown in Fig. 31a finds common use in microwave bandpass filter design. A gap cut in the center conductor of a coaxial line normal to its axis essentially introduces series capacity into the line. A number of such gaps may be used to couple together intervening sections of transmission line to form a filter. To design these filters without the need for time consuming laboratory trial-and-error methods demands a knowledge of the network properties of the gap region.
Although the main purpose of the gap is always to introduce series capacity, between reference planes $T$ and $T'$ in Fig. 31a its complete representation requires the $\Pi$ equivalent network as in Fig. 31b. It is necessary to determine the series and shunt arms of this equivalent network. If the dimensions of the gap region remain small compared with the wavelength of RF excitation a quasistatic solution is acceptable, that is, the problem may be treated electrostatically as follows.

Consider the circuit shown in Fig. 31c in which a length of coaxial line is terminated in a "half-gap," that is, a gap bisected by the plane $AA'$ midway between $T$ and $T'$. Denoting the series and shunt arms of the equivalent network of the complete gap by $C_1$ and $C_2$, respectively, if a perfect shorting plane is inserted at $AA'$ the line section is effectively terminated to ground through a capacity $2C_1 + C_2$. Removal of the shorting plane and substitution of an open circuit plane (magnetic conductor) leaves the line grounded through a capacity $C_2$.

To compute these capacities the total capacity of a length of line terminated alternately as proposed above is calculated. To obtain valid answers the length of the section must be great enough to ensure that the disturbance to the normally purely radial field has become insignificant. Some preliminary runs have shown that a line length equal to or greater than the diameter of the outer conductor is enough to ensure this. The capacity of an "undisturbed" section (that is, one in which the gap is neglected) of this length is also determined.

Deduction of the open circuit capacity from the short circuit capacity leads to the equivalent series arm, while deduction of the undisturbed component from the open circuit capacity yields the shunt arm. Capacities of gaps of varying width in a range of lines of varying diameter ratios have been calculated and the results for the series and shunt arms are given in Table X; both are normalized in terms of capacity per unit outer conductor circumference (pf/cm). This table is thought to cover most usual design needs.

2. Comparison with Small Aperture Theory

A related problem to the above has been given an approximate theoretical treatment by Marcuvitz [69] using "small aperture" theory in which the
### Table X
SERIES (C1) AND SHUNT (C2) ARMS OF EQUIVALENT NETWORK FOR GAPS IN COAXIAL LINE

<table>
<thead>
<tr>
<th>Diameter ratio</th>
<th>10:9</th>
<th>4:3</th>
<th>5:3</th>
<th>2:1</th>
<th>2.3:1</th>
<th>3:1</th>
<th>5:1</th>
<th>7:1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap ratio</td>
<td>C1</td>
<td>C2</td>
<td>C1</td>
<td>C2</td>
<td>C1</td>
<td>C2</td>
<td>C1</td>
<td>C2</td>
</tr>
<tr>
<td>0.05</td>
<td>0.367</td>
<td>0.0354</td>
<td>0.275</td>
<td>0.0143</td>
<td>0.188</td>
<td>0.00823</td>
<td>0.138</td>
<td>0.00610</td>
</tr>
<tr>
<td>0.075</td>
<td>0.238</td>
<td>0.0486</td>
<td>0.183</td>
<td>0.0206</td>
<td>0.127</td>
<td>0.0120</td>
<td>0.0946</td>
<td>0.00893</td>
</tr>
<tr>
<td>0.10</td>
<td>0.173</td>
<td>0.0598</td>
<td>0.136</td>
<td>0.0265</td>
<td>0.0960</td>
<td>0.0156</td>
<td>0.0719</td>
<td>0.0116</td>
</tr>
<tr>
<td>0.15</td>
<td>0.106</td>
<td>0.0767</td>
<td>0.0858</td>
<td>0.0366</td>
<td>0.0623</td>
<td>0.0221</td>
<td>0.0474</td>
<td>0.0166</td>
</tr>
<tr>
<td>0.20</td>
<td>0.0718</td>
<td>0.0890</td>
<td>0.0598</td>
<td>0.0450</td>
<td>0.0443</td>
<td>0.0277</td>
<td>0.0340</td>
<td>0.0210</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0516</td>
<td>0.0985</td>
<td>0.0436</td>
<td>0.0520</td>
<td>0.0328</td>
<td>0.0327</td>
<td>0.0254</td>
<td>0.0248</td>
</tr>
<tr>
<td>0.30</td>
<td>0.0383</td>
<td>0.106</td>
<td>0.0328</td>
<td>0.0579</td>
<td>0.0249</td>
<td>0.0369</td>
<td>0.0194</td>
<td>0.0281</td>
</tr>
</tbody>
</table>

*Entries in this table are in picofarads per centimeter of outer conductor circumference.*
principal mode under RF excitation is treated correctly and all higher order modes are approximated. The geometry of this problem, which consists of a coaxial line closed at one end with a shorting plate from which the inner conductor is shorter by a gap width $s/2$, is shown in Fig. 32. Referred to the reference plane $T$, the inner conductor is shown to be terminated to ground through a capacity

$$C = \frac{\pi a^2 \varepsilon}{2s} + 2a\varepsilon \ln \frac{b-a}{s}$$

(51)

This formula is said to be valid under the restrictions

$$\lambda \gg b-a$$

(52a)

$$s \ll b-a$$

(52b)

and is readily seen to be composed of two distinct parts, a portion giving the parallel plate capacity between the face of the inner conductor and the cover plate and a “fringing” term.

It will be clearly seen that $C$ is to be equated with the group $2C_1 + C_2$ and, while owing to the approximate nature of Eq. (51) no direct check is possible, a useful comparison may be made. This shows that for small gaps ($s/a < 0.1$) the agreement is good even for a line diameter ratio of 10:9, which clearly violates inequality (52b), the error in this case being only 5%. For a 50-$\Omega$ line it is less than 1.5%.

For very small gaps the computation time for a numerical solution climbs steeply even on a large digital machine (IBM 7090), as to obtain reasonable accuracy it is necessary to maintain an adequate number of nodes in the gap region, and this in turn leads to very large over-all problems. Use may be made of the close agreement between the numerical work and the approximate

![Fig. 32. Marcuvitz problem. (a) Cavity with foreshortened inner conductor; (b) equivalent circuit.](image)
theory to cover these cases. For very small gaps the shunt component may be taken as negligible for many purposes and Eq. (51) considered as giving the series arm [70]. Where an estimate of the shunt component must be obtained it can be interpolated from Table X, as it must go to zero at zero gap.

F. CAPACITIVE GAPS IN STRIP LINE

1. Approximate Numerical Solution

The equivalent circuit of a gap in the inner conductor of strip line will now be considered. For a strip line with a single, zero thickness, inner conductor,

shown in Fig. 11a, this problem has been considered by Oliner [71] on the basis of an approximate parallel plate model, having an inner conductor widened sufficiently to take up the fringing capacity at the strip edges as equivalent parallel plate capacity. An analogy can then be drawn with the E plane slit coupling of rectangular waveguides and the solution written down by con-
sulting Ref. [72], the equivalent circuit referred to reference planes at the gap edges being a \( \Pi \) network of capacitors.

To illustrate the use of the numerical approach to this problem attention will be given in this section to a gap in the inner conductor of the strip line cross section shown in Fig. 11b [73]. This example is useful not just for the data presented but because it shows the shortcomings of the purely two-dimensional finite difference method.

As with a coaxial line and other forms of strip line, dielectric supported strip line bandpass filters often rely on series capacitive gaps to provide coupling between adjacent resonators, and for design purposes it is clearly necessary to have an accurate knowledge of the network properties of these gaps. On the basis of the edge equivalent circuit for a gap in strip line with a single inner conductor [71] it is reasonable to assume that gaps in dielectric supported strip lines may also be represented by a \( \Pi \) equivalent network of capacitors. The problem is to find the elements of this network as a function of gap width.

When, as must always be the practical case, the width of the inner conductor is finite, the fringing field in the gap region needs three coordinates for its complete description. However, if, following Oliner's attack, the line is considered for mathematical convenience to be part of an infinitely wide flat "sandwich" the field in the gap becomes two dimensional and is amenable to solution by the numerical methods thus far described.

Consider the structure depicted in Fig. 33 in which a gap is sawed through the inner conductor normal to the direction of propagation. The computation of the arms of the equivalent network for unit width of this line is carried out in exactly the same way as in coaxial line in Section IV, E, I by considering an electric and then a magnetic conducting plane to be introduced alternately into the gap center plane. As before an attempt may then be made to apply these data to a practical line by proportionate scaling using an equivalent parallel plate model with an inner conductor widened to take account of the fringing capacity of the actual strip. In this way the arms of a capacitive network representing a gap in 50-\( \Omega \), Rexolite 2200 supported line are derived; these are plotted as a function of gap width in Fig. 34.

2. Experimental Determination

While this method can give an approximation to the answer, more accurate values are required to achieve useful filter designs. Values for the 50-\( \Omega \) Rexolite supported stripline have been derived by experimental measurement, in S band, of the responses of single cavity filters [73]. The results are shown superimposed on Fig. 34.

In comparing the theoretical and experimental results it will be seen that for the more important series component the agreement is only rough, an
Fig. 34. Gap capacities for sawed capacitive gap in 50-Ω strip line on Rexolite 2200 card.

effective strip width nearer the actual physical width giving better answers. However, a useful agreement appears in the shunt component for which the data have been smoothed by a least squares process [73]. Nonetheless, in both cases experimental and theoretical data are seen to exhibit similar trends.
Multicavity filters designed using the experimentally derived data have been found to have accurately predictable responses and considerable confidence can be expressed in their accuracy. The failure of the approximate theory to yield good results is no doubt due to its crudity. However, it will be seen that there is no equivalent width of inner conductor which is capable simultaneously of reconciling series and shunt components of capacity. This suggests the need to operate numerically in three dimensions.

V. CONCLUDING REMARKS

This chapter has shown with a number of examples how numerical methods based on the philosophically very simple finite difference theory, much of which is of comparatively long standing, may be used to obtain accurate answers to many microwave problems involving TEM-mode lines which otherwise cannot be solved exactly at all, or are tractable only after involved mathematical development. There seems to have been little effort to apply purely numerical methods in the microwave field [74] but this can be expected to change under the impact of the large digital computer.

The finite difference method, which has been shown to be capable of considerable generalization within the framework of the one machine program, requires virtually no preliminary analysis for setting up. The principal disadvantage is the rather large amount of computer time consumed but, as has been shown by suitable comparisons, this defect is probably inherent in almost any purely numerical method of solution. If it were not so, there would scarcely be any profit in analysis, at least from a purely engineering viewpoint.

The numerical treatment has the further disadvantage that, up to the present, it has not been possible to perform other than analysis. This often leads to the necessity to interpolate through several results to obtain the answer sought. At the same time this weakness is shared with many of the mathematical methods.

The last example given on series gaps in strip line points to the desirability of extending into three dimensions (using, say, a seven point operator). Further thought would then need to be given to the methods to be used to solve for the potential field, as comparatively simple problems would then involve the possibility of sufficient nodes to overrun the high speed store capacity of even the most modern machines.

It requires little imagination to see other ways in which these basic ideas may be developed. The application to a machine of certain of Kron’s ideas [75] on the network representation of Maxwell’s equations should open the way to a solution of electrodynamic problems. Extension to cover other than TEM-mode devices would then appear possible.

Finally, while the computation time for a finite difference or other numerical
solution may be objectionable in some instances, these methods are admirably suited to take advantage of all future advances in computer technology even without themselves undergoing further refinement, which is unlikely. It should not be too much to hope for a speedup of at least an order.

ACKNOWLEDGMENTS

The author wishes to thank the principals of the firm of Rhode and Schwarz, Germany, and the editors of the IEEE Transactions on Microwave Theory and Techniques and the journal Microwaves for permission to use copyrighted material. A special debt of gratitude is owed to C. T. Carson, who not only carefully read the manuscript and made many valuable suggestions but was always ready to shore up flagging enthusiasm. The author also wishes to thank B. P. McDowall, who has materially assisted him with help and advice on the programming aspects of this work and has contributed a number of sound suggestions. Special thanks are also due to the author's wife for help in preparation of the manuscript and to Miss Raelene Golding for her careful typing of it. Finally the author wishes to thank the Australian Department of Supply for their support of this work and the Chief Scientist for permission to publish it.

REFERENCES

Harry E. Green


Author Index

Numbers in parentheses are reference numbers and indicate that an author’s work is referred to although his name is not cited in the text. Numbers in italic show the page on which the complete reference is listed.

A

Aarons, M. W., 4(9), 38
Adler, R., 90(4), 162
Adler, R. B., 95(17), 162
Allen, C. M., 95(19), 162
Allen, J. W., 70(37), 87
Anderson, A. J., 47(1), 86
Anderson, G. M., 349(23), 350, 391
Anderson, W. W., 25, 40
Anditi, M., 354(35), 365(54), 391, 392
Aron, R., 97(28), 163
Assadourian, F., 365(53), 392

B

Baird, J. R., 16, 39
Bandler, J. W., 25, 40
Baranowski, J. J., 74(50), 75(50), 88
Barber, M. R., 32, 33, 40, 69(35), 87
Barnell, E., 99(36), 163
Barnes, C. E., 99(32), 163
Barnes, S. H., 47(6), 86
Barrett, R. M., 354(33), 391
Batdorf, R. L., 71(44), 72(44), 84(58), 85(58), 87, 88
Bertram, W. J., 99(32), 163
Bisbee, D. L., 238(5), 324
Black, K. G., 365(52), 392
Blake, C., 90(10), 162
Blicher, A., 4(10), 38
Bloomberger, N., 89(1), 161
Bode, H. W., 222, 236, 241(10), 243, 324, 325
Boldlajan, J. T., 261(18), 263(18), 271(20), 310(34), 325, 326
Boothroyd, A. R., 173(7), 226(7), 227(7), 235
Bowles, L. W., 90(10), 162
Brandi, A., 16, 18, 39
Brand, F. A., 74(50), 75(50), 88
Breitner, D. J., 32(55), 33, 40
Bricker, C. H., 58(17), 61(17), 86
Brinton, J. B., Jr., 144(59), 164
Brown, G. S., 3(5), 38
Burckhardt, C. B., 231, 232, 236
Burrus, C. A., 8, 12(17b), 19, 29, 35, 38, 39, 40, 41, 62(22), 64(27), 74(49), 86, 87, 88
Bush, R. R., 31(53), 40
Butcher, P. N., 69(34), 87
Butzien, P. E., 1(3c), 38

C

Cambrell, G. K., 352(29), 391
Carley, D. R., 47(3), 86
Carlson, R. O., 122(48), 123, 163
Carre, B. A., 341, 390
Carter, P. S., 261(18), 263(18), 271(20), 325
Casterline, E., 6(12), 18(26), 28(26), 38, 39
Caulton, M., 47(8), 86
Chang, K. K. N., 1(3, 3a, 3b), 24(3), 34, 35, 38, 40, 41
Chase, P. E., 35, 41
Chen, T. S., 349(24), 351(24), 353(24), 391
Chen, Y. S., 70(37), 87
Chisholm, R. M., 355(39), 368(39), 391
Chisholm, S. H., 31, 36(50a), 40
Christensen, B., 31, 40
Churney, M., 146(24), 147(24), 162
Chynoweth, A. G., 70(36), 71, 82(54), 83 (54), 87, 88
Coale, F. S., 238(1, 2), 324
Cohen, B. G., 71(42), 73(42), 74(42), 87
Cohn, S. B., 238(2), 261(18), 263(18), 271(20), 311(35), 324, 325, 326, 347(19), 354(32), 355(32, 37), 356(41), 358, 360, 361(41), 391, 392
Collin, R. E., 328(1), 344(1), 345(15), 355(1), 356(1), 358(1), 390, 391
Comstock, R. L., 128, 129(13), 131, 132, 133(52), 153, 155, 163, 164
Cook, J. S., 1(3c), 38
Cooke, H. F., 47(1), 86
Cooper, H. W., 358(45), 392
Coriell, A. S., 70(36), 87
Cowan, M. J., 99(31, 32), 163
Cristal, E. H., 254(16), 272(22, 24), 277(24), 284(24), 287(24), 291, 292(24), 296(25), 297, 298(22), 301(22), 307(30), 31, 32, 312(22), 314(30), 318(31), 325, 326, 361 (48), 364, 392
Cruzan, O. R., 350, 391
Culver, R. C., 347(21), 391

D
Damiano, R., 134(55), 158, 164
Dansky, A. H., 20, 39
Davidson, L. A., 25(40), 39
Davidson, U. S., 4(9), 38
Davis, J. A., 198, 201(10), 202, 203(10), 235
Davis, R. E., 6, 21(12b), 30(12b), 38, 63(23), 87
Degen, W. von, 17, 39
De Jager, J. T., 90(11), 97, 98(22), 155, 162
De Loach, B. C., 57(20), 71(42), 72(45), 73, 74(42, 45), 86, 87
Diamond, B. L., 227, 230, 236
Dishal, M., 301(27), 326
Dow, D. G., 69(30), 87
Duncan, J. W., 356(42), 361(42), 392

E
Early, J. M., 48, 49, 86
Edson, W. A., 15(21), 39
Ehrenreich, H., 70(38), 87
Eisele, K. M., 146(24), 147(24), 162
Eldred, W. N., 355(38), 391
Elward, J. P., 90(12), 146(12), 152, 162
Eng, S. T., 35(62), 41
Engelmann, H. F., 365(51), 392
Erkan, G., 47(6), 86
Ernst, R. L., 47(8), 86
Esaki, L., 1(1), 4(8), 9(1), 38, 62, 86

F
Fay, C. E., 128, 129, 130, 131, 132, 133(52), 153, 155, 163, 164
Feist, W. M., 22(35), 39
Ferris, W. R., 31(53), 40
Fink, H. J., 90(9), 95(9, 20), 96(20), 118(43), 162, 163
Flammer, C. C., 261(18), 263(18), 271(20), 325
Fledermann, W. L., 70(36), 87
Follmer, W. C., 35(61), 40
Fortgang, M. M., 19(29), 39
Foster, K., 357, 392
Fox, L., 370(10), 341, 348, 390, 391
Foyt, A. F., 68(32), 69(32), 87
Francken, J. C., 332(7), 390
Frisch, J. T., 25(41), 39
Fromm, W. E., 354(34), 391
Fukada, M., 97(25), 162

G
Gallagher, R. D., 28(44b), 40
Gambling, W. A., 31, 34, 36(50b), 40
Gandy, R. W. H., 343, 390
Garbrecht, K., 106, 112, 113, 163
Garver, R. V., 350, 391
Getsinger, W. J., 3(4a), 24, 38, 39, 261(18), 263(18), 271(20), 307(30), 312(36), 314 (30), 325, 326, 355(36), 358, 361(36), 363(36), 391
Gibbons, G., 6, 21(12b), 30(12b), 38, 63(23), 87
Gifford, W. E., 134(57), 164
Glicksman, R., 3, 4(10), 6(12), 11, 38
Goldman, R. L., 35
Grayzel, A. J., 205, 215, 218(19), 221(19), 224(19), 225(19), 235, 236
Green, H. E., 329(4), 349(25), 351(25), 357 (4, 44), 358(4), 360(4), 377(4), 378(4), 380(66), 387(73), 388(73), 390, 391, 392, 393
Greenspan, M., 190, 192(8), 235
Grieg, D. D., 365(51), 392
Griemsmann, J. W. E., 346(17), 391
Grimes, E. S., Jr., 146(65), 155(65), 156(65), 164
Grinter, L. E., 329(2), 389(74), 390, 393
Gudikunst, J. B., 144(60), 164
Guillemot, E. A., 241(8), 242(11), 243(13), 244(14), 246(15), 324, 325
Gunn, J. B., 67, 69, 82, 83, 87, 88

H
Hakki, B. W., 68(31), 69(31, 33, 35), 87
Hamasaki, J., 26, 28(44a), 40
Hanson, D. C., 90(9), 95(9), 146(24), 147 (27), 162
Hauer, W. B., 7, 19, 38
Haus, H. A., 95(17), 162
AUTHOR INDEX

397

Hayt, W. H., 355, 391
Hedderly, D. L., 207, 209(15), 326
Heffner, H., 90(5), 162
Heilmeyer, G. H., 34(59), 40
Heinlein, V. W., 97(26), 98(26), 162
Heinz, W. W., 133, 134(54), 164
Helgesson, A. L., 219(20), 236
Henoch, B. T., 25, 40, 97(29), 163
Herold, E. W., 31(53), 40
Herrmann, G. F., 90(6), 9, 97(21), 162
Higgins, T. J., 349(26), 365(52), 391, 392
Higgins, V. J., 74(50), 75, 88
Hilsung, C., 69(34), 70, 87
Hinding, H. J., 19(29), 39
Hines, M. E., 25, 39, 40
Holifield, R., 99(36), 163
Holonyak, N., Jr., 4(9), 38
Hornbostel, D. H., 19(31a), 39
Hutson, A. R., 70(36), 87
Hyltin, T. M., 365(55), 392

J

Irvin, J. C., 72(46), 87, 117, 118(44), 119, 163
Ishii, J., 286(24a, 24b), 325

K

Jamieson, H. W., 369, 392
Jasik, H., 368(58), 392
Jayaramon, A., 70(36), 87
Johnson, E. O., 48, 49, 50, 86
Johnson, K. M., 99(35), 163, 226(22), 227, 236
Johnston, R. L., 71, 72(45), 73, 74(42, 45, 51), 75, 87, 88
Jones, E. M., 238(3), 240(7), 254(16a), 260(7), 261(18, 19), 263(16a, 18), 266(3), 268(3), 271(20), 299(26), 301(28), 312(7), 324, 325, 326
Josenhaus, J. G., 72(48), 88, 99(33), 101(37), 102, 146(23, 64), 162, 163, 164

K

Kach, A., 215, 236
Kahng, D., 115(42), 163
Kaminsky, G., 71(44), 72(44), 87
Kane, E. O., 3, 38
Kaylie, H. L., 232, 236
Kennedy, D. P., 119(45), 122(45), 163
Kim, C. S., 16, 18, 31(49), 32(49), 33, 39, 40

L

King, B. G., 10, 12(17a), 38
Klapper, J., 33, 40
Kleimack, J. J., 84(58), 85(58), 88
Kliphuis, J., 98(30), 99(30), 134(55), 158, 163, 164
Knight, S., 69(35), 87
Korn, H. A., 340(11), 390
Korn, T. M., 340(11), 390
Kotzebue, K. L., 19(31), 39, 90(5), 162
Kraus, A., 379, 381, 393
Krauss, H. L., 341(13), 390
Kron, G., 389, 393
Kuh, E. S., 97(25), 162
Kurokawa, K., 93(15), 125(51), 162, 163
Kvalrna, Y., 25, 40

M

Lance, D., 47(2), 86
Laybourn, P. J. R., 183(28), 184(28), 186(28), 189(28), 236
Lee, C. A., 71, 72, 84(58), 85(58), 87, 88
Lee, C. W., 30, 40
Leenov, D., 205, 235
Leeson, D. B., 191, 235
Lepoff, J. H., 29, 40
Lesk, I. A., 4(9), 38
Liegay, P. R., 95(19), 162
Lim, J. T., 24, 39
Loeb, L. B., 84(59), 88
Logan, R. A., 84(58), 85(58), 88
Lueck, A., 6, 38
Luettzenau, G. G., 47(6), 86

Marcatili, E. A., 238(5), 324
Marcuvitz, N., 346(16), 353(31), 369, 383, 387(72), 391, 392, 393
Marmiani, A., 6(12a), 38
Martin, D. W., 340(9), 390
Martin, L. T., Jr., 14, 38
Matthaei, G. L., 97(27), 162, 238(3), 240(7), 254(16, 16a, 16b, 17), 260(7), 261
AUTHOR INDEX

M
Melchior, H., 17(24a), 39
Meier, D. M., 3, 38
Mezger, P. G., 97(26), 98(26), 162
Minton, R. M., 3, 4(10), 11, 18(26), 28(26), 38, 39
Misaiva, T., 72, 74(52), 84(52), 88
Miyahara, Y., 4(8), 38
Moreno, T., 55, 86, 376(64), 392
Morin, F. J., 118, 120(46), 163
Morrow, W., 146(65), 155(65), 156(65), 164
Mosher, C. H., 69(30), 87
Motz, H., 329(3), 390
Mouw, R. B., 35(66), 41
Murphy, E. A., 58(14), 86

N
Nelson, D. E., 14, 16, 17(19), 18, 20(19), 28(26), 38, 39
Neuf, D., 99(36), 163
Newton, A., 33(57), 40
Nielsen, E. G., 25, 40

O
Oberhettinger, F., 350(28), 391
Okean, H. C., 30, 40
Okonogi, T., 637, 642
Okwit, S., 133, 134(54), 164
Oliver, A. A., 386, 387(71), 393
Ordung, R. F., 341(13), 390
Ozaki, H., 286(24a, 24b), 325

P
Page, C. H., 205, 236
Parker, G., 47(2), 86
Pearson, G. L., 70(37), 87
Penfield, P. Jr., 92(14), 125(50), 162, 163, 167, 169, 172(4, 5, 6), 176, 180, 185, 186, 190, 191, 211, 230, 235, 236
Peterson, D. G., 32(56), 40
Peterson, E., 89(2), 161
Posner, W., 58(14), 86
Prager, H. J., 34(59), 40
Presser, A., 18(26), 28(26), 29, 31(50), 32, 33(58), 34, 36(50), 39, 40
Prior, A. C., 82, 83, 88
Pucel, R. A., 9(17), 31, 32(48), 38, 40, 66, 87
Pyle, J. R., 357(44), 381, 387(73), 388(73), 392, 393

Q
Quist, T. M., 68(32), 69(32), 87

R
Rabinovici, B., 33(57), 40
Rafuse, R. P., 92(14), 125(50), 162, 163, 167, 176, 180, 185, 186, 190, 191, 211, 230, 231, 232, 235, 236
Ragan, G. L., 376(63), 392
Ramo, S., 6, 38
Read, W. T., 71, 72, 75, 76, 87
Reich, H. J., 341(13), 390
Renkowitz, D., 58(14), 86
Richardson, L. F., 347(20), 391
Ridley, B. K., 70, 87
Rimai, E., 365(53), 392
Ringenbach, M. E., 358(45), 392
Robbins, T. E., 369(62), 392
Robinson, B. J., 90(11), 162
Roswell, A. E., 18, 39
Roulston, D. J., 173(7), 206, 226(7), 227(7), 235, 236
Rowe, H. E., 91, 162
Rucker, C. T., 146(65), 155, 156, 164
Rulison, R. L., 58(17), 61(17), 86, 95(20), 96(20), 162
Ryder, E. J., 83(55), 88

S
Salzberg, B., 90(7), 95(19), 162
Sard, E. W., 90(7), 162
Scanlan, J. O., 24, 39, 183(28), 184(28), 186 (28), 189(28), 236
Schiffman, B. M., 261(18), 263(18), 271(20), 272(23), 277(23), 284(23), 286(23), 287(23), 297, 298(23), 307(31), 318(31), 325, 326
Schlapfer, M. F., 17(24a), 39
Schultz, W., 6(12a), 38
Schneider, M. V., 8, 21(15), 38, 64(27), 87
Schumacher, F. N., 35(66), 41
AUTHOR INDEX

Scott, R. B., 134(56), 164
Seidel, H., 94(16), 162
Sharpe, G. E., 10, 12(17a), 38
Sharpless, W. M., 95(18), 162
Shaw, R. C., 19(27a), 39, 64(27), 87
Shimizu, J. K., 261(18), 263(18), 271(20), 325
Shockley, W., 163
Shyam, M., 70(37), 87
Sibilia, J. T., 145(63), 164
Sie, J. J., 22(34), 39
Silverman, S. J., 122(48), 123(48), 163
Skalnik, J. G., 341(13), 390
Skiles, J. J., 349(26), 391
Slack, H. A., 122(48, 49), 123, 124, 163
Smolen, L. 1., 24, 39
Sobol, H., 47(8), 86
Solomon, A. H., 33(58), 40
Sommers, H. S., Jr., 1(2), 3(5), 38
Southwell, R. V., 329(5), 343, 347, 390
Southworth, G. C., 328(4), 324
Steinhoff, R., 8(14), 21(14), 25(14), 27, 38
Sterzer, F., 8(14), 14, 15(22), 16, 17(19), 20 (19, 32), 21(14), 25(14), 27, 31(50), 32, 33, 34, 36(50), 38, 39, 40, 47(4), 86
Stillwell, A. L., 146(24), 147(24), 162
Stratton, J. A., 364(50), 392
Strutt, M. J. O., 17(24a), 32, 39, 40
Swan, C. B., 58(16), 86

T
Tateno, I., 637, 642
Tee, G. J., 340(9), 390
Thim, H. W., 69(35), 87
Thompson, J. J., 353(30), 356(30), 326(30), 391
Trambarulo, R., 19, 29, 39, 40

U
Uenoahara, M., 58(17), 61(17), 69(35), 86, 87, 90(6), 8, 9, 12, 93(15), 94(16), 95 (9, 18), 146(12, 23, 24), 147, 152, 162
Uhlir, A., Jr., 90(6), 104, 115(41), 162, 163, 168, 205, 208, 210, 214, 225, 233, 236
Uhlir, A. R., 58(18), 86
Utsunomiya, T., 233, 236

V
van der Pol, B., 12, 16(23), 38, 39, 38, 39, 38, 39
Van der Ziel, A., 90(3), 161
Vane, A. B., 69(30), 87
Varettoni, L., 6, 38
Vincent, B. T., Jr., 99(34), 163

W
Wallmark, J. T., 20, 39
Watkins, T. B., 70, 87
Webster, R. R., 47(7), 86
Weinberg, L., 240(6), 241(9), 312(6), 324
Weinreb, S., 191, 235
Weller, D. B., 254(16), 325
Wenzel, R. J., 272, 325
Wheatley, G. H., 120(47), 163
Wheeler, G. J., 29, 40
Wheeler, H. A., 365, 368, 392
Whelan, J. M., 120(47), 163
Whinnery, J. R., 6, 38, 369, 392
Whooley, W. B., 355(38), 391
Wiegmann, W., 71(44), 72(44), 84(58), 85 (58), 87, 88
Wild, N. R., 360(47), 392
Willoughby, E. O., 332(8), 390
Wolf, H., 47(2), 86
Wolfson, R. I., 211, 212(17), 215, 216(17), 225, 236
Wright, M. L., 19(28), 30, 39
Wright, R. O., 35, 41
Wu, T. T., 346(18), 365, 391

Y
Yamashita, E., 16, 39
Yariv, A., 1(3c), 38
Youla, D., 24, 39
Young, D. T., 19, 39, 64(27), 87, 332, 390
Young, L., 238(3), 240(7), 254(16a), 260 (7), 261(19), 263(16a), 266(3), 268(3), 299(26), 301(28), 312(7), 324, 325, 326, 386(70), 393
Yuan, S., 233, 236

Z
Zener, C., 52, 86
Subject Index

A
Abrupt-junction diode, 115, 170, 190
Abrupt junction doubler, 180, 200
Accelerating factor, 340
Acceptor, 53
concentration, ionized, 115
Accuracy, 340, 349, 357, 364, 368
of numerical solutions, 348
Active circuit element, 12
Active converters, 33
Acoustical vibrations, 82
Admittance, input, 242
complementary, 246
Admittance, real and imaginary parts of the input, 243, 305
Admittance inverter, 311
Advancing to a finer net, 347
Air cooling, 139
Alloy dot, 6
Alloyed diode, 6
All-stop structure, 310
Alumina, 9
Aluminium-germanium alloy junction, 35
Ambient temperature, 95
Ammonia, 135
Amplifier, 21
broad-band, 9, 154
circulator-type, 92
cooled varactor parametric, 89, 164
degenerate, 92
integrated, 30
L-band, 150
liquid helium cooled, 152
liquid nitrogen cooled, 149
narrow band, 152
nondegenerate, 92
performances of cooled parametric, 157
reflection-type, 21
transmission-type, 21
tunnel diode, 21
X-band, 155
Amplifier impedance, 100
Analysis, orthonormal block, 349
Angle dielectric, 337
Annular surfaces, 342
Annulling network, 263, 305, 315, 322
reactance, 254
resonant frequency of, 262
susceptance, 251, 258
Annulling stub, drawing of, 320
Anomalous noise behaviour, 104
Approximations, small signal, 191
Area of diode junction, 21
Arsenic, 117
Atomic energy, 329
Attenuation, 262, 303
Attenuation constant, 346
Audio frequency, 35
Autodyne converter, 33
Avalanche breakdown, 56, 59
in gases, 84
Avalanche delay, 76
Avalanche multiplication, 51, 80
in solids, 83
Avalanche process, “snapshots” in time of, 77
Avalanche region, 75
Avalanche and transit time effects, power limitations of a dielectric varactor harmonic generator due to, 61
Avalanche transit time oscillators, 71
frequency limitations of, 78
Avalanching cathode, 76
Axis of symmetry, 342

B
Backward diode, 35
Balanced strip-line, 8
Balanced-varactor configuration, 98
Ball alloying, 63
Bandgap, 66
values, 63
Bandpass filter, 237, 254, 277, 298, 361
comb-line, 310
dielectric supported strip line, 381
interdigital, 272, 296, 363
stub-resonator, 254
Bands, conduction, 3
valence, 3
Bandstop filter, 238, 276, 297, 298
<table>
<thead>
<tr>
<th>Subject</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandstop filter—cont.</td>
<td>296</td>
</tr>
<tr>
<td>Bandstop filter designs for use in diplexers</td>
<td>284</td>
</tr>
<tr>
<td>Bandwidth, 24, 29, 32, 97, 159</td>
<td></td>
</tr>
<tr>
<td>maximum, 220</td>
<td></td>
</tr>
<tr>
<td>narrow to moderate, 299</td>
<td></td>
</tr>
<tr>
<td>optimum, 24</td>
<td></td>
</tr>
<tr>
<td>Bandwidth of circulator, 130, 134</td>
<td></td>
</tr>
<tr>
<td>Bandwidth in multipliers without idlers, 211</td>
<td></td>
</tr>
<tr>
<td>Bandwidth spreading, 319</td>
<td></td>
</tr>
<tr>
<td>Barrier, Schottky, 106</td>
<td></td>
</tr>
<tr>
<td>Barrier capacitance, 115</td>
<td></td>
</tr>
<tr>
<td>Barrier elastance, 115</td>
<td></td>
</tr>
<tr>
<td>Barrier layer, 101</td>
<td></td>
</tr>
<tr>
<td>Base resistance, 6</td>
<td></td>
</tr>
<tr>
<td>Base widening in transistor, 49</td>
<td></td>
</tr>
<tr>
<td>Bead supports, 372, 375, 376</td>
<td></td>
</tr>
<tr>
<td>Beads with undercut faces, 377</td>
<td></td>
</tr>
<tr>
<td>Beam lead, 63</td>
<td></td>
</tr>
<tr>
<td>Beam-lead diode, 30</td>
<td></td>
</tr>
<tr>
<td>Bend, mitered, 353</td>
<td></td>
</tr>
<tr>
<td>Bias, 32</td>
<td></td>
</tr>
<tr>
<td>forward, 198</td>
<td></td>
</tr>
<tr>
<td>negative, 125</td>
<td></td>
</tr>
<tr>
<td>reverse, 78</td>
<td></td>
</tr>
<tr>
<td>Bias charge, 177</td>
<td></td>
</tr>
<tr>
<td>Bias on varactor, 58</td>
<td></td>
</tr>
<tr>
<td>Bias voltage, 10, 111, 115, 122, 179</td>
<td></td>
</tr>
<tr>
<td>forward, 125</td>
<td></td>
</tr>
<tr>
<td>Biasing magnetic field, 130</td>
<td></td>
</tr>
<tr>
<td>Biasing resistors, nonlinear, 20</td>
<td></td>
</tr>
<tr>
<td>Bifurcated parallel plate line, 353, 356</td>
<td></td>
</tr>
<tr>
<td>Bode's relationship, 244, 247, 250, 295</td>
<td></td>
</tr>
<tr>
<td>Boron, 116</td>
<td></td>
</tr>
<tr>
<td>Breakdown, 85</td>
<td></td>
</tr>
<tr>
<td>dielectric, 80</td>
<td></td>
</tr>
<tr>
<td>Breakdown in gases, avalanche, 84</td>
<td></td>
</tr>
<tr>
<td>Breakdown limit, 180</td>
<td></td>
</tr>
<tr>
<td>Breakdown voltage, 118</td>
<td></td>
</tr>
<tr>
<td>reverse, 64, 196</td>
<td></td>
</tr>
<tr>
<td>Bridge doubler, 223</td>
<td></td>
</tr>
<tr>
<td>with extra diode to eliminate fourth harmonic, 225</td>
<td></td>
</tr>
<tr>
<td>Bridge-type doubler, 208</td>
<td></td>
</tr>
<tr>
<td>Broadband amplifier, 9, 154</td>
<td></td>
</tr>
<tr>
<td>Broadband matching networks, 24</td>
<td></td>
</tr>
<tr>
<td>Broadband receiver, 159</td>
<td></td>
</tr>
<tr>
<td>Broadband varactor doublers, 222</td>
<td></td>
</tr>
<tr>
<td>Burnout, 34, 36</td>
<td></td>
</tr>
<tr>
<td>Butt transitions in coaxial line, 380</td>
<td></td>
</tr>
<tr>
<td>Capacitance, 59</td>
<td></td>
</tr>
<tr>
<td>barrier, 115</td>
<td></td>
</tr>
<tr>
<td>charge on variable, 198</td>
<td></td>
</tr>
<tr>
<td>distributed, 280, 298</td>
<td></td>
</tr>
<tr>
<td>effective, 175</td>
<td></td>
</tr>
<tr>
<td>incremental, 198</td>
<td></td>
</tr>
<tr>
<td>junction, 2, 5, 31, 114</td>
<td></td>
</tr>
<tr>
<td>lumped, 310</td>
<td></td>
</tr>
<tr>
<td>mutual, 279</td>
<td></td>
</tr>
<tr>
<td>package, 2, 9, 122</td>
<td></td>
</tr>
<tr>
<td>stray, 122</td>
<td></td>
</tr>
<tr>
<td>zero bias, 155</td>
<td></td>
</tr>
<tr>
<td>Capacitance modulation factor, 95</td>
<td></td>
</tr>
<tr>
<td>Capacitance of varactor, 105</td>
<td></td>
</tr>
<tr>
<td>Capacitance variation, 108</td>
<td></td>
</tr>
<tr>
<td>Capacitance-voltage curve for varactor, 169</td>
<td></td>
</tr>
<tr>
<td>Capacities at the butt junction, discontinuity, 381</td>
<td></td>
</tr>
<tr>
<td>Capacitive gaps, 338</td>
<td></td>
</tr>
<tr>
<td>in coaxial line, 382</td>
<td></td>
</tr>
<tr>
<td>in strip line, 386</td>
<td></td>
</tr>
<tr>
<td>Capacitor, shunt, 252, 257</td>
<td></td>
</tr>
<tr>
<td>Capacity, 341, 347, 353</td>
<td></td>
</tr>
<tr>
<td>discontinuity, 374</td>
<td></td>
</tr>
<tr>
<td>fringing, 362</td>
<td></td>
</tr>
<tr>
<td>parallel plate, 354, 362</td>
<td></td>
</tr>
<tr>
<td>Capacity in coaxial line, discontinuity, 371</td>
<td></td>
</tr>
<tr>
<td>Capacity of gap, 383</td>
<td></td>
</tr>
<tr>
<td>Cardan's solution, 197</td>
<td></td>
</tr>
<tr>
<td>Carre method, 341</td>
<td></td>
</tr>
<tr>
<td>Carrier</td>
<td></td>
</tr>
<tr>
<td>density, 83</td>
<td></td>
</tr>
<tr>
<td>fluctuation, 101</td>
<td></td>
</tr>
<tr>
<td>freeze-out, 118</td>
<td></td>
</tr>
<tr>
<td>velocity, 52, 56, 71</td>
<td></td>
</tr>
<tr>
<td>saturation, 51, 81, 83</td>
<td></td>
</tr>
<tr>
<td>Carrier, lifetime of, 104</td>
<td></td>
</tr>
<tr>
<td>minority, 101</td>
<td></td>
</tr>
<tr>
<td>multiple transit, 104</td>
<td></td>
</tr>
<tr>
<td>single transit, 104</td>
<td></td>
</tr>
<tr>
<td>Cartesian coordinates, 329</td>
<td></td>
</tr>
<tr>
<td>Cartesian equation, 336</td>
<td></td>
</tr>
<tr>
<td>Cascading, 27</td>
<td></td>
</tr>
<tr>
<td>Cathode, avalanching, 76</td>
<td></td>
</tr>
<tr>
<td>Cavity, 17</td>
<td></td>
</tr>
<tr>
<td>conical, 8</td>
<td></td>
</tr>
</tbody>
</table>
Cavity—cont.
re-entrant, 18
Cavity resonators, 329
C-band, 27, 28
Center frequency, 303
Channel-separating unit, 294, 301
Characteristic cutoff frequency, 55
Characteristic impedance, 18, 22, 345
of square coaxial line, 351
of strip line, 357, 358
Charge, bias, 177
Charge, maximum, 78
Charge, stored, 202
Charge-capacitance-voltage relationships of varactors, 169
Charge density, 53
Charge excursion, 177
Charge-field dynamics of Gunn oscillator, 71
Charge, resistance and frequency in varactor doubler, 203
Charge-storage diode, 167, 205
multiplier, 205
Charge-storage effects, minority carrier, 62
Charge swing, peak-to-peak, 199
Charge on variable capacitance, 198
Charge-voltage characteristic, 205
Charge-voltage curve for varactor, 169
Chebyshev filter, 240, 250, 260, 308
low pass prototype, 256
Chebyshev response, 24, 29
Circuits, equivalent, 369
Circular rod inner conductor, 355
Circulator, 21, 22, 92, 99, 126, 151
bandwidth of, 130, 134
cooled, 144
cryogenic, 127, 130, 132, 134
dipolar mode in, 128
ferromagnetic resonance loss in, 130
insertion loss of, 134
internal magnetic field in, 130
isolation of, 134
liquid helium temperature, 133
low field loss in, 131
low temperature, 127, 130, 132, 134
rotating modes in, 129
switchable latching, 22
temperature insensitive, 134
Y-junction, 128
Circulator impedance, 100
Circulator mode, $Q$ of, 130
Circulator-type amplifier, 92
Civil engineering, 329
Claude cycle refrigerator, 138
Closed-cycle refrigerator, 90, 134, 138
Coarse net, 347
Coaxial circuit, 9
Coaxial filter, low pass, 267
Coaxial line, 17, 28
bead in, 375, 376
butt transitions in, 380
capacitive gaps in, 382
discontinuity capacity in, 371
series gap in, 383
square, 349
steps in, 369, 370
Coaxial plate line, discontinuity in, 373
Coaxial rectangular conductors, 349
Coaxial-to-strip line transition, 358
Cohn structure, 114
Column interface, 337
Comb-line filter, 310
Commercial refrigerators, 144
Compensated junctions, 379
Complementary diplexer, 246, 254, 263, 266
Complementary filters, 296
Complementary input admittance, 246
Complementary input impedance, 246
Complex frequency, 241
Complex permittivity, 346
Compressor, 141
Compression space, 139
Computed attenuation of three-channel multiplexer, 316
Computed diplexer performance, 261, 284
Computed potential field, 341
Computer diode, 72
Computing time, 352, 354, 356, 371, 389
Concentric prisms, 349
Conductance, conversion, 31
image, 32
incremental, 198
internal, 32
load, 32
Conducting surface, 364
Conduction band, 3, 66, 70
Conduction forward, 101
Conduction process, multivalley, 70
Conductivity, electrical, 143
thermal, 143
Conductor, inner, 354, 355
Conductor, magnetic, 351, 354
  thickness correction for strip line inner, 358
Conductor (series) losses, 346
Confidence, degree of, 348
Conformal transformation, 328, 349, 355
Conical cavity, 8
Conical waveguide resonators, 19
Connector, type N, 271, 380
Corrosion, 143
Constant resistance property, 238
Contact, liquid gallium-indium, 73
  ohmic, 83, 107
Contact potential, 5, 169, 198
Contact resistance, 6
Contiguous channels, multiplexers having, 305
Continuous dielectric media, 334
Convergence, 332, 340
Conversion conductance, 31
Conversion gain, 31, 33, 34
Conversion losses, 32, 226
Converter, active, 33
  autodyne, 33
  frequency, 21
  nonoscillating, 34
  passive, 32, 34
  self-oscillating, 33
  tunnel diode, 30
Cooled circulator, 114
Cooled parametric amplifiers, performances of, 157
Cooled varactor, 122
Cooled varactor parametric amplifiers, 89–164
Coordinates, Cartesian, 329
  cylindrical, 329
Coordinate system, 341
Corners, reentrant, 348
Corner point, 336
Corner region, 351
Counterflow heat exchanger, 138
Coupled strip lines, 356, 361
Coupling, high-impedance wire type, 311
  magnetic 310
Critical field, 52, 68
Crossover frequency, 305
Cryogenic circulator, 127
Cryogenic system design, 142
Cryogenic temperature, 95
Cryogenic temperature circulator, 134
Cryogenic value, 142
Cryostat, 147
Crystal, gallium arsenide, 120 (see also Gallium arsenide and Silicon)
  highly doped, 21
  silicon, 120
Crystal diode, point-contact, 30
Crystal mixers, 33
Crystal-mixer diodes, 35
Cubic approximation, 16
Cubic equation, 174, 223
Cubic I-V characteristic, 12
Current, displacement, 60
  forward, 96
  junction, 3, 14
  particle, 60, 77
  peak, 6, 26, 165
  reverse, 96
  transport, 60, 77
Current crowding, 46
Current generator, zero-admittance, 239
Current swing, 55
Current in transistor, maximum, 49, 52
Current in tunnel diode, maximum, 66
Cutoff frequency, 10, 21, 25, 123, 168, 272
  characteristic, 55
  effective, 169
  resistive, 11, 62
  transit time, 59, 62
Cutoff frequency (of filters), 250
CW operation of Gunn oscillator, 68
Cycle, Gauss-Seidel, 340
Cylindrical coordinates, 332, 329
Cylindrical equation, 336
Cylindrical geometries, 6
Cylindrical rod inner conductor, 363

D
Deferred approach to the limit, 332, 347
Degenerate amplifier, 92
Degree of confidence, 348
Delay, avalanche, 76
  drift, 78
Delay lines, 361
Density, doping, 3
  peak-current, 3
  valley-current, 3
Depletion layer, 5, 167, 169
Depletion layer capacity, 58
Depletion layer varactor, 58
Depletion region, 71
Derivative, normal, 343
Detectors, video, 35
Detuning of doubler, 216
Dewar, 90
Dewar system, open, 132
Die-drawn tubing, 143
Dielectric, 356
  angle, 337
  inhomogeneous, 328
  lossy, 346
  mixed, 357
  piecewise homogeneous, 330
Dielectric bead, 378
Dielectric breakdown, 49, 80
Dielectric constant, 5, 376
  effective, 380
Dielectric interface, 337
Dielectric media, continuous, 334
Dielectric supported shielded strip line, 357
Dielectric supported strip line bandpass filters, 387
Dielectric varactor, 58, 61
Diffused varactor, 103
  GaAs, 116
Diffusion potential, 115
Diffusion voltage, 169
Digital computer, 329
Diode
  abrupt-junction, 115, 170, 190
  alloyed, 6
  backward, 35
  charge-storage, 167, 205
  computer, 72
  crystal-mixer, 35
  etched, 6
  graded-junction, 115
  high-current, 21
  junction, 6
  maximum peak-current, 21
  pn junction, 115
  point-contact crystal, 30
  Read, 71, 75
  Schottky barrier, 30, 115, 159
  semiconductor-metal surface barrier, 115
  series, 226
  step-recovery, 167, 205
  tunnel, 62, 165
  varactor, 6
Diode junction, area of, 21
Diode multiplier, charge-storage, 205
Diplexer, 237
  1- to 4-Gc band, 266
  1.5-Gc, 301
  4- to 12.4-Gc, 271, 273
  bandstop filter designs for use in, 284
  complementary, 246, 254, 263, 266
  computed performance of, 261, 284
  interdigital filter for use in, 279
  measured response of, 263, 271, 303
  measured transmission, 270
  partly complementary, 251
  wide-band, 261
Diplexer junction, 268
Dipolar mode in circulator, 128
Directional couplers, 361
Directional filters, 238
Directivity, 28
Discontinuities, neighboring, 374
Discontinuity capacity, 374
  in coaxial line, 371
  at the butt junction, 381
Displacement current, 60
Dissipation, pump power, 106
Distributed capacitance, 280, 298
Divalent iron, 132
Dominant error, 343
Donor, 53
  concentration, ionized, 115
Doped crystals, highly, 21
Doping, 52, 60
  linearly graded, 170
Doping density, 3
Doping profile, 75, 95
Dot, alloy, 6
Double-side-band noise figure, 35
Double step in transmission line, 372
Doubler, 56
  abrupt junction, 180, 200
  bridge-type, 208
  broadband varactor, 222
  detuning of, 216
  efficiency of, 216
  graded-junction, 190, 200
Doubler with single tuned load, 211
Doubly terminated filter, 308
  foreshortened, 245, 250, 263
  lowpass prototype, 240, 322
Drift delay, 78
Drift field, 73
Drift region, 48
Drift space, 76
Drift velocity, 81
Drive level, 180, 202, 231
Duality principle, 348
Dynamic impedance, 94
Dynamic quality factor, 93, 105, 123, 156
Dynamic range, 23, 28, 36
Dynamics of Gunn oscillator, 71

Engine speed, refrigerator, 139
Entropy, 135
Epitaxial film, 106, 119
Epitaxial layer, 47, 73
Epitaxial pn junction, 72
Epitaxial technology, 45, 56
Epitaxial varactor, 103, 116, 121, 122
silicon, 121, 122

Equation
Cartesian, 336
cubic, 223
cylindrical, 336
integral, 364
and finite-difference equations method compared, 364
Johnson’s, 49, 55
Laplace’s, 328, 332
McKay’s, 51
Poisson’s, 60
quartic, 223

Equations
Euler-Lagrange, 345
finite, difference, 334
linear, 329, 332, 364
Maxwell’s, 389
scaled, 196
simultaneous, 332, 335, 364
special finite difference, 336
Equivalent circuit, 13, 21, 369
jumped-component, 166
Equivalent circuit of tunnel diode, 2, 13, 14, 23, 25, 30
Equivalent linear elements, 15
Equivalent network, 353
for gaps in coaxial line, 384
Equivalent temperature, 107
Etched diode, 6
Etching, 6, 119
electrolytic, 5
Euler-Lagrange equations, 345
Euler’s constant, 353
Evanescent higher order modes, 369
Even mode, 361
Excess noise, 103
Exchange gas, 114
Expansion-engine refrigerator, 137
Expansion space, 139
Experimental results on tunnel diode oscillators, 17
Experiments on tunnel diode converters, 33
Fabrication, 12, 21
Fermi level, 3
Ferrimagnet, 128
Ferrimagnetic material, 128
Ferrite disk, 129, 130
Ferromagnetic resonance loss, 128 in circulator, 130
Fiberglass, 143, 366, 368
Field-distance profile, 72
Figure of merit, 122
Films, epitaxial, 119
Filters,
bandpass, 237, 254, 277, 298
interdigital, 254, 272, 296, 298, 363
with stubs, 254
bandstop, 238, 276, 297, 298
with stubs, 272
Chebyshev, 240, 250, 260, 308
low pass prototype, 256
comb-line, 310
complementary, 296
dielectric supported strip line bandpass, 381
diplexer (see Diplexer)
directional, 238
doubly terminated, 308
low pass prototype, 240, 322
foreshortened, 323
doubly loaded, 245, 250, 263
highpass, 251, 254
interdigital, 254, 272, 296, 298, 363
lowpass, 251, 254, 272
coaxial, 267
lumped-element, 240
low-pass Chebyshev prototype, 280
maximally flat, 240
modified form of interdigital, 277
multiplexer (see Multiplexer)
parallel-coupled-resonator bandstop, 296
prototype, 252
singly terminated, 250, 308
low pass prototype, 239
stub-resonator bandpass, 254
stub-type bandstop 272
tuning of, 301
with $\lambda_0/4$ stubs and $\lambda_0/4$ connecting lines, design equations for, 264
Filter and interconnection, low pass, 257
Fine net, 347

Finger (contacting), 6
Finite difference equations, 334
special, 336
Finite difference theory, 329
Five point operator, 332
Fixed-frequency oscillators, 17
Flux linkage, 8
Folding process, 280
Forbidden energy states, 66
Foreshortened doubly terminated filter, 245, 250, 263, 323
Forward bias, 198
Forward bias direction of tunnel diode, 65
Forward bias voltage, 125
Forward conduction, 101
Forward current, 96
Forward-driven varactor multiplier, 198
Fourier coefficients, 33, 172, 186, 228
Freeze-out, carrier, 118
Frequency, 13, 28
audio, 35
center, 263, 303
characteristic cutoff, 55
complex, 241
crossover, 305
cutoff, 10, 21, 25, 123, 168, 250, 272
effective, 169
fundamental, 208
idler, 92
image, 31, 35
impedance level and, 50
intermediate, 30
low, 35
operating, 21, 25, 130
temperature dependence of, 128
pump, 92, 97
radian, 241
reactive cutoff, 11
real, 241
resistive cutoff, 11, 62
resonant, 262
self-resonant, 10, 11, 18, 21, 98, 155
sideband, 30, 97
signal, 92
spurious, 62
sideband, 97
temperature dependance of operating, 128
transit time cutoff, 59, 62
Frequency convertor, 21
| Frequency doubler, varactor, 147 | Graded junction multiplier, 233 |
| Frequency guard band, 305 | Gradient, impurity, 122 |
| Frequency of Read diode, optimum operating, 79 | Green's function, 356 |
| Frequency source, electronically swept, 301 | Ground plane mode, 271 |
| Frequency stabilization, thermal, 147 | Guard band, frequency, 305 |
| Frequency translation, 31 | Gunn effect, 67–71 |
| Fringing capacity, 356, 362 | Gunn oscillation, critical fields for, 70 |
| Fringing field, 310, 355, 374 | Gunn oscillator, 62, 67, 81, 166 |
| Fundamental oscillations of Gunn oscillator, 69 | charge-field dynamics of, 71 |
| Gain, 25, 29, 35 | CW operation of, 68 |
| Gain-bandwidth limitation, 97 | electron dynamics of, 71 |
| Gain-bandwidth product, 24, 28, 91, 98, 123 | fundamental oscillations of, 69 |
| Gain compression, 29 | heat problem in, 68 |
| Gain ripple, 24 | maximum current swing of, 71 |
| Gain sensitivity, 159 | modes of operation of, 69 |
| Gain stability, 91, 100 | power-frequency limitations of, 69 |
| Gallium antimonide, 3, 10 | pulse operation of, 68 |
| Gallium arsenide, 3, 17, 20, 63, 95, 103, 120, 151 | |
| Gallium arsenide varactor, 96, 120, 121, 153 | |
| Gallium-indium contact, liquid, 73 | |
| Gap capacities, 388 | |
| Gap in coaxial line, 383 | |
| equivalent network for, 384 | |
| Garnet, 19, 128 | |
| Gas contamination, 142 | |
| Gas, exchange, 114 | |
| Gas expansion, isenthalpic and isentropic, 134, 135, 137 | |
| Gas impurities, 138 | |
| Gauss-Seidel cycle, 340 | |
| Gauss’s theorem, 341 | |
| Generator, varactor harmonic, 134 | |
| Generator impedance, 113 | |
| Germanium, 3, 11, 14, 20, 53 | |
| Gifford-McMahon refrigeration cycle, 135, 140 | |
| Graded-junction diode, 115 | |
| Graded-junction doubler, 190, 200 | |
| H | |
| Harmonic generator, 172 | |
| varactor, 56, 134, 165 | |
| Heat exchanger, 137 | |
| counterflow, 138 | |
| periodic-flow, 139 | |
| Heat problem in Gunn oscillator, 68 | |
| Heat reservoir, 139 | |
| Heat sink, 107 | |
| Heating, varactor, 106, 113 | |
| Helium, liquid, 90, 106, 114, 135 | |
| High-current diode, 21 | |
| High-impedance wire type coupling, 311 | |
| High-mobility valley, 81 | |
| High-temperature load resistance, 97 | |
| Highly doped crystals, 21 | |
| Higher power tunnel diode oscillators, 20 | |
| Highpass filter, 254 | |
| Hole electron pair, 83 | |
| Holes, ionization rate for, 84 | |
| h-parameters, 284–292, 297, 298 | |
| Hybrid conformal transformation variational approach, 357 | |
| Hybrid ring, 22 | |
| Hydrostatic pressure, 70 | |
| Hyperabrupt impurity, 75 | |
| Hypergeometric series, 172 | |
| Hysteresis, 214 | |
| I | |
| Idler circuit, 227 | tunable, 151 |
Idler frequency, 92
Idler load resistance, 110
Idler loading factor, 110
Idlers, 167
  bandwidth in multipliers without, 211
  varactor harmonic generators with, 27
  varactor harmonic generators without, 172
IEEE standards on tunnel diodes, 2
Image conductance, 32
Image frequency, 31, 35
Image impedance, 263, 379
Image rejection, 34
IMPATT oscillators, 62
  pn and pin, 79—(see also Avalanche
  transit time oscillator)
  power-frequency performance of, 72
Impedance, amplifier, 100
  characteristics, 22, 345
  circulator, 100
  complementary input, 246
  constant-resistance input, 293
  dynamic, 94
  generator, 113
  image, 263, 379
  input, 241
    real and imaginary parts of, 243
    intrinsic, 347
    inverter, 311
    minimum, 55
    pump circuit, 100
    varactor, 100
Impedance elements, parasitic, 92
Impurity gradient, 122
Impedance level, 79
  and frequency, 50
  and tunnel diode, 66
  and varactor, 61
Impedance and power in transistor, 55
Impedance of tunnel diode, low, 63
Impedance variations, 100
Improved tunnel diode converters, 36
Impurities, gas, 138
Impurity, hyperabrupt, 75
Impurity scattering, 82
Incremental capacitance, 198
Incremental conductance, 198
Indium antimonide, 3
Indium O ring, 149
Indium phosphide, 3
Inductance, 8, 21
Inductance—cont.
  series, 2, 6, 122
  ultra low, 21
Infinite array of parallel-coupled rods, 364
Initial solutions, 347
Inner conductor, 354, 355
  cylindrical rod, 363
  rectangular bar, 363
  step in, 371, 377
  thickness correction for strip line, 358
  undercut in the, 372
Input admittance, 242
  real part of, 305
Input drive level, 180
Input impedance, 241
Input noise temperature, 94
Input resistance, 180
Insertion loss, 100, 126
  temperature dependence of, 131
Insertion loss of circulator, 134
Integrals, elliptic, 350
Integral equation, 364
Integrated amplifier, 30
Interaction effects (between filters), 282
Interdemodulation, 29
Interdigital (bandpass) filter, 254, 272, 296, 298, 363
  modified form of, 277
Interdigital filter for use in diplexer, 279
Interface, column, 337
  dielectric, 337
  row, 337
Intermediate point, ordinary, 336
Intermediate frequency, 30
  low, 35
Intermodulation noise, 159
Internal conductance, 32
Internal field emission current, 52
Internal magnetic field in circulator, 130
Interpolation, linear, 347
Intrinsic impedance, 347
Inversion temperature, 135
Inverter, admittance, 311
Inverter, impedance, 311
Ionization rate for electrons, 84
Ionization rate for holes, 84
Ionized acceptor concentration, 115
Ionized donor concentration, 115
Isenthalpic expansion, 134, 135
Isentropic expansion, 137
Isolation of circulator, 134
Isolator, 34
Iterative solution, 340

J
Johnson's equation, 49, 55
Johnson's study on transistors, 49
Joule-Thompson refrigeration cycle, 135
Junction, aluminium-germanium alloy, 35
area of diode, 21
compensated, 379
line, 21
multiple, 21
p-n, 5
resistance, 3
Junction capacitance, 2, 5, 31, 114
Junction current, 3, 14
Junction diode, 6
pn, 115
Junction resistance, 2, 4
Junction voltage, 3, 14

K
Kirchoff's first circuitual law, 332
Kron's method, 389
Ku band, 28
Kuroda's identity, 283, 286

L
L band, 20, 28, 29, 34
L-band amplifier, 150
Laplace's equation, 328, 332
Latching circulator, switchable, 22
Latent root, 341
Lattice scattering, 82, 83
Lattice vibrations, optical mode, 82
Lifetime of carrier, 104
Light-sensitive emulsion, 47
Line, characteristic impedance of square coaxial, 351
Line junction (in diode), 21
Lines with small losses, 345, 346
Linear algebraic equation, 329 (see also Equations)
Linear equation, simultaneous, 332, 364
(see also Equations)
Linear interpolation, 347
Linearized resonance condition, 175, 230
Linear graded doping, 170 (see also Doping)
Linewidth, 128
Linewidth of polycrystalline YIG, 131
Liquid gallium-indium contact, 73
Liquid helium cooled amplifiers, 152
Liquid helium temperature, 121, 127
circulator, 133 (see also Helium)
Liquid nitrogen, 10
Liquid nitrogen cooled amplifier, 149
Liquid nitrogen temperature, 100, 127
Loads, filter-type, 25
Load conductance, 32
Local oscillator, 30, 33
Loss,
conversion, 32, 226
ferromagnetic resonance, 128
insertion, 100, 126, 134
temperature dependence of, 131
Loss in circulator,
ferromagnetic resonance, 130
insertion, 134
low-field, 131
Losses, conductor (series), 346
Losses in series resistance, 21
Lossy dielectric, 346
Lossless lines, 345
Low-field loss in circulator, 131
Low-frequency efficiency, 197
Low impedance of tunnel diode, 63
Low intermediate frequency, 35
Low-mobility valley, 81
Lowpass filter, 251, 254, 272
Chebyshev prototype, lumped-element, 280
coaXial, 267
doubly terminated prototype, 240
Low pass prototype filter, doubly terminated, 322
singly terminated prototype, 239
Lowpass prototype circuits, 222
Lowpass filter and interconnection, 257
Low-temperature circulator, 130, 132
Lower sideband, 91
Lower sideband up-converter, 92
Lubricants, 142
Lumped capacitance, 310
Lumped-element filter, 240
Lumped-component equivalent circuits, 166
Lumped-element low-pass Chebyshev prototype filter, 280
Magnet, permanent, 133
   superconducting, 132, 133
Magnetic conductors, 351, 354, 383
Magnetic coupling, 310
Magnetic field, 159
   biasing, 130
   operating, 130
Magnetic field in circulator, internal, 130
Magnetic planes, 378
Magnetic saturation, 130
Magnetostatic considerations, 369
Maintenance cost of refrigerator, 160
Mapped prototype, 283
Maser, 89, 134, 159
   traveling wave, 159
Mass, effective, 70
Materials constants, 48
Material uniformity, 55
Match, 28
Matching network, 218
Matrix, 341
Maximally flat filter, 240
Maximum charge, 78
Maximum current in transistor, 49, 52
Maximum current in tunnel diode, 66
Maximum current swing of Gunn oscillator, 71
Maximum efficiency, 180, 183, 190
Maximum peak-current diode, 21
Maximum power, 20, 55
Maximum power convertible in varactor, 60
Maximum power output, 20
Maximum voltage, 78
Maxwell's equations, 389
McKay's equation, 51
Measured attenuation of three-channel multiplexer, 321
Measured performance of diplexer, 263, 270, 271, 303
Mechanical vibration, 100
Media, stepwise continuous, 334
Mesa varactor, 106
Mesa (construction), 120
Mesa varactor, 106, 117
Mesh, 114, 333
   wire, 7
Microplasma, 73
Microplasma noise, 96
Microstrip transmission line, 365
Microwave generators, solid state, 165
Microwave printed circuitry, 63
Microwave power, 12
Microwave generators, solid state, 43
Microwave tubes, 329
Millimeter waves, 11, 19
Minimum dissipation, Van der Pohl's principle of, 16
Minimum impedance, 55
Minimum reactance, 241, 295, 305
Minimum susceptance, 241, 295, 305
Minority carriers, 101
Minority carrier charge storage effects, 62
Mitered bend in parallel-plate line, 353
Mixed dielectric, 357
Mixer, 31
   passive, 33
Mobile charges, 55
Mobility, effective, 70
Mode, 19
   even and odd, 361
   ground plane, 271
Mode-matching method, 370
Modes of operation of Gunn oscillator, 69
Model, transistor, 48
Model of a varactor, 168
Modulation factor, 109
   capacitance, 95
Multiplexer, 237, 301, 322
   computed attenuation of three-channel, 316
   measured attenuation of three-channel, 321
   parallel-connected, 305, 307
   series connected, 296, 305
   three-channel, sketch of, 318
Multiplexer filters, 322
Multiplexer of three channels, 305
Multiplexers having contiguous channels, 305
Multiplexers having noncontiguous channels, 293
Multiple junction, 21
Multiple transit carrier, 104
Multiplication, order of, 208
Multiplication factor, 85
Multiplier circuit configuration, 223
Multipliers, varactor, 166, 198, 205, 211, 223
<table>
<thead>
<tr>
<th>Subject</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multivalley conduction process</td>
<td>70</td>
</tr>
<tr>
<td>Mutual capacitance</td>
<td>279</td>
</tr>
<tr>
<td>Noise, 62</td>
<td></td>
</tr>
<tr>
<td>excess</td>
<td>103</td>
</tr>
<tr>
<td>intermodulation</td>
<td>159</td>
</tr>
<tr>
<td>microplasma</td>
<td>96</td>
</tr>
<tr>
<td>1/f</td>
<td>35</td>
</tr>
<tr>
<td>shot</td>
<td>96, 105</td>
</tr>
<tr>
<td>thermal</td>
<td>94</td>
</tr>
<tr>
<td>Noise behaviour, anomalous</td>
<td>104</td>
</tr>
<tr>
<td>Noise conductance</td>
<td>33</td>
</tr>
<tr>
<td>Noise constant</td>
<td>11, 25</td>
</tr>
<tr>
<td>Noise current</td>
<td>9, 33</td>
</tr>
<tr>
<td>Noise-current generator, 3</td>
<td></td>
</tr>
<tr>
<td>Noise figure, 11, 25, 28, 29, 32, 33, 35, 36, 92</td>
<td></td>
</tr>
<tr>
<td>double-side-band</td>
<td>35</td>
</tr>
<tr>
<td>radiometer</td>
<td>35</td>
</tr>
<tr>
<td>single-side-band</td>
<td>35</td>
</tr>
<tr>
<td>Noise measure</td>
<td>95</td>
</tr>
<tr>
<td>Noise power</td>
<td>104</td>
</tr>
<tr>
<td>Noise sources, 2</td>
<td></td>
</tr>
<tr>
<td>Noise temperature, 91, 152</td>
<td></td>
</tr>
<tr>
<td>input</td>
<td>94, 114</td>
</tr>
<tr>
<td>Noise voltage</td>
<td>9</td>
</tr>
<tr>
<td>Noise-voltage generator, 3</td>
<td></td>
</tr>
<tr>
<td>Noncontiguous channels, multiplexers having</td>
<td>293</td>
</tr>
<tr>
<td>Nondegenerate amplifiers, 90, 92</td>
<td></td>
</tr>
<tr>
<td>Nonlinear biasing resistors, 20</td>
<td></td>
</tr>
<tr>
<td>Nonlinear element, 15</td>
<td></td>
</tr>
<tr>
<td>Nonlinear negative resistance, 12</td>
<td></td>
</tr>
<tr>
<td>Nonminimum-reactance network, 242</td>
<td></td>
</tr>
<tr>
<td>Nonoscillating converters, 34</td>
<td></td>
</tr>
<tr>
<td>Normal derivative</td>
<td>343</td>
</tr>
<tr>
<td>Ocean floor, 89</td>
<td></td>
</tr>
<tr>
<td>Octuplers, 231</td>
<td></td>
</tr>
<tr>
<td>Odd mode</td>
<td>361</td>
</tr>
<tr>
<td>Ohmic contact, 83, 107</td>
<td></td>
</tr>
<tr>
<td>Ohmic regions, 55</td>
<td></td>
</tr>
<tr>
<td>Open-circuit edges</td>
<td>335</td>
</tr>
<tr>
<td>Open-circuited stub</td>
<td>260</td>
</tr>
<tr>
<td>Operating frequency</td>
<td>21, 25, 130</td>
</tr>
<tr>
<td>Operating magnetic field</td>
<td>130</td>
</tr>
<tr>
<td>Optical mode lattice vibrations</td>
<td>82</td>
</tr>
<tr>
<td>Optical phonon, 83</td>
<td></td>
</tr>
<tr>
<td>collisions</td>
<td>66</td>
</tr>
<tr>
<td>Optimization of efficiency</td>
<td>179, 197</td>
</tr>
<tr>
<td>Optimization of output power</td>
<td>179</td>
</tr>
<tr>
<td>Optimum bandwidth</td>
<td>24</td>
</tr>
<tr>
<td>Optimum load resistance</td>
<td>182</td>
</tr>
<tr>
<td>Optimum operating frequency of Read diode</td>
<td>79</td>
</tr>
<tr>
<td>Optimum pump frequency</td>
<td>108</td>
</tr>
<tr>
<td>Order of multiplication</td>
<td>208</td>
</tr>
<tr>
<td>Ordinary interior point</td>
<td>336</td>
</tr>
<tr>
<td>Ordinary point</td>
<td>336</td>
</tr>
<tr>
<td>Orthonormal block analysis, 349</td>
<td></td>
</tr>
<tr>
<td>Orthogonal set</td>
<td>334</td>
</tr>
<tr>
<td>Oscillations of Gunn oscillator, fundamental</td>
<td>69</td>
</tr>
<tr>
<td>Oscillations, sinusoidal</td>
<td>13</td>
</tr>
<tr>
<td>spurious</td>
<td>18</td>
</tr>
<tr>
<td>Oscillator, 21</td>
<td></td>
</tr>
<tr>
<td>avalanche transit time</td>
<td>71 (see also Avalanche transit time oscillator)</td>
</tr>
<tr>
<td>electronically tuned tunnel-diode</td>
<td>19</td>
</tr>
</tbody>
</table>
Oscillator—cont.
fixed-frequency, 17
Gunn, 62, 67, 81, 166 (see also Gunn)
IMPATT, 62, 72, 79 (see also Avalanche transit time oscillator)
tunable, 17, 19
tunnel-diode, 12 (see also Tunnel diode)
Outer conductor, overcut in the, 372
rectangular, 354
step in, 371, 377
Outer space, 89
Output power, 19, 197, 202
optimization of, 179
Overcut in the outer conductor, 372
Overdriving, 205
Overlay transistor, 45
Oxide, 47

P
Package, 107
pill, 6, 9
Package capacitance, 2, 9, 122
Package resistance, 6
Parallel-connected diplexer, 254, 272
Parallel-connected multiplexer, 307
Parallel-coupled-resonator bandstop filter, 296
Parallel-coupled rods, infinite array of, 364
Parallel-plate capacity, 354, 362
Parallel-plate line, bifurcated, 353, 356
discontinuity in, 373
Parameters, tunnel-diode, 2
Parametric amplifiers, cooled, 89–164
performances of, 157
reliability of, 134
Parasitic elements, 13, 21, 32, 92, 98, 123
Parasitic impedance, 92
Parasitic resistance, 25
Partly complementary diplexer, 251–324
Particle current, 60, 77
Particle injection rate, 77
Passband, 251
second, 310, 311
Passive converters, 32, 34
Passive input resistance, 95
Passive mixer, 33
Peak-charge, 199
Peak current, 6, 21, 26, 165
Peak-current density, 3
Peak-current diode, maximum, 21

Peak field, 79
Peak voltage, 3, 26
forward, 58
reverse, 59
Performances of cooled parametric amplifiers, 157
Periodic-flow heat exchanger, 139
Permanent magnet, 133
Permittivity, 115
complex, 346
Permutation, 336
Phase angles, 232
Phase response, 24
Phase velocity, 357
Phased array, 29
Phenolic, 143
Phonon collisions, optical, 66
Phosphorus, 116
Photoresist, 56
Photolytic methods, 6
Physical temperature, 107
Pill package, 6, 9, 18
Piecewise homogenous dielectrics, 330
pn junction, 5, 48
pn junction diodes, 115
pn and pin IMPATT oscillators, 79
Polder tensor components, 130
Polycrystalline YIG, linewidth of, 131
Polyiron, 271
Polygonal lines, 349
regular, 352
Polynomials, 241
Point-contact crystal diode, 30
Point, corner, 336
ordinary, 336
Poisson's equation, 60
Post in waveguide, 9
Potential, contact, 5, 169, 198
Potential field, computed, 341
Power convertible in varactor, maximum, 60
Power deliverable to carriers, upper limit on, 55
Power delivered to the load, 13
Power dissipation, pump, 106
Power-frequency capabilities of varactors, 56
Power-frequency limitations of avalanche transit time oscillators, 78
Power-frequency limitations of Gunn oscillator, 69
Subject Index

Power-frequency limitations of transistors, 48
Power-frequency limitations of tunnel diodes, 66
Power-frequency performance of Gunn oscillator, 68
Power-frequency performance of IMPATT oscillator, 72
Power-frequency performance of transistors, 47
Power gain, reflection, 94
Power-handling capability, 27, 231
Power and impedance in transistor, 55
Power limitation of a dielectric varactor harmonic generator due to avalanche and transit time effects, 61
Power maximum, from transistor, 55
Power, microwave, 12
Power output, 15, 16, 18, 20, 25, 28, 29, 36, 55, 197, 202
Power series expansion, 172
Power sources, solid state, 12
Power supply, 33
Power varactor, 62
Preamplifier, 126
Pressure, hydrostatic, 70
Principle of minimum dissipation, Van der Pohl's, 16
Printed circuitry, microwave, 63
Problems, well-behaved (without reentrant corners), 348
Proximity effects, 374
Proximity factors, 375
Principle of minimum dissipation, Van der Pohl's, 16
PTFE, 357
Pulse operation of Gunn oscillator, 68
Pulse radar, 22
Pump, 91
  reliability of, 134
Pump circuit impedance, 100
Pump cycle, 104
Pump frequency, 92, 97, 159
  optimum, 108
Pump power, 109, 154
Pump power dissipation, 106
Push-pull amplifier, 30

Q

Q (factor)—cont.
  negative, 13
Q (factor) of circulator mode, 130
Quadrant (geometry in finite-difference equations), 337
Quadruplers, 231
Quality factor (see Q factor)
Quartic equation, 223
Quartz, 9
Quasi-static method, 369
Quintuplers, 231

R

Radar, 35, 89, 91
  pulse, 22
Radar receiver, 146
Radian frequency, 241
Radio astronomy, 89, 91
Radio observatory, 90
  radiometer, 146
  Radiometer, 92
  noise figure, 35
Radiometric systems, 90
RC product, 67
Reactance, minimum, 241, 295
  series, 243
Reactance-annulling network, 254 (see also Annulling network)
Reactive cutoff frequency, 11
Read diode, 71, 75
  optimum operating frequency of, 79
Real frequency, 241
Real and imaginary parts of the input admittance, 243
Real and imaginary parts of the input impedance, 243
Real part of input admittance, 305
Recovery time, 159
Rectangular conductors, coaxial, 349
  inner, 355, 363
  outer, 354
Rectangular net, 354
Rectification, 198
Re-entrant cavity, 18
Reentrant corners, 348
Reflection power gain, 94
Reflection-type tunnel-diode amplifier, 21, 23
Reflectometer, 301
Refrigeration, 95
Refrigeration cycle, Gifford-McMahon, 135
Joule-Thompson, 135
Sterling, 135
Refrigerator, 134
Claude cycle, 138
closed-cycle, 134, 138
commercial, 144
efficiency of, 140
expansion-engine, 137
Gifford-McMahon cycle, 140
maintenance cost of, 160
Sterling cycle, 139
Refrigerator engine speed, 139
Refrigerator systems, closed-cycle, 90
Regenerator, 139
Regular polygonal line, 352
Relaxation process, 343
Relaxation, successive over, 340
Reliability of cooled parametric amplifiers, 134
Reliability of the pump, 134
Resistance, 21
  base, 6
  charge and frequency in varactor doubler, 203
  contact, 6
  high temperature load, 97
  idler load, 110
  input, 180
  junction, 2, 3, 4
  negative, 2, 10, 12, 25, 64, 71, 76
  optimum load, 182
Resistance, package, 6
  parasitic, 25
  passive input, 95
  series, 2, 5, 92, 95, 114, 118, 122
  shunt, 101
  spreading, 117, 119
  thermal, 106, 122 (see also Resistor)
Resistance property, constant, 238
Resistive cutoff frequency, 11, 62
Resistive losses, 67
Resistivity, surface, 143
Resistor, nonlinear biasing, 20
  stabilizing, 12, 17, 18 (see also Resistance)
Resonance condition, linearized, 175, 230
Resonance mode, 130
Resonance situation, 202
Resonant frequency of annulling network, 262
Resonant solution, 174, 230
Resonator, conical waveguide, 19
  series bandstop, 300
  shunt bandpass, 300
  shunt bandstop, 300
Resonator lines, 310
Reverse bias, 78
  in tunnel diode, 20
Reverse breakdown voltage, 64, 196, 198
Reverse currents, 96
Rexolite, 357
Richardson's method, 332, 347
Rieke diagram, 19, 20
Room temperature, 95, 100
Rotating modes in circulator, 129
Rotational symmetry, 329
Row interface, 337

S

S band, 18, 20, 34
Satellite communications, 89
Satellite ground station, 91
  receiver, 146
Saturated drift velocity, 49, 52
Saturation, 23, 26
  carrier velocity, 51, 81, 83
  magnetic, 130
Saturation field, velocity, 77
Saturation power, 159
Scaled equations, 196
Scattering, impurity, 82
  lattice, 82, 83
Scattering cross section, 83
Schwarz Christoffel conformal transformation, 350, 356
Schottky barrier, 106
Schottky barrier diode, 30, 115, 159
Second passband, 310, 311
Self bias, 100, 216
Self-oscillating converter, 33
Self-purging engines, 142
Self-resonant frequency, 10, 11, 18, 21, 98, 155
Semiconductor-metal surface barrier diode, 115
Semiconductor wafer, 20, 107
Semilumped elements, 251
Series bandstop resonator, 300
Series connected diplexer, 263
Series connected multiplexer, 296
Series diode, 226
Series gap in coaxial line, 383
Series inductance, 2, 6, 122
Series model of varactor harmonic generator, 173
Series reactance, 243
Series resistance, 2, 5, 92, 95, 114, 118, 122
  losses in, 21
Series stub, 265
Set, orthogonal, 334
Shielded strip line, 354, 355
Shot noise, 9, 33, 91, 96, 101, 105
Shunt bandpass resonator, 300
Shunt bandstop resonator, 300
Shunt capacitor, 252, 257
Shunt model of varactor harmonic generator, 173
Shunt resistance, 101
Shunt susceptance, 243
Sideband, lower, 91
  upper, 91
Side-band frequencies, 30
Signal, weak, 89
Signal frequency, 92
Signal frequency ratio, optimum idler to, 109
Signal quality, 62
Silicon, 3, 47, 95, 103
Silicon crystal, 120
Silicon varactor, 62, 121, 153
  epitaxial, 121, 122
Simultaneous equations, 332, 335, 364
Single-side-band noise figure, 35
Single transit carrier, 104
Single-varactor configuration, 98
Singly terminated filter, 250, 308
  low-pass prototype, 239, 283
Sinusoidal oscillations, 13
Skin depth, 120
Skin effect, 6, 347
Slope factors, 98
Slope parameter, 312
Slotted cross section of strip line, 360
Slotted line section, 358
Small aperture theory, 383
Small signal approximations, 191, 233
Small-signal gain, 21
  “Snapshots” in time of avalanche process, 77
Solid state microwave generators, 12, 43, 165
Solution, energy integral, 344
  initial, 347
  iterative, 340
SOR, 340
Special finite difference equations, 336
Spinel, 128
Splitting, frequency, 129
Spreading resistance, 117, 119
Spurious frequencies, 62
Spurious sideband frequency, 97
Spurious oscillations, 18
Square nets, 332
Square coaxial line, 349
  characteristic impedance of, 351
Step discontinuities in coaxial lines, 369, 370
Stepped conductor (inner or outer), 371, 372, 377
Step-recovery diode, 167
Steps in both conductors simultaneously, 372
Stepwise continuous media, 334
Strip line, 8, 17, 28, 29, 296, 301, 349, 354, 361, 365, 386
  balanced, 8
  capacitive gaps in, 386
  characteristic impedance of, 358
  coupled, 356
  dielectric supported shielded, 357
  micro, 365
  shielded, 355
  slotted cross section of, 360
Strip line bandpass filters, dielectric supported, 387
Strip line construction, 360
Strip line inner conductor, thickness corrections for, 358
Stability, 24
  frequency, 19
  gain, 100
Stabilizing network, 28
Stabilizing resistor, 12, 17, 18
Stagger tuning, 99
Stainless steel waveguide, 148
Standard temperature, 95
Sterling cycle refrigerator, 139
Sterling refrigeration cycle, 135
Stop band, 251
Storage effect, 104
Stray capacitance, 122
Stub, open-circuited, 260
  series, 265
SUBJECT INDEX

Stub-type bandstop filter, 272
Styrofoam, 148
Substrate, 107
Subsurfaces, 342
Successive over-relaxation, 340
Sulfur dioxide, 135
Superconducting magnet, 132, 133
Surface, annular, 342
Surface barrier varactor, 103
gallium arsenide, 153
Surface integral, 344
Surface of integration, 342
Surface resistivity, 143
Susceptance-annulling network, 251, 258
(see also Annulling network)
Susceptance, minimum, 241
shunt, 243
Switchable latching circulator, 22
Symmetry, 352
axis of, 342
rotational, 329
Symmetry in transmission lines, 328

T

Taylor expansion, 333
Teflon, 148, 357, 368, 379
Tellurium copper, 155
TEM mode transmission lines, 329
Temperature, 4, 10, 20, 25, 28, 33
ambient, 95
cryogenic, 95
equivalent, 107
input noise, 94
inversion, 135
liquid helium, 121, 127
liquid nitrogen, 100, 127
noise, 152
physical, 107
room, 95, 100
standard, 95
Temperature dependence of insertion loss, 131
Temperature dependence of operating frequency, 128
Temperature-insensitive circulator, 134
Tensor components, Polder, 130
Theory of tunnel diode
amplifiers, 21
converters, 30
oscillator, 12
Thermal conductivity, 123, 143
Thermal design (of transistors), 46, 55
Thermal frequency stabilization (of parametric-amplifier pump), 147
Thermal noise, 94
Thermal path, 143
Thermal radiation, 143
Thermal resistance, 106, 122
Thermodynamic cycle, 134
Thickness correction for strip line inner conductor, 358
Time, computing, 352, 354, 356, 371, 389
Time-varying noise conductance, 33
Transistors, 45, 165
base widening in, 49
definition of, 46
Early–Johnson approach to, 48
impedance level and power of, 55
Johnson’s study on, 49
maximum current in, 49, 52
overlay, 45
power and impedance level of, 55
power-frequency limitations of, 48
power-frequency performance of, 47
tunneling in, 52
Transistor model, 48
Transit time, 68, 80
Transit time cutoff frequency, 59, 62
Transit time effects, power limitation of a
dielectric varactor harmonic generator
due to avalanche and, 61
Transmission lines, 143, 349
plated, 143
Transmission line constants, 345
Transmission-type amplifier, 21
Transport current, 60, 77
Trapezoidal rule, 343
Traveling wave maser, 159
Traveling-wave tube, 89
Tripler, 56, 231
Tripler, 318, 323
TR tube leakage, 22
Tubes, microwave, 329
Tunable idler circuit, 151
Tunable oscillators, 17, 19
Tuned tunnel-diode oscillators, electronically, 19
Tuning of filter, 301
Tunnel diode, 1, 62, 165
built-in potential of, 65
Tunnel diode—cont.
equivalent circuit of, 2, 23, 25, 30
forward bias direction of, 65
impedance level of, relative to microwave
circuit, 66
low impedance of, 63
maximum current in, 66
maximum voltage on, 66
power-frequency limitations of, 66
reverse-biased, 20
unstable, 34
Tunnel diode amplifiers, 21
experimental, 27
improved, 30
reflection-type, 23
theory of, 21
Tunnel diode converters, 30
experiments on, 33
improved, 36
theory of, 30
Tunnel diode devices, 1
Tunnel-diode oscillators, 12
electronically tuned, 19
experimental results on, 17
higher power, 20
theory of, 12
Tunnel-diode parameters, 2
Tunneling current, 66
Tunneling in transistors, 52
Turns ratio, 280
Type-N connector, 271, 380

V
Valence bands, 3
Valley, 70
  high-mobility, 81
  low-mobility, 81
Valley-current density, 3
Valley voltage, 3
Van der Pohl’s principle of minimum
dissipation, 16
Vapor pressure, 137
Varactor, 19, 75, 90, 92, 115
  bias on, 58
  capacitance of, 105
  capacitance-voltage curve for, 169
  charge-capacitance, voltage relationships
  of, 169
  charge-voltage curve for, 169
  cooled, 122
  depletion layer, 58
  dielectric, 58, 61
  diffused GaAs, 116
  elastance of, 171
  epitaxial, 103, 116
  silicon, 121, 122
gallium arsenide, 96, 120, 121, 153
high-Q-factor, 214
impedance level and, 61
maximum power convertible in, 60
  mesa, 106, 117
  model of, 168
  power, 62
  power-frequency capabilities, 56
  silicon, 62, 121, 153
  surface barrier, 103
  underdriven, 196
Varactor chain, 62
Varactor characteristics, 114
Varactor configuration, single and balanced, 98
Varactor dependence of capacitance on
  voltage, 167
Varactor diode, 6
Varactor doubler, 147
  broadband, 222
  resistance, charge and frequency in, 203
Varactor energy transfer, 60
Varactor harmonic generator, 56, 134, 165
  series and shunt model of, 173
Varactor harmonic generators without
  idlers, 172
Varactor heating, 106, 113
Varactor impedance, 100
Varactor model, 167
Varactor multiplier, 166, 198, 205, 207, 211, 223
forward-driven, 198
Varactor package, 98
Varactor parametric amplifiers, cooled, 89–164
Variation, capacitance, 108
gain, 100
Variational method, 328, 349, 356
hybrid conformal transformation, 357
Velocities of holes and electrons, 83
Velocity of carriers, 49
Velocity ratio, 357
Velocity saturation, carrier, 51, 81, 83
Velocity saturation field, 77
Video detectors, 35
Vibration, 139
acoustical, 82
mechanical, 100
optical mode lattice, 82
Voltage
bias, 10, 111, 115, 122, 179
forward, 125
breakdown, 118
diffusion, 169
junction, 3, 14
maximum, 78
peak, 3, 26
forward, 58
Voltage, peak—cont.
reverse, 59
reverse breakdown, 64, 196, 198
valley, 3
Voltage generator, zero-impedance, 239
Voltage swing, 13
Voltage on tunnel diode, maximum, 66

W
Wafer, semiconductor, 107
Water cooling, 139
Waveguide, 9, 17, 19, 28
stainless steel, 148
Waveguide resonators, conical, 19
Weak signals, 89
Well-behaved problems, 348
Wide-band diplexer, 261
Wire mesh, 7

XYZ
X-band, 17
amplifier, 155
YIG, linewidth of polycrystalline, 131
Y-junction circulator, 128
Young's property A, 332
Yttrium iron garnet (see YIG)
Zero-admittance current generator, 239
Zero bias capacitance, 155
Zero-impedance voltage generator, 239
Complete in three volumes

Microwave Scanning Antennas

Edited by R. C. Hansen

VOLUME 3: ARRAY SYSTEMS
1966, 421 pp., $16.50

CONTENTS:
W. H. Kummer, Feeding and Phase Scanning
N. A. Begovich, Frequency Scanning
Jesse L. Butler, Digital, Matrix and Intermediate-Frequency Scanning
Robert W. Bickmore, Time vs. Space in Antenna Theory
Donald L. Margerum, Self-Phased Arrays

AUTHOR INDEX — SUBJECT INDEX.

VOLUME 2: ARRAY THEORY AND PRACTICE
1966, 401 pp., $15.50

CONTENTS:
R. S. Elliott, The Theory of Antenna Arrays
A. A. Oliner and R. G. Malech, Radiating Elements and Mutual Coupling
A. A. Oliner and R. G. Malech, Mutual Coupling in Infinite Scanning Arrays
A. A. Oliner and R. G. Malech, Mutual Coupling in Finite Scanning Arrays

AUTHOR INDEX — SUBJECT INDEX.

VOLUME 1: APERTURES
1964, 442 pp., $16.00

"Volume 1 represents the best and most complete treatment that I have seen on its subject.

"It has a unity and a coverage that should commend it to those wishing a serious introduction to studies and advances made in a field that is important in radio astronomy, in space communications, and in defense."

—JOURNAL OF THE FRANKLIN INSTITUTE

CONTENTS:
R. C. Hansen, Aperture Theory
L. K. De Size and J. F. Ramsay, Reflecting Systems
R. C. Johnson, Optical Scanners
Hsien Ching Ko, Radio-Telescope Antennas
J. A. Vitale, Large Radomes
Appendix I: Tables of Taylor Line Source Distributions
Appendix II: Tables of Taylor Circular Aperture Distributions

AUTHOR INDEX — SUBJECT INDEX.

ACADEMIC PRESS, New York and London
111 Fifth Avenue, New York 10003
Berkeley Square House, London, W. 1